Boosting Learning Efficiency for Tuning a Powered Knee Prosthesis: Offline-Online Policy Iteration

Minhan Li
Department of Biomedical Engineering
NC State University
UNC-Chapel Hill
NC, 27695, USA
mli37@ncsu.edu

Jennie Si
Department of Electrical, Computer, and Energy Engineering
Arizona State University
AZ, 85281, USA
si@asu.edu

He (Helen) Huang
Department of Biomedical Engineering
NC State University
UNC-Chapel Hill
NC, 27695, USA
hhuang11@ncsu.edu

I. Introduction

The current powered knee prostheses are widely relying on the finite-state machine impedance control (FSM-IC) to regulate the dynamics between users, devices and environments. To maximize benefits of the power to individual users, a personalized impedance tuning is indispensable, yet challenging. Recent studies have attempted various approaches to automate the tuning process, and achieved promising results [1]–[3]. However, few work has been focusing on enhancing the tuning efficiency in high-dimension parameter space with human-in-the-loop, resulting in a limited scalability for clinical applications. To address the remaining problem, we are in need of a novel automatic method that can efficiently obtain a tuning policy, and reduce the time and labor efforts required in online training that involves human users.

To fulfill such a requirement, this study developed a novel reinforcement learning based algorithm by taking the framework of policy iteration, which enabled the training of tuning policy in both offline and online manners. Enjoying the advantages from the both training schemes, the method could significantly enhance the efficiency in policy learning.

II. Methods

In this study, the proposed algorithm is designed to find out an optimal policy, acting as a supplementary parameter tuner to the powered knee prosthesis control, to have users restore a near-normal knee motion. The algorithm is implemented together with a well-established FSM-IC framework [4]. The whole learning process involves two types of updates: the impedance update is executed by the parameter tuner following some specific policies to adjust impedance for the FSM-IC; whereas the policy update is carried out by the novel policy iteration algorithm where policy evaluation and policy improvement were alternately performed until meeting stopping conditions. To test the algorithm, we recruited two participants, one able-bodied individual and one transfemoral amputee, and each was asked to perform the treadmill walking task at the speed of 0.6 m/s.

This work was supported in part by National Science Foundation #1563454, #1563921, #1808752 and #1808898.

III. Results and Conclusions

We tested the online training of prosthesis tuning policy on the participants while their walking. The test was conducted under two different initial policy conditions (1) randomly initialized, and (2) offline pre-trained. For the offline pre-trained policy, the training data was already gathered from other sources before the test. The initial impedance parameters for the test were random selected. We evaluated the efficiency by quantifying the number of impedance updates to meet the stopping criterion for prosthesis tuning.

The comparison results between two conditions revealed that online tuning starting with pret-rained policies, albeit not perfect, were significantly more efficient than those starting with random policies. That is to say, pre-trained cases were observed to have less overall number of impedance updates than random cases did as to meet the stopping criterion. On average, pre-trained cases ended up with 38 impedance updates (about 4.7 minutes), whereas random cases took 96 impedance updates to complete the tuning, which was equivalent to about 12 minutes. In addition, at the end of training, all yielded policies presented a strong stability and robustness against perturbations or task changes to keep reaching the tuning goal.

In conclusion, we implemented the proposed novel algorithm with a real human-prosthesis system, and validated its exceptional performances including efficiency, stability and robustness. Such findings demonstrated promising potentials of the proposed method to be used for, in the future, clinical applications at scales.

References

[1] Y. Wen, J. Si, A. Brandt, X. Gao, and H. Huang, "Online reinforcement learning control for the personalization of a robotic knee prosthesis," *IEEE Trans. Cybern.*, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8613842

[2] M. Li, X. Gao, Y. Wen, J. Si, and H. Huang,"Offline policy iteration based reinforcement learning controller for online robotic knee prosthesis parameter tuning," *ICRA 2019*, May 2019, pp. 2831-2837.

[3] H. Huang, D. L. Crouch, M. Liu, G. S. Sawicki, and D. Wang,"A cyber expert system for auto-tuning powered prosthesis impedance control parameters," *Ann. Biomed. Eng.*, vol. 44, no. 5, pp. 1613-1624, 2016.

[4] F. Sup, A. Bohara, and M. Goldfarb, "Design and control of a powered transfemoral prosthesis,"*Int. J. Robot. Res*, vol. 27, no. 2, pp. 263–273, 2008.