An approach for robotic leaning inspired by biomimetic adaptive control

Article (Accepted Version)

An Approach for Robotic Leaning Inspired by Biomimetic Adaptive Control

Chao Zeng, *Member, IEEE*, Hang Su, *Member, IEEE*, Yanan Li, *Senior Member, IEEE*, Jing Guo, and Chenguang Yang, *Senior Member, IEEE*

***Abstract*—How to enable robotic compliant manipulation has become a critical problem in the robotics field. Inspired by a biomimetic adaptive control strategy, this work presents a novel representation model named human-like compliant movement primitives (Hl-CMPs) which could allow a robot to learn human-like compliant behaviours. The state-of-the-art approaches can hardly learn complete compliant profiles for a specific task. Comparatively, our model can encode task-specific parametric movement trajectories, correspondingly associated with dynamic trajectories including both impedance and feedforward force profiles. The compliant profiles are learned based on a biomimetic control strategy derived from the human motor learning in the muscle space, enabling the robot to simultaneously learn the impedance and the force while executing the movement trajectories obtained from human demonstration. Furthermore, both the kinematic and the dynamic profiles are learned in the parametric space, thus enabling the representation of a skill using corresponding parameters (i.e, task-specific parameters). Hl-CMps can allow the robot to automatically learn compliant behaviours in an online manner after kinematic demonstration. Our approach is validated by an insertion task and a cutting task based on a KUKA LBR iiwa robot.**

***Index Terms*—Adaptive impedance/force control; Robot control; Impedance learning; Human-robot interaction; Robotics**

I. Introduction

Humans have an amazing ability of compliantly performing tasks, although we have not completely understood the deep reasons behind this yet. It is, therefore, natural for us to ask two questions: *can we enable a robot to learn such an ability from humans, and how?* For the first question the answer is arguably yes. Actually, it has become an agreement that learning from demonstration can be an efficient way to transfer human behaviours to robots [1–4], and is being widely utilized for robot programming. For the second one, however, it still remains quite challenging, especially when it comes to compliant manipulation skills [5]. This work aims to take one step forward toward this goal.

This work was in part supported by National Nature Science Foundation (NSFC) under Grants 61861136009 and 61811530281, the Fellowship of the China Postdoctoral Science Foundation under Grant 2020M682613, the Opening Project of Shanghai Robot R&D and Transformation Functional Platform, and the Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/S001913. Corresponding author: C. Yang. Email: cyang@ieee.org

C. Zeng and Jing Guo are with School of Automation, Guangdong University of Technology, Guangzhou 510641, China. Y. Li is with the Department of Engineering and Design, University of Sussex, Brighton, BN1 9RH, U.K. H. Su is with the Department of Electronics, Information and Bioengineering, Politecnico di Milano, 20133 Milan, Italy. C. Yang is with Bristol Robotics Laboratory, University of the West of England, Bristol, BS16 1QY, UK.

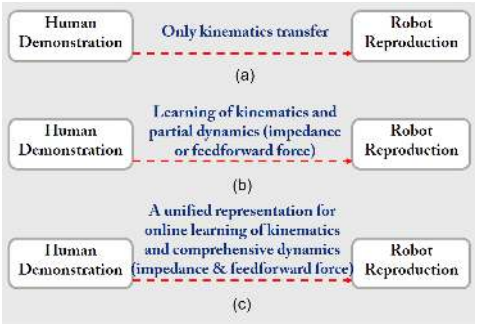


Figure 1: General description of state-of-the-art approaches which are focused on (a) kinematics transfer and (b) learning of partial dynamics profiles (impedance or feedforward force), and (c) the proposed approach. Our work seeks to provide a unified representation for online learning of kinematics and dynamics profiles. After the kinesthetic demonstration of a task, our approach allows a robot to compliantly reproduce the task by automatically learning the dynamics profiles during the execution of the movement trajectories.

Although many efforts have been recently made to allow robots to perform tasks in a compliant manner based on adaptive control strategies [6–10], so far few works have ever considered to integrate the human arm adaptive control strategy in the muscle space into a robotic learning system so that the robots can behave in a human-like way to deal with *force-relevant* tasks [11]. More specifically, most of the previous studies in this domain focused on the learning of robotic impedance while leaving alone the force, or the opposite way. As a matter of fact, it has already been made clear that in a dynamic task we humans can simultaneously learn the arm impedance and the feedforward force to meet the task requirements [12]. This work, instead, will integrate this principle found in human motor learning into the robotic skill learning model.

A general description of the state-of-the-art and the proposed approaches is shown in Fig. 1. In this paper, we propose a unified approach for the parametric representation (i.e., Hl-CMPs) of a robotic compliant skill including not only the kinematic movement trajectories, but also the impedance and force profiles, in order to facilitate the learning of human-like behaviours. The impedance and the feedforward force are adapted in the parameter space instead of directly learned in the trajectory level, which is different from previous works. Our approach includes a two-step procedure: the first step is to

obtain the task-specific movement trajectories by demonstration, for which there are several ways that can be utilized here including teleoperation and kinesthetic demonstration through physical interaction. The parameters for these demonstration movement trajectories are then learned through supervised learning. During the second step, the robot can reproduce the demonstrated task in a compliant manner since the robot can automatically learn the compliant behaviour along the execution of the reference movement trajectories which are encoded by the dynamical movement primitives (DMPs) in advance, thanks to the proposed Hl-CMPs built on the biomimetic control.

The contributions of our work are summarized as: i) a novel robotic compliant skill learning approach that integrates the DMP-based motion trajectory encoding model and the biomimetic adaptive control strategy is first established; and ii) The adaptation of the compliant profiles including impedance, damping and force is extended from the previous work in trajectory-space to parametric space in this work. The contributions will be comparatively explained in detail in the following section.

II. Related work

A. Learning of robotic complaint movements

Robotic compliant movement profiles mainly include kinematic trajectories and dynamic (i.e., impedance and force/torque) trajectories [13]. During the last two decades a number of models have been developed to address the encoding of kinematic trajectories. Among these models, DMPs has been widely used in a number of imitation learning tasks. Similar as in [14], in this work we take DMPs as the representation constituting the kinematic part of the Hl-CMPs, thanks to its many advantageous properties such as adaptation and generalization.

In order to achieve compliant manipulation, variable impedance based torque control has been recently proposed and validated for robotic task execution. Variable impedance control is a very wide topic in this domain, but the core question in robotic learning is *how to learn proper impedance profiles for a specific task.*

One solution to this question is to enable the robot to learn impedance from humans. To this end, the human user's limb impedance is first estimated by processing Electromyography (EMG) signals extracted from the human user during demonstration [15], and then transferred to the robotic arm through the impedance mapping between the human and the robot. Human-robot impedance transfer has successfully enabled the robots to learn/generalize several compliant skills such as co-carrying [16] and pushing [17]. However, one drawback of this approach lies in its requirement of an off-line complex parametric identification process and the need for force and EMG sensors for the estimation of the human arm impedance.

