Robust 3D Bipedal Locomotion through Reinforcement Learning and Feedback Regulation

1st Guillermo A. Castillo
Electrical and Computer Engineering
The Ohio State University
Columbus, OH, USA
castillomartinez.2@osu.edu

2nd Bowen Weng
Electrical and Computer Engineering
The Ohio State University
Columbus, OH, USA
weng.172@osu.edu

3rd Wei Zhang
SUSTech Institute of Robotics
Southern University of Technology (SUSTech)
Shenzhen, China
zhangw3@sustech.edu.cn

4th Ayonga Hereid
Mechanical and Aerospace Engineering
The Ohio State University
Columbus, OH, USA
hereid.1@osu.edu

I. Background

Robust bipedal locomotion remains one of the main challenges in the field of legged robotics [1]. In practice, the robustness of the control policy is determined by (i) the capability of handling various external disturbances, (ii) maintaining stable gaits while operating under various terrain conditions, and (iii) accomplishing the sim-to-real transfer with as little effort as possible. To achieve robust bipedal locomotion, model-based methods, learning-based (model-free) methods, and a combination of both have been proposed in the literature. While model based methods often rely on a simplified model and require extensive control gain tuning, most learning-based solutions require a large amount of data. In this work, we build upon the work in [2] to implement a cascade control structure that couples both learning-based and model-based approaches to realize robust locomotion in 3D bipedal robots.

II. Methods

The proposed framework, shown in Fig. 1, addresses the 3D walking problem from two hierarchical levels, trajectory planning and feedback regulation. The first block of the framework uses a neural network-based trajectory planner that takes a reduced-dimensional representation of the robot states to compute a set of coefficients $\boldsymbol{\alpha}$ that parameterize joint reference through 5th Bézier Polynomials. The decoupled structure of the network allows to run the trajectory planning and low-level controller at different frequencies, which is important for real-time implementation of the neural network policy in hardware.

The feedback regulation uses the robot's pelvis velocities and torso orientation to modify the joint reference trajectories. It compensates for the uncertainty in the robot's model and the environment, improving the robustness of the walking gait. The joint-level PD controllers are used to track the compensated reference joint trajectories.

III. Results and Discussion

The proposed framework allows the transference of walking policies learned in simulation to the real hardware with minimal tuning and enhanced robustness. We show that this controller structure is robust enough to mitigate the uncertainty in the robot's dynamics caused by the mismatch between the simulation and real hardware. By thorough experiments on hardware, we show the robustness of the learned controller against adversarial disturbances applied to the robot. In addition, we show in Fig.1 (b) the policy's capability to adapt to various terrains without need of training for such challenging scenarios.

To the best of our knowledge this is the first time that a policy learned in simulation is successfully transferred to hardware to realize stable and robust dynamic walking gaits for the Digit robot[1].

IV. Conclusions

By combining a sample efficient learning structure with intuitive but powerful feedback regulations in a cascade structure, we decouple the learning problem into two stages that work at a different frequency to facilitate the implementation of the controller in the real hardware. The end result are robust policies learned from scratch that are transferred successfully to hardware with minimal tuning. Experimental results show the controller realizes stable walking gaits that are robust to external disturbances and challenging terrains without being trained under those conditions.

References

[1] J. W. Grizzle, C. Chevallereau, A. D. Ames, and R. W. Sinnet, "3D bipedal robotic walking: models, feedback control, and open problems," in IFAC Symposium on Nonlinear Control Systems, Bologna, Sep. 2010

[2] G. A. Castillo, B. Weng, W. Zhang and A. Hereid, "Hybrid Zero Dynamics Inspired Feedback Control Policy Design for 3D Bipedal Locomotion using Reinforcement Learning," *2020 IEEE ICRA*, 2020, pp. 8746-8752, doi: 10.1109/ICRA40945.2020.91971

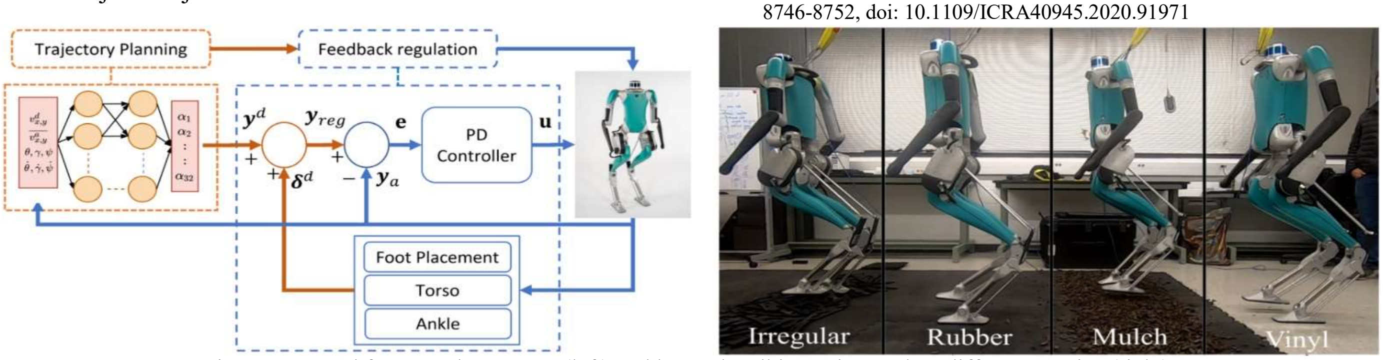

Fig. 1. Proposed framework structure (left) and learned walking gait tested on different terrains (right).

[1]https://www.youtube.com/watch?v=j8KbW-a9dbw