EPFL

Extreme Parkour with Legged Robots

Group 24

Teo Halevi
William Martin
Loïs Marchioni

Article

2024 IEEE International Conference on Robotics and Automation (ICRA), May 13-17, 2024. Yokohama, Japan

Xuxin Cheng, Kexin Shi, Ananye Agarwal, Deepak Pathak ; Carnegie Mellon University ; University of Zurich

Core Idea & Motivation

- **Goal:** Enable low-cost legged robots to perform precise, dynamic parkour
- **Challenge:** Imprecise actuation & low-quality sensing
- **Inspiration:** Humans master parkour through experience, not perception
- **Aim:** Replicate learning-driven adaptability in affordable robots
- Classical control methods → fail under such uncertainty

Article

Key aspects

Robot characteristics

Robot: Unitree A1 – low-cost 12-joint quadruped, capable of bipedal motion

Dimensions: thigh joint height ≈ 26 cm, body length ≈ 40 cm

Some extreme parkour maneuvers:
- High jumps over obstacles 2× height
- Long jumps across gaps 2× length
- Handstands with seamless transitions
- Running over tilted ramps

Onboard hardware:
- Intel RealSense D435 depth camera (~10 Hz)
- Jetson NX computer
- Depth backbone at 10 Hz, base policy at 50 Hz
- Depth frames (58 × 87) cropped, downsampled, passed to policy with latency

Training: End-to-end policy trained on a single RTX 3090 GPU in < 20 hours

Key aspects

Learning Approach

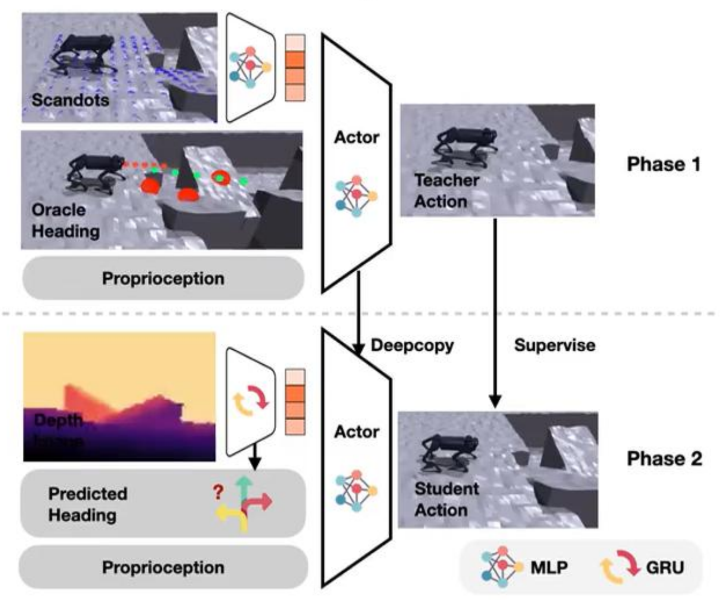
- Single end-to-end neural policy trained via large-scale RL in simulation
- Maps front-facing depth images → motor commands
- No separate perception or planning modules
- Robust to noisy, low-frequency sensing & laggy actuation

Technical aspects:

- Dual distillation method: combines agile motor imitation + heading prediction from depth images
- Unified reward principle: inner-product formulation promoting diverse parkour skills
- Automatic terrain curriculum for easier RL exploration
- Uses scandots as terrain-agnostic privileged signals
- Obstacle-aware heading control at deployment

Reinforcement Learning

Method – Teacher-student distillation



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

Phase 1
Teacher learns locomotion policy using privileged data :

- Environment parameters
- Scandots
- Heading from waypoints

Phase 2
Student learns heading using only depth camera data, under the teacher's supervision.

Reinforcement Learning

Unified Reward

Velocity tracking reward :
Aligns robot velocity with next waypoint heading.

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

Safety near edges reward :
Safety margin before gaps.