Another approach is based on reinforcement learning (RL). Policy improvement with path integrals (PI^2) is one of the RL-based algorithms that can be utilized to learn robotic variable impedance profiles [18, 19]. It has been applied to several robotic tasks such as opening a door [18] and robust grasping

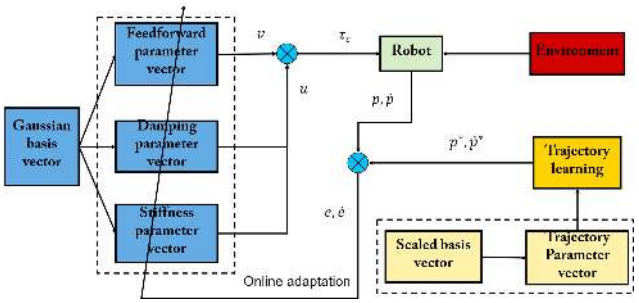


Figure 2: The diagram of the proposed biomimetic control strategy which is adapted from the previous work [24]. The yellow blocks represent the trajectory learning part completed before the reproduction of the task; the blue blocks represent the online adaptation of the feedforward force and the impedance in order to deal with the interaction force from the environment.

[20]. RL-based approaches optimize the impedance profiles, often initialized with constant stiffness/damping trajectories, based on predefined cost objective functions. A number of trials may be required until the cost objective functions are eventually minimized/maximized.

Variable impedance profiles can also be estimated by extracting the mapping between the sensed force and the corresponding position variation, which has been used in some task scenarios [5, 21]. A force sensor is required and mounted onto the endpoint of the robotic arm to collect the force signals for stiffness estimation, thus inevitably increasing the total cost of the robotic system. Furthermore, the estimated stiffness strongly depends on the sensed force signals.

In [22], the stiffness profiles are adapted along with the execution of the known movement trajectories based on perturbations imposed by a human teacher through physical interaction, which means human inference is required during the learning process. A linear quadratic tracking approach is used in [23] to obtain stiffness, damping and feedforward force profiles depending on the precomputed model parameters. Very recently, compliant movement primitives (CMPs) is proposed for the learning of robotic compliant behaviours. Different from the approaches mentioned above, CMPs aims to acquire not the variable impedance profiles but the feedforward dynamic torque trajectories, referred as task-specific torque primitives (TPs) [14]. TPs is obtained through the execution of the desired movement trajectories by the robot under a high-gain control mode which may be harmful to the robotic platform. CMPs is further improved in [13] to avoid this problem through autonomous learning of the TPs profiles.

Unlike the work mentioned above, this paper aims to enable the learning of both impedance and feedforward torque profiles simultaneously and automatically during the robot reproduction of a task.

B. Biomimetic control based on human motor learning

Biomimetic control for robotic compliant manipulation can be inspired by human motor learning and control. With the aid of the progress made in neuroscience study (e.g., [12]),

actually, several biomimetic controllers have been developed for the transfer of control strategies in human muscle space to a robotic arm in joint/Cartesian space, in order to learn human-like compliant behaviours. Typically, [25] derived a biomimetic controller from human motor control by constructing a cost function with *muscle activation parameters*, based on which using gradient descent the impedance and the force profiles can be learned concurrently. This controller is further improved and integrated into motion planning for compliant interaction, and this approach was evaluated in a simulated two planar robot collaborative lifting task.

Another biomimetic robotic controller was developed in [26] that can allow the robot to deal with not only stable tasks but also unstable tasks during the interaction with its environment or a human user. Recently, [24] improved the controller for the learning of dynamic profiles, and beyond that movement adaptation was also considered in the controller. Their approach has been validated by several tasks such as drilling and haptic identification [24]. In our work, the adaptation of movement trajectory is neglected because it can be directly learned from human demonstration. In order to facilitate robotic learning of human-like skills in a unified representation, we do not directly adapt the dynamic profiles (impedance, damping and force) but update the task parameters representing these dynamic profiles, namely, they are learned in the parametric space.

To summarize, we seek to provide a compact and unified representation for learning all the dynamics profiles including impedance and feedforward force. Inspired by the concept CMPs proposed in [13, 14], here we name our approach human-like compliant movement primitives (Hl-CMPs) indicating that our approach is derived from the human motor learning strategy.

III. Preliminary

A. Human motor learning

This subsection briefly presents some neuroscience findings in human motor learning.

In [12], it has been already shown that humans have the ability to adapt both impedance and feedforward force in unknown dynamical environments under the control of the central nervous system (CNS), which is a step forward from the previous finding that only the force is adapted to deal with stable tasks.

More importantly, it is observed that the impedance $u_{mus.}$ and the force $v_{mus.}$ are learned in an independent way, and the sum of the two terms forms the motor command $\tau_{mus.}$ in the human muscle space, i.e.,

$$\tau_{mus.} = u_{mus.} + v_{mus.} \tag{1}$$

Furthermore, it has also demonstrated that the feedforward command, i.e. $v_{mus.}$ and the feedback command, i.e. $u_{mus.}$ are concurrently learned to guarantee the minimization of both movement error and effort during a specific task. Based on these findings in human motor learning, several biomimetic controllers have been developed recently for the robotic compliant control, e.g, [25]. We would like not to explain too much about human motor learning in this work, so please refer to [24, 26] for the details.

B. Robotic control

The dynamics of a rigid-body robot in the task space can be given by

$$M(q)\ddot{x} + C(q,\dot{q})\dot{x} + G(q) = f_c + f_{ext} \tag{2}$$

where q is the joint angle, and $\dot{q}$ and $\ddot{q}$ are corresponding joint velocity and acceleration, respectively. x represents the position of the robotic arm endpoint. $M(q)$, $C(q,\dot{q})$ and $G(q)$ represent inertia, the Coriolis and centrifugal forces, and gravitational force, respectively. The sum of these three forces is considered as f_{dyn} which is assumed to be known in this paper. f_c is the control input, and f_{ext} represents the external force exerted on the robot by its environment or a human user.

According to the human muscle motor learning principles, the control input can be split into two parts

$$f_c = -u - v \tag{3}$$

where u and v represent impedance and feedforward force, respectively.

And impedance is usually given by

$$u = K_s e + K_d \dot{e} \tag{4}$$

with

$$\begin{cases} e = x - x_d \\ \dot{e} = \dot{x} - \dot{x}_d \end{cases} \tag{5}$$

where K_s and K_d are stiffness and damping, respectively. e and $\dot{e}$ are position and velocity errors, respectively. x_d and $\dot{x}_d$ are desired position and desired velocity, respectively.

IV. Methodology

First, DMPs model is briefly introduced. We then present how to reform DMPs as Hl-CMPs. Finally, we will elaborate how to learn Hl-CMPs based on biomimetic motor control.

