Open access • Journal Article • DOI:10.1109/TII.2020.2984482

Simultaneously Encoding Movement and sEMG-Based Stiffness for Robotic Skill Learning — Source link

Chao Zeng, Chenguang Yang, Hong Cheng, Yanan Li ...+1 more authors

Institutions: South China University of Technology, University of Electronic Science and Technology of China, University of Sussex

Published on: 01 Feb 2021 - IEEE Transactions on Industrial Informatics (IEEE)

Topics: Impedance control, Robot and Stiffness

Simultaneously encoding movement and sEMG-based stiffness for robotic skill learning

Article (Accepted Version)

Simultaneously Encoding Movement and sEMG-based Stiffness for Robotic Skill Learning

Chao Zeng, *Student Member, IEEE*, Chenguang Yang, *Senior Member, IEEE*,
Hong Cheng, *Senior Member, IEEE*, Yanan Li, *Member, IEEE*, Shi-Lu Dai, *Member, IEEE*

***Abstract*—Transferring human stiffness regulation strategies to robots enables them to effectively and efficiently acquire adaptive impedance control policies to deal with uncertainties during the accomplishment of physical contact tasks in an unstructured environment. In this work, we develop such a physical human-robot interaction (pHRI) system which allows robots to learn variable impedance skills from human demonstrations. Specifically, the biological signals, i.e., surface electromyography (sEMG) are utilized for the extraction of human arm stiffness features during the task demonstration. The estimated human arm stiffness is then mapped into a robot impedance controller. The dynamics of both movement and stiffness are simultaneously modeled by using a model combining the hidden semi-Markov model (HSMM) and the Gaussian mixture regression (GMR). More importantly, the correlation between the movement information and the stiffness information is encoded in a systematic manner. This approach enables capturing uncertainties over time and space and allows the robot to satisfy both position and stiffness requirements in a task with modulation of the impedance controller. The experimental study validated the proposed approach.**

***Index Terms*—Adaptive Impedance Control; Multimodality; Human-robot interaction systems**

I. Introduction

Programming by demonstration (PbD) is regarded as one of the most promising ways to enable robots to efficiently acquire the ability of performing tasks by transferring human dexterous manipulation skills to them [1–3]. Especially, for in-contact tasks where force profiles in addition to positional profiles need to be regulated [4], PbD allows relaxing the analytical burden required for the process of human-to-robot physical skills transfer [5]. One of the challenges is to enable a robot to learn human-like behaviours with flexibility and impedance adaptation [6–9]. Especially for force-dominant tasks [10], this challenge needs to be addressed urgently.

A potential way to equip the robots with a high level of interaction capabilities is to explore human's underlying sensorimotor principles and integrate multimodal information into the robotic control policies [11–13]. Neurological research has shown that humans can adapt limb impedance subconsciously to deal with different situations when performing tasks thanks to the central nervous system (CNS). The development of a PbD system for transferring such a kind of impedance regulation mechanism from a human instructor to a robot, therefore, can to a large extent facilitate the improvement of robot learning dexterous skills. Actually, a number of studies inspired by the research in human neuroscience and biomechanics have sought to transfer the sEMG-based human limb stiffness regulation skills to the robots and achieved encouraging results [14, 15]. By utilizing the sEMG signals collected from the human limb for representing the activation level of the human muscles, the human arm joint or end-effector stiffness can be estimated and in real-time extracted during the task execution [16–19].

One of the most significant benefits of the bio-inspired human-to-robot impedance feature transfer is that adaptive impedance control for robotic arms can be realized, which has demonstrated better performance than position control or invariant impedance control for in-contact tasks by a number of works (e.g., [4, 14, 15, 20, 21]). In [22], a learning framework was established for achieving variable impedance control for robots. However, the variable impedance profiles are obtained via a time-consuming process which may limit the framework's functionality available to real world applications. In [4, 20], the variable stiffness profile in task space is computed partially based on the measured forces through an external high-accuracy force sensor mounted on the robot end-effector, thus increasing the cost of the HRI system. Compared with these methods, the sEMG-based human-to-robot impedance transfer has the following advantages: i) the cost is no longer a problem since the human limb sEMG signals can usually be collected via some cheap devices (e.g., MYO Armhand); ii) the human limb stiffness profiles can be extracted in a real time manner, guaranteeing the efficiency of the realization of the variable impedance control [23]; and iii) in this way, most importantly, the human factors (e.g., flexibility and adaptability) are taken into account within the process of the human-to-robot skill transfer, which can facilitate the interactions between humans and robots [24] and robotic dexterous manipulations [25].

One issue of this impedance skill transfer is to represent the stiffness profile during robot task execution. This is because simply reproducing the learned control policies from humans is often insufficient for the robot to successfully accomplish a specific task, especially when dealing with the task situations different from the demonstration. One solution to this issue is to equally treat movement trajectories and stiffness profiles by simultaneously encoding them in a systematic manner [26]. In [27, 28], a framework was proposed to achieve this goal by

*This work was partially supported by National Nature Science Foundation (NSFC) under Grants 61861136009, 61811530281 and 61973129.

C. Zeng, C. Yang and S.-L. Dai are with School of Automation Science and Engineering, South China University of Technology.

H. Cheng is with Center for Robotics, University of Electronic Science and Technology of China, Chengdu, China.

Y. Li is with Department of Engineering and Design, University of Sussex, Brighton BN1 9RH, U.K.

Corresponding author: C. Yang. Email: cyang@ieee.org

encoding movement and stiffness in parallel using dynamic movement primitives (DMP). DMP encodes each dimension of movement and stiffness separately, allowing the learning of the control policies in each dimension without changes to the basic approach [29]. Modeling the dimensions independently, however, can not leave room for exploiting the correlation between movement trajectories and stiffness profiles [4].

In [30], the framework combining Hidden Markov Model (HMM) and GMR was proposed to generate a probabilistic model of demonstrated data, by modeling a joint probability density function between the position and the velocity using HMM, and by generalizing the learned skills through regression using GMR. In [31], the hidden semi-Markov model (HSMM) was further used to encode the duration information of each HMM state such that the duration and position information can be encapsulated in a robust manner with parameterization on the involvement of both temporal and spatial constraints. In [20], GMM was used to model both movement and force patterns for robot learning impedance behaviors. In [4], the HSMM-GMR model was proposed to encode the demonstration data, plus the force profiles sensed at the robot end-effector. Inspired by the encouraging results of the use of these models, our work develops a PbD approach based on the HSMM-GMR model, encoding the demonstrated data including the stiffness profiles extracted from the human instructor.

The contributions of this work are summarized as follows:

i) We develop a PbD system for human-robot variable impedance skill transfer, which enables the robots to physically interact with the environment by directly adapting the human instructor's arm impedance profiles without force sensing at the robotic manipulator.

ii) A novel approach is proposed to encode both the demonstrated movement trajectories and the stiffness profiles, considering the correlation between the physical information and the biological information in a systematic manner. The evolution of the stiffness profiles depends on the position information, which can improve the performance of the robot's task execution.