$$r_{clearance} = -\sum_{i=1}^{4} c_i M[p_i]$$

Style reward :
Enables aesthetic gaits, e.g. handstand

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

Impact of some parameters

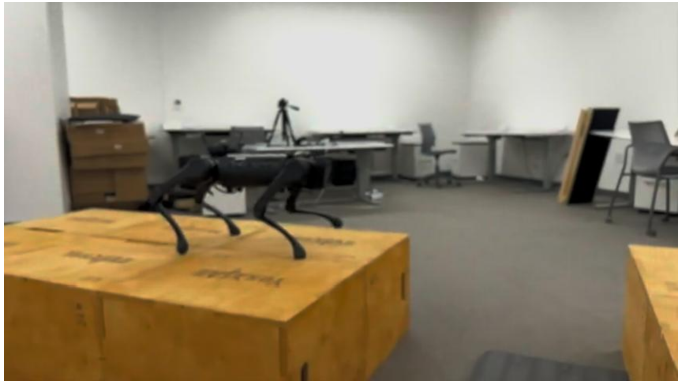
Simulation Baselines			
Reward	*NoInner*	No direction alignment	→ robot avoids obstacles instead of jumping
Reward	*NoClear*	No edge penalty	→ unsafe footholds near edges
Reward	*Noisy*	Noisy elevation maps	→ inconsistent policy, relies on collisions
Heading	*Both*	Always uses predicted heading	→ drift in control due to compounding error
Heading	*Mask*	No heading command in training	→ fails to learn turns
Heading	*Oracle*	Ground-truth heading	→ **best possible** benchmark for comparison

Real world		
NoClear	No edge clearance penalty at training	→ Robot steps too close to edge
NoDir	Human controls heading via joystick	→ Too slow for sudden changes

Only the full method (*inner-product reward, clearance penalty, and learned heading*) **succeeds reliably in both simulation and real-world parkour.**

Results

Quite impressive performance…



Performance metrics

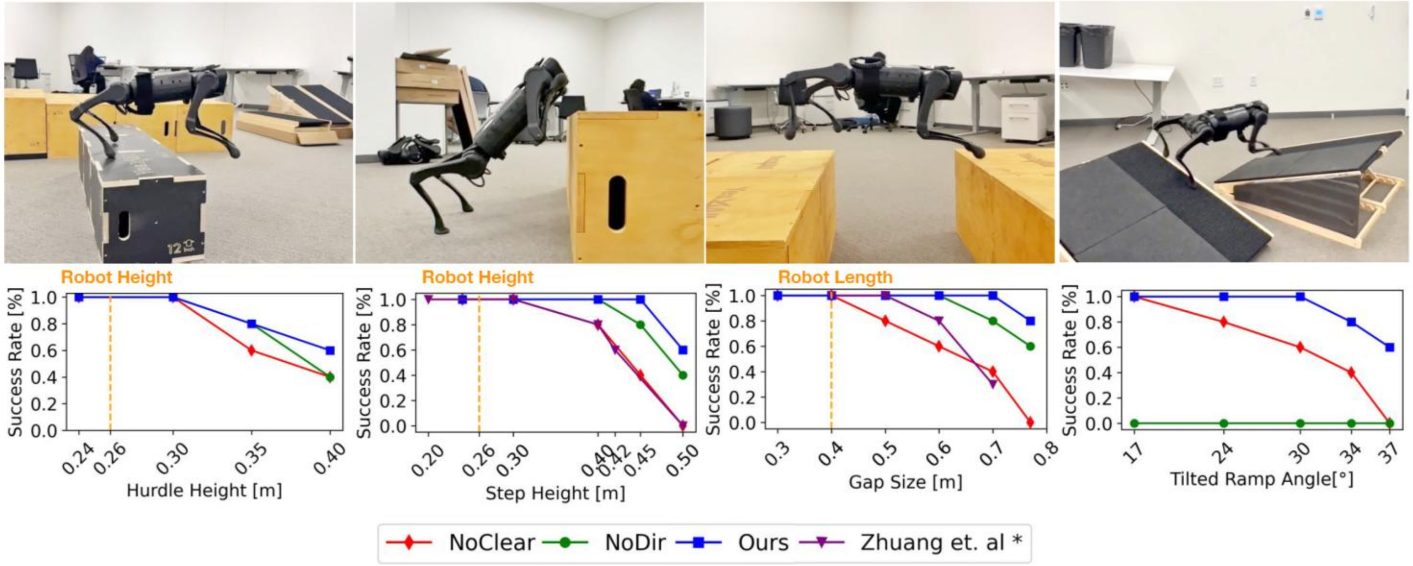
Perform parkour with a *low-cost* robot

- Imprecise actuation
- Imprecise sensing
- Single depth camera:
  - Low-frequency, Jittery, Prone to artifacts

Performance

High jump	2x the height of the hip joints.	0.5m
Long jump	2x its length.	0.8m
Handstand	Transition quadrupedal to bipedal walking	
Ramp	37°	