A. Recall of DMPs

DMPs has been widely used to encode robotic motion trajectories in joint or Cartesian space. Generally, we can encode a 1-DOF (degree of freedom) point-to-point trajectory by the following equations.

$$\tau\dot{z} = \alpha(\beta(p_{goal} - p) - z) + f(s) \tag{6}$$

$$\tau\dot{p} = z \tag{7}$$

$$\tau\dot{s} = -\alpha_s s \tag{8}$$

$$f(s) = \frac{\Sigma_{n=1}^{N}\omega_n(s)\theta_n}{\Sigma_{n=1}^{N}\omega_n(s)} s(p_{goal} - p_0) \tag{9}$$

with

$$\omega_n(s) = \exp(-0.5h_n(s - c_n)^2) \tag{10}$$

where p is motion position and z is the corresponding velocity. p_{goal} is the goal of the motion trajectory. α and β are constant parameters which are set properly in advance before learning of the model. τ is a temporal scaling coefficient that can

Table I: The learning procedure of Hl-CMPs

I Demonstration
A human user demonstrates the robot to perform a task, during which the robotic motion states are recorded.
II Learning of motion trajectory parameter θ_p
Fit the DMP model based on the recorded data using LWR
III Learning of stiffness, damping, force parameters θ_k, θ_d, θ_v
Initialize these parameters, $\theta_k(t_0)$, $\theta_d(t_0)$ and $\theta_v(t_0)$
Set constant matrices/coefficients, e.g., Q_k, Q_d and Q_v
For each time step t, **do**
Sense robot current states and compute sliding error (Eq. 29)
Compute the increment rates $\dot{\theta}_k(t)$, $\dot{\theta}_d(t)$, $\dot{\theta}_v(t)$ (Eq. 31)
Update these parameters, i.e., $\theta_k(t)$, $\theta_d(t)$, $\theta_v(t)$
Update the impedance and feedforward force
Calculate the desired joint torque commands using Eq. 3
Send the generated commands to the robotic actuators
End For

modify the execution duration of the motion trajectory. s is phase variable determined by Eq. 8 with a constant coefficient α_s. The non-linear force term $f(s)$ can modify the shape of the motion trajectory, and it is defined by Eq. 9 that can be seen as a linear combination of a set of basis functions. Usually, radial basis function (Eq. 10) is chosen to form the force term with a width h_n and a center c_n. N is the number of radial basis functions.

The system can be guaranteed to converge to the goal by choosing proper constants, i.e., α, β, α_s and τ. For details about the stability and other characteristics of this model we suggest to refer to [27].

B. Representation of Hl-CMPs

First, for the encoding of motion data we rewrite DMPs in the following equations

$$\tau\dot{z} = \alpha(\beta(p_{goal} - p) - z) + \theta_p \cdot g_p \tag{11}$$

with motion parameters θ_p. $(\cdot)$ denotes dot product operation.

$$\theta_p = [\theta_{p,1}, \theta_{p,2}, \cdots, \theta_{p,n}, \cdots, \theta_{p,N}]$$

and the basis functions g_p.

$$g_p = [g_{p,1}, g_{p,2}, \cdots, g_{p,n}, \cdots, g_{p,N}]^T$$

The n-th element in g_p is determined by

$$[g_p]_n = \frac{\omega_n(s)s}{\sum_{n=1}^{N}\omega_n(s)}(p_{goal} - p_0) \tag{12}$$

Then, we use the following equations for the encoding of compliant movement profiles.

For the m-th DOF, the stiffness:

$$K_{s,m} = \theta_{k,m} \cdot g \tag{13}$$

with

$$\theta_{k,m} = [\theta_{k,1}, \theta_{k,2}, \cdots, \theta_{k,n}, \cdots, \theta_{k,N}]$$

The damping term:

$$K_{d,m} = \theta_{d,m} \cdot g \tag{14}$$

with

$$\theta_{d,m} = [\theta_{d,1}, \theta_{d,2}, \cdots, \theta_{d,n}, \cdots, \theta_{d,N}]$$

and the feedforward force:

$$v_m = \theta_{v,m} \cdot g \tag{15}$$

with

$$\theta_{v,m} = [\theta_{v,1}, \theta_{v,2}, \cdots, \theta_{v,n}, \cdots, \theta_{v,N}]$$

They share the same basis vector g that is determined by

$$[g]_n = \frac{\omega_n(s)}{\sum_{n=1}^{N}\omega_n(s)} \tag{16}$$

Thus, a robotic compliant skill can be represented by specific parameters as below

$$\theta = \{\theta_{p,i}, \theta_{k,i}, \theta_{d,i}, \theta_{v,i}\}_{i=1}^{M} \tag{17}$$

where M is the number of DOFs.

C. Learning of Hl-CMPs

The learning of Hl-CMPs parameters θ is divided into two parts. The first part is to learn the motion trajectories, i.e., the motion parameters θ_p. In the second part, then, the other parameters are learned along the execution of the learned motion trajectories.

1) Learning of motion parameters: Eq. 7 is first substituted into Eq. 11 to yield a second-order system

$$\theta_{p,m} \cdot g_p = \tau^2\ddot{p} + \alpha\tau\dot{p} - \alpha\beta(p_{goal} - p) \tag{18}$$

Given the m-th DOF reference trajectory position $p_{m,t}$, and its corresponding first-order and second-order derivatives $\dot{p}_{m,t}$ and $\ddot{p}_{m,t}$, the target force function is then derived as

$$f_{target,i}(s_t) = \tau^2\ddot{p}_{m,t} + \alpha\tau\dot{p}_{m,t} - \alpha\beta(p_{goal} - p_{m,t}) \tag{19}$$

with

$$p_{goal} = p_{m,T}, \quad t = [1, 2, \ldots, T] \tag{20}$$

Then, θ_p can be obtained by

$$\min \|f_{target}(s) - \theta_p \cdot g_p\|^2 \tag{21}$$

which is considered as a supervised learning problem which can be efficiently solved by using regression algorithms, e.g., Locally Weighted Regression (LWR) and Gaussian Process Regression (GPR).

2) Learning of parameters of compliant profiles: In the following we will derive the learning law of θ_k, θ_d and θ_v based on a recently developed biomimetic controller [24, 26].

First, $\theta_k^*(t)$, $\theta_d^*(t)$ and $\theta_v^*(t)$ are assumed to be the desired parameters of the corresponding desired stiffness, damping and feedforward force that can maintain stability, and they are represented as follows

$$\Phi^*(t) = [\bar{\theta}_k^{*T}, \bar{\theta}_d^{*T}, \bar{\theta}_v^{*T}]^T \tag{22}$$

where $\bar{\theta}_k^*$, $\bar{\theta}_d^*$ and $\bar{\theta}_v^*$ are the row average vectors of θ_k^*, θ_d^* and θ_v^*, respectively.

According to the human motor learning, we can assume that CNS could automatically adapt the stiffness, damping and feedforward force to approach the desired ones by minimizing the cost function

$$J_c = \frac{1}{2}\tilde{\Phi}^T Q^{-1}\tilde{\Phi} \tag{23}$$

Table II: Dimension of the main parameters.

