A Learning Framework of Adaptive Manipulative Skills From Human to Robot

Chenguang Yang, *Senior Member, IEEE*, Chao Zeng, Yang Cong, *Senior Member, IEEE*, Ning Wang, *Member, IEEE*, and Min Wang, *Member, IEEE*

***Abstract*—Robots are often required to generalize the skills learned from human demonstrations to fulfil new task requirements. However, skill generalization will be difficult to realize when facing with the following situations: the skill for a complex multistep task includes a number of features; some special constraints are imposed on the robots during the process of task reproduction; and a completely new situation quite different with the one in which demonstrations are given to the robot. This work proposes a new framework to facilitate robot skill generalization. The basic idea lies in that the learned skills are first segmented into a sequence of subskills automatically, then each individual subskill is encoded and regulated accordingly. Specifically, we adapt each set of the segmented movement trajectories individually instead of the whole movement profiles, thus, making it more convenient for the realization of skill generalization. In addition, human limb stiffness estimated from surface electromyographic signals is considered in the framework for the realization of human-to-robot variable impedance control skill transfer, as well as the generalization of both movement trajectories and stiffness profiles. Experimental study has been performed to verify the effectiveness of the proposed framework.**

***Index Terms*—Adaptive stiffness control, human-robot skill transfer, skill segmentation, stiffness generalization.**

I. Introduction

ROBOT programming by demonstration (PbD) has recently received much attention because it can enable robots to be fast programmed [1]. Through PbD robots are able to efficiently learn manipulation skills for accomplishing tasks from human guidance. Compared to conventional programming methods, PbD has many advantages listed as follows: 1) it does not particularly require a human instructor to have expert knowledge of skills; and 2) most importantly, human characteristics such as flexibility and compliance can be taken into account in PbD systems, which will to a large extent facilitate the performance of task accomplishment.

Some advanced adaptive control strategies have recently been proved effective for the improvement of the control performance [2]–[4]; thus, it would be necessary to develop adaptive control techniques in PbD systems. One possible way is to capture the features of human limb stiffness adaptation during a specific task, and then transfer these features to a robot. For instance, a robot can perform tasks in a human-like manner by learning adaptive impedance control [5]–[8]. Particularly, transferring human arm stiffness estimated based on electromyographic (EMG) signals to a robot can achieve variable impedance control for the robot arm, which has demonstrated a better performance than without human-to-robot stiffness transfer [9], [10]. According to [11]–[14], force profiles in addition to positional profiles should also be regulated, especially for in-contact tasks. EMG-based variable impedance transfer is able to enable the robot to adapt the force profiles to different task situations with a natural regulation process [10], [15].

Generally, human demonstrations are often simply treated as trajectories (usually positional profiles and velocities) for a robot to imitate. Thus, the robot is only able to mimic the behaviours as demonstrated. From the perspective of programming efficiency, however, the robot is expected to be capable of generalizing the learned skills to new given task situations. To do so, the trajectories (or behaviours) can be encoded by several models such as dynamical movement primitives (DMP) [16], which has been widely applied to a number of tasks such as ball-in-a-cup [17], table tennis [18], and grasping [19]. In most of the studies, DMP is only used as PbD method to encode movement trajectories. In [20] and [21], DMP is also used for encoding force profiles obtained from demonstrations. In this work, we will use DMP to encode both movement trajectories and stiffness profiles to achieve variable impedance skill transfer and generalization.

Manuscript received December 6, 2017; revised February 16, 2018; accepted March 30, 2018. Date of publication April 12, 2018; date of current version February 1, 2019. This work was supported in part by National Nature Science Foundation (NSFC) under Grants 61473120, 61773169 and U1613214, in part by the Science and Technology Planning Project of Guangzhou under Grant 201607010006, in part by the State Key Laboratory of Robotics and System (HIT) under Grant SKLRS-2017-KF-13, and in part by the Fundamental Research Funds for the Central Universities under Grant 2017ZD057. Paper no. TII-17-2926. *(Corresponding author: Chenguang Yang.)*

C. Yang, C. Zeng, and M. Wang are with the Key Laboratory of Autonomous Systems and Networked Control, College of Automation Science and Engineering, South China University of Technology, Guangzhou 510640, China (e-mail: cyang@ieee.org; mjzengchao@163.com; auwangmin@scut.edu.cn).

Y. Cong is with the State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China (e-mail: congyang@sia.cn).

N. Wang is with Center for Robotics and Neural Systems, Plymouth University, Plymouth PL4 8AA, UK. (e-mail: nwang0902@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TII.2018.2826064

For complex multistep tasks, it is quite difficult to well generalize the learned policies and adopt them to new task situations. One promising way to address this problem is to first segment the learned skill into a set of subskills. Each subskill can be then individually modeled and generalized in order to fulfil the requirements of a specific task, and it can also be easily reused in other similar task situations. However, skill segmentation by hand is faced with a large number of problems such as the lack of flexibility, low efficiency, and heavy workload. Most importantly, to divide a skill into a set of subskills is often difficult since an effective segmentation technique often requires the knowledge of the robots kinematic properties [22].

Recently, the Beta Process Autoregressive HMM (BP-AR-HMM) has been proposed in [23], which can automatically divide multimodel human behaviours into a sequence of features. The work [22], [24] has introduced this algorithm in a PbD framework to sparse the demonstrated trajectories into several sequences. This model has a number of merits, as follows: 1) no *priori* knowledge is in advance specifically required for the selection of the number of the features. This largely facilitates the segmentation process, and makes it possible to potentially build a repository with infinite number of skills, and 2) it is able to discover the same features from multiple demonstrations. This is of great significance because it is often necessary to teach a robot several times for a specific task. In this work, we take one step forward from [24]: the estimated human stiffness profiles from demonstrations will be segmented along with the demonstrated movement trajectories.

To summarize, the main contributions of this work lie in that: 1) we develop a PbD framework with a demonstration–segmentation–alignment–generalization process for human-to-robot skills transfer. Although these aspects have been addressed individually, we integrate them into a unified framework to address these issues in a systematic approach; and 2) the segmentation of both movement trajectories and the EMG-based stiffness profiles are simultaneously taken into account in the proposed framework, thus, making it more convenient and efficient for robot to acquire and generalize human-like adaptive stiffness manipulative skills. To the best of our knowledge, it has never been reported in the literature before.

The rest of this paper is organized: the methodology will be detailed in Section II. Next, the experimental study is presented in Section III, and Section IV finally concludes this paper.

