

“© 2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

Environment-adaptive Interaction Primitives for Human-Robot Motor Skill Learning

Yunduan Cui^{1*}, James Poon^{2*}, Takamitsu Matsubara¹, Jaime Valls Miro²,
Kenji Sugimoto¹ and Kimitoshi Yamazaki³

Abstract—In complex environments where robots are expected to co-operate with human partners, it is vital for the robot to consider properties of their collaborative activity in addition to the behavior of its partner. In this paper, we propose to learn such complex interactive skills by observing the demonstrations of a human-robot team with additional external attributes. We propose Environment-adaptive Interaction Primitives (EalPs) as an extension of Interaction Primitives. In cooperation tasks between human and robot with different environmental conditions, EalPs not only improve the predicted motor skills of robot within a brief observed human motion, but also obtain the generalization ability to adapt to new environmental conditions by learning the relationships between each condition and the corresponding motor skills from training samples. Our method is validated in the collaborative task of covering objects by plastic bag with a humanoid Baxter robot. To achieve the task successfully, the robot needs to coordinate itself to its partner while also considering information about the object to be covered.

I. INTRODUCTION

There has been a significant increase in interest concerning human interactivity with ‘social’ robots over the last decade [1], with emphasis on systems implemented within domestic and industrial settings. Collaborative handling of objects is something commonly done by groups of people, for example moving bulky furniture or manipulating large deformable objects such as a table-cloth. In the domain of socially interactive autonomous systems it is also desirable for tasks routinely accomplished by multiple people to be performed by a cohesive human-robot team. This shift towards interactive robots, particularly in the industrial manufacturing sector, reflects industry trends in the gradual phasing-out of lower-level positions in assembly and fabrication that were traditionally fulfilled by human workers. It is estimated that between 2015 and 2018 approximately 1.3 million new industrial robots will be installed into factories worldwide [2]. Hand-tuning parameters and task details to suit each task in the ever-increasing scope of activities robots will be expected to participate in, is simply infeasible for objectives incompletely defined for the duration of a task. To address this, demonstration learning provides robots the ability to generalize their own task parameters from previous experiences. This is particularly applicable to humanoid

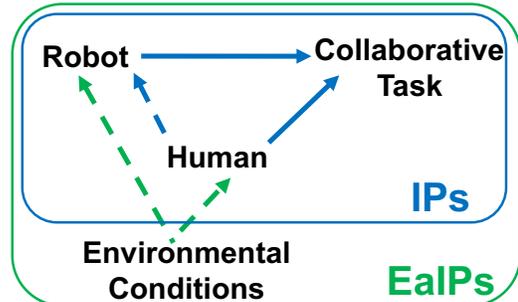


Fig. 1: Schematic diagram of EalPs: an extension of IPs to adapt to environmental conditions. In IPs, a robot predicts trajectory parameters to cooperate with a human partner after observing a brief movement period. EalPs enable robots to consider additional environmental conditions during trajectory prediction.

robotic platforms with regards to performing tasks normally done by people, as there is an intuitive connection between desirable behaviors of a human and a surrogate robot.

Our purpose in this study is to design a learning-from-demonstration framework for human-robot cooperative tasks with additional environmental conditions (e.g, the size of target object). After learning from training samples of both human and robot movement to finish a task under different environmental settings, the framework must be able to predict suitable robot motor skills to satisfy both a short partner observation period and novel environmental conditions.

The work in this paper builds upon the Interaction Primitives [3] (IPs) framework, which allow for a robot and human to perform collaborative tasks by converging upon a suitable parameter set for the execution of Dynamic Movement Primitives [4] (DMPs) after observing some initial period of human partner movement. In order to allow IPs to function in more complex situations, we integrate environmental parameters about the task to be accomplished into the parameter inference step for Environment-adaptive Interaction Primitives (EalPs) (Fig 1), which give inferences that consider both partner behavior and parameters describing environmental conditions. This is more aligned with the approach a human would undertake when collaborating with another person; information about the best action for them to perform must not come solely from his partner, but also from their shared environment. The impact of this contribution is twofold; the first being that the delay caused by partner observation can be

¹ Y. Cui, T. Matsubara and K. Sugimoto are with the Graduate School of Information Science, Nara Institute of Science and Technology, Japan

² J. Poon and J. Valls Miro are with the Faculty of Engineering and IT, University of Technology Sydney, Australia.

³ K. Yamazaki is with the Faculty of Engineering, Shinshu University, Japan

* These authors contributed equally to this paper.

significantly reduced, which allows for more immediate and fluent robot motion in situations where the human actions are ambiguous in the initial moments. The second effect is the capacity of adapting different environmental conditions in prediction. When faced with novel environmental parameters the prediction is inevitably tuned to parameters that allow for a similar basic trajectory structure, whereas conventional IPs require additional training data in scenarios where scaling a DMP on human motion no longer ensures safety. These improvements allow the EalPs framework to be robust against poor inferences from noisy observations of partner behavior.

