
Phase Estimation for Fast Action Recognition and Trajectory Generation in Human-Robot Collaboration

Journal Title
XX(X):1–16
©The Author(s) 2015
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/


Guilherme Maeda¹, Marco Ewerton¹, Gerhard Neumann², Rudolf Lioutikov¹ and Jan Peters^{1,3}

Abstract

This paper proposes a method to achieve fast and fluid human-robot interaction by estimating the progress of the movement of the human. The method allows the progress, also referred to as the phase of the movement, to be estimated even when observations of the human are partial and occluded; a problem typically found when using motion capture systems in cluttered environments. By leveraging on the framework of Interaction Probabilistic Movement Primitives (ProMPs), phase estimation makes it possible to classify the human action, and to generate a corresponding robot trajectory before the human finishes his/her movement. The method is therefore suited for semi-autonomous robots acting as assistants and coworkers. Since observations may be sparse, our method is based on computing the probability of different phase candidates to find the phase that best aligns the Interaction ProMP with the current observations. The method is fundamentally different from approaches based on Dynamic Time Warping (DTW) that must rely on a consistent stream of measurements at runtime. The phase estimation algorithm can be seamlessly integrated into Interaction ProMPs such that robot trajectory coordination, phase estimation, and action recognition can all be achieved in a single probabilistic framework. We evaluated the method using a 7-DoF lightweight robot arm equipped with a 5-finger hand in single and multi-task collaborative experiments. We compare the accuracy achieved by phase estimation with our previous method based on DTW.

Keywords

Class file, $\text{\LaTeX} 2_{\mathcal{E}}$, SAGE Publications

1 Introduction

Assistive and collaborative robots must have the ability to physically interact with the human, safely and synergistically. However, pre-programming a robot for a large number of tasks is not only tedious, but unrealistic, especially if tasks are added or changed constantly. Moreover, conventional programming methods do not address semi-autonomous robots—robots whose actions depend on the actions of a human partner. Once deployed, for example in a domestic or small industrial environment, a semi-autonomous robot must be easy to program, without requiring the need of a dedicated expert. For this reason, this paper proposes the use of interaction learning, a data-driven approach based on the use of imitation learning (Schaal 1999) for learning tasks that involve human-robot interaction. In this paper, we exploit the benefits of augmenting the interaction learning method with a temporal model of the distribution of phases of a human movement.

An important aspect of collaborative robots is the ability to recognize the action of the human, and to quickly

generate a corresponding robot trajectory that matches the predicted human trajectory. As illustrated in Figure 1, by observing the movement of the human, a robot partner must not only decide if it should hand over a screw, a plate, or hold a screwdriver; but once the action is recognized, the robot must spatially coordinate its trajectory w.r.t the human movement. Probabilistic models have been used to address either action recognition or trajectory generation, but for their realization, many methods require the time alignment of training data such that the spatial correlation can be properly captured (Calinon et al. 2007; Ye and Alterovitz

¹ Intelligent Autonomous Systems Group, TU Darmstadt, Germany

² Computational Learning for Autonomous Systems Group, TU Darmstadt, Germany

³ Max Planck Institute for Intelligent Systems, Tuebingen, Germany

Corresponding author:

Guilherme Maeda, Intelligent Autonomous Systems Group, TU Darmstadt, Germany

Email: maeda@ias.tu-darmstadt.de

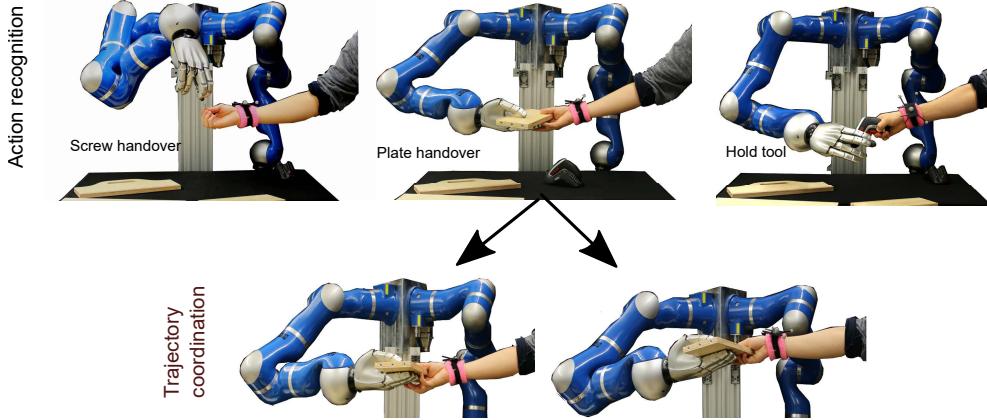


Figure 1. A robot coworker must recognize the intention of the human before deciding which action to take. In this case, to hand over a screw, to hand over a plate, or to hold the tool for the human partner. Once the action decision is made, the robot must then coordinate its trajectory with the trajectory of the human, such that a handover of a plate is successful, for example. The goal of this paper is to allow the robot to quickly take such actions at the early stages of the human movement, possibly under partial occlusion of the human movement.

2011; Dong and Williams 2012; Ben Amor et al. 2014; Perez-D’Arpino and Shah 2015). Under severe changes in the duration of human trajectories, such models require reliable measurements of positions for online alignment of the observations before the interaction model can be queried. This requirement poses a problem of practical importance since occlusions and interrupted streams of position are prone to occur in many of the collaborative environments of interest such as hospitals, homes and factories. Other methods will be discussed in the related work section.

Instead of systematically computing distances between trajectories for an exact time-alignment as in dynamic time warping (DTW), the principle of our proposed method is to test phase candidates by computing the likelihood of models with different durations. This approach allows for querying the model with a minimum number of observations. The general idea is illustrated in Figure 2 where training data collected with different durations are first normalized to a single nominal duration. A probabilistic representation of movement primitives (Paraschos et al. 2013) is used to learn the parameters of the distribution under the normalized time. During execution, the method then computes temporal variations of the model, which are governed by the phase. From Figure 2, it is clear that given the same observations, the model with duration T_1 is more likely to represent the current human motion than the model with duration T_2 .

To achieve coordination with a robot partner, the method of Interaction ProMPs are used (Maeda et al. 2016), which allows the robot trajectory to be generated in conjunction with the prediction of the human motion. Basically, an Interaction ProMP provides a model that correlates the

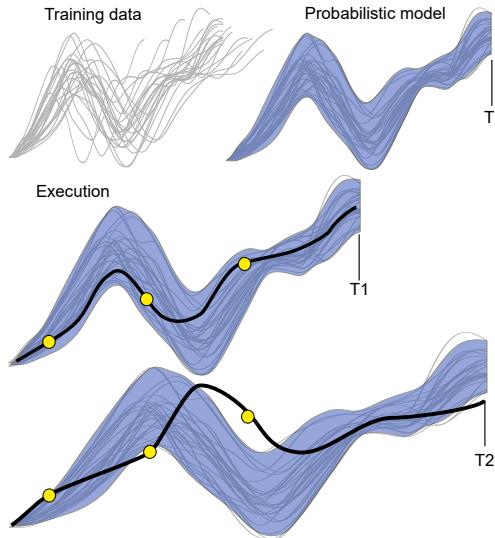


Figure 2. A cartoon of the phase estimation problem: under sparse observations there may be many ways in which the observation can be explained by variants of a model that differ by a temporal scaling. Each variation leads to different human prediction and collaborative robot trajectories. The most probable time scaling is assumed to generate the correct human-robot coordination.

parameters that describe elementary trajectories of a human and a robot when executing a task in collaboration. Using a probabilistic treatment, the trajectory of the robot is generated by inference, where the Interaction ProMP is conditioned on the observations of the human. The robot states are subsequently sampled from the resulting posterior distribution.

The contribution of this paper is a single probabilistic framework that allows a robot to estimate the phase of the human movement online, and to associate the outcome of the estimation to address different tasks and their respective robot motions. Phase estimation not only allows the robot to react faster—as the trajectory inference can be done with partial and even occluded observations—but also eliminates the need for time-alignment. An initial version of this paper was presented in Maeda et al. (2015) where the algorithm for phase estimation was introduced. Here, we provide a more elaborated an extended version with new experimental results and a significant number of additional evaluations. The full probabilistic framework including its relation with Interaction ProMPs is also explained in more detail.

This paper is organized as follows. Section 2 addresses the relevant literature related to the framework with an emphasis on phase and time representations. Section 3 describes the proposed method with a brief background on ProMPs, followed by Interaction ProMPs, phase estimation, and action recognition. Finally, Section 4 provides experiments and discussions on the application of the method in a single-task handover experiment and in a multi-task assembly scenario.

2 Related Work

This section initially discusses approaches that address the problem of robot movement generation or human action recognition under temporal variations. The usual approach is based on time alignment of trajectories. Other works, however, address time as transition probabilities, or consider phase as a dynamical system. An important difference with our work, is that the great majority either address the problem of movement generation or action recognition, thus lacking a principled, unified way to connect these two problems. Due to the practical application, this section also briefly discusses the handover literature.

Dynamic Time Warping (DTW) (Sakoe and Chiba 1978) has been used in many robotics applications for temporally aligning trajectories. For example, as part of an algorithm that estimates the optimal, hidden trajectory provided by multiple expert demonstrations (Coates et al. 2008) or as a pre-processing step for the generation of probabilistic representations, notably Gaussian Mixture Models (GMM) (Calinon et al. 2007; Ye and Alterovitz 2011) but also other forms of representations (Dong and Williams 2012; Perez-D’Arpino and Shah 2015). Although DTW is suitable for off-line processing of data, its online application is challenging as its principle mechanism relies on systematic distance computations between full trajectories. Dong and Williams (2012) and Ben Amor et al. (2014) presented DTW formulations for online applications useful when the stream of data is consistent.

Under partial/occluded observations, as illustrated in Figure 2, the use of DTW becomes impractical.

