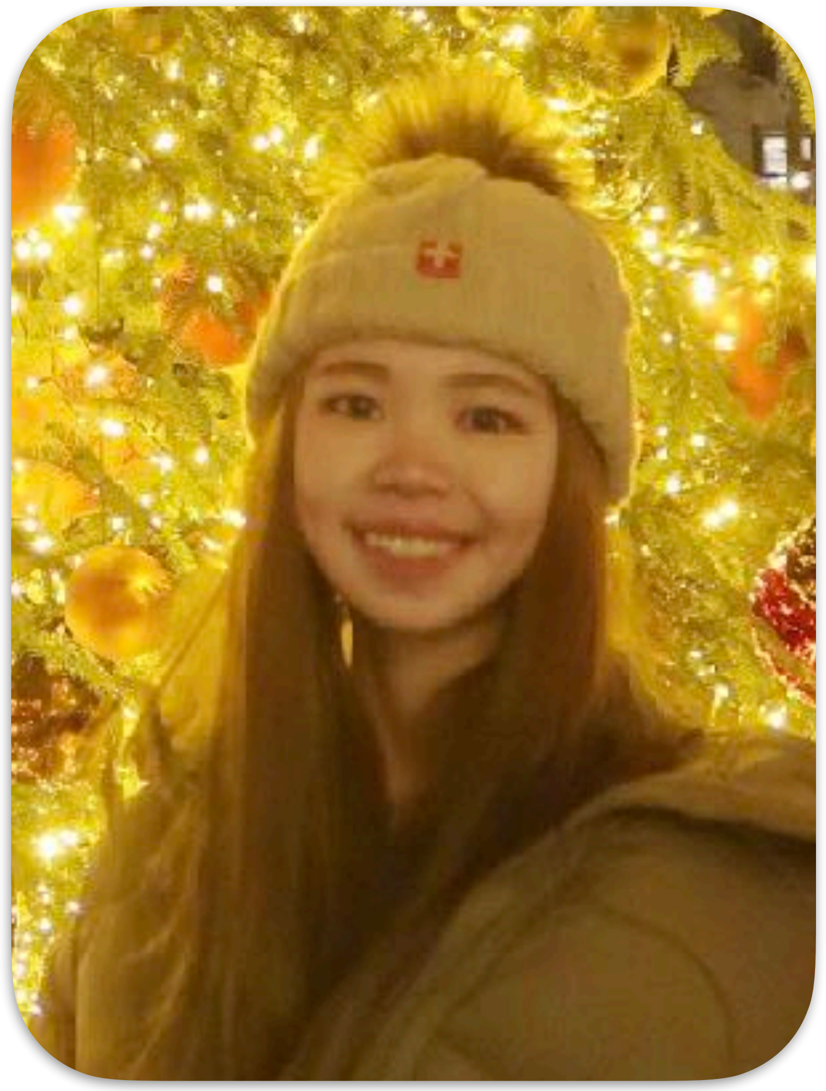
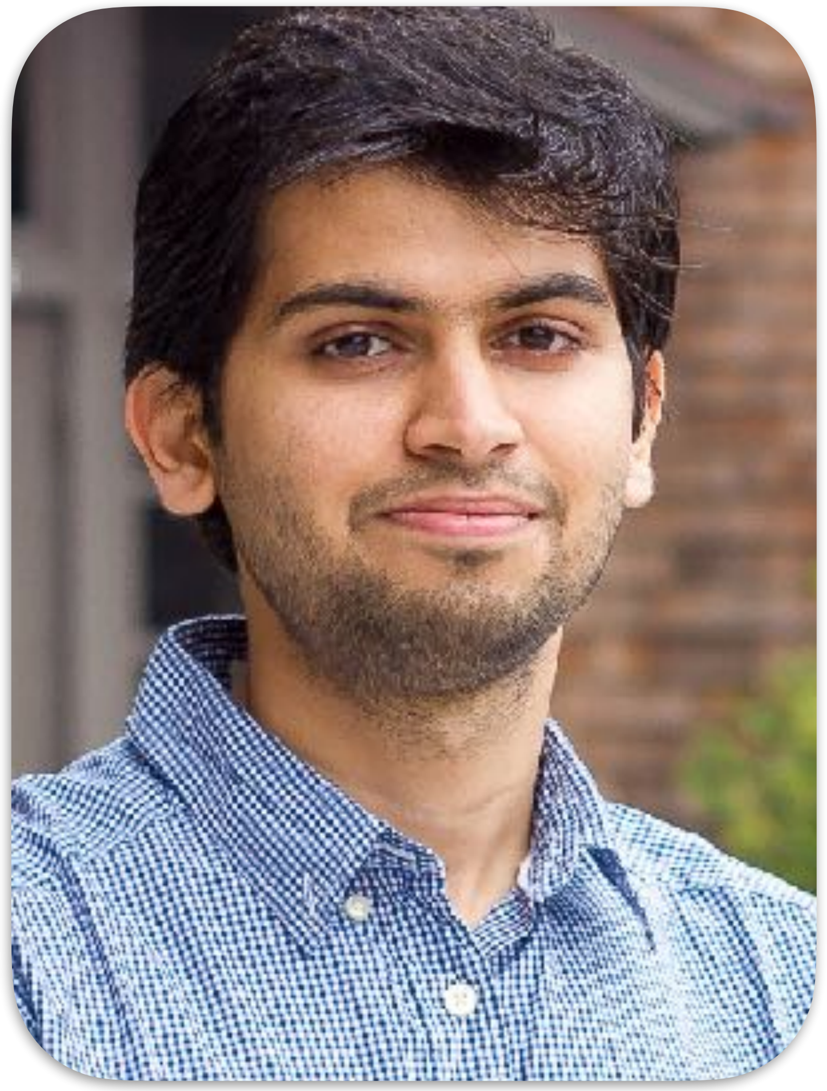
Extreme Parkour with Legged Robots