II. Methodology

This section will first provide an overview of the proposed framework. Then, a brief introduction to the basic models integrated into our framework will be presented. We will also show the modifications of the relevant algorithms to make them suitable to our framework.

A. Overview of the Framework

The proposed framework includes four phases: demonstration, segmentation, alignment, and generalization. The overview of the framework is shown in Fig. 1.

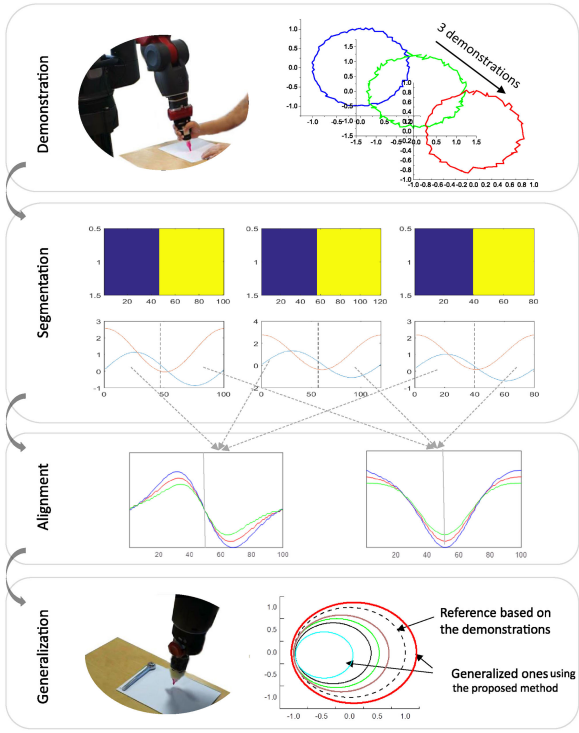


Fig. 1. Graphical representation of the overview of the proposed framework.

Demonstration: Conventionally, a human instructor demonstrates the skills to accomplish one specific task, during which the movement trajectories are recorded for subsequent usage. In our framework, the instructor's arm EMG signals are also extracted for the purpose to extract the stiffness features. In order to capture as many features as possible, several times of demonstrations are often performed for one task. The dual-arm control-based teleoperation method described in [10] is used for demonstration.

Segmentation: The skills represented by movement trajectories together with stiffness profiles obtained from the demonstrations are divided into sequences of subskills. In this way, a repository of features of one specific task is established.

Alignment: The demonstration profiles with different time durations often need to be temporally aligned. Additionally, we also need to align the coordinate points from different segments. To this end, the generalized time warping (GTW) algorithm [25] is utilized in this work.

Generalization: The segmented and aligned sequences of movement trajectories and stiffness profiles are then encoded by DMP. Finally, each set of the movement trajectories as well as the stiffness profiles is determined whether to be generalized or not according to the requirements of the task situations. Apparently, the robot can simply mimic the demonstrations when it is unnecessary to generalize the learned skills.

B. Skill Segmentation Using BP-AR-HMM

Hidden Markov Model (HMM) is a statistical method that is appropriate for time series analysis. It has been applied for

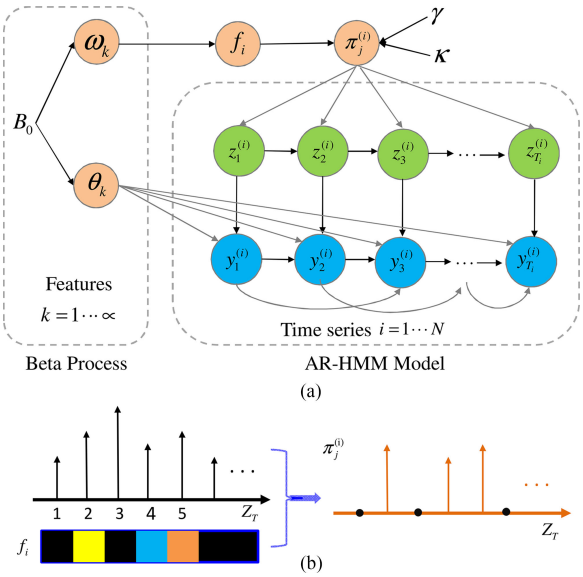


Fig. 2. Graphical representation of the BP-AR-HMM model (a) and the feature selection mechanism (b).

encoding trajectory in robot skill acquisition. An HMM is a dual stochastic process that is characterized by a Markov chain of a sequence of hidden state variables and a corresponding sequence of observation data. Generally, a transition function is defined to describe the probability of each state at time t given the last state. Based on a set of observed data, the parameters of an HMM can be efficiently estimated by using the forward–backward or Viterbi algorithms, and the most likely sequence of states that generated the observed data can also be determined. However, a *priori* knowledge is often needed to chose an implicit number of states, easily leading to the problem of over fitting or low-fitting. This drawback largely limits the utilization of the model, especially when dealing with movement segmentation of a complex task. Additionally, the conventional HMM is not well suitable to address the problem of segmenting complex behaviours according to multiple observed data because the observations are independently considered in the model and the dependency between different series of observations are neglected.

The Beta Process Autoregressive HMM (BP-AR-HMM) can address these two drawbacks with the HMM model as mentioned above. A graphical representation is shown in Fig. 2(a). It mainly contains two parts, i.e., the BP and the AR-HMM model. Here, we briefly introduce the basic idea of the algorithm.

First, a draw B from a BP defines the global weights for the states, which can correspondingly encode a number of different behaviours, see (1) and (2). The feature-inclusion probabilities ω_k and state-specific parameters $\theta_k = \{A_k, \Sigma_k\}$ are generated via this process [23]

$$B|B_0 \sim \mathrm{BP}(1, B_0) \tag{1}$$

$$B = \sum_{k=1}^{\infty} \omega_k \theta_k . \tag{2}$$

Then, a Bernoulli Process (BeP) parameterized by B is performed to generate an X_i for each time series i, as shown in (3). Each X_i is utilized to generate a binary vector $\boldsymbol{f_i} = [f_{i1}, f_{i2}, ...]$ indicating which of the global features are shared in the ith time series. For instance, if the kth element in the vector $\boldsymbol{f_i}$ is 1, i.e., $f_{ik} = 1$, then the ith time series includes the jth feature. Therefore, this enables the possibility of sharing of global features amongst multiple time series, meanwhile, it allows variability for each time series. The graphical representation of the feature selection mechanism is shown in Fig. 2(b).

