Learning from Demonstration Facilitates Human-Robot Collaborative Task Execution

Maria Koskinopoulou, Stylianos Piperakis and Panos Trahanias
Institute of Computer Science, Foundation for Research and Technology - Hellas (FORTH)
and
Department of Computer Science, University of Crete
Heraklion, Crete, Greece
Email: {mkosk,spiperakis,trahania} @ics.forth.gr

***Abstract*—Learning from Demonstration (LfD) is addressed in this work in order to establish a novel framework for Human-Robot Collaborative (HRC) task execution. In this context, a robotic system is trained to perform various actions by observing a human demonstrator. We formulate a latent representation of observed behaviors and associate this representation with the corresponding one for target robotic behaviors. Effectively, a mapping of observed to performed actions is defined, that abstracts action variations and differences between the human and robotic manipulators, and facilitates execution of newly-observed actions. The learned action-behaviors are then employed to accomplish task execution in an HRC scenario. Experimental results obtained regard the successful training of a robotic arm with various action behaviors and its subsequent deployment in HRC task accomplishment. The latter demonstrate the validity and efficacy of the proposed approach in human-robot collaborative setups.**

Keywords: Learning from Demonstration; observation space; latent space; Gaussian Process; human-robot collaboration.

I. INTRODUCTION

Learning new behaviors enables robots to expand their repertoire of actions, learn new tasks and effectively interact with other agents and their environment. In the current work, we introduce a novel methodology to accomplish Learning from Demonstration (LfD) in robotic systems. Accordingly, action behaviors are systematically learned by observing a human actor, and are subsequently utilized in a Human-Robot Collaborative (HRC) task execution scenario.

Choosing and executing an appropriate action at every stage of an HRC scenario can be formulated as a problem of learning mappings between world states and actions. Such mappings, called policies, enable a robot to select actions based on the current world state. The development of such policies manually is often very challenging, so machine learning techniques have been proposed to cope with the inherent complexity of the task [1]. A distinct approach to policy learning regards the so-called Learning from Demonstration (LfD), also referred in the literature as Imitation Learning, or Programming by Demonstration [2].

LfD is treated generically in our approach, in that we attempt to abstract from the differences present in individual demonstrations and summarize the learned task in an executable robotic behavior. An appropriate latent space representation is employed that effectively encodes demonstrated actions. A transformation from the actor's latent space to the robot's latent space dictates the action to be executed by the robotic system. Learned actions are in turn used in HRC scenarios to facilitate fluent cooperation between two agents (human and robotic system). Extensive experimentation with the developed approach has been conducted in a realistic collaborative scenario. The obtained results attest for the appropriateness and viability of the proposed formulation for HRC task execution.

The rest of the paper is organized as follows. The next section presents a brief literature review, and then Section III describes the proposed LfD methodology and its implementation for HRC task execution. Experimental results obtained with real robotic systems are presented in Section IV, and the paper concludes with a summary and ideas for future work in Section V.

II. RELATED WORK

A learning process involves the extraction of salient information from demonstrations, and its subsequent encoding in a motion model that can be used to reproduce a given task. Using a set of basic learned motion primitives, more advanced robot motions can be achieved by combining and adapting the different motion primitives. Various learning techniques and motion models have been proposed, including polynomial and spline-based methods [3], nonlinear regression techniques, time-dependent dynamical systems [4], [5], and autonomous dynamical systems [6], [7]. These methods have been successfully developed to learn motion primitives, such as discrete (point-to-point) motions and their extensions to obstacle avoidance, rhythmic and hitting motions, etc. [8].

In the current work we develop an LfD approach that relies heavily on a latent space representation of the actors' configurations. An advantage of the proposed formulation over the state of the art methods for imitation learning is that the mapping between the observed and the action spaces is established independently of the teacher's and the robot's kinematics. At the same time, LfD methods are usually faced with the correspondence problem across the two (teacher-learner) spaces [2], [8]. The latter is greatly facilitated in our approach with the employment of the latent space representation. The formulation of the latter is accomplished with the employment

978-1-4673-8370-7/16/$31.00 © 2016 European Union

of a recently introduced procedure, namely Gaussian Process Latent Variable Model (GPLVM) [9], [10].

Recent research works assuming latent-space representations via GPLVM have appeared in various fields, such as computer vision for body pose tracking [11], [12], [13], computer graphics for graphical model representation (character animation) [14], and speech processing for voice conversion [15]. GPLVM is a non-linear generalization of Principal Component Analysis (PCA) [16], which provides a probabilistic compact transformation of a given high dimensional dataset to a low dimensional one. Recent works of Shon et al. [17] and Grimes et al. [18], [19] introduce a latent space representation in robotic imitation tasks. Given a dataset of human animated characters they project each animation pose to a latent space that also corresponds to a robot's pose. Nevertheless, the previous approaches have not been applied to real imitation learning scenarios. Training methods have also been used in human-robot interaction. In [20] a robotic system learned by its own experience how to interact with a human in a simple collaborative task. The work in [21] employs Gaussian mixture models to accomplish robot training, and in turn facilitate human-robot joint action execution.

In this paper we formulate LfD based on GPLVM for learning robotic action behaviors. Besides successful reproduction of demonstrated actions, the proposed approach is also capable to learn novel action behaviors. Additionally, learned behaviors are credibly applied in real HRC scenarios for cooperative task execution. The latter constitutes the main contribution of our work, demonstrating effective human-robot collaboration.

III. LEARNING ACTIONS FROM DEMONSTRATIONS

The proposed HRC task execution is based on learned robot action behaviors. LfD is employed to train a robotic system, and acquired action skills facilitate HRC. In LfD, robots are taught to perform tasks by observing a set of demonstrations provided by a "teacher" (human or robot). In this work, a human actor assumes the role of the teacher, and the robotic system is expected to learn appropriate actions by observing action demonstrations.