The remainder of the paper is structured as follows. Related work is outlined in Section II. Details for IP and EalP are available under Section III. Our validation task is the collaborative covering of large objects with a plastic bag, first tested in a toy simulation problem in Section IV and then on NAIST’s humanoid Baxter robot (Fig. 5) in Section V. Section VI concludes with avenues for future investigation.

II. RELATED WORK

Owing to the increasing interest in physical human-robot collaboration, there have been numerous proposed methodologies from prior works for a mapping of robot behaviors from observation of the human partner’s behavior. Our work is inspired from the stylized DMPs of Matsubara et al. [5], where dimensionality reduction was used to obtain style parameters for DMPs when provided with multiple training examples. Here we attempt to replicate the robustness of latent style parameters through additional environmental parameters, without needing explicit transformation of training data. There have also been works focusing on our target level of action abstraction where the robot does not directly observe the human; instead actions must be derived from other observations upon the shared object being manipulated by both agents. Force feedback is also utilized by Lawitzky et al. [6] for motion generation when the human leads a robot around by means of manipulating a rigid object held between them. A similar concept is explored by Kruse et al. [7] but with a sheet of deformable fabric, utilizing either fold detection or force detection to infer actions that aim to maintain tension across the material’s surface.

These frameworks aim to achieve fluent and natural behaviors in collaborative tasks that require intimate human-robot co-operation to succeed. Higher-level task abstraction such as those in the frameworks that follow can allow a robot to exercise its own behavioral protocols based more on intentions rather than lower-level partner movement, however in scenarios involving physical communication it is crucial for the robot to behave with the approximate fluency of a human partner in order to achieve believable and natural human-robot interaction.

Koskinopoulou et al. [8] used RANSAC to map a latent space of human behavior to a latent space of a robot arm, both derived from Gaussian Process Latent Variable Models in order to accumulate a collection of skills such as opening doors on furniture objects. Patel et al. [9] used Hierarchical Hidden Markov Models to allow representation

of human behaviors at multiple levels of abstraction for the modular transfer of task knowledge to a robot. Anticipatory control is also explored by Huang and Mutlu in [10] with a gaze/speech-based approach to human behavior estimation utilizing Support Vector Machines, in a mock food-preparation scenario. An approach by Mühlig et al. [11] learns links between objects to be interacted with, to assist in the sequencing of more complex collaborative tasks. Ude et al. [12] utilized local weighting via Gaussian Process regression to allow for the synthesis of novel DMPs from demonstrated ones given query information about the goal and other task-specific details. Calinon et al. [13] assemble a task-specific trajectory model database to connect trajectory parameters and query information such as object pose.

III. APPROACH

A. Dynamic Movement Primitives

Encoding trajectory of human or robot movements, DMP [4] is formally written as a dynamic system:

$$\ddot{y}(t) = \left(\alpha_y \left(\beta_y (g - y(t)) - \frac{\dot{y}(t)}{\tau} \right) + f(x_t) \right) \tau^2 \quad (1)$$

where α_y and β_y are constants, y is the state variable of the trajectory, g is the target position, τ is a time constant and t is the time step. $f(x_t)$ is the forcing function built by M Gaussian basis functions and a corresponding $M \times 1$ weights vector \mathbf{w} :

$$f(x_t) = \frac{\sum_{i=1}^M \psi_i(x_t) w_i x_t}{\sum_{j=1}^M \psi_j(x_t)} = \phi(x_t)^T \mathbf{w}, \quad (2)$$

x follows a canonical system: $\dot{x} = -\alpha_x x \tau$ where $x_0 = 1$.

To learn a weight vector \mathbf{w} of DMP encoding a T step trajectory $\mathbf{y} = [y(t), \dot{y}(t), \ddot{y}(t)]_{t=1:T}^T$, the forcing function that reproduces the sample trajectory from the t -th step is calculated according to Eq. (1):

$$f(x_t) = \frac{1}{\tau^2} \ddot{y}(t) - \alpha_y \left(\beta_y (g - y(t)) - \frac{\dot{y}(t)}{\tau} \right). \quad (3)$$

The system can be resolved with $\mathbf{f} = \Phi \mathbf{w}$ where $\Phi = [\phi(x_1), \dots, \phi(x_T)]^T$ and $\mathbf{f} = [f(x_1), \dots, f(x_T)]^T$. Its least squares solution follows:

$$\mathbf{w} = (\Phi^T \Phi)^{-1} \Phi^T \mathbf{f} \quad (4)$$

B. Interaction Primitives in human-robot cooperation tasks

According to [3], applying Interaction primitives (IPs) to human-robot cooperation tasks has two steps: 1. Estimating phase of observed human movement. 2. predicting robot motor skills with a partially observation of only human’s movement.