II. Methodology

A. The PbD system overview

The overview diagram of the proposed PbD system is shown in Fig. 1, which includes three phases:

Demonstration: In a conventional PbD system, a human instructor demonstrates the skills to accomplish one specific task, during which the demonstration trajectories (sometimes including forces) are recorded for subsequent usage. In our PbD system, the instructor's arm sEMG signals are also extracted for stiffness estimation of the human arm (see Subsection II-B). The teleoperation based on a dual-arm control strategy is used for task demonstration, in which a haptic feedback mechanism is introduced for easy regulation of the instructor's muscle activations [15].

Model training: The second phase is model training (see Subsection II-C 1). The demonstration profiles are fitted into the HSMM model and the model parameters are accordingly

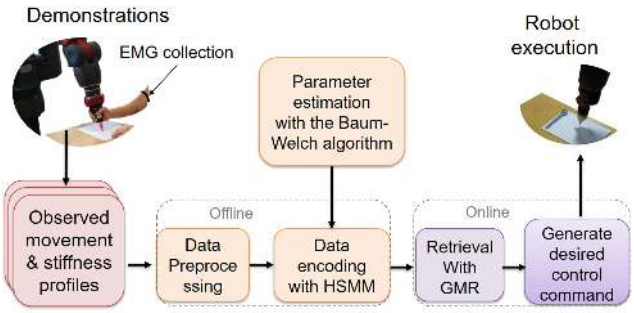


Fig. 1: Workflow diagram of learning from demonstration for human robot variable impedance skill transfer.

computed. In this way, the dynamics of the movement and the stiffness are modeled and learned.

Robot task execution: Finally, the robot executes the learned task based on the desired control policies, which are generated by GMR in accordance with the estimated parameters and the current measured robot state (see Subsection II-C 2).

B. Stiffness Extraction

1) Human arm Cartesian impedance model: Generally, the dynamic behaviour of the human arm during human-robot interaction is usually described as a mechanical impedance which relates to the desired force $\boldsymbol{F}$ of the limb endpoint to a deviation from the desired position $\boldsymbol{x}$. It is defined as below

$$\boldsymbol{F} = \boldsymbol{I}_H\ddot{\boldsymbol{x}} + \boldsymbol{D}_H\dot{\boldsymbol{x}} + \boldsymbol{K}_c\boldsymbol{x} \tag{1}$$

where $\boldsymbol{I}_H$, $\boldsymbol{D}_H$ and $\boldsymbol{K}_c$ are inertia, damping and endpoint Cartesian stiffness matrices of human arm, respectively. By ignoring the negligible influence of the muscle mass distribution on $\boldsymbol{I}_H$ in the vicinity of the predefined posture [14], the control objective (1) can be simplified by dropping the inertia term in the control loop with the following form

$$\boldsymbol{F} = \boldsymbol{D}_H\dot{\boldsymbol{x}} + \boldsymbol{K}_c\boldsymbol{x} \tag{2}$$

where the limb endpoint stiffness matrix $\boldsymbol{K}_c$ takes the form as below

$$\boldsymbol{K}_c = \begin{bmatrix} \boldsymbol{K}_P & \boldsymbol{K}_{PR} \\ \boldsymbol{K}_{RP} & \boldsymbol{K}_R \end{bmatrix} \tag{3}$$

with $\boldsymbol{K}_P$ relating forces to positional profiles, $\boldsymbol{K}_{PR}$ relating forces to rotational profiles, $\boldsymbol{K}_{RP}$ relating torques to positional profiles, and $\boldsymbol{K}_R$ relating torques to rotational profiles. For simplicity, the off-diagonal components (owing to the existence of cross-joint muscles) can be assumed as $\boldsymbol{K}_{PR} = \boldsymbol{K}_{RP} = 0$ during the experiments [24]. It means that positional profiles only result in force corrections, and rotational profiles only result in torque corrections. Correspondingly, the damping matrix $\boldsymbol{D}_H$ is defined in the same structure as the endpoint stiffness matrix, i.e.,

$$\boldsymbol{D}_H = \begin{bmatrix} \boldsymbol{D}_P & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{D}_R \end{bmatrix} \tag{4}$$

In our work, the stiffness $\boldsymbol{K}_c$ is estimated based on the sEMG signals extracted from the human instructor's limb. The robot end-effector force which can be seen as a consequence

of the stiffness modulation [32], therefore, can adapt to task requirements.

2) sEMG-based human arm endpoint stiffness acquisition from demonstrations: This module is developed to enable the robot learning of the impedance regulation strategy from the human instructor's demonstrations. To this end, the human instructor's arm endpoint stiffness is estimated first.

The relation between the end-effector stiffness of the human arm and the joint stiffness is described as [23]

$$\boldsymbol{K}_c(p,x) = \boldsymbol{J}^{+T}(x)[\boldsymbol{K}_J(p,x) - \boldsymbol{G}_J(x)]\boldsymbol{J}^{+}(x) \tag{5}$$

with

$$\boldsymbol{G}_J(x) = \frac{\partial \boldsymbol{J}(x)\boldsymbol{f}_0}{\partial x} + \frac{\partial \boldsymbol{\tau}_g(x)}{\partial x} \tag{6}$$

and

$$K_J(p,x) = c(p)K_J^{min} \tag{7}$$

where $\boldsymbol{K}_c(p,x)$ and $\boldsymbol{K}_J(p,x)$ are the Cartesian stiffness matrix and the joint stiffness matrix of the human instructor, respectively, with p and x denoting the muscle activity (considered as the stiffness indicator in this work) and the joint angle vector, respectively. $\boldsymbol{J}(x)$ represents the human arm Jacobian matrix. $\boldsymbol{G}_J(x)$ takes into account the effect of arm geometry in the presence of external force $\boldsymbol{f}_0$ and gravity load $\boldsymbol{\tau}_g(x)$. (5) and (6) suggest that the Cartesian stiffness profile depends on the joint stiffness (through the muscle activities of contraction and co-contraction), the exerted external force and gravity. In our case, for simplicity $\boldsymbol{f}_0$ and $\boldsymbol{\tau}_g(x)$ are dropped within the identification of the human arm parameters as suggested in [23]. $c(p)$ is a variable coefficient that will be introduced later. K_J^{min} is the minimal joint stiffness.

The muscle activity indicator is obtained based on the extraction of the human instructor's sEMG signals. The collected raw sEMG signals are first smoothed using a moving average process, subsequently filtered by a low-pass filter (2nd order low-pass filter with cutoff frequency 2.5 Hz). In this way, an envelope from the raw sEMG signals is extracted.

