Learning Human Postural Strategies: Reinforcement Learning Approach Incorporating a CoP Constraint

Seongwoong Hong, and Sukyung Park
Dept. of Mechanical Engineering, KAIST, Daejeon, Rep. of Korea
Email: sukyungp@kaist.ac.kr

Summary

This study suggests that reinforcement learning, employing a trial-and-error-based learning approach with a CoP constraint reward, can account for the human ankle-hip strategy transition during external perturbations.

Introduction

Humans maintain balance using an ankle strategy for small perturbations, switching to a hip strategy for larger ones. While each of these strategies have been well-studied, the mechanisms governing their transition remain unclear[1-2]. Therefore, this study aims to demonstrate that the transition from ankle to hip strategies in response to external perturbations can be learned through reinforcement learning (RL), with the center of pressure (CoP) constraint during upright standing considered as part of the reward.

Methods

We collected postural control data from ten healthy subjects who maintained an upright stance on a backward-moving force platform. Ground reaction forces, joint angles, and angular velocities were recorded using a force platform and motion capture system; joint forces and torques were calculated via inverse dynamics.

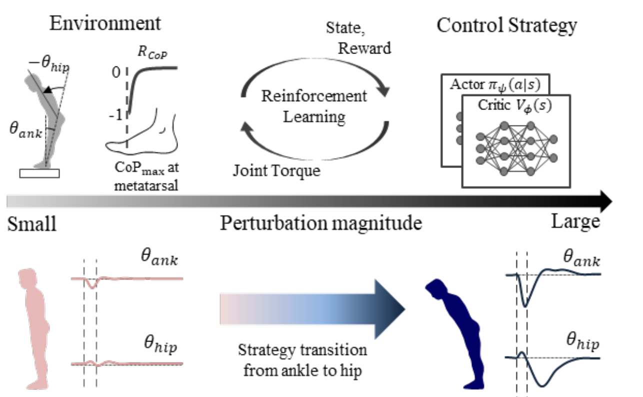

Figure 1: RL framework for learning human postural control strategy transitions in response to backward perturbations

We constructed an RL framework (Figure 1) to learn the ankle-hip strategy transition. The reward function was designed to encourage upright posture and penalize high joint torques, while crucially incorporating a CoP constraint: the reward decreased sharply as the CoP approached its limits. An optimal control policy was learned to maximize this reward, and the resulting RL agent's behavior was compared with human data under identical perturbation conditions.

Results and Discussion

The RL agent, trained with the CoP constraint incorporated into the reward, successfully learned to ankle-hip strategy transition in response to increasing magnitudes of backward perturbations. In contrast, when the CoP constraint was not included in the reward, the agent maintained the ankle strategy regardless of the perturbation magnitude (Figure 2).

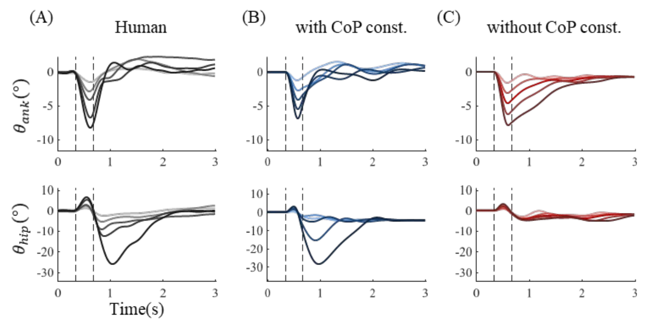

Figure 2: Ankle and hip joint angles of (A) human, (B) RL model with CoP, and (C) without CoP constraints, during backward perturbations. Dashed lines: perturbation start/end. Color gradient: increasing perturbation magnitude.

This indicates that the strategy transition is learned from a reward function that incorporates the CoP constraint. For small perturbations, the CoP constraint has a minimal impact on the reward, leading to the learning of an ankle strategy that quickly restores upright posture. However, for large perturbations, the CoP frequently approaches its limits, causing a significant reduction in reward. Therefore, the agent learns a hip strategy, which maintains smaller CoP movements compared to the ankle strategy, to maintain balance. These results suggest that the ankle-hip strategy transition in human can be learned through a trial-and-error process that concurrently optimizes posture stabilization and CoP constraint rewards.

Conclusions

This study demonstrates that a trial-and-error-based learning mechanism, incorporating the CoP constraint, can reproduce the strategic transition of humans. These findings that shifts in the relative importance of CoP constraints in response to perturbations are a key factor in learning the ankle-to-hip strategy transition of humans.

Acknowledgments

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. RS-2024-00356657) and Basic Research Funds from the Department of Mechanical Engineering at KAIST (A0601003029).

References

[1] Afschrift J et al. (2016). *J. Neurophysiol*, **116**: 1937-1945.
[2] Wouwe V. et al. (2022). *J. Neurophysiol,* **125**: 586-598