Xuxin Cheng*   Kexin Shi*   Ananye Agarwal   Deepak Pathak

* denotes equal contribution

Carnegie Mellon University

THE ROBOTICS INSTITUTE

https://extreme-parkour.github.io/

Robot Parkour

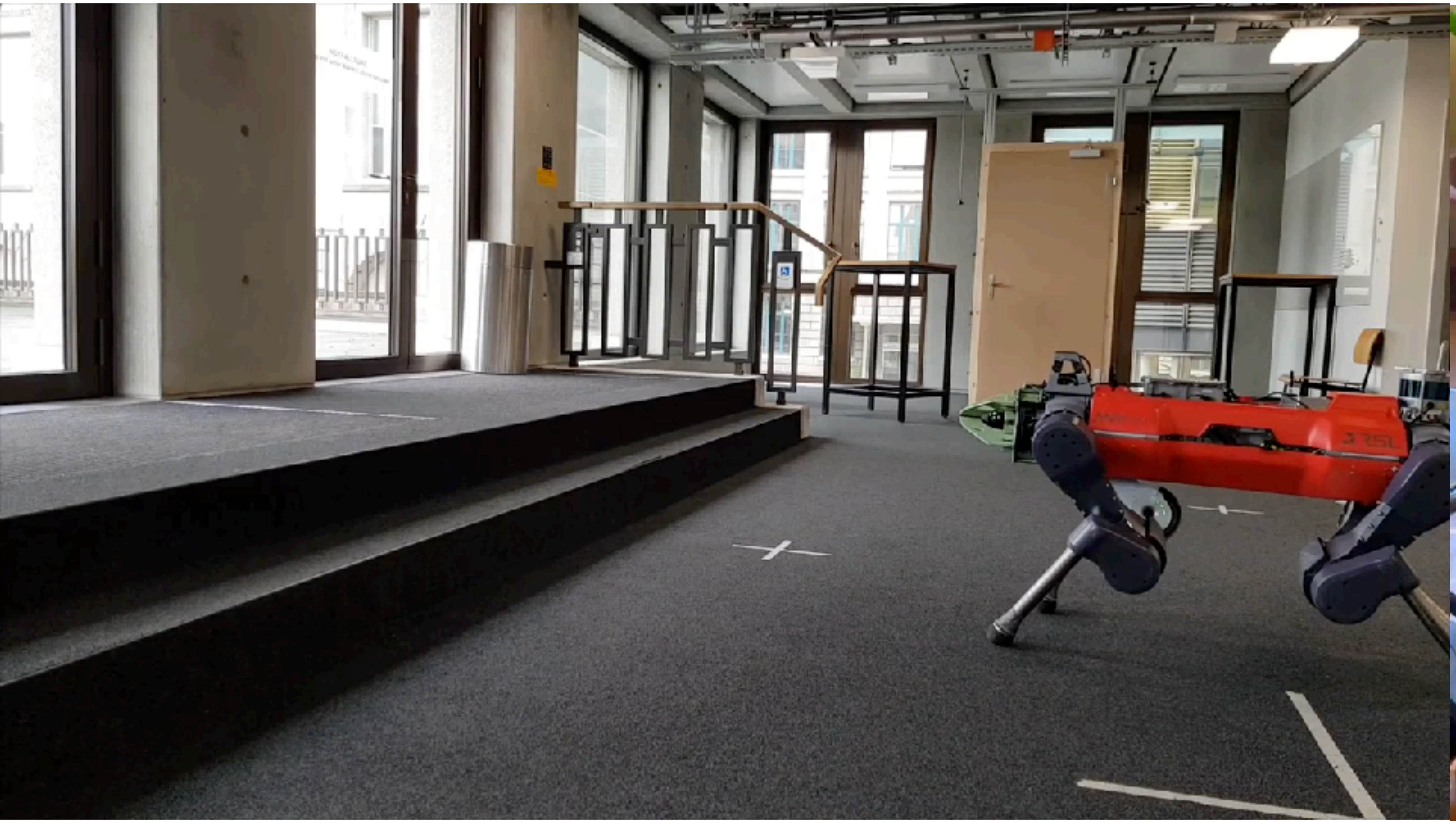
- Classical model-based control.
- Should know the layout of obstacles in advance.
- Can not generalize to different layout.

Human Parkour

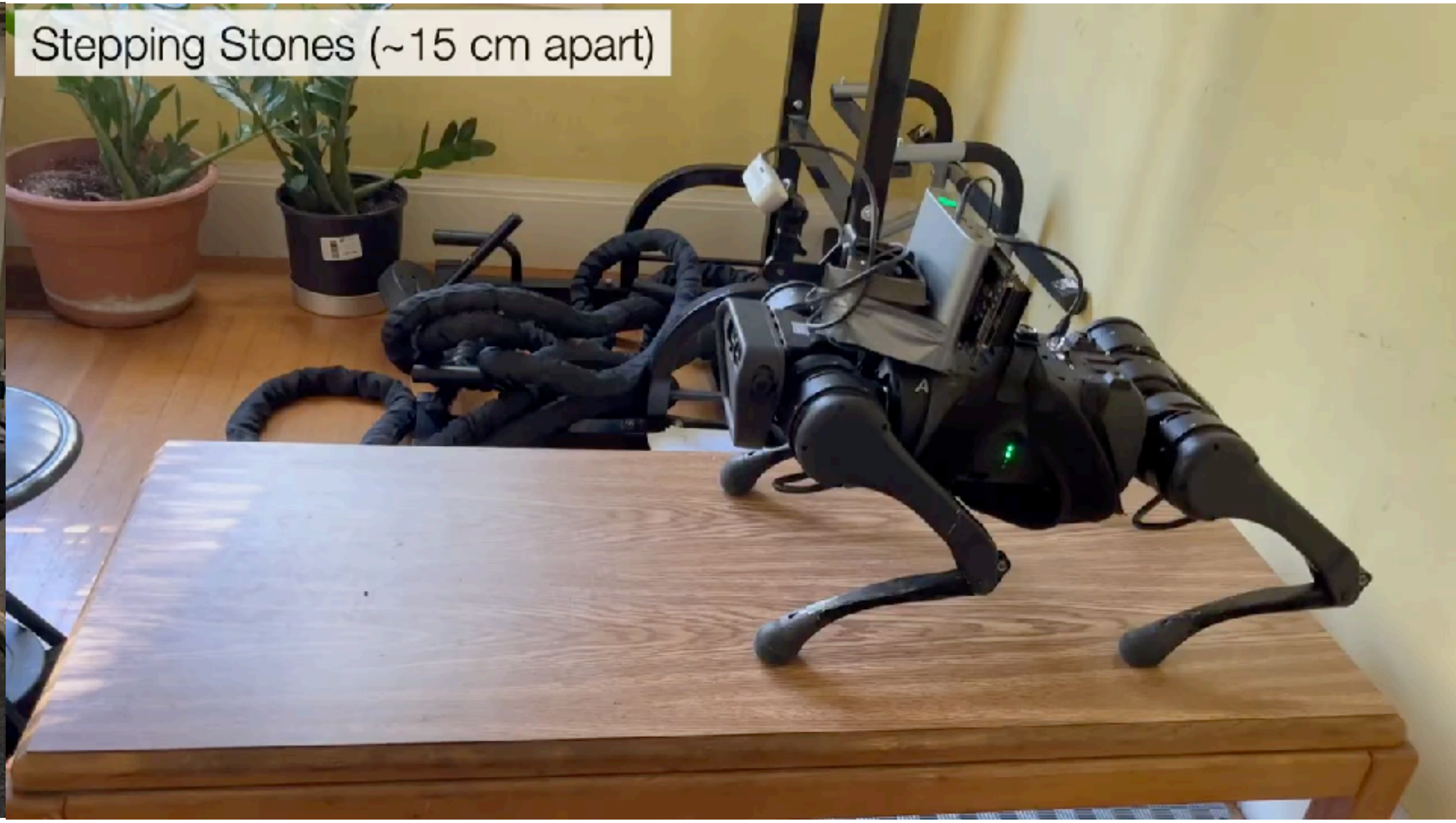
- Leverage egocentric vision information
- Develop a eye-muscle coordination
- Can self-adjust to different terrains