Dynamic Time Warping (DTW) [14] is employed to estimate the phase of observed human movement. Given one partially observed human movement $[\mathbf{y}_1^*, \dots, \mathbf{y}_{T'}^*]^T$ and one reference movement $[\mathbf{y}_1, \dots, \mathbf{y}_T]^T$, the full human movement during the original demonstration of the task. DTW measures the similarity between these two temporal sequences and provides the index t^* reflecting the frame in the reference movement which produces minimal costs with respect to

the observed query movement, i.e., $[\mathbf{y}_1^*, \dots, \mathbf{y}_{T'}^*]^T$ is close to $[\mathbf{y}_1, \dots, \mathbf{y}_{t^*}]^T$. The estimated phase of partially observed human movement is therefore:

$$x^* = \exp\left(-\alpha_x \left(\frac{t^*}{T}\right) \tau\right). \quad (5)$$

For the prediction of robot motor skills with a partially observation of human's movement using IPs, we firstly prepare S sets of N DoFs trajectories that are time-scaled to the same length T as the training samples:

$$\mathbf{Y} = [\mathbf{Y}_{human}, \mathbf{Y}_{robot}] = \begin{bmatrix} \mathbf{y}_1^1 & \dots & \mathbf{y}_N^1 \\ \vdots & \ddots & \vdots \\ \mathbf{y}_1^S & \dots & \mathbf{y}_N^S \end{bmatrix} \quad (6)$$

where N is the totally number of DoFs for both human and robot. Defining \mathbf{y}_i^j , \mathbf{w}_i^j and g_i^j as the trajectory, weights vector and target position of the i -th DoF in the j -th demonstration respectively, $\boldsymbol{\theta}^{[j]} = [\mathbf{w}_1^j, g_1^j, \dots, \mathbf{w}_N^j, g_N^j]^T$, $j = 1, \dots, S$ is the DMPs parameter vector learned from $[\mathbf{y}_1^j, \dots, \mathbf{y}_N^j]$. Thus $p(\boldsymbol{\theta})$, the distribution among the parameter vector samples $\boldsymbol{\theta}^{[j]}$, $j = 1, \dots, S$, follows:

$$p(\boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{\theta} | \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta), \quad (7)$$

$$\boldsymbol{\mu}_\theta = \frac{\sum_{j=1}^S \boldsymbol{\theta}^{[j]}}{S}, \boldsymbol{\Sigma}_\theta = \frac{\sum_{j=1}^S (\boldsymbol{\theta}^{[j]} - \boldsymbol{\mu}_\theta)(\boldsymbol{\theta}^{[j]} - \boldsymbol{\mu}_\theta)^T}{S}. \quad (8)$$

Note that $\boldsymbol{\theta} = [\boldsymbol{\theta}_{human}, \boldsymbol{\theta}_{robot}]^T$ contains the parameter vectors of both human and robot.

Partially observing human's movement and estimating its phase x^* according to a reference movement by DTW, the trajectories $\mathbf{Y}_{human}^* = [\mathbf{y}_1^*, \dots, \mathbf{y}_n^*]^T$ are resampled from the observed movement where $n < N$ is the DoFs of human movement. The unavailable trajectories of robot \mathbf{Y}_{robot}^* are set to $\mathbf{0}$. Defining $\mathbf{Y}^* = [\mathbf{Y}_{human}^*, \mathbf{Y}_{robot}^*]$, the prediction of both human and robot's parameter vector is represented by:

$$p(\boldsymbol{\theta} | \mathbf{Y}^*) \propto p(\mathbf{Y}^* | \boldsymbol{\theta}) p(\boldsymbol{\theta}). \quad (9)$$

The likelihood $p(\mathbf{Y}^* | \boldsymbol{\theta})$ is modeled by a Gaussian distribution of the forcing function:

$$p(\mathbf{Y}^* | \boldsymbol{\theta}) \sim \mathcal{N}(\mathbf{F}^* | \boldsymbol{\Omega}\boldsymbol{\theta}, \sigma^2 \mathbf{I}) \quad (10)$$

where \mathbf{F}^* has two parts: $\mathbf{F}_{human}^* = [\mathbf{f}_1^*, \dots, \mathbf{f}_n^*]^T$ is the observed forcing function of \mathbf{Y}_{human}^* , its element is given by:

$$f_i^*(x_t) = \frac{1}{\tau^2} \ddot{y}_i^*(t) - \alpha_y (-\beta_y y_i^*(t) - \frac{\dot{y}_i^*(t)}{\tau}). \quad (11)$$

\mathbf{F}_{robot}^* is the unavailable forcing function of robot and set as $\mathbf{0}$. The matrix $\boldsymbol{\Omega}\boldsymbol{\theta}$ contains the forcing function with relationship to $\tilde{\boldsymbol{\Phi}}_t = [\phi(x_t)^T, \alpha_y \beta_y]$ over learning samples for $1 \leq t \leq t^*$:

$$\boldsymbol{\Omega}\boldsymbol{\theta} = \begin{bmatrix} \tilde{\boldsymbol{\Phi}} & 0 & \dots & \dots \\ 0 & \tilde{\boldsymbol{\Phi}} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & 0 \end{bmatrix} \begin{bmatrix} \mathbf{w}_1 \\ g_1 \\ \vdots \\ \mathbf{w}_N \\ g_N \end{bmatrix} \quad (12)$$

with the $\tilde{\boldsymbol{\Phi}}$ related to $\boldsymbol{\theta}_{robot}$ in $\boldsymbol{\Omega}$ being set to 0. σ^2 is the observation noise variance.