A schematic representation of the proposed LfD methodology is illustrated in Fig. 1. According to this methodology, latent (hidden) spaces are employed to associate corresponding actions performed by the teacher and the robot. More specifically, learning of action behaviors is accomplished in the following steps:

- The human teacher and the robotic system perform physically similar action behaviors. The trajectory of each individual action is tracked, giving rise to a workspace representation of the relevant action.
- An appropriate dimensionality reduction scheme is employed to transform the above representation into a low dimensional, latent space representation. The latter results in compact, latent representations of demonstrated and learned actions.
- By establishing a correspondence across the two latent spaces, we achieve to properly associate learned actions with the demonstrated ones.

In this work, we are interested in, and focus our study on actions performed by an arm. The execution of similar action behaviors by a human and a robotic arm is assured by physically attaching the two systems. The compliance feature of the robotic arm is exploited in order to let the human arm steer the robotic one in an action behavior.

To obtain the workspace representations of the trajectories of the above mentioned arm actions, we rely on the physical joints of each arm system. Accordingly, for the human arm we utilize the 3D coordinates of each of the three arm joints, namely shoulder, elbow, wrist. Similarly, the joint-configuration of the robotic arm is used to unambiguously represent the followed trajectory. A conceptual illustration of the above is shown in Fig. 2.

The GPLVM algorithm is employed in order to derive the latent space representation for both the human and the robotic arm. GPLVM realizes dimensionality reduction, much like PCA [16], in a probabilistic formulation. Effectively, it achieves to compress information in a small number of parameters, respecting at the same time non-linearities in the initial representations.

The mapping of the two extracted latent spaces is accomplished through a linear geometric transformation. The well-known RANSAC algorithm (RANdom SAmple Consensus) [22], [23] is used to establish the mentioned mapping due to its appropriateness in associating relevant features across two sets. In the current work, we adopt the term hyperspace to jointly denote the two latent spaces along with the corresponding mapping.

In the rest of this section we present the mathematical set-up of the building blocks of the proposed methodology, namely GPLVM and RANSAC. The section concludes with a rigorous formulation of the proposed action imitation framework.

A. Latent space representation - GPLVM

As already mentioned, GPLVM is used to derive the latent space representation [9]. This algorithm can be considered as

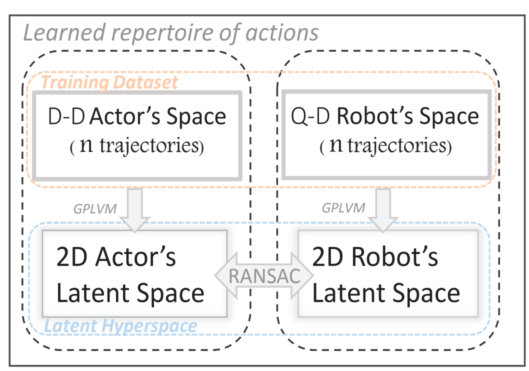


Fig. 1: Block Diagram of the proposed LfD methodology.

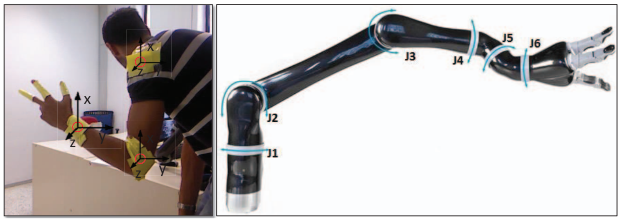


Fig. 2: Arm joints are employed to characterize and represent the trajectories of performed actions.

a non-linear generalization of PCA. Effectively, it provides a more accurate representation of an abstract multidimensional space projected to a lower dimensional one. Given an observation data set $\mathbf{X} = \{x_1, x_2, ..., x_n\}$, $x_n \in \mathbb{R}^D$, we search for an L - dimensional representation, where usually $L \ll D$. The derived representation, $\mathbf{Y} = \{y_1, y_2, ..., y_n\}$, $y_n \in \mathbb{R}^L$, termed as latent variables, is obtained by:

$$\mathbf{x}_n = \mathbf{W}\mathbf{y}_n + \boldsymbol{\eta}_n \tag{1}$$

where the matrix $\mathbf{W} \in \mathbb{R}^{D \times L}$ specifies the linear relationship between the data space and the latent space, and the noise values $\boldsymbol{\eta_n}$ follow a Gaussian distribution $\mathcal{N}(\boldsymbol{\eta}_n \mid \mathbf{0}, \beta^{-1}\mathbf{I})$ with unit covariance, and β denoting the inverse variance. To this end, the conditional probability of a data point $\mathbf{x}_n$ will also follow a Gaussian distribution as shown in (2):

$$p(\mathbf{x}_n \mid \mathbf{y}_n, \mathbf{W}, \beta) = \mathcal{N}(\mathbf{x}_n \mid \mathbf{W}\mathbf{y}_n, \beta^{-1}\mathbf{I}) \tag{2}$$

To obtain the marginal likelihood, an integration over the latent variables is needed, as

$$p(\mathbf{x}_n \mid \mathbf{W}, \beta) = \int p(\mathbf{x}_n \mid \mathbf{y}_n, \mathbf{W}, \beta) p(\mathbf{y}_n) d\mathbf{y}_n \tag{3}$$

Then, the conditional probability of $\mathbf{X}$ can be written as

$$p(\mathbf{X} \mid \mathbf{W}, \mathbf{Y}, \beta) = \prod_{n=1}^{N} \mathcal{N}(\mathbf{x}_n \mid \mathbf{W}\mathbf{y}_n, \beta^{-1}\mathbf{I}) \tag{4}$$