Vision Locomotion

Miki et.al "Learning robust perceptive locomotion for quadrupedal robots in the wild" Science Robotics 2022



Agarwal et.al "Vision Locomotion using Egocentric Vision" CoRL 2022



Could we learn a single vision-based policy for agile parkour?

Challenges

Vision Locomotion:

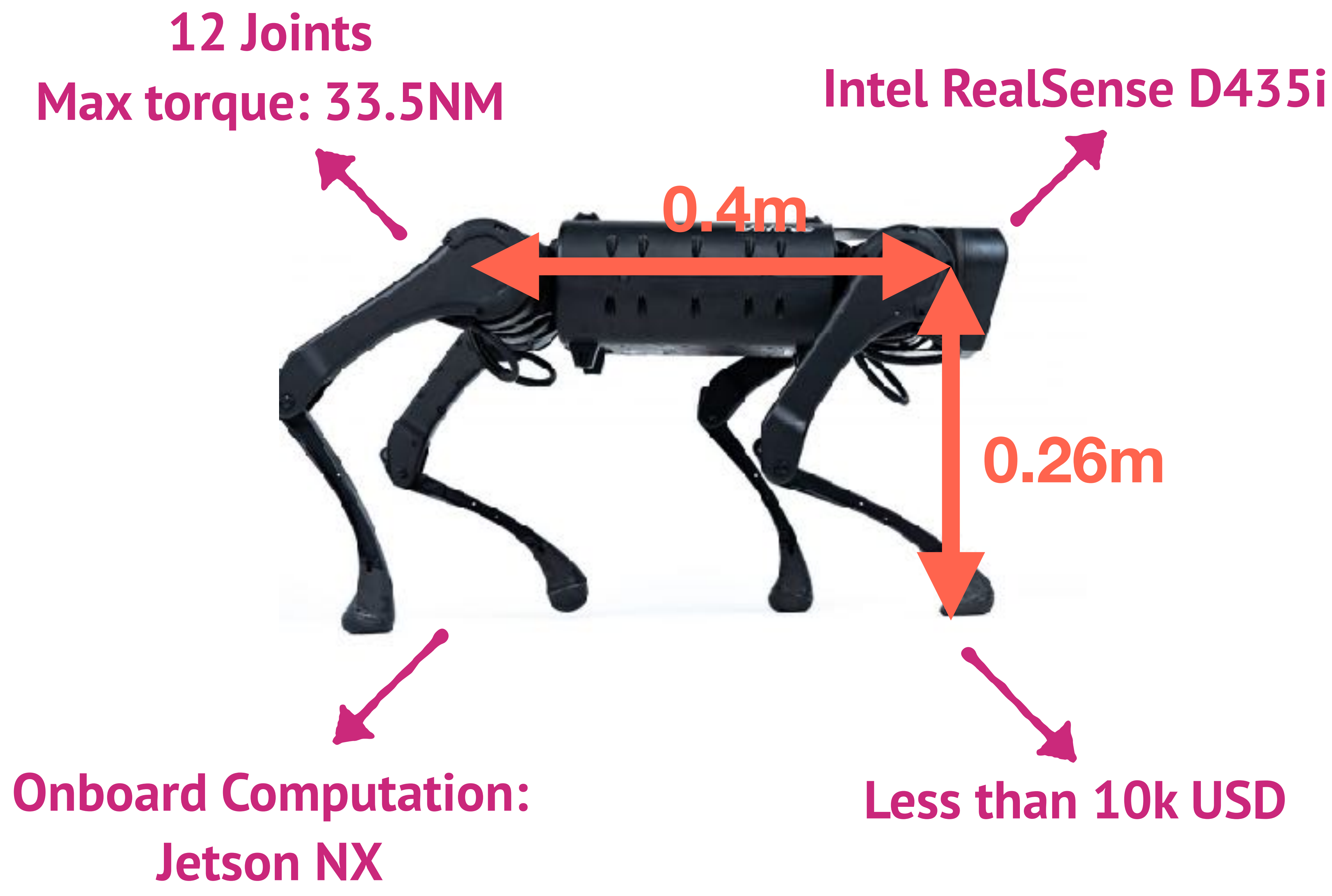
- The action is noisy and laggy.
- The camera has artifacts, latency and jitter.

