2016 IEEE International Conference on Robotics and Automation (ICRA)
Stockholm, Sweden, May 16-21, 2016

Hierarchical Interactive Learning for a Human-Powered Augmentation Lower Exoskeleton

Rui Huang[1], Hong Cheng[1], *Senior Member, IEEE*, Hongliang Guo[2], Qiming Chen[1], Xichuan Lin[1]

Abstract— **Learning by demonstration methods have gained considerable interest in human-coupled robot control. It aims at modeling the goal motion trajectories through human demonstration. However, in lower exoskeleton control, the physical human-robot interaction is changing from pilot to pilot or even for one pilot in different walking patterns. This characteristic requires that the exoskeletons should have the ability to learn and adapt the motion trajectories as well as controllers online. This paper presents a novel Hierarchical Interactive Learning (HIL) strategy which reduces the complexity of the exoskeleton sensory system and is able to handle varying interaction dynamics. The proposed HIL strategy is composed of two learning hierarchies, namely, high-level motion learning and low-level controller learning. The Dynamic Movement Primitives (DMPs) combined with Locally Weighted Regression (LWR) are employed to model and learn the motion trajectories, while reinforcement learning (RL) is used to learn the model-based controller. We demonstrate the efficiency of proposed HIL strategy on a single degree-of-freedom (DOF) platform as well as a HUman-powered Augmentation Lower EXoskeleton (HUALEX) system. Experimental results indicate that the proposed HIL strategy is able to handle the varying interaction dynamics with less interaction force between the pilot and the exoskeleton when compared to traditional model-based control algorithms.**

Index Terms - Interactive Learning, Reinforcement Learning, Physical Human-Robot Interaction, Dynamic Movement Primitive (DMP), Lower Exoskeleton, HUman-powered Augmentation Lower EXoskeleton (HUALEX).

I. INTRODUCTION

Many lower extremity exoskeletons have been developed as load carriers for human augmentation related applications [1], [2], [3]. These kinds of exoskeletons should be able to track the pilot's motion with little interaction force between the exoskeletons and the pilot. The design philosophy of the exoskeleton controllers can be categorized into two types: one is sensor-based controller and the other is model-based controller.

Sensor-based controllers take as input the (real time) measurement of sensors and apply many variations of control strategies such an impedance control. For example, Y. Sankai *et al.* apply impedance control method for human enhancement based on Eletro-Myo-Graphical (EMG) sensors [4]. The method has been implemented on the Hybrid Assistive Limb (HAL) system [5]. In the experiment, EMG sensors are utilized to calculate the reference patterns of the pilot through the direct measurement of human-exoskeleton interactions. Furthermore, active-impedance control [6] and fuzzy-based impedance control [7] are proposed to adapt to the changing interaction dynamics based on acceleration sensors. Sensor-based controllers are generally model free, however, they rely heavily on complex sensor systems, which, due to cost issues, prohibits its extensive application to exoskeletons.

On the other hand, model-based controllers, which are able to simplify the sensory system of exoskeletons, are employed in exoskeletons for human augmentation applications. Sensitivity Amplification Control (SAC) algorithm is one of model-based controllers which is proposed in the Berkeley Lower Extremity Exoskeleton (BLEEX) [8]. Through an indirect augmentation control algorithm, SAC reduces the interaction force between the pilot and exoskeleton without measuring it directly. Without the need of the interaction force between the pilot and the exoskeleton, SAC reduces the complexity of sensory system. However, it needs a complicated system identification process [9] to obtain the accurate dynamic models, hence it is quite sensitive (not robust) to model imperfections.

This paper aims at bridging the research gap between sensor-based controllers and model-based controllers. We propose a Hierarchical Interactive Learning (HIL) strategy which inherits both advantages of sensor-based and model-based controllers. On the one hand, HIL does not rely on complex sensory systems to operate, on the other hand, the model that is used in HIL can be adapted online, and hence does not need the complicated identification process.

HIL is composed of two learning hierarchies, namely, high-level motion learning and low-level controller learning. For high-level motion learning, Dynamic Movement Primitive (DMP) [14], [15] is utilized to model the motion trajectories. In order to update the DMP incrementally, we utilize Locally Weighted Regression (LWR) [10], [11] with demonstrated motion trajectories via physical human-exoskeleton interaction. For low-level controller learning, reinforcement learning [12], [13] is employed to let the exoskeleton behave according to the reshaped motion trajectories. In essence, HIL is a hierarchical *model-based* controller, but the model itself can be learned online through learning hierarchies.

The contributions of this paper can be summarized as follows.

1) We bring about a hierarchical model-based controller (HIL), whose model part (motion planning) can be adapted online;
2) Different from canonical model-based controller [8], which presets a fixed sensitivity amplification factor,

[1]Rui Huang, Hong Cheng, Qiming Chen, Xichuan Lin are with Center for Robotics, School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu, China ruihuang2010@gmail.com

[2]Hongliang Guo is with School of Computer Engineering, Nanyang Technological University, Singapore

we utilize reinforcement learning as the model-based controller to dynamically learn/optimize the sensitivity amplification factor;

3) Incremental DMP is used to enable exoskeleton controllers to adapt to varying interaction dynamics from different pilots or even one pilot during different gait cycles;

The proposed HIL strategy is first validated on a single degree-of-freedom (DOF) exoskeleton platform, and then applied on a human-powered augmentation lower exoskeleton(HUALEX) system. Experimental results indicate that the proposed HIL strategy has a good performance in dealing with variable interaction dynamics from different pilots or even the same pilot in different gait patterns.