Calinon et al. (2007, 2012) propose explicitly encoding time dependency in a mixture model. Thus, by directly conditioning the model on time, smooth temporal solutions using Gaussian Mixture Regression can be achieved. In our work, however, observations are provided by the human—as opposed to the movements of a controlled robot. Due to the many ways the human can change the speed, the time index does not reflect the phase of the movement, hindering the possibility to condition the model directly. In fact, an important difference of our work is the coordination between human and robot, rather than the single agent robot scenario. Hidden Markov Models (HMMs) have also been used in (Calinon et al. 2006) and in (Lee and Ott 2011) particularly to avoid the pre-processing of data during online execution. While the temporal information is addressed by the transition probabilities of the HMM, HMMs alone does not suffice to completely represent continuous distribution of trajectories, often requiring additional mechanisms to overcome its discrete nature.

The temporal alignment or the phase estimation problem can be alleviated when velocities or the goal of the human trajectory are known. The measurement or estimation of velocity, for example, by differentiation of a consistent stream of positions, eliminates the ambiguity illustrated in Figure 2 and allows for the realization of online algorithms that cope with very fast dynamics (Kim et al. 2010, 2014). Such methods, however, rely on a planned environment free from occlusions and fast tracking capabilities; requirements difficult to achieve in environments where semi-autonomous robots are expected to make their biggest impact, such as in small factories, hospitals and home care facilities. Englert and Toussaint (2014) presented a method to reactively adapt trajectories of a motion planner due to changes in the environment. This was achieved by measuring the progress of a task with a dynamic phase variable. While this method is suited for cases where the goal is known—as the phase is estimated from the distance to the goal—a semi-autonomous robot is not provided with such information: the goal must be inferred from the observation of the human movement, which in turn requires an estimate of the phase.

Dynamical Movement Primitives (DMPs) have the ability to modulate temporal variations with a phase variable (Ijspeert et al. 2013). The phase variable is used to govern the spread of a fixed number of basis functions that encode parameters of a forcing function. Recently, a modified form of DMPs where the rate of phase change is related to the speed of movement has been presented by Vuga et al. (2014). The method uses Reinforcement Learning and Iterative Learning Control to speed up the

execution of a robot's movement without violating pre-defined constraints such as carrying a glass full of liquid without spilling it. A similar form of iterative learning was used to learn the time mapping between demonstrated trajectories and a reference trajectory (Van Den Berg et al. 2010). With their approach, a robot was able to perform a surgical task of knot-tie faster than the human demonstrator. Although such methods exploit phase or durations to adapt a velocity profile, they do not address the inverse problem of estimating the phase itself. ProMPs use the concept of phases in the same manner as DMPs, with the difference that the basis functions are used to encode positions. This difference is fundamental for the tractability of Interaction Primitives since estimating the forcing function of the human (a DMP requirement) is nontrivial in practice.

Several works have addressed the action recognition problem. Graphical models, in particular, have been widely used. In human-robot interaction, HMMs have been used hierarchically to represent states and to trigger low-level primitives (Lee et al. 2010). HMMs were also applied to predict the positions of a coworker in an assembly line for tool delivery (Tanaka et al. 2012) while in Koppula and Saxena (2013), Conditional Random Fields were used to predict the possible actions of a human. The prediction of actions and movements of human coworkers have been addressed by many authors using probabilistic approaches to model trajectories (Dong and Williams 2012; Mainprice and Berenson 2013; Perez-D'Arpino and Shah 2015). In common, the cited methods do not explicitly address robot trajectory generation as part of the model, usually treating the design of the robot motion as an independent step that must be executed once the action is recognized. Robot trajectories were pre-programmed for the recognized actions Koppula and Saxena (2013), or generated with motion planners Mainprice and Berenson (2013); Hayne et al. (2016). We exploit the use of correlated movement primitives to generate the robot trajectory such that action recognition and movement generation are both given by the same probabilistic model and solved by similar computations.

Although some of the scenarios here presented ultimately lead to the handover of objects, handovers are not the only application of Interaction ProMPs and our method is not intend to be a self-contained solution to the whole handover problem. A handover is comprised of a complex series of combined physical and social interaction steps. As previously investigated by Strabala et al. (2013), these steps range from (1) the social-cognitive cues that establish the connection between the giver and the taker, (2) the coordination of the location and the resulting trajectory as a function of preferences and socially acceptable movements (Sisbot and Alami 2012), and (3) the final physical transfer that comprises interaction forces and compliances

(Kupcsik et al. 2015). While Interaction ProMPs does not encode or output trajectories based on such vast amount of information, as an imitation learning method, they implicitly encode user preferences from demonstrations which appears to be more suited for human interaction than pure motion planning approaches (Cakmak et al. 2011). Human-robot handover offers an interesting application for methods that emphasize the fast response of the robot, and therefore, it suits our investigations in phase estimation. In the handover context and under the assumption of a single task, this paper shares similar challenges faced in Yamane et al. (2013) where the robot trajectory had to be coordinated according to the observation of the human partner during the passing of an object. Yamane et al. (2013) encoded the demonstrations in a tree-structured database as a hierarchy of clusters, which then poses the problem of searching matching trajectories given partial observations using a sliding window. In contrast, our method uses a flat representation which requires less computation while allowing the recognition of different tasks.

This paper uses Gaussian distributions to encode the joint distribution of human and robot movement primitives over *entire* trajectories. This makes action recognition straightforward since each Gaussian represents one task. Also, this representation allows to quickly infer the most appropriate robot movement primitive by conditioning the joint distribution on the human task. Note that this is fundamentally different from the GMMs approach that encode *local* variations across demonstrated trajectories (e.g. Calinon et al. (2012)). In the body of work of Calinon et al., the flexibility provided by multiple components can be exploited, for example, by a stiffness controller which adapts the robot behavior as a function of the local uncertainty over the distribution of demonstrated trajectories.

This paper consolidates our recent efforts in different aspects of semi-autonomous robots. It leverages on our developments in human-robot interaction (Ben Amor et al. 2014) and the ability to address multiple tasks (Ewerthon et al. 2015b; Maeda et al. 2016). While our previous interaction models were explicitly time-dependent and reliant on DTW, here, we introduce a phase-dependent method that is free from time alignment processing and allows for fast robot action recognition and trajectory generation.

3 Probabilistic Movement Primitives for Human-Robot Interaction

This section introduces the basic concepts of ProMPs on a one dimensional case, followed by the multi-dimensional ProMP case. As it will become clear, there is a natural

transition from a multi-dimensional ProMP to the human-robot case with Interaction ProMPs. This section also addresses how to compute the probabilities of a task when many Interaction ProMPs are used to represent different tasks. This section finishes by introducing phase estimation, which also provides means to recognize human actions in multiple-task scenarios under partial human observation.

3.1 Probabilistic Movement Primitive for a Single Dimension

For each time step t a position is represented by y_t and a trajectory of T time steps as a sequence $\mathbf{y}_{1:T}$. A parameterization of y_t in a lower dimensional weight space is proposed as

$$y_t = \psi_t^T \mathbf{w} + \epsilon_y, \quad (1)$$

$$p(\mathbf{y}_{1:T} | \mathbf{w}) = \prod_1^T \mathcal{N}(\mathbf{y}_t | \psi_t^T \mathbf{w}, \sigma_y), \quad (2)$$

where $\epsilon_y \sim \mathcal{N}(0, \sigma_y)$ is zero-mean i.i.d. Gaussian noise, $\mathbf{w} \in \mathbb{R}^N$ is a weight vector that parameterizes the trajectory, and $\psi_t = [(\psi_t)_1, \dots, (\psi_t)_N]^T \in \mathbb{R}^{N \times 1}$ has the values of each of the basis function at time t . The weights are computed with linear regression with a time-dependent design matrix

$$\mathbf{w} = (\Psi_{1:T}^T \Psi_{1:T})^{-1} \Psi_{1:T}^T \mathbf{y}_{1:T} \quad (3)$$

where

$$\Psi_{1:T} = \begin{bmatrix} (\psi_1)_1 & \dots & (\psi_1)_N \\ \vdots & \ddots & \vdots \\ (\psi_T)_1 & \dots & (\psi_T)_N \end{bmatrix}, \quad (4)$$

The number of Gaussian bases N is often much lower than the number of trajectory time steps*. The number of bases is a design parameter that must be matched with the desired amount of detail to be preserved during the encoding of the trajectory.

Assume M trajectories are obtained via demonstrations; their parameterization leading to a set of weight vectors $\mathbf{W} = \{\mathbf{w}_1, \dots, \mathbf{w}_i, \dots, \mathbf{w}_M\}$ (the subscript i as in \mathbf{w}_i will be used to indicate a particular demonstration when relevant, and will be omitted otherwise). Define $\boldsymbol{\theta}$ as a parameter to govern the distribution of the weight vectors in the set \mathbf{W} such that $\mathbf{w} \sim p(\mathbf{w}; \boldsymbol{\theta})$. The model $p(\mathbf{w}; \boldsymbol{\theta})$ is assumed as a Gaussian with mean $\boldsymbol{\mu}_w \in \mathbb{R}^N$ and covariance $\boldsymbol{\Sigma}_w \in \mathbb{R}^{N \times N}$, that is $\boldsymbol{\theta} = \{\boldsymbol{\mu}_w, \boldsymbol{\Sigma}_w\}$. This prior model allows us to sample trajectories with

$$\begin{aligned} p(y_t; \boldsymbol{\theta}) &= \int p(y_t | \mathbf{w}) p(\mathbf{w}; \boldsymbol{\theta}) d\mathbf{w} \\ &= \mathcal{N}(y_t | \psi_t^T \boldsymbol{\mu}_w, \psi_t^T \boldsymbol{\Sigma}_w \psi_t + \sigma_y). \end{aligned} \quad (5)$$

The Gaussian assumption is restrictive in the sense that the training data must be time-aligned, for example by DTW, such that the spatial correlation can be captured appropriately. In Section 3.4 we will introduce a method to avoid such alignment.