$$X_i|B \sim \mathrm{BeP}(B) \tag{3}$$

$$X_i = \sum_{k=1} f_i \delta_{\theta_k} . \tag{4}$$

Third, the transition distributions $\pi_j^{(i)}$ is constructed via a Dirichlet distribution with the hyper-parameters γ and κ, given the feature-indicating vector $\boldsymbol{f_i}$ for each time series. Thus, the ith time series can select features from a global library with the feature-constrained transition distributions $\pi_j^{(i)}$.

$$\pi_j^{(i)}|\boldsymbol{f_i}, \gamma, \kappa \sim \mathrm{Dir}([\gamma, ..., \gamma + \kappa, \gamma, ...] \otimes \boldsymbol{f_i}). \tag{5}$$

Fourth, at each time step t the state $z_t^{(i)}$ is generated based on the last time step, as shown in (6).

$$z_t^{(i)} \sim \pi^{(i)}_{z_{t-1}^{(i)}} . \tag{6}$$

Finally, the observations $y_t^{(i)}$ can be obtained using (7), which shows a Vector Autoregression Model (VAR). It demonstrates that given the VAR order r, the observation is calculated as the sum of linear transformations of the previous r observations of the model, plus a state-specific noise [23].

$$y_t^{(i)} = \sum_{j=1}^{r} A_{j,z_t^{(i)}} y_{t-j}^{(i)} + \mathrm{e}_t^{(i)}(z_t^{(i)}). \tag{7}$$

The BP-AR-HMM model is able to provide reliable inference with only a few open parameters. A BP prior is used to determine the total number of the potential features representing a complex task in a fully Bayesian way, without the need of manual intervention as mentioned above. More importantly, it is able to recognize global features among multiple demonstrations, as well as to leave room for variability of each individual time series.

In this work, we will use the model to segment not only the movement trajectories obtained during multiple demonstrations but also the estimated human arm stiffness profiles based on EMG signals which are obtained from the human instructor. To do so, we encapsulate the movement information x and the stiffness information p into the observations: $y_t^{(i)} = \{x_t^{(i)}, p_t^{(i)}\}$ such that they can be segmented in parallel.

C. Demonstration Profiles Alignment Using the Generalized Time Warping Algorithm

The dynamical time warping (DTW) algorithm [26] has been widely used for the alignment of time series data. Considering

two time sequences $U = [u_1, \cdots, u_P]$ and $V = [v_1, \cdots, v_Q]$ of size P and Q, respectively, DTW is the technique to obtain optimal correspondences between U and V such that a given sum-of-square cost error is minimized

$$J_{\text{dtw}} = \|UR_x - VR_y\|^2 \tag{8}$$

where R_x and R_y are binary replication matrices. This process can be seen as finding an alignment between the reference sequence U and the observed subsequence V.

In the context of robot learning skills from demonstrations, the obtained profiles are often multidimensional time series with a large number of data points, in which case DTW might be restrictive due to the quadratic both time and space computational complexity. Additionally, extending DTW for the alignment of multiple demonstration sequences is usually difficult due to the combinatorial explosion of possible warping paths.

To realize multimodal alignment of human motion, the work [25] proposed the GTW algorithm which can address the issues above. Compared to the DTW, GTW has following three advantages.

(1) It uses a Gauss–Newton algorithm with linear complexity in the length of the sequence to optimize the time warping function.
(2) It considers the differences in dimensionality using multiset canonical correlation analysis (mCCA).
(3) it uses a more flexible warping model parametrized by a set of monotonic bases to compensate for changes in time space.

Considering m demonstration profiles, $\{U_1, \cdots, U_m\}$, with $U_i = [u_1^i, \cdots, u_{n_i}^i]$, GTW minimizes the cost function [25]

$$J_{\text{gtw}} = \sum_{i=1}^{m}\sum_{j=1}^{m}\frac{1}{2}\|V_i^T U_i W_i - V_j^T U_j W_j\|^2 + \left(\sum_{i=1}^{m}\varphi(W_i) + \phi(V_i)\right) \tag{9}$$

where W_i and V_i are the nonlinear temporal transformation and the low-dimensional spatial embedding, respectively. $\varphi(\cdot)$ and $\phi(\cdot)$ are regularization functions. Equation (9) suggests that for each U_i, GTW finds a W_i and a V_i such that the sequence $V_i^T U_i W_i$ is well aligned with the others in the least-squares sense. To optimize the cost function is a nonconvex optimization problem with respect to the alignment (i.e., W_i) and projection matrices (i.e., V_i), which are solved by Gauss–Newton algorithm and mCCA, respectively [25].

D. Skill Representation and Generalization Using DMP

1) DMP Model: Typically, a point-to-point movement DMP for a one-dimensional system is represented by three differential equations [16]

$$\tau\dot{v} = K(x_g - x) - Dv + (x_g - x_0)f(s;\boldsymbol{w}) \tag{10}$$

$$\tau\dot{x} = v \tag{11}$$

$$\tau\dot{s} = -\alpha_1 s. \tag{12}$$

Here, we ignore the time variable, i.e., we denote x_t as x. Equation (10) represents a transform system that consists of a linear dynamical system acting like a spring-damping system perturbed by a nonlinear forcing function, i.e., $f(s;\boldsymbol{w})$. K and D are constant stiffness coefficient and constant damping coefficient, respectively. And, D is usually set to be $D = 1/4K$ such that the linear dynamical system is critically damped. x_g and x_0 are the goal and initial point of a trajectory, respectively. τ is a scaling factor that shared by all these equations. x and v denote the position and velocity of the trajectory, respectively, and they are related as shown in (11). Finally, s is the phase variable of the transform system, and it is represented by another differential equation, see (12), and the α_1 is a predefined positive constant.

The nonlinear forcing function $f(s;\boldsymbol{w})$ takes the following representation:

$$f(s;\boldsymbol{w}) = \boldsymbol{w}^T\boldsymbol{g} \tag{13}$$

where $\boldsymbol{g}$ is a time-parameterized kernel vector, and $\boldsymbol{w}$ is policy parameter vector, which affects the shape of the learned trajectory. The element of the kernel vector is defined as

$$[\boldsymbol{g}]_n = \frac{\varphi_n(s)s}{\sum_{n=1}^{N}\varphi_n(s)} \tag{14}$$

with a normalized basis function $\varphi_n(s)$, which is usually defined as a radial basis function kernel

$$\varphi_n(s) = \exp(-h_n(s - c_n)) \tag{15}$$

where c_n and h_n are the centres and widths of these kernel functions, respectively. Generally, c_n are equispaced in time during of the whole trajectory, and h_n are selected based on experience.

