SCIENCE CHINA
Information Sciences

• LETTER •

November 2021, Vol. 64 219201:1–219201:3
https://doi.org/10.1007/s11432-019-2754-9

Adaptive compensation for time-varying uncertainties in model-based control of lower-limb exoskeleton systems

Guangkui SONG[1], Rui HUANG[1*], Hong CHENG[1], Jing QIU[1,2] & Shuai FAN[2]

[1]Center for Robotics, University of Electronic Science and Technology of China, Chengdu 611731, China;
[2]School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China

Received 19 April 2019/Revised 2 September 2019/Accepted 26 December 2019/Published online 30 July 2020

Citation Song G K, Huang R, Cheng H, et al. Adaptive compensation for time-varying uncertainties in model-based control of lower-limb exoskeleton systems. Sci China Inf Sci, 2021, 64(11): 219201, https://doi.org/10.1007/s11432-019-2754-9

Dear editor,

Exoskeletons are widely used in human locomotion assistance and strength augmentation to enhance strength and endurance, wherein strategies based on model-based control are extremely sensitive to errors in the dynamic models. Additionally, the unpredictable human-exoskeleton interaction force can unbalance the pilot. Therefore, learning methods were utilized to improve the accuracy of the dynamic model [1, 2]. The modeling of dynamic model was decoupled based on the direction of interest [3]. Nevertheless, the time-varying uncertainty of the dynamic model [4] forced the learning algorithm to learn this model at all times. In this study, a novel strategy called adaptive compensation learning (ACL) is proposed to consider the extra interaction forces caused by the time-varying uncertainty, in which the model-based controller is used to amplify the forces caused by the pilot [5], and the compensation strategy generator adaptively generates a compensation strategy for the extra interaction forces. Reinforcement learning can determine the optimal coefficients of the ACL strategy for different pilots and walking patterns. In the learning process, we consider the interaction forces and their changing velocities. The efficiency of the ACL is verified through the comparison experiments, and the experimental results demonstrate that the ACL can achieve better performance.

Our algorithm. The proposed ACL strategy in Figure 1(a) is designed for the swing phase with high-frequency, and is detailed with a single degree of freedom (DOF) exoskeleton platform. The single DOF exoskeleton comprises a thigh and a shank, which are connected by a knee joint. The dynamic model of the single DOF exoskeleton is defined as follows:

$$\begin{cases} J\ddot{\theta} + B\dot{\theta} + mgl \cdot \sin\theta = \tau_e + \tau_h, \\ \tau_e = \hat{\tau}_e + \tau_c, \end{cases} \tag{1}$$

where J, B, m and l represent the physical parameters of the single DOF exoskeleton; a joint state vector $(\theta, \dot{\theta}, \ddot{\theta})$ is used to describe the motion states. During the locomotion, the exoskeleton is driven by a resultant torque τ_e provided by an actuator, and the torque τ_h is imposed by the pilot. The extra unpredictable interaction forces resulting from the errors in the dynamic model are reflected in the torque τ_h and considered in a compensation strategy by the compensation strategy generator, which outputs the compensation torque τ_c. A limit-cycle oscillator dynamic movement primitives (DMP) model that focuses on the imitation learning of rhythmical movements [6] is used to learn and estimate the current motion of the pilot. The estimated motion $(\hat{\theta}_h, \dot{\hat{\theta}}_h, \ddot{\hat{\theta}}_h)$ is mapped to the joint torque τ_c through a proportion derivative (PD) gain controller [7] with a proportion gain P and a derivative gain D, due to its simplicity and stable performance. The main assistive torque $\hat{\tau}_e$ is provided by a model-based controller as follows:

$$\tau_c = P(\theta - \hat{\theta}_h) + D(\dot{\theta} - \dot{\hat{\theta}}_h), \tag{2}$$

$$\hat{\tau}_e = mgl \cdot \sin\theta + (1 - \alpha^{-1})(\hat{J}\ddot{\theta} + \hat{B}\dot{\theta}), \tag{3}$$

where $\hat{J}$ and $\hat{B}$ are the physical parameters estimated offline. α is an amplification number greater than unity, which has been optimized in advance in this study.