The structure of this paper is organized as follows. We first introduce the HIL framework in Section II, and then describe the methodology details in Section III. Experimental results on both the single DOF platform and the HUALEX system are presented and analyzed in Section IV. This paper ends with conclusion and future works in Section V.

II. Hierarchical Interactive Learning with Physical Human-Robot Interaction

Physical Human-Robot Interaction (pHRI) plays a significant role in human-coupled robots, since robots can learn and imitate human motion from interactions caused by human [21], [16], [22]. To imitate the demonstrated motion from human accurately, motion trajectory modeling and adaptation have become critical issues in human-coupled robot's control [15], [17], [19], [20]. In human augmentation applications of lower exoskeletons, the demonstration process happens naturally as the pilot moves the exoskeleton forcibly whenever the pilot wants to change his/her motion patterns. However, the pHRI is changing from pilot to pilot even for one pilot in different walking patterns. The property requires the controllers of exoskeletons to be able to handle changing dynamics online. Therefore, the motion trajectories and controllers should be learned simultaneously to adapt to the variation of interaction dynamics.

In order to learn the motion trajectories and controllers simultaneously, we propose a Hierarchical Interactive Learning (HIL) strategy, as shown in Fig. 1. The proposed HIL strategy comprises of two learning hierarchies: 1) high-level motion trajectory learning; 2) low-level controller learning.

In the motion trajectory learning hierarchy, pHRI from the pilot is utilized to demonstrate the exoskeleton online. If the pilot wants to move with different walking patterns, he/she will correct the exoskeleton's motion forcibly in the next gait cycle. Since the pilot's movement is rhythmic in most cases, the expected motion trajectory can be modeled through DMPs [15]. In order to incrementally learn the motion trajectory online, an incremental learning module with Locally Weighted Regression (LWR) is applied to update DMPs with historical information. The reshaped motion trajectory is regarded as the goal trajectory in the controllers learning hierarchy.

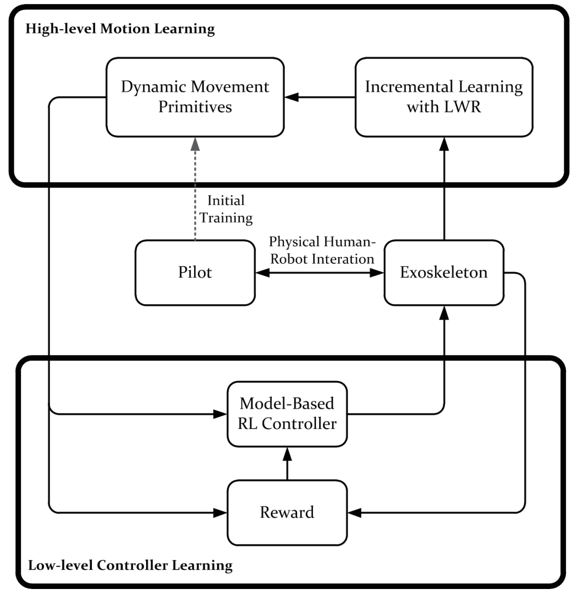


Fig. 1. The proposed hierarchical interactive learning strategy for lower exoskeleton control.

In the controller's learning hierarchy, we adopt model-based controller with sensitivity factors. Reinforcement Learning (RL) is utilized to learn the sensitivity factors online. The methodology details are introduced in Section III.

III. Methodologies of the Learning Hierarchies

This section presents the methodology details of the two learning hierarchies. Section III-A lays down the basic concepts of DMP and LWR, then we present how we merge incremental LWR into DMP so as to have the online adaptive feature of motion learning. Section III-B first introduces the basic dynamics of single DOF exoskeletons, then we present how to map the model parameters to Markov Decision Process (MDP) elements, i.e, state, action, reward and then apply Q-learning to learn the sensitivity amplification factor.

A. Learning Motion Trajectories with Dynamic Movement Primitives

DMP has gained considerable interest in robotic applications in the area of learning by human demonstration [18]. In this paper, we use a limit-cycle oscillator DMP [15] to describe the motion trajectory as shown in Eq. (1):

$$\begin{cases} \tau\dot{z} = -\dfrac{\mu}{E_0}(E-E_0)z - k^2u \\ \tau\dot{u} = z, \ \text{where } E = (z^2+k^2u^2)/2 \end{cases} \tag{1}$$

where μ, k, E_0 are positive parameters, and τ determines the frequency of the oscillator. Based on the limit cycle oscillator, the desired motion trajectory θ is calculated via Eq. (2):

$$\tau\dot{\theta} = \beta(\theta_m - \theta) + f, \tag{2}$$

where f is a nonlinear function as defined in Eq. (3):

$$f = \frac{\sum_{i=1}^{N}\Psi_i\omega_i^T\widetilde{\mathbf{z}}}{\sum_{i=1}^{N}\Psi_i}, \ \widetilde{\mathbf{z}} = [z, \sqrt{E_0}]^T, \tag{3}$$

TABLE I

INCREMENTAL LEARNING PROCESS USING LOCALLY WEIGHTED REGRESSION.