However, instead of denoting a prior distribution over latent space $p(\mathbf{Y})$, and seek to marginalize out $\mathbf{Y}$, a dual form of PCA can be considered that appoints a Gaussian distribution of weights:

$$p(\mathbf{W}) = \prod_{n=1}^{D} \mathcal{N}(\mathbf{w}_n \mid 0, \mathbf{I}) \tag{5}$$

Marginalizing out $\mathbf{W}$

$$p(\mathbf{X} \mid \mathbf{Y}, \beta) = \int \prod_{n=1}^{D} p(\mathbf{x}_n \mid \mathbf{y}_n, \mathbf{W}, \beta) p(\mathbf{W}) d\mathbf{W} \tag{6}$$

we get

$$p(\mathbf{X} \mid \mathbf{Y}, \beta) = \frac{1}{(2\pi)^{DN/2} |K|^{D/2}} \exp(-\tfrac{1}{2} \operatorname{Tr}(\mathbf{K}^{-1}\mathbf{X}\mathbf{X}^\top)) \tag{7}$$

where $\mathbf{K}$ is the covariance matrix defined by

$$\mathbf{K} = \mathbf{W}\mathbf{W}^\top + \beta^{-1}\mathbf{I} \tag{8}$$

Gaussian Process properties allow us to rewrite the previous equation as a product of D independent processes

$$p(\mathbf{X} \mid \mathbf{Y}, \beta) = \prod_{i=1}^{D} \frac{1}{(2\pi)^{N/2} |K|^{1/2}} \exp(-\tfrac{1}{2}(\mathbf{x}_{:,n}^\top \mathbf{K}^{-1} \mathbf{x}_{:,n})) \tag{9}$$

where $\mathbf{x}_{:,n}$ is the n_{th} column of $\mathbf{X}$. Therefore, a map from the latent to the observed space is created by the Gaussian process. Learning this mapping via maximum likelihood with respect to $\mathbf{Y}$, is what we define GPLVM. Since, linear kernels are unable to encode the non-linear features of the input data, we replace them by non-linear kernels. In this work, the Radial Basis Functions (RBFs) are used, given by

$$\mathbf{k}(\mathbf{y}_i, \mathbf{y}_j) = \alpha \exp(-\frac{\gamma}{2}(\mathbf{y}_i - \mathbf{y}_j)^\top(\mathbf{y}_i - \mathbf{y}_j) + \beta^{-1}\boldsymbol{\delta}_{ij}) \tag{10}$$

where $\boldsymbol{\delta}_{ij}$ is the Kronecker delta function, and α, β, γ are the RBF Kernel's parameters. According to [24], [10], the latter are set as $a = 1$ (unweighed function), $\beta = 1$ (unit width) and $\gamma = 2$. Finally, (4) is considered as the likelihood for $\mathbf{X}$ over $\mathbf{Y}$, therefore, the latent data $\mathbf{Y}$ can be computed by the Maximum Likelihood Estimation (MLE), which minimizes the corresponding negative log-likelihood $\mathbf{L}$ by making its gradient ($\frac{\partial \mathbf{L}}{\partial \mathbf{y}}$) close to zero. Thus, according to chain rule

$$\frac{\partial \mathbf{L}}{\partial \mathbf{K}} = \mathbf{K}^{-1}\mathbf{X}\mathbf{X}^\top\mathbf{K}^{-1} - \mathbf{D}\mathbf{K}^{-1} \tag{11}$$

$$\frac{\partial \mathbf{k}(\mathbf{y}, \mathbf{y}')}{\partial \mathbf{y}} = \gamma(\mathbf{y} - \mathbf{y}')\mathbf{k}(\mathbf{y}, \mathbf{y}') \tag{12}$$

Accordingly, the partial gradients of the kernel with respect to the latent points are computed and used in combination with (7) in a non-linear optimizer to obtain a latent variable representation of the data. Moreover, the computed gradients with respect to the parameters of the covariance matrix $\mathbf{K}$ are used to optimize the latent variables $\mathbf{Y}$.

B. Mapping across Latent Spaces - RANSAC

Herewith, the mapping between the two derived latent spaces is considered. The employed approach originates from the feature based image matching domain. In the latter, the RANdom SAmple Consensus (RANSAC) algorithm [22], [23] has been proposed as an effective and robust technique to compute feature correspondences. The RANSAC algorithm is a learning technique for estimating model parameters by randomly sampling the input data. This algorithm assumes that the noisy data points will not vote consistently for any single model and there exists a specific structure in the data to agree on a globally accepted model. Our goal is to find the homography $\mathbf{T}$ that best transforms points in the actor's latent space to the robot's latent space, minimizing the objective function J, as

$$J = \min_{\mathbf{T}} \sum_{n=1}^{N} \|\mathbf{y}_n - \mathbf{T}\mathbf{x}_n\| \tag{13}$$

The algorithm used is summarized in Alg. 1. At first, a sample subset containing minimal data items is randomly selected

Algorithm 1 RANSAC

while *Maximum Iterations not reached* **do**
    Select randomly a data group from the input data set
    Compute a homography
    Find the inlier group to this transformation
    Score the number of inliers
    Keep the transformation with the largest inlier group
end while

from the input dataset. A homography is readily computed from this data group. Then, the number of data points of the entire dataset that are consistent with the computed model, termed inliers, is calculated. Data points which do not fit the computed model within some error threshold, defined as a maximum deviation due to noise effect of the input data, are considered outliers.

C. Robotic Action Replication

The proposed LfD approach (Fig. 1), as explained in detail above, can be employed to associate a human action behavior with a robotic one. Effectively, a correspondence is established in the two latent action representations. In order to reproduce a demonstrated action by the robotic arm, a final step is needed to transform the representation in the robot's latent space to the robot's physical configuration space. This is schematically illustrated in Fig. 3, where INV-Tr denotes the latter transformation that gives rise to the actual reproduction of the demonstrated act.