3.2 Correlating Human and Robot Movements with Interaction ProMPs

Interaction ProMPs model the correlation of multiple dimensions, here each dimension given by each of the degrees-of-freedom (DoFs) of multiple agents. Let us define the state vector as a concatenation of the P number of observed DoFs of the human, followed by the Q number of DoFs of the robot

$$\mathbf{y}_t = [y_{1,t}^H, \dots, y_{P,t}^H, y_{1,t}^R, \dots, y_{Q,t}^R]^T, \quad (6)$$

where the upper scripts $(\cdot)^H$ and $(\cdot)^R$ refer to the human and robot DoFs, respectively. Similar to the one dimensional case, trajectories of each DoF are parameterized as weights such that

$$p(\mathbf{y}_t | \bar{\mathbf{w}}) = \mathcal{N}(\mathbf{y}_t | \mathbf{H}_t^T \bar{\mathbf{w}}, \boldsymbol{\Sigma}_y), \quad (7)$$

where $\mathbf{H}_t^T = \text{diag}((\psi_t^T)_1, \dots, (\psi_t^T)_P, (\psi_t^T)_1, \dots, (\psi_t^T)_Q)$ has $P+Q$ diagonal entries. Each collaborative demonstration now provides $P+Q$ training trajectories. The weight vector $\bar{\mathbf{w}}_i$ of the i -th demonstration is now a concatenation of all weight vectors involved in the i -th demonstration. Thus, the many DoFs involved in the interaction will be correlated

$$\bar{\mathbf{w}}_i = [(w_1^H)^T, \dots, (w_P^H)^T, (w_1^R)^T, \dots, (w_Q^R)^T]^T. \quad (8)$$

A normal distribution from a set of M demonstrations $\bar{\mathbf{W}} = \{\bar{\mathbf{w}}_1, \dots, \bar{\mathbf{w}}_M\}$ with $\boldsymbol{\mu}_w \in \mathbb{R}^{(P+Q)N}$ and $\boldsymbol{\Sigma}_w \in \mathbb{R}^{(P+Q)N \times (P+Q)N}$ can be computed.

The fundamental operation to infer the robot trajectory is to compute a posterior probability distribution of the weights $\bar{\mathbf{w}} \sim \mathcal{N}(\boldsymbol{\mu}_w^{new}, \boldsymbol{\Sigma}_w^{new})$ conditioned on the observations of the human. Since the robot is not observed we denote the observation vector as

$$\mathbf{y}_t^o = [y_{1,t}^H, \dots, y_{P,t}^H, 0_{1,t}^R, \dots, 0_{Q,t}^R]^T. \quad (9)$$

To contrast with a complete sequence $[t : t']$ we use the notation $[t - t'] \in \mathbb{R}^{S \times P}$ to indicate that the sequence of S observations in the interval is incomplete, that is, some measurements in between are missing. The

*In the particular case of the experiments here reported, trajectories have an average time of 3 seconds, sampled at 50 Hz. We used 20 basis functions and thus, for a 1 DoF case the dimensionality is decreased from $3 \times 50 = 150$ samples to a weight vector $\mathbf{w} \in \mathbb{R}^{20}$.

Interaction ProMP is updated with $\mathbf{y}_{t-t'}^o$ using closed-form conditioning

$$\begin{aligned}\boldsymbol{\mu}_w^{new} &= \boldsymbol{\mu}_w + \mathbf{K}(\mathbf{y}_{t-t'}^o - \mathbf{H}_{t-t'}\boldsymbol{\mu}_w), \\ \boldsymbol{\Sigma}_w^{new} &= \boldsymbol{\Sigma}_w - \mathbf{K}(\mathbf{H}_{t-t'}\boldsymbol{\Sigma}_w),\end{aligned}\quad (10)$$

where $\mathbf{K} = \boldsymbol{\Sigma}_w \mathbf{H}_{t-t'}^T (\boldsymbol{\Sigma}_y^o + \mathbf{H}_{t-t'} \boldsymbol{\Sigma}_w \mathbf{H}_{t-t'}^T)^{-1}$ and $\boldsymbol{\Sigma}_y^o$ is the measurement noise. The upper-script $(\cdot)^{new}$ is used for values after the update. The observation matrix $\mathbf{H}_{t-t'}$ is obtained by concatenating the bases at the corresponding observation steps, where the Q unobserved states of the robot are represented by zero entries in the diagonal. For a single observation at time t ,

$$\mathbf{H}_t = \left[\begin{array}{ccc|cc} (\psi_t^T)_1 & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & (\psi_t^T)_P & \mathbf{0} & \dots & \mathbf{0} \\ \hline \mathbf{0} & \dots & \mathbf{0} & \mathbf{0}_P & \dots & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0}_Q \end{array} \right] \quad (11)$$

with $\mathbf{H}_t \in \mathbb{R}^{(P+Q) \times (P+Q)N}$.

Trajectory distributions that predict human and robot movements are obtained by integrating out the weights of the posterior distribution

$$p(\mathbf{y}_{1:T}; \boldsymbol{\theta}^{new}) = \int p(\mathbf{y}_{1:T} | \bar{\mathbf{w}}) p(\bar{\mathbf{w}}; \boldsymbol{\theta}^{new}) d\bar{\mathbf{w}}. \quad (12)$$

3.3 Multiple Interaction Patterns

To address multiple tasks we repeat the procedure described in Section 3.2 for each task. The set of demonstrated trajectories for each type of collaboration is labeled for action recognition purposes and are trained independently. Given that a K number of tasks are presented, and given the observation vector as defined in (9) the probability of each task k can be computed with

$$\begin{aligned}p(k | \mathbf{y}_{t-t'}^o) &\propto p(\mathbf{y}_{t-t'}^o | k) p(k) \\ &\propto p(\mathbf{y}_{t-t'}^o | \boldsymbol{\theta}_k) p(k),\end{aligned}\quad (13)$$

where $p(k)$ is a prior probability of the task. The action is recognized by selecting the posterior with the highest probability

$$k^* = \arg \max_k p(k | \mathbf{y}_{t-t'}^o), \quad (14)$$

whose corresponding model is $\boldsymbol{\theta}_{k^*}^{new}$.

To compute the likelihood in (13) note that each component k is governed by $\boldsymbol{\theta}_k$, therefore

$$\begin{aligned}p(\mathbf{y}_{t-t'}^o; \boldsymbol{\theta}_k) &= \\ &= \int p(\mathbf{y}_{t-t'}^o | \mathbf{H}_{t-t'}^T \bar{\mathbf{w}}, \boldsymbol{\Sigma}_y^o) p(\bar{\mathbf{w}}; \boldsymbol{\theta}_k) d\bar{\mathbf{w}} \\ &= \mathcal{N}(\mathbf{y}_{t-t'}^o | \mathbf{H}_{t-t'}^T \boldsymbol{\mu}_w, \mathbf{H}_{t-t'}^T \boldsymbol{\Sigma}_w \mathbf{H}_{t-t'} + \boldsymbol{\Sigma}_y^o).\end{aligned}\quad (15)$$

The most likely robot trajectory given the human action is inferred by conditioning the most probable interaction model $\boldsymbol{\theta}_{k^*}^{new}$ with (10) on $\mathbf{y}_{t-t'}^o$ and by integrating out the weights

$$p(\mathbf{y}_{1:T}; \boldsymbol{\theta}_{k^*}^{new}) = \int p(\mathbf{y}_{1:T} | \bar{\mathbf{w}}) p(\bar{\mathbf{w}}; \boldsymbol{\theta}_{k^*}^{new}) d\bar{\mathbf{w}}. \quad (16)$$

3.4 Estimating Phases on Multiple Tasks

Previous works (Ewerthon et al. 2015b; Maeda et al. 2016) have only addressed spatial variability, but not temporal variability of demonstrated movements. However, when demonstrating the same task multiple times, a human demonstrator will inevitably execute movements at different speeds, thus changing the phase at which events occur. Previously, this problem was mitigated during the training phase by time-aligning the demonstrated trajectories using a DTW method. Time alignment ensures that the weights of each demonstration can be regressed using the same feature $\psi_{1:T}$ for all demonstrations. As a consequence, during execution, the conditioning (10) can only be used when the phase of the human demonstrator coincides with the phase encoded by the time-aligned model, which is unrealistic in practice.

Under temporal variability and unobserved velocities, the problem is to retrieve the corresponding basis ψ_t for an observation \mathbf{y}_t^o such that the conditioning can be computed correctly. While time alignment allows for the encoding of spatial variability during training, it poses a difficult problem during execution since the observations must be aligned while the human is moving. Dong and Williams (2012) and Ben Amor et al. (2014) have proposed online variants of DTW given stream of positions. Such approaches, however, are not suitable when only a few sparse points are measured as the estimation of distances between trajectory segments is compromised. A practical solution to the alignment problem is to condition only at the final position of the human movement, since only for this particular case, the corresponding basis function is known to be the last one ψ_T . This, however, causes a significant lag on the robot response as the robot has to wait for the human to finish his movement first.

To solve this problem, we propose incorporating the temporal variance as part of the model by learning a distribution over phases from the same demonstrations previously used to create the Interaction ProMP. This enriched model not only eliminates the need for time-alignment, but also allows for faster robot reactions as the conditioning (10) can be applied before the end of the human movement. Initially, we replace the original time indexes of the basis functions with a phase variable $z(t)$. Thus, a trajectory of duration T is now computed relative

to the phase

$$p(\mathbf{y}_{1:T}|\mathbf{w}) = \prod_1^T \mathcal{N}(\mathbf{y}(z_t)|\boldsymbol{\psi}^T(z_t)\mathbf{w}, \Sigma_y). \quad (17)$$

Define a nominal sequence $\{1 : T_{nom}\}$, for example, by taking the average final time of the demonstrations. Assuming that each of the i -th demonstrations has a constant temporal change in relation to the nominal duration, we define a scaling factor

$$\alpha_i = T_i/T_{nom}, \quad (18)$$

such that all demonstrations can be indexed by the same nominal time index. We then define fixed Gaussian bases spread over the nominal duration $\boldsymbol{\psi}_{1:T_{nom}}$. The weights of all demonstrations can then be regressed from the same bases to obtain the parameters of the distribution $\theta = \{\mu_w, \Sigma_w\}$ using a single design matrix with nominal duration $\Psi_{1:T_{nom}}$.