Parameter description	Notations	Dimension
position and velocity	x, $\dot{x}$, x_d, $\dot{x}_d$, e, $\dot{e}$, ε	M × 1
force and torque	f_c, f_{ext}, u, v	M × 1
stiffness and damping	K_s, K_d	M × M
basis function vector	g_p, g	N × 1
Hl-CMP model para.	$\theta_p, \theta_k, \theta_d, \theta_v$	M × N
cost functions	J_{all}, J_c, J_e	1 × 1
learning rates.	Q_k, Q_d, Q_v	M × M

where

$$\tilde{\Phi} = \Phi - \Phi^* = [\tilde{\theta}_k^T, \tilde{\theta}_d^T, \tilde{\theta}_v^T]^T \tag{24}$$

with

$$\Phi = [\bar{\theta}_k^T, \bar{\theta}_d^T, \bar{\theta}_v^T]^T \tag{25}$$

and

$$\tilde{\theta}_k = \bar{\theta}_k - \bar{\theta}_k^*, \tilde{\theta}_d = \bar{\theta}_d - \bar{\theta}_d^*, \tilde{\theta}_v = \bar{\theta}_v - \bar{\theta}_v^* \tag{26}$$

where $\bar{\theta}_k$, $\bar{\theta}_d$ and $\bar{\theta}_v$ are the row average vectors of θ_k, θ_d and θ_v, respectively.

The constant matrix Q is defined by

$$Q = \text{diag}(Q_k \otimes \text{I}, Q_d \otimes \text{I}, Q_v \otimes \text{I}) \tag{27}$$

where Q_k, Q_d and Q_v are symmetric positive-definite matrices, corresponding to the stiffness, damping and the feedforward force, respectively, i.e., $Q_k = diag\{Q_{k,m}\}$, $Q_v = diag\{Q_{v,m}\}$, $Q_v = diag\{Q_{v,m}\}$, $m \in [0, \cdots, M]$. And I is the identity matrix.

For minimization of the motion tracking error, the following cost function can be considered

$$J_e = \frac{1}{2}\varepsilon^T M(q)\varepsilon \tag{28}$$

where the sliding error

$$\varepsilon = \dot{e} + \pi e \tag{29}$$

with a positive constant coefficient π.

Therefore, the overall cost function is the sum of the two above-mentioned cost functions

$$J_{all} = J_c + J_e \tag{30}$$

The minimization of the cost function can be achieved through the following learning laws. They are adapted at each time step for the $m-$th ($m \in [0, \cdots, M]$) DOF as:

$$\begin{aligned} \dot{\theta}_{k,m}^T &= Q_{k,m}\varepsilon_m e_m g \\ \dot{\theta}_{d,m}^T &= Q_{d,m}\varepsilon_m \dot{e}_m g \\ \dot{\theta}_{v,m}^T &= Q_{v,m}\varepsilon_m g \end{aligned} \tag{31}$$

A brief proof for Eq. 31 can be found in the Appendix. Until now, all the task parameters in Eq. 17 can be obtained. The diagram of the biomimetic control strategy is shown in Fig. 2. The whole learning procedure of Hl-CMPs is summarized in Table I. And Table II summarizes the size of the main parameters used for the updating law.

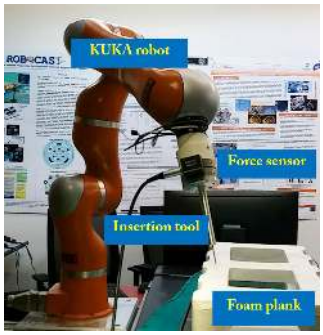


Figure 3: The experimental setup for the insertion task.

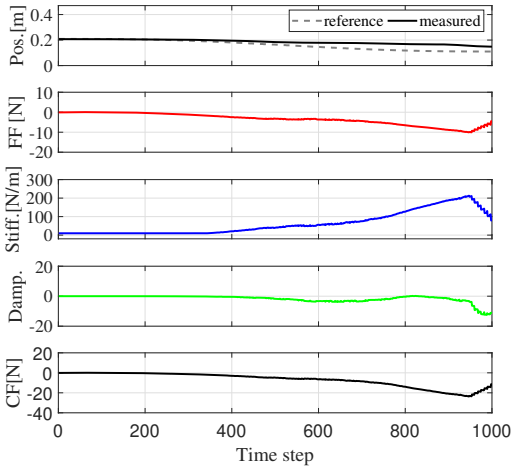


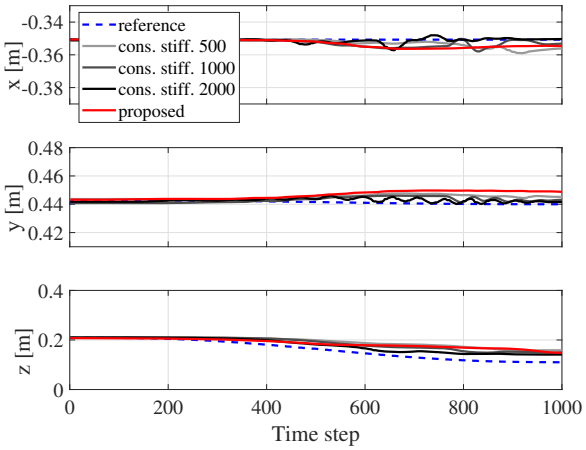
Figure 4: The online learning results of the compliant profiles while executing the reference movement trajectory in z axis. FF represents the feedforward force, and CF represents the commanded force.

D. Insertion task

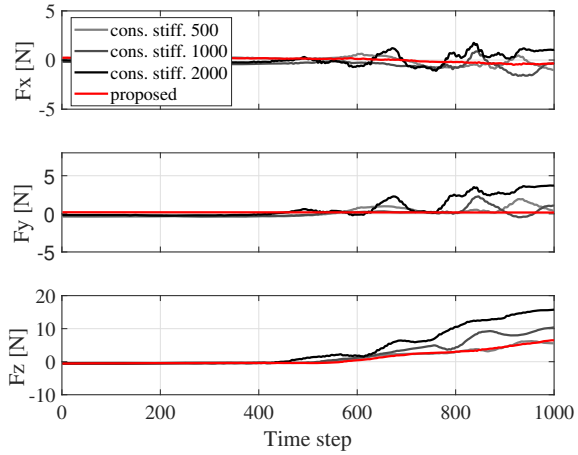
V. Experimental Study

In this section we show that our approach is validated by an insertion task. The experimental set-up is shown in Fig. 3. The KUKA LWR4+ robot is controlled in the task space under the force control mode. A force sensor (M8128C6, SRI) is mounted between the robotic endpoint and the tool that is used to collect the interaction force signals. Two tasks, i.e., insertion and cutting, are conducted to verify the performance of the proposed method.