Formally speaking, INV-Tr should be implemented as the inverse of GPLVM. Still, a rigorous mathematical derivation of the latter is not possible, due to the highly non-linear and, at the same time, probabilistic nature of GPLVM. The latter is not an obstacle in our approach, given that an acceptable approximation of the inverse GPLVM solution can be estimated. In our implementation this is addressed by computing off-line the latent representations of a sufficiently dense population of physical configurations of the robotic arm. In practice, we iterate over all arm-DoFs (degrees of freedom), and for each arm configuration the latent space representation is computed via the GPLVM. The above iterations are performed by (a) respecting the physical limits of each arm-joint, and (b) employing an appropriate iteration step. A small value in the latter results in a denser representation of the "configuration space - latent space" pairings. For points in the latent space that are not included in the above pre-computed pairs, the corresponding points in the physical robot space are derived by interpolation. Experimentally it has been established that neither the step-value nor the actual method of interpolation were critical. This is due to the fact that the sought inverse transformation should not provide an accurate (exact) replica of a demonstrated act, but rather a robotic action behavior sufficiently similar to the latter.

IV. EXPERIMENTAL RESULTS

The proposed framework has been implemented and validated in realistic HRC scenarios. The conducted experiments fall in two categories. In the first, novel action imitation was tested, whereas the second regards HRC task execution. In all our experiments two robotic systems were used, namely the JACO robotic arm and the NAO humanoid robot. JACO, by Kinova Robotics [25], is a six-joint arm manipulator. Its joints can be controlled independently either with position-control or with torque-control. JACO's compliant mode of operation is particularly useful in order to appropriately steer the arm and hence have it execute (learn) a desired action behavior. NAO robot by Aldebaran Robotics [26] comprises a small humanoid with 25 DoFs. Its joints are position-controlled, using closed-loop PID controllers. For the tasks at hand, robot localization was performed as in [27], posture control was accomplished as in [28], whereas only one arm of the humanoid with six DoFs was used. Evidently, the kinematics of the two systems, i.e. the JACO arm and the NAO arm, are completely different. The latter allowed us to validate the abstraction property of the latent space representation.

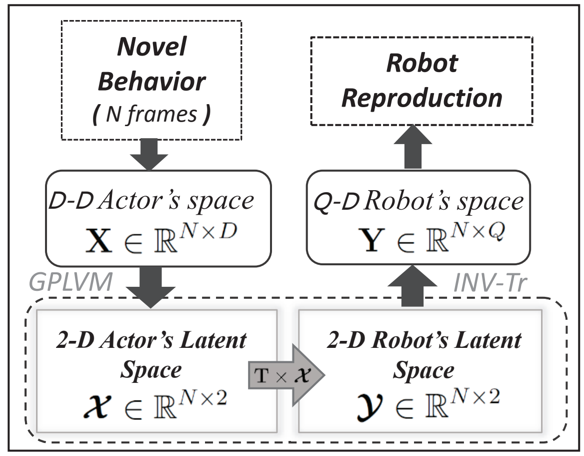


Fig. 3: The proposed Imitation Framework.

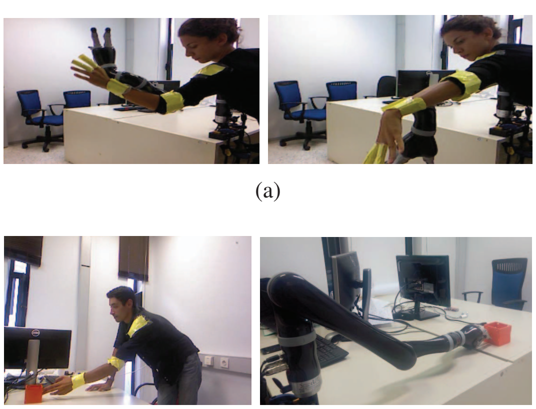
A. Imitation Framework Evaluation

Extensive experimentation has been performed to assess the proposed LfD methodology. The relevant experimental set-up involved kinesthetic teaching in order to demonstrate certain action behaviors to the robotic arm. More specifically, the compliant mode of the JACO arm was used in order to physically steer the arm to execute various motion trajectories. Each trajectory corresponded to a specific reaching action, such as pushing forward, reaching sidewards, reaching from the top, etc.

In the above trials, the physical spaces of the demonstrator (human) and the robot were recorded by utilizing information for the joints of the respective arms. In the case of the human actor the three arm joints were used, namely shoulder, elbow and wrist, and for each one of them the 3D coordinates (x, y, z) were derived. Prominent color markers were placed on each joint, and a Kinect RGB-D camera was utilized to unambiguously track the joints and extract their 3D coordinates (see Fig. 4a) [29], [30], [31]. By tracking the three arm joints of the teacher as above, a $9D$ representation of the observed physical space is formed. The above procedure, besides being simple and robust, facilitated the execution of multiple demonstrations.

Regarding the robot's physical space, we obtained the angles of the arm's six joints during a performed act via the available encoders. This resulted in a $6D$ physical representation for the robotic system. The above described procedure, i.e. action demonstration, action replication by the robotic arm, registration of the physical space parameters for both arms, was repeatedly executed for an adequate number of reaching actions. More specifically, the training phase was completed with 30 actions, of 30 fps each. In our experiments we concluded that a number of 30 actions was utterly satisfactory, by gradually increasing the training data set, starting from a rather small number of 10 actions. Although 10 actions might seem a small number, experimentally it was verified that they span densely the action space and, moreover, are adequate for the considered HRC tasks. By gradually increasing the action count to 30, algorithm's scalability has been verified. It is noted at this point that actions with different durations are regarded different in our implementation, since time is not directly taken into account in the proposed method but rather is considered as an intrinsic feature of the motion-action captured. Examples of the performed actions by the human teacher are shown in Fig. 4a. In the same figure, the physical attachment of the robotic arm to the human arm is also illustrated. Accordingly, the robotic arm performs a motion act similar to the one performed by the human arm.

