Learning Compliant Robotic Movements based on Biomimetic Motor Adaptation

Chao Zeng†1, Xiongjun Chen†2, Ning Wang[3], Chenguang Yang*3

Abstract

It is one of the great challenges for a robot to learn compliant movements in interaction tasks. The robot can easily acquire motion skills from a human tutor by kinematics demonstration, however, this becomes much more difficult when it comes to the compliant skills. This paper aims to provide a possible solution to address this problem by proposing a two-stage approach. In the first stage, the human tutor demonstrates the robot how to perform a task, during which only motion trajectories are recorded without the involvement of force sensing. A dynamical movement primitives (DMPs) model which can generate human-like motion is then used to encode the kinematics data. In the second stage, a biomimetic controller, which is inspired by the neuroscience findings in human motor learning, is employed to obtain the desired robotic compliant behaviours by online adapting the impedance profiles and the feedforward torques simultaneously. Several tests are conducted to validate the effectiveness of the proposed approach.

Keywords: Compliant Robotic Movements; Biomimetic Motor Control; Impedance Adaptation; Learning from Demonstration (LfD); Human-Robot Interaction and Collaboration.

1. Introduction

Nowadays, an industrial robot is most likely to be programmed to perform tasks in structured environments. The robot is controlled under a fixed position control mode without much flexibility and adaptatibility. This kind of robotic manufacturing systems can not gradually meet the increasing requirements of *High-Mix*, *Low-Volume* and *Short-Cycle* production in the market [1]. One promising solution to this problem is to integrate human factors into the robotic manufacturing systems in order to construct human-in-the-loop human-cyber-robot-systems (HCRS) [2]. By taking the advantages of both robots (e.g., good-repeatability) and humans (e.g., flexibility and adaptability), it has a great potential to improve the-state-of-art robotic-based production and to remove the barriers toward the new generation of intelligent manufacturing.

A number of approaches have been recently developped for the enhancement of robot learning in order to improve the robotic manipulation abilities (e.g., [3, 4]). Specifically, learning from human demonstration has been considered as an effective and efficient way to bring together humans' and robots' advantages [5, 6]. LfD allows to conveniently transfer human skills to a robot without the need of an expert's specific knowledge. LfD has been widely utilized for robotic skill learning in the field of human-robot interactions and human-robot collaborations in the last decades [7, 8, 9, 10, 11].

Most of previous studies in LfD have only concentrated on the learning of motion movements. These approaches can be utilized to address the encoding of motion profiles in a specific task. However, for force-dominant tasks these approaches may be insufficient. Even in a simple robotic pick-and-place task, for instance, when it comes to the consideration of the task dynamics compliant manipulation not only the motion planning should be addressed. Very recently, some researchers in the society of robotics have developed force/impedance-based approaches to enable the learning of compliant behaviours from humans [12, 13]. What should be emphasized is that the variable impedance control strategy has nearly become a common view that could help to achieve this point [14, 15]. However, it is not easy and continent to obtain variable impedance profiles, and a time-consuming complex process is often required.

In this work, we propose an approach based on the human biomimetic motor adaptation to address the above problem. The overview of the proposed approach is shown in Fig. 1. It basically consists of there steps: a human user first demonstrates the robot to perform a task during which the motion data, i.e., position and velocity trajectories, are recorded. The interaction force information is unnecessary and thus the force sensor is not needed in our approach. This step is optional since there are other ways to obtain the motion profiles; Then, the motion encoding model is fitted using the data obtained in the first step; Subsequently, during the robotic reproduction of the task the outputs of the model are used as reference position profiles, along which the compliant profiles (impedance and feedforward force) are learned based on the biomimetic controller.

Our contribution lies in the integration of biomimetic control into a robotic skill learning framework. With our approach compliant skills including impedance profiles can be efficiently

*Corresponding author

Email address: cyang@ieee.org (Chenguang Yang*)

[1]School of Automation, Guangdong University of Technology, Guangzhou 510006, China.

[2]School of automation science and engineering, South China University of Technology, Guangzhou 510641, China.

[3]Bristol Robotics Laboratory, University of the West of England, Bristol, BS16 1QY, UK.

This work was partially supported by National Nature Science Foundation (NSFC) under Grant 62003096 and Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/S001913.

† Contributed Equally

learned along the movement trajectories during the task executions, which can greatly facilitate the learning of human-like motor skills. Human-robot interactive and collaborative tasks have been performed and validated the proposed approach.

The rest of this paper is structured as follows. The related work is summarized in Section 2. The methodology is presented in Section 3. The experimental evaluation and discussion are detailed in Section 4. Section 5 finally concludes this paper.

2. Related Work

Until now, there are four main ways in the literature to obtain proper stiffness trajectories for robotic variable impedance control, i.e., the EMG-based; the optimization-based; the force-based; and the biomimetic control approaches, which are separately introduced as below.

i) *EMG-based*: The EMG (electromyography) signals detected from human arms can be utilized to extract human limb stiffness features. Therefore, the human arm stiffness profile can be estimated based on EMG during the interactions with robots. A number of studies have reported their results on this point. Typically, [14] proposed an EMG-based tele-impedance concept which could enable to transfer the human arm stiffness to a teleoperated robot. [16] and [17] proposed an EMG-based human-robot stiffness transfer interface that could allow robots to imitate both motion and impedance behaviours from humans. [18] and [19] proposed to use EMG signals to adapt impedance in the human-robot collaboration and coordination control scenarios.