The policy parameter $\boldsymbol{w}$ can be efficiently learned using supervise learning algorithms such as weighted linear regression. To put it simply, we need to find a proper parameter vector by minimizing the following error [16]:

$$\min\sum(f_{\text{target}} - f(s))^2 \tag{16}$$

$$f_{\text{target}} = \frac{\tau\dot{v} + Dv - K(x_g - x)}{x_g - x} \tag{17}$$

which can be computed based on a demonstrated trajectory $x_t, \dot{x}_t, \ddot{x}_t$, with $t = 1, \cdots, \text{T}$ and $x_g = x(T)$.

2) Extended DMP for Encoding Stiffness: In order to simultaneously encode both movement trajectories and stiffness profiles obtained from demonstrations, inspired by [16]–[18], we have extended a DMP framework as below [27]

$$\begin{cases} \dot{s} = h(s) & (18)\\ \dot{x} = g_1(x,\omega,s) & (19)\\ \dot{p} = g_2(p,\gamma,s) & (20) \end{cases}$$

where (18) is the Canonical system formally represented by (12). Equations (19) and (20) are two transform systems that one is used for movement trajectories (x) and another for stiffness profiles (p). These two transform systems share the identical form as shown in (10) but with different parameters ω and γ. They are driven by the same phase variable s such that the phase

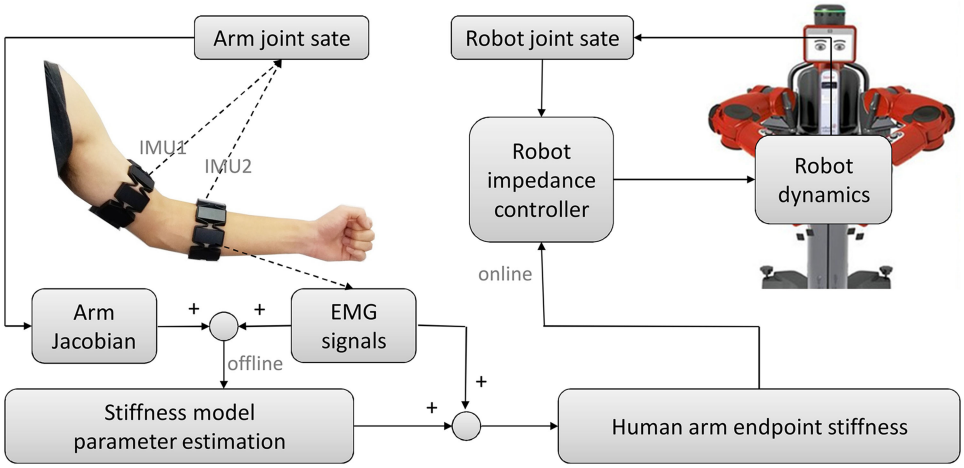


Fig. 3. Mapping of human arm endpoint stiffness to robot joint impedance controller.

synchronization can be strictly guaranteed. The estimation of stiffness parameters γ is very similar to the estimation of ω as described in (13)–(17).

E. Mapping of Human Arm Endpoint Stiffness to Robot Impedance Controller

To capture the variable features of the human arm muscles during the task execution, the instructor's arm EMG signals are monitored in a real-time manner. The human arm stiffness is then estimated based on the EMG signals. Finally, the stiffness is mapped to the robot joint impedance controller. The graphical of this mapping is shown in Fig. 3.

An EMG armband is used for the collection of the raw EMG signals. Then, the EMG signals are processed with the moving average technique and a low-pass filter [10]. There is an inertial measurement unit in the armband. Therefore, by using two armbands the human arm joint angles can be determined: one is worn on the upper arm close to the shoulder joint and the other on the forearm close to the elbow joint, respectively [28].

The human arm triangle model is utilized to describe the arm configuration in this work, based on which the human arm Jacobian is calculated [29]. Subsequently, combining the human arm Jacobian and processed EMG signals the parameters of the human arm joint-to-endpoint stiffness model is estimated offline as introduced in [9]. Then, the human arm endpoint stiffness is mapped to the robot arm impedance controller (see [10]) online during the demonstration phase, in which only one armband is used to collect EMG signals. In this way, the human arm variable stiffness is transferred to the robot, and the robot arm impedance is regulated to overcome external uncertainties in a specific task.

III. Experimental Study

A. Experimental Setup

The humanoid dual-arm Baxter Research Robot (Baxter R) developed by Rethink Robotics is used in our experimental study. Each arm of the robot has seven DoFs (degrees of freedom), i.e., two shoulder joints (S0 and S1), two elbow joints (E0 and E1), and three wrist joints (W0, W1, and W2). Joint stiffness and damping can be modified under torque control mode, which is simplified as PD impedance control law. The commercial EMG armband named MYO is used to collect raw EMG signals for monitoring the human instructor's muscle activities and stiffness estimation. The water-lifting task has been performed in this work. The experimental setup is shown in Fig. 4. During demonstration, the master arm of the robot is directly collected to one of the human instructor's arms through a physical coupling interface. The instructor demonstrates the water-pouring skill and the robot slave arm follows the motion of the master arm because of the virtual spring-damping system attached between these two arms, see [10] for details.

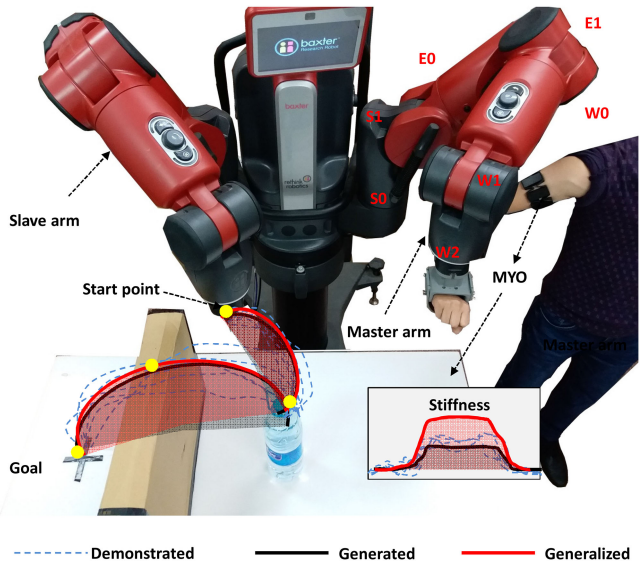


Fig. 4. Experimental setup for the water-lifting task.

The MYO worn on the instructor's upper arm is used for collecting the EMG signals in a real-time manner. The EMG signals are sent to a computer using bluetooth transmission and then processed for the estimation of human arm endpoint stiffness.