According to [23, 33], we use a single joint stiffness as the estimation of the human instructor's arm stiffness along the relevant axis in the Cartesian space. It is reasonable to do so because a stiffening pattern can be observed among different antagonistic pairs thanks to the human arm muscle activations following a synergistic way [34]. In our work, the antagonistic muscles Biceps and Triceps are utilized to compute the muscle stiffness indicator as below

$$p = \frac{1}{W}\Big(\sum_{k=1}^{W-1} E_B(t-k) + \sum_{k=1}^{W-1} E_T(t-k)\Big) \tag{8}$$

where W is the predefined window size. $E_B(\cdot)$ and $E_T(\cdot)$ denote the amplitudes of the enveloped sEMG signals of Biceps and Triceps, respectively. t and k represent the current sampling time and the sample point, respectively.

Then, the muscle stiffness indicator is mapped to the coefficient $c(p)$ in the model of the human arm endpoint stiffness by

$$c(p) = 1 + \frac{\gamma_1[1 - e^{-\gamma_2 p}]}{1 + e^{-\gamma_2 p}} \tag{9}$$

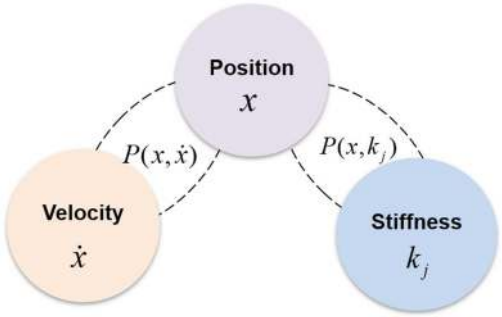


Fig. 2: Graphical representation of the two joint Gaussian distributions $\mathcal{P}(x,\dot{x})$ and $\mathcal{P}(x,k_j)$, respectively, encoded in each of two continuous HSMMs of K states. The output distribution of each state of these two HSMMs is represented by a Gaussian locally encoding variation and correlation information between position and velocity, and position and stiffness, respectively.

where γ_1 and γ_2 are predefined constant coefficients affecting the amplitude and the shape of $c(p)$, respectively. The estimation of the desired Cartesian stiffness profile can be then obtained in accordance with [14] requiring an off-line identification and calibration process before demonstrations, which will not be detailed in this paper.

Subsequently, the normalized human arm endpoint stiffness is mapped to the robotic arm endpoint stiffness $\boldsymbol{K}_r$ by

$$\boldsymbol{K}_r = \begin{cases} \boldsymbol{K}_r^{max} & K_r > K_r^{max} \\ \gamma \boldsymbol{K}_c & K_r^{min} < K_r < K_r^{max} \\ \boldsymbol{K}_r^{min} & K_r < K_r^{min} \end{cases} \tag{10}$$

where $\boldsymbol{K}_r^{max}$ and $\boldsymbol{K}_r^{min}$ are the pre-set values of the robotic arm endpoint stiffness. γ is a pre-set constant enabling the robot to work with an endpoint stiffness within a proper range. To summarize, after the parameters of the human arm stiffness model are estimated, during one specific demonstration the human arm endpoint stiffness can be extracted by (9), (7) and (5), and further mapped to the robot endpoint through (10).

C. Demonstration data modeling with the HSMM model

Our developed method based on the HSMM model and the GMR model applied to PbD is first proposed in [30] and [31], respectively. The work [4] combines these two models considering forces for in-contact tasks. In order to achieve human robot variable impedance skill transfer, our method takes into account stiffness profiles recorded during demonstrations.

Our method includes two basic steps. First, the HSMM is used to model the dynamics of the demonstration trajectories with a continuous Gaussian observation probability distribution assigned to each HSMM state. The parameters of this model are learned with the Baum-Welch algorithm from the demonstration trajectories in an offline manner. Then, based on the estimated parameters the GMR is utilized to compute the desired stiffness profiles at each time step [35].

1) Encoding Demonstration Data with HSMM: Given a set of M demonstrations, each demonstration $m \in \{1, \ldots, M\}$ consists of a set of T_{max} samples representing the robot movement trajectories in joint space and stiffness extracted from the human instructor's arm. The HSMM with K states is parametrized by:

$$\Theta = \{\{a_{i,j}\}_{j=1,j\neq i}^{K}, \pi_i, \mu_i^D, \Sigma_i^D, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^{K} \tag{11}$$

where π_i is the initial probability of the ith state. a_{ij} is the transition probability from state j to i. μ_i^D and Σ_i^D are sets of mean values and variances, respectively, modelling the K Gaussian parametric duration distributions. $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$ are sets of mean vectors and covariance matrices of the K Gaussian joint observation probabilities, respectively.

The ith state duration probability density function is defined as

$$p_i^D(t) = \mathcal{N}(t; \mu_i^D, \Sigma_i^D) \tag{12}$$

with $t = 1, \ldots, t_{max}$. t_{max} is the maximum allowed duration of a HSMM state which is usually determined by

$$t_{max} = \eta \frac{T_{max}}{K} \tag{13}$$

where η is a scaling factor that is set $2 \sim 3$ so as to guarantee that $p_i^D(t)$ is well defined even if EM converges to poor local optima [31].

The observation probability at each time step t for the ith state is defined by

$$p_i(z_t) = \mathcal{N}(z_t; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \tag{14}$$

where ${}^1\boldsymbol{z}_t = [\boldsymbol{x}_t^T \ \dot{\boldsymbol{x}}_t^T]^T$ and ${}^2\boldsymbol{z}_t = [\boldsymbol{x}_t^T \ \boldsymbol{k}_{j_t}^T]^T$ are the concatenation of the observed variables at each time step t. $\boldsymbol{x}_t$ and $\dot{\boldsymbol{x}}_t$ denote the robot joint angles and velocities, respectively. k_{j_t} denotes the estimated human arm stiffness during demonstration. The mean vector $\boldsymbol{\mu}_i$ and the covariance matrix $\boldsymbol{\Sigma}_i$ are defined as [35]

$$\begin{cases} {}^1\boldsymbol{\mu}_i = \begin{bmatrix} \mu_i^x \\ \mu_i^{\dot{x}} \end{bmatrix} \\ {}^1\boldsymbol{\Sigma}_i = \begin{bmatrix} \Sigma_i^{xx} & \Sigma_i^{x\dot{x}} \\ \Sigma_i^{\dot{x}x} & \Sigma_i^{\dot{x}\dot{x}} \end{bmatrix} \end{cases} \tag{15}$$

and

$$\begin{cases} {}^2\boldsymbol{\mu}_i = \begin{bmatrix} \mu_i^x \\ \mu_i^{k_j} \end{bmatrix} \\ {}^2\boldsymbol{\Sigma}_i = \begin{bmatrix} \Sigma_i^{xx} & \Sigma_i^{xk_j} \\ \Sigma_i^{k_jx} & \Sigma_i^{k_jk_j} \end{bmatrix} \end{cases} \tag{16}$$

parametrize the joint Gaussian distribution $\mathcal{P}(x, \dot{x})$ and $\mathcal{P}(x, k_j)$, respectively. This indicates that we model these two joint Gaussian distributions in parallel (see Fig. 2). These parameters Θ are learned over the demonstration dataset using the Baum-Welch algorithm [36], which is a variant of EM algorithm.