The joint distribution $p(\mathbf{Y}^*, \boldsymbol{\theta})$ is also a Gaussian distribution given the likelihood $p(\boldsymbol{\theta} | \mathbf{Y}^*)$:

$$\mathcal{N}\left(\begin{bmatrix} \mathbf{F}^* \\ \boldsymbol{\theta} \end{bmatrix} \middle| \begin{bmatrix} \boldsymbol{\Omega}\boldsymbol{\theta} \\ \boldsymbol{\mu}_\theta \end{bmatrix}, \begin{bmatrix} \mathbf{A} & \boldsymbol{\Sigma}_\theta \boldsymbol{\Omega}^T \\ \boldsymbol{\Omega} \boldsymbol{\Sigma}_\theta^T & \boldsymbol{\Sigma}_\theta \end{bmatrix}\right) \quad (13)$$

where $\mathbf{A} = \sigma^2 \mathbf{I} + \boldsymbol{\Omega} \boldsymbol{\Sigma}_\theta \boldsymbol{\Omega}^T$. The mean and variance of conditional distribution $p(\boldsymbol{\theta} | \mathbf{Y}^*)$ is derived as:

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{\theta} | \mathbf{Y}^*} &= \boldsymbol{\mu}_\theta + \boldsymbol{\Sigma}_\theta \boldsymbol{\Omega}^T \mathbf{A}^{-1} (\mathbf{F}^* - \boldsymbol{\Omega} \boldsymbol{\mu}_\theta), \\ \boldsymbol{\Sigma}_{\boldsymbol{\theta} | \mathbf{Y}^*} &= \boldsymbol{\Sigma}_\theta - \boldsymbol{\Sigma}_\theta \boldsymbol{\Omega}^T \mathbf{A}^{-1} \boldsymbol{\Omega} \boldsymbol{\Sigma}_\theta. \end{aligned} \quad (14)$$

After obtaining $\boldsymbol{\theta}$, the robot motor skills are operated by running DMPs with parameter vector $\boldsymbol{\theta}_{robot}$ with estimated phase x^* .

C. Learning Environment-adaptive Interaction Primitives

In the inference of human-robot cooperation tasks using IPs, a parameter vector $\boldsymbol{\theta}$ of robot movement primitives is required to suitably finish the task together with the human after a brief observation period. However if the task has different environmental conditions (e.g. object size and shape) and the observed trajectories of each training object are not sufficiently different through their initial steps, accurate and punctual prediction of DMP parameters is infeasible.

EalPs add the environment parameter e representing the different features of objects into IPs along with human observation trajectories $p(\mathbf{y}^*)$, i.e. calculating a joint distribution $p(\mathbf{Y}^*, \boldsymbol{\theta}, e)$. For recording trajectories with the environment parameters, a new training sample is defined as \mathbf{Y}_e with an additional dimension for $\mathbf{E} = [e^1, \dots, e^S]^T$:

$$\mathbf{Y}_e = [\mathbf{Y}_{human}, \mathbf{Y}_{robot}, \mathbf{E}] = \begin{bmatrix} \mathbf{y}_1^1 & \dots & \mathbf{y}_N^1 & e^1 \\ \vdots & \ddots & \vdots & \vdots \\ \mathbf{y}_1^S & \dots & \mathbf{y}_N^S & e^S \end{bmatrix}. \quad (15)$$

Applying DMPs to learn $\boldsymbol{\theta}$ from \mathbf{Y}_e , The likelihood $p(e | \boldsymbol{\theta})$ could be calculated by a supervised learning algorithm like least squares regression:

$$p(e | \boldsymbol{\theta}) \sim \mathcal{N}(e | \boldsymbol{\theta} \mathbf{w}_e, \sigma_e^2 \mathbf{I}). \quad (16)$$

Given t^* steps observing trajectories and environmental parameter \mathbf{Y}^* , e^* , we combine the observing forcing function and environmental parameter to $\mathbf{F}_e^* = [\mathbf{f}_1^*, \dots, \mathbf{f}_N^*, e^*]^T$ and get a distribution similar to Eq. (13):

$$\mathcal{N}\left(\begin{bmatrix} \mathbf{F}_e^* \\ \boldsymbol{\theta} \end{bmatrix} \middle| \begin{bmatrix} \boldsymbol{\Omega}_e \boldsymbol{\theta} \\ \boldsymbol{\mu}_{\boldsymbol{\theta}_p} \end{bmatrix}, \begin{bmatrix} \mathbf{A}_e & \boldsymbol{\Sigma}_\theta \boldsymbol{\Omega}_e^T \\ \boldsymbol{\Omega}_e \boldsymbol{\Sigma}_\theta^T & \boldsymbol{\Sigma}_\theta \end{bmatrix}\right) \quad (17)$$

$$\boldsymbol{\Omega}_e \boldsymbol{\theta} = \begin{bmatrix} \tilde{\boldsymbol{\Phi}} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \\ & & \mathbf{w}_e \end{bmatrix} \begin{bmatrix} \mathbf{w}_1 \\ g_1 \\ \vdots \\ \mathbf{w}_N \\ g_N \end{bmatrix}, \quad (18)$$

$$\mathbf{A}_e = \begin{bmatrix} \sigma^2 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & \sigma^2 & 0 \\ 0 & \dots & 0 & \sigma_e^2 \end{bmatrix} + \mathbf{\Omega}_e \mathbf{\Sigma}_\theta \mathbf{\Omega}_e. \quad (19)$$

The mean and variance of conditional distribution $p(\boldsymbol{\theta}|\mathbf{Y}_e^*)$ is derived by plugging Eqs. (18,19) into Eq. (14).