Performance metrics

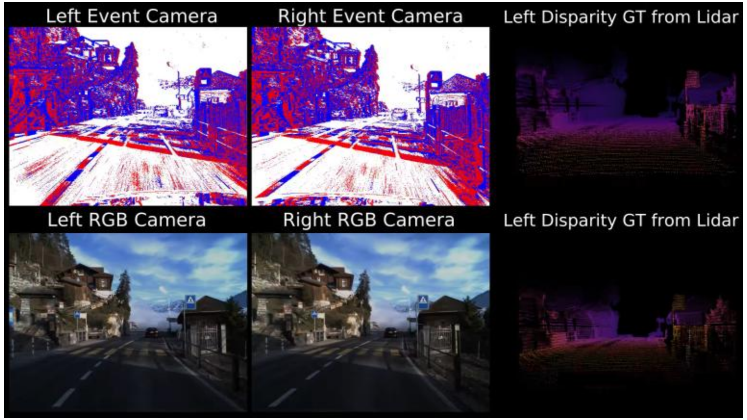


Performance		
High jump	2x the height of the hip joints.	0.5m
Long jump	2x its length.	0.8m
Handstand	Transition quadrupedal to bipedal walking	
Ramp	37°	

Impact on other articles

2 alternative approaches

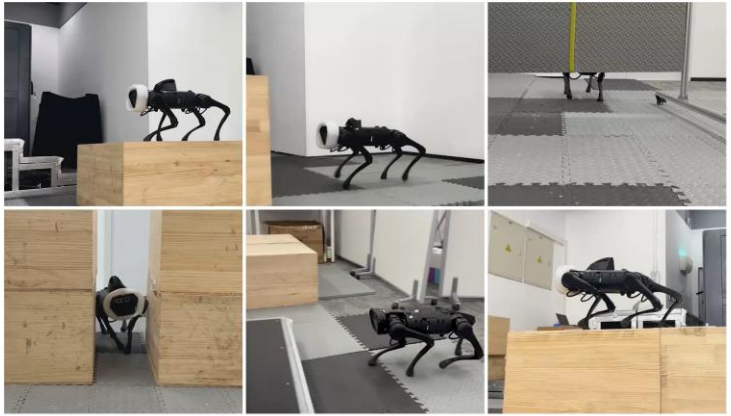
Event camera and spiking neural network :



https://www.youtube.com/watch?v=W4yW78y4F7A

https://ieeexplore.ieee.org/document/11209556

World model-based perception :



https://wmp-loco.github.io/

https://ieeexplore.ieee.org/document/11128762

Conclusion

Pros / Cons

PROS

- High performances with low cost and imprecise hardware
- baseline method
- Emergent/adaptive behavior in challenging environments
- Minimal intervention
- Model-free approach

CONS

- Require offline training
- Finding suitable reward function is a challenging/tedious task
- Policy's success depends on sim-to-real gap.
- Hard to guarantee safety
- Might face suboptimal convergence
- Depth camera limitation : lightning, low-frequency, …

Conclusion

Pros / Cons

PROS

- High performances with low cost and imprecise hardware
- baseline method
- Emergent/adaptive behavior in challenging environments
- Minimal intervention
- Model-free approach

CONS

- Require offline training
- Finding suitable reward function is a challenging/tedious task
- Policy's success depends on sim-to-real gap.
- Hard to guarantee safety
- Might face suboptimal convergence
- Depth camera limitation : lightning, low-frequency, …

EPFL

Conclusion

Pros / Cons

PROS

- High performances with low cost and imprecise hardware
- baseline method
- Emergent/adaptive behavior in challenging environments
- Minimal intervention
- Model-free approach

CONS

- Require offline training
- Finding suitable reward function is a challenging/tedious task
- Policy's success depends on sim-to-real gap.
- Hard to guarantee safety
- Might face suboptimal convergence
- Depth camera limitation : lightning, low-frequency, …

Conclusion

Pros / Cons

PROS

- High performances with low cost and imprecise hardware
- baseline method
- Emergent/adaptive behavior in challenging environments
- Minimal intervention
- Model-free approach

CONS

- Require offline training
- Finding suitable reward function is a challenging/tedious task
- Policy's success depends on sim-to-real gap.
- Hard to guarantee safety
- Might face suboptimal convergence
- Depth camera limitation : lightning, low-frequency, …

Repartition du temps

- Summary and Key aspects – 2min30 – William
  - Material (robot Unitree, nb of joints, type of camera)
  - Key aspects
  - Videos
- Details about the article – 2min30 – Lois
  - RL methods (rewards, scandots and distillation)
  - Show phases image
  - Show
- Results, Limits, Comparison with articles – 2min30 – Teo
  - Results : gif to show the results, talk about the performance
  - Limits : impact of removing one parameter
  - Comparison with articles