Most of the studies utilized the EMG signals to estimate the diagonal elements in the human arm endpoint stiffness matrix. In [20], a model based estimator was developed to extract human arm complete joint stiffness. However, EMG-based approaches need a complex process to estimate the parameters of the EMG-impedance mapping model, which may be sometimes time-consuming. The parameters vary from one human user to another due to the different arm characteristics, and it is quite difficult to learn a general model for multiple different human demonstrators. Besides, the human arm configuration would have a large effect on the estimation results.

ii) *Optimization-based*: The optimization-based approaches prefer to learn a proper stiffness profile for variable impedance control by using optimization techniques such as reinforcement learning [21], black-box evolution [22], and adaptive control [23], etc. A constant reference stiffness trajectory is used for the initialization of there models, and then a number of trials are often required to learn a decent stiffness profile. The disadvantage of this method is that it is sometimes not easy to define a good reward/cost function, especial for a complex task, resulting in the need of a large number of trial and error which could be harmful to the robotic platforms.

iii) *Force-based*: Force-based approaches refers to use a force sensor mounted onto the robotic endpoint during demonstration to measure the interaction force, based on which the stiffness is estimated. Typically, [7] used Gaussian mixture model to encode the joint dataset (position and force) and then used Gaussian mixture regression to get the stiffness profile based on the learned model. [24] extended the work to use hidden semi-Markov model to model the correction between the position/rotation and the force/torque. In [25], the covariance matrix of the force data was first computed, then the stiffness was estimated based on the eigenvalues/eigenvectors of the covariance matrix.

Obviously, these approaches need at least one force to estimated the stiffness which could increase the expenses of the robotic systems. More importantly, the stiffness estimation strongly depends on the force signals, making it suffering from noises and the performance of the force sensor.

The common shortcoming of the methods discussed above is that they could not enable the robot to automatically adapt the impedance in an online manner, making it quite inefficient for the learning of compliant movements.

and iv) *Biomimetic control*: Biomimetic control approaches are inspired by the human motor learning. It argues that the impedance and feedforward torque/force should be concurrently adapted in order to deal with stable and unstable situations in unknown environments [26, 27]. In [28], a biomimetic controller was proposed based on this argument and implemented on a robot with one DOF (degree of freedom). [29] extended this controller for dual robotic arms (each with 2 DOFs) collaboration task in simulation. So far few studies have been reported to integrate biomimetic control into a robot learning framework until this work. Here we propose an approach based on this bio-inspired controller which can enable a robot with high DOFs to learn compliant motor skills from the human demonstration and from the human-robot collaboration.

3. Methodology

In this section, we first introduce the task representation model used in our approach, i.e., the dynamical movement primitives. Then, the impedance control model is simply presented. Finally, the biomimetic controller is given for the adaptation of the impedance, as well as the feedforward torque.

3.1. Learning a task using DMPs

DMPs is a well-known model which is able to efficiently represent a skill/task, and has been widely used in a large number of articles. It can model and generate human-like movements. For the sake of completeness, here we give a brief introduction to DMPs. For more details, please refer to [30, 31, 32]

Basically, the DMPs model can be separated into the following two parts.

3.1.1. Constructing a second-order non-linear system

First, a second-order non-linear system is constructed to model a specific motion trajectory. Based on the motion types, i.e., the *Discrete* movements and the *Rhythmic* movements, different non-linear systems are needed for different kinds of tasks. Here we are interested in the former type. For a one-DOF

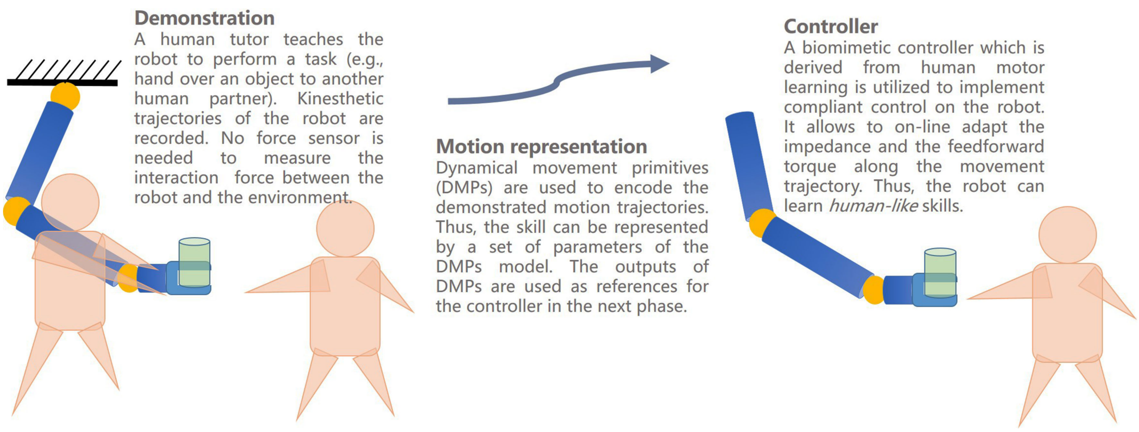


Figure 1: The overview of the proposed approach for the learning of compliant skills.

movement trajectory, the system is defined by the following equations [31].

$$\tau \dot{y} = k(g - x) - d\dot{x} - k(g - x_0)s + kf(s) \tag{1}$$

$$\tau \dot{x} = y \tag{2}$$

$$\tau \dot{s} = -\alpha_s s \tag{3}$$

$$f(s) = \frac{\sum_{i=1}^{N} \gamma_i \phi_i(s) s}{\sum_{i=1}^{N} \phi_i(s)} \tag{4}$$

$$\phi_i(s) = \exp(-h_i(s - c_i)^2) \tag{5}$$

where x and y represent the angle in joint space (or the position in Cartesian space) and corresponding velocity of the one-DOF movement trajectory. The velocity is often obtained by the direct numerical differentiation operation over the angle trajectory x. x_0 and g represent the initial value and the goal (i.e., the last value) of the angle trajectory, respectively. Eq. 1 can also be considered as a spring-damping system with the spring parameter k and the damping parameter d, respectively, which are often properly chosen in advance as $d = 2\sqrt{k}$. τ is the temporal constant which is used to control the evolution duration of the system. The whole system is driven by the phase variable s generated from Eq.3 instead of directly using time such that the evolution of the system can be efficiently edited. $s \in (0, 1]$ starts from 1 and monotonically converges to 0 along with the duration of the motion trajectory, granting that the motion finally converges to goal point. α_s is a pre-defined constant coefficient.