IV. SIMULATION RESULTS

EalPs are firstly applied to a toy simulation that simplifies our task of covering objects with a plastic bag in three-dimensional space, to passing over rectangular objects of different size in a two-dimensional space. As shown in Fig. 2 there are three rectangular objects, each allocated an environmental attribute as a measure of their height: 125, 215 and 305 pixels respectively. 30 training trajectories (blue) and 15 testing trajectories (each with 500 steps) are generated by hand. All trajectories in simulation have two DoFs (X and Y axis). We compare the inference accuracy of IPs and EalPs between two situations: a full observation (both horizontal and vertical movement) and a partial observation (horizontal movement only) over 15 testing trajectories with different lengths of observed trajectories. The inference accuracy is represented by the average dynamic time warping distance between the predicted and original trajectories.

According to the results shown in Fig. 4, EalPs considerably improve the inference accuracy in this simulation. In the full observation case where both X and Y axis trajectories are observable, both IPs and EalPs have better prediction with the increase of the observed trajectories' length. On the other hand EalPs predict better than IPs with shorter observed trajectories ($t^* \leq 100$ steps) because the environment parameter indicates the target object's class, whereas the early steps in both horizontal and vertical movements contain less information to determine which object is being covered. As all trajectories in X axis are very similar in Fig. 2, the prediction is more challenging in the partial observation case where only horizontal movements are available. IPs cannot predict accurately even after observing 60% of the trajectory while EalPs have almost identical performance to the full observation case, showing the environment parameter's capability of improving prediction performance with partially observed trajectories that lack sufficient information.

The environment parameter not only improves the inference accuracy, but also enables EalPs to generalize trajectories for tasks with unknown environmental setting according to training samples. Given a 100 steps test trajectory in the single observation case, trajectories based on different environmental parameters are shown in Fig. 3 where 100 trajectories generated by $\mathcal{N}(\boldsymbol{\theta}|\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta)$ for each object are represented by transparent lines, and solid lines represent the mean trajectories. We add new objects D and E that are not included in training samples to investigate EalPs' generalization ability. This shows the EalPs' capability of generating

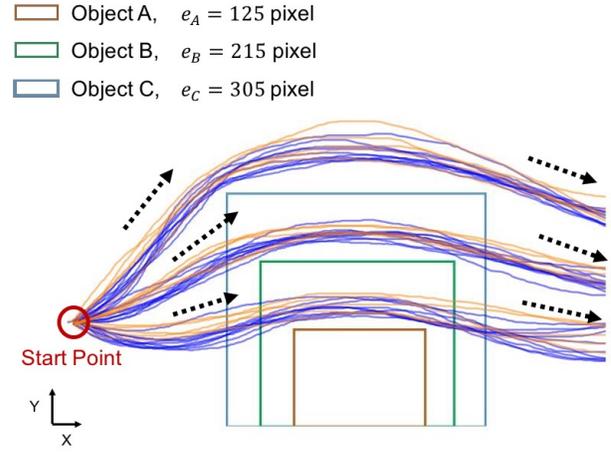


Fig. 2: Trajectories of training samples (blue) and testing samples (orange) to cross three objects on a 2D plan in simulation.

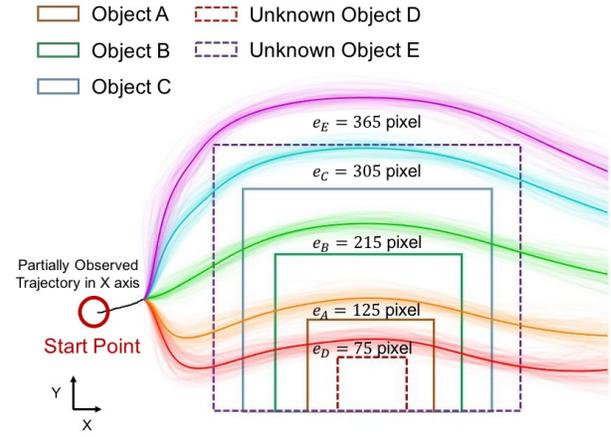
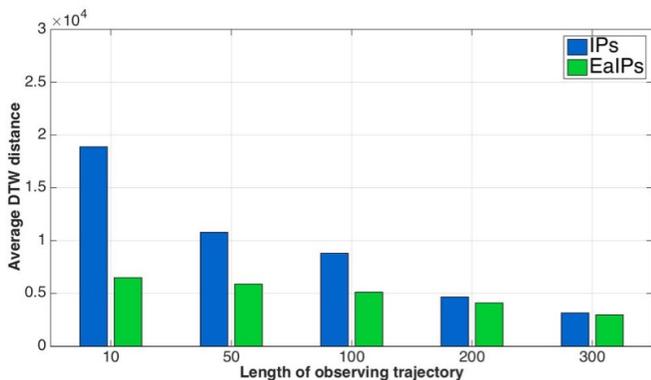


