Emergence of Natural and Robust Bipedal Walking by Learning from Biologically Plausible Objectives

Pierre Schumacher[1,2], Thomas Geijtenbeek[3], Vittorio Caggiano[4], Vikash Kumar[4,5], Syn Schmitt[6], Georg Martius[1], Daniel F. B. Haeufle[1,2]

[1]Department of Computer Science, University of Tübingen, Germany, [2]Hertie Institute for Clinical Brain Research and Center for Integrative Neuroscience, Tübingen, Germany, [3]Goatstream, The Netherlands, [4]Myolab Inc., New York, USA, [5]Carnegie Mellon University, Pittsburgh, USA, [6]Institute for Modelling and Simulation of Biomechanical Systems, University of Stuttgart, Germany

Email: pierre.schumacher@tuebingen.mpg.de

Summary

Humans show unparalleled ability when maneuvering diverse terrains. While reinforcement learning (RL) has shown great promise for musculoskeletal simulations, current approaches rely on motion data for policy training. Our combination of RL with a bio-inspired reward can learn robust locomotion controllers for 4 different musculoskeletal models and with up to 90 muscles by leveraging an adaptive reward function. This enables motor control and biomechanics communities to generate complex behaviors with RL, without relying on large amounts of motion data.

Introduction

Predictive simulations have reproduced human locomotion behavior in simple models, but high-dimensional 3D movement on uneven terrains has not been realized. In robotics, reinforcement learning (RL) methods can achieve highly stable (and efficient) locomotion, but the generation of human-like walking requires extensive use of motion capture data. A robust RL controller for high-dimensional musculoskeletal systems that can produce natural motion is still missing.

Methods

We use DEP-RL [1] a recent RL algorithm for musculoskeletal control known for its robustness and flexibility. RL training produces a controller aiming to maximize a numerical reward function. The starting point for our reward leverages results from previous studies [2], using a combination of task-, energy-, and torque-based costs aiming to induce behaviors that are close-to-natural motion:

$$r = w_1 r_{task} - w_2 c_{effort} - w_3 c_{pain},$$

where the learning reward r is constituted of r_{task}, incentivizing the controller to keep a running velocity of 1.2 m/s, c_{effort} is a combination of muscle effort and activity smoothness costs, c_{pain} reduces joint overextensions and ground-reaction forces, and w_i are weighting coefficients.

It is, however, observed that this reward function cannot adequately reduce the used muscle effort, even after optimizing the weighting coefficients of the respective terms. We therefore propose to use an adaptive reward schedule that adapts only the effort coefficient during training, making our method more performant and easily applicable to different models.

Results and Discussion

The combination of a bio-inspired reward function, a performant RL agent and an adaptive effort term result in gaits that match experimental data, for musculoskeletal models of varying complexity, without the need to adjust the reward function or agent hyperparameters.

We also observe that our controllers are more robust under unseen terrain variations than comparable reflex-based controller baselines [3] and are applicable to complex humanoid models with up to 90 muscles.

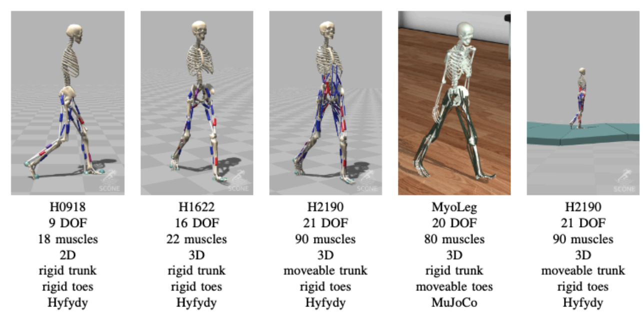

Figure 1: This study uses a variety of different models of varying complexity: 2D and 3D movement, between 18 and 90 muscles, between 9 and 21 degrees-of-freedom and across two simulators.

Conclusions

Our approach to musculoskeletal control provides a solid foundation for the development of neuromechanical simulations with robust adaptive controllers for complex tasks and environments.

Acknowledgments

This work was supported by the Cyber Valley Research Fund (CyVy-RF-2020-11 to DH and GM)

References

[1] Schumacher et al. (2023). *ICLR*, DEP-RL
[2] Weng et al. (2023). *IEEE-RAL*, 6(2): 4156-4162
[3] Song et al. (2015). *J Physiol.*, 1593(16): 3493–3511