……

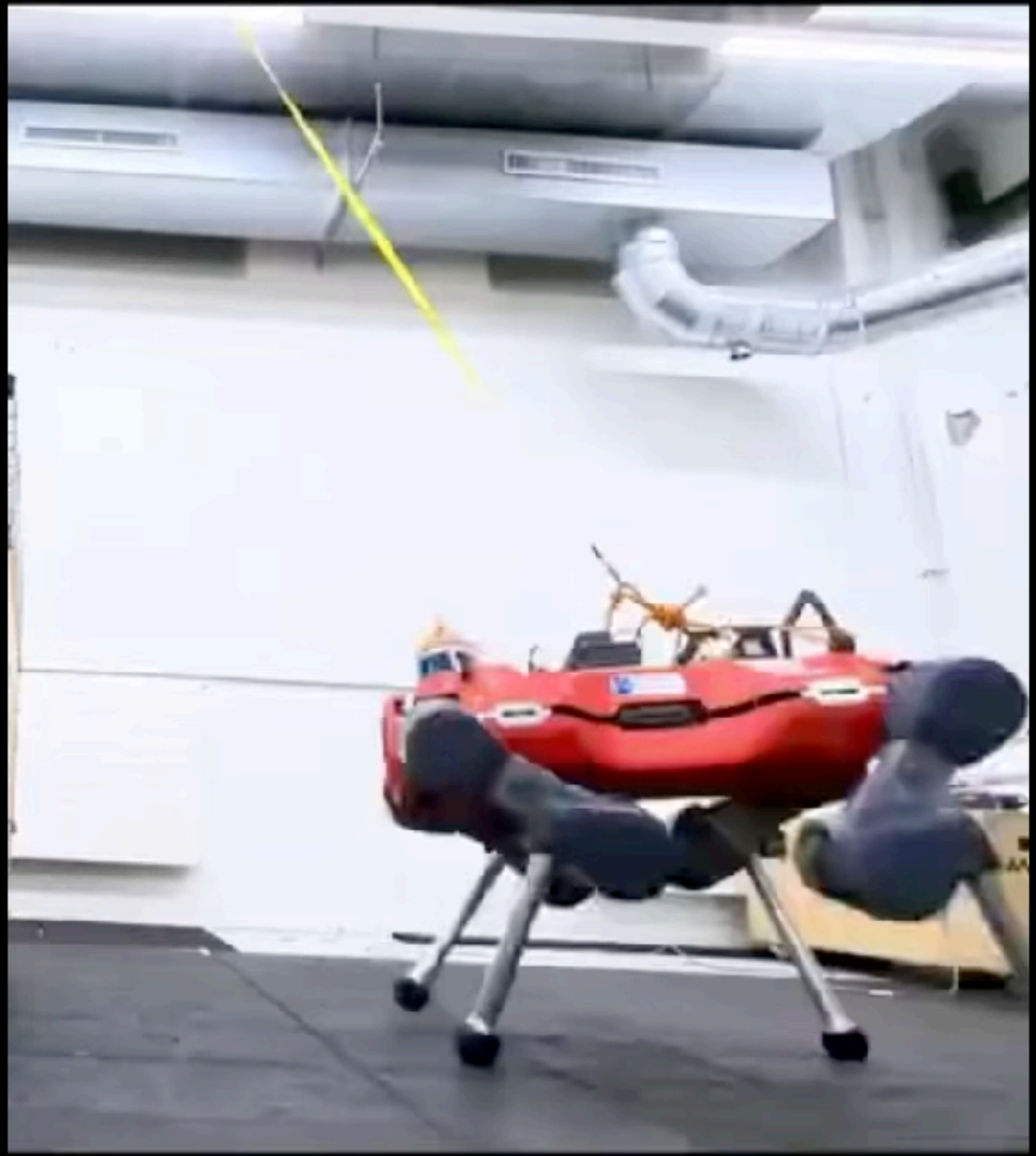
Agile Parkour:

- Extreme motion needs precise control.
- The heading should be adjusted by robot itself.
- Walking in different styles is still under exploration.

……



Rudin et.al "ANYmal Parkour: Learning Agile Navigation for Quadrupedal Robots" 2023



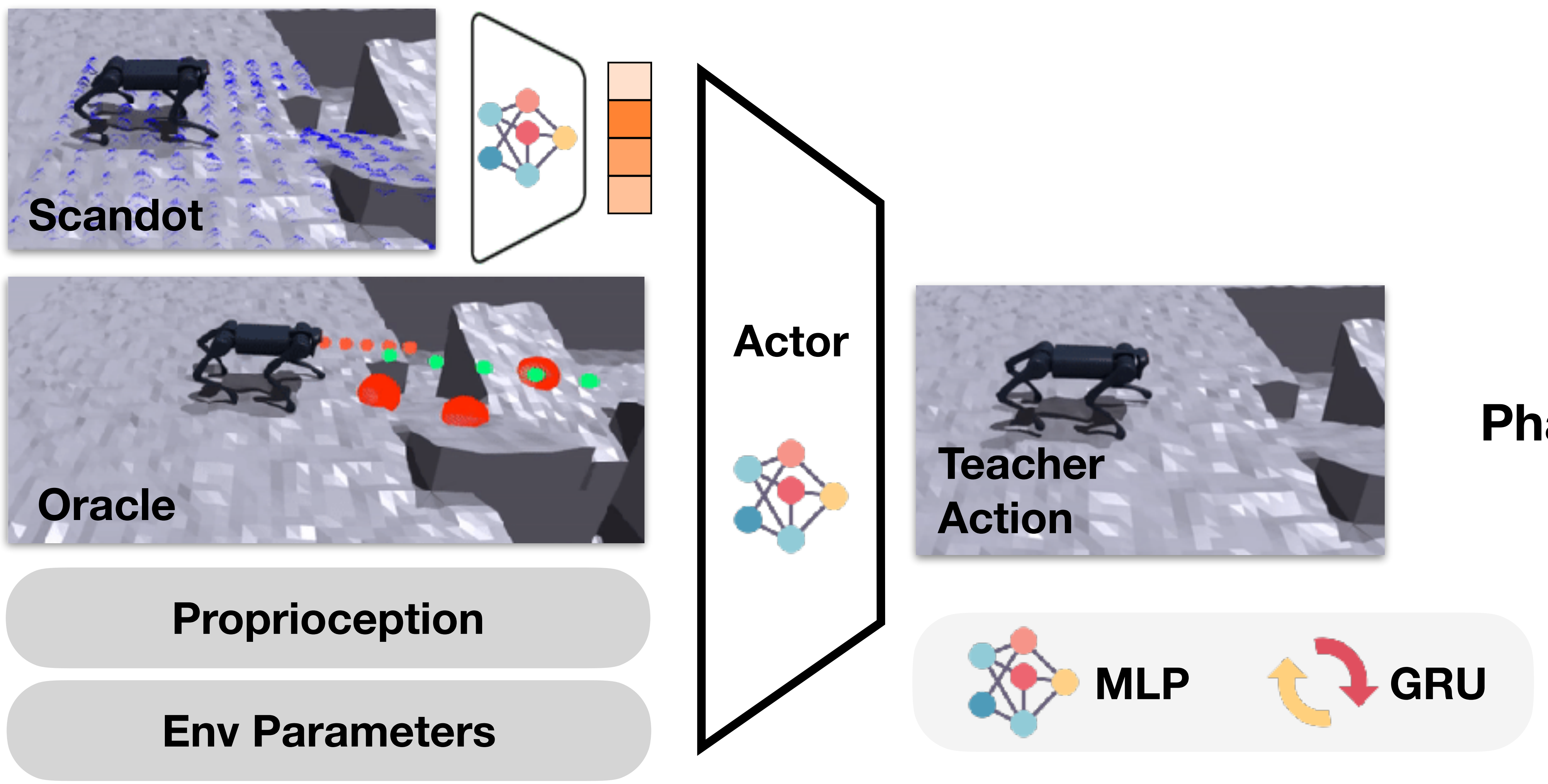
- Task-specific policies.
- A high-level navigation module to select skill.
- Rely on elevation maps reconstructed by 6 depth sensors.
- 2x its height and 1.5x its length.