The non-linear force term $f(s)$ in Eq. 1 is determined by Eq. 4. $\phi_i(s)$ represents the widely used Gaussian basis functions with the width $h_i > 0$, and the center c_i which is evenly distributed along with the phase variable s. N represents the total number of the Gaussian basis which needs to be set in advance. γ_i represents the parameters of the DMPs model, which can be utilized to regulate the shape of the force term and thus to regulate the shape of the motion trajectory. It can be seen that the specific task/skill can be parametrized by a set of parameters associated with corresponding motion variables (e.g., the starting point, the goal and the duration of the movement).

Note that Eqs. 1, 2 and 4 are used for each separate DOF, which Eq. 3 is shared across all the DOFs. For example, for the encoding of a 7-DOFs robot arm movements, all the 7 movement trajectories (represented by Eq. 1) are driven by the same phase variable such that the duration synchronization of the whole system can be strictly guaranteed. Furthermore, Eqs. 1-3 can be coupled with additional spatial and temporal terms for specific usages [33, 34, 35].

3.1.2. Learning the DMPs model

The learning of the DMPs model here refers to the learning of the parameters γ_i as described above, which can be basically considered as a supervised learning problem [6].

Given one demonstration data consisting of a movement angle and a velocity trajectory $\{x_i, \dot{x}_i, \ddot{x}_i\}_{i=1}^{T}$, the following three steps are performed accordingly to adapt the parameters of the DMPs model. If the joint velocity and acceleration are not available, we can directly derivative the joint angle x at each time step.

i) *First step*: $s(t)$ is computed by integrating the canonical system as shown in Eq. 3.

ii) *Second step*: we construct a target function f_{target} based on Eq. 1.

and iii) *Third step*: locally weighted linear regression is utilized to solve the following equation, and thus to obtain the model parameter γ_i.

$$\min(\sum (f_{target}(s) - f(s))^2) \tag{6}$$

We choose the DMPs as the task representation model thanks to its a number of advantages. The first one of these lies in that it can be efficiently learned and generalized to other similar task situations. The second one is that it can represent any shape of trajectories theoretically. Furthermore, the optimization of the parameters can be easily formed as a reinforcement learning problem [21, 36, 37], which, however, will not be considered in this work.

3.2. Impedance Control Model

Considering a robotic arm with n DOFs, it's dynamics can often be expressed in joint space as follows [38].

$$M(x)\ddot{x} + C(x,\dot{x})\dot{x} + G(x) = \tau_c + J^T F \tag{7}$$

where x, $\dot{x}$ and $\ddot{x}$ represent the joint angle, velocity and acceleration, respectively. $M(x)$ represents the inertia matrix[4]. $C(x,\dot{x})$ denotes the Coriolis and Centrifugal forces, and $G(x)$ is the gravity force. F represents the force applied by the environment (including a human operator) in a specific interaction. The robotic arm dynamics $\tau_{dyn} = M(x)\ddot{x} + C(x,\dot{x})\dot{x} + G(x)$ are assumed known, they are provided by the robot manufacturer, or they are identified based on nonlinear adaptive control techniques (see e.g., [39]). J represents the robotic arm Jacobian matrix. τ_c represents the input control torque which will be detailed in the following section.

3.3. Learning of Compliant Movement Profiles based on Biomimetic Control

3.3.1. Robotic Compliant Movement Representation

Given the above robotic arm dynamics, we separate the control input τ_c into two parts. Inspired by the human arm motor learning regulations, the control command can be represented by the sum of a feedforward command and a feedback command [40, 41]:

$$\tau_c = u + v \tag{8}$$

where u represents the feedforward torque vector, and v represents the impedance (i.e., the feedback command vector) which is defined as a PD form in this work.

$$v = Ke + D\dot{e} \tag{9}$$

with the angle error and the velocity error:

$$e = x_r - x \tag{10}$$

$$\dot{e} = \dot{x}_r - \dot{x} \tag{11}$$

where x_r and $\dot{x}_r$ represent the reference joint angle and the reference joint velocity, respectively, which are the outputs of the DMPs model as explained in section 3.1. K and D represent the stiffness matrix and the damping matrix, respectively. The stiffness is a diagonal matrix, i.e.,

$$K = diag\{k_1, k_2, \cdots, k_n\} \tag{12}$$

where each of the elements corresponds to each joint stiffness of the robotic arm, and will be adapted according to the task requirements. The damping matrix is also a diagonal matrix determined by

$$D = diag\{d_1, d_2, \cdots, d_n\} \tag{13}$$

[4]For simplicity we do not use bold formatting in this work.

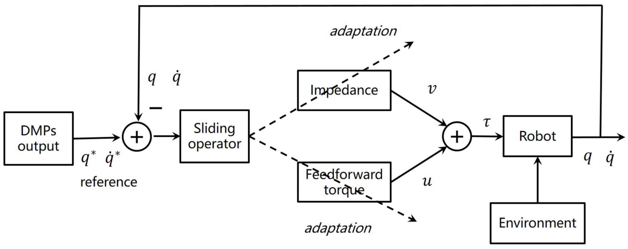


Figure 2: The control digram for the learning of impedance and feedforward force, they are simultaneously learned based on the errors between the reference and the current robotic motion ststes. This figure is adapted from [26].

which is determined by

$$d_i = 2\sqrt{k_i} \tag{14}$$

Until now, the compliant movements in our work include the movement trajectories, the stiffness profiles and the feedforward torque profiles. We conclude the compliant movements as below

$$\Omega = \{x_i, \dot{x}_i, K_i, v_i\}_{i=1}^T \tag{15}$$

3.3.2. Adaptation Law

The adaptation strategy of the variable impedance control is shown in Fig. 2. It shows that the feedforward torque and the impedance need to be updated at the same time within one control loop.

In the human motor learning, the goal is to minimize the movement error and the effort. Accordingly, we consider the following cost function [28].

$$J_{cost} = \frac{\alpha}{2} v^T v + \gamma \sum_{i=1}^{N} u_i \tag{16}$$

where the forward term is the cost for the movement feedback, and the last term is the cost for the feedforward. α and γ are positive constant coefficients which will be later extended as vectors for our usage.

