Learning Bipedal Locomotion in Task Space

Helei Duan, Jeremy Dao, Kevin Green, Alan Fern, Jonathan Hurst
Collaborative Robotics and Intelligent Systems Institute, Oregon State University, Corvallis, Oregon, 97331
Email: {duanh, daoje, greenkev, alan.fern, jonathan.hurst}@oregonstate.edu

***Abstract*—Recent reinforcement learning (RL) approaches have demonstrated the possibility of generating continuous control policies for bipedal robots. However, most efforts focus on imitating trajectories in the complete robot configuration space, which can make it challenging for learning agents to incorporate primary control interests in task space, such as desired reaction forces or foot placements. In this work, we present an RL approach that leverages domain knowledge to directly learn the mapping from bipedal robot states to desired task space setpoints given an expert trajectory. Specifically, the agent learns to output foot positions with regard to the floating base, which are then mapped to joint torques by projecting inverse dynamics into task space. We evaluated the approach using bipedal robot Cassie in simulation and showed the learned policy produces stable gaits and robustness under various disturbances.**

I. Introduction

Legged locomotion has been extensively studied in terms of reduced-order models. These models often parameterize locomotion dynamics in a particular action space, such as leg angle, actuator position (setpoint), and impedance along the actuation direction, so that the models can generate a set of proper dynamics between the floating base and the contacts.

Taking inspiration from biomechanics observations and theories from reduced order models, interfacing learning-based control policies with variable setpoints (foot trajectories or placements) or even impedances in task space might speed up the learning process and produce natural and robust behaviors.

II. Methods

Previous task space control techniques on bipedal robots, such as in [1], require an optimization problem that takes in desired task space commands to generate torques. Adding such an optimization process in the training loop can significantly reduce the sampling speed. We apply the analytical projection-based methods [2] to obtain torques given desired task space trajectories of the foot when considering the mechanism constraints and underactuation. Here, we only consider the foot commands given the center of mass and foot trajectories. Adding control of the floating-base will be added to the learning framework in future work.

We use Proximal Policy Optimization [3], a model-free policy gradient actor-critic method. The action space represents the Cartesian foot positions residuals, which will be added with reference foot positions from the expert trajectory and sent into the task space control.

III. Results and Discussions

The learned policy is able to maintain balance during walking and track the desired task space control interests in simulation. The simulation results also exhibit robustness under various impulse disturbances to the pelvis during the execution of the policy.

Fig. 1: The policy learns to choose actions in Cassie's task space in terms of foot positions, which essentially preserves the robust action space based on reduced-order models. Along with the analytical controller that maps actions into joint torques, the proposed learning approach facilitates natural action representations for what the bipedal locomotion policy should learn in the most meaningful way.

Further work will focus on how to add center of mass control as a part of the action space into the current learning framework. The resulting action space will include both positions for center of mass (pelvis) and the two feet. In this way, control policies can use setpoints with impedances as control actions to interact with the world through practical actuators. Both setpoints and impedances together can create fast and approximately desirable responses, along with other higher-level planning and control methods to re-plan and re-adjust in according to the full state information of the robot.

Acknowledgments

This work is supported by the NSF Grant No. IIS-1849343, DGE-1314109, and DARPA Contract W911NF-16-1-0002.

References

[1] Taylor Apgar, Patrick Clary, Kevin Green, Alan Fern, and Jonathan Hurst. Fast Online Trajectory Optimization for the Bipedal Robot Cassie. *Robotics: Science and Systems XIV*, 2018.

[2] Michael Mistry and Ludovic Righetti. Operational space control of constrained and underactuated systems. *Robotics: Science and Systems*, 2012.

[3] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal Policy Optimization Algorithms. 2017. URL http://arxiv.org/abs/1707.06347.