1	Obtain reshaped motion trajectories from pilot's demonstration
2	Update τ based on the time of current gait cycle
3	Initialize ω_i with historic DMPs
4	Calculate f_g based on the reshaped trajectories according Eq. (5)
5	**For** all $i \in [1,N]$ (N is number of Gaussian kernels)
6	**Repeat**
7	$\omega_i \leftarrow \omega_i + P_i \widetilde{\mathbf{z}} e_i$
8	$P_i \leftarrow \frac{1}{\lambda}(P_i - \frac{P_i \widetilde{\mathbf{z}}\widetilde{\mathbf{z}}^T P_i}{\frac{\lambda}{\Psi_i} + \widetilde{\mathbf{z}}^T P_i \widetilde{\mathbf{z}}})$
9	**Until** $e_i = \mid f_g - \omega_i \widetilde{\mathbf{z}} \mid < \varepsilon$ (ε is a small positive number)
10	Calculate reshaped DMPs according Eq. (2)

Ψ_i is defined in Eq. (4):

$$\Psi_i = \exp(-0.5 h_i (\phi - c_i)^2) \tag{4}$$

and β is a positive constant, θ_m determines the baseline of desired position θ. The basis function is defined as Gaussian kernel functions with variable $\phi = \text{atan2}(z, ku)$. The parameters h_i and c_i represent the location and width of the bases, respectively. Hence, for any given motion trajectory, the parameters ω_i can be optimized to achieve good approximation.

Locally Weighted Regression (LWR) method [23] is used to learn the parameters ω_i incrementally. The goal trajectory (f_g) of nonlinear function f is calculated based on the desired motion trajectory θ_d as in Eq. (5):

$$f_g = \tau \dot{\theta}_d - \beta(\theta_m - \theta_d). \tag{5}$$

The incremental learning process is presented in Table I. Before the start of process, the original DMPs are trained offline using normal walking trajectories and the trained parameters are set as the initial values the first gait cycle.

Subsequent DMPs can be learned incrementally through the pilot's compulsive demonstration. The parameter τ is updated based on the time of current gait cycle, which will normalize the original DMPs in current gait cycle.

In order to accelerate the learning process, the parameters ω_i are initialized with the historical DMPs from last gait cycle. Afterwards, the LWR process runs until the nonlinear function f converges to goal trajectories f_g. Finally, the reshaped DMPs can be calculated based on the approximated trajectory f_g according Eq. (2).

B. Reinforcement Learning for Controller Adaptation

As a human-powered robot, the exoskeleton should have the ability to deal with changing interaction dynamics from different pilots or even one pilot over time. Before introducing the online reinforcement learning controller, we first spend three subsections to introduce the inherent dynamics of the lower exoskeleton, how a pilot influences the exoskeleton's dynamics and how sensitivity amplifier influences the exoskeleton's dynamics, respectively. Note that in this section, the reinforcement learning controller is introduced in a single DOF exoskeleton settings, as for the HUALEX system, RL can be easily applied once the mapping from model parameters to the MDP elements (state, action, reward) is done (the mapping process is quite similar to the single DOF case).

1) Inherent Single DOF Lower Exoskeleton Dynamics: Due to the compliant connections at lower links, the pilot is able to make contact with the exoskeleton flexibly and impose forces onto the exoskeleton leg. The equivalent torque on the exoskeleton leg, which results from the pilot's applied force, can be represented as τ_h. The dynamics of the one DOF exoskeleton including the pilot dynamics can then be represented as:

$$J\ddot{\theta} + B\ddot{\theta} + mgl \cdot sin\theta = \tau_e + \tau_h \tag{6}$$

where J,B,m and l represent inertial moment, viscous friction coefficient, exoskeleton shank mass and the length of the one DOF exoskeleton, respectively. $(\theta, \dot{\theta}, \ddot{\theta})$ are the joint states, representing the angle, angular velocity, and angular acceleration of the knee joint. g refers to the gravitational constant.

2) Impedance Model (Human Interactions with the Lower Exoskeleton): When the pilot's leg move with the exoskeleton leg, the interaction torque τ_h is directly applied to the exoskeleton leg. During the analysis of controller stability and simulation, the interaction dynamics can be modeled by an impedance model:

$$\tau_h = K_s(\theta - \theta_h) + K_d(\dot{\theta} - \dot{\theta}_h) \tag{7}$$

where K_s and K_d are positive quantities, θ_h and $\dot{\theta}_h$ are the pilot's angle and angular velocity, respectively.

