Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16)

Fast Motion Prediction for Collaborative Robotics

Claudia Pérez-D'Arpino and Julie A. Shah
Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology
{cdarpino, julie_a_shah} @csail.mit.edu

Abstract

The efficient and safe performance of collaborative robots requires advancements in perception, control, design and algorithms, among other factors. With regard to algorithms, representing the structure of collaborative tasks and reasoning about progress toward task completion in an on-line fashion enables a robot to be a fluent and safe collaborator based on its ability to predict the next actions of a human agent. With this goal in mind, we focus on real-time target prediction of human reaching motion and present an algorithm based on time series classification. Results from on-line testing involving a tabletop task with a PR2 robot yielded 70% prediction accuracy with $400msec$ of observed trajectory.

1 Introduction

The execution of manipulation tasks for which humans and robots must share a workspace represents new challenges for robotics in manufacturing environments. A robot's motion planning algorithm must allow for awareness of the human pose and, ideally, of their next actions to be able to compute trajectories in a timely fashion that synchronize naturally with human motion. We believe that accurate prediction of human actions will be a key enabler of collaborative robotics, as it allows a robot to plan in advance according to these predictions, as opposed to using a purely reactive approach. Here, we focus on the problem of predicting the target of human reaching motion during a set of tabletop human-robot collaborative tasks within a manufacturing environment. In this collaborative setting, a robot can assume different roles while sharing the workspace with the human, including:

Facilitating tools and parts: The robot should provide the human with the right tool at the right time by maintaining a convenient posture when offering the tool.

Asynchronous assistance: Both the robot and human agents work toward accomplishing their own tasks, with spatio-temporal constraints with inter-agent dependencies.

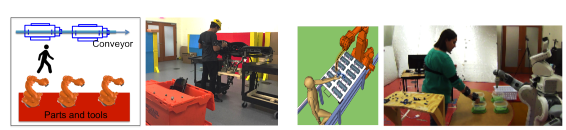
Physical collaboration: Interaction between the human and robot involves force and contact created by both agents.

We focus on the first two roles, which involve no force or contact. In these circumstances, the assistance of the robot is mainly intended to reduce the time a human would need to spend on *non-value added work* if completing the task alone. For example, on car manufacturing assembly lines, the main body and larger components of a car are typically transported along a main conveyor line while associates retrieve parts from the periphery of the conveyor and manually install them, as illustrated in Fig.1a. This installation process requires expert skill, but the time spent retrieving parts could be saved by incorporating a robotic assistant, improving efficiency, avoiding errors during repetitive tasks and increasing the physical comfort of the human. Similarly, kitting and assembly processes are examples of circumstances in which asynchronous assistance from a robot would be beneficial (Fig.1b).

We have developed a prediction algorithm that incorporates a data-driven dual approach to exploit knowledge about human actions at both the motion and task levels in order to predict the intended target of a human reaching motion in real time [Pérez-D'Arpino and Shah., 2015]. Using time series where each time step is encoded as a multivariate Gaussian distribution in feature space, this algorithm builds a library of motions from human demonstrations, based on a statistical representation of the degrees-of-freedom of the human arm across multiple tasks. This library of motions is then used to perform time series classification in real time to predict the target at the motion level using the initial partial segment of the trajectory of the human arm.

2 Related Literature

At a high level, there are two main corpus of literature that focus on motion prediction for collaborative robotics. One approach is based on the concept of inverse reinforcement learning (IRL), while the other is based on statistical data modeling and likelihood computation. The IRL approach involves learning a cost function that explains the data and that can



(a) (b)

Figure 1: Illustration of collaborative manipulation tasks: (a) facilitating tools and parts, (b) asynchronous assistance.

be used to generate similar motions, which allows for generalization to new scenarios [Dragan and Srinivasa, 2013]. However, IRL also involves the design of this objective function that can capture the correct dynamics of human motion, which remains a manual process. Recent work using the IRL approach has incorporated human motion data during collaborative tasks to learn a cost function for human-human collaboration [Mainprice *et al.*, 2015].

The statistical approach does not involve assumptions or domain-specific information encoded on cost functions, but it is less clear how to achieve generalization for prediction in new scenarios using this model. A common technique for this approach is to use Gaussian mixture models (GMM) to model motion classes within the data using the expectation maximization (EM) algorithm [Calinon *et al.*, 2007; Mainprice and Berenson, 2013; Ewerton *et al.*, 2015]. Our work lies on the statistical side of the spectrum, as we adapt a Bayesian technique developed for action recognition and exploit its ability to compute partial results on-line [Dong and Williams, 2011; 2012].

