Human Motion Trajectory Prediction in Human-Robot Collaborative Tasks

Shiqi Li, Haipeng Wang, Shuai Zhang, Shuze Wang and Ke Han

School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan, China

*wanhai463100@126.com

Abstract. A method is introduced to predict human motion trajectory in the process of human-robot collaboration (HRC). In the method, the human-robot distances are assumed to be a Gaussian Process (GP). To achieve this, a human-robot handover task is conducted by a human and a collaborative robot, while the positions of the human hand and the robot end-effector are recorded. Some of the recorded data are used for the Gaussian Process Regression (GPR), a GP and a 95% confidence convince about the GP are obtained by the GPR. Experimental results show that about 80% of the testing data are included in the 95% confidence convince. The method and results here are useful to other human-robot collaborative tasks where existing human-robot relative motions, especially, the method is able to predict the human motion trajectory with varying initial position of the human hand and varying locations of the robot end-effector.

1. Introduction

A natural human-robot collaboration (HRC) flow is important to reduce human's waiting time, increase efficiency and enhance naturalness [1, 2, 3]. To achieve that, it is critical for the robot to comprehend their human peers’ on-going actions and predict their behaviours in the near future, to start motion planning and execution early enough to smooth the turn-taking transition. So, not only the human motion target, but also the human motion trajectory over time is required.

The robots make prediction of human motion mainly for the purpose of collision avoidance [4, 5] or adapting to the human [6, 7]. The realm of human motion prediction can be divided into two categories [8]: the unsupervised approaches and the supervised approaches. In terms of the unsupervised approaches, the Hidden Markov Model (HMM) [9], the Support Vector Machine (SVM) [10], the Partially Observable Markov Decision Processes (POMDPs) [11] and the social force model [12] are always used, in recent years, the brain-machine interface (BMI) have demonstrated success by means of the sparse linear regression to predict hand trajectory [13, 14]. In terms of the supervised approaches, the inverse optimal control [15, 16] and the neural network, especially the Long Short-Term Memory network [17, 18, 19] are always used, in which a series of labelled demonstrations of specific tasks are required, the labelled demonstrations may be a series of motions of a specific task [20], or a motion library of basic human motion [21, 22].

However, it is found that most of the researches are focused on the human motion target prediction or the human behaviour recognition, the problem with this is that the robot is usually able to predict what the human is about to do, but the motion trajectories in a future time is still unknown.
Figure 1. Illustration of the human motion trajectory: in these scenarios, the starting and ending points are known, while the human motion trajectory is unknown. The knowledge of the human motion trajectory can help the robot make early motion planning and perform more natural and smooth actions.

The human always moves with considerable uncertainty, even to a specific task, the human motion trajectories varies, based on this issue, a method of human trajectory prediction in HRC is proposed in this paper, in a human-robot handover task (see Figure 1), the distances between the human hand and the robot end-effector alone three axis of coordinates are assumed to be Gaussian Processes as time goes on, the human motion trajectories can be described by the combination of the three Gaussian Processes, experimental results show that the proposed method can effectively predict the human trajectories.

2. Proposed Approach

2.1. Task description

In many collaborative tasks, such as the human-robot object handover, the human and the robot should move close to each other and make physical contact, the task can be divided into three phase:

**Phase 1: Preparing**
- **Task Description:** The human takes the bottle from the table
- **Collaboration Status:** No collaboration

**Phase 2: Approaching**
- **Task Description:** The robot moves its left hand to a fixed place near the human, the human moves the bottle close to the robot
- **Collaboration Status:** Collaboration

**Phase 3: Handing over**
- **Task Description:** The human puts the bottle in the robot’s hand and the robot grasps it
- **Collaboration Status:** Physical collaboration

Figure 2. The overall task flow

In the first phase (see Figure 2), the human holds the bottle from the table, the initial position of the bottle on the table varies in each demonstration. In the second phase, the robot moves its hand to a fixed place which is in the scope of the human hand's movement space, the human moves the bottle close to the robot end-effector, a Microsoft Kinect is used to track the motion trajectory of the human during the entire task. In the last step, the robot grasps the bottle.
2.2. The Gaussian Process (GP)
A series of pre-experiments show that the human moves with considerable random, even for a same task—the human moves a bottle from a fixed position to another fixed position, the trajectories usually vary. The Gaussian Process is used in this paper to predict the human motion trajectories, in which the human hand position probability, not the accurate position, is predicted. It is assumed that at time $t_i$, the distance between the human hand and the robot end-effector $f(t)$ obey the Gaussian distribution, and the change of human–robot distance with time obey the Gaussian Process:

$$
\begin{bmatrix}
f(t_1) \\
\vdots \\
f(t_n)
\end{bmatrix} \sim N
\begin{bmatrix}
m(t_1) \\
\vdots \\
m(t_n)
\end{bmatrix}
, 
\begin{bmatrix}
k(t_1, t_1) & \cdots & k(t_1, t_n) \\
\vdots & \ddots & \vdots \\
k(t_n, t_1) & \cdots & k(t_n, t_n)
\end{bmatrix}
$$