Given a time step t from the nominal sequence, a trajectory with different duration is found by using $z_t = \alpha t$ in (17). This temporal model implicitly assumes that the human movement is governed by a single phase and therefore, we will refer to α as a phase ratio. Note from (18) that each i -th demonstration in the training set provides a value of α_i . We will assume that the phase ratios of different demonstrations vary according to a normal distribution, that is, $\alpha \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)$, where

$$\begin{aligned} \mu_\alpha &= \text{mean}(\{\alpha_1, \dots, \alpha_i, \dots, \alpha_M\}), \\ \sigma_\alpha^2 &= \text{var}(\{\alpha_1, \dots, \alpha_i, \dots, \alpha_M\}), \end{aligned}$$

and M is the number of demonstrations. Despite the simplicity of this model, experiments have shown that these assumptions (single phase, normal distribution) hold in practice for simple, short stroke movements typical of handovers[†].

Given a sparse partial sequence of observations $\mathbf{y}_{t-t'}^o$, a posterior probability distribution over phases is given as

$$p(\alpha|\mathbf{y}_{t-t'}^o, \theta) \propto p(\mathbf{y}_{t-t'}^o|\alpha, \theta)p(\alpha), \quad (19)$$

where $p(\alpha)$ is the prior probability of the phase.

For a specific α value the likelihood is

$$\begin{aligned} p(\mathbf{y}_{t-t'}^o|\alpha, \theta) &= \int p(\mathbf{y}_{t-t'}^o|\bar{\mathbf{w}}, \alpha)p(\bar{\mathbf{w}})d\bar{\mathbf{w}} \\ &= \mathcal{N}(\mathbf{y}_{t-t'}^o|\mathbf{A}(z_{t-t'})^T \boldsymbol{\mu}_w, \\ &\quad \mathbf{A}(z_{t-t'})^T \Sigma_w \mathbf{A}(z_{t-t'}) + \Sigma_y^o), \end{aligned} \quad (20)$$

where

$$\mathbf{A}(z_{t-t'}) = \begin{bmatrix} \boldsymbol{\psi}(z_{t-t'})_1^T & \dots & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \boldsymbol{\psi}(z_{t-t'})_P^T \end{bmatrix}, \quad (21)$$

is the partition of the full matrix \mathbf{H} in (11) that contains the basis functions of the observed positions of the human—however, now indexed by the phase $z_t = \alpha t$.

Given the observations $\mathbf{y}_{t-t'}^o$, the likelihood of each sampled α candidate is computed with (20), and the most probable value

$$\alpha^* = \arg \max_{\alpha} p(\alpha|\mathbf{y}_{t-t'}^o, \theta) \quad (22)$$

is selected. Intuitively, the effect of different phases is to stretch or compress the temporal axis of the prior (unconditioned) distribution proportionally to α . The method then compares which scaling value generates the model with the highest probability given the observation $\mathbf{y}_{t-t'}^o$. Once the most probable scaling value α^* is found, its associated observation matrix $\mathbf{H}(z)$ can be used in (10) to condition and predict the trajectories of both human and robot. To efficiently estimate the phase during execution, one approach is to sample a number of values of α from the prior $p(\alpha)$ and precompute and store, for each of them, the associated matrices of basis functions $\mathbf{A}(z_{1:T})$ and $\mathbf{H}(z_{1:T})$ beforehand.

Figure 3 summarizes the workflow of the Interaction ProMP with phase estimation. During the training phase, trajectories are scaled to a nominal time T_{nom} and the distribution of the scaling values is encoded as a normal distribution. The set of normalized trajectories are used to learn the parameter $\theta = \{\mu_w, \Sigma_w\}$ that correlates the trajectories of human and robot. In the figure, the distribution modelled by θ is abstracted as a bivariate Gaussian where each of the two dimensions are given by the distribution over the weights of the human and robot trajectories. During execution, the assistive trajectory of the robot is predicted by integrating out the weights of the posterior distribution $p(\bar{\mathbf{w}}; \theta^{new})$. The operation of conditioning is illustrated by the slicing of the prior, at the current observation of the position of the human \mathbf{y}_t^o . The conditioning requires finding the correct temporal scaling of the prior model that best fits the observations in time. Thus, the probability of many phase ratio candidates are tested on the sparse human observations and the most probable value is assumed to provide the optimal time indexing of the observations.

In a multi-task scenario, we can now address the recognition of the task given the positions $\mathbf{y}_{t-t'}^o$ and

[†]For problems where this assumption does not hold, e.g. when the movement accelerates and decelerates along and between demonstrations, the estimation of multiple phases must be addressed. Initial investigations reported in (Ewerthon et al. 2015a) show that movements with multiple phases on synthetic data can be treated under the assumption that different phases must be correlated along the motion. This assumption, however, is too stringent for realistic purposes. We found that real human experiments with short strokes are better addressed by the single phase method presented here.

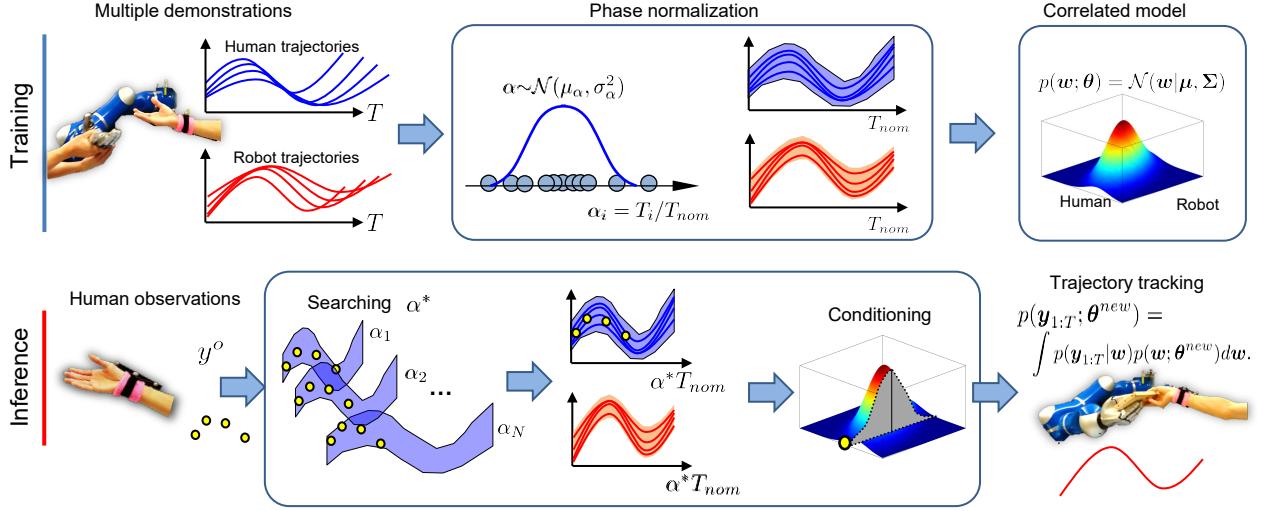


Figure 3. The workflow of Interaction ProMP (after Maeda et al. (2016)) augmented with a step to learn the distribution over phases during training, and a step to infer the phase during execution. The figure illustrates a single task case where the distribution of human-robot parameterized trajectories is abstracted as a bivariate Gaussian. The conditioning step is shown as the slicing of the distribution at the observation of the human position. In the real case, this distribution has more dimensions.

unknown phase ratio. Let the parameter θ_k represent a collaborative task independently trained as an Interaction ProMP. For each task k , the most probable α_k^* is first searched with (22) rendering the set $\{\alpha_k^*, \theta_k\}$. The task recognition is given by the procedure described in Section 3.3 with the difference that the likelihood (13) is now

$$p(k|y_{t-t'}^o) \propto p(y_{t-t'}^o|\alpha_k^*, \theta_k)p(k). \quad (23)$$

Since $\{\alpha_k^*, \theta_k\}$ is the solution of the phase search for each task, task recognition demands only the computation of (23). The two optimizations—to search for the correct phase, and to recognize the task—lead to an algorithm that scales linearly in the number of sampled α 's and in the number of tasks.

4 Experiments with a Semi-Autonomous Robot

Experiments were conducted using a 7-DoF lightweight KUKA arm equipped with a 5-finger hand. To train the primitives the joint encoder trajectories of the robot were recorded by kinesthetic teaching. For each demonstration, the set of human-robot measurements were stored with a sampling rate of 50 Hz.

The first set of experiments evaluates the responsiveness of the method in a single task scenario where the robot tries to predict where the human will handover a cup. The Cartesian coordinates of the cup were tracked with a motion capture system by placing a marker directly on it. In the second set of experiments, a multi-task scenario

of collaborative box assembly was used to evaluate the problem of task recognition under partial observations. In this case motion capture was used to track the Cartesian coordinates of the wrist of the human.

4.1 Predictive Handover

An important characteristic for fluid human-robot interaction is the capability of the robot to preemptively react to the human partner. In this experiment we evaluate this capability in an illustrative scenario of a cup handover. As shown in Figure 4, initially, a set of paired trajectories was collected by simultaneously moving the robot in kinesthetic teaching mode while the human was bringing the cup towards the robot. A total of 20 pairs of Cartesian trajectories of the cup and robot joint encoders were recorded. The demonstrator enforced temporal variability by changing the speed of the demonstration, with fast trajectories taking approximately 2 seconds and slower trajectories taking approximately 3 seconds.

