



上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

可泛化人形机器人运控控制和移动操作

上海人工智能实验室

庞江淼

2025.04.13

机器人学习与 人形机器人

2019年之前
波士顿动力



传统控制方法
WBC+MPC

2019
ETH Zurich



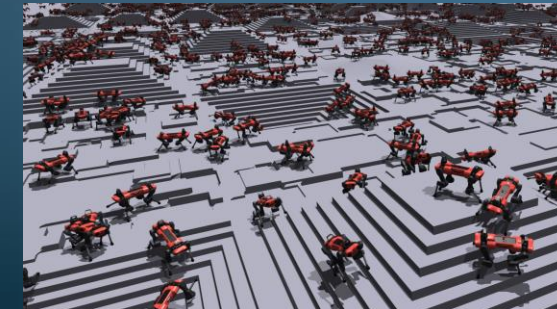
Science Robotics
首次强化学习 Sim2Real 超越传统方法

2020
UC Berkeley



RSS 2020 Best Paper
使用宇树机器人 Laikago

2021
Nvidia



Isaac Gym
GPU 仿真训练效率提升 2-3 个数量级

2025 百花齐放



机器人学习使人形机器人不断突破能力上限，且更新速度还在加快

关键目标

本体泛化

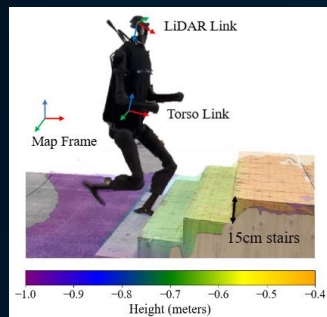
技能泛化

场景泛化

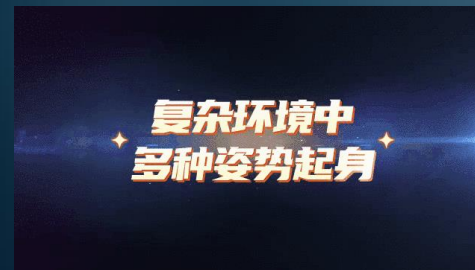
人形机器人：从运动控制到移动操作



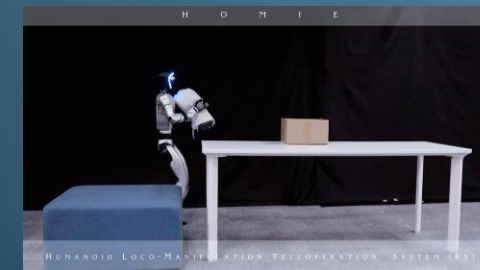
基础行走



感知融合



全身控制



人景交互 & 移动操作

强化学习？传统控制？ Motion Priors？ 大规模数据？

HIM

传统内模控制

融合强化学习框架

仿真训练一小时

高效可泛化

PIM

复杂地形可泛化行走

BeamDojo

走梅花桩

VC-Com

处理动态场景

HugWBC

多步态全身控制方法

HoST

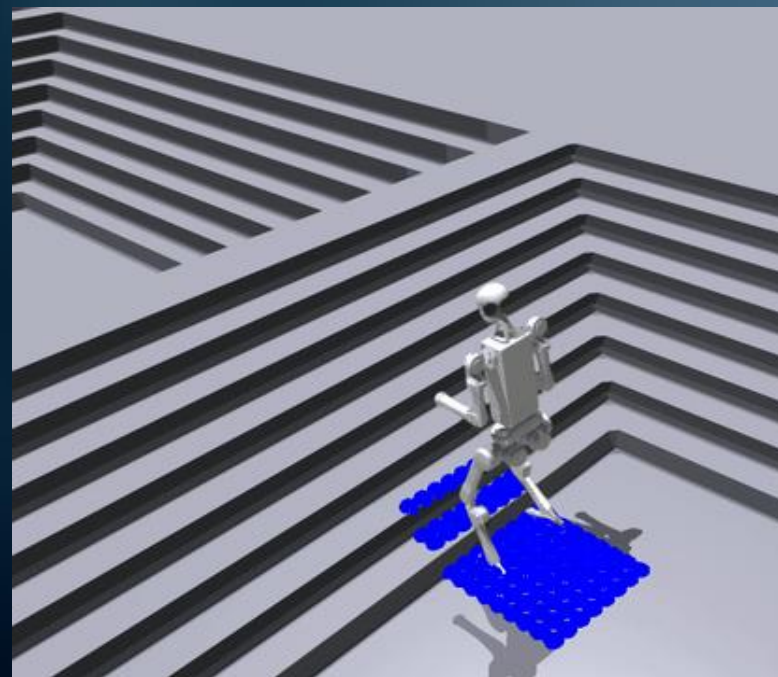
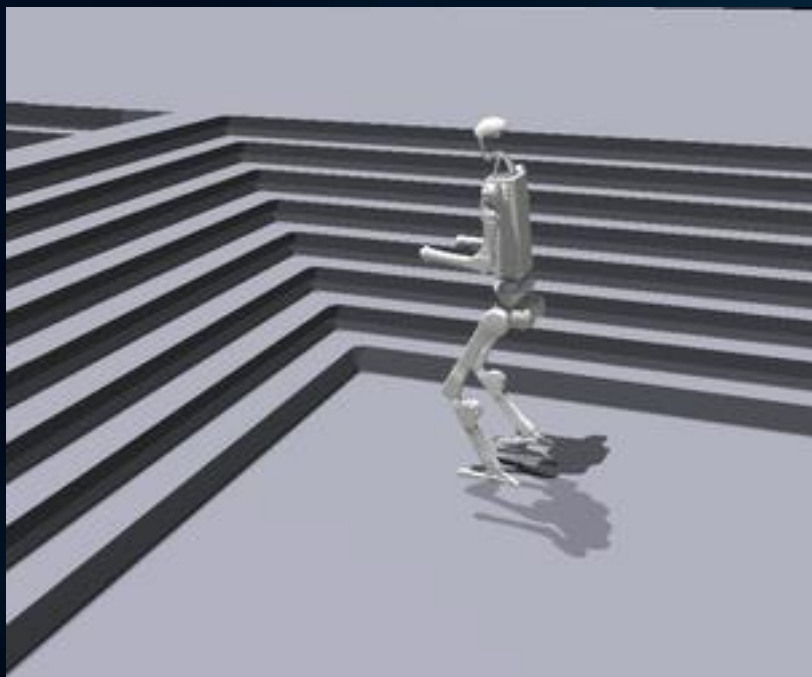
可泛化自主起立方法

HOMIE

人形机器人驾驶舱

PIM: 感知融合驱动复杂地形可泛化运动控制

关键问题: 人形机器人本身结构“不稳定”，复杂地形可泛化运动控制必须融合感知信息



感知融合运动控制的关键难点:

难点1: 感知信息很难完全精准, 会不可避免的给控制系统带来扰动

难点2: 感知信息 Sim-to-Real Gap 大, 会增大虚实迁移的困难程度

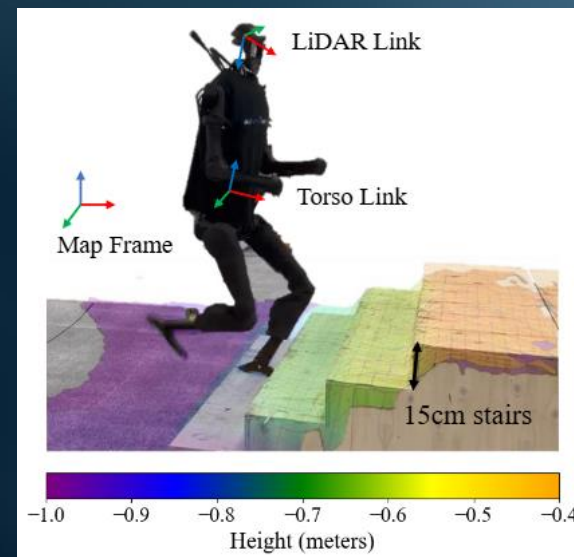
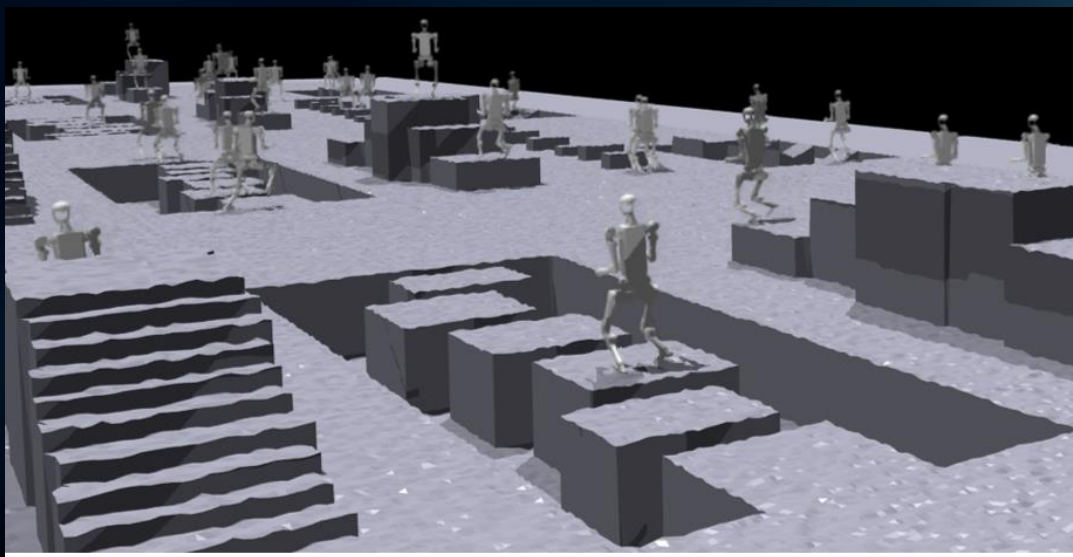
难点3: 需要渲染感知图像, 会影响控制策略在仿真中的学习效率

PIM: 感知融合驱动复杂地形可泛化运动控制

解决思路:

仿真训练直接使用**障碍地形的绝对高度**进行策略学习

真机推理采用使用 **onboard, continuously updated elevation maps** centered around the robot 进行感知




方法优势1: HIM 范式, Batch-level Contrastive Learning, 高效训练

方法优势2: 训练无需 encode/render RGBD 信息/原始点云, 仅需 RTX 4090 上 3 小时训练时间

方法优势3: 持续对机器人周围全局信息的感知, 能够抵抗传感器的移动/噪声

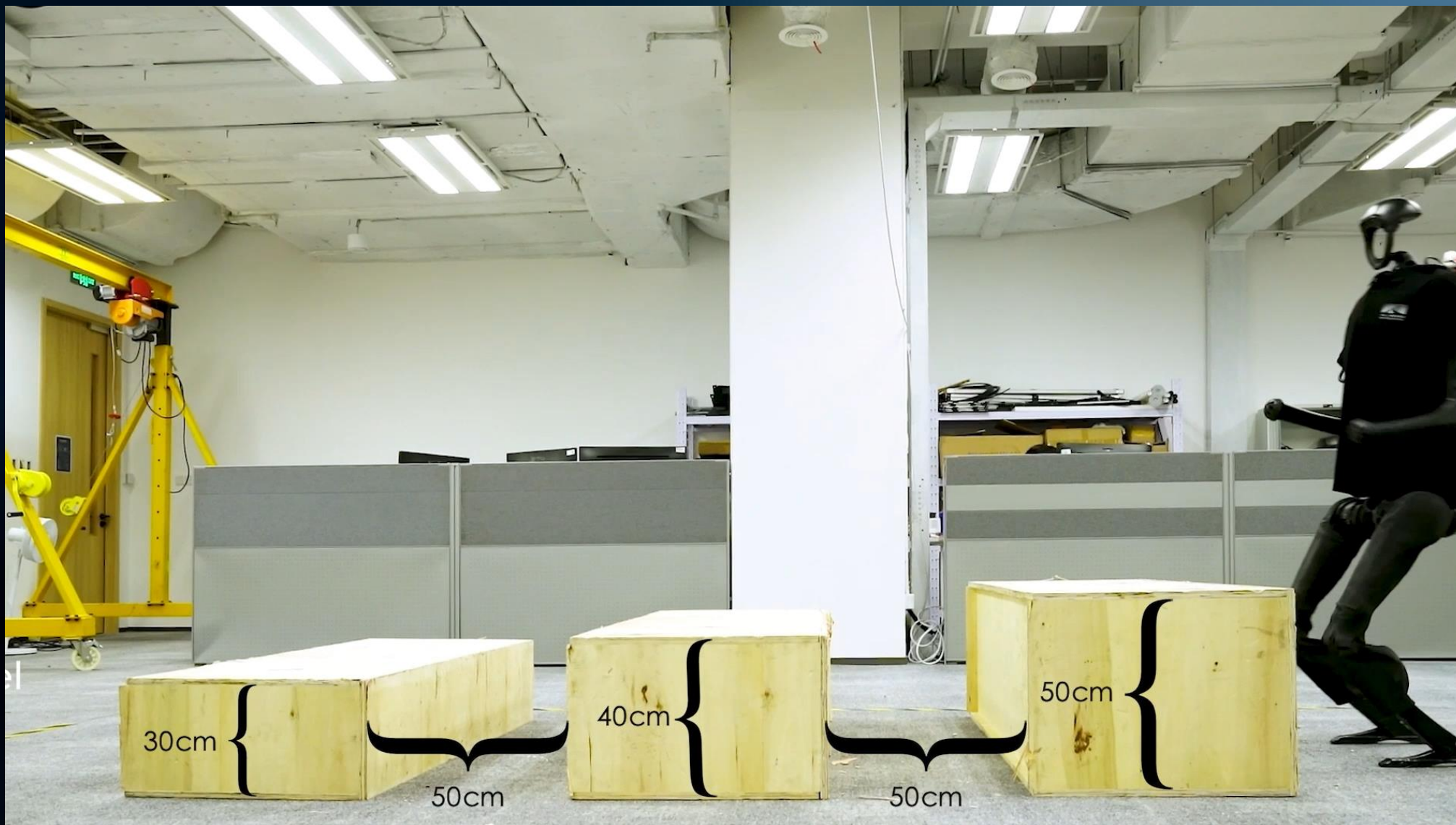
系统优势: 首个【多种人形机器人】【室内外地形】【上下台阶、跳 50 cm 高台、跳 Gap】的策略

PIM: 感知融合驱动复杂地形可泛化运动控制



Continuous up and down stairs
(single-shot view)

PIM: 感知融合驱动复杂地形可泛化运动控制



BEAMDOJO: 人形机器人在稀疏落足点上的稳定行走

问题: 人形机器人的运动控制如何处理稀疏落足点的复杂地形 (如梅花桩、独木桥等)



Lipschitz-Constrained [Chen et al. 2024]



PIM [Long et al. 2024]

关键难点:

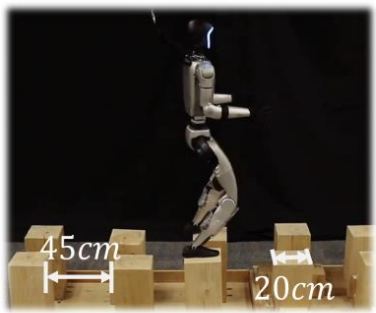
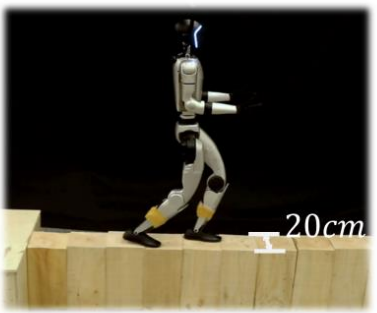

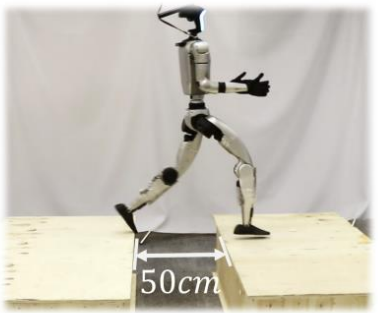
难点1: 稀疏的落足奖励, 引入值函数奖励分配的难题

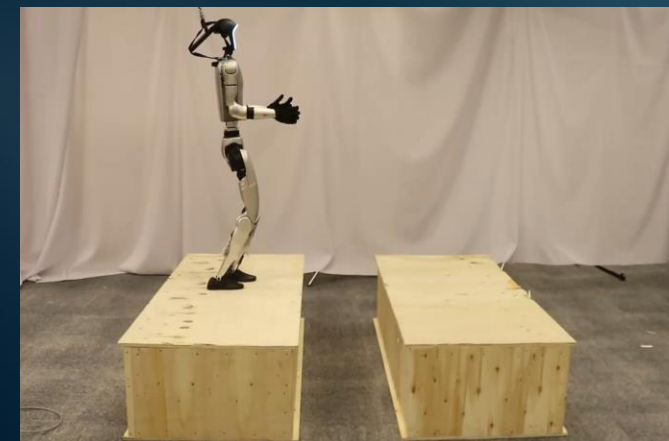
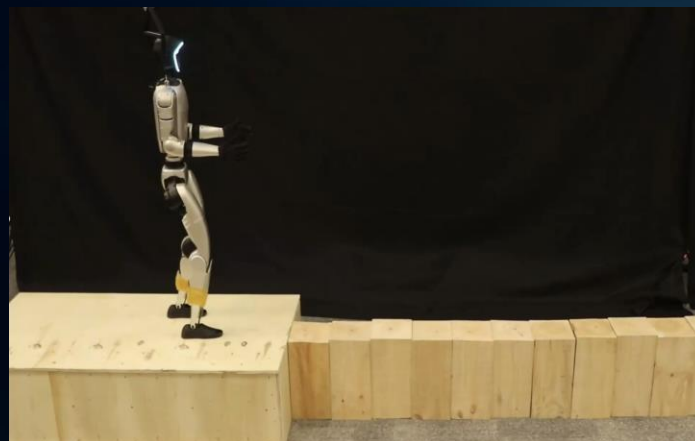
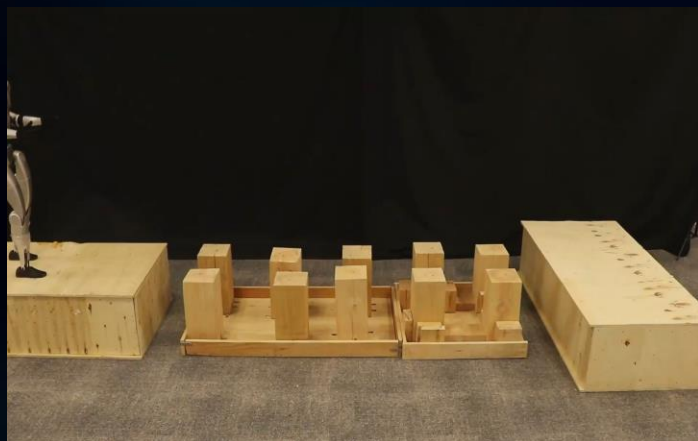
难点2: 在学习过程中容易跌倒导致早停, 无法有效收集成功轨迹并学习

难点3: 真实世界中缺乏用于控制细粒度行走的有效可靠的感知信息

BEAMDOJO: 人形机器人在稀疏落足点上的稳定行走

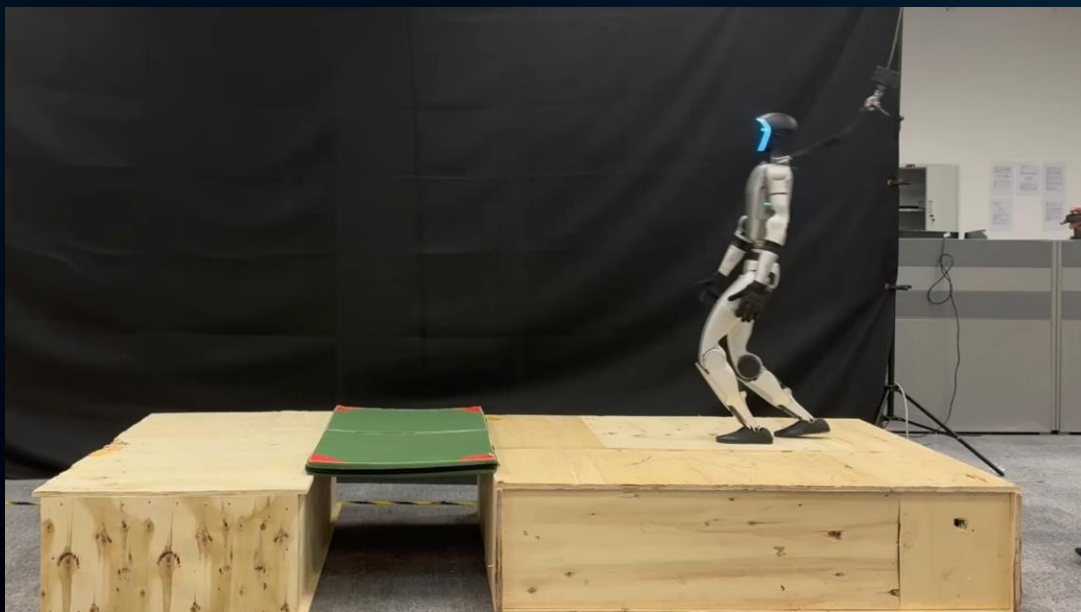
真机实验: BEAMDOJO 可以在梅花桩地形上连续5次测试达到4/5成功率, 并且可以有效零样本泛化到跳木、50cm间隔Gap上

			
Stepping Stones	Balancing Beams	Stepping Beams	Gaps
R_{succ}	R_{succ}	R_{succ}	R_{succ}
R_{trav}	R_{trav}	R_{trav}	R_{trav}
Ours w/o HR	Ours w/o HR	Ours w/o HR	Ours w/o HR
BEAMDOJO	BEAMDOJO	BEAMDOJO	BEAMDOJO
1/5	0/5	1/5	3/5
38.20%	12.37%	30.00%	60.00%
4/5	4/5	3/5	5/5
92.18%	88.16%	70.00%	100.00%



VB-COM: 视野遮挡和动态干扰下的稳定行走

关键问题: 基于感知的高动态运动控制难以应对真实环境中大量存在的**视觉失效**以及**动态地形/障碍物**。



现有方法



VB-Com

关键难点:

难点1: 仿真中**难以模拟**真实环境中存在的可变地形等复杂干扰。

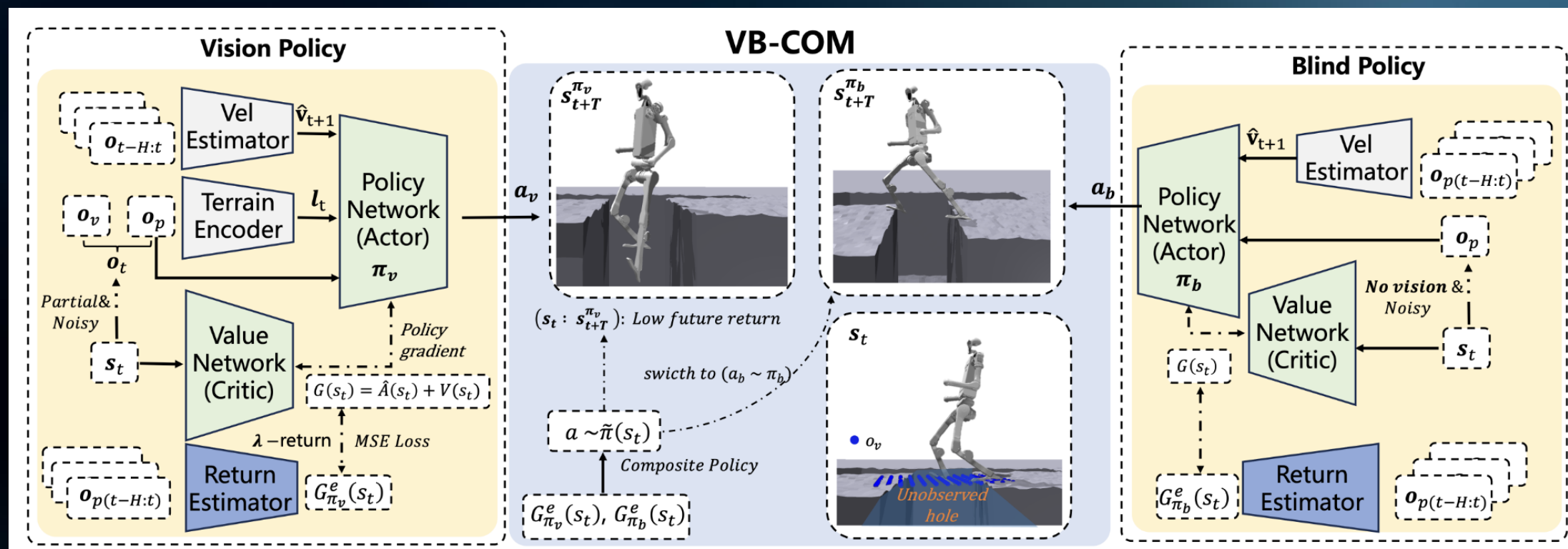
难点2: 高动态干扰场景的应对**快速**的响应时间。

VB-COM: 视野遮挡和动态干扰下的稳定行走

解决思路:

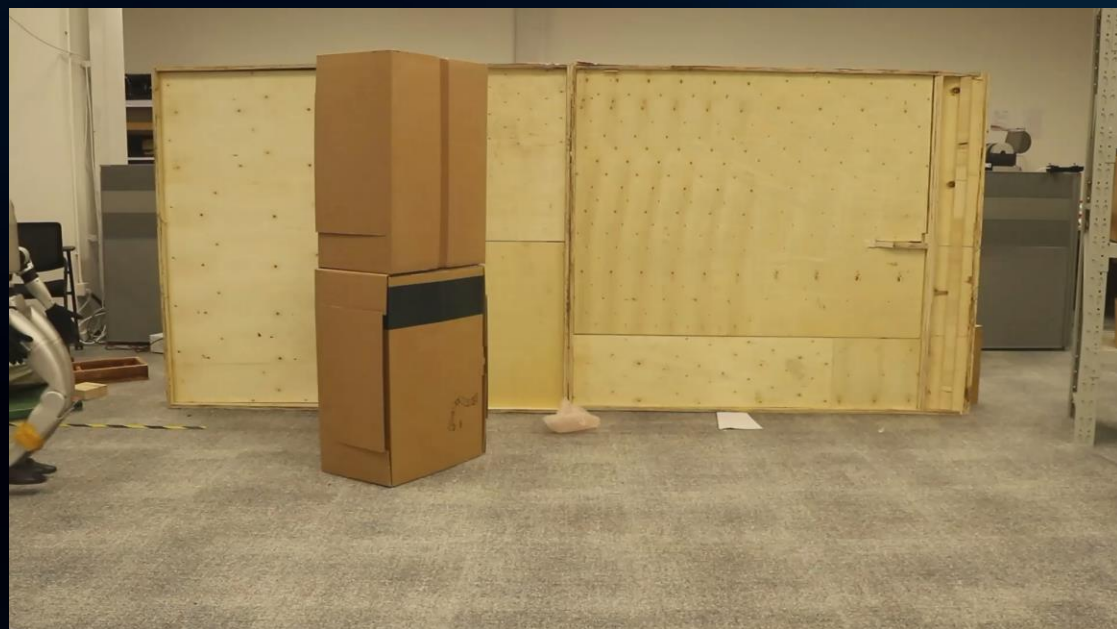
双策略自动切换：稳定态使用视觉感知，非稳定态使用本体感知。

实现方式：利用机器人的本体感知信息实现在线回报估计（Return Estimation），自动判断状态是否稳定



视觉+控制——VBCOM

基于视觉的动态控制策略选择



VB-Com



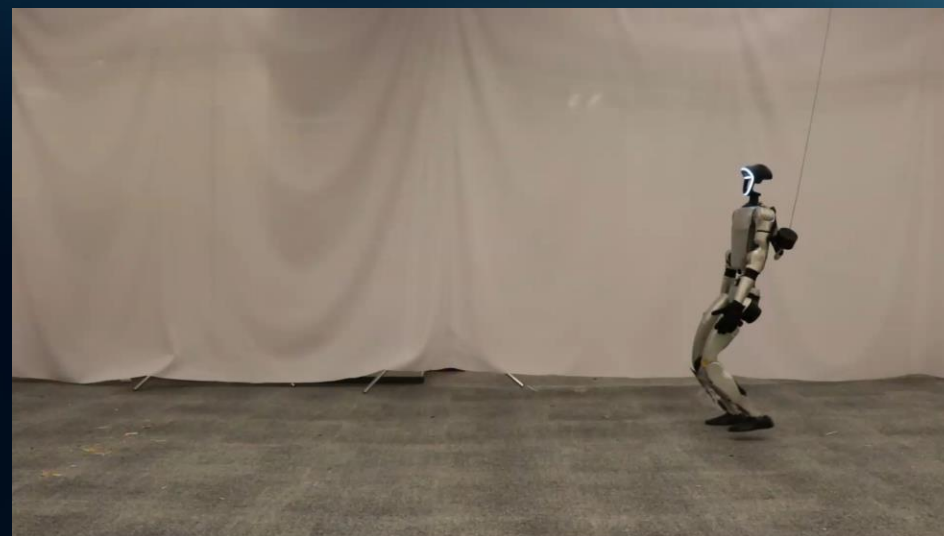
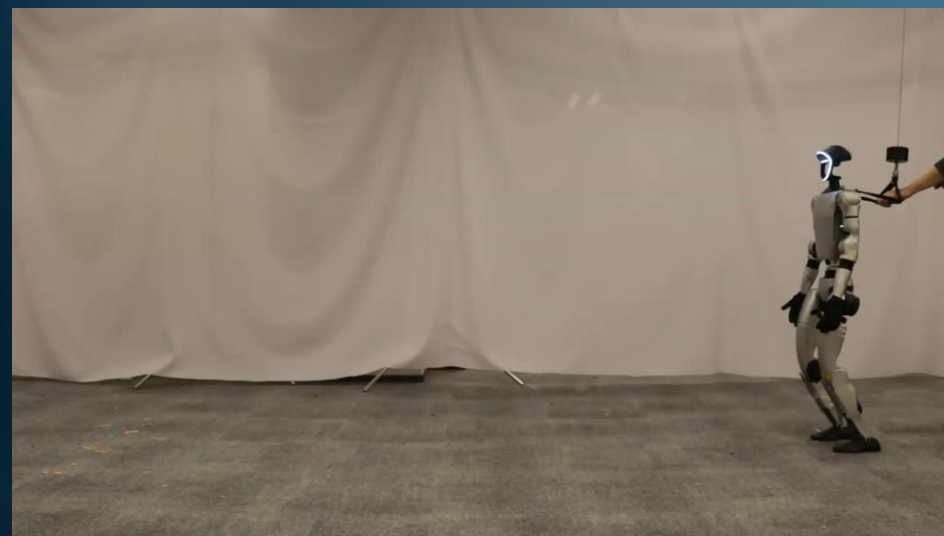
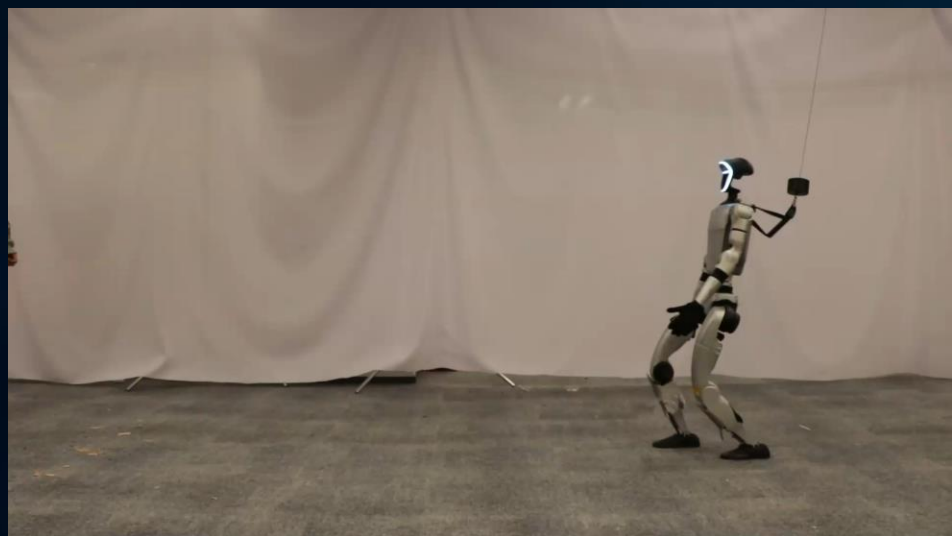
Blind Only



Vision Only

视觉+控制——VBCOM

基于视觉的动态控制策略选择

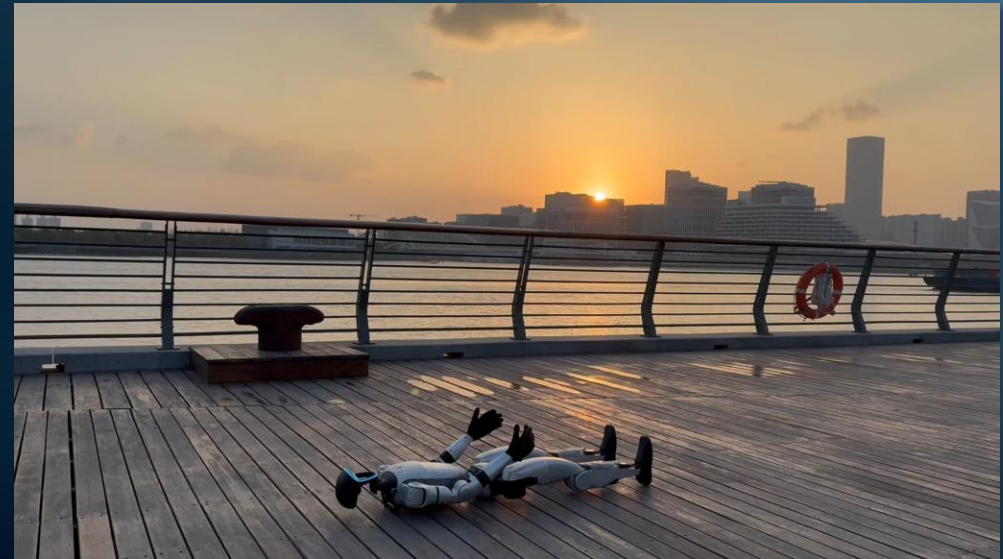


慢速物体躲避：用视觉

快速物体躲避：用本体

HoST: 多姿势下的人形机器人站起控制

关键问题: 平滑、鲁棒、可泛化的站起控制策略学习



人形机器人站起控制的关键难点:

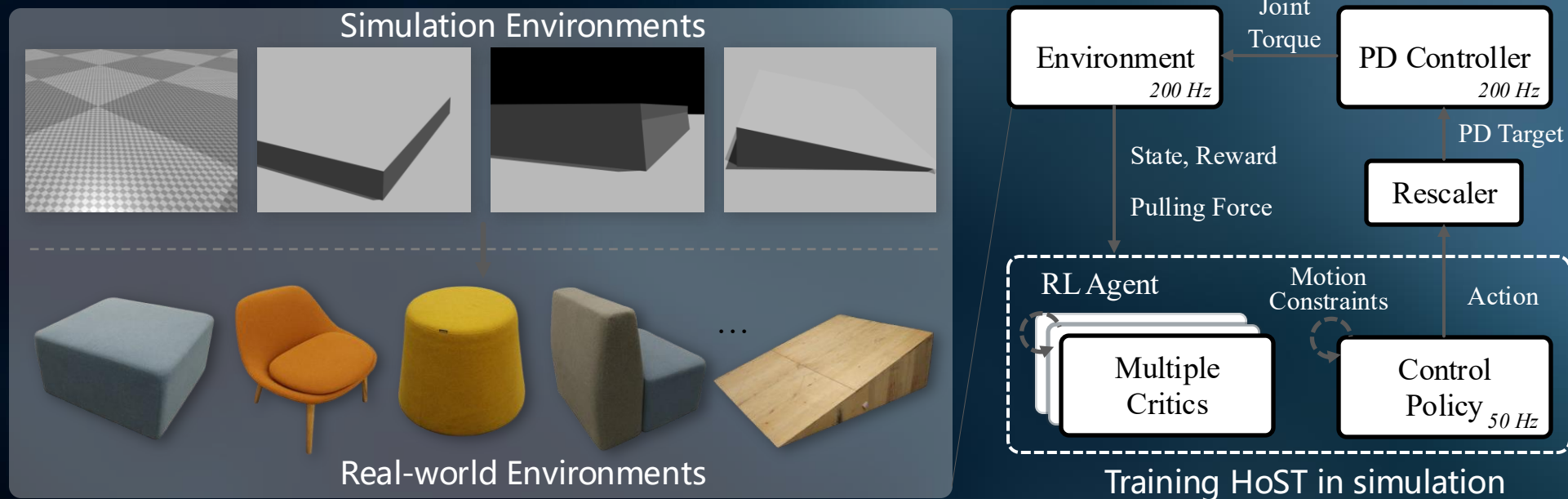
难点1: 站起动作过于剧烈、抖动, 影响真机部署的可能性

难点2: 仅通过奖励函数, 探索且掌握平滑的站起动作很困难

难点3: 站起控制是多接触、全身、高动态动作, 奖励函数本身难以设计和优化

HoST: 多姿势下的人形机器人站起控制

解决思路: 仿真训练引入**多维动作约束**、**多种地形**、**多个值函数**进行策略学习



方法优势1: 受启发于婴儿的技能学习, 引入辅助外力促进站起行为探索

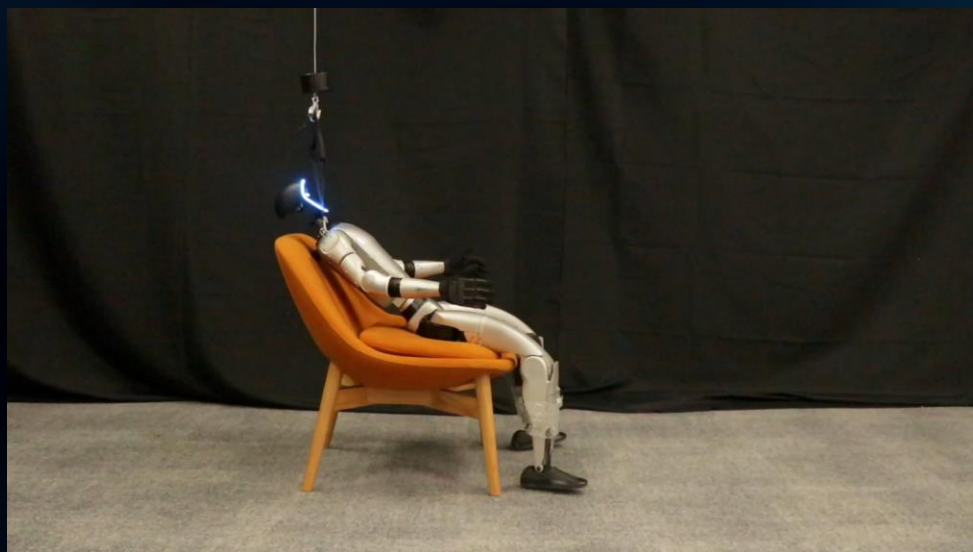
方法优势2: 引入action rescaler限制运动速度, smoothness regularization规避真机抖动

方法优势3: 多值函数对20多项奖励函数“分而治之”, 减小调参难度与奖励函数间的冲突

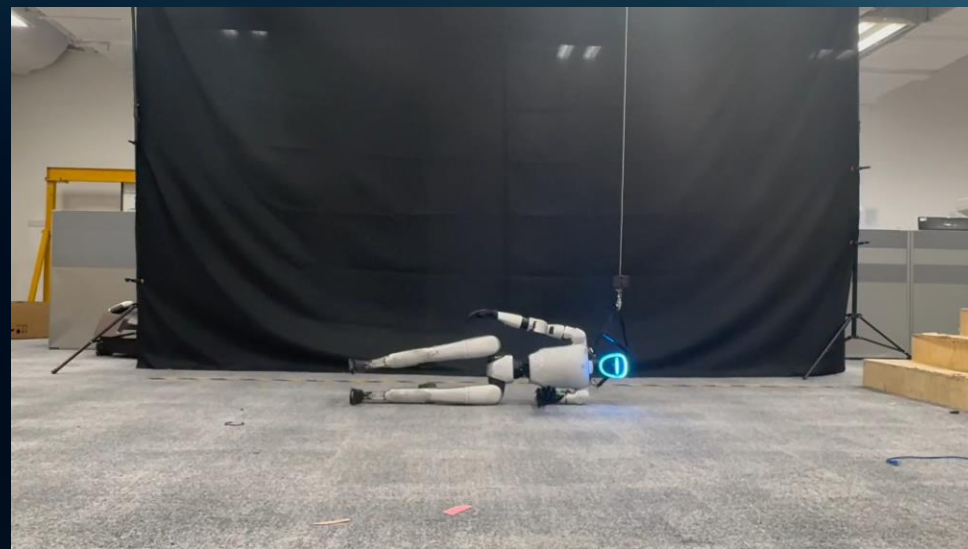
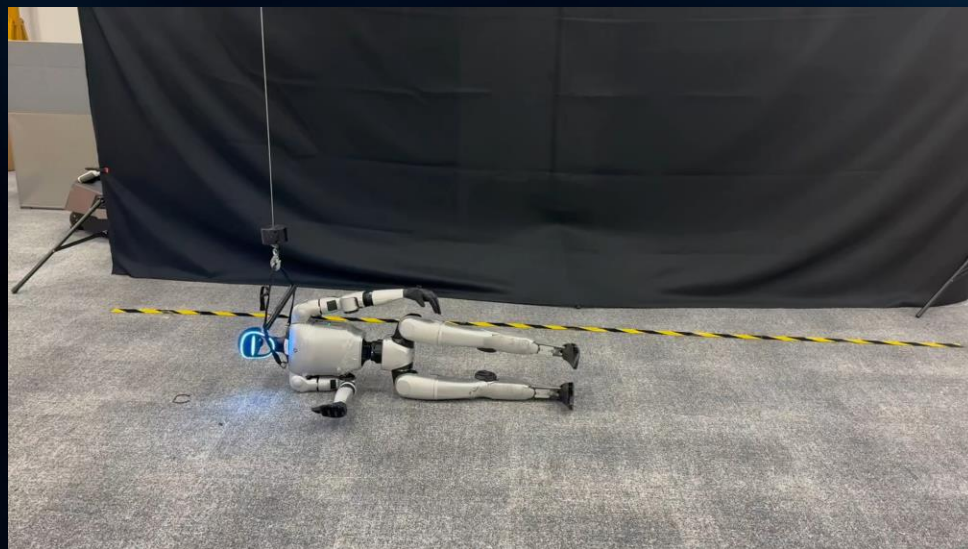
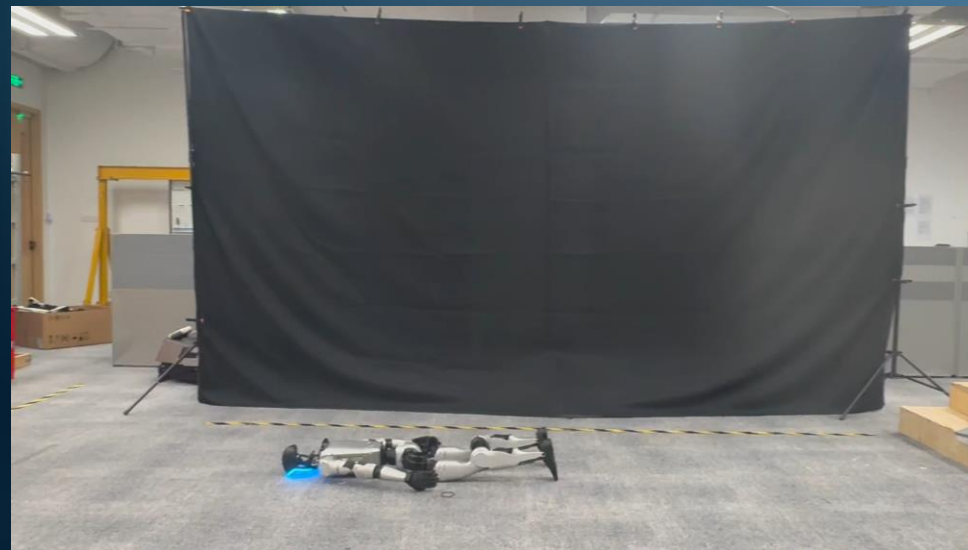
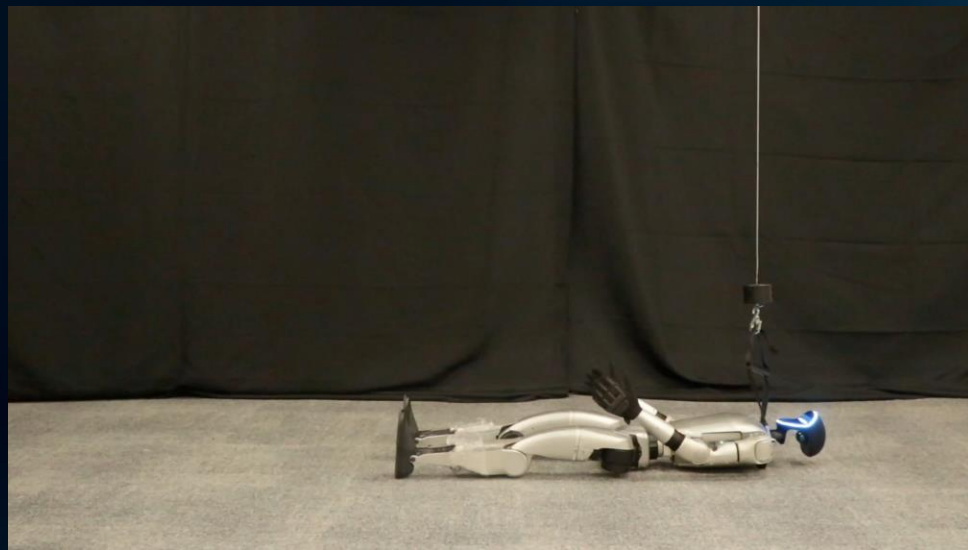
方法优势4: 建立多种仿真地形模拟现实姿势, 加入域随机化实现zero-shot真机部署