Robot Parkour Learning

- Complex curriculum (3 stages).
- Individual locomotion policies then distilled into one vision policy.
- Simplified obstacle abstractions as privileged visual information.
- 1.5x its height and 1.5x its length.

twitter: @ziwenzhuang_leo
@zipengfu

Phase 1 - Reinforcement Learning from Scandots



Phase 1

Proprioception

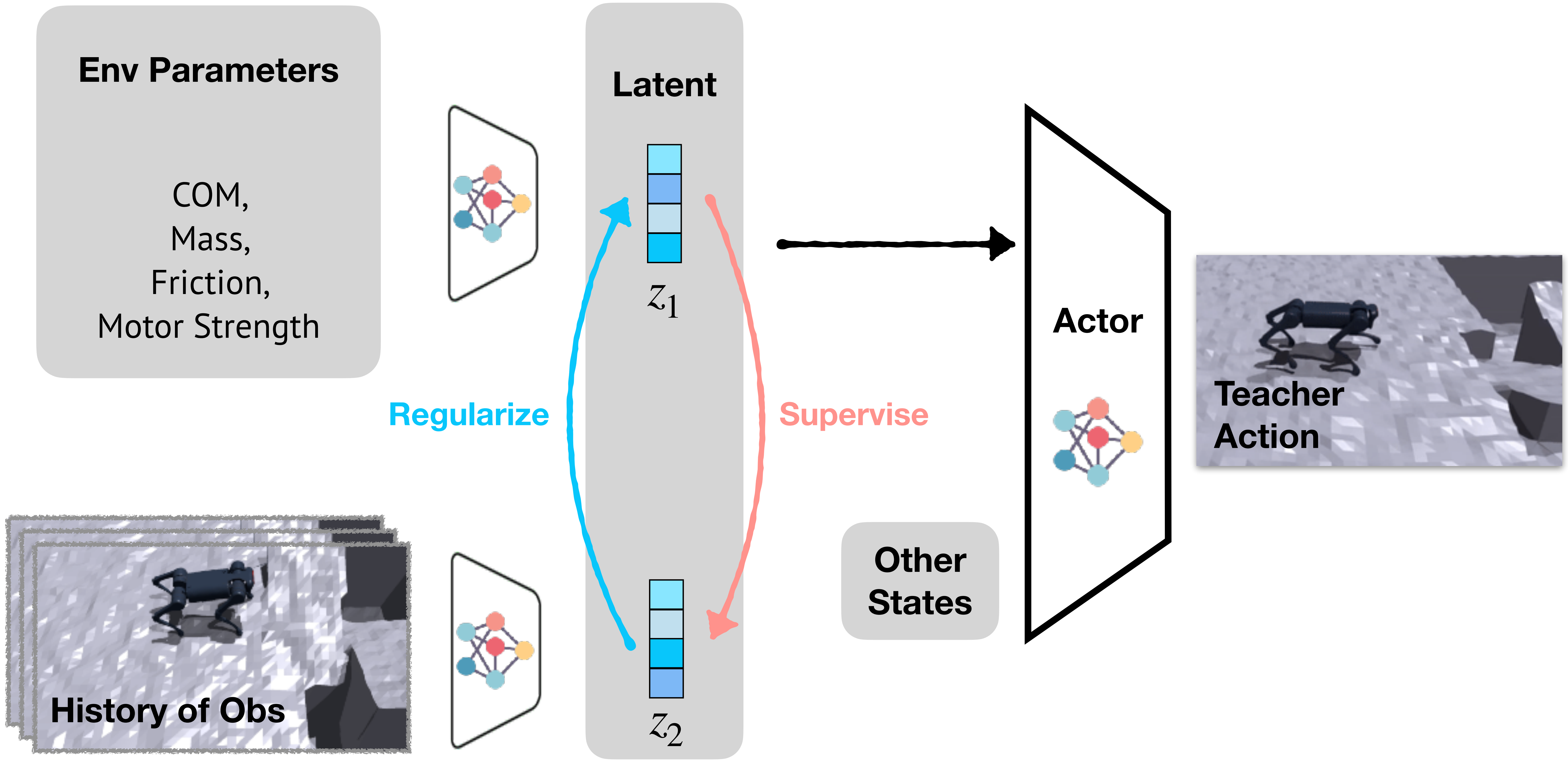
Angular Velocity,
Roll, Pitch,
Walking Flag,
Joint Position,
Joint Velocity,
Last Action,
Contact...