In the first task, the robot is controlled to insert the tool into the foam plank during demonstration. The main setting of the parameters for the Hl-CMPs is chosen by heuristics: $Q_K = diag\{3000, 3000, 3000\}$, $Q_D = diag\{90, 90, 90\}$, $Q_v = diag\{9, 9, 9\}$. The parameters θ_k, θ_d and θ_v are initialized with zero. An example of the online learning of the compliant profiles (i.e., stiffness, damping and feedforward force) is shown in Fig. 4. It shows the adaptation of these profiles along the movement trajectory in z axis. During the reaching phase (from the very beginning to about 400th time steps), the feedforward force and the stiffness stay at comparatively lower values, thus allowing the robot to compliantly contact with the foam plank. Subsequently, they slowly



(a)



(b)

Figure 5: The measured endpoint (a) position trajectories and (b) force profiles in x, y and z axes during robot reproduction under the different control conditions.

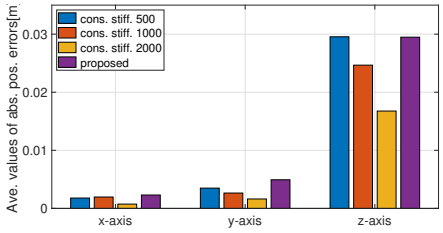


Figure 6: The average absolute values of the position errors between the reference and measured trajectories.

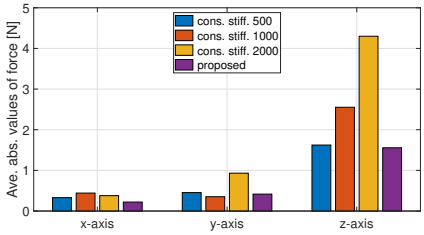


Figure 7: The average absolute values of the force profiles.

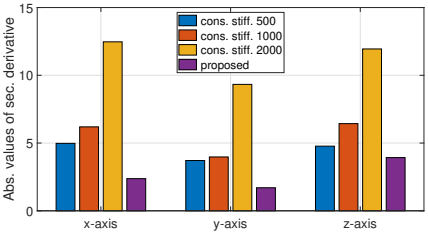


Figure 8: The sum of the absolute values of the second derivative of the force profiles.

increase during the insertion phase, resulting in the increasing of the command force, in order to deal with the external resistance applied onto the tool. After the tool piercing the foam plank, during the third phase the force and the impedance start to decrease since the external resistance becomes smaller. It is interesting to find that in our experiment the feedforward force and the command force almost share the same shape, and more interestingly, the value of the feedforward force at each time step is almost half of that of the command force, i.e., $u \approx v \approx \frac{1}{2}\tau_c$. Of course, this might not be always true for other tasks since the dynamics vary largely in different situations. For instance, different inserting objects would result in different task dynamics [24].

Conventionally, the robot is often controlled in a constant impedance mode to perform a specific task. For comparison, the robot is also enabled under this mode to perform the insertion task, with different constant stiffness values (i.e., 500, 1000 and 2000[N/m]) for each axis. The measured endpoint position trajectories and force profiles during robot reproduction are shown in Fig. 5. To quantitatively compare the performances, we visualize i) the average absolute values of the position errors between the reference and measured trajectories in Fig. 6; ii) average absolute values ($\frac{1}{N}\sum_i^N \|f_i\|$) of the force profiles in Fig. 7; and iii) the sum of the absolute values of the second derivative ($\sum_i^N \|\ddot{f}_i\|$) of the force profiles in Fig. 8, which can approximately illustrate the smoothness of force profiles. We can see that larger stiffness would result in a better performance of position tracking. However, it is also observed that the constant impedance control would easily cause unstable interactions during the insertion process even using a comparatively low stiffness value (e.g., 500 [N/m]). Our approach, on the other hand, can allow the robot to learn compliant behaviours resulting in more smooth force profiles which are slowly and stably adapted during the interaction between the robot and the environment. Furthermore, the force profiles are almost kept zero in x and y axes, which means the proposed Hl-CMPs model can automatically decouple the dynamics profiles in different axes and then selectively adapt the task-related ones, but the fixed impedance mode could not do so.

The direct visual results are shown in Fig. 9. We can see the surface of the foam plank tends to be more easily damaged by the tool under the constant impedance mode than using the Hl-CMPs, due to the lack of the adaptability.

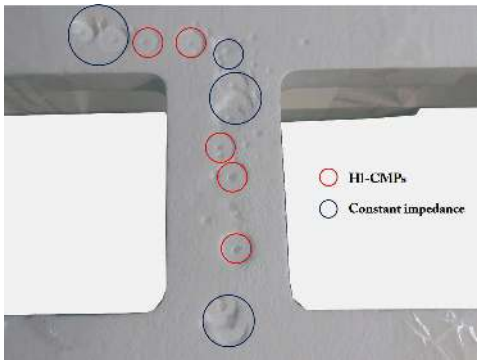
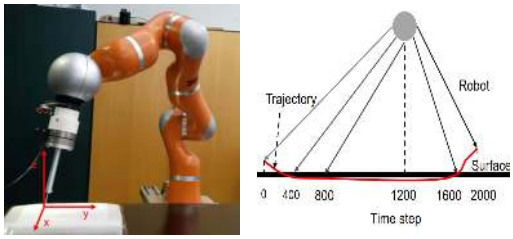


Figure 9: The visual results of the reproduction of the insertion task, using the proposed HI-CMPs and constant impedance.



(a) (b)

Figure 10: (a) The experimental setup for the cutting task, and (b) a cartoon illustration of the cutting process.

A. *Cutting task*

The second task, i.e., a cutting task is further performed to verify the generalizability of our approach in different task situations. The task setting is sown in Fig. 10(a), and a cartoon illustration of the cutting process is sown in Fig. 10(b). During this task the robot is first demonstrated by a human user to lightly cut the surface of an object, following three steps: reaching and contacting the object from the start point; cutting it in $x-y$ plane; and finally lifting the robot arm. Then, the robot is controlled to reproduce the process with the proposed HI-CMPs algorithm, and with constant stiffness values 500 and 1000[N/m] for each DOF.

The parameter settings for the HI-CMPs in this task are: $Q_K = diag\{4000, 4000, 4000\}$, $Q_D = diag\{50, 50, 50\}$, $Q_v = diag\{9, 9, 9\}$. The learning results are shown in Fig 11. The three subplots show all the learned profiles including the position trajectories, feedforward force, stiffness and damping profiles, as well as the commanded forces in the Cartesian space along x, y and z axes, respectively. From the curves, we can see that during the first step, i.e., the free-motion reaching stage, the impedance and feedforwrd force do not change much in x and y directions, but in z axis they gradually increase to enable the robot endpoint to contact the object as if there was a virtual force attached to it by a human user. During the cutting step (approximately from 400 to 1600 time steps), the compliant profiles in x and z directions do not adapt as obviously as in y direction. This may be explained by the fact that the task dynamics in y axis is more difficult than that in the other two axes, and thus the impedance and feedforward force profiles have to be largely adapted to deal with it. And during the third step, it is similar to the first step that the impedance and force in z direction are accordingly adapted, and that in the other directions almost keep constant. In this task it can be observed again that our approach can enable the robot to adapt the dynamics profiles selectively in different axes according to the task requirement. The measured force of the robot endpoint in z axis during the task is shown in Fig. 12(a), and the sum of absolute values of the second derivative of these force profiles is shown in 12(b). Compared with the constant stiffness impedance control, our approach enables a more compliant cutting behaviour which is illustrated by a more smooth force profile.