The brief distribution is first updated over the K HSMM states, based on which we compute the desired control parameters: angles, velocities, and stiffness profiles in joint space. The brief represents the probability to be in state i at time step t given the partial observation $\boldsymbol{z}_{1:t} = \{\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_t\}$ and defined as

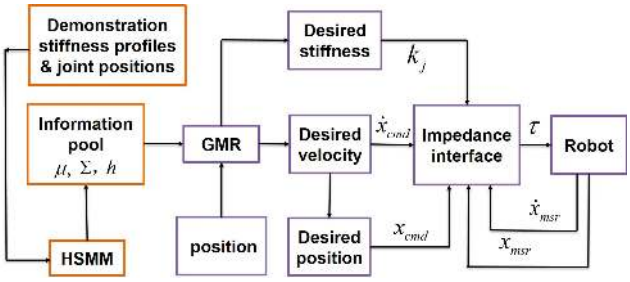


Fig. 3: The control diagram during the reproduction phase.

$$h_{i,t} = \mathcal{P}(s_t = i; z_{1:t}) = \frac{a_{i,t}}{\sum_{\kappa=1}^{K} a_{\kappa,t}} \tag{17}$$

with the forward variable $a_{i,t}$ recursively computed by

$$a_{i,t} = \sum_{j=1}^{K} \sum_{d=1}^{\min(t_{max}, t-1)} a_{j,t-d} a_{j,i} p_i^D(d) \prod_{s=t-d+1}^{t} \mathcal{N}(x_s; \mu_i^x, \Sigma_i^{xx}) \tag{18}$$

and initiation in each sate given by

$$a_{i,1} = \pi_i \mathcal{N}(\boldsymbol{x}_1; \mu_i^x, \Sigma_i^{xx}) \tag{19}$$

with $\boldsymbol{x}_1$ denoting the starting position.

The brief is a normalized version of the forward variable, and it is seen as the weight assigned to each HSMM state at time step t.

2) Task Reproduction with GMR: We compute the desired control parameters using the GMR model at each time step t. Their expectations are based on the current HSMM state weights $h_{i,t}$ given the reference position, i.e.,

$$\dot{\boldsymbol{x}}_t^* = \sum_{i=1}^{K} h_{i,t} [\mu_i^{\dot{x}} + \Sigma_i^{\dot{x}x} (\Sigma_i^{xx})^{-1} (\boldsymbol{x}_t - \mu_i^x)] \tag{20}$$

$$\boldsymbol{k}_{j_t}^* = \sum_{i=1}^{K} h_{i,t} [\mu_i^{k_j} + \Sigma_i^{k_jx} (\Sigma_i^{xx})^{-1} (\boldsymbol{x}_t - \mu_i^x)] \tag{21}$$

The desired velocities and the desired stiffness profiles are computed in (20) and (21), respectively, based on the reference positions $\boldsymbol{x}_t$, the estimated parameters of the HSMM's states, and assuming that the distributions of these variables are Gaussian. According to [4], the forward variable depends only on the observed positions [see (18)], which suggests that during reproduction the evolution of the HSMM states do not directly depend on the velocities or stiffness profiles.

It can be seen from Eqs. (15), (16), (20) and (21) that two HSMM models are used in parallel to encode position-velocity and position-stiffness rather than encoding them in one HSMM model. We would like to learn the stiffness step by step depending on the evolution of position trajectory. This is consistent with our experience, that is, we adapt arm stiffness based on position information to complete a task and we usually do not adapt it depending on the moving speed.

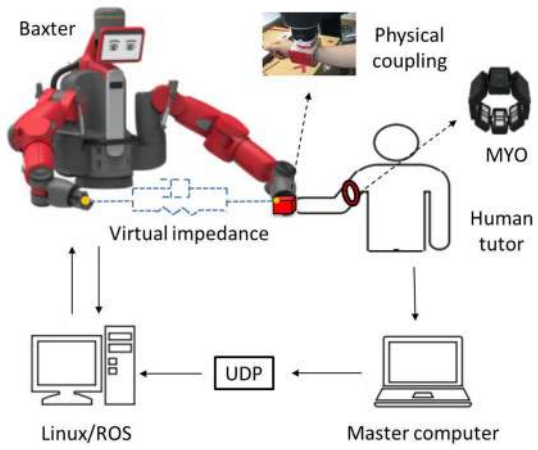


Fig. 4: The experimental setup for skill demonstration

More specifically, the adaptation of stiffness should be directly related to the position not to the velocity.

D. Impedance controller

We use an impedance controller with variable stiffness for different DOFs to control the robotic arm in joint space. In this work, the controller is designed as

$$\begin{aligned}\boldsymbol{\tau}_{cmd} =& \boldsymbol{K}_j(\boldsymbol{x}_{cmd} - \boldsymbol{x}_{msr}) + \boldsymbol{D}_j(\dot{\boldsymbol{x}}_{cmd} - \dot{\boldsymbol{x}}_{msr}) \\ &+ \boldsymbol{\tau}_{dyn}(x, \dot{x}, \ddot{x})\end{aligned} \tag{22}$$

with

$$\boldsymbol{K}_j = diag(\boldsymbol{k}_j) = \boldsymbol{J}_r^T \boldsymbol{K}_r \boldsymbol{J}_r \tag{23}$$

where $\boldsymbol{K}_j$ denotes the diagonal joint stiffness matrix with the elements of $\boldsymbol{k}_j = [k_1, k_2, \cdots, k_7]$ on the main diagonal, and $\boldsymbol{D}_j$ the corresponding damping matrix, computed based on $\boldsymbol{k}_j$ to make the controller critically damped. $\boldsymbol{J}_r$ denotes the robot arm Jacobian matrix. $\boldsymbol{x}_{cmd}$ and $\boldsymbol{x}_{msr}$ are the desired and the measured joint positions, respectively. $\dot{\boldsymbol{x}}_{cmd}$ and $\dot{\boldsymbol{x}}_{msr}$ are the desired and the measured joint velocities, respectively. $\boldsymbol{\tau}_{dyn}(x, \dot{x}, \ddot{x})$ represents the model of the arm compensating for dynamical forces, i.e., the gravity, the inertia and the Coriolis forces.

We compute the desired velocities $\dot{\boldsymbol{x}}_{cmd}$ based on the HSMM-GMR model from (20), and the desired joint stiffness matrix $\boldsymbol{K}_j$ from (21). The control diagram during the task reproduction phase is shown in Fig. 3.