Subsequently, the estimated stiffness profile is properly mapped into the robot arm impedance controller using UDP protocol, to generate robot motion control commands. Meanwhile, the robot arm state and stiffness are recorded with a sample rate of 100 Hz.

B. Setting

A BP-AR-HMM MATLAB implementation was first made available by E. Fox to segment sets of motion features collected from multiple demonstrations. Then, Niekum modified it to make it more suitable for the segmentation of movement trajectories in PbD. During the segmentation process, only the demonstrated movement trajectories and stiffness profiles in joint space are given without any other knowledge about the task. In accordance with [23], the obtained demonstration profiles are preprocessed such that the variance of the first differences in each dimension is 1, and the average value is 0. There are total 8 dimensions: 7-DoFs movement trajectories and 1-DoFs stiffness profile. Here, the variability of the stiffness is represented by a stiffness indicator, see [10]. All the demonstrated profiles are also subsampled down to 20 Hz and smoothed as suggested in [22].

We choose the autoregressive order of 1 in AR-HMM model, and adopt the other parameters the same with the parameters chosen by Emily Fox. Considering the principled statistical technique of the BP-AR-HMM model, the Metropolis–Hastings and Gibbs sampler is run 10 times for 1200 iterations each to segment the demonstration profiles, producing 10 segmentations. We chose the one which has the highest log likelihood of the feature settings among the 10 runs and chose the segments as the movement primitives for the follow-up treatment.

C. Results and Analyses

In the demonstration phase, the human instructor demonstrates the water-lifting task six times. Unlike [22], [30], in this work the start point and the goal are both fixed in all demonstrations such that the movement trajectories are not largely different in shape, this is because we are only interesting in considering the stiffness variability. To this end, the robot is taught to perform the task under following two conditions: 1) to lift *half bottle of water* four times, and 2) to lift *one third bottle of water* twice. It is understandable that the six demonstrations have different lengths in time space. Additionally, the stiffness profiles are not exactly the same with each other in shape even under the same task condition due to the uncertainty of the human instructor's muscle activation.

The result of the segmentation is shown in Fig. 5. The top row shows segmenting the demonstration profiles into subskills marked with different coloured bars, and the bottom row shows the corresponding divisions overlaid on a subplot of each of the 8-DoFs demonstration profiles. The first, third, fifth, and sixth subplots correspond to the first task condition, while the second and the fourth subplots correspond to the second task condition. It can be seen that the BP-AR-HMM can successfully segment the water-lifting skill into four sequences basically corresponding to the multiple steps of the task: reach the bottle; pick up the handle of the bottle; get close to the obstacle; and lift the bottle to pass over the obstacle and place it to the goal. The first and third phases are considered as the same subskill (coloured with light-green). This is mainly because these two steps of the task have similar characteristics, only differing in reaching to different objectives. The BP-AR-HMM is able to consistently recognize the repeated subskills across the multiple demonstrations, even though they occur at different positions and with different stiffness. The segments are formally identical with minor defects in the first and the fourth demonstrations that might be caused by the sudden changes of movement trajectories. These minor defects can be directly neglected since they do not affect much the performance of the segmentations.

Fig. 6 shows the result of the alignment of the demonstration profiles. The GTW can align each dimensional profiles all at once, thus, resulting in a high alignment efficiency. The gray lines and the red lines denote the aligned profiles and the generated profiles, respectively. The aligned profiles are averaged and then modelled with the DMP model. One may also utilize some statistical approaches to encode the multiple profiles (see, e.g., [31]). It is seen that all the demonstration profiles, the 7-DoFs movement trajectories and the 1-DoF stiffness are well aligned in time space. The visual inspection of the lines in Fig. 6 suggests that the generated ones can capture the features in each dimension across the six demonstrations. Additionally, GTW is also able to align all lengths of the profiles to a particular value, which can facilitate task plan.

Subsequently, we examine the task generalization to the following task situations: i) to lift a bottle *full of water* to pass over the obstacle and place it to the goal; and ii) *to keep low stiffness control as demonstrated* when high stiffness is undesired, namely, the robot is required to be under low-gain control during the first two steps. In this case, it is usually unlikely to accomplish the task meeting these two requirements simultaneously, while it can be handed with the proposed framework.

Once the segments of the stiffness profile are obtained, each sequence can be efficiently adapted using DMP model. To deal with the new task situation, the third and the fourth segments of the stiffness are planned by generalizing the goal of the third sequence, i.e., the initial point of the fourth sequence to a certain value. Fig. 7(a) shows three generalized stiffness profiles (red lines) based on the reference one (dark line), namely, the generated profile from the last phase. Then, through experience we select the stiffness profile with a proper goal for the task described above. The position profiles shown in Fig. 6 and their first derivative profiles (velocities) are utilized as the reference profiles in the variable impedance controller.

The new task is performed successfully with the generalized stiffness profile. The collected force profiles of the robot endpoint in z-axis are shown in Fig. 7(b). The gray lines denote the force profiles measured from the six demonstrations, and the blue line represents one typical force profile obtained form the generalization phase. For the better visualization, all the force lines are aligned in the time coordinate. The force profiles and the stiffness profiles have a very similar shape, it means