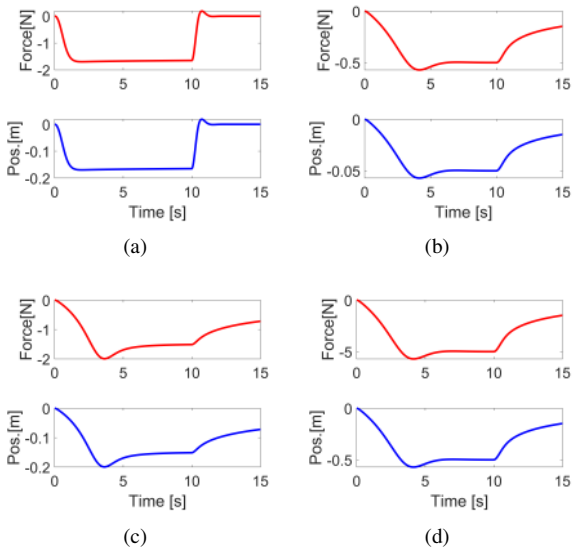


Figure 13: The simulation results using (a) the proposed approach and the adaptive impedance controller in [21] with different settings (b) $f_d = 0.5$N, (c) $f_d = 1.5$N, and (d) $f_d = 5$N.

B. *Comparison with variable impedance controller in [21]*

In this subsection, we illustrate the performance of our control strategy in simulation, compared with the variable impedance controller in [21]. In the simulation task, a 1-DOF robot is controlled to cut an object with the desired cutting depth 0.15[m], starting from the start point at the timestamp $t = 0$, i.e., $x_0 = 0$, and $x_d = -0.15$[m]. And from the timestamp $t = 10$[s], the robot is required to return back to the start point, i.e., $x_0 = -0.15$[m], and $x_d = 0$.

The stiffness of the environment is set as 10[N/m], and the inertia of the robot $M = 5$. The parameters of our control strategy are set as $Q_K = 2000$, $Q_D = 500$, $Q_v = 0$. The variable impedance controller developed in [21] requires an additional input variable, i.e., the desired interaction force f_d. To illustrate the comparative results, we set three values for this variable, i.e., $f_d = 0.5$N, $f_d = 1.5$N, and $f_d = 5$N.

The simulation results are shown in Fig. 13. We can observe that our approach can enable the robot to efficiently cut into the object, and complete the task more smoothly [see Fig.

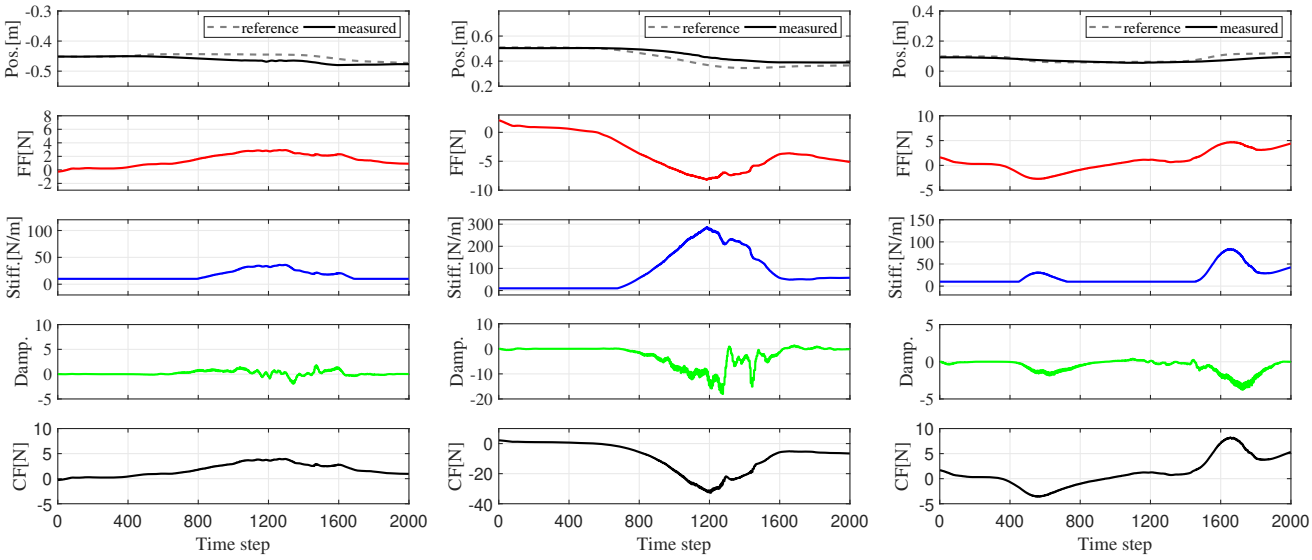


Figure 11: The online learning results of the compliant profiles in the cutting task while executing the reference movement trajectory in x, y and z axes.

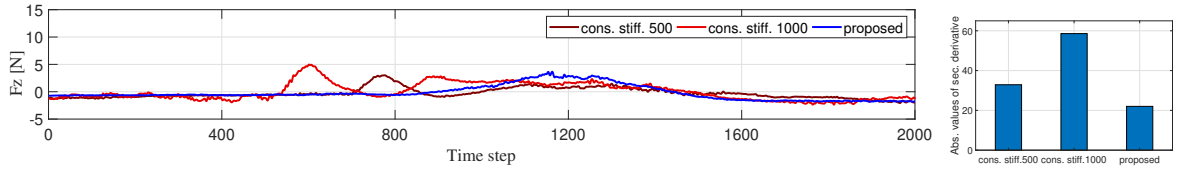


Figure 12: (a) The measured endpoint force profiles in z axis during robot reproduction under the different control conditions, and (b) the sum of absolute values of the second derivative of these force profiles.

13(a)], compared with the controller in [21]. Moreover, the predefined desired force f_d has a large influence on the control performance and an improper value may result in the failures of tracking of the desired position. Even with an proper f_d that allows the desired cutting depth, it cannot achieve a fast response compared to our approach. This suggests that our approach has an advantage to deal with this kind of tasks. Despite these results, we would like to point out that we cannot claim that our approach is better than the one in [21] for all tasks. The controller in [21] may outperform ours in other tasks where explicit force control is required.