With [28], each element in the feedback vector is assumed as a linear function increasing in both directions.

$$v_i = \varepsilon_{i,+} + \zeta\varepsilon_{i,-}, \quad \zeta \in (0,1) \tag{17}$$

where $\varepsilon_{i,+}$ and $\varepsilon_{i,-}$ represent the positive part and the negative part, respectively.

The sliding error is defined as

$$\varepsilon_i = \pi(e_i + \delta\dot{e}_i) \tag{18}$$

with the positive constant coefficients π and δ

The learning problem can be solved by the gradient descent law

$$\triangle u^t = \alpha v^t - \gamma \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_N \tag{19}$$

Then, based on the assumption Eq. 17, the above equation can be split into three parts, i.e.,

$$\triangle u^t = \frac{\alpha}{2}(1-\zeta)\varepsilon^t + \frac{\alpha}{2}(1+\zeta) \mid \varepsilon^t \mid -\gamma \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_N \tag{20}$$

with

$$\mid \varepsilon \mid = (\mid \varepsilon_1 \mid, \mid \varepsilon_2 \mid, \cdots, \mid \varepsilon_N \mid) \tag{21}$$

Finally, yielding the following update law[28]

$$\triangle K^t = \beta \mid \varepsilon^t \mid -\gamma \tag{22}$$

$$\triangle u^t = \alpha\varepsilon^t - (1-\mu)u^t \tag{23}$$

where β is a positive constant gain coefficients, and $\mu \in (0,1)$ is a relaxation factor. The stiffness K_i may become negative, therefore, K_i are limited into a proper range $[K_{i,min}, K_{i,max}]_{i=1}^N$

In this work, the following three aspects are modified for our usage:

i) For the convenient control of a robotic manipulator with multiple DOFs, we first extend the constant coefficients to vectors. α, β and γ are shared for all the muscles in the motor learning. However, the joints are separated and not coupled together for the robot arm. Accordingly, the objective function is thus adapted to [42]

$$J_{obj} = \min(\frac{\alpha}{2}v^T v + \sum_{i=1}^{N} \gamma_i u_i) \tag{24}$$

with N dimension vectors α and γ.

ii) The last term of Eq. 22 is adjusted based on the sliding error instead of constant values by

$$\gamma_i = \frac{a}{1 + b \mid \varepsilon_i \mid} \tag{25}$$

where a and b are pre-defined positive constant coefficients. With this formulation, γ_i can regulate the increment impedance of the corresponding joint.

and iii) The relaxation factor is also not fixed but adapted based on the error. Eq. 23 is accordingly modified as:

$$\triangle u^t = \alpha\varepsilon^t - \frac{1}{\exp(\mid \varepsilon \mid)}u^t \tag{26}$$

The stiffness and feedforward torque are updated by using Eqs. 22, 25 and 26 at each time step along the movement trajectories.

4. Experimental Validation

In order to verify the effectiveness of the proposed approach, the following three experiments have been performed. For all experiments, the robotic arm is controlled in joint space under the torque control model with a sampling rate of 2000Hz.

4.1. Simulation Task

The first experiment is a simulation task performed based on a simulated Baxter robot in the Gazebo environment[5]. The Baxter robot have two arms, each of them has 7 joints, i.e., 2-DOF shoulder joint (S0, S1), 2-DOF elbow joint (E0, E1), and 3-DOF wrist joint (W0, W1, W2). The task is a simulation "water-pouring" movement in which all the joints of the robot are involved. In the simulation, the robotic arm is controlled under a free motion manner [see Fig. 3(h)], i.e., no external force is applied onto the manipulator.

The parameter settings for the DMPs model are: $\tau = 1$, $\alpha_s = 1$, $k = 100$. The parameter settings for this simulation task are as below: $\pi = 1.2$, $\delta = 0.008$, $\beta = [0.3, 0.3, 0.3, 0.3, 0.3, 0.3, 0.5]^T$, $\alpha = [5.0, 5.0, 5.0, 5.0, 5.0, 5.0, 8.0]^T$, $a = 0.05$, $b = 10$, and $[K_{i,min}, K_{i,max}] = [2, 200]_{i=1}^7$.

The simulation results are shown in Fig. 3(a-g). It shows the movement trajectories, the stiffness and the feedforward torque profiles of the 7 joints. The straight dark lines are the trajectories learned by the DMPs model, and the dash ones are the measured angle trajectories during the reproduction of the task. It also shows the adaptation of both stiffness and feedforward during the evolution of the movement trajectories. The movement, impedance and force/torque of all the joints are adapted. Almost all the stiffness profiles follow the same pattern: increasing from a small value and then decreasing to a certain value, which is basically consistent with the human experience when performing this kind of tasks. Besides, the adaptation in time coordinate is also demonstrated as expected. Taking the last joint (W2) for an example, the stiffness and feedforward keeps constant during the reaching phase and thereafter they adapt to complete the "pouring" step (starting from about 2.5s)

4.2. Handover Task

The second task is implemented on a real-word Baxter robot which has the same structure with the simulated Baxter robot in the Gazebo. First, a human demonstrator teaches the robot how to hand over an object to another human partner, during which the robot arm states are recoded. The recorded data are then modelled by DMPs with the same parameters used in the first task. Subsequently, the robot play back the handover movement of the handover task without the human guidance again [see Fig. 4(h)], during which the stiffness and the feedforward are learned at each time step.

The parameter settings for this task are as below: $\pi = 1.3$, $\delta = 0.008$, $\beta = [4.0, 2.5, 1.0, 3.5, 0.3, 0.3, 0.3]^T$, $\alpha = [5.0, 5.0, 5.0, 5.0, 5.0, 5.0, 5.0]^T$, $a = 0.8$, $b = 10$, and $[K_{i,min}, K_{i,max}] = [15, 200]_{i=1}^7$.