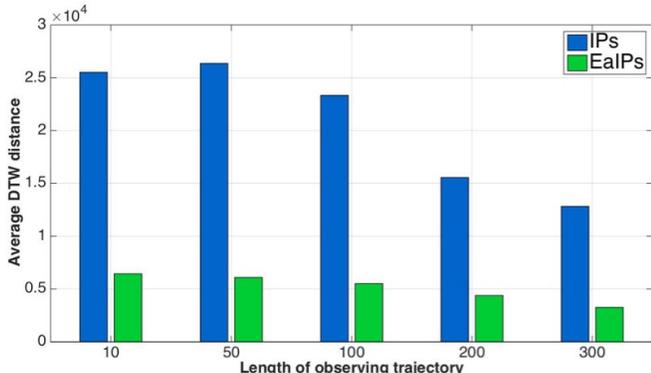
Fig. 3: Predicted trajectories of one testing samples ($t^* = 100$, only X axis is observable) with different environmental parameters. (Objects D and E are not included in training samples)

suitable trajectories to new situations; EalPs successfully obtain movement style from samples and automatically adjust the scale of trajectories to fit new environments according to the relationship between the trained and new environmental parameters.

These simulation results show the potential of EalPs in the early stages of movement inference of human-robot tasks. EalPs are able to give accurate prediction with short and partially observed trajectories from a very brief observation and can generate suitable co-operative robot movement. Moreover, EalPs adjust their prediction to fit new environments based on the knowledge learned from training samples. It reduces the requirement of training samples to cover many possible cases, while still allowing robots to intelligently co-operate with humans when dealing with various objects.



(a) The average inference accuracy of full observation.



(b) The average inference accuracy of partial observation.

Fig. 4: Comparison of trajectory prediction accuracy between IPs and EalPs. DTW distance (vertical axis) is a unitless measure of error between two temporally aligned signals.

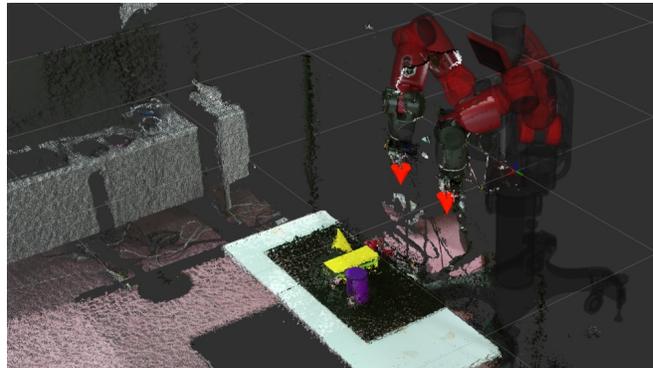
V. EXPERIMENTAL RESULTS

For validation in a real experiment, the NAIST Baxter robot (Fig 5) was made to hold one corner of a large plastic bag stationary with its right gripper, while its left gripper held another corner of the bag and swept it over a large object to cover it in tandem with his partner, who moves both hands. The NAIST Baxter has three Kinect V2 sensors (Fig 5a) providing visualization of the robot from left, right and birds-eye views (Fig 5b). Brightly colored cleaning gloves are worn by the interaction partner throughout gathering of training data and testing, so the Kinects can track the partner’s hands in real-time. Data communication and logging was managed via the Robot Operating System (www.ros.org) middleware.

Five training trajectories were gathered for each of three objects (Fig 6a): a 75 cm high stool, an 80 cm high office chair and a 90 cm high cabinet. Sample trajectories (Fig 6b-6c) containing the poses of the Baxter’s grippers and the positions of the partner’s hands were then cut and interpolated in Matlab to be of uniform length for EalP training. For ease of differentiability, the imaginary axes of gripper orientation Quaternions were learned instead of their Euler equivalents. Figure 7 shows the Baxter’s left gripper



(a) The Baxter research robot with three Kinect V2.



(b) Point cloud generated by Kinects.

Fig. 5: NAIST Baxter research robot learning system.

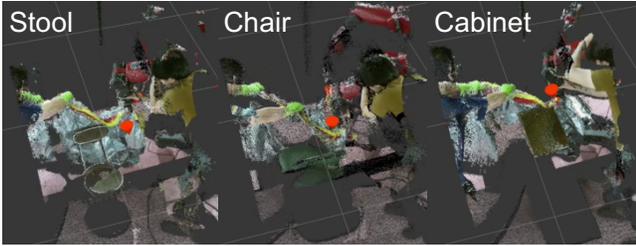
trajectory in the execution of EalP for covering each of the three known objects with a new interaction partner, as well as a novel fourth object: the 90 cm high cabinet elevated by an additional 10 cm. It can be seen that each trajectory is suitably changed to satisfy each environmental parameter, and is robust to both inconsistent starting positions of the partner’s hands and significant noise from glove tracking throughout their motion. Only 50 partner observations (1.5 seconds) were required to converge on the parameter sets that resulted in the trajectories shown. Figure 8 shows object coverage by the interaction partner and the EalP-generated trajectories from Figure 7, showing appropriate movements within the Baxter arm’s range of motion.