3) Model-based Controller with Sensitivity Factors: The adaptive sensitivity amplification controller is defined as:

$$\tau_e = mgl \cdot \sin\theta + (1 - \alpha^{-1})(\hat{J}\ddot{\theta} + \hat{B}\dot{\theta}) \tag{8}$$

where $\hat{J}$ and $\hat{B}$ are the estimated inertial moment and viscous friction coefficient of the exoskeleton leg, respectively. α is the sensitivity amplifier factor.

4) Adaptive Reinforcement Learning Controller: As the targeted trajectory has been computed through online DMP in the previous section, the current question is how to design an adaptive online controller which enables the lower exoskeleton robot to track the targeted trajectory and in the meanwhile copes with real time human interventions. Since reinforcement learning is adaptive to real time input signals, we propose an online adaptive reinforcement learning controller. The basic elements of the RL controller are defined as follows:

- State: S is a set of all possible joint angle, angular velocity, angular acceleration of the exoskeleton leg. We identify the model at 100Hz, so that the time interval of s_t and s_{t+1} is 0.01 seconds.
- Action: A is the selection range of the sensitivity factor α as is used in Eq. (8).
- Reward: The reward function R is defined as $R = -[\beta_1(\theta - \theta^*)^2 + \beta_2(\dot{\theta} - \dot{\theta}^*)^2]$,where β_1 and β_2 are weighted parameters.

The reinforcement learning process is as follows:

$$Q_{t+1}(s_t,a_t) = (1-\eta_t)Q_t(s_t,a_t) + \eta_t\left(R_{t+1} + \gamma \max_a Q_t(s_{t+1},a)\right) \tag{9}$$

where η_t is the learning rate, here we define it as $\eta_t = \frac{1}{\sqrt{t}}$. Through iteratively updating the Q-values in Eq. (9), the low-level controller is learned online.

IV. Experiments and Discussions

In this section, we validate the proposed HIL strategy through experiments on both the single DOF platform and the HUman-powered Augmentation Lower EXoskeleton (HUALEX) system. The next two subsections (IV-A and IV-B) will detail the experiment setup, results and corresponding analysis for the Single DOF platform and HUALEX respectively.

A. Numerical Simulation in Single DOF exoskeleton

1) Model of Single DOF exoskeleton: The schematic diagram of a single DOF exoskeleton is shown in Fig. 2. In single DOF exoskeleton, the pilot's leg is attached with the single DOF exoskeleton, which is flexible *only* in swing movement.

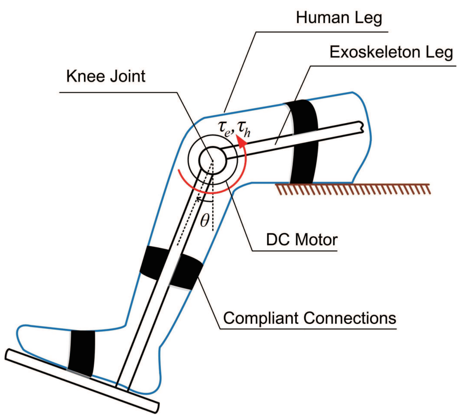


Fig. 2. The schematic diagram of a single DOF exoskeleton interacting with the pilot's leg.

The exoskeleton is powered by a DC motor, which provides τ_e at the knee joint, and the internal dynamics of the motor itself will be ignored since we focus on the controller of the exoskeleton in this paper. Due to the compliant connections at lower links, the pilot can make contact with the exoskeleton flexibly and impose forces on the exoskeleton.

As discussed in the previous section, the dynamics of the single DOF exoskeleton including the pilot dynamics can be represented as Eq. (6). Based on the interaction dynamics which is modeled as an impedance model in Eq. (7), the controller of single DOF exoskeleton is defined in Eq. (8) which has a sensitivity amplification factor.

2) Numerical Simulation Experiments: In the numerical simulation on single DOF exoskeleton platform, the pilot's target motion is set as periodic sine waves but with different frequencies and amplitudes. The parameters $\{J,B,m,l,g\}$ of the dynamic models in single DOF exoskeleton are set as one meter length rotary link with suitable values. Before controlling the exoskeleton, the original DMP is trained through sampling pilot motion trajectories. As a tradeoff between accuracy of approximation and real-time application, the number of Gaussians in DMP is set as 25, and ε in Tab. I is chosen as 0.5. Fig. 3 shows the normalized Mean Square Error of nonlinear function f during the incremental learning process. The results show that the motion trajectory can be approximated accurately after 20 iterations via learning weighted parameters of DMP which is described in Eq. (3).