(1)

Where $m(x_i)$ is the mean value of the human–robot distance at time $t_i$ and $k(\cdot, \cdot)$ is the covariance matrix. At time $t_i$, the human–robot distance $f(t_i)$ obey a Gaussian distribution with an expectation of $m(t_i)$ and a variance of $k(t_i, t_i)$. And in different time $t_i$ and $t_j$, the $f(t_i)$ and $f(t_j)$ are not statistically independent, the $k(t_i, t_j)$ (not zero) is used to represent the mutual relationship.

3. Algorithms
Our approach to predicting human motion trajectories consists of three phases (see Figure 3). In the first phase we provide demonstrations of human–robot handover and gather data (the position trajectories of the human and the robot), the data is processed in the second phase, including the coordinate transformation and the data segmentation, in the third phase the training data is used for the Gaussian process regression, and then the GP is used to predict the human trajectories.

3.1. Data Collection
The human–robot handover demonstrations are performed by a human and a collaborative robot, A Microsoft Kinect was used to track the motion trajectory of the human (the data acquisition cycle is 0.03s). The bottle is placed on the table, the human holds the bottle and moves it to the robot, 30 demonstrations were made, the initial positions of the bottle on the table, and the positions of the robot end-effector are different in each demonstration.

3.2. Data Process
**Data segmentation:** A total of 30 groups of experiments were conducted, 20 of which were used for training, and the others were used for testing. In this paper, the relative distance between the human and the robot along three coordinate axes were studied separately.

3.3. Data training and testing
**Gaussian Process Regression:** At any time step, the mean distance between the human hand and the
At time $t_i$, $n_i$ is the total number of data collected, $d_{ij}$ is the human-robot distance of each point. There is a strong correlation between the human-robot distance and time, so the covariance at different time is not zero, a Radial Basis Function (RBF) is used to represent the covariance matrix. The first reason is that the covariance matrix is required to be semi-positive definite matrices, the RBF meets that need, the other is that the RBF providing a fast approximation of the covariance matrix [23], has good generalization ability and a fast learning convergence speed:

\[
    k(t_i, t_j) = \delta^2 e^{-\frac{(t_i-t_j)^2}{2\tau^2}}
\]

Where $k(t_i, t_j)$ are the covariance of human-robot distance between time $t_i$ and $t_j$. $\delta$ and $l$ are the parameters to be iterated optimization, the constraint condition is to maximize $p(Y|X)$ while $f(x)$~$N(m(x), k(x,x^T))$, the object function is set to:

\[
    logp(Y|X) = logN(\mu, K)
\]

**Prediction:** the mean value and the covariance matrix are calculated based on the algorithms above and the training data, and the distance between the human and the robot at the other time also subordinate to the Gaussian Process, so:

\[
    \begin{bmatrix} f \end{bmatrix} \sim N\left( \begin{bmatrix} m \end{bmatrix}, \begin{bmatrix} k & k^* \\ k^* & k^{**} \end{bmatrix} \right) \]

Where $f$ are the distance at the other time, $k^*$ and $k^{**}$ are the covariance matrix which can be calculated by (5).

\[
    f(t_i)^{upper} = m(t_i) + 1.96 \cdot \sqrt{k(t_i, t_i)}
\]

\[
    f(t_i)^{lower} = m(t_i) - 1.96 \cdot \sqrt{k(t_i, t_i)}
\]

Where $f(t_i)^{upper}$ and $f(t_i)^{lower}$ are the upper and lower limits of the 95% confidence interval respectively.

4. Experiments and Results

Figure 4. some of the measured trajectories and the predicted trajectories. The pink and thick curves represent the measured trajectories and the blue and thin curves represent the predicted trajectories.

Three of the tested trajectories is shown in a virtual 3D scenario (see Figure 4), as can be shown, two of the prediction trajectories coincidence with the measured trajectory, the other shows a relative big difference.
In all 10 testing trajectories (see Figure 5), 8 of the trajectories (about 80%) are included in the 95% confidence interval, the remaining shows a relatively bigger difference. On the whole, the 95% confidence interval include most of the trajectories, which can be used as a prediction interval of human trajectories in the human-robot handover.