VI. CONCLUSION AND FUTURE WORK

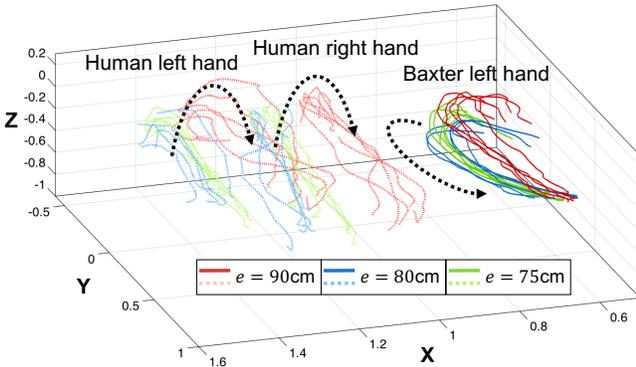
This work presents an extension of the Interaction Primitives framework to allow for more natural human-robot interaction. By taking into account environmental conditions along with parameters describing human behavior, the robot can quickly and confidently determine suitable parameters of motor skills even when faced with ambiguous behavior from the human partner during initial moments of their collaborative activity. The additional environmental parameter allows some knowledge of behavioral ‘style’ desirable of the robot to be learned independently from the human. Experimental results in both simulation of a toy problem and on a real humanoid platform show that these allow for fast



(a) Objects for training samples with e as height.



(b) Examples of training sample generation, captured by Kinect V2 sensors.



(c) 15 training sets (includes human left/right hand, and Baxter left hand gripper).

Fig. 6: Training trajectories for three objects.

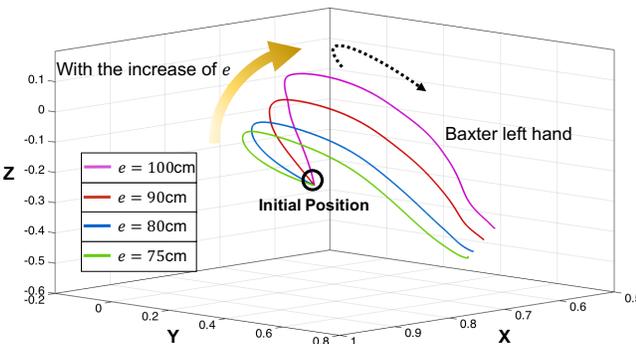


Fig. 7: Baxter left gripper trajectory from EalP across various objects.

and accurate task generalizations despite poorly informative early partner behavior. For future investigation we intend to pursue the application of dimensionality reduction techniques for learning more rigidly structured trajectories [5], and the tracking of more of the human partner's body for achieving increasingly complex collaborative human-robot activities with stochastic environmental parameters from sensor data. The consideration of force/torque measurements from the shared object may allow for additional behavioral adaptations.

VII. ACKNOWLEDGMENT

We gratefully acknowledge the support from the New Energy and Industrial Technology Development Organization (NEDO) for this research. We also thank Mr. Haifeng Han for his assistance with real experiment.

REFERENCES

- [1] M. Vircikova, P. Smolar, and P. Sincak, "Current trends in human-robot interaction: towards collaborative & friendly machines," in *12th Scientific Conference of Young Researchers*, 2012.
- [2] International Federation of Robotics, "World robotics 2015 industrial robots," tech. rep., International Federation of Robotics, Germany, 2015.
- [3] H. Ben Amor, G. Neumann, S. Kamthe, O. Kroemer, and J. Peters, "Interaction primitives for human-robot cooperation tasks," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pp. 2831–2837, IEEE, 2014.
- [4] A. J. Ijspeert, J. Nakanishi, H. Hoffmann, P. Pastor, and S. Schaal, "Dynamical movement primitives: learning attractor models for motor behaviors," *Neural computation*, vol. 25, no. 2, pp. 328–373, 2013.
- [5] T. Matsubara, S.-H. Hyon, and J. Morimoto, "Learning parametric dynamic movement primitives from multiple demonstrations," *Neural Networks*, vol. 24, no. 5, pp. 493–500, 2011.
- [6] M. Lawitzky, J. R. Medina, D. Lee, and S. Hirche, "Feedback motion planning and learning from demonstration in physical robotic assistance: differences and synergies," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3646–3652, 2012.
- [7] D. Kruse, R. J. Radke, and J. T. Wen, "Collaborative human-robot manipulation of highly deformable materials," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3782–3787, 2015.
- [8] M. Koskinopoulou, S. Piperakis, and P. E. Trahanias, "Learning from demonstration facilitates human-robot collaborative task execution," in *The Eleventh ACM/IEEE International Conference on Human Robot Interaction, HRI '16*, pp. 59–66, 2016.
- [9] M. Patel, J. V. Miro, D. Kragic, C. H. Ek, and G. Dissanayake, "Learning object, grasping and manipulation activities using hierarchical hmms," *Autonomous Robots*, vol. 37, no. 3, pp. 317–331, 2014.
- [10] C. M. Huang and B. Mutlu, "Anticipatory robot control for efficient human-robot collaboration," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 83–90, 2016.
- [11] M. Mühlig, M. Gienger, and J. J. Steil, "Interactive imitation learning of object movement skills," *Autonomous Robots*, vol. 32, no. 2, pp. 97–114, 2012.
- [12] A. Ude, A. Gams, T. Asfour, and J. Morimoto, "Task-specific generalization of discrete and periodic dynamic movement primitives," *IEEE Transactions on Robotics*, vol. 26, no. 5, pp. 800–815, 2010.
- [13] S. Calinon, T. Alizadeh, and D. G. Caldwell, "On improving the extrapolation capability of task-parameterized movement models," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 610–616, 2013.
- [14] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE transactions on acoustics, speech, and signal processing*, vol. 26, no. 1, pp. 43–49, 1978.