We evaluated the effect of the number of candidates of phase ratios α w.r.t the approximation error of the true duration of the movement. Also, we evaluated the variation of the same parameter in relation to the accuracy of the prediction of the robot trajectory. To create a sparse limited observation, only the first half of each test trajectory was used, from which five states were randomly drawn along the time axis as observations. One instance is shown in Figure 5 (a) where ten phase ratios candidates were evaluated. The plot at the left shows the Cartesian coordinate of the marker on the cup and randomly drawn

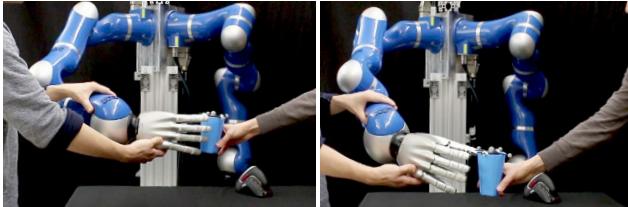


Figure 4. Collection of training data to create a single Interaction ProMP for the handover of a cup. The marker is placed on the top of the cup. Both cup and robot arm are moved towards the same position, leading to pairs of collaborative trajectories. This figure shows two of such demonstrations at different positions.

observations represented as the circles. The blue patch represents ± 2 standard deviations around the mean and the traced curve represents the ground truth. The plot at the right shows the predicted trajectory distribution of one of the joints of the robot arm. The many gray curves at the background show samples that were drawn from the trained model to illustrate the variability in both time and position. In (a), the fact that the predicted distribution has roughly the same duration as the test trajectory indicates that a reasonably good approximation to the true phase ratio could be found. The bottom row shows the same case, however when only two phase ratio candidates were evaluated leading to a coarse approximation of the true trajectory duration.

This procedure was repeated systematically with leave-one-out cross validation (LOOCV) on 20 test trajectories. The number of phase ratio candidates was varied from 1 to 15 for each test trajectory, resulting in a total of 15×20 runs. To systematically evaluate the different number of phases under repetitive conditions, we selected the candidates on an equally spaced grid within the interval covering 95 percent ($\pm 2\sigma$) of the normal distribution of phases. Figure 6 (a) shows the error in predicting the correct phase ratio as the number of sampled candidates increases. The results show that, on average, no improvement is achieved beyond 6 sampled candidates.

Due to the nature of the handover task, it is impractical to evaluate the accuracy of the robot in online experiments as the human will invariably adapt to the robot's response. Thus, we used the same LOOCV procedure to evaluate the accuracy of the robot end-effector with respect to the corresponding recorded human trajectory. This result is shown in Figure 6 (b), which shows the error of the final predicted positioning of the robot hand as the number of sampled candidates increases. The final position error of the robot hand was computed via forward kinematics of the arm given the mean of the predicted joint angle distributions. There is a lower limit of 3 cm in the accuracy of the robot prediction. This value suffices for tasks such as handovers

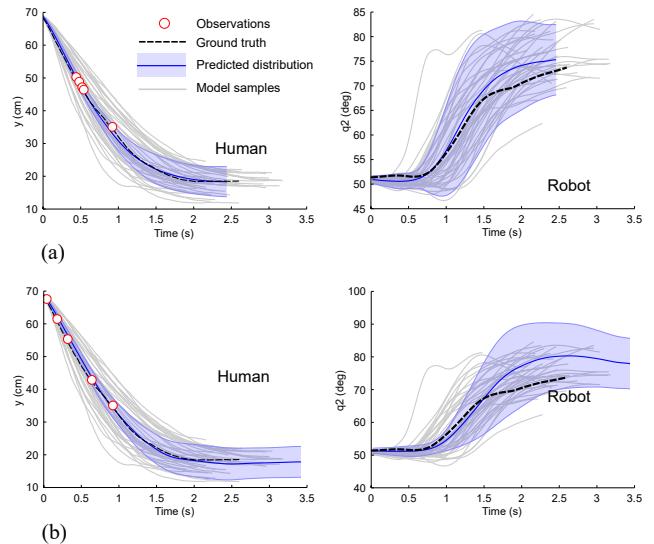


Figure 5. Examples of two different phase estimations under the same test data and five observations up to 1 second of the human movement. (a) Ten phase candidates evaluated. The position error of the robot hand is 4.31 cm. (b) Two phase candidates evaluated leading to a erroneous prediction of the trajectory duration. The position error of the robot hand was of 5.20 cm.

of objects and may be affected by the quality of the setup, such as the positioning of the markers and the non-ideal demonstrations that invalidate the Gaussian assumption. As it will be discussed in Section 5, direct feedback tracking of the marker may be used to increase the accuracy of the robot.

It is also important to evaluate the effect of the duration of the observations of the human before the robot attempts to infer its own trajectory. The sooner the robot can predict the correct human trajectory, and as a consequence the final handover position, the quicker it can react to provide assistance. On the other hand, trying to predict the human motion too early may lead to poor coordination between the human and robot final positions. Figure 7 shows snapshots where the robot observed only 0.2 seconds of the human movement (approx. 8 % of the whole trajectory length) before generating its own trajectory after evaluating ten phase candidates. The final position of the robot hand does not match the final position of the human hand. Empirically, for this particular task we noticed that observing the human for at least 0.5 seconds (approx. 20 % of the full trajectory length) while evaluating ten phase candidates lead to quite satisfactory coordination between the final positions of the human and robot hands.

For a quantitative analysis, we used the same 20 test trajectories of the previous analysis as ground truths and fixed the number of phase ratio candidates to ten. Each

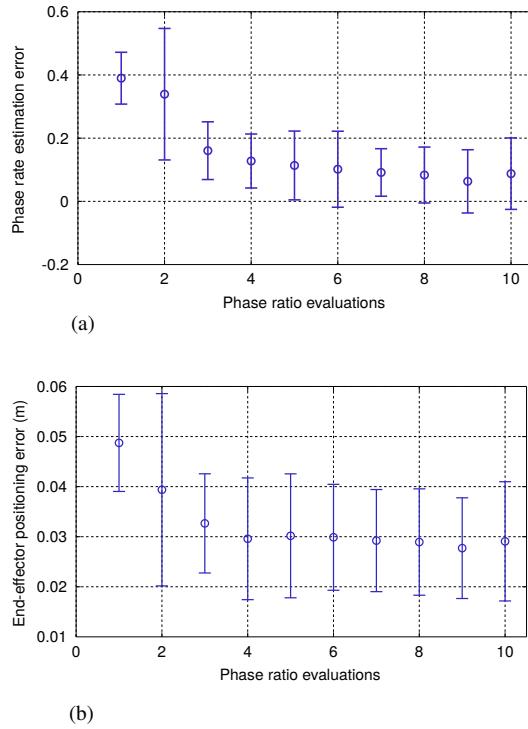


Figure 6. Effects of the number of phase ratio evaluations on (a) the approximation to the true phase ratio value, and (b) the accuracy of the prediction of the final position of the robot hand.

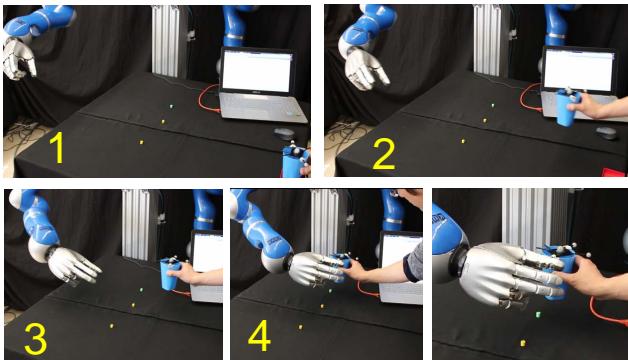


Figure 7. Snapshots of a handover trajectory where the robot reacts too early to predict the final human position accurately. In this example the robot observed only the first 0.2 seconds of the human movement before attempting to generate its own corresponding trajectory.

of the test trajectories was used to generate ten randomly drawn human observations. The span in which these observations were drawn varied between the first 10 % to 99 % of the total test trajectory length. Figure 8 (a) shows two examples where the trajectory of the human is predicted from different observation durations. On the left, only 5 % of the trajectory was observed; on the

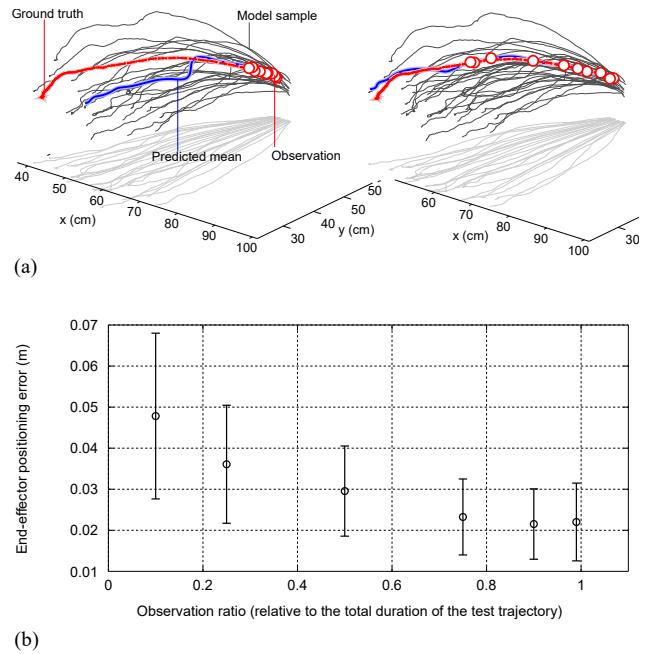


Figure 8. Effect of the duration of the observation of the human position on the cup handover task. (a) Prediction of the human trajectory in the Cartesian space given ten randomly sampled observations. Left: 5 % of the trajectory observed. Right: 50 % of the trajectory observed. (b) Accuracy of the final robot hand position as a function of increasing durations of the human observation.

right, 50 % of the trajectory was observed. As expected, the predicted human trajectory approximates the ground truth as the duration of observation increases. As a full state estimator, not only the trajectory of the human is computed from the observations, but also the corresponding robot trajectories. Figure 8 (b) shows the effect of different observation durations—indicated in the horizontal axis as the ratio between the observed length and the total length of the trajectory—on the prediction of the location of the robot hand. Similar to the previous analysis, the final position of the robot hand was computed by using forward kinematics of the arm given the predicted joint angles. As expected, the longer the human is observed before the prediction is attempted, the better the final accuracy of the robot, at the expense of lag in responsiveness.