5. Conclusion
An approach of human trajectories prediction is proposed in this paper, in which the human motion trajectories along three coordinate axes are fitted into Gaussian Process respectively, and the Gaussian Processes are then used for prediction. Experimental results show that the method can make prediction of human motion trajectories in a relatively high accuracy. The method and results presented in this paper could be extended to other human-robot collaborative tasks when the human and the robot move closer to each other.

Acknowledgments
We acknowledge the support received for the Manned Aerospace Research Project of China (Grant No. 060601).

References
[1] Mainprice J, Berenson D. Human-robot collaborative manipulation planning using early prediction of human motion[C]//2013 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2013: 299-306.
[2] Zhou T, Wachs J P. Early prediction for physical human robot collaboration in the operating room[J]. Autonomous Robots, 2018, 42(5): 977-995.
[3] Ajoudani A, Zanchettin A M, Ivaldi S, et al. Progress and prospects of the human–robot collaboration[J]. Autonomous Robots, 2018, 42(5): 957-975.
[4] Ragaglia M, Zanchettin A M, Rocco P. Safety-aware trajectory scaling for human-robot collaboration with prediction of human occupancy[C]//2015 International Conference on Advanced Robotics (ICAR). IEEE, 2015: 85-90.
[5] Luo R, Hayne R, Berenson D. Unsupervised early prediction of human reaching for human–robot collaboration in shared workspaces[J]. Autonomous Robots, 2018, 42(3): 631-648.
[6] Lawitzky M, Medina J R, Lee D, et al. Feedback motion planning and learning from demonstration in physical robotic assistance: differences and synergies[C]//2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2012: 3646-3652.
[7] Nikolaidis S, Kuznetsov A, Hsu D, et al. Formalizing human-robot mutual adaptation: A bounded memory model[C]//The Eleventh ACM/IEEE International Conference on Human Robot Interaction. IEEE Press, 2016: 75-82.
[8] Kong Y, Fu Y. Human action recognition and prediction: A survey[J]. arXiv preprint arXiv:1806.11230, 2018.
[9] Liu H, Wang L. Human motion prediction for human-robot collaboration[J]. Journal of Manufacturing Systems, 2017, 44: 287-294.
[10] Huang C M, Andrist S, Sauppé A, et al. Using gaze patterns to predict task intent in collaboration[J]. Frontiers in psychology, 2015, 6: 1049.
[11] Pellegrinelli S, Admoni H, Javdani S, et al. Human-robot shared workspace collaboration via hindsight optimization[C]//2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2016: 831-838.
[12] Luber M, Stork J A, Tipaldi G D, et al. People tracking with human motion predictions from social forces[C]//2010 IEEE International Conference on Robotics and Automation. IEEE, 2010: 464-469.
[13] Nakanishi Y, Yanagisawa T, Shin D, et al. Mapping ECoG channel contributions to trajectory and muscle activity prediction in human sensorimotor cortex[J]. Scientific reports, 2017, 7: 45486.
[14] Nakanishi Y, Yanagisawa T, Shin D, et al. Prediction of three-dimensional arm trajectories based on ECoG signals recorded from human sensorimotor cortex[J]. PloS one, 2013, 8(8): e72085.
[15] Mainprice J, Hayne R, Berenson D. Goal set inverse optimal control and iterative replanning for predicting human reaching motions in shared workspaces[J]. IEEE Transactions on Robotics, 2016, 32(4): 897-908.
[16] Mainprice J, Hayne R, Berenson D. Predicting human reaching motion in collaborative tasks using inverse optimal control and iterative re-planning[C]//2015 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2015: 885-892.
[17] Alahi A, Goel K, Ramanathan V, et al. Social lstm: Human trajectory prediction in crowded spaces[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 961-971.
[18] Cheng Y, Zhao W, Liu C, et al. Human motion prediction using adaptable neural networks[J]. arXiv preprint arXiv:1810.00781, 2018.
[19] Fernando T, Denman S, Sridharan S, et al. Soft+ hardwired attention: An lstm framework for human trajectory prediction and abnormal event detection[J]. Neural networks, 2018, 108: 466-478.
[20] Vogt D, Stepputti S, Grehl S, et al. A system for learning continuous human-robot interactions from human-human demonstrations[C]//2017 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2017: 2882-2889.
[21] Luo R, Berenson D. A framework for unsupervised online human reaching motion recognition and early prediction[C]//2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2015: 2426-2433.
[22] Pérez-D’Arpino C, Shah J A. Fast target prediction of human reaching motion for cooperative human-robot manipulation tasks using time series classification[C]//2015 IEEE international conference on robotics and automation (ICRA). IEEE, 2015: 6175-6182.
[23] Curtó J D, Zarza H C, Kim T. MKernel: Approximate Kernel Expansions in Log-linear Time through Randomization[J]. arXiv preprint arXiv:1702.08159, 2017.