Following the above, the latent space representation has been extracted for both cases, human and corresponding robot trajectories. A home-made implementation of GPLVM has been utilized for this task. Experimentally it was established that a latent space dimension of 2 was appropriate for the tested actions. Nonetheless, theoretical verification of the above has also been obtained via maximum-likelihood estimation of the intrinsic input data dimension, as proposed in [32]. For illustrative purposes, four sample motions have been identified, and the corresponding latent space representations are provided in Fig. 5, with the four motions being marked as $D1 \dots D4$. In Fig. 5a the named representations are given for the case of the demonstrator, whereas Fig. 5b shows the respective plots for the robotic arm. As can be observed, the obtained representations are appropriate for the given task, since different actions result in distinct trajectories in the latent spaces. Moreover, the neighbouring property of GPLVM is effectively demonstrated, resulting in adjacent latent points for adjacent points in the physical space. As a last step, RANSAC was used to establish the correspondence across the two latent spaces and form the latent hyperspace.



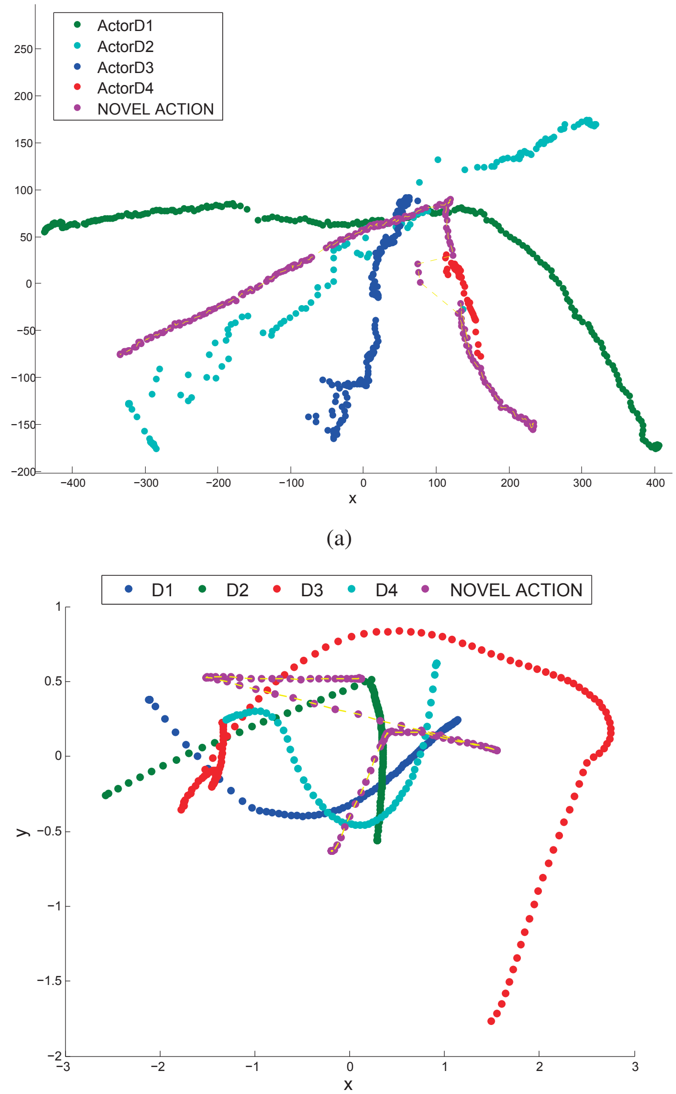
(a) (b)

Fig. 4: (a) Kinesthetic training of the robotic arm. (b) A novel human motion behavior (left), reproduced by the robotic arm (right).

To examine the validity of the extracted latent hyperspace, a set of experiments was conducted whereby a newly presented action behavior was asked to be reproduced by JACO. This behavior was either performed by a different human actor, or by the same one like above but this time performing a behavior not included in the set of learned actions. Fig. 4b shows a sample experiment, where a different actor performs a "docking action" (pushing the orange box). This 120-frame sequence of $9D$ vectors is projected to the formatted actor's latent space of Fig. 5a, marked as *NOVEL ACTION*. By left multiplying the latter with the transformation matrix obtained by RANSAC, the initial actor's trajectory is projected to the JACO's latent space, shown in Fig. 5b. Finally, the representation of the indicative movement in JACO's joints angles is computed by applying INV-Tr, resulting in a representation in the actual joint space of JACO. This action can be effectively reproduced by JACO; a snapshot of this behavior is shown in Fig. 4b.