III. Experimental Study

A. Experimental setup

An experimental platform based on a Baxter robot is set up for the validation of the proposed method. Fig. 4 shows the dual arm teleoperation system used for skill demonstration. The MYO is used as a sEMG detection device to collect the human tutor's upper arm sEMG signals. Then, the raw sEMG signals are collected at 200 Hz and sent to the master computer for stiffness estimation. Then, the estimated endpoint stiffness is sent through UDP to the slave computer (Linux/ROS). The generated control command is finally sent to the robot at 100 Hz. The robot has two arms, each of which has 7 degrees of freedom (DOFs). The master arm is physically connected to the human tutor's hand through a mechanical module. A virtual spring system is attached between the two arms of the robot, which enables the slave arm to follow the movement of the master arm. In this way, the tutor is able to demonstrate the skills. See [15] for the details of the experimental system.

Fig. 5: Successful task reproduction of the button-pressing task. From left to right: the starting pose, moving to the button, pressing the button and finally leaving the button.

During the demonstration, the human demonstrator guides the robot to press a button or push a box at a reachable distance from the robot arm. The robot joint state and the human arm muscle sEMG signals are simultaneously recorded for subsequent model training. The changes (e.g. drift) that may appear in sEMG sensors during demonstrations are not considered in this paper because they will not significantly affect the performance of the proposed method. Only several demonstrations are usually needed for most tasks and the long time usage of sEMG sensors will be unnecessary, and the human tutor can control the arm's moving speed and configuration in a proper range.

Two tasks are performed in this section to verify the effectiveness of the proposed method, and they are detailed as below.

B. Button-pressing task

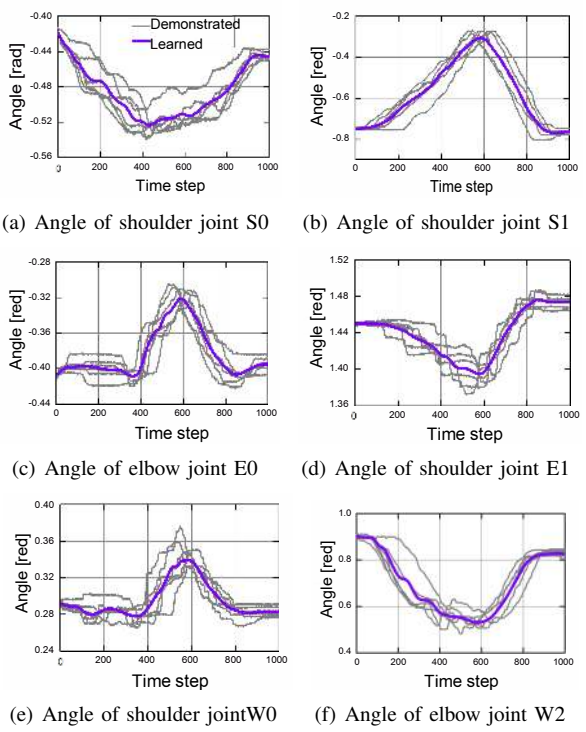
1) Setting: During the button-pressing experiment, the maximum and minimum joint stiffness of the robot arm joint are respectively set as: $K_r^{max} = [80, 80, 80, 60, 30, 20, 10]\frac{\text{Nm}}{\text{rad}}$ and $K_r^{min} = [10, 10, 10, 10, 1, 1, 0.5]\frac{\text{Nm}}{\text{rad}}$. The constant γ is chosen as 15.

For the sEMG processing, the window size W is set 40 in this work. A set of M = 6 demonstrations are obtained and then used to train the HSMM model. The number of states of the HSMM model K is manually chosen as 15 and 20 for the learning of the observed variables ${}^1\boldsymbol{z}_t = [\boldsymbol{x}_t^T \ \dot{\boldsymbol{x}}_t^T]^T$ and ${}^2\boldsymbol{z}_t = [\boldsymbol{x}_t^T \ \boldsymbol{k}_{j_t}^T]^T$, respectively.

For comparison, the following four experimental conditions were considered. The code we have utilized is mainly based on the implementation provided by S. Calinon's group[1].

Condition 1: force-free control mode. The human demonstrator taught the button-pressing task with the built-in functionality of the robot by grabbing the flange of the robot arm and moving it to approach the button. Then, the robot reproduced the task under the position control mode without involving stiffness regulation.

[1]The implementation can be found at http://calinon.ch/codes.htm

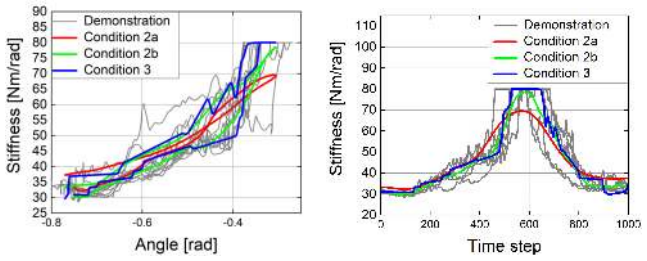


(a) Angle of shoulder joint S0 (b) Angle of shoulder joint S1

(c) Angle of elbow joint E0 (d) Angle of shoulder joint E1

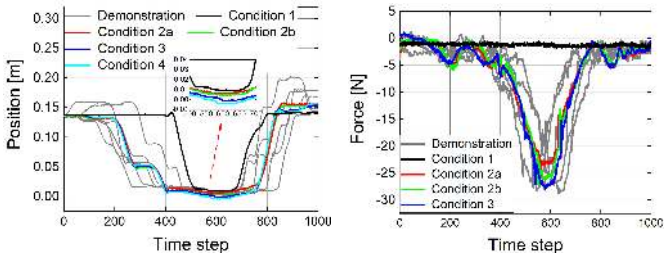
(e) Angle of shoulder jointW0 (f) Angle of elbow joint W2

Fig. 6: The demonstrated and commanded angle profiles of the six joints.



(a) (b)

Fig. 7: The demonstrated and learned stiffness profiles of joint S1 with respect to (a) joint angle and (b) time step.



(a) (b)

Fig. 8: The measured (a) position profiles and (b) force profiles of robot endpoint in z direction during task reproduction.

Condition 2a: the human demonstrated the skill using the dual arm demonstration for better collection of sEMG signals. The position and velocity control variables (i.e., ${}^{1}\boldsymbol{z}_t$) were estimated using the HSMM model as described above. The stiffness control variables, however, were learned using dynamic movement primitives (DMP) model from the demonstrated stiffness profiles. Then, the robot reproduced the task under the torque control mode with impedance adaptation.

Condition 2b: the procedure was the same with Condition 2a only with one modification: the stiffness was modelled using the GMM model instead of DMP. Under conditions 2a and 2b, position and stiffness were modelled in a separate manner, which means that the stiffness adaptation is independent of the movement information.