Env Parameters

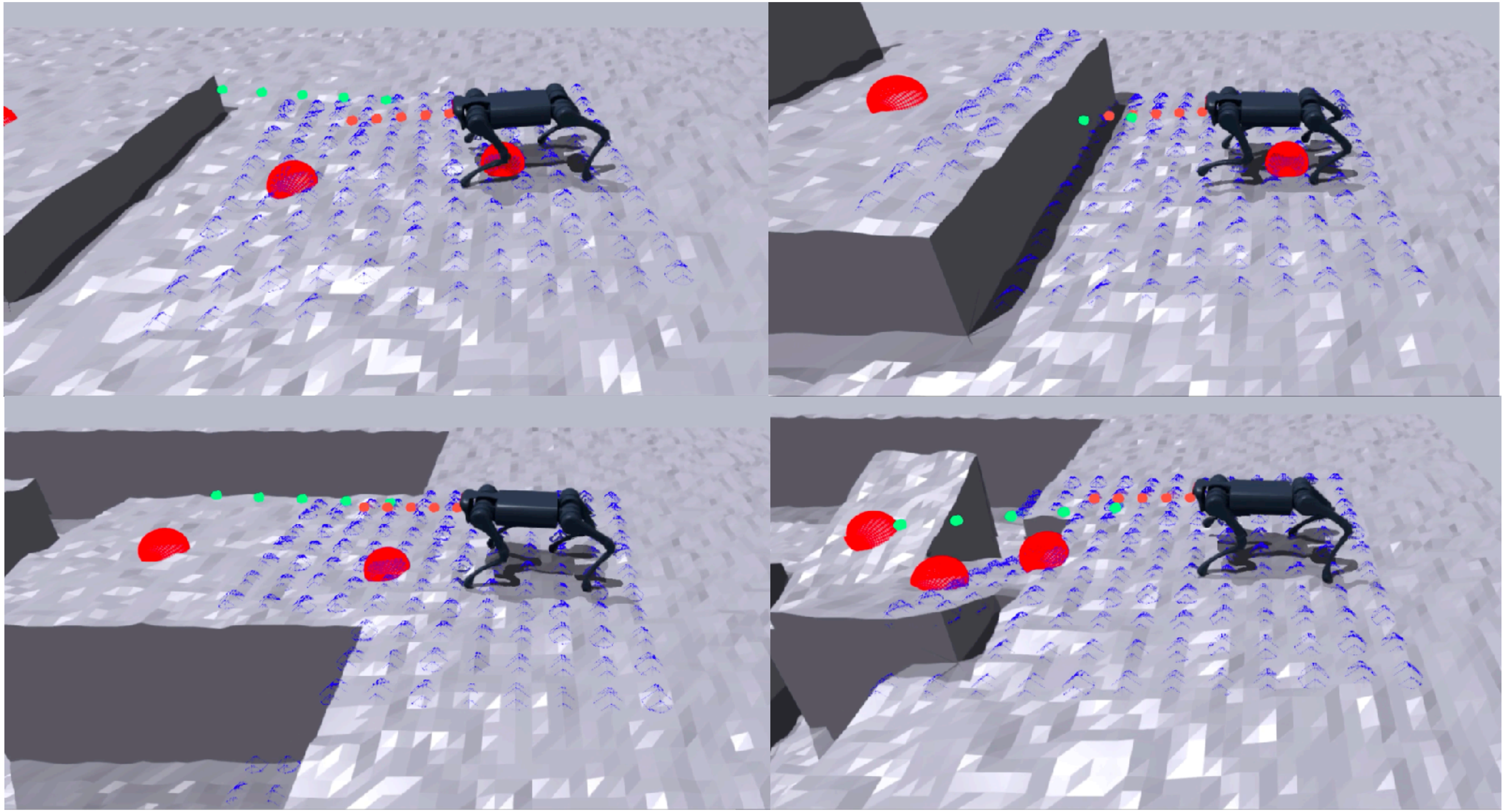
COM,
Mass,
Friction,
Motor Strength

How could we estimate environment parameters in real world?

Regularized Online Adaption (ROA)