B. HRC Task Execution

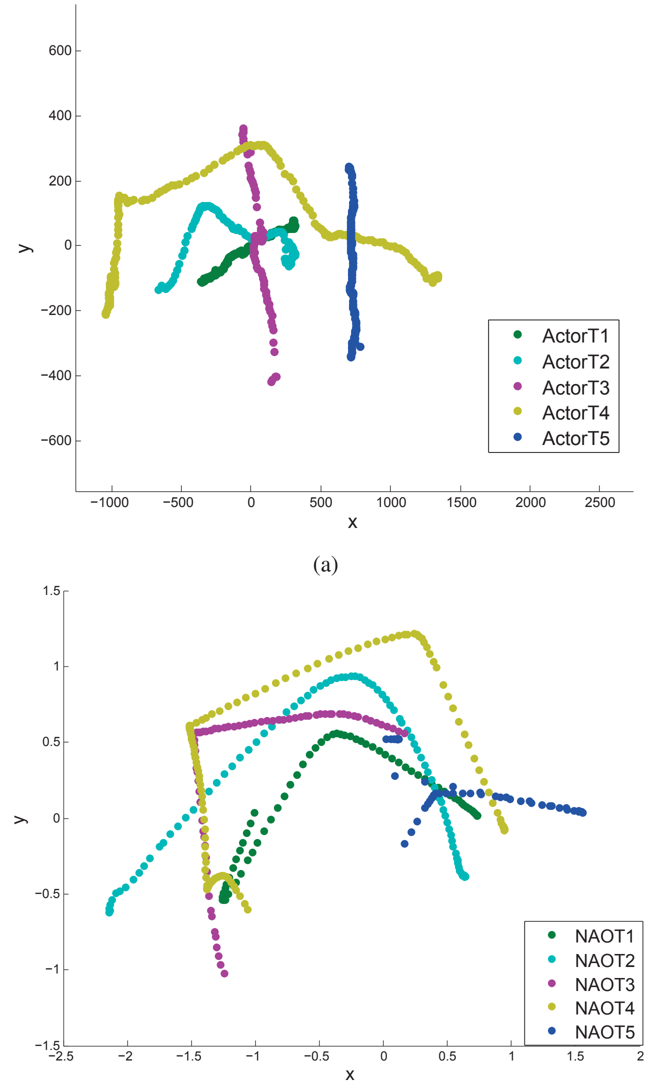
The second category of experiments regards the validation of realistic HRC scenarios. Accordingly, the learned actions by a robotic system were utilized in a real HRC context, where the robotic system was expected to perform specific actions. Experimentation in this case was performed with both robotic systems, i.e. the JACO arm and the NAO humanoid; in the case of the latter, only the arm was controlled to perform action behaviors. The actual experiments conducted assumed HRC scenarios, whereby a human was opening a closed compartment and removing an object from it. Subsequently, the robot was expected to perform the correct action to properly close the compartment. Example cases of compartments tested involve drawers, closets and cabinets. The respective actions that the robotic systems should perform consisted of "forward pushing" (close a drawer), "sideward reaching" (close an open closet door) and "reaching from the top" (close a cabinet). Furthermore, the robots were trained with two additional actions namely "hello-waving" and "sidewards pointing". Those additional actions were included in order to furnish the robots with richer repertoire of actions and hence operate more realistically in the HRC scenarios.



(a)

(b)

Fig. 5: Extracted Latent Hyperspace formulated by five sample trajectories marked as $D1-D4$ and *NOVEL ACTION*, accomplished by actor and JACO, respectively: (a) Actor's Latent Space, and (b) JACO's Latent Space.



(a)

(b)

Fig. 6: Extracted NAO's Latent Hyperspace, formulated by five sample trajectories $T1-T5$, accomplished by actor and NAO, respectively: (a) Actor's Latent Space, and (b) NAO's Latent Space.

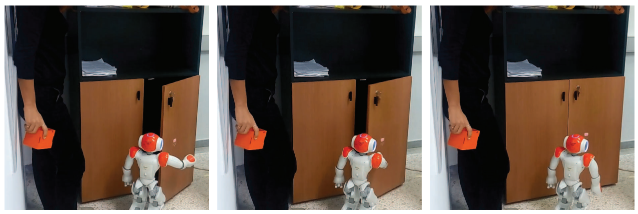
The above actions were demonstrated by a human actor and learned by both robotic systems, in a similar way as already explained in Section IV-A. Figure 6 presents the latent spaces that were registered for the human actor and the NAO arm in this training session (due to space limitations, the latent space representation for the JACO arm is not provided). In the actual execution of the HRC scenario, visual recognition of the relevant compartment was used in order to index in the latent space and retrieve the correct action representation. In addition, the robotic action was triggered by visual detection of the removed object based on color information [33]. Interestingly, in all HRC experiments conducted, both robotic systems collaborated successfully and capably closed the respective compartment by employing the correct action behavior. Example cases are illustrated in Fig. 7 for both robotic systems employed in the relevant experiments.

C. Quantitative Evaluation

According to [34], the definition of imitation metrics in LfD is task specific. In other words, the task under consideration may give rise to appropriate metrics for quantitatively characterizing execution of the learned behavior. A general formula has been proposed in [34], as:

$$J = \|\xi_a^\theta - \xi_{rob}^\theta\|^2_{\mathbf{W}^\theta} + \|\xi_a^x - \xi_{rob}^x\|^2_{\mathbf{W}^x} + \|\xi_a^y - \xi_{rob}^y\|^2_{\mathbf{W}^y} \quad (14)$$

(a)



(b)

Fig. 7: HRC scenario whereby a robot assumes a learned action behavior to close an open closet door. (Top) JACO arm, (Bottom) NAO humanoid robot arm.

where, $\xi_a^{\theta,x,y}$ are the actor's variables that are considered in the demonstrated action, e.g. 3D-coordinates of the arm's joints. Similarly, $\xi_{rob}^{\theta,x,y}$ are the corresponding robot's ones, and $\mathbf{W}^{\theta,x,y}$ are appropriate weights that effectively scale the contribution of each variable in the final result. In order to quantitatively assess the reproduction of learned actions by a robotic platform, we have instantiated (14) as follows:

$$ERROR_{Traj} = \|\xi_a^{EE} - \xi_{rob}^{EE}\|^2 \qquad (15)$$

where ξ_a^{EE} and ξ_{rob}^{EE} stand for the end-effector's 3D trajectories ($3D$ coordinates) accomplished by actor and robot, respectively. For the set of experiments conducted in our study, (15) captures the error between the demonstrator's and the robot's $3D$-trajectories and therefore quantitatively evaluates task reproduction. Fig. 8 illustrates the end-effectors' $3D$-trajectories followed by the three agents involved, namely Actor, NAO and JACO, during the two HRI motion acts "*Close an open closet*" and "*Close an open drawer*", respectively.