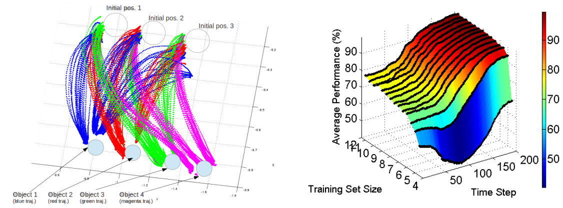
3 Prediction Performance

Our algorithm was validated off-line using real human motion data and tested on-line using a cooperative table-top manipulation task with a PR2 robot [Pérez-D'Arpino and Shah., 2015]. This task required the collection of parts from a table, with four possible targets located along a single axis and three possible initial positions, for a total of 12 possible motion classes. Fig.2a depicts the 3D trajectories of the right hand of the human operator performing 20 demonstrations of each motion class. Validation was performed using 25 random libraries of motions per training set to record the percentage of correct classification per time step across all 12 motion classes using five random test trajectories, for a total of 125 tests per class. The results are plotted in Fig. 2b as a performance surface, indicating the evolution over time of the average correct classification per time step as a function of the number of demonstrations.

Our findings show considerable improvement over current techniques in early prediction, with the algorithm achieving 70% or higher average correct classification having observed the first third of the trajectory ($\sim 400msec$), enabling on-line prediction in a timely fashion [Pérez-D'Arpino and Shah., 2015]. When compared to a GMM approach, we observed that our method improved classification accuracy by 15% at most in the early part of the trajectory, when the number of Gaussian components was limited to run in real time, comparable to our method. We also tested the prediction algorithm in the domain of human walking and found that it generalizes well to other time series dynamics and different sets of features [Pérez-D'Arpino* *et al.*, 2015].

4 Summary, Contributions and Future Work

We presented a prediction algorithm for human motion and developed a multithreading, real-time implementation [Pérez-D'Arpino and Shah., 2015]. Unlike previous work, our algorithm is based on time series modeling, shows improvements over GMM (particularly early in the trajectory)



(a) (b)

Figure 2: (a) 3D trajectories of the right hand during task execution, (b) Results of prediction performance.

and allows for processing in real time. We tested the system using a tabletop pick-and-place task of the asynchronous assistance type involving a PR2 robot, video available at https://youtu.be/OT3ybT-e6l0. The algorithm has also been tested in a walking prediction application with features derived from biomechanical anticipatory signals [Pérez-D'Arpino* *et al.*, 2015]. In future work, we will explore the generalization to unseen motion classes and further develop the system for action prediction at both the motion and task levels using a task that requires a robot to provide parts and tools to a human worker.

References

[Calinon *et al.*, 2007] Sylvain Calinon, Florent Guenter, and Aude Billard. On learning, representing, and generalizing a task in a humanoid robot. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 37(2):286–298, 2007.

[Dong and Williams, 2011] Shuonan Dong and Brian Williams. Motion learning in variable environments using probabilistic flow tubes. In *IEEE ICRA*, pages 1976–1981, 2011.

[Dong and Williams, 2012] Shuonan Dong and Brian Williams. Learning and recognition of hybrid manipulation motions in variable environments using probabilistic flow tubes. *International Journal of Social Robotics*, 4(4):357–368, 2012.

[Dragan and Srinivasa, 2013] Anca Dragan and Siddhartha Srinivasa. Generating legible motion. In *RSS*, 2013.

[Ewerton *et al.*, 2015] Marco Ewerton, Gerhard Neumann, Rudolf Lioutikov, Heni Ben Amor, Jan Peters, and Guilherme Maeda. Learning multiple collaborative tasks with a mixture of interaction primitives. In *IEEE ICRA*, 2015.

[Mainprice and Berenson, 2013] Jim Mainprice and Dmitry Berenson. Human-robot collaborative manipulation planning using early prediction of human motion. In *IEEE/RSJ IROS*, pages 299–306, 2013.

[Mainprice *et al.*, 2015] Jim Mainprice, Rafi Hayne, and Dmitry Berenson. Predicting human reaching motion in collaborative tasks using inverse optimal control and iterative re-planning. In *IEEE ICRA*, 2015.

[Pérez-D'Arpino and Shah., 2015] Claudia Pérez-D'Arpino and Julie A. Shah. Fast target prediction of human reaching motion for cooperative human-robot manipulation tasks using time series classification. In *IEEE ICRA*, 2015.

[Pérez-D'Arpino* *et al.*, 2015] Claudia Pérez-D'Arpino*, Vaibhav Unhelkar*, Leia Stirling, and Julie A. Shah. Human-robot co-navigation using anticipatory indicators of human walking motion. In *IEEE ICRA*, 2015.