A side-by-side comparison between short and long observations of the human movement and its effect on the robot responsiveness and accuracy can be watched in the accompanying video and also by following the link <https://youtu.be/bUVn0AwAb1U>. To control the robot motion, the mean of the posterior distribution over trajectories for each joint of the robot was used, and tracked

by the standard, compliant joint controller provided by the robot manufacturer[‡].

4.2 A Multi-Task Semi-Autonomous Robot Coworker

As it was motivated at the introduction of this paper, we applied our method on a multi-task scenario where the robot plays the role of a coworker that helps a human assembling a toolbox. This scenario was previously proposed in Ewerthon et al. (2015b). The assembly consisted of three different collaborative interactions. In one of them, the human extends his hand to receive a plate. The robot fetches a plate from a stand and gives it to the human by predicting the location of the marker on his wrist. In a second interaction, the human fetches the screwdriver and the robot grasps and gives a screw to the human as a pre-emptive collaborator would do. The third type of interaction consists of the robot receiving a screwdriver such that the human coworker can have both hands free (the same primitive representing this interaction is also used to give the screwdriver back to the human). Each interaction of “plate handover”, “screw handover” and holding the screwdriver was demonstrated 15, 20, and 13 times, respectively.

The upper row of Figure 9 shows snapshots of the training phase for each of the interactions and their respective multiple demonstrated trajectories in Cartesian space. The markers were placed on the wrist of the human. The bottom row shows the execution of the learned Interaction ProMPs from the demonstrations. On purpose, all human demonstrations started roughly at the same position in order to make the problem of action recognition apparent.

We used the original training data previously presented in Ewerthon et al. (2015b). The original data set did not present sufficient variability of phases and the correct phase could be reasonably well estimated with only two to three sample candidates. Thus, the durations of each demonstration was scaled by sampling different final times out of a normal distribution centered at the mean final time of the training set and with a standard deviation of 0.5 seconds. This perturbation acts as a surrogate of a demonstrator moving at different speeds at each demonstration.

We evaluate the effect of observing different durations of the human trajectory before attempting to recognize the action. Similar to the experiments in Section 4.1, each of the test trajectories was used to generate five, sparse randomly drawn, human observations. A total of ten phase ratio candidates were fixed. The duration from which these observations were drawn varied between the first 5 % to 75 % of the total duration of the test trajectory. Figure 10 shows, as circles, the observations of the “plate handover” trajectory drawn from the first 35 % of the trajectory length.

Note that while these observations clearly do not fit the first task of holding the tool, they can incorrectly explain the task of “screw handover”. Selecting the tasks by comparing their probabilities provides a principled mechanism to make such a decision.

Figure 11 (a) shows the action recognition accuracy according to the ground truth trajectory labeled with the corresponding action. The square marks represent the case when the trajectory of the “hold tool” task was used as a test data. The circle marks represent the case when the trajectory of “plate handover” task was the correct action, and cross marks represent the case when the “screw handover” was the correct action. The data measured from the motion capture system was used as a noiseless case. From the plot it is observed that even when only 5 % of the task was observed, the “plate handover” could be correctly recognized in all tests. The subplot (b) repeats the same procedure but Gaussian noise with zero mean and standard deviation of 3.5 cm on the observed points were added. As expected the recognition deteriorates as the few noisy observations could overlap significantly onto the incorrect distribution. Figure 11 (c) shows the case where the Gaussian noise has a standard deviation of 7.0 cm, in which case, practically the full human trajectory has to be observed such that a proper action recognition can be made. All cases in Figure 11 were evaluated using ten phase candidates.

As mentioned in Section 3.4, in previous works (Ewerthon et al. 2015b), the training data had to be time-aligned with a method based on DTW. During execution, the robot was then conditioned at the final position of the human as a way to overcome the problem of aligning partial and sparse observations online. With phase estimation not only considerable pre-processing effort is avoided—the use of DTW on both training and observations is not needed—but, the robot can now predict the collaborative trajectory before the human finishes moving, leading to a faster robot response. Figure 12 shows the difference in uncertainty and accuracies between the time-aligned method using DTW (a) with two cases of phase estimation that differ in regards to the durations of observations (b-c) for the task of “plate handover”. As shown in (a), since the Interaction ProMP is conditioned at the final position of the human hand—the principal state of interest for predicting the final position of the robot—inferece of the robot trajectory in this case provides a baseline to compare with the phase estimation method.

Using LOOCV over the training set, we quantified the accuracy of the prediction of the phase estimation

[‡]Although not used in this paper, the ProMP framework also provides means to compute the feedback controller and the interested reader is referred to Paraschos et al. (2013).

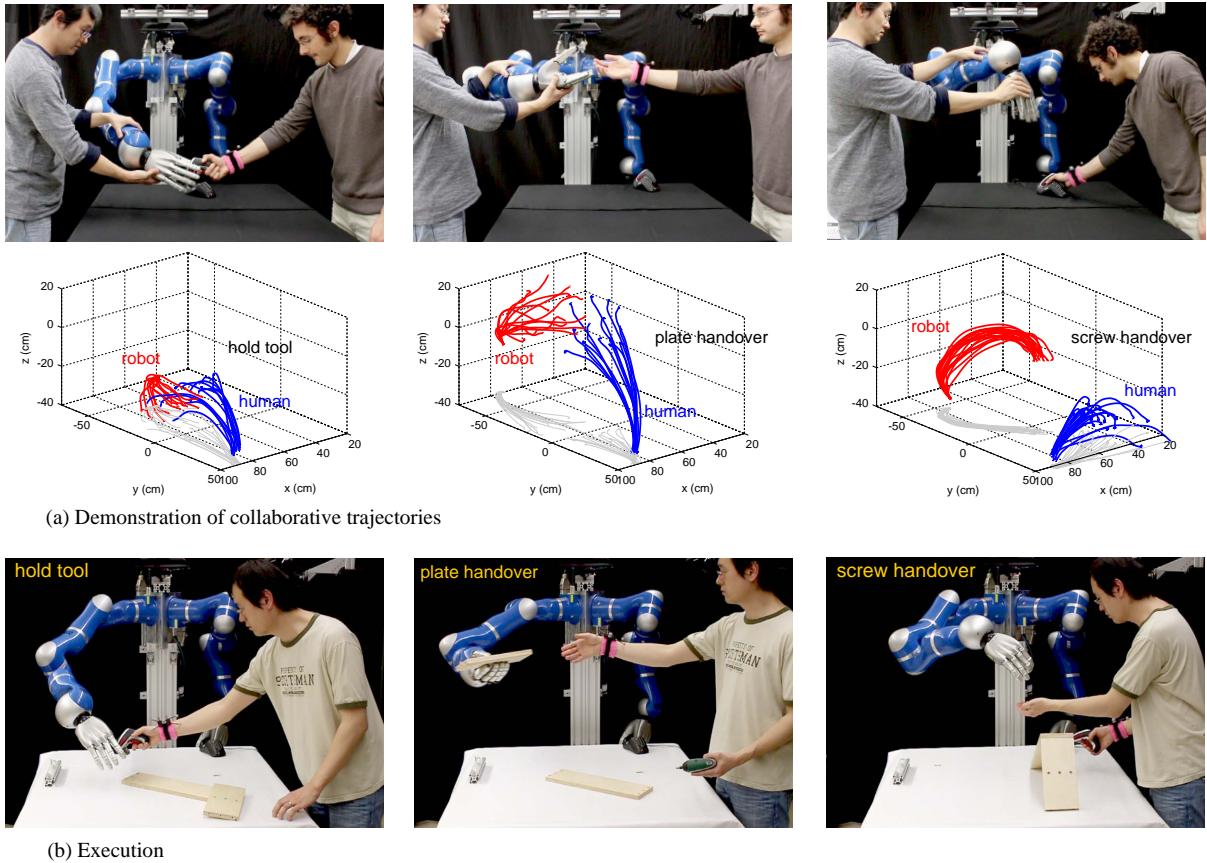


Figure 9. (a) Demonstrations of the three different interactions and their respective trajectories. (b) Human-robot collaboration using the trained Interaction ProMP for each task. The robot actions are conditioned on the partial observations of the human movement. The scenario and training data appeared previously in Ewerthon et al. (2015b) and Maeda et al. (2016) where training data had to be time-aligned by DTW.