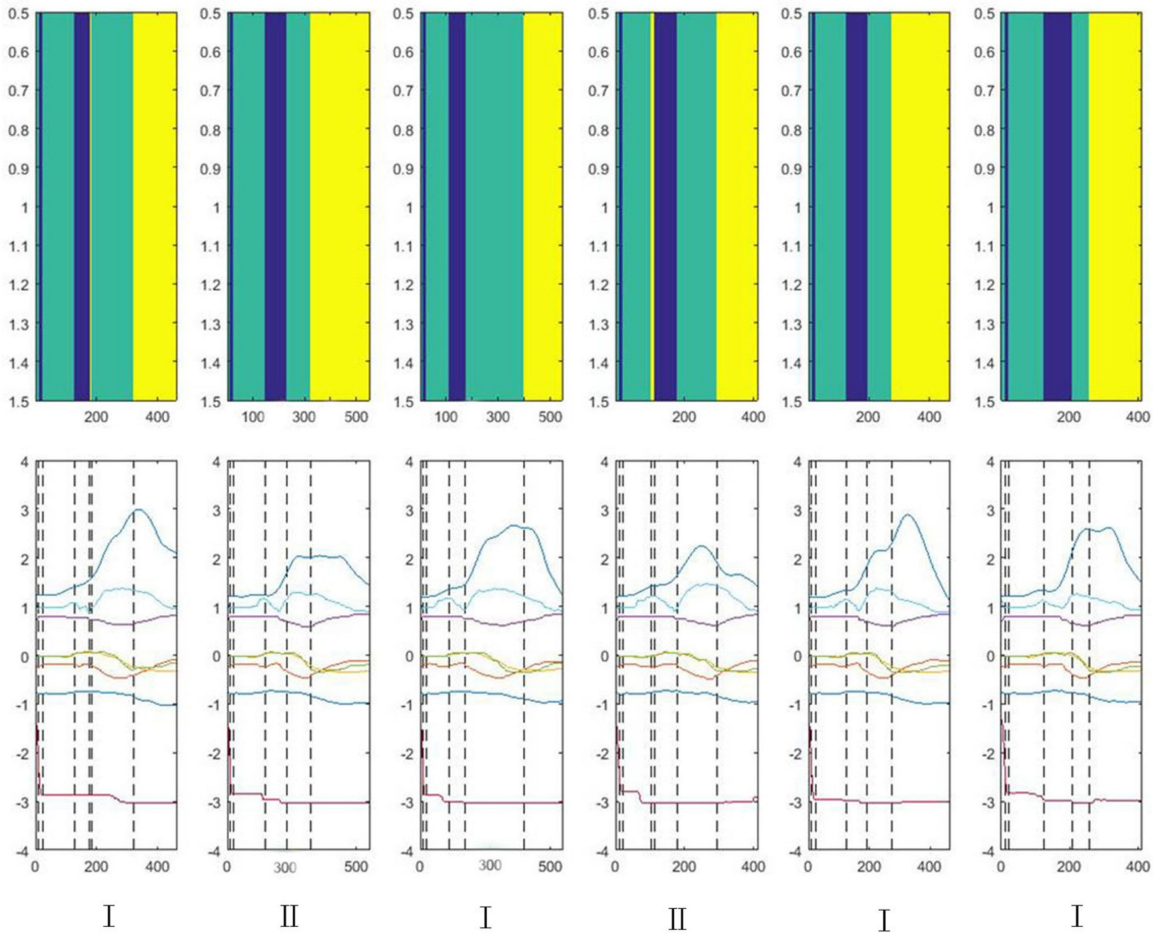


Fig. 5. Result of segmentation of the 7-DfoFs movement trajectories and the 1-DoF stiffness profiles. The first line in each of the six subplots denotes the stiffness, and the others represent the joint angles. I and II correspond to the first and the second task conditions, respectively.

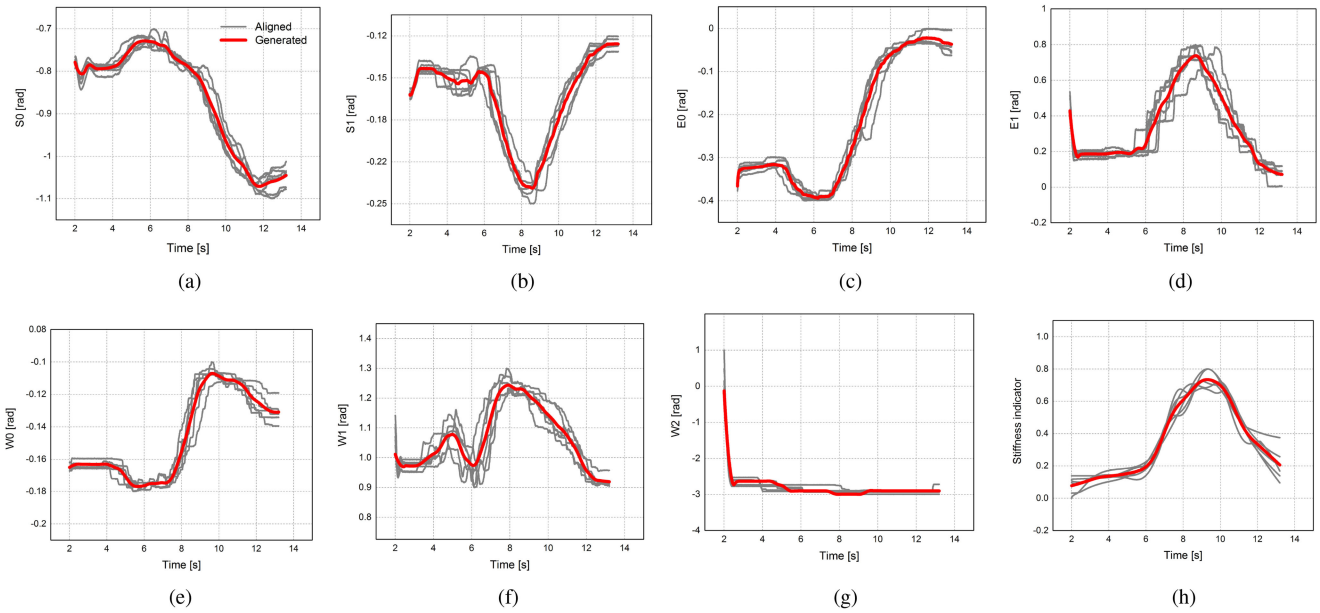


Fig. 6. Result of the alignment of the 7-DoFs movement trajectories (a–g) and the 1-DoF stiffness profile (h). (a) Angle of joint S0. (b) Angle of joint S1. (c) Angle of joint E0. (d) Angle of joint E1. (e) Angle of joint W0. (f) Angle of joint W1. (g) Angle of joint W2. (h) Stiffness indicator.