系统优势: 首个【多室内外场景】【超越平地】【强泛化性与鲁棒性】的站起控制策略

HoST: 多姿势下的人形机器人站起控制



HoST: 多姿势下的人形机器人站起控制

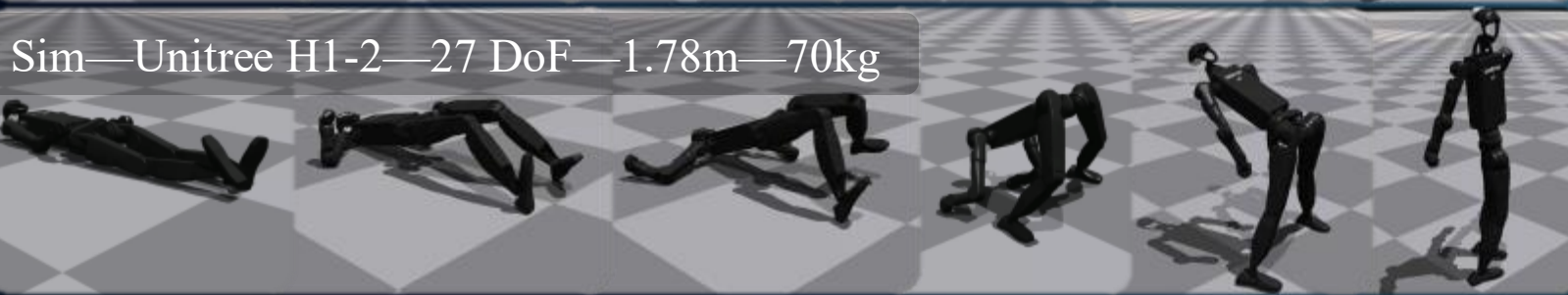


HoST: 多姿势下的人形机器人站起控制

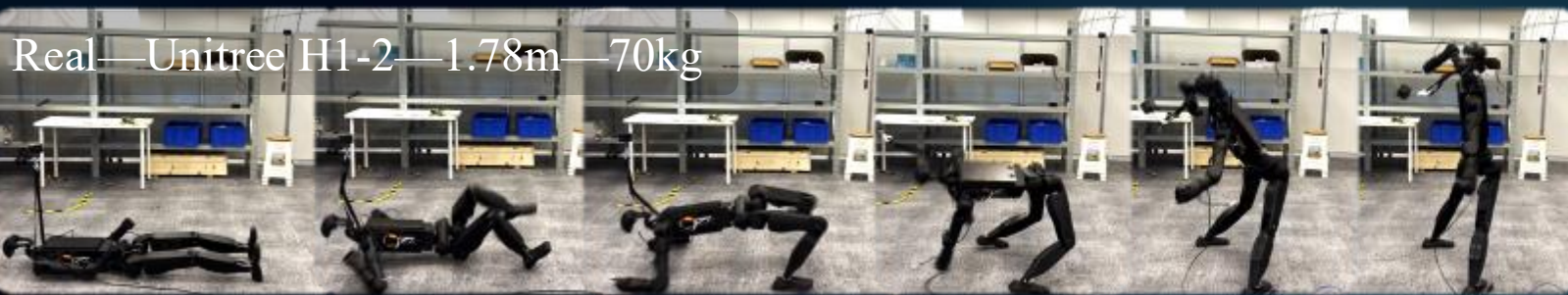
Sim—Unitree H1—19 DoF—1.8m—47kg



Sim—Unitree H1-2—27 DoF—1.78m—70kg



Real—Unitree H1-2—1.78m—70kg



HugWBC: 多用途的人形机器人细粒度控制

关键问题: 当前的人形机器人仅能产生单一的运动行为, 缺乏可扩展性, 难以产生像人类般的多样化运动 (如奔跑、跳跃、单脚跳跃等)



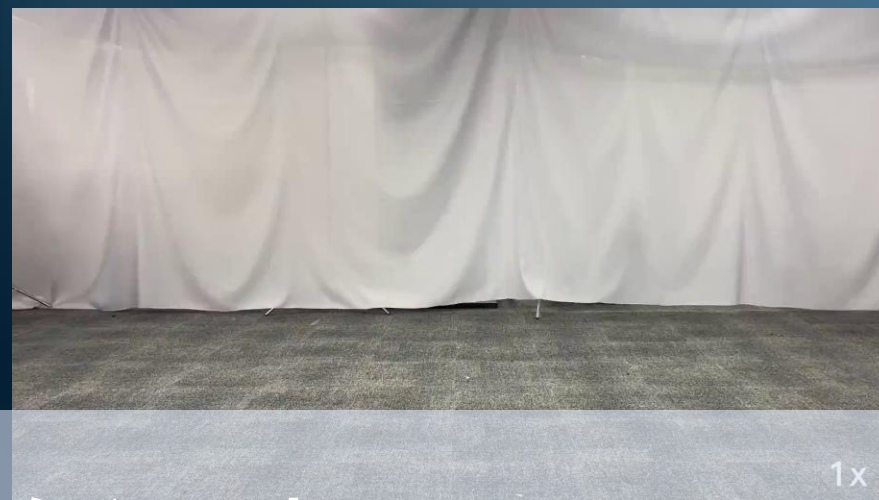
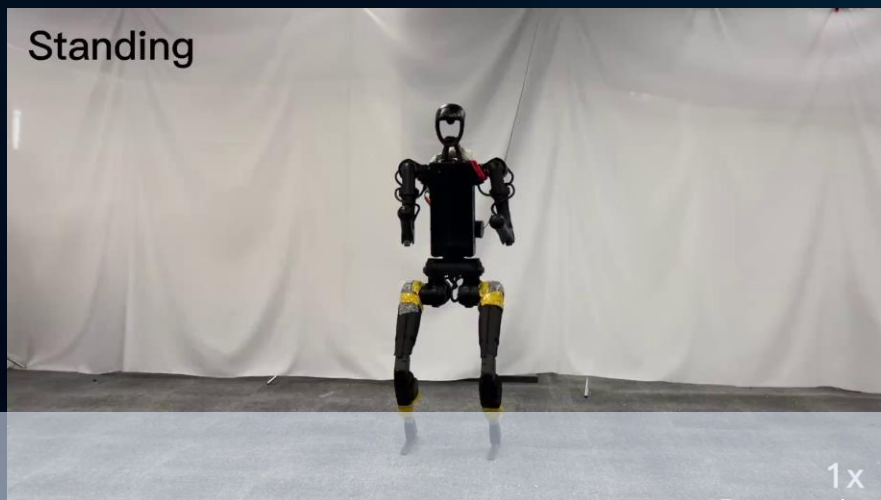
多用途细粒度运动控制的关键难点:

难点1: 步态和细粒度控制会引入多目标优化问题, 策略容易陷入局部次优, 产生不自然的运动

难点2: 上肢的控制会给控制系统引入额外扰动, 影响控制策略在仿真中的学习效率

难点3: 人形机器人静态不稳定特性, 会增大动态性运动在真实环境迁移的难度

HugWBC: 多用途的人形机器人细粒度控制

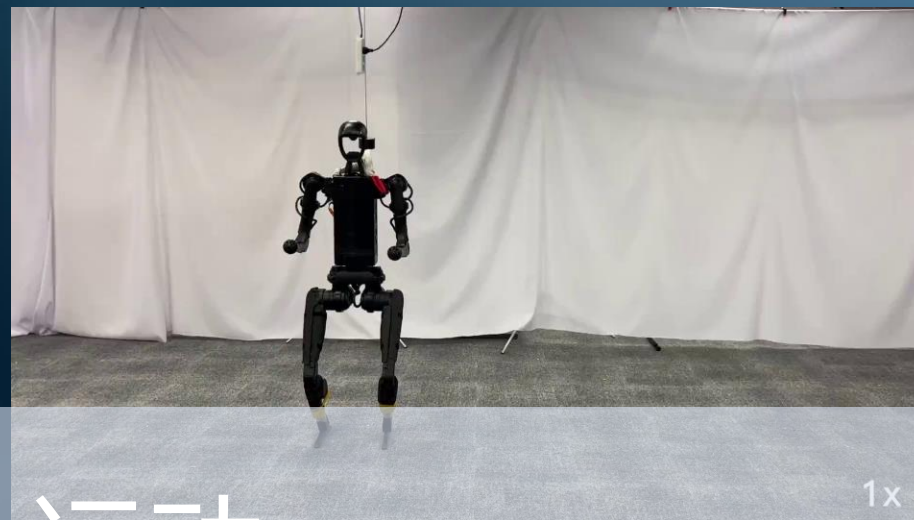


丰富全身运动



系统优势: 首个【单一控制策略】【实现多种拟人步态】的控制器, 包括【奔跑, 跳跃, 站立, 行走】

HugWBC: 多用途的人形机器人细粒度控制



动态性运动

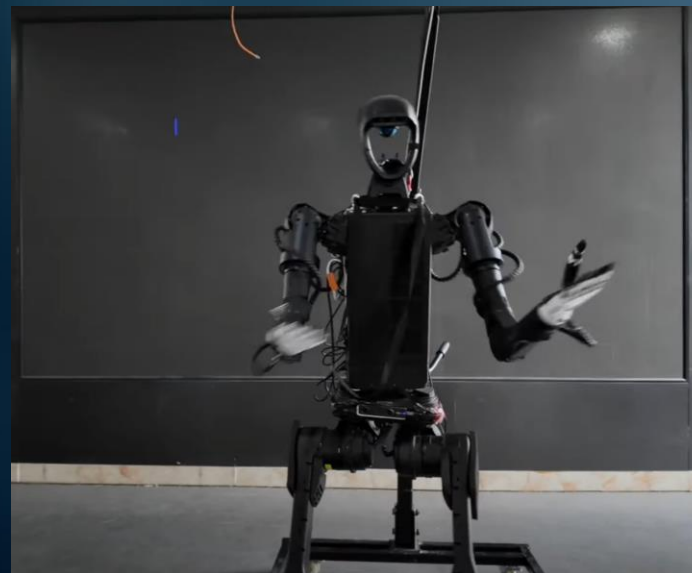


HOMIE: 人型机器人移动操作驾驶舱

关键问题: 让人类能够丝滑准确操控人形机器人完成移动操作任务



Mobile-TeleVision [Lu et al. 2024]



OmniH2O [He et al. 2024]

人型机器人完成移动操作任务的关键难点:

难点1: “大脑”侧——缺乏准确且丝滑的人型机器人全身遥操作方案

难点2: “小脑”侧——需要稳定的控制策略, 在上半身任意姿态下保持平衡

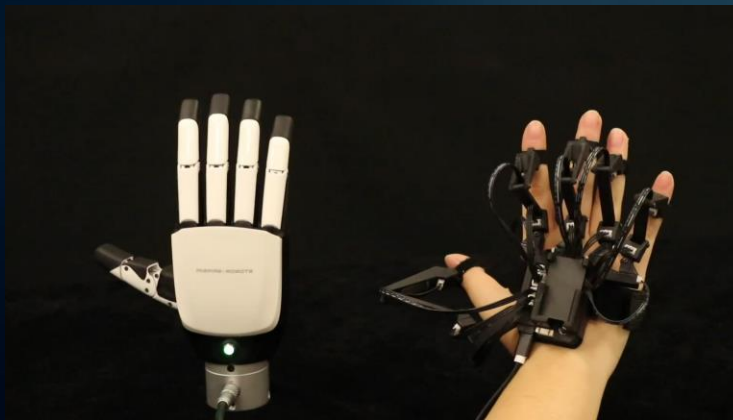
HOMIE: 人型机器人移动操作驾驶舱

“大脑” —— 全新人形机器人全身遥操作方案



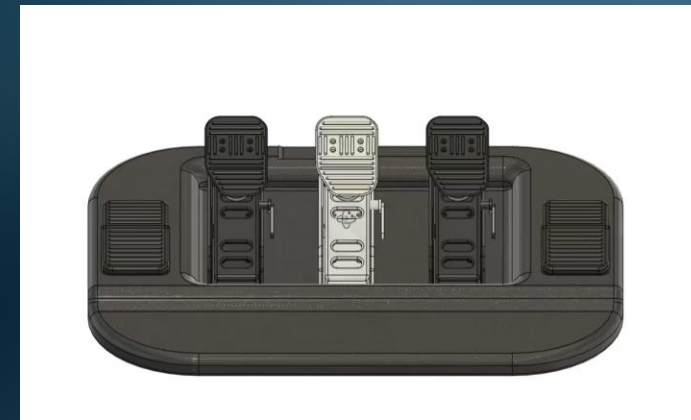
1、同构外骨骼:

与机器人上半身相同构型
避免逆运动学解算误差



2、动捕手套:

基于霍尔传感器的15DoF手套
实现准确高效的关节旋转采集

