ETH zürich

CRL Computational Robotics Lab

Trajectory Optimization and Policy Distillation for Agile Control of Hybrid Mobile Robots

Motivation:

Hybrid mobile robots — such as those combining wheels and legs — offer powerful locomotion capabilities by uniting the efficiency of wheeled travel with the versatility of legged movement. However, their control poses unique challenges due to underactuated dynamics, terrain interactions, and the tight coupling between balance, propulsion, and steering.

While trajectory optimization (TO) and model predictive control (MPC) methods can generate high-quality control actions that adapt to the environment, their computational demands can limit practical deployment. On the other hand, neural policies are fast at inference but often struggle with training and reward shaping. This project proposes to combine the best of both worlds: use TO or MPC to control the system and simultaneously distill a policy that mimics these control signals, enabling efficient, deployable policies trained directly on real robot rollouts.

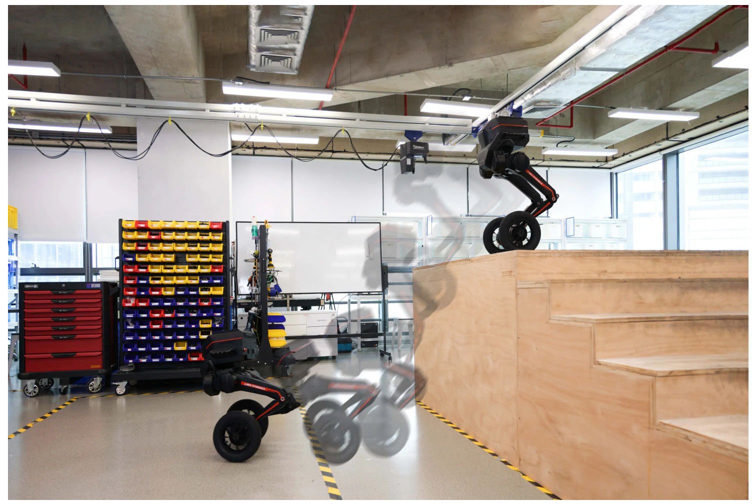

Figure 1: LIMX Dynamics TRON1 robot—a bipedal wheeled platform with 3 actuated joints and one wheel per leg (4 DoFs per leg total)

Approach:

The project will explore a control-learning pipeline that couples trajectory optimization or model predictive control (TO/MPC) with online policy distillation. First, a TO or MPC controller will be implemented to generate high-quality control actions. This controller can be either sampling-based (e.g., MPPI [1], CEM [2], or diffusion-based annealing [3, 4]) for use in non-differentiable simulators such as MuJoCo Playground [5] or Brax [6], or gradient-based when using differentiable simulators, such as warp [7]. Task-specific cost functions will be designed to support stable locomotion, agile maneuvering, and robustness to disturbances. While the TO/MPC controller runs, its trajectories and actions will be logged in real time and used to train a neural policy through supervised learning, following principles similar to the PODS framework [8]. The policy aims to imitate the optimizer's output, and the training process may alternate between trajectory-space and parameter-space updates to enhance convergence and stability.

General Details:

Applicants should have programming experience in both Python and C++, along with a basic understanding of robot simulation and trajectory optimization. Familiarity with reinforcement learning is a plus but not strictly required. We will be using LIMX Dynamics TRON1 robot as the main development platform Fig. 1.

Interested?

The project will be supervised by Jin Cheng and Prof. Dr. Stelian Coros. Please reach out to Jin (jin.cheng@inf.ethz.ch) with your CV and transcripts.

References

[1] Juan Alvarez-Padilla, John Z Zhang, Sofia Kwok, John M Dolan, and Zachary Manchester. Real-time whole-body control of legged robots with model-predictive path integral control. *arXiv preprint arXiv:2409.10469*, 2024.

[2] Taylor Howell, Nimrod Gileadi, Saran Tunyasuvunakool, Kevin Zakka, Tom Erez, and Yuval Tassa. Predictive sampling: Real-time behaviour synthesis with mujoco. *arXiv preprint arXiv:2212.00541*, 2022.

[3] Haoru Xue, Chaoyi Pan, Zeji Yi, Guannan Qu, and Guanya Shi. Full-order sampling-based mpc for torque-level locomotion control via diffusion-style annealing. *arXiv preprint arXiv:2409.15610*, 2024.

[4] Chaoyi Pan, Zeji Yi, Guanya Shi, and Guannan Qu. Model-based diffusion for trajectory optimization. *Advances in Neural Information Processing Systems*, 37:57914–57943, 2024.

[5] Kevin Zakka, Baruch Tabanpour, Qiayuan Liao, Mustafa Haiderbhai, Samuel Holt, Jing Yuan Luo, Arthur Allshire, Erik Frey, Koushil Sreenath, Lueder A Kahrs, et al. Mujoco playground. *arXiv preprint arXiv:2502.08844*, 2025.

[6] C Daniel Freeman, Erik Frey, Anton Raichuk, Sertan Girgin, Igor Mordatch, and Olivier Bachem. Brax–a differentiable physics engine for large scale rigid body simulation. *arXiv preprint arXiv:2106.13281*, 2021.

[7] Miles Macklin. Warp: A high-performance python framework for gpu simulation and graphics. `https://github.com/nvidia/warp`, March 2022. NVIDIA GPU Technology Conference (GTC).

[8] Miguel Angel Zamora Mora, Momchil Peychev, Sehoon Ha, Martin Vechev, and Stelian Coros. Pods: Policy optimization via differentiable simulation. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 7805–7817. PMLR, 18–24 Jul 2021. URL `https://proceedings.mlr.press/v139/mora21a.html`.