Condition 3: the proposed method was used in this condition. Both the observations were estimated using HSMM. Thus, the correlation between the position and the stiffness can be obtained. The robot was also controlled under the torque control mode with varying impedance.

Condition 4: in order to further test the abilities of our method, we introduced small perturbations into the experiment environment by placing the button 5 mm lower in the z axis. The experimental procedure is the same as condition 3.

2) Results and analyses: For all the conditions, task reproductions were conducted several times and no significant variance was obtained between the reproductions under each condition.

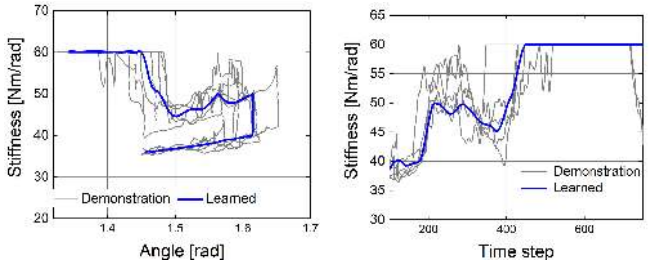
Under condition 1, the task's goal could not be achieved. This can be explained by the fact that only position control can not deal with this force-dominant task which requires stiffness regulation during the physical interaction with the environment.

For conditions 2a-4, the learned joint angle command profiles of joints S0-W1 are shown in Fig. 5. The joint W2 is fixed during task demonstration and reproduction for the convenience of mounting the tool. Reference trajectories which are estimated from the six demonstrations for these joints are needed in this work since the master arm is not used again during the task reproductions. Fig 6 shows the position commands of each joint learned from demonstrations. It shows that our method can generate decent commands although there are significant variances between the different demonstrations.

Task reproduction has also not been achieved successfully under conditions 2a and 2b. This can be explained by the fact that the stiffness can not be modeled well enough by using DMP and GMM compared with the HSMM model. Under conditions 3 and 4 the task has been successfully performed even when there exist small perturbations. An example of the successful reproduction is shown in Fig. 5 (also see the supplementary video). The Spearman correlation coefficient (SCC) between stiffness and position can be coded and increased with the proposed method (see Table I). Thus the dependence of the evolution of the stiffness on position is obtained, resulting in the better performance of representation

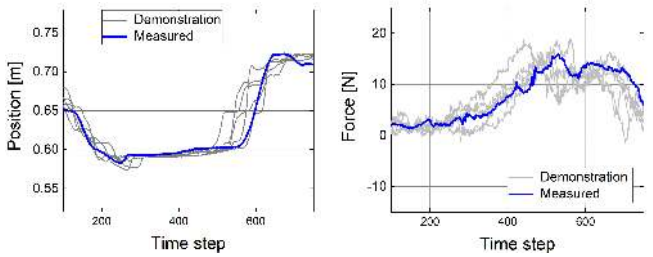
TABLE I: The Spearman correlation coefficient (SCC) between the joint angles and the corresponding stiffness profiles.

SCC	Condition 2a	Condition 2b	Condition 3
	0.9102	0.9380	0.9609



(a) (b)

Fig. 9: The demonstrated and learned stiffness profiles of joint E1 with respect to (a) joint angle and (b) time step.



(a) (b)

Fig. 10: The measured (a) position profiles and (b) force profiles of robot endpoint in y direction during task reproduction.

of the stiffness regulation features.

Take joint S1 for example. Fig. 7 shows the learned stiffness profiles with respect to the demonstrations and the time coordinate. The visual inspection of the lines in Fig. 7 suggests that the HSMM model can capture most of the features across the demonstrations. Fig. 8 shows the measured position and force profiles of the robot endpoint in z axis during task reproduction under these conditions. The position and force profiles meet the expectation of the task reproductions regarding the stiffness profiles in Fig. 7.

C. Box-pushing task

Another type of task, i.e., the box-pushing task has also been performed based on the proposed method. In this task the robot was demonstrated to push a box with a weight of 2.4 Kg placed on the surface of a table along y axis.

1) Setting: For this task, the maximum and minimum joint stiffness of the robot arm joint are respectively set as: $K_r^{max} = [100, 90, 80, 60, 30, 20, 10]\frac{\text{Nm}}{\text{rad}}$ and $K_r^{min} = [10, 10, 10, 10, 1, 1, 0.5]\frac{\text{Nm}}{\text{rad}}$.

A set of M = 5 demonstrations are obtained and then used to train the HSMM model. The number of states of the HSMM model K is manually chosen as 12 and 15 for learning the observed variables ${}^1\boldsymbol{z}_t = [\boldsymbol{x}_t^T \ \dot{\boldsymbol{x}}_t^T]^T$ and ${}^2\boldsymbol{z}_t = [\boldsymbol{x}_t^T \ \boldsymbol{k}_{j_t}^T]^T$, respectively. The parameters for sEMG processing are set the same as in the button-pressing task.

2) Results and analysis: This task has been successfully replayed using our method. It has also been performed several times and there is no obvious variance observed from these reproductions. The result shows that the stiffness profiles can be well modelled by coding the correlation between them and the position trajectories. As an example, Fig. 9 shows the learned stiffness of the joint E1 with respect to the joint angle and time. The SCC between the position and the stiffness of this joint for this task is 0.85. Fig. 10 shows the measured position and force profiles of the robot endpoint in y axis during the task reproduction, which is basically consistent with the demonstrated ones. Table II shows the RMSE values of the learned stiffness of the joint E1, the measured force and position profiles in y axis, which are computed between the demonstrations and reproductions for the ensemble of trajectories.

Note that the proposed method enables the robot to perform the button-pressing and the box-pushing tasks by modeling the stiffness instead of directly modeling the force. To model the force profiles is usually difficult and needs to equip force sensors in robotic systems. This experiment suggests that variable stiffness regulation can be used as an impedance modulating strategy for the tasks that do not require precise force control.

TABLE II: The RMS error for the box-pushing task of the profiles with respect to the demonstrations.

RMSE	Position[m]	Stiffness[Nm/rad]	Force[N]
	0.0156	3.230	2.425

D. Discussion

Learning a task by human demonstration such as pressing a button (see [4]) and pushing an object is sometimes difficult for a lightweight robot. Although the two tasks are quite easy for a human or a traditional heavy-load industrial robot, they are indeed not as easy as expected for a current collaborative robot, e.g., the Baxter robotic arm equipped with Series Elastic Actuators (SEAs) as joint actuators. Furthermore, it becomes more difficult when it comes to the learning of the impedance-based skills where both movement and stiffness/force constraints are required to be satisfied simultaneously. Our approach has the capability of addressing this issue by enabling the robot to learn the motor skills including both movement and stiffness information from the human demonstration.