VI. Discussion and Conclusion

A robotic learning approach (i.e., Hl-CMPs) is introduced in this paper, which can provide a novel strategy to address the question *how to enable the robot to learn compliant behaviours*. Our approach is based on a biomimetic control strategy inspired by the recent neuroscience findings of the human motor learning principles in the muscle space. Two typical robot-environment physical interaction tasks, i.e., insertion and cutting tasks, have been conducted to validate the performance of the proposed approach. From the experiment results, we can observe that the constant impedance control mode (including high and low stiffness values) may result in unstable and large contact force profiles [see Figs. 5(b), 7 and 8 for the insertion task and Fig. 12 for the cutting task]. On the contrary, the proposed approach could achieves a more smooth profile. It means that Hl-CMPs allows the robot to deal with this kind of physical interaction tasks with a versatile manipulation skill

Compared with the state-of-the-art approaches in the learning of robotic compliant movement, as stated above, the most important advantages of our approach lie in two aspects: i) all the compliant profiles (stiffness, damping and feedforward/task-specific force) are simultaneously obtained, instead of a separate identification/calibration procedure; and ii) they are all online learned during the execution of a specific task, without the need for an offline estimation process.

In our experiments, for simplicity the learning rates Q_K, Q_D and Q_v are set the same values for all the DOFs. Hl-CMPs leaves the flexibility that one can particularly set different values for different DOFs in a specific task. Furthermore, if the parameters are too small it would be too slow to obtain the proper dynamics profiles; but if they are set too large, it may cause system instability. One can start by setting relatively low values, and then increase them progressively until an acceptable range of impedance parameters is achieved. Furthermore, it could be explored to improve our approach in future work

by estimating these parameters from the demonstration data as suggested in [22, 23].

Compared with the high impedance control, the performance in position tracking of the proposed approach is not good enough. In the future work, we will consider to include demonstration force information into the Hl-CMPs model, namely, the demonstration force will also be encoded during the learning process, to address this issue. The Hl-CMPs model may be further improved by generalizing the learned dynamics profiles (including impedance and feedforward force) to novel tasks. For the sake of unification, generalization should be achieved in the parametric space. A possible method is utilizing regression algorithms (see, e.g., [13]) to model the dynamics parameters corresponding to task parameters (i.e., queries), and then generalizing these dynamics parameters (i.e., θ_k, θ_d and θ_v) according to the variations of the queries to deal with new tasks.

APPENDIX

First, we assume that the external force f_{ext} as

$$f_{ext} = K_s^* e + K_d^* \dot{e} + v^* \tag{32}$$

Then, Eq. 3 is adapted by adding another term f_0 that considers the dynamics of the robot

$$f_c = f_0 - u - v \tag{33}$$

where

$$f_0 = M\ddot{q}_e + C\dot{q}_e + G - L\varepsilon \tag{34}$$

with a symmetric positive-definite matrix L with minimal eigenvalue, and $\dot{x}_e$ is an auxiliary variable, $\dot{x}_e = \dot{x}_d - \pi e$.

Using Eq. 33 and Eq. 2, we obtain

$$M\dot{\varepsilon} + C\varepsilon = -u - v + f_{ext} - L\varepsilon \tag{35}$$

Combining with Eq. 32 further yields

$$M\dot{\varepsilon} + C\varepsilon = -\tilde{K}_s e - \tilde{K}_d \dot{e} - \tilde{v} - L\varepsilon \tag{36}$$

The first derivative of J_c is

$$\dot{J}_c = \tilde{\Phi}^T Q^{-1} \dot{\tilde{\Phi}} \approx \tilde{\theta}_k^T Q_k^{-1} \dot{\tilde{\theta}}_k + \tilde{\theta}_d^T Q_d^{-1} \dot{\tilde{\theta}}_d + \tilde{\theta}_v^T Q_v^{-1} \dot{\tilde{\theta}}_v \tag{37}$$

And the first derivative of J_e is

$$\begin{aligned} \dot{J}_e =& \varepsilon^T M\dot{\varepsilon} + \frac{1}{2}\varepsilon^T \dot{M}\varepsilon = \varepsilon^T (M\dot{\varepsilon} + C\varepsilon) \\ =& -\left[\varepsilon^T (\tilde{K}_s e + \tilde{K}_d \dot{e} + \tilde{v})\right] - \varepsilon^T L\varepsilon \end{aligned} \tag{38}$$

with

$$\tilde{K}_s = diag\{\tilde{\theta}_k \cdot g\}, \tilde{K}_d = diag\{\tilde{\theta}_d \cdot g\}, \tilde{v} = \tilde{\theta}_v \cdot g \tag{39}$$

Subsequently, the first derivative of the overall cost is

$$\begin{aligned} \dot{J}_{all} =& \dot{J}_c + \dot{J}_e = \tilde{\theta}_k^T Q_k^{-1} \dot{\tilde{\theta}}_k - \varepsilon^T diag\{\tilde{\theta}_k \cdot g\} e \\ &+ \tilde{\theta}_d^T Q_d^{-1} \dot{\tilde{\theta}}_d - \varepsilon^T diag\{\tilde{\theta}_d \cdot g\} \dot{e} \\ &+ \tilde{\theta}_v^T Q_v^{-1} \dot{\tilde{\theta}}_v - \varepsilon^T (\tilde{\theta}_v \cdot g) - \varepsilon^T L\varepsilon \end{aligned} \tag{40}$$

The stiffness, damping and feedforward force are then adapted to minimize the cost. For the stiffness, we have

$$\tilde{\theta}_k^T Q_k^{-1} \dot{\tilde{\theta}}_k = \varepsilon^T diag\{\tilde{\theta}_k \cdot g\} e \tag{41}$$

Here we assume that the dynamics of the environment is time-invariant, i.e., $\dot{\tilde{\theta}}_k = \dot{\theta}_k - \dot{\theta}_k^* = \dot{\theta}_k$.

For the m-th DOF, we design the following adaptation law to satisfy Eq. 41

$$\dot{\theta}_{k,m}^T = Q_{k,m} \varepsilon_m e_m g \tag{42}$$

The updating law for damping and feedforward force are obtained in a similar way.

REFERENCES

[1] Y. Liu, Z. Li, H. Liu, and Z. Kan, "Skill transfer learning for autonomous robots and human-robot cooperation: A survey," *Robotics and Autonomous Systems*, p. 103515, 2020.

[2] A. G. Billard, S. Calinon, and R. Dillmann, "Learning from humans," in *Springer handbook of robotics*. Springer, 2016, pp. 1995–2014.

[3] W. He, Z. Li, and C. P. Chen, "A survey of human-centered intelligent robots: issues and challenges," *IEEE/CAA Journal of Automatica Sinica*, vol. 4, no. 4, pp. 602–609, 2017.

[4] F. Qin, D. Xu, D. Zhang, and Y. Li, "Robotic skill learning for precision assembly with microscopic vision and force feedback," *IEEE/ASME Transactions on Mechatronics*, vol. 24, no. 3, pp. 1117–1128, 2019.

[5] S. Xu, Y. Ou, J. Duan, X. Wu, W. Feng, and M. Liu, "Robot trajectory tracking control using learning from demonstration method," *Neurocomputing*, vol. 338, pp. 249–261, 2019.

[6] 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.

[7] Z. Zhang, Z. Yan, and T. Fu, "Varying-parameter rnn activated by finite-time functions for solving joint-drift problems of redundant robot manipulators," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 12, pp. 5359–5367, 2018.