The experimental results of this task are shown in Fig. 4(a-g). Again, it shows the robot is able to complete the task while keeping as compliant as possible: increasing stiffness if needed to compensate for the movement error, and keeping low if not necessary. Unlike in the first task, not all the joints are needed to adjust their impedance and feedforward values. If one joint

[5] http://sdk.rethinkrobotics.com/wiki/Baxter_Simulator

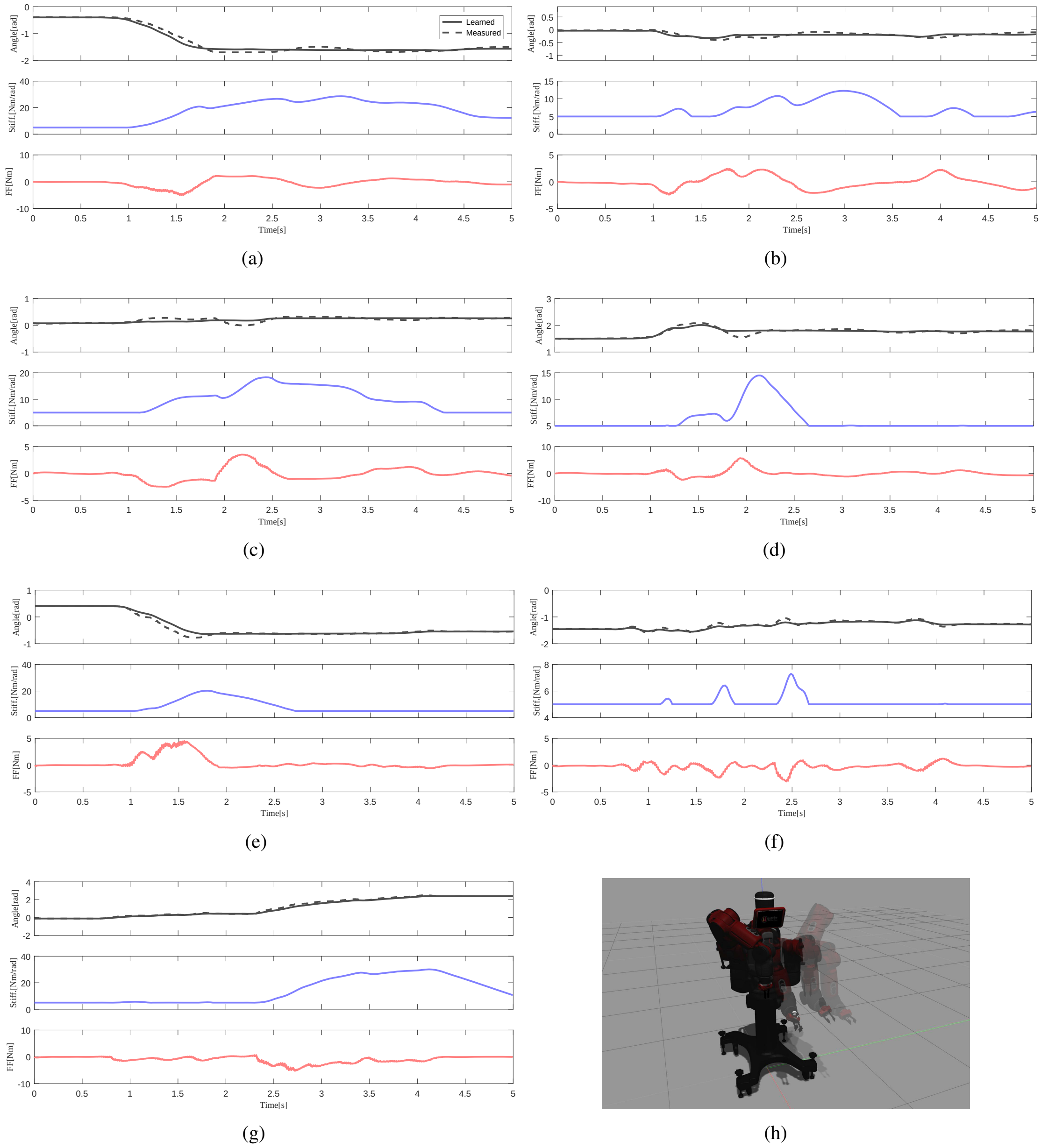


Figure 3: The experimental results of the simulation task.