Phase 1 - Reinforcement Learning from Scandots

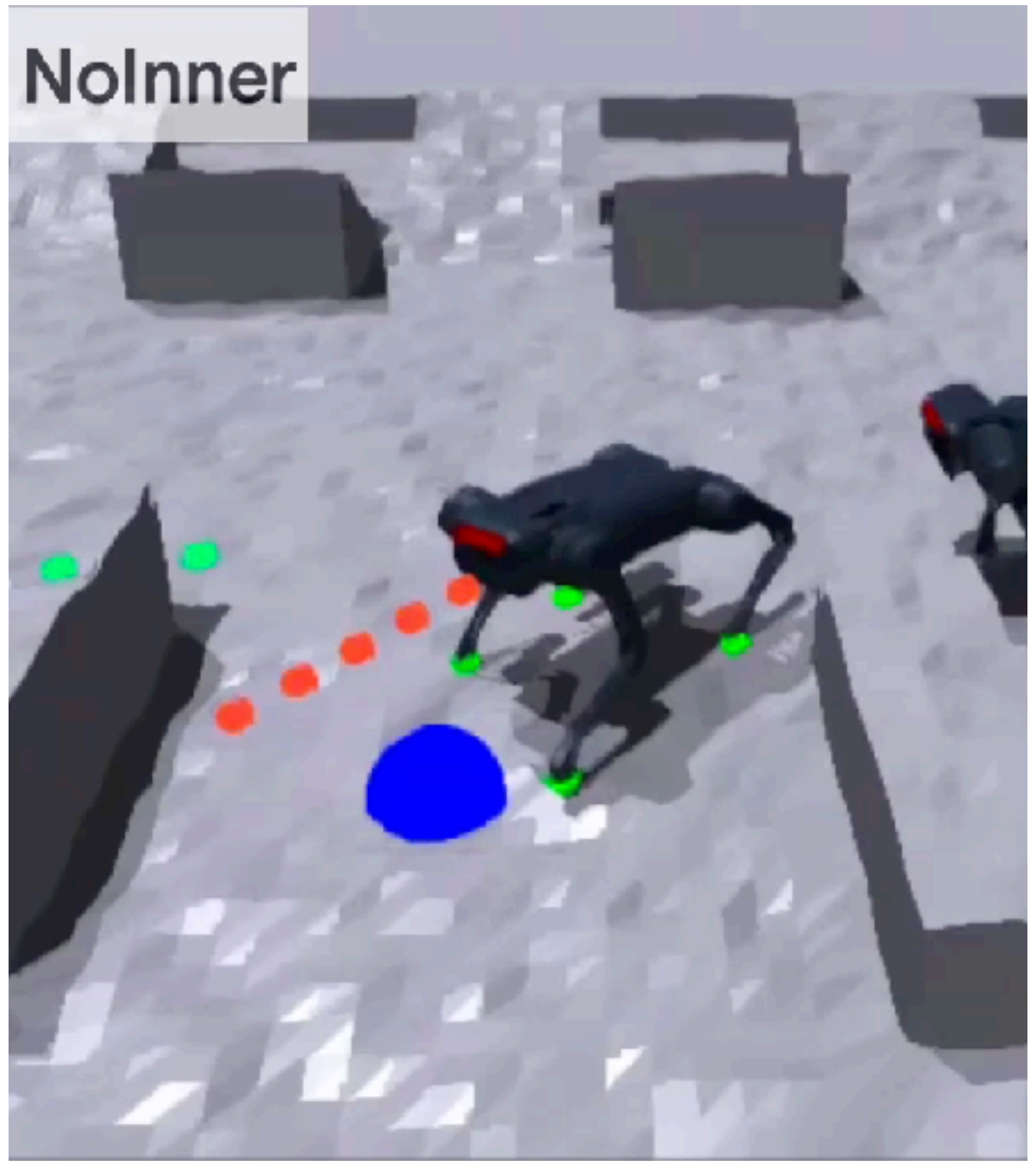


Rewards

- **Velocity Tracking Reward:** Encourage to track commanded linear velocity

$$r_{tracking} = v_x^{\text{cmd}} - \left|v_x^{\text{cmd}} - v_x\right| - \left|\omega_z^{\text{cmd}} - \omega_z\right|$$

Agarwal et.al "Vision Locomotion using Egocentric Vision" CoRL 2022

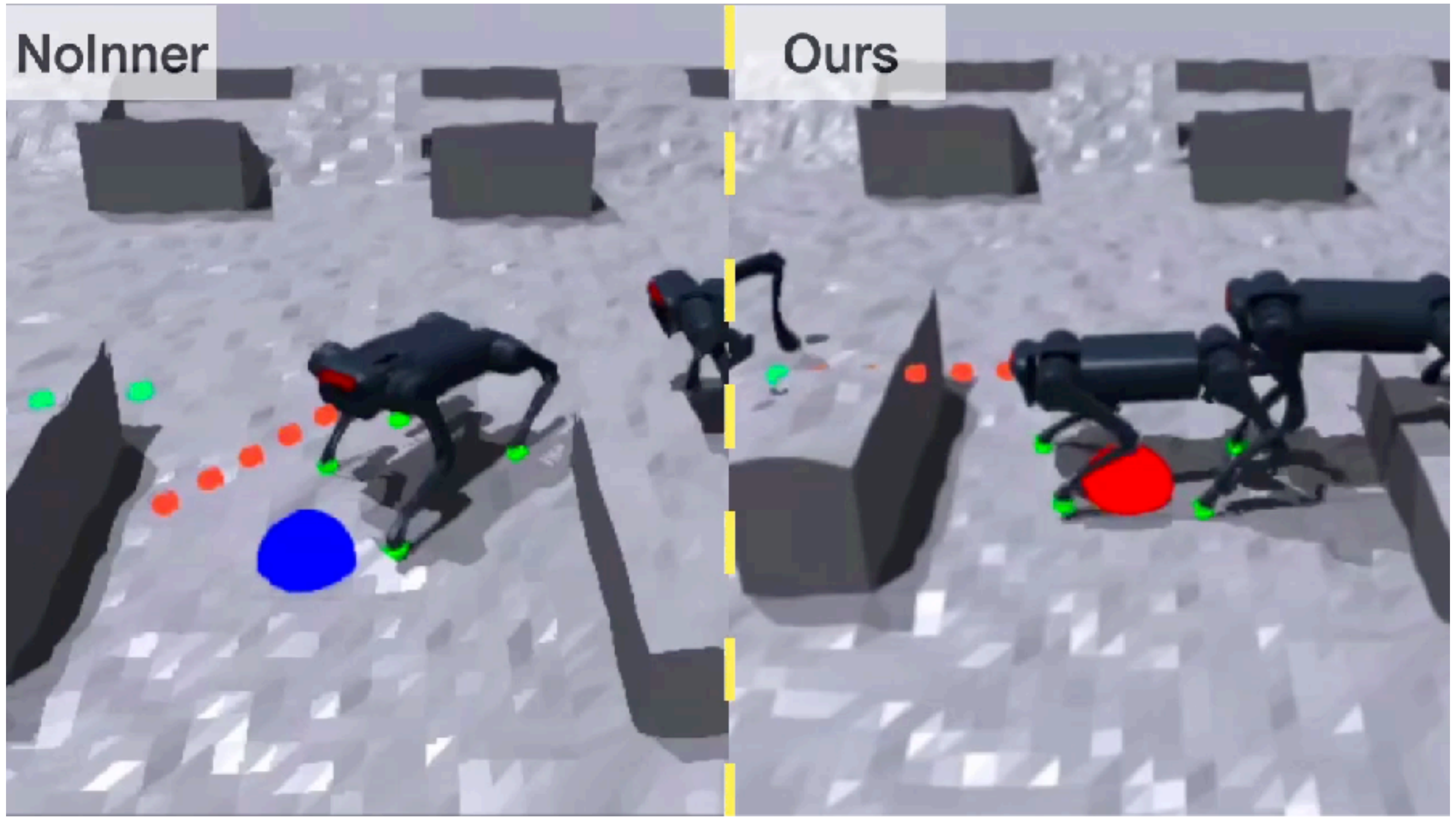


Unified Rewards

- **Velocity Tracking Reward:** Encourage to track heading

$$r_{tracking} = \min(\langle \mathbf{v}, \hat{\mathbf{d}}_w \rangle, v_{cmd})$$

where $\hat{\mathbf{d}}_w = \frac{\mathbf{p}-\mathbf{x}}{\|\mathbf{p}-\mathbf{x}\|}$ is the next target direction.



Unified Rewards

- **Velocity Tracking Reward:** Encourage to track heading

$$r_{tracking} = \min(\langle \mathbf{v}, \hat{\mathbf{d}}_w \rangle, v_{cmd})$$

where $\hat{\mathbf{d}}_w = \frac{\mathbf{p}-\mathbf{x}}{\|\mathbf{p}-\mathbf{x}\|}$ is the next target direction.

- **Clearance Reward:** Penalize dangerous footprints

$$r_{clearance} = -\sum_{i=0}^{4} c_i \cdot M[p_i]$$

where M is a boolean function which is 1 iff the point p_i lies within 5cm of an edge.