An additional error metric can be defined similarly to (15), by considering Mahalanobis distances of the respective trajectories in the latent spaces (actor's and robot's):

$$ERROR_{Latent} = \|y_a - y_{rob}\|^2_{\mathbb{S}^{-1}} \qquad (16)$$

where y_a, y_{rob} are latent-space points (actor's and robot's, respectively), and $\mathbb{S}$ is the covariance matrix derived from GPLVM. Please note that direct comparison of raw latent-space trajectories is meaningless, and therefore points from one latent space are first transformed by the resultant RANSAC transformation matrix, and are effectively mapped to the other latent space. This step allows two latent-space trajectories from different latent spaces (actor's and robot's) to be quantitatively compared.

Table I and Table II summarize the quantitative evaluation results, as derived from (15) and (16), for each of the five sample motion acts that were used to formulate the latent

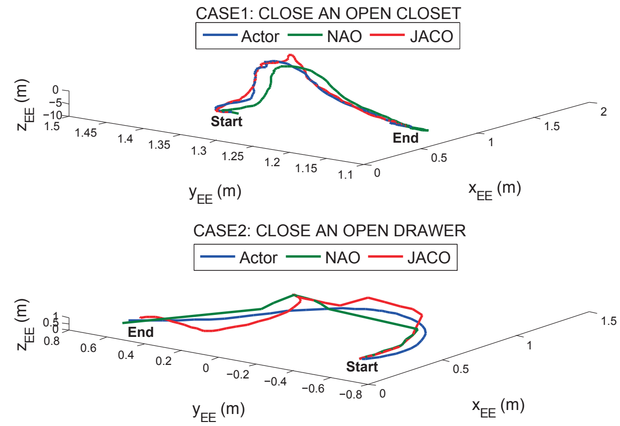


Fig. 8: Trajectories of end-effectors' movement during the two aforementioned HRC tasks.

TABLE I: JACO Performance Metrics

Action	$ERROR_{Traj}(m)$	$ERROR_{Latent}$
T1	0.086	0.099
T2	0.078	0.095
T3	0.049	0.073
T4	0.088	0.091
T5	0.064	0.098

TABLE II: NAO Performance Metrics

Action	$ERROR_{Traj}(m)$	$ERROR_{Latent}$
T1	0.089	0.096
T2	0.068	0.099
T3	0.063	0.093
T4	0.040	0.088
T5	0.056	0.099

hyperspace of Fig. 6. The first two rows in each table correspond to the assumed trajectories for the "*Close an open closet*" and "*Close an open drawer*", respectively (see Fig. 8). As can be observed from the tabulated figures in Tables I and II, both robotic systems performed the learned actions with high accuracy. The $3D$-trajectory errors are very small (less than 10 centimeters in all experiments). More importantly, the computed errors are negligible for the studied imitation tasks, and do not affect the behavior expressed by the robotic system. Moreover, the low error values in the latent-space trajectories are commendable. As such they serve as the experimental verification of the adopted means to obtain the latent space representation and effectively justify the parameter choices in our implementation.

V. CONCLUSIONS

In this work we have studied and validated an LfD methodology and have successfully applied this in HRC task execution scenarios. LfD is approached by employing a latent space representation for both demonstrated and learned action behaviors. This intermediate representation effectively records significant features of motion acts, abstracting at the same time irrelevant differences that stem from different embodiments and arm kinematics. The employed GPLVM transformation from the actor's physical space to the actual latent space was proven very effective in appropriately encoding the salient, pertinent information from a motion act. Learned action behaviors were in turn employed to accomplish HRC task completion. Our results demonstrate the successful involvement of the robotic system in relevant tasks, whereby learned behaviors are appropriately executed.

The so far encouraging results attest for the validity and effectiveness of the proposed approach. Nevertheless, open issues for investigation still exist and constitute promising areas of future work. In the current work an action behavior was deemed successful provided that the reproduced trajectory was appropriate for the task at hand. For a more systematic assessment of learned actions we plan to introduce suitable error measures to cater for mismatches in the reproduced trajectories. Moreover, we intend to study additional relevant parameters, such as velocity, acceleration, etc. Such a holistic approach to LfD may clearly support more complex HRC scenarios, such as cases where action timing plays a role in human-robot cooperation. The latter scenarios, augmented with the presence of additional agents (humans and/or robots), constitute interesting research directions.

Acknowledgment

This work has been partially supported by the EU FET Proactive grant (GA: 641100) TIMESTORM - Mind and Time: Investigation of the Temporal Traits of Human-Machine Convergence.

References

[1] B. Argall, S. Chernova, M. Veloso, and B.Browning, "A survey of robot learning from demonstration", *Rob. Auton. Syst.*, vol. 57, no. 5, pp. 469–483, 2009.

[2] S. Schaal, A. Ijspeert, and A. Billard, "Computational approaches to motor learning by imitation," *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, vol. 358, pp. 537–547, 2003.

[3] M. Matsushima, T. Hashimoto, M. Takeuchi,and F. Miyazaki, "A Learning Approach to Robotic Table Tennis," *IEEE Transactions on Robotics*, vol. 21, pp. 767771, 2005.

[4] A. Ijspeert, J. Nakanishi, and S. Schaal, "Movement imitation with nonlinear dynamical systems in humanoid robots", *Proc. 2002 IEEE Int. Conf. Robot. Autom. (Cat. No.02CH37292)*, vol. 2, pp. 1–6, 2002.