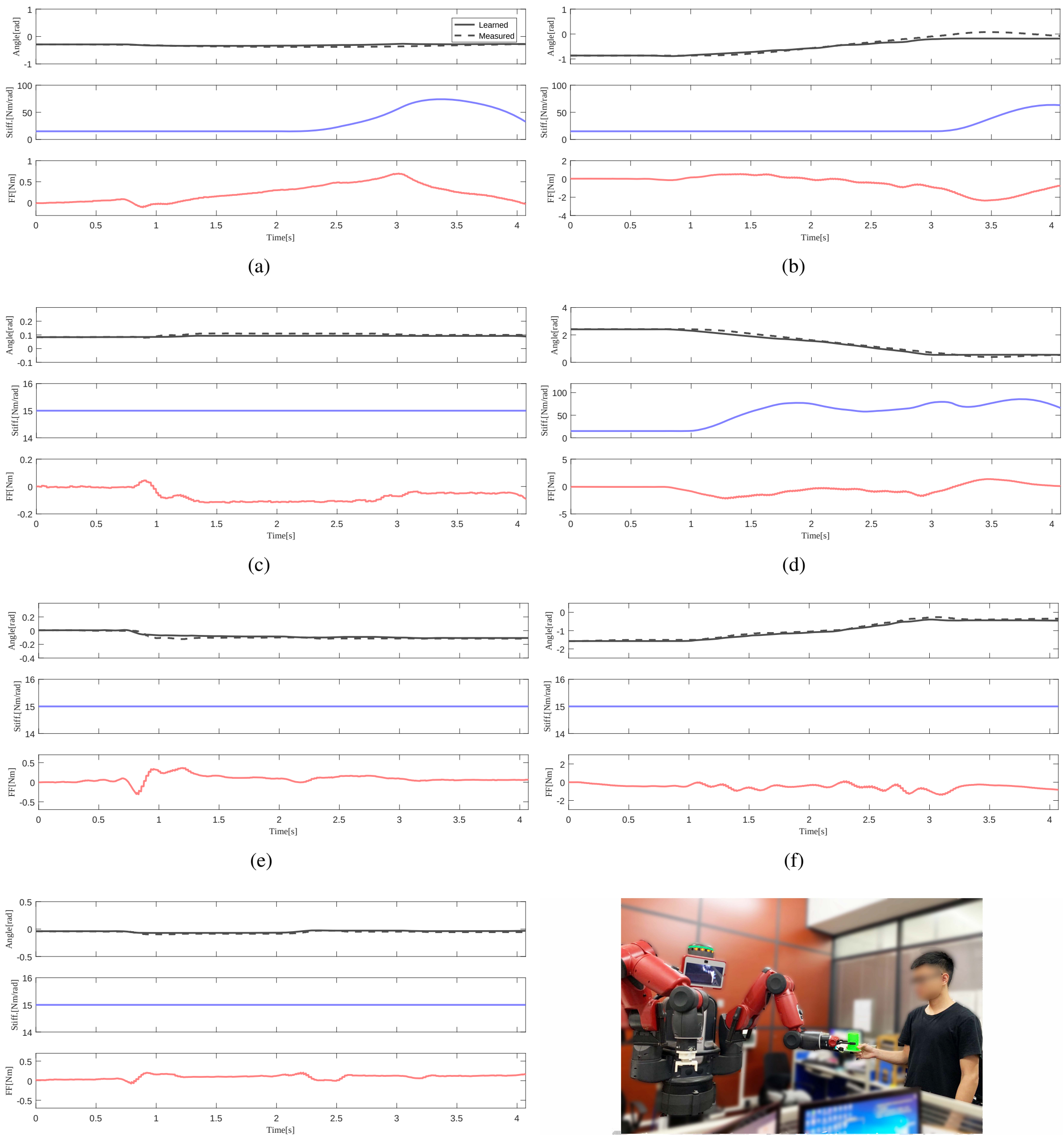
is not particularly involved, its impedance keeps at the smallest value.

4.3. Sawing Task

The third task is the human-robot collaborative sawing task. The setup for this task is shown in Fig. 5(e). A saw is connected to one of the robotic endpoints through a specifically designed module. The robot and the human partner collaborate to saw a piece of wood which is mounted onto the table. In this task, the reference angles remain unchanged and the reference velocities remain zero.

The settings for the sawing task are given as below: $\pi = 1.3$, $\delta = 0.01$, $\beta = [5.0, 2, 0.4, 0.75, 0.4, 0.6, 0.75]^T$, $\alpha = [5.0, 5.0, 5.0, 5.0, 5.0, 5.0, 5.0]^T$, $a = 0.6$, $b = 12$, and $[K_{i,min}, K_{i,max}] = [5, 200]_{i=1}^{7}$.

The experimental results of this task are shown in Fig. 5(a-d). It shows the measured angles, the joint torques and the stiffness of the 7 joints. There are three joints (i.e., S1, E1 and W1) mainly involved during the task execution, while the others (i.e., S0, E0, W0 and W2) almost keep constant(see Fig. 5(d)) since these joint angles do not change much during the sawing periods. It can be seen that the stiffness profiles of the three joints could be automatically adapted to the human partner during the sawing process. When the human partner increases his strength to pull the saw, the robot arm impedance increases gradually. When the robot arm impedance becomes large to some extent, the robot would start to pull it back while the human partner



(a) (b) (c) (d) (e) (f) (g) (h)

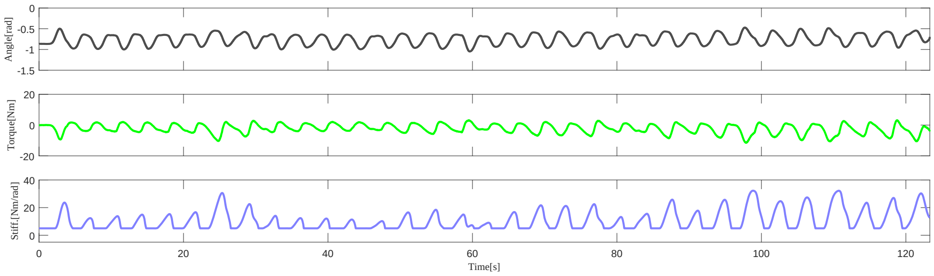
Figure 4: The experimental results and setup of the handover task.

loosening this arm strength. This period then repeats over and over until the task is finished finally.

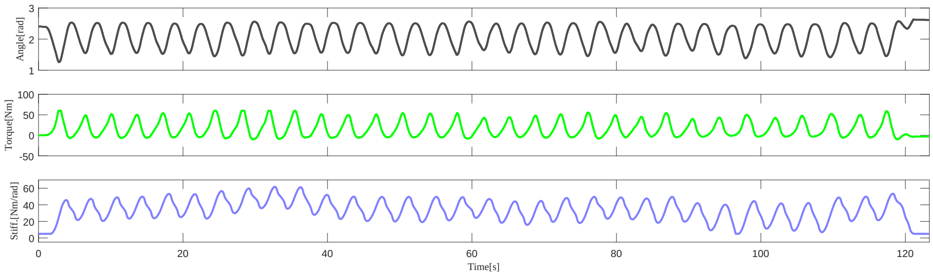
4.4. Discussion

From the experimental results, we can conclude that it is meaningful and helpful for a robot in a dynamic environment (e.g., handing over an object to a human partner, or collectively sawing with its human partner) to continuously adapt the interaction force/torque/impedance to satisfy the requirements of the task situations. Even for a *free-force* motion task (e.g., the simulation task in this work), the adaptation of impedance and feedforward torque can indeed help to obtain compliant robotic behaviours. The impedance and the task-specific torque profiles are simultaneously obtained without the need of learning the interaction dynamics.