The motion trajectory $\hat{\theta}_h$ is described by a limit-cycle oscillator as

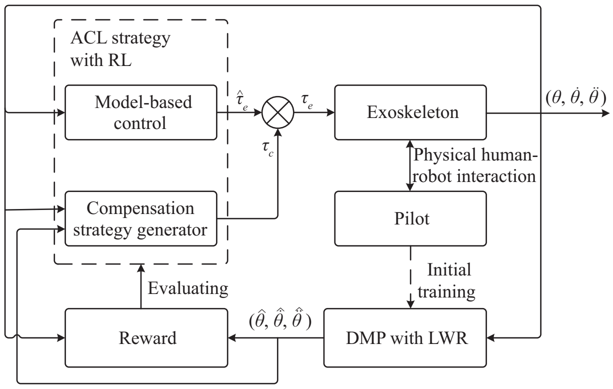
$$\tau\dot{\hat{\theta}}_h = \beta(\theta_m - \hat{\theta}_h) + f, \tag{4}$$

where β is set as a positive constant, τ relates to the frequency of the oscillator, and θ_m determines the baseline around which the motion trajectory $\hat{\theta}_h$ oscillates. The nonlinear function f is defined as

$$f = \frac{\sum_{i=1}^{N} \psi_i \omega_i^{\mathrm{T}} \tilde{z}}{\sum_{i=1}^{N} \psi_i}, \quad \tilde{z} = [z, \sqrt{E_0}]^{\mathrm{T}}. \tag{5}$$

* Corresponding author (email: ruihuang@uestc.edu.cn)

© Science China Press and Springer-Verlag GmbH Germany, part of Springer Nature 2020



(a)

	Walking speed = 0.5 m/s			Walking speed = 1 m/s			Walking speed = 2 m/s		
	ACL	HIL	SAC	ACL	HIL	SAC	ACL	HIL	SAC
PilotA	0.005	0.017	0.034	0.013	0.024	0.049	0.026	0.034	0.088
PilotB	0.005	0.014	0.027	0.014	5.100	0.421	0.030	0.030	0.100
PilotC	0.04	0.011	0.027	0.018	0.019	0.041	0.027	0.018	0.078

(b)

Figure 1 (a) Control diagram of ACL strategy; (b) comparison of ACL, HIL and SAC strategies with nMSE (rad) on HUALEX.

The curve represented by the nonlinear function f is adjusted by learning the appropriate weights ω_i of the Gaussian kernels $\psi_i = \exp -0.5h_i(\phi - c_i)^2$ for enabling DMP to model arbitrary trajectories. Before learning, the learning goal f_g is calculated by

$$f_g = \tau\dot{\theta}_d - \beta(\theta_m - \theta_d) \tag{6}$$

using recorded motion trajectories θ_d, and locally weighted regression (LWR) is used to learn the weights ω_i. Finally, the estimated motion trajectories $\hat{\theta}_h$ can be obtained by (4).

For different pilots and walking patterns, the compensation torque must be adaptively tuned [8], which is provided by the introduction of reinforcement learning (Q-learning). The exoskeleton executes an action $a \rightarrow (P, D)$ in the current state $s \rightarrow (\theta, \dot{\theta}, \ddot{\theta})$, calculates immediate reward r by

$$\begin{aligned} r = &- [k_1(\theta - \hat{\theta}_h)^2 + k_2(\dot{\theta} - \dot{\hat{\theta}}_h)^2] \\ &- [k_3((\theta(i) - \hat{\theta}_h(i)) - (\theta(i-1) \\ &- \hat{\theta}_h(i-1))) \\ &+ k_4((\dot{\theta}(i) - \dot{\hat{\theta}}_h(i)) - (\dot{\theta}(i-1) \\ &- \dot{\hat{\theta}}_h(i-1)))], \end{aligned} \tag{7}$$

and observes the next state s'. The action-value function Q is updated by

$$\begin{aligned} Q_{k+1}(s,a) =& (1-\alpha_k)Q_k(s,a) \\ &+ \alpha_k\Big[r + \gamma \max_{a\in A} Q_k(s',a')\Big], \end{aligned} \tag{8}$$

where α can adjust the learning rate; the immediate reward is decayed by γ. During learning, the maximization of the term $-[k_1(\theta - \hat{\theta}_h)^2 + k_2(\dot{\theta} - \dot{\hat{\theta}}_h)^2]$ is the process of minimizing the effects of the time-varying uncertainty, and the term $-[k_3((\theta(i) - \hat{\theta}_h(i)) - (\theta(i-1) - \hat{\theta}_h(i-1))) + k_4((\dot{\theta}(i) - \dot{\hat{\theta}}_h(i)) - (\dot{\theta}(i-1) - \dot{\hat{\theta}}_h(i-1)))]$ is used for smoothing the exoskeleton output. k_1, k_2, k_3 and k_4 are used to weight the terms. If the admissible pairs of state-action are updated infinitely, and α_k decays with increasing k while obeying the usual stochastic approximation condition, Q_k converges to optimal Q^* with a probability of 1.