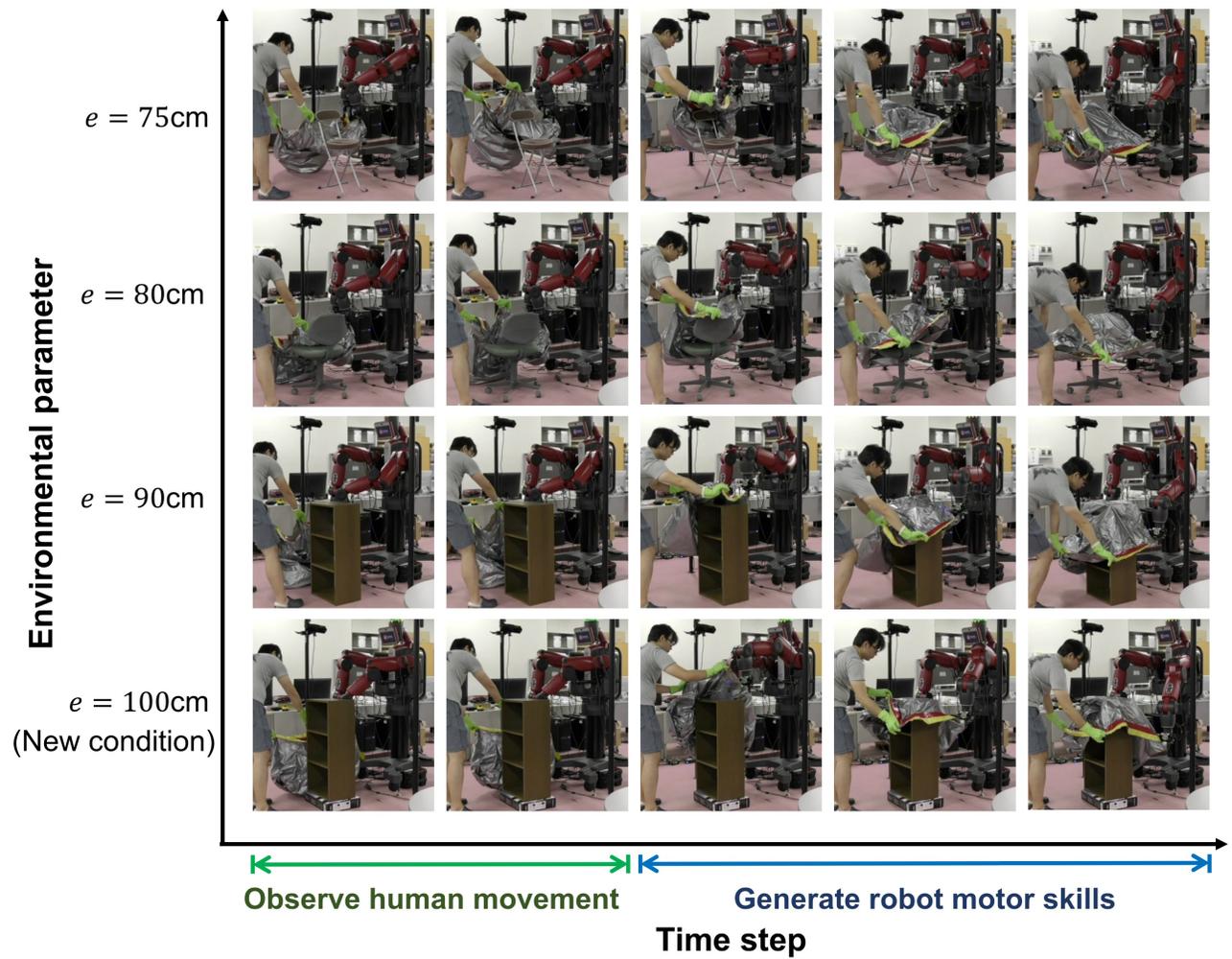


Fig. 8: Results of EalP trajectory generation in cooperative covering task with different environmental conditions.