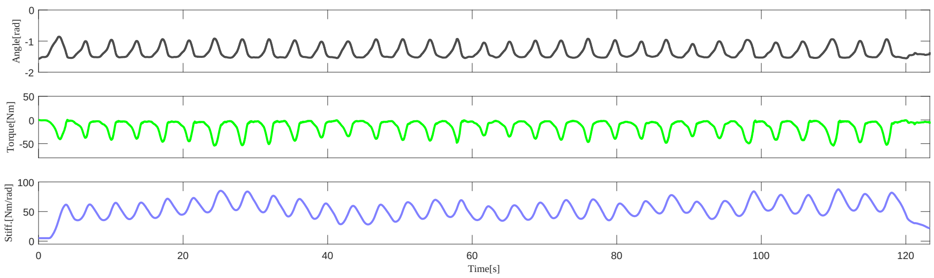
As stated before, several methods can be utilized to obtain variable impedance profiles. The EMG-based stiffness estimation methods (see e.g., [16, 17, 18, 43]) need an offline time-consuming process to identify the human arm impedance model. Another typical way to acquire proper impedance profiles is to refine them through interacting the environment based on reinforcement learning (RL) [37] or black-box (BB) optimization [44]. It usually needs a number of trials to finally learn a proper stiffness trajectory which may also be time-consuming for some complex tasks. Furthermore, compared with the force-based stiffness regulation methods (e.g., [24, 25], no additional force sensor is needed in this work which could reduce the cost of the robotic system.



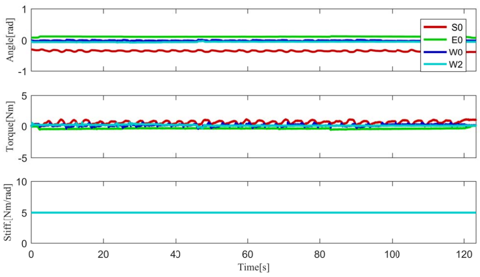
(a)



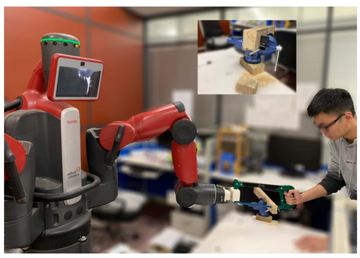
(b)



(c)



(d)



(e)

Figure 5: The experimental results and the setup of the collaborative sawing task.

There are at least two open strategies that may be used to improve our approach: i) Although the impedance and the torque profiles are learned online and therefore it is free of the time-consuming problem, however, it is difficult to always obtain a good tracking performance because many coefficients need to be properly set in advance. One possible solution is to balance the efficiency and the tracing accuracy by combination of the proposed approach and the RL-based or BB-based optimization techniques. ii) In this work, two-stage is needed for the representation of the movement and the learning of impedance as well as feedforward torque. The movement is encoded in parametric space, while the impedance and the feedforward are updated at the trajectory level. It would be much better to develop a unified representation for all the compliant profiles Ω (Eq. 15), which will largely be advantageous to the robotic skill learning.

5. Conclusion

In this work we propose a approach that can enable robots to learn compliant motor skills. Specifically, the compliant profiles include movement trajectories, impedance/stiffness profiles and joint feedforward torques. The DMPs model is utilized as the representation model for the encoding of the movement trajectories. The impedance and the feedforward profiles can be efficiently obtained based on a biomimetic controller during the task executions, which is derived from the human motor learning principles. The proposed approach has been verified by three tasks: the simulation task, the handover task on the Baxter robot and the human-robot collaboration task. In the future work, we will continue to improve the proposed approach as discussed above, and to implement our approach on more complex tasks.

References

[1] F. Chen, K. Sekiyama, F. Cannella, T. Fukuda, Optimal subtask allocation for human and robot collaboration within hybrid assembly system, IEEE Transactions on Automation Science and Engineering 11 (4) (2014) 1065–1075.

[2] Z. Ji, P. Li, Y. Zhou, B. Wang, J. Zang, M. Liu, Toward new-generation intelligent manufacturing, Engineering 4 (1) (2018) 11–20.

[3] H. Liu, F. Sun, X. Zhang, Robotic material perception using active multimodal fusion, IEEE Transactions on Industrial Electronics.

[4] H. Liu, F. Wang, F. Sun, B. Fang, Surface material retrieval using weakly paired cross-modal learning, IEEE Transactions on Automation Science and Engineering 16 (2) (2018) 781–791.

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

[6] B. D. Argall, S. Chernova, M. Veloso, B. Browning, A survey of robot learning from demonstration, Robotics and autonomous systems 57 (5) (2009) 469–483.

[7] L. D. Rozo, S. Calinon, D. Caldwell, P. Jiménez, C. Torras, Learning collaborative impedance-based robot behaviors, in: AAAI Conference on Artificial Intelligence, 2013.

[8] B. Ti, Y. Gao, Q. Li, J. Zhao, Human intention understanding from multiple demonstrations and behavior generalization in dynamic movement primitives framework, IEEE Access 7 (2019) 36186–36194.

[9] B. Nemec, N. Likar, A. Gams, A. Ude, Human robot cooperation with compliance adaptation along the motion trajectory, Autonomous Robots 42 (5) (2018) 1023–1035.

[10] B. Fang, X. Wei, F. Sun, H. Huang, Y. Yu, H. Liu, Skill learning for human-robot interaction using wearable device, Tsinghua Science and Technology 24 (6) (2019) 654–662.

[11] B. Fang, F. Sun, H. Liu, D. Guo, W. Chen, G. Yao, Robotic teleoperation systems using a wearable multimodal fusion device, International Journal of advanced robotic systems 14 (4) (2017) 1729881417717057.

[12] S. Calinon, P. Kormushev, D. G. Caldwell, Compliant skills acquisition and multi-optima policy search with em-based reinforcement learning, Robotics and Autonomous Systems 61 (4) (2013) 369–379.