It should be mentioned that there are other approaches for obtaining variable stiffness profiles. One of them is to derive a stiffness profile based on the force signals by placing a force sensor at the robotic wrist (see, e.g., [37]). Furthermore, this approach assumes that the stiffness is heavily dependent on the force and should be learned along the force trajectory. In the human motor learning, however, it has been validated that the stiffness and the feedforward force are learned separately [38, 39]. Our approach can be extended to simultaneously encode stiffness and force. Some tasks may require very delicate force and position control performances, in which cases the dynamics of force need to be well modelled and learned. One possible way to address this is to add another component to consider force information based on the proposed method.

One weakness of our approach is the accuracy of the estimated stiffness since so far it is difficult to precisely calculate the human arm stiffness based on the sEMG signals.

Although this is not a problem for most tasks, we will improve our approach for enabling the robot to learn more human-like impedance adaptability. The dynamics of the sEMG-based stiffness are often complex (see Figs. 7 and 9). The stiffness profiles should be more complex in a more complex task situation, in which case it may affect the learning performance; and it would increase computing cost with a larger number of model states and more computing time. Therefore, another direction to improve our approach is to enable effective and efficient learning of stiffness from demonstration data for complex tasks. Furthermore, in this paper the stiffness is encoded as a diagonal matrix, which may limit the flexibility of the impedance controller in a more complex manipulation task. The complete joint stiffness will be considered in future work as suggested in [16, 32].

IV. Conclusion

In this work, we proposed a programming-by-demonstration method for force-dominant tasks which enables robots to learn both movement and stiffness regulation features from humans. The hidden semi-Markov model is utilized to model the dynamics of the motion trajectories as well as the stiffness profiles, and Gaussian Mixed Regression is used to generate control commands based on the learned information of HSMM. The human stiffness is estimated directly based on the extraction of human limb muscle sEMG signals during task demonstration. This can realize a more complete skill transfer process than only considering movement demonstration. Our method integrates the bio-inspired impedance control into a robot learning system in a unified manner. The real-world experiments have verified the capacities of the method. Our future work will concentrate on the improvement of our approach as discussed above.

References

[1] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, "Robot programming by demonstration," in *Springer handbook of robotics*. Springer, 2008, pp. 1371–1394.

[2] J. Duan, Y. Ou, J. Hu, Z. Wang, S. Jin, and C. Xu, "Fast and stable learning of dynamical systems based on extreme learning machine," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, no. 99, pp. 1–11, 2017.

[3] B. Ti, Y. Gao, Q. Li, and J. Zhao, "Human intention understanding from multiple demonstrations and behavior generalization in dynamic movement primitives framework," *IEEE Access*, vol. 7, pp. 36 186–36 194, 2019.

[4] M. Racca, J. Pajarinen, A. Montebelli, and V. Kyrki, "Learning in-contact control strategies from demonstration," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2016, pp. 688–695.

[5] Z. Deng, H. Guan, R. Huang, H. Liang, L. Zhang, and J. Zhang, "Combining model-based q -learning with structural knowledge transfer for robot skill learning," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 11, no. 1, pp. 26–35, March 2019.

[6] Y. Li, G. Ganesh, N. Jarrass, S. Haddadin, A. Albu-Schaeffer, and E. Burdet, "Force, impedance, and trajectory learning for contact tooling and haptic identification," *IEEE Transactions on Robotics*, vol. 34, no. 5, pp. 1170–1182, Oct 2018.

[7] J. Duan, Y. Gan, M. Chen, and X. Dai, "Adaptive variable impedance control for dynamic contact force tracking in uncertain environment," *Robotics and Autonomous Systems*, vol. 102, pp. 54–65, 2018.

[8] W. He and Y. Dong, "Adaptive fuzzy neural network control for a constrained robot using impedance learning," *IEEE transactions on neural networks and learning systems*, vol. 29, no. 4, pp. 1174–1186, 2018.

[9] L. Kong, W. He, C. Yang, Z. Li, and C. Sun, "Adaptive fuzzy control for coordinated multiple robots with constraint using impedance learning," *IEEE transactions on cybernetics*, 2019.

[10] X. Gao, J. Ling, X. Xiao, and M. Li, "Learning force-relevant skills from human demonstration," *Complexity*, vol. 2019, 2019.

[11] H. Liu, C. Deng, A. Fernández-Caballero, and F. Sun, "Multimodal fusion for robotics," 2018.

[12] W. Zheng, H. Liu, B. Wang, and F. Sun, "Cross-modal surface material retrieval using discriminant adversarial learning," *IEEE Transactions on Industrial Informatics*, 2019.

[13] H. Liu, F. Sun, and X. Zhang, "Robotic material perception using active multi-modal fusion," *IEEE Transactions on Industrial Electronics*, pp. 1–1, 2018.

[14] A. Ajoudani, N. Tsagarakis, and A. Bicchi, "Tele-impedance: Teleoperation with impedance regulation using a body–machine interface," *The International Journal of Robotics Research*, vol. 31, no. 13, pp. 1642–1656, 2012.

[15] C. Yang, C. Zeng, P. Liang, Z. Li, R. Li, and C.-Y. Su, "Interface design of a physical human–robot interaction system for human impedance adaptive skill transfer," *IEEE Transactions on Automation Science and Engineering*, vol. 15, no. 1, pp. 329–340, 2018.

[16] C. Fang, A. Ajoudani, A. Bicchi, and N. G. Tsagarakis, "Online model based estimation of complete joint stiffness of human arm," *IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 84–91, 2018.

[17] M. Li, J. Deng, F. Zha, S. Qiu, X. Wang, and F. Chen, "Towards online estimation of human joint muscular torque with a lower limb exoskeleton robot," *Applied Sciences*, vol. 8, no. 9, p. 1610, 2018.

[18] G. Li, J. Li, Z. Ju, Y. Sun, and J. Kong, "A novel feature extraction method for machine learning based on surface electromyography from healthy brain," *Neural Computing and Applications*, vol. 31, no. 12, pp. 9013–9022, 2019.

[19] R. Wu, H. Zhang, T. Peng, L. Fu, and J. Zhao, "Variable impedance interaction and demonstration interface design based on measurement of arm muscle co-activation for demonstration learning," *Biomedical Signal Processing and Control*, vol. 51, pp. 8–18, 2019.

[20] L. Rozo, S. Calinon, D. Caldwell, P. Jiménez, and C. Torras, "Learning collaborative impedance-based robot behaviors," in *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence*, 2013, pp. 1422–1428.

[21] F. Ficuciello, L. Villani, and B. Siciliano, "Variable impedance control of redundant manipulators for intuitive human–robot physical interaction," *IEEE Transactions on Robotics*, vol. 31, no. 4, pp. 850–863, 2015.

[22] Z. Li, T. Zhao, F. Chen, Y. Hu, C.-Y. Su, and T. Fukuda, "Reinforcement learning of manipulation and grasping using dynamical movement primitives for a humanoidlike mobile manipulator," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 1, pp. 121–131, 2018.