Experimental results. Our strategy is validated on a single DOF exoskeleton and a human-powered augmentation lower exoskeleton (HUALEX) system. In the first experiment, we manually set the time-varying uncertainty of the dynamic model to verify the performance of the ACL strategy on the single DOF exoskeleton. The errors of the dynamic model are set as 10% and 20% (e.g., the inaccurate rate of 20% implies that we set the estimated parameters $\hat{J}$ and $\hat{B}$ as $\hat{J} = 0.8J$ and $\hat{B} = 0.8B$). The optimal $\alpha = 100$ obtained from offline learning is used in the experiments, and the parameters are set as $k_1 = 1$, $k_2 = 2$, $k_3 = 0.5$ and $k_4 = 1$ for experience. After the learning, the action vector (P, D) is updated with appropriate values for different pilots and walking patterns. The experimental results demonstrate a good performance of the proposed ACL strategy for handling dynamic models with uncertainties (normalized mean squared error (nMSE): 0.009 rad at 10% model errors and 0.010 rad at 20% model errors). The quantitative comparison of the ACL, SAC and HIL with nMSE in single DOF platform experiments demonstrates that the proposed ACL strategy can handle an inaccurate dynamic model well; especially in higher model error situations, the advantages of the ACL strategy are more evident.

The HUALEX system is used to amplify the strength and endurance of the pilot using a pair of wearable robotic legs. There are four active joints on the hips and knees to provide the assistive torques. A storing energy mechanism is integrated in the ankle joint to store energy in stance phase and release it in the swing phase.

In the HUALEX experiments, we obtain a set of practical dynamic parameters of the HUALEX from the design parameters. The parameters $k_1 = 1$, $k_2 = 2$, $k_3 = 0.5$ and $k_4 = 1$ of the ACL strategy is the same as that in the single DOF exoskeleton experiments. Three randomly selected pilots operate the exoskeleton at different speeds (0.5, 1 and 2 m/s) for 3 min. We set four inertial measurement unit (IMU) sensors on the lower links of the pilot to capture the states of each joint, and calculate the nMSE to illustrate the performance of the ACL strategy instead of the interaction force obtained by modifying the mechanical structure. The comparison of the performances of the SAC, HIL and ACL at the left knee joint is presented in Figure 1(b). Although the nMSE of the three pilots increases with walking speed, it is still smaller than that in other algorithms. The HIL algorithm achieves better following performance than SAC due to the optimization of its sensitivity factor of SAC algorithm. However, the HIL algorithm does not consider the time-varying uncertainty. Therefore, the ACL strategy exhibits a better performance. The existence of the time-varying uncertainty enables the ACL strategy to further improve the exoskeleton algorithm. The nMSE increases with walking speed, which can be attributed to the increased estimation errors in the DMP model.

Acknowledgements This work was supported by National Key Research and Development Program of China (Grant No. 2017YFB1302300), National Natural Science Foundation of China (Grant Nos. 6150020696, 61503060), Sichuan Science and Technology Major Projects of New Generation Artificial Intelligence (Grant No. 2018GZDZX0037), and Fundamental Research Funds for the Central Universities (Grant No. ZYGX2015J148).

Supporting information Appendix A. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Atkeson C G, Moore A W, Schaal S. Locally weighted learning for control. Lazy Learn, 1997, 11: 75–113
2 Nguyen-Tuong D, Peters J R, Seeger M. Local Gaussian process regression for real time online model learning. In: Proceedings of Advances in Neural Information Processing Systems, 2009. 1193–1200
3 Yue M, Ning Y G, Yu S Z, et al. Composite following control for wheeled inverted pendulum vehicles based on human-robot interaction. Sci China Inf Sci, 2019, 62: 050206
4 Li Z, Ma W, Yin Z, et al. Tracking control of time-varying knee exoskeleton disturbed by interaction torque. ISA Trans, 2017, 71: 458–466
5 Kazerooni H, Racine J-L, Huang L, et al. On the control of the Berkeley lower extremity exoskeleton (BLEEX). In: Proceedings of IEEE International Conference on Robotics and Automation (ICRA), 2005. 4353–4360
6 Song G K, Huang R, Cheng H, et al. Learning continuous coupled multi-controller coefficients based on actor-critic algorithm for lower-limb exoskeleton. Sci China Inf Sci, 2021, 64: 159203
7 Kelly R. PD control with desired gravity compensation of robotic manipulators. Int J Robot Res, 1997, 16: 660–672
8 Huang R, Cheng H, Guo H, et al. Hierarchical learning control with physical human-exoskeleton interaction. Inf Sci, 2018, 432: 584–595