[13] M. Deniša, A. Gams, A. Ude, T. Petrič, Learning compliant movement primitives through demonstration and statistical generalization, IEEE/ASME transactions on mechatronics 21 (5) (2016) 2581–2594.

[14] A. Ajoudani, N. Tsagarakis, A. Bicchi, Tele-impedance: Teleoperation with impedance regulation using a body–machine interface, The International Journal of Robotics Research 31 (13) (2012) 1642–1656.

[15] F. Ficuciello, L. Villani, B. Siciliano, Variable impedance control of redundant manipulators for intuitive human–robot physical interaction, IEEE Transactions on Robotics 31 (4) (2015) 850–863.

[16] C. Yang, C. Zeng, P. Liang, Z. Li, R. Li, 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 15 (1) (2018) 329–340.

[17] C. Yang, C. Zeng, C. Fang, W. He, Z. Li, A dmps-based framework for robot learning and generalization of humanlike variable impedance skills, IEEE/ASME Transactions on Mechatronics 23 (3) (2018) 1193–1203.

[18] L. Peternel, L. Rozo, D. Caldwell, A. Ajoudani, A method for derivation of robot task-frame control authority from repeated sensory observations, IEEE Robotics and Automation Letters 2 (2) (2017) 719–726.

[19] Z. Li, Y. Kang, Z. Xiao, W. Song, Human–robot coordination control of robotic exoskeletons by skill transfers, IEEE Transactions on Industrial Electronics 64 (6) (2017) 5171–5181.

[20] C. Fang, A. Ajoudani, A. Bicchi, N. G. Tsagarakis, Online model based estimation of complete joint stiffness of human arm, IEEE Robotics and Automation Letters 3 (1) (2018) 84–91.

[21] J. Buchli, F. Stulp, E. Theodorou, S. Schaal, Learning variable impedance control, The International Journal of Robotics Research 30 (7) (2011) 820–833.

[22] F. Stulp, O. Sigaud, Robot skill learning: From reinforcement learning to evolution strategies, Paladyn, Journal of Behavioral Robotics 4 (1) (2013) 49–61.

[23] W. He, Y. Dong, Adaptive fuzzy neural network control for a constrained robot using impedance learning, IEEE transactions on neural networks and learning systems 29 (4) (2018) 1174–1186.

[24] M. Racca, J. Pajarinen, A. Montebelli, 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.

[25] J. Duan, Y. Ou, S. Xu, Z. Wang, A. Peng, X. Wu, 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.

[26] C. Yang, G. Ganesh, S. Haddadin, S. Parusel, A. Albu-Schaeffer, E. Burdet, Human-like adaptation of force and impedance in stable and unstable interactions, IEEE transactions on robotics 27 (5) (2011) 918–930.

[27] Y. Li, G. Ganesh, N. Jarrassé, S. Haddadin, A. Albu-Schaeffer, E. Burdet, Force, impedance, and trajectory learning for contact tooling and haptic identification, IEEE Transactions on Robotics (99) (2018) 1–13.

[28] G. Ganesh, A. Albu-Schäffer, M. Haruno, M. Kawato, 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.

[29] E. Gribovskaya, A. Kheddar, A. Billard, Motion learning and adaptive impedance for robot control during physical interaction with humans, in: 2011 IEEE International Conference on Robotics and Automation, IEEE, 2011, pp. 4326–4332.

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

[31] P. Pastor, H. Hoffmann, T. Asfour, S. Schaal, Learning and generalization of motor skills by learning from demonstration, in: 2009 IEEE International Conference on Robotics and Automation, IEEE, 2009, pp. 763–768.

[32] A. J. Ijspeert, J. Nakanishi, H. Hoffmann, P. Pastor, S. Schaal, Dynamical movement primitives: Learning attractor models for motor behaviors, Neural Computation 25 (2) (2013) 328–373. `doi:10.1162/NECO_a_00393`.

[33] J. Kober, B. Mohler, J. Peters, Learning perceptual coupling for motor primitives, in: Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on, IEEE, 2008, pp. 834–839.

[34] A. Gams, B. Nemec, L. Zlajpah, M. Wächter, A. Ijspeert, T. Asfour, A. Ude, Modulation of motor primitives using force feedback: Interaction with the environment and bimanual tasks, in: 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2013, pp. 5629–5635.

[35] A. Gams, B. Nemec, A. J. Ijspeert, A. Ude, Coupling movement primitives: Interaction with the environment and bimanual tasks, IEEE Transactions on Robotics 30 (4) (2014) 816–830.

[36] F. Stulp, J. Buchli, A. Ellmer, M. Mistry, E. A. Theodorou, S. Schaal, Model-free reinforcement learning of impedance control in stochastic environments, IEEE Transactions on Autonomous Mental Development 4 (4) (2012) 330–341.

[37] Z. Li, T. Zhao, F. Chen, Y. Hu, C.-Y. Su, T. Fukuda, Reinforcement learning of manipulation and grasping using dynamical movement primitives for a humanoidlike mobile manipulator, IEEE/ASME Transactions on Mechatronics 23 (1) (2018) 121–131.

[38] L. Sciavicco, B. Siciliano, Modelling and control of robot manipulators, Springer Science & Business Media, 2012.

[39] E. Burdet, A. Codourey, L. Rey, Experimental evaluation of nonlinear

adaptive controllers, IEEE control systems magazine 18 (2) (1998) 39–47.

[40] E. Burdet, R. Osu, D. W. Franklin, T. E. Milner, M. Kawato, The central nervous system stabilizes unstable dynamics by learning optimal impedance, Nature 414 (6862) (2001) 446–449.

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

[42] Z. Chao, Y. Chenguang, C. Zhaopeng, Bio-inspired robotic impedance adaptation for human-robot collaborative tasks, Science China Information Sciences 63 (7) (2020) 170201.

[43] C. Yang, C. Zeng, Y. Cong, N. Wang, M. Wang, A learning framework of adaptive manipulative skills from human to robot, IEEE Transactions on Industrial Informatics 15 (2) (2019) 1153–1161.

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