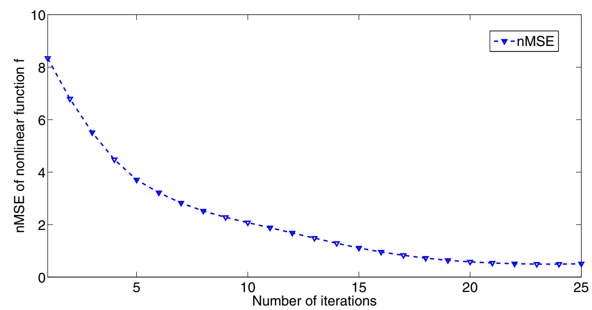


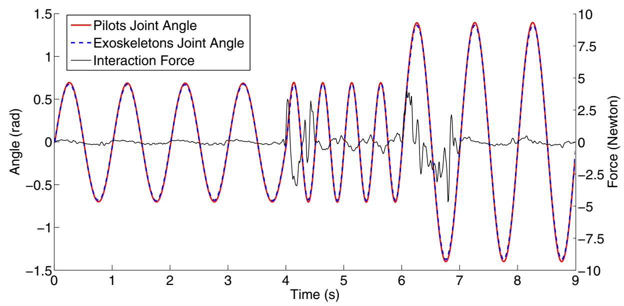
Fig. 3. The nMSE of nonlinear function f during the incremental learning process.

According to the framework of proposed HIL strategy presented in Fig. 1, the original DMP will be regarded as reference motion trajectory until the pilot moves the exoskeleton with pHRI. If the pilot changes the motion patterns with the exoskeleton, the reference motion trajectory will be learned incrementally with the reshaped DMP. In the experiments of single DOF exoskeleton, the pilot's motion is randomly chosen from, in total, 11 gait cycles which are all periodic sine waves but with different frequencies and amplitudes.

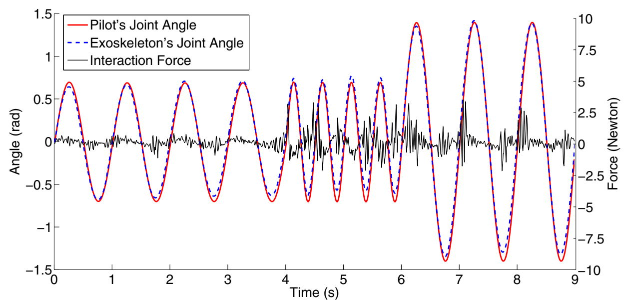
We compare the performance of the HIL strategy with SAC algorithm [8] in the experiment. In the simulation experiments, the frequencies of both of the control system are set at 100Hz.

The control performances of proposed HIL strategy and SAC algorithm are illustrated in Fig. 4(a) and Fig. 4(b), respectively. As shown in Fig. 4, the proposed HIL strategy achieve better performance than SAC algorithm except in the first gait cycle when the pilot change his/her motion patterns. The black curve as shown in Fig. 4 shows the interaction forces between the pilot and the exoskeleton which is calculated via Eq. (7).

As shown in Fig. 4(a) for HIL strategy, the interaction forces become larger when the pilot's motion pattern is changed since the exoskeleton is moved forcibly in the first gait cycle of changed patterns. After obtaining the reshaped motion trajectory, HIL strategy reduces the interaction forces much more than the SAC algorithm. The normalized mean square error (nMSE) of angle trajectories also indicate that the proposed HIL strategy achieve better performance in lower exoskeleton control (0.016 rad compared with 0.045 rad).

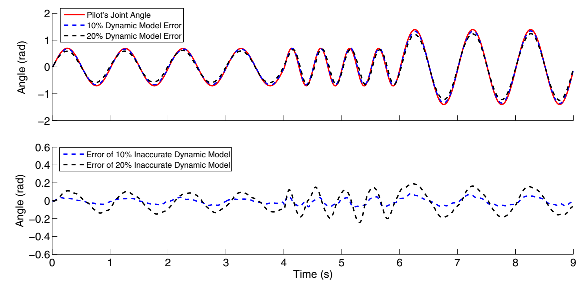


(a) HIL strategy

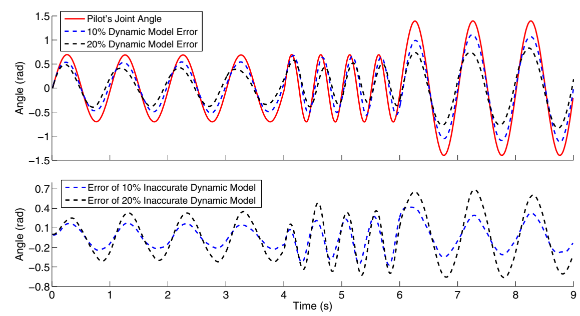


(b) SAC algorithm

Fig. 4. Comparison of control performances of (a) HIL strategy and (b) SAC algorithm with accurate dynamic models. The normalized mean squared error (nMSE) of the angle trajectories: (a) nMSE(HIL)$= 0.016$ *rad*, (b) nMSE(SAC)$= 0.045$ *rad*.



(a) HIL strategy



(b) SAC algorithm

Fig. 5. Comparison of control performances of (a) HIL strategy and (b) SAC algorithm with inaccurate dynamic models.

In model-based control algorithms, a fundamental problem is that the dynamic model of the system (exoskeleton in our case) must be accurate enough. To evaluate the robustness of the proposed HIL strategy against inaccurate dynamic models, another experiment is carried out and compared with the SAC algorithm.

