Deep Whole-Body Control
Learning a Unified Policy for Manipulation and Locomotion
Zipeng Fu*  Xuxin Cheng*  Deepak Pathak

Motivation: whole-body >> modular, low-cost >> expensive hardware

TL;DR: learning an end-to-end unified policy for whole-body control of a custom-built low-cost quadruped mobile manipulator

Regularized Online Adaptation for Sim-to-Real Transfer

$$L(\theta_\tau, \theta_\mu, \theta_\phi) = -J(\theta_\tau, \theta_\mu) + \lambda ||z^\mu - sg(z^\phi)||_2 + ||sg(z^\mu) - z^\phi||_2$$

Advantage Mixing for Policy Learning

$$J(\theta_\tau) = \frac{1}{|D|} \sum (s_t, a_t^\tau) \in D \log \pi(a_t^\tau | s_t) (A_{manip} + \beta A_{lco}) + \log \pi(a_t^{EE} | s_t) (\beta A_{manip} + A_{lco})$$

Hardware Setup Comparison

Spot with Arm
- > $100K
- No low-level control APIs

Anymal B + Kinova
- Bully (~40kg)
- Expensive

Ours (Got + WidowX 250a)
- $6.3K hardware cost
- Fully onboard compute + power

Diverse Task Set

Large Whole-Body Workspace

Videos & Code
https://maniploco.github.io