[23] A. Ajoudani, C. Fang, N. G. Tsagarakis, and A. Bicchi, "A reduced-complexity description of arm endpoint stiffness with applications to teleimpedance control," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2015, pp. 1017–1023.

[24] Z. Li, Y. Kang, Z. Xiao, and W. Song, "Human–robot coordination control of robotic exoskeletons by skill transfers," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 6, pp. 5171–5181, 2017.

[25] F. Chen, M. Selvaggio, and D. G. Caldwell, "Dexterous grasping by manipulability selection for mobile manipulator with visual guidance," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 2, pp. 1202–1210, 2019.

[26] Y. Hu, X. Wu, P. Geng, and Z. Li, "Evolution strategies learning with variable impedance control for grasping under uncertainty," *IEEE Transactions on Industrial Electronics*, 2018.

[27] C. Yang, C. Zeng, C. Fang, W. He, and Z. Li, "A dmps-based framework for robot learning and generalization of human-like variable impedance skills," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 3, pp. 1193–1203, 2018.

[28] C. Yang, C. Zeng, Y. Cong, N. Wang, and M. Wang, "A learning framework of adaptive manipulative skills from human to robot," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 2, pp. 1153–1161, 2019.

[29] S. Schaal, J. Peters, J. Nakanishi, and A. Ijspeert, "Learning movement primitives," in *Robotics Research. The Eleventh International Symposium*. Springer, 2005, pp. 561–572.

[30] S. Calinon, F. D'Halluin, E. L. Sauser, and D. G. Caldwell, "Learning and reproduction of gestures by imitation," *Robotics and Automation Magazine IEEE*, vol. 17, no. 2, pp. 44–54, 2010.

[31] S. Calinon, A. Pistillo, and D. G. Caldwell, "Encoding the time and space constraints of a task in explicit-duration hidden Markov model," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sept 2011, pp. 3413–3418.

[32] A. Ajoudani, C. Fang, N. Tsagarakis, and A. Bicchi, "Reduced-complexity representation of the human arm active endpoint stiffness for supervisory control of remote manipulation," *The International Journal of Robotics Research*, vol. 37, no. 1, pp. 155–167, 2018.

[33] L. Peternel, N. Tsagarakis, and A. Ajoudani, "A human-robot co-manipulation approach based on human sensorimotor information," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2017.

[34] M. Ison and P. Artemiadis, "The role of muscle synergies in myoelectric control: trends and challenges for simultaneous multifunction control," *Journal of neural engineering*, vol. 11, no. 5, p. 051001, 2014.

[35] C. Zeng, C. Yang, J. Zhong, and J. Zhang, "Encoding multiple sensor data for robotic learning skills from multimodal demonstration," *IEEE Access*, vol. 7, pp. 145 604–145 613, 2019.

[36] S.-Z. Yu and H. Kobayashi, "Practical implementation of an efficient forward-backward algorithm for an explicit-duration hidden Markov model," *IEEE Transactions on Signal Processing*, vol. 54, no. 5, pp. 1947–1951, 2006.

[37] J. Duan, Y. Ou, S. Xu, Z. Wang, A. Peng, X. Wu, and W. Feng, "Learning compliant manipulation tasks from force demonstrations," in *2018 IEEE International Conference on Cyborg and Bionic Systems (CBS)*. IEEE, 2018, pp. 449–454.

[38] E. Burdet, R. Osu, D. W. Franklin, T. E. Milner, and M. Kawato, "The central nervous system stabilizes unstable dynamics by learning optimal impedance," *Nature*, vol. 414, no. 6862, pp. 446–449, 2001.

[39] E. Burdet, G. Ganesh, C. Yang, and A. Albu-Schäffer, "Interaction force, impedance and trajectory adaptation: by humans, for robots," in *Experimental Robotics*. Springer, 2014, pp. 331–345.

Chao Zeng (S'18) received the M.S. degree in precision instrument and mechanics from Shanghai University, Shanghai, China, in 2016, and the Ph.D. degree in pattern recognition and intelligent system at the School of Automation Science and Engineering, South China University of Technology, Guangzhou, China, in 2019. He visited TAMS group, University of Hamburg, Germany, from October 2018 to October 2019. His current research interests include human robot interaction, programming by demonstration and human robot skill transfer.

Chenguang Yang (M'10-SM'16) received the B.Eng. degree in measurement and control from Northwestern Polytechnical University, Xi'an, China, in 2005, and the Ph.D. degree in control engineering from the National University of Singapore, Singapore, in 2010. He received postdoctoral training at Imperial College London, UK. He is the recipient of the Best Paper Award from the IEEE Transactions on Robotics and a number of international conferences. His research interests lie in robotics and automation.

Hong Cheng (M'06-SM'14) received the Ph.D. degree in pattern recognition and intelligent systems from Xian Jiaotong University, in 2003. He was a Visiting Scholar with the School of Computer Science, Carnegie Mellon University, USA, from 2006 to 2009. He has been an Executive Director of the Center for Robotics, since 2014, has been with UESTC, since 2010, and has been an Associate Professor with Xian Jiaotong University, since 2005. He is currently a Full Professor with the School of Automation and Engineering, University of Electronic Science and Technology of China (UESTC). He has over 100 academic publications, including two books: Digital Signal Processing (Tsinghua University Press, 2007) and Autonomous Intelligent Vehicles: Theory, Algorithms and Implementation (Springer, 2011). His current research interest includes machine learning in humanrobot hybrid systems. He served/is serving as a General Chair of VALSE 2015, a Program Chair of CCPR 2016, and a General Chair for CCSR 2016.

Yanan Li (S'10-M'14) received the BEng and MEng degrees from the Harbin Institute of Technology, China, in 2006 and 2008, respectively, and the PhD degree from the National University of Singapore, in 2013. Currently he is a Lecturer in Control Engineering with the Department of Engineering and Design, University of Sussex, UK. From 2015 to 2017, he has been a Research Associate with the Department of Bioengineering, Imperial College London, UK. From 2013 to 2015, he has been a Research Scientist with the Institute for Infocomm Research (I2R), Agency for Science, Technology and Research (A∗STAR), Singapore. His general research interests include human-robot interaction, robot control and control theory and applications.

Shi-Lu Dai (S'09-M'11) received his B.Eng. degree in thermal engineering in 2002, and his M.Eng. and Ph.D. degrees in control science and engineering in 2006 and 2010, respectively, from the Northeastern University, Shenyang, China.

He was a Visiting Student in the Department of Electrical and Computer Engineering, National University of Singapore, Singapore, from November 2007 to November 2009, and a Visiting Scholar at the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN, USA, from October 2015 to October 2016. Since 2010, he has been with the School of Automation Science and Engineering, South China University of Technology, Guangzhou, China, where he is currently a Professor. His current research interests include adaptive and learning control, distributed cooperative systems.