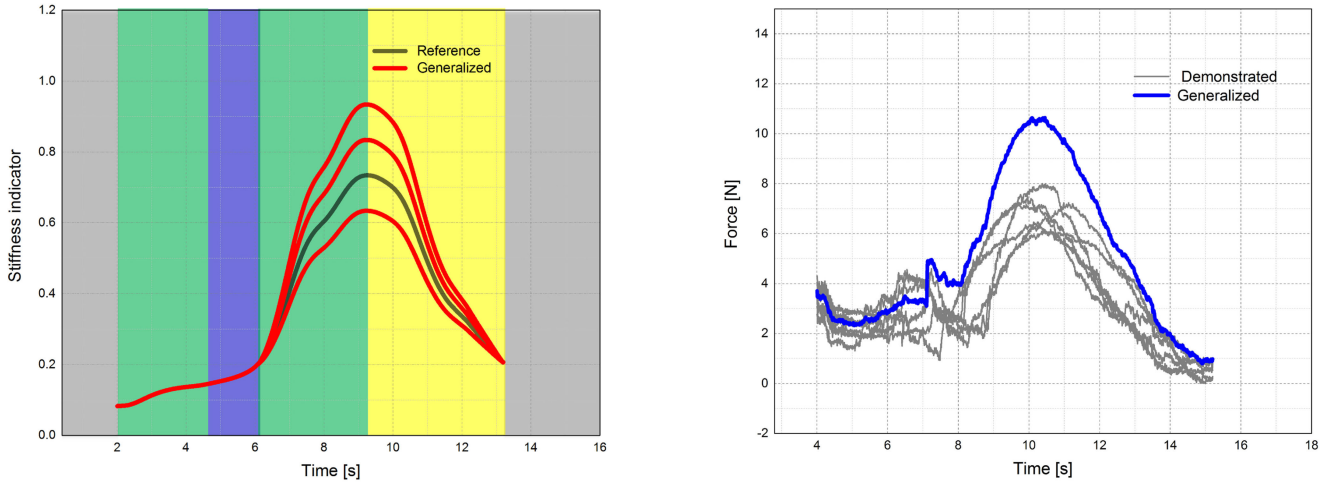


Fig. 7. (a) Planned stiffness profiles and (b) measured force profiles in z-axis from the six demonstrations (gray lines) and the generalization phase (blue line).

the human-to-robot stiffness transfer has been achieved successfully. More importantly, the force during the generalization phase shows the behaviour as expected: keeping low in the first and the second steps and then increasing when the robot lifts the bottle full of water, which can demonstrate the effectiveness of the proposed method.

IV. Conclusion

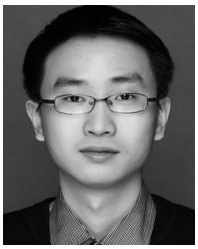
In this work, we proposed a novel PbD framework for the robot learning and generalizing human-like variable impedance skills that integrates several models for encoding, segmenting, aligning, and generalizing the positional trajectories as well as the EMG-based stiffness profile. By integrating BP-AR-HMM, GTW, and DMP into the framework, it allows the potential of our framework to learn a large library of skills including an infinite number of both movement primitives and stiffness primitives, and it also enables to efficiently regulate each individual segment of the demonstrated profiles that can greatly facilitate the generalization of the learned skills. Additionally, by considering the EMG-based stiffness estimation of the instructor, our framework allows robots to learn variable impedance skills from humans and to generalize the human-like skills to adapt to novel task situations. The experimental study shows that our method is able to learn and generalize a multistep task on the Baxter robot manipulator. Our framework can be applied to other robotic platforms thanks to the model-free models.

This work makes a preliminary investigation toward the highly effective PbD system, leaving a large number of directions open for the future work as follows: 1) a more effective and efficient stiffness estimator will be essential for robots to learn highly human-like features from humans; and 2) some reinforcement learning techniques can be integrated into the framework for the improvement of the DMP policy. In addition, the policy improvement can be effectively achieved by deriving a proper reward/cost function only for the key subskills instead of the whole trajectories.

References

[1] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, "Robot programming by demonstration," in *Springer Handbook Robot*. Berlin, Germany: Springer-Verlag, 2008, pp. 1371–1394.
[2] Y. Pan and H. Yu, "Composite learning from adaptive dynamic surface control," *IEEE Trans. Autom. Control*, vol. 61, no. 9, pp. 2603–2609, Sep. 2016.
[3] N. Sun, Y. Wu, Y. Fang, and H. Chen, "Nonlinear antiswing control for crane systems with double-pendulum swing effects and uncertain parameters: Design and experiments," *IEEE Trans. Autom. Sci. Eng.*, 2017, in press, doi: 10.1109/TASE.2017.2723539.
[4] Y. Pan, T. Sun, and H. Yu, "Composite adaptive dynamic surface control using online recorded data," *Int. J. Robust Nonlinear Control*, vol. 26, no. 18, pp. 3921–3936, 2016.
[5] 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 Trans. Robot.*, vol. 27, no. 5, pp. 918–930, Oct. 2011.
[6] F. Ficuciello, L. Villani, and B. Siciliano, "Variable impedance control of redundant manipulators for intuitive human–robot physical interaction," *IEEE Trans. Robot.*, vol. 31, no. 4, pp. 850–863, Aug. 2015.
[7] Z. Li, Z. Huang, W. He, and C.-Y. Su, "Adaptive impedance control for an upper limb robotic exoskeleton using biological signals," *IEEE Trans. Ind. Electron.*, vol. 64, no. 2, pp. 1664–1674, Feb. 2017.
[8] Z. Ju, G. Ouyang, M. Wilamowska-Korsak, and H. Liu, "Surface EMG based hand manipulation identification via nonlinear feature extraction and classification," *IEEE Sens. J.*, vol. 13, no. 9, pp. 3302–3311, Sep. 2013.
[9] A. Ajoudani, N. Tsagarakis, and A. Bicchi, "Tele-impedance: Teleoperation with impedance regulation using a body–machine interface," *The Int. J. Robot. Res.*, vol. 31, no. 13, pp. 1642–1656, 2012.
[10] C. Yang, C. Zeng, P. Liang, Z. Li, R. Li, and C.-Y. Su, "Interface design of a physical human–robot interaction system for human impedance adaptive skill transfer," *IEEE Trans. Autom. Sci. Eng.*, vol. 15, no. 1, pp. 329–340, Jan. 2018.
[11] M. Racca, J. Pajarinen, A. Montebelli, and V. Kyrki, "Learning in-contact control strategies from demonstration," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst, (IROS)*, Oct. 2016, pp. 688–695.
[12] A. I. Ahmed, H. Cheng, H. Liu, X. Lin, and M. J. Atieno, "Interaction force convex reduction for smooth gait transitions on human-power augmentation lower exoskeletons," in *Proc. Int. Conf. Cogn. Syst. Signal Proc.* Springer-Verlag, 2016, pp. 398–407.
[13] A. Cirillo, F. Ficuciello, C. Natale, S. Pirozzi, and L. Villani, "A conformable force/tactile skin for physical human–robot interaction," *IEEE Robot. Autom. Lett.*, vol. 1, no. 1, pp. 41–48, Jan. 2016.
[14] H. Liu, F. Sun, B. Fang, and F. Long, "Material identification using tactile perception: A semantics-regularized dictionary learning method," *IEEE/ASME Trans. Mechatronics*, 2017.