[5] S. Schaal, J. Peters, and J. Nakanishi, "Learning movement primitives", *Robot. Res.*, pp. 1–10, 2005.

[6] S. M. Khansari-Zadeh and A. Billard, "Learning Stable Non-Linear Dynamical Systems with Gaussian Mixture Models", *IEEE Trans. Robot.*, vol. 27, pp. 943–957, 2011.

[7] M. S. Khansari-Zadeh and A. Billard, "A dynamical system approach to realtime obstacle avoidance", *Auton. Robots*, vol. 32, no. 4, pp. 433–454, 2012.

[8] P. Pastor, H. Hoffmann, T. Asfour, and S. Schaal, "Learning and generalization of motor skills by learning from demonstration", *IEEE Int. Conf. Robot. Autom. 2009. ICRA '09*, pp. 763–768, 2009.

[9] N. Lawrence and A. Hyvrinen, "Probabilistic non-linear principal component analysis with gaussian process latent variable models", *Journal of Machine Learning Research*, vol. 6, pp. 1783–1816, 2005.

[10] M. K. Titsias and N. D. Lawrence, "Gaussian process latent variable models for visualisation of high dimensional data", in *Adv. in Neural Inf. Proc. Sys*, 2004.

[11] S. Hou, A. Galata, F. Caillette, N. Thacker, and P. Bromiley, "Articulated pose estimation in a learned smooth space of feasible solutions", in *IEEE International Conference on Computer Vision (ICCV)*, 2007.

[12] T.-P. Tian, R. Li, and S. Sclaroff, "Articulated pose estimation in a learned smooth space of feasible solutions", *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, vol. 0, p. 50, 2005.

[13] A. Gupta, T. P. Chen, F. Chen, D. Kimber, and L. S. Davis, "Context and observation driven latent variable model for human pose estimation", in *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008), 24-26 June 2008, Anchorage, Alaska, USA*, 2008.

[14] S. Quirion, C. Duchesne, D. Laurendeau, and M. Marchand, "Comparing gplvm approaches for dimensionality reduction in character animation", *Journal of WSCG*, vol. 16, no. 1-3, pp. 41–48, 2008.

[15] X. Sun, X. Zhang, T. Cao, J. Yang, and J. Sun, "Voice conversion using a two-factor gaussian process latent variable model", *Pr. Elektrotechniczny*, vol. 0, pp. 318–324, 2012.

[16] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis", *Chemom. Intell. Lab. Syst.*, vol. 2, pp. 37–52, 1987.

[17] A. Shon, D.B.Grimes, C. Baker, and R. P. N. Rao, "A Probabilistic Framework for Model-Based Imitation Learning", *Proc. Twenty-Sixth Annu. Conf. Cogn. Sci. Soc.*, pp. 1237–1242, 2004.

[18] D. B. Grimes and R. P. N. Rao, "Learning actions through imitation and exploration: Towards humanoid robots that learn from humans", *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 5436, pp. 103–138, 2009.

[19] D. Grimes, D. Rashidand, and R. P. N. Rao, "Learning Nonparametric Models for Probabilistic Imitation", *Adv. Neural Inf. Process. Syst. 19*, vol. 19, pp. 521–528, 2007.

[20] Monica N. Nicolescu and Maja J. Mataric, "Learning and Interacting in Human-Robot Domains", *IEEE Transactions on Systems, Man, and Cybernetics*, Part A, vol. 31, 419430, 2001.

[21] P. Evrard, E. Gribovskaya, S. Calinon, A. Billard, and A. Kheddar, "Teaching physical collaborative tasks: object-lifting case study with a humanoid", in *Humanoids*, IEEE, 2009, pp. 399–404.

[22] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography", *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[23] D. A. Forsyth and J. Ponce, *Computer Vision A Modern Approach*. Prentice Hall, 2003.

[24] M. Mongillo, "Choosing basis functions and shape parameters for radial basis function methods", *SIAM Undergrad. Res. Online*, pp. 190-209, 2011.

[25] Kinova Robotics, "Jaco documentation", only available online: http://kinovarobotics.com/products/jaco-robotics/.

[26] Aldebaran Robotics, "Nao documentation", 2012, only available online: www.aldebaran-robotics.com/documentation.

[27] H. Baltzakis and P. Trahanias, "Hybrid Mobile Robot Localization using Switching State-Space Models", *IEEE Intl. Conf. Robotics and Autom*, ICRA2002, pp.366-373, Washington D.C., USA, May 11-15, 2002.

[28] S. Piperakis, E. Orfanoudakis, and M. Lagoudakis, "Predictive Control for Dynamic Locomotion of Real Humanoid Robots", in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*, Sept 2014, pp. 4036–4043.

[29] H. Baltzakis, M. Pateraki, P. Trahanias, "Visual tracking of hands, faces and facial features of multiple persons", *Machine Vision and Applications*, pp. 1-17, 2012.

[30] M. Sigalas, M. Pateraki, and P. Trahanias, "Robust articulated upper body pose tracking under severe occlusions", in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Chicago, USA, Sep. 2014.

[31] M. P. M. Sigalas and P. Trahanias, " Full-Body Pose Tracking – The Top View Reprojection Approach", *IEEE Trans. Pattern Anal. Mach. Intell. (T-PAMI)*, vol. 38, 2016.

[32] E. Levina and P. J. Bickel, "Maximum Likelihood Estimation of Intrinsic Dimension," in *NIPS*, 2004.

[33] G. Bradski, "opencv toolbox", *Dr. Dobb's Journal of Software Tools*, 2000.

[34] S. Calinon, F. Guenterlorent, and A. Billard, "On learning, representing, and generalizing a task in a humanoid robot", *IEEE Trans. Syst. Man, Cybern. Part B Cybern.*, vol. 37, pp. 286–298, 2007.