In this experiment, the error of two inaccurate dynamic models are set as 10% and 20%. Fig. 5 presents the control performance of HIL strategy and SAC algorithm with inaccurate dynamic models. Table II represents the quantitative comparison of HIL strategy and SAC algorithm in single DOF simulation experiments. The experimental results show that the proposed HIL strategy obtained a better performance for managing changing interaction dynamic and inaccurate dynamic models.

TABLE II
COMPARISON OF HIL STRATEGY AND SAC ALGORITHM IN SINGLE DOF SIMULATION.

nMSE (rad)	accurate dynamic model	10% model error	20% model error
HIL	0.016	0.035	0.058
SAC	0.045	0.114	0.295

B. Experiments on a Human-Powered Augmentation Lower Exoskeleton

1) Introduction to HUALEX: The HUman-powered Augmentation Lower EXoskeleton (HUALEX) is designed as an ergonomic, robust and lightweight equipment for human augmentation applications. As shown in Fig. 6, HUALEX has, in total, four active joints to provide driving torques, which are activated by DC servo motors. The ankle joints of HUALEX are designed as an energy-storage mechanism which stores energy in the stance phase and release it in the swing phase during walking. Besides the joints and rigid links, many compliant connections at waist, thighs, shanks and feet are provided to connect HUALEX to the pilot in a semi-rigid way.

We embedded a distributed control system in HUALEX, which is composed of a main controller and four node controllers. The main controller is set at the backpack, which runs the control algorithm. Near each active joint, a node controller is installed for two purposes: 1. collecting sensor data for the main controller; 2. executing commands from the main controller.

In order to ensure the real-time performance of control algorithm, node controllers communicate with the main controllers via Controller Area Network (CAN). HUALEX has a total of three kinds of sensors for its current state monitoring. Encoders are integrated in joint actuators, which measures the current state of each joint. An accelerometer is set at the backpack to record the walking velocity of the pilot. Plantar sensors are placed in the sole for measuring the foot pressure.

2) Dynamic Models of HUALEX: In the HUALEX system, the dynamic models are expanded up to the whole lower extremities. A general form is utilized to describe HUALEX

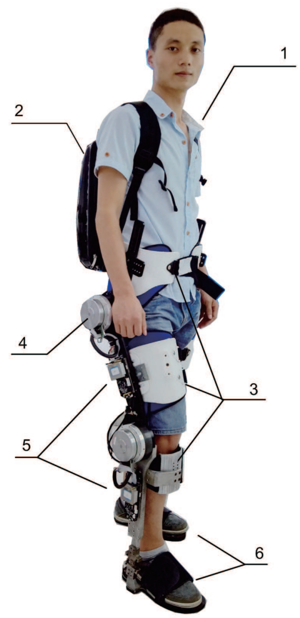


Fig. 6. HUman-powered Augmentation Lower EXoskeleton (HUALEX) with the pilot. 1: The pilot; 2: The load backpack with the power unit and main controller (rigid connection with the HUALEX spline); 3: Semi-rigidly connecting HUALEX to the pilot (waist, thighs, shanks and feet); 4. Active joints with DC servo motors (hip joints and knee joints); 5. Node controllers for active joints; 6. Smart shoes with plantar sensors.

dynamics:

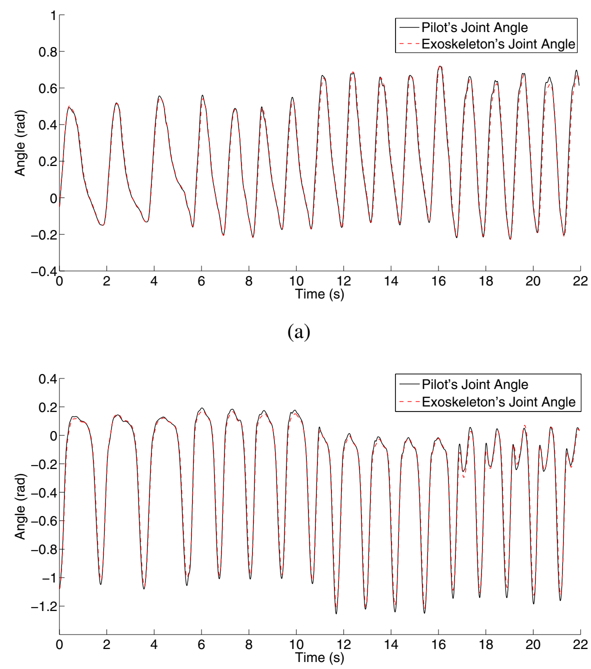
$$M(\Theta)\ddot{\Theta}+C(\Theta,\dot{\Theta})\dot{\Theta}+G(\Theta)=T_e+T_d \tag{10}$$

where Θ is a vector of active joints' angle, T_e and T_d represent the input torques from HUALEX and the pilot, respectively. The whole extremities are separated in swing leg and stance leg:

- Swing Leg: The swing leg of HUALEX is modeled as a three-link swing mechanism in the sagittal plane with a base coordinate fixed on the back of it, where the parameters of $M(\Theta)$ and $C(\Theta,\dot{\Theta})$ are 3×3 matrices, and $\{\Theta,T_e,T_d,G(\Theta)\}$ are 3×1 vectors.
- Stance Leg: The stance leg is considered as a four DOFs serial link mechanism with a base coordinate fixed on the ankle joint. The parameters of $M(\Theta)$ and $C(\Theta,\dot{\Theta})$ are 4×4 matrix, and $\{\Theta,T_e,T_d,G(\Theta)\}$ are 4×1 vectors.