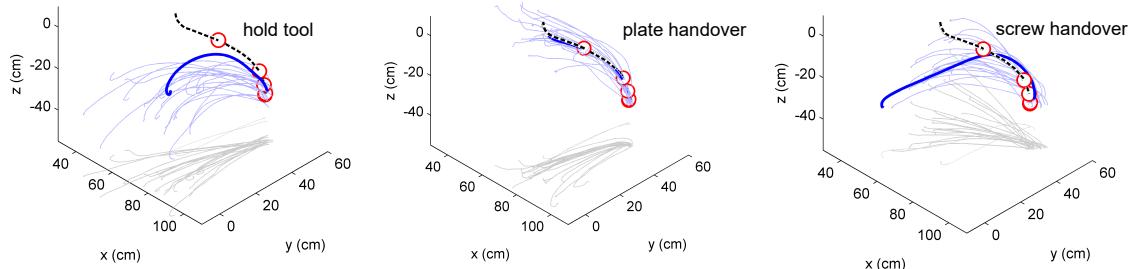


Figure 10. Each of the three tasks conditioned on the same observations sampled from the “plate handover” task. The same observations can fit the task of “screw handover” with a lower probability. The thick solid line shows the predicted human trajectory. The dashed lines is the “plate handover” ground truth, and the circles are observations.

method and the time-aligned model conditioned at the end of the observation (as illustrated in Figure 12). These results are summarized in Figure 13 where the proposed method is represented as the circles and the previous time-aligned method is shown by the square marker. The results show that for the “plate handover” and “hold tool” tasks the accuracy improves as longer observations are

made, as expected. At about 75 % of the observation, the accuracy achieved by the phase estimation method is comparable to the time-aligned trajectories using DTW. This experiment show that the single phase assumption is reasonably met, otherwise worst accuracy would be expected in relation to the time-aligned trajectories. For the task of “screw handover” no improvement is observed as

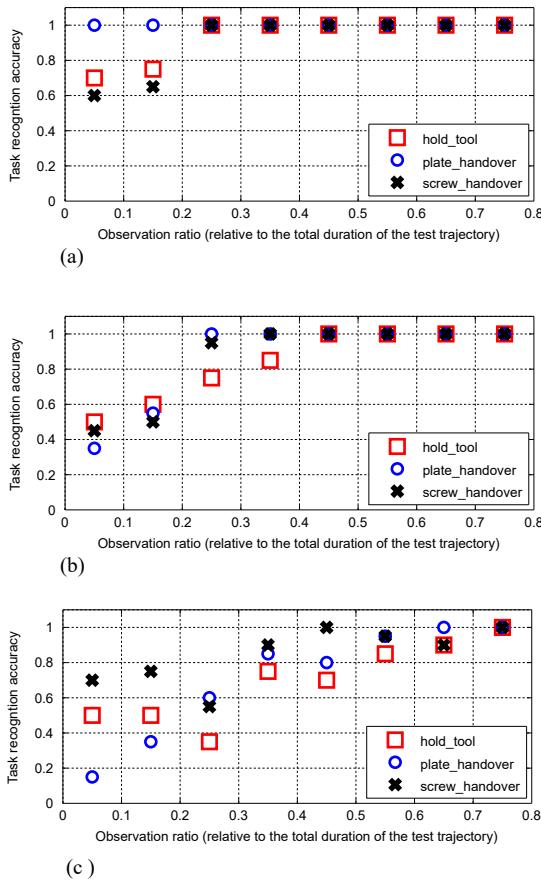


Figure 11. Improvement in action recognition as the duration of the observation increases under different amounts of noise. (a) Clean observations. (b) Observations added with zero-mean Gaussian noise and standard deviation 3.5 cm. (c) Observations added with zero-mean Gaussian noise and standard deviation 7.0 cm.

the trajectories of the human and robot are not correlated by the demonstration. In this case, the robot brings the screw to the mean of the unconditioned distribution (i.e. the prior mean) regardless of the position where the human grasps the screwdriver on the table.

5 Discussion of the Experiment and Limitations

Since the experiments aimed at exposing the predictive trajectory capability of the Interaction ProMPs no direct feedback tracking of the marker on the human wrist was made. However, even humans are prone to infer the wrong location when attempting to grasp an object from a partner too quickly. We naturally use visual feedback for corrections. In the same way, the Interaction ProMP framework may potentially benefit when used in

combination with a feedback controller that tracks the markers directly. In this case, the predictive trajectory generation that results from the method provides a rich nonlinear behavior as a movement primitive learned from demonstrations. Direct tracking of the marker could then act in a corrective/reactive manner to account for the mistakes in prediction. Note, however, that it is not possible to completely replace an Interaction ProMP by a tracking controller. In a multi-task scenario Interaction ProMPs are essential for action recognition. Moreover, a feedback controller does not easily provide the flexibility and richness of the trajectories that can be encoded in a primitive learned from human demonstrations.

The presented method scales linearly in the number of tasks and in the number of phase ratio candidates to be evaluated. In practice, this scalability imposes an upper limit on the number of tasks and samples α 's that can be supported, which can be empirically evaluated. The constraint is that the total time required to compute the probability of all sampled α 's for each possible tasks must be less than the duration of the human movement. Otherwise, one can simply use the final state of the human trajectory, as it was done in previous works with time aligned trajectories, to recognize the action and compute the corresponding collaborative robot trajectory. Therefore, the duration of the human movement and the implementation efficiency dictate to which granularity the phases can be estimated, and how many tasks can be recognized. In our preliminary evaluations on the assembly scenario, a total of 25 candidates of phase ratios α for each of the three tasks were used. This setting required 75 (25 samples \times 3 tasks) calls to the computation of the probabilities (19) while the human was moving his arm. The whole process, from the prediction of the full trajectory with (12), observed during the first second of the human movement took in average 0.20 seconds using Matlab code on a conventional laptop (Core i7, 1.7 GHz). In contrast, the duration of a handover stroke can vary from 1 to 2.5 seconds.

In our previous work (Maeda et al. 2016), for the same collaborative task, we evaluated the deterioration of action recognition when the duration of the observation differed from that of the time-aligned model. The frequency of the correct recognition of a “plate handover” decreased from 100 % to 50 % when the observed trajectory had a duration 1.25 longer than the duration of the time-aligned model. This sensitivity to temporal misalignments is an evidence that, unless the human partner is very consistent in terms of the speed of his/her movement, phase estimation is essential for action recognition. In the same work, we also compared the accuracy of a single-task Interaction ProMP with a baseline nearest-neighbor (NN) method on the task of plate handover. The Interaction ProMP presented twice

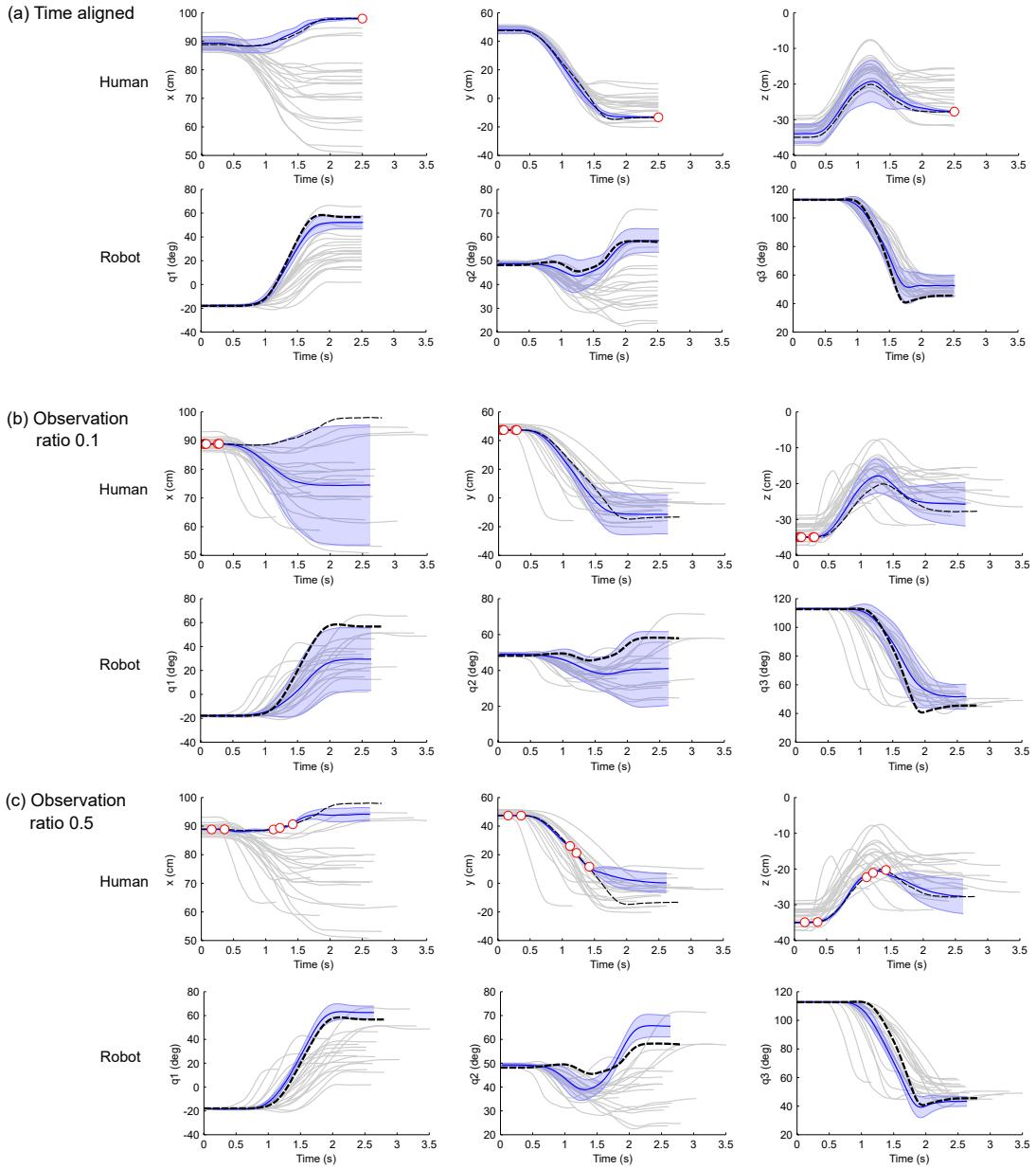


Figure 12. The patch represents the posterior distributions (± 2 standard deviations) for different methods to condition on the human observations. (a) Time-aligned trajectories (using DTW) conditioned at the final position. In (b) and (c), phase estimation with an observation ratio of 10 % and 50 % relative to the total duration of the test data, respectively.

the accuracy of the nearest-neighbor given the same training set.