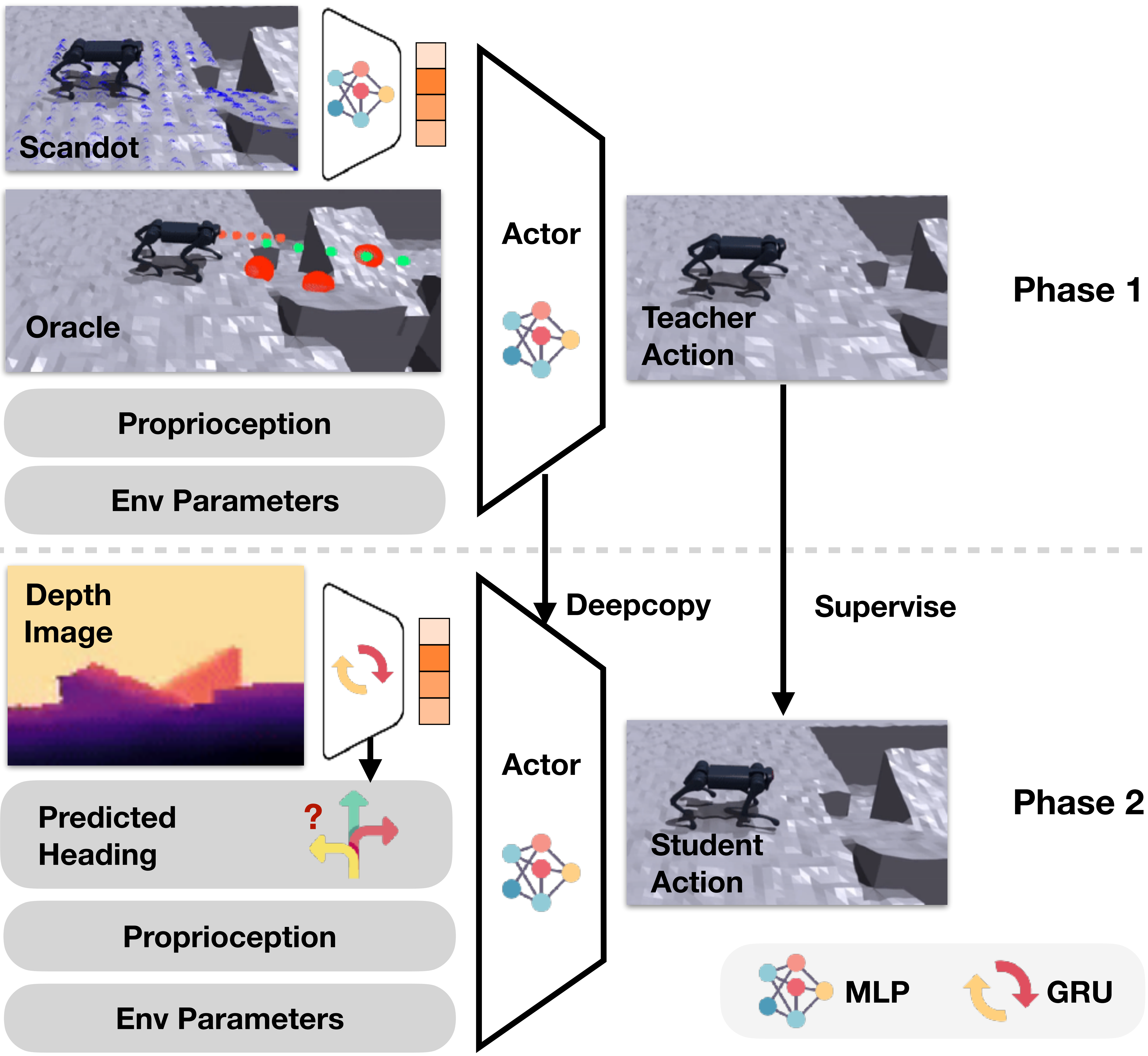
- **Stylized Reward:** Encourage handstand walk

$$r_{stylized} = W \cdot \left[0.5 \cdot \langle \hat{\mathbf{v}}_{fwd}, \hat{\mathbf{c}} \rangle + 0.5\right]^2$$

where W is sampled randomly in $\{0, 1\}$ in training and controlled via joystick in deployments.

$\hat{\mathbf{c}} = [0, 0, -1]^T$ when do a handstand.

Phase 2 - Distilling Direction and Exteroception



Mixture Teacher & Student (MTS)

$$obs_\theta = \begin{cases} \theta pred, & \text{if } |\theta pred - \hat{\mathbf{d}}_w| < 0.6 \\ \hat{\mathbf{d}}_w, & \text{otherwise} \end{cases}$$

- Bootstrap imitation learning
- Better convergence

Extreme Results: Long Jump


Emergent Behavior: Feet Adjustment before Jump


Long Jump without Clearance Reward

Extreme Results: High Jump


Emergent Behavior: Remedy from Rear Leg
VENTION
VENTION
16 inch
16 inch


Handstand on Grass


Blind Handstand Downstair


Parkour Course
Generalize to combination of different obstacles

Without direction distillation

Control via Joystick


3min Pressure Test
8X

Step in Trail


Consecutive Jumps
AI

Consecutive Straw Hurdles

More: extreme-parkour.github.io

Quantitative Results

Conclusions

- A novel **dual distillation** method for distilling both **agile motor commands** and rapidly fluctuating **heading directions** from depth images.

- A simple yet effective **inner-product reward** design principle for general robot base motion acquisition.

- A new landmark in learning-driven parkour with **extreme behaviors**. (High jumps that are **2x** the height of the robot, long jumps that are **2x** the length of the robot, **handstand**, and jumping over **titled ramps**.

- **Less than 20h** training on **a single RTX 3090**.

Future Works

- Tightly couple perception, locomotion and local navigation for other applications, e.g. obstacle avoidance.

- Attach an arm on the body to do locomotion and manipulation simultaneously instructed a vision-language-action model.

Standing on the Shoulders of Giants

Huge Thanks to my Great Collaborators and Inspiring Friends!


Davide Scaramuzza


Deepak Pathak


Yunlong Song


Mathias Gehrig


Nico Messikommer


Robert Penicka


Songyou Peng


Ananye Agarwal


Paul-Edouard Sarlin


Jiacheng Qiu


Chunwei Xing


Xinyu Sun


Rulin Shao


Jiawei Fu


Jiaxu Xing


Chao Ni


Xuxin Cheng


Tianyi Zhang


Heng Yu


Kangni Liu


Yifei Liu


Weirong Chen


Yidan Gao


Fenglong Song

Thank you!