[15] P. Liang, C. Yang, Z. Li, and R. Li, "Writing skills transfer from human to robot using stiffness extracted from sEMG," in *Proc. IEEE Int. Conf. Cyber Technol. Autom., Control, Intell. Syst, (CYBER)*, Jun. 2015, pp. 19–24.
[16] A. J. Ijspeert, J. Nakanishi, and S. Schaal, "Movement imitation with nonlinear dynamical systems in humanoid robots," in *Proc. IEEE Int. Conf. Robot. Autom., (ICRA)*, May 2002, pp. 1398–1403.
[17] J. Kober, B. Mohler, and J. Peters, "Learning perceptual coupling for motor primitives," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2008, pp. 834–839.
[18] K. Muelling, J. Kober, and J. Peters, "Learning table tennis with a mixture of motor primitives," in *Proc. 10th IEEE-RAS Int. Conf. Humanoid Robots*, Dec. 2010, pp. 411–416.
[19] 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 Trans. Mechatronics*, vol. 23, no. 1, pp. 121–131, Feb. 2018.
[20] P. Kormushev, S. Calinon, and D. G. Caldwell, "Imitation learning of positional and force skills demonstrated via kinesthetic teaching and haptic input," *Adv. Robot.*, vol. 25, no. 5, pp. 581–603, 2011.
[21] F. Steinmetz, A. Montebelli, and V. Kyrki, "Simultaneous kinesthetic teaching of positional and force requirements for sequential in-contact tasks," in *Proc. IEEE-RAS 15th Int. Conf. Humanoid Robots (Humanoids)*, Nov. 2015, pp. 202–209.
[22] S. Niekum, S. Osentoski, G. Konidaris, and A. G. Barto, "Learning and generalization of complex tasks from unstructured demonstrations," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2012, pp. 5239–5246.
[23] E. B. Fox *et al.*, "Joint modeling of multiple time series via the beta process with application to motion capture segmentation," *The Ann. Appl. Statist.*, vol. 8, no. 3, pp. 1281–1313, 2014.
[24] S. Niekum, S. Osentoski, G. Konidaris, S. Chitta, B. Marthi, and A. G. Barto, "Learning grounded finite-state representations from unstructured demonstrations," *Int. J. Robot. Res.*, vol. 34, no. 2, pp. 131–157, 2015.
[25] F. Zhou and F. D. la Torre, "Generalized time warping for multi-modal alignment of human motion," in *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, Jun. 2012, pp. 1282–1289.
[26] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. 26, no. 1, pp. 43–49, Feb. 1978.
[27] C. Yang, C. Zeng, C. Fang, W. He, and Z. Li, "A DMPs-based framework for robot learning and generalization of human-like variable impedance skills," *IEEE/ASME Trans. Mechatronics*, to be published, doi: 10.1109/TMECH.2018.2817589.
[28] C. Yang, P. Liang, A. Ajoudani, Z. Li, and A. Bicchi, "Development of a robotic teaching interface for human to human skill transfer," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst., (IROS)*, Oct. 2016, pp. 710–716.
[29] C. Fang and X. Ding, "A set of basic movement primitives for anthropomorphic arms," in *Proc. IEEE Int. Conf. Mechatronics Autom.*, Aug. 2013, pp. 639–644.
[30] M. Chi, Y. Yao, Y. Liu, Y. Teng, and M. Zhong, "Learning motion primitives from demonstration," *Advances Mech. Eng.*, vol. 9, no. 12, 1687814017737260, 2017.
[31] S. Calinon, "A tutorial on task-parameterized movement learning and retrieval," *Intell. Service Robot.*, vol. 9, no. 1, pp. 1–29, 2016.



Chenguang Yang (M'10–SM'16) received the B.Eng. degree in measurement and control from Northwestern Polytechnical University, Xi'an, China, in 2005, and the Ph.D. degree in control engineering from the National University of Singapore, Singapore, in 2010.

He received the Postdoctoral training at Imperial College London, London, U.K. His research interests include robotics and automation.

Dr. Yang was the recipient of the Best Paper Award from the IEEE TRANSACTIONS ON ROBOTICS and a number of international conferences.

Chao Zeng received the M.S. degree in precision instrument and mechanics from Shanghai University, Shanghai, China, in 2016. He is currently working toward the Ph.D. degree in pattern recognition and intelligent system at the College of Automation Science and Engineering, South China University of Technology, Guangzhou, China.

His research interests include human robot interaction, programming by demonstration, and human robot skill transfer.

Yang Cong (M'11–SM'15) received the B.Sc. degree from Northeast University, Boston, MA, USA, in 2004, and the Ph.D. degree from the State Key Laboratory of Robotics, Chinese Academy of Sciences, Shenyang, China, in 2009.

He is currently a Full Professor with the Chinese Academy of Sciences, Shenyang, China. From 2009 to 2011, he was a Research Fellow with the National University of Singapore and Nanyang Technological University, Singapore, respectively; and a Visiting Scholar with the University of Rochester, NY, USA. He has authored more than 50 technical papers. His research interests include image processing, compute vision, machine learning, multimedia, medical imaging, data mining, and robot navigation.

Dr. Cong has served on the Editorial Board of the *Journal of Multimedia*.

Ning Wang (M'10) received the B.Eng. degree in measurement and control technologies and devices from the College of Automation, Northwestern Polytechnical University, Xi'an, China, in 2005, the M.Phil. and Ph.D. degrees in electronic engineering from the Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong, China, in 2007 and 2011, respectively.

From 2011 to 2013, she was a Postdoctoral Fellow at the Department of Computer Science and Engineering, The Chinese University of Hong Kong, Hong Kong, China. Her research interests include signal processing and machine learning, with applications in robust speaker recognition, biomedical pattern recognition, intelligent data analysis, and human–robot interaction.

Min Wang (M'09) received the B.Sc. and M.Sc. degrees from the Department of Mathematics, Bohai University, Jinzhou, China, in 2003 and 2006, respectively; the Ph.D. degree from the Institute of Complexity Science, Qingdao University, Qingdao, China, in 2009.

From November 2017 to present, she is an Academic Visitor with the Department of Information Systems and Computing, Brunel University London, Uxbridge, U.K. She is currently an Associate Professor with the School of Automation Science and Engineering, South China University of Technology, Guangzhou, China. She has authored or co-authored nearly 30 papers in top international journals. She is a very active Reviewer for many international journals.

Dr. Wang was a recipient of the Excellent Doctoral Dissertations Award of Shandong Province in 2010, the Science and Technology New Star of Zhujiang, Guangzhou, in 2014, and the youth talent of Guangdong Tezhi Plan in 2016. Her research interests include nonlinear systems, intelligent control, robot control, and dynamic learning.