Measurement setup systems that allow for reliable velocity measurements, or its estimation via position differentiation, can greatly simplify the problem of phase estimation. On the other hand, to become widely accepted, the deployment of semi-autonomous robots in the field must cope with occluded and cluttered environments. Sparse

position measurements must be taken into account in realistic scenarios, where noisy measurements are often provided by low-cost sensors such as Kinect cameras. The results in Figure 11 show promise in regards to the robustness of the method to large amounts of noise.

When compared to representations based on multiple reference frames such as Dynamical Systems (Calinon et al. 2012), and forcing functions as in DMPs, ProMPs have

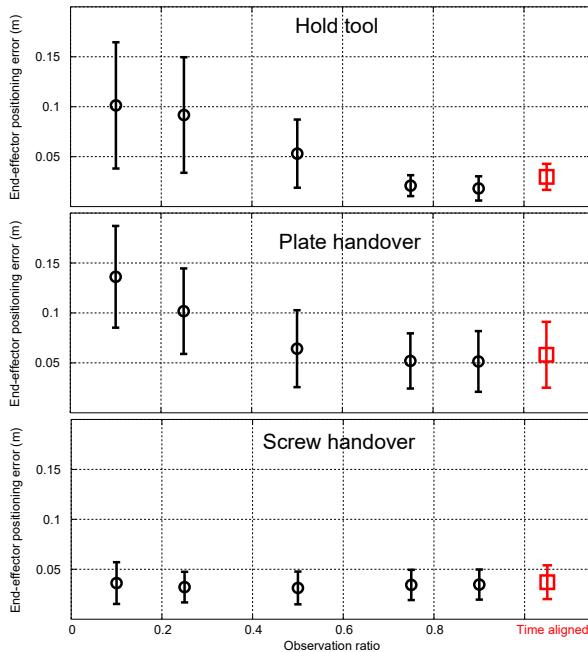


Figure 13. Effect of the duration of the observation on the predicted accuracy of the final robot hand position for each of the tasks. The time-aligned case represents the time-aligned model conditioned at the final position of the human measurement. In the “screw handover” case the position of the human hand is not correlated with the location in which the robot gives the screw.

the limitation to only operate within the demonstrated set. Extensions to generalize ProMPs to other robot kinematics and environments are currently being investigated.

6 Conclusion

This paper presented a method suited for collaborative and assistive robots whose movements must be coordinated with the trajectories of a human partner moving at different speeds. This goal was achieved by augmenting the previous framework of Interaction ProMPs with a prior model of the phases of the human movement, obtained from demonstrated trajectories. The encoding of phases enriches the model by allowing the alignment of the observations of the human in relation to the interaction model, under an intermittent stream of positions. We experimentally evaluated our method in an application where the robot acts as a coworker in a factory. Phase estimation allowed our robot to predict the trajectories of both interacting agents before the human finishes the movement, resulting in a faster interaction. The duration of a handover task could be decreased by more than 50 % in the single task case, and our current evaluation on the multi-task scenario decreases time by 25 %.

A future application of the method is to use the estimated phase of the human to adapt the velocity of the robot. A slowly moving human suggests that the robot should also move slowly, as an indication that a delicate task is being executed. Conversely, if the human is moving fast, the robot should also move fast as its partner may want to finish the task quickly. This paper initially appeared in Maeda et al. (2015), and here an extended version with additional evaluations was presented.

Acknowledgements

The authors would like to thank the reviewers of the first version of this paper, initially presented at ISRR 2015. Their comments were invaluable to improve the quality of this extended article, part of the special issue on the same conference.

The research leading to these results has received funding from the European Community’s Seventh Framework Programmes (FP7-ICT-2013-10) under grant agreement 610878 (3rdHand) and from the project BIMROB of the Forum für interdisziplinäre Forschung (FiF) of the TU Darmstadt.

References

- Ben Amor H, Neumann G, Kamthe S, Kroemer O and Peters J (2014) Interaction primitives for human-robot cooperation tasks. In: *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*.
- Cakmak M, Srinivasa SS, Lee MK, Forlizzi J and Kiesler S (2011) Human preferences for robot-human hand-over configurations. In: *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*. IEEE, pp. 1986–1993.
- Calinon S, Guenter F and Billard A (2006) On learning the statistical representation of a task and generalizing it to various contexts. In: *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 2978–2983.
- Calinon S, Guenter F and Billard A (2007) 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.
- Calinon S, Li Z, Alizadeh T, Tsagarakis NG and Caldwell DG (2012) Statistical dynamical systems for skills acquisition in humanoids. In: *Proceedings of the IEEE/RAS International Conference on Humanoid Robots (HUMANOIDS)*. pp. 323–329.
- Coates A, Abbeel P and Ng AY (2008) Learning for control from multiple demonstrations. In: *Proceedings of the 25th international conference on Machine learning*. ACM, pp. 144–151.
- Dong S and Williams B (2012) Learning and recognition of hybrid manipulation motions in variable environments using

- probabilistic flow tubes. *International Journal of Social Robotics* 4(4): 357–368.
- Englert P and Toussaint M (2014) Reactive phase and task space adaptation for robust motion execution. In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. pp. 109–116.
- Ewerton M, Maeda G, Peters J and Neumann G (2015a) Learning motor skills from partially observed movements executed at different speeds. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. pp. 456–463.
- Ewerton M, Neumann G, Lioutikov R, Ben Amor H, Peters J and Maeda G (2015b) Learning multiple collaborative tasks with a mixture of interaction primitives. In: *Proceedings of the International Conference on Robotics and Automation (ICRA)*. pp. 1535–1542.
- Hayne R, Luo R and Berenson D (2016) Considering avoidance and consistency in motion planning for human-robot manipulation in a shared workspace. In: *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*.
- Ijspeert AJ, Nakanishi J, Hoffmann H, Pastor P and Schaal S (2013) Dynamical movement primitives: learning attractor models for motor behaviors. *Neural computation* 25(2): 328–373.
- Kim S, Gribovskaya E and Billard A (2010) Learning motion dynamics to catch a moving object. In: *Proceedings of the IEEE/RAS International Conference on Humanoids Robots (HUMANOIDS)*. pp. 106–111.
- Kim S, Shukla A and Billard A (2014) Catching objects in flight. *IEEE Transactions on Robotics* 30(EPFL-ARTICLE-198748).
- Koppula HS and Saxena A (2013) Anticipating human activities using object affordances for reactive robotic response. In: *Robotics: Science and Systems*.
- Kupcsik A, Hsu D and Lee S (2015) Learning dynamic robot-to-human object handover from human feedback. In: *International Symposium on Robotics Research (ISRR)*. IFRR.
- Lee D and Ott C (2011) Incremental kinesthetic teaching of motion primitives using the motion refinement tube. *Autonomous Robots* 31(2-3): 115–131.
- Lee D, Ott C and Nakamura Y (2010) Mimetic communication model with compliant physical contact in human-humanoid interaction. *The International Journal of Robotics Research* 29(13): 1684–1704.
- Maeda G, Neumann G, Ewerton M, Lioutikov R, Kroemer O and Peters J (2016) Probabilistic movement primitives for the coordination of multiple human-robot collaborative tasks. In press: *Autonomous Robots. Special Issue on Assistive and Rehabilitation Robotics* : 1–20.
- Maeda G, Neumann G, Ewerton M, Lioutikov R and Peters J (2015) A probabilistic framework for semi-autonomous robots based on interaction primitives with phase estimation. In: *Proceedings of the International Symposium of Robotics Research (ISRR)*.
- Mainprice J and Berenson D (2013) Human-robot collaborative manipulation planning using early prediction of human motion. In: *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*. IEEE, pp. 299–306.
- Paraschos A, Daniel C, Peters J and Neumann G (2013) Probabilistic movement primitives. In: *Advances in Neural Information Processing Systems (NIPS)*. pp. 2616–2624.
- Perez-D'Arpino C and Shah JA (2015) Fast target prediction of human reaching motion for cooperative human-robot manipulation tasks using time series classification. In: *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, pp. 6175–6182.
- Sakoe H and Chiba S (1978) Dynamic programming algorithm optimization for spoken word recognition. *Acoustics, Speech and Signal Processing, IEEE Transactions on* 26(1): 43–49.
- Schaal S (1999) Is imitation learning the route to humanoid robots? *Trends in cognitive sciences* 3(6): 233–242.
- Sisbot EA and Alami R (2012) A human-aware manipulation planner. *Robotics, IEEE Transactions on* 28(5): 1045–1057.
- Strabala KW, Lee MK, Dragan AD, Forlizzi JL, Srinivasa S, Cakmak M and Miceli V (2013) Towards seamless human-robot handovers. *Journal of Human-Robot Interaction* 2(1): 112–132.
- Tanaka Y, Kinugawa J, Sugahara Y and Kosuge K (2012) Motion planning with worker's trajectory prediction for assembly task partner robot. In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 1525–1532.
- Van Den Berg J, Miller S, Duckworth D, Hu H, Wan A, Fu X, Goldberg K and Abbeel P (2010) Superhuman performance of surgical tasks by robots using iterative learning from human-guided demonstrations. In: *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. pp. 2074–2081.
- Vuga R, Nemec B and Ude A (2014) Speed profile optimization through directed explorative learning. In: *Proceedings of the IEEE/RAS International Conference on Humanoids Robots (HUMANOIDS)*. IEEE, pp. 547–553.
- Yamane K, Revfi M and Asfour T (2013) Synthesizing object receiving motions of humanoid robots with human motion database. In: *Robotics and Automation (ICRA), 2013 IEEE International Conference on*. IEEE, pp. 1629–1636.
- Ye G and Alterovitz R (2011) Demonstration-guided motion planning. In: *International symposium on robotics research (ISRR)*, volume 5.