[8] L. Han, W. Xu, P. Kang, and H. Yuan, "Unified neural adaptive control for multiple human-robot-environment interactions," *IEEE Transactions on Industrial Informatics*, 2020.

[9] Y. Yuan, Z. Li, T. Zhao, and D. Gan, "Dmp-based motion generation for a walking exoskeleton robot using reinforcement learning," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 5, pp. 3830–3839, 2020.

[10] Z. Zhang, S. Chen, X. Zhu, and Z. Yan, "Two hybrid end-effector posture-maintaining and obstacle-limits avoidance schemes for redundant robot manipulators," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 2, pp. 754–763, 2020.

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

[12] 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, p. 446, 2001.

[13] T. Petrič, A. Gams, L. Colasanto, A. J. Ijspeert, and A. Ude, "Accelerated sensorimotor learning of compliant movement primitives," *IEEE Transactions on Robotics*, vol. 34, no. 6, pp. 1636–1642, 2018.

[14] M. Deniša, A. Gams, A. Ude, and T. Petrič, "Learning compliant movement primitives through demonstration and statistical generalization," *IEEE/ASME transactions on mechatronics*, vol. 21, no. 5, pp. 2581–2594, 2015.

[15] L. Bi, C. Guan *et al.*, "A review on emg-based motor intention prediction of continuous human upper limb motion for human-robot collaboration," *Biomedical Signal Processing and Control*, vol. 51, pp. 113–127, 2019.

[16] F. Bian, D. Ren, R. Li, P. Liang, K. Wang, and L. Zhao, "An extended dmp framework for robot learning and improving variable stiffness manipulation," *Assembly Automation*, vol. 40, no. 1, pp. 85–94, 2019.

[17] C. Zeng, C. Yang, H. Cheng, Y. Li, and S.-L. Dai, "Simultaneously encoding movement and semg-based stiffness for robotic skill learning," *IEEE Transactions on Industrial Informatics*, 2020.

[18] J. Buchli, F. Stulp, E. Theodorou, and S. Schaal, "Learning variable impedance control," *The International Journal of Robotics Research*, vol. 30, no. 7, pp. 820–833, 2011.

[19] 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*, vol. 66, no. 10, pp. 7788–7799, 2018.

[20] 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, 2017.

[21] 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.

[22] K. Kronander and A. Billard, "Online learning of varying stiffness through physical human-robot interaction," in *2012 IEEE International Conference on Robotics and Automation*. Ieee, 2012, pp. 1842–1849.

[23] A. K. Tanwani and S. Calinon, "Learning robot manipulation tasks with task-parameterized semitied hidden semi-markov model," *IEEE Robotics and Automation Letters*, vol. 1, no. 1, pp. 235–242, 2016.

[24] Y. Li, G. Ganesh, N. Jarrassé, 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, 2018.

[25] G. Ganesh, A. Albu-Schäffer, M. Haruno, M. Kawato, and E. Burdet, "Biomimetic motor behavior for simultaneous adaptation of force, impedance and trajectory in interaction tasks," in *2010 IEEE International Conference on Robotics and Automation*. IEEE, 2010, pp. 2705–2711.

[26] C. Yang, G. Ganesh, S. Haddadin, S. Parusel, A. Albu-Schaeffer, and E. Burdet, "Human-like adaptation of force and impedance in stable and unstable interactions," *IEEE transactions on Robotics*, vol. 27, no. 5, pp. 918–930, 2011.

[27] A. J. Ijspeert, J. Nakanishi, and S. Schaal, "Movement imitation with nonlinear dynamical systems in humanoid robots," in *Proceedings 2002 IEEE International Conference on Robotics and Automation*, vol. 2. IEEE, 2002, pp. 1398–1403.

Chao Zeng (S'18-M'20) received the Ph.D. degree in pattern recognition and intelligent systems from South China University of Technology, Guangzhou, China, in December 2019. He visited Department of Informatics, University of Hamburg (UHH), Germany, from October 2018 to October 2019. He is a research fellow in Guangdong University of Technology where he did this research from January 2020. He is currently a research associate at UHH from September 2020. His research interests include human-robot physical interaction, robot control, programming by demonstration and human-robot skill transfer.

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. His general research interests include human-robot interaction, robot control and control theory and applications.

Hang Su (S'12-M'19) received the M.Sc. degree in control theory and control engineering in South China University of Technology, Guangzhou, China, in 2015 and the Ph.D. degree in Bioengineering from Politecnico di Milano, Milano, Italy, in 2019. He participated in the EU funded project (SMARTsurg) in the field of Surgical Robotics. Dr. Hang Su is currently working in the Department of Electronics, Information and Bioengineering (DEIB) of Politecnico Di Milano. He is also working as an Ass Prof in the Institute of Advanced Technology, University of Science and Technology of China, China. He is currently a Program Chair of IEEE International Conference on Advanced Robotics and Mechatronics (ICARM 2021). He also severs as the Associate Editor for IEEE ICRA and the IEEE/RSJ IROS, IEEE RO-MAN and IEEE ICARM, and the Guest Associate Editor for a couple of journals, like IEEE Robotics and Automation Letters, Complexity, Actuators, Mathematical Problems in Engineering, Sensors, Frontiers in Robotics and AI, etc..

He was a recipient of the Best Paper Award in Advanced Robotics at the IEEE ICARM 2020, and the ICRA Travel Award funded by the IEEE Robotics and Automation Society in 2019. His main research interests include control and instrumentation in medical robotics, human-robot interaction, surgical robotics, deep learning, bilateral teleoperation, etc..

Jing Guo obtained his Ph.D degree from LIRMM, CNRS-University of Montpelier, France in 2016. He received his master and bachelor degree from Guangdong University of Technology in 2009 and 2012 respectively. He has been research fellow at National University of Singapore (NUS) during 2016-2018. He is an associate proferssor affilicated with Guangdong University of Technology. His current research interests include robotic control and learning, haptic bilateral teleoperation, and surgical robotics, has served as guest editor for IEEE RA-L, Frontiers in Robotics and AI, etc.

Chenguang Yang (M'10-SM'16) received the Ph.D. degree in control engineering from the National University of Singapore, Singapore, in 2010, and postdoctoral training in human robotics from the Imperial College London, London, U.K. He was awarded EU Marie Curie International Incoming Fellowship (individual) grant and UK EPSRC UKRI Innovation Fellowship grant. He won the Best Paper Award of the IEEE Transactions on Robotics as well as over ten international conference Best Paper Awards. He is a Fellow of British Computer Society and a Fellow of Higher Education Academy. He is a Co-Chair of IEEE Technical Committee on Collaborative Automation for Flexible Manufacturing (CAFM) and a Co-Chair of IEEE Technical Committee on Bio-mechatronics and Bio-robotics Systems (B2S). He serves as Associate Editors of a number of international top journals including seven IEEE Transactions. He is also supervisor of an EU H2020 Marie Skłodowska-Curie Standard European Individual Fellow. His research interest lies in human robot interaction and intelligent system design.