3) Experimental Setup: Three pilots (A, B, C) with different heights (170cm, 176cm, 180cm) are chosen to operate the HUALEX system for one minute in the experiment. The controller of each joint in the HUALEX is designed based on the dynamic models presented in Eq. (10), with extra sensitivity factors. The parameters of the dynamic models are identified through Solidworks. In order to validate the control performance of proposed HIL strategy, a wearable sensory system (including five inclinometers) is utilized to capture the joints' state of the pilot.

4) Results and Discussions: The control performance of the HIL strategy in HUALEX with pilot C is shown in Fig. 7 (a period of the whole experiment). The pilot's waling speed in Fig. 7, is varying from 0.4 m/s to 1.2 m/s. Experimental results indicate that through incremental DMP and reinforcement learning, HUALEX is able to follow the pilot's motion after one gait cycle's correction. Fig. 8 shows the control performance of the proposed HIL strategy with three pilots at different walking speeds. The nMSE value is measured in left knee joints. The results show that proposed HIL strategy obtained a good performance (largest nMSE is 0.032 rad). As it is shown in Fig. 8, for lower walking speed, the proposed HIL strategy has better performance.



(a)

(b)

Fig. 7. Control performance of the proposed HIL strategy on HUALEX with pilot C in 22 seconds walking (period of whole experiments): (a) left hip joint; (b) left knee joint.

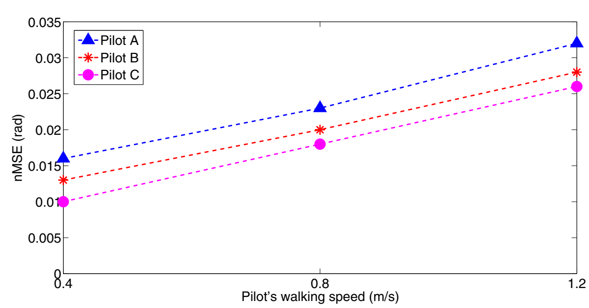


Fig. 8. Control performance of proposed HIL strategy with three pilots in different walking speed (in left knee joint).

Table III shows the comparison of SAC and HIL strategy at the left knee joint. In Table III, we can see that the proposed HIL strategy achieves better performance in HUALEX robotic system than SAC, especially in higher walking speed situations. Taking the experiments with pilot

C as an example, the SAC algorithm has almost three times nMSE when compared to the HIL strategy (0.076 rad compare to 0.026 rad).

The experimental results on both single DOF platform and HUALEX show that, through simultaneously learning the pilot's motion and controller, the HIL strategy is able to cope with the changing interaction dynamics from the pilot.

TABLE III
COMPARISON OF SAC AND HIL STRATEGY IN HUALEX.

nMSE (rad) (HIL|SAC)	0.4 m/s		0.8 m/s		1.2 m/s	
pilot A	0.016	0.033	0.023	0.048	0.032	0.086
pilot B	0.013	0.026	0.02	0.045	0.028	0.08
pilot C	0.01	0.026	0.018	0.04	**0.026**	**0.076**

V. CONCLUSION AND FUTURE WORKS

This paper has proposed a Hierarchical Interactive Learning (HIL) strategy to control a HUman-powered Augmentation EXoskeleton robotic system. The HIL strategy has two learning hierarchies, namely, high-level motion learning and low-level controller learning. We aim at learning the motion trajectories and the controllers simultaneously. For the motion learning hierarchy, we employ the incremental learning method based on LWR to optimize the weighted parameters of DMPs. After obtaining the reshaped motion trajectories, RL method is utilized to update the sensitivity factors of controllers to reduce the interaction forces. Control performances of the HIL strategy are evaluated on a single DOF exoskeleton platform as well as the HUALEX robotic system. Experimental results show that the proposed HIL strategy is able to cope with changing interaction dynamics in different gait patterns.

In the future, we will investigate in methods which can predict the pilot's short term intention. In this case, the HUALEX is able to adjust corresponding parameters in advance and proactively serve the pilot/user. Moreover, we are also building a hydraulic exoskeleton system which has extensive applications in load carrying scenarios.

VI. ACKNOWLEDGMENTS

This work was made possible by support from NSFC (No. 6150020696, 6157021026), Fundamental Research Funds for the Central Universities (ZYGX2014Z009) and SRF for ROCS, SEM. Moreover, the authors greatly acknowledge the contribution of the reviewer's comments.

REFERENCES

[1] H. Kazerooni, A. Chu, R. Steger, "That Which Does Not Stabilize, Will Only Make Us Stronger", *International Journal of Robotics Research*, vol. 26, 2007, pp 75-89.
[2] Y. Sankai, "HAL: Hybrid Assistive Limb based on Cybernics", *Robotics Research*, 2011, pp 25-34.
[3] C. J. Walsh, D. Paluska, et al., "Development of a Lightweight, Underactuated Exoskeleton for Load-Carrying Augmentation", In *IEEE International Conference on Robotics and Automation (ICRA)*, 2006, pp 3485-3491.
[4] S. Lee, Y. Sankai, "Power Assist Control for Walking Aid with HAL-3 Based on EMG and Impedance Adjustment around Knee Joint", In *International Conference on Intelligent Robots and Systems (IROS)*, 2002, pp 1499-1504.
[5] T. Hayashi, H. Kawamoto, Y. Sankai, "Control Method of Robot Suit HAL Working as Operator's Muscle Biological and Dynamical Information", In *International Conference on Intelligent Robots and Systems (IROS)*, 2005, pp 3063-3068.
[6] A.O. Gabriel, J.E. Colgate, M.A. Peshkin, A. Goswami, "Active-Impedance Control of a Lower-Limb Assistive Exoskeleton", In *IEEE International Conference on Rehabilitation Robotics*, 2007, pp 188-195.
[7] H.T. Tran, H. Cheng, M.K. Duong, H. Zheng, "Fuuzy-based Impedance Regulation for Control of the Coupled Human-Exoskeleton System", In *IEEE International Conference on Robotics and Biomimetics*, 2014.
[8] H. Kazerooni, J.L Racine, L. Huang, R. Steger, "On the Control of the Berkeley Lower Extremity Exoskeleton (BLEEX)", In *International Conference of Robotics and Automation (ICRA)*, 2005, pp 4353-4360.
[9] J. Ghan, R. Steger, H. Kazerooni, "Control and System Identification for the Berkeley Lower Extremity Exoskeleton", *Advanced Robotics*, vol. 20, no. 9, 2006.
[10] H.T. Tran, H. Cheng, X. Lin, R. Huang, "The Relationship between Physical Human-Exoskeleton Interaction and Dynamic Factors: Using a Learning Approach for Control Applications", *Science China Information Science*, vol. 57, no. 12, 2014.
[11] C.G. Atkeson, A.W. Moore, S. Schaal, "Locally Weighted Learning for Control", *Lazy Learning*, 2014, pp 75-113.
[12] A.Y. Ng, A. Coates, M. Diel, V. Ganapathi, et al., "Autonomous Inverted Helicopter Flight via Reinforcement Learning", *Experimental Robotics*, 2006, pp 363-372.
[13] R. Huang, H. Cheng, Q. Chen, et al., "Interative Learning for Sensitivity Factors of a Human-Powered Augmentation Lower Exoskeleton", In *IEEE International Conference on Intelligent Robots and Systems (IROS)*, 2015 (accepted).
[14] A.J. Ijspeert, J. Nakanishi, S. Schaal, "Movement Imitation with Nonlinear Dynamical Systems in Humanoid Robots", In *IEEE International Conference on Robotics and Automation (ICRA)*, 2002, pp 1398-1430.
[15] A.J. Ijspeert, J. Nakanishi, S. Schaal, "Learning Rhythmic Movements by Demonstration using Nonlinear Oscillators", In *IEEE International Conference on Intelligent Robots and Systems (IROS)*, 2002, pp 958-963.
[16] J. Nakanishi, J. Morimoto, G. Endo, et al., "Learning from Demonstration and Adaptation of Biped Locomotion", *Robotics and Autonomous Systems*, vol. 47, no. 2, 2004.
[17] H.B. Amor, G. Neumann, S. Kamthe, et al., "Interation Primitives for Human-Robot Cooperation Tasks", In *IEEE International Conference on Robotics and Automation (ICRA)*, 2014, pp 2831-2837.
[18] A. Gams, B. Nemec, A.J. Ijspeert, "Coulpling Movement Primitives: Interaction with the Environment and Bimanual Tasks", *IEEE Transaction on Robotics*, vol. 30, no. 4, 2014.
[19] F. Stulp, S. Schaal, "Hierarchical Reinforcement Learning with Movement Primitives", In *IEEE-RAS International Conference on Humanoid Robots*, 2011, pp 231-238.
[20] P. Kormushev, S. Calinon, D.G. Caldwell, "Robot Motor Skill Coordination with EM-based Reinforcement Learning", In *IEEE International Conference on Intelligent Robots and Systems (IROS)*, 2010, pp 3232-3237.
[21] K. Kronander, M. Khansari, A. Billard, "Incremental Motion Learning with Locally Modulated Dynamical Systems", *Robotics and Autonomous Systems*, vol. 70, 2015.
[22] D. Kulic, C. Ott, D. Lee, et al., "Incremental Learning of Full Body Motion Primitives and Their Sequencing through Human Motion Observation", *International Journal of Robotics Research*, 2011.
[23] S. Schaal, C.G. Atkeson, "Constructive Incremental Learning from Only Local Information", *Neural Computation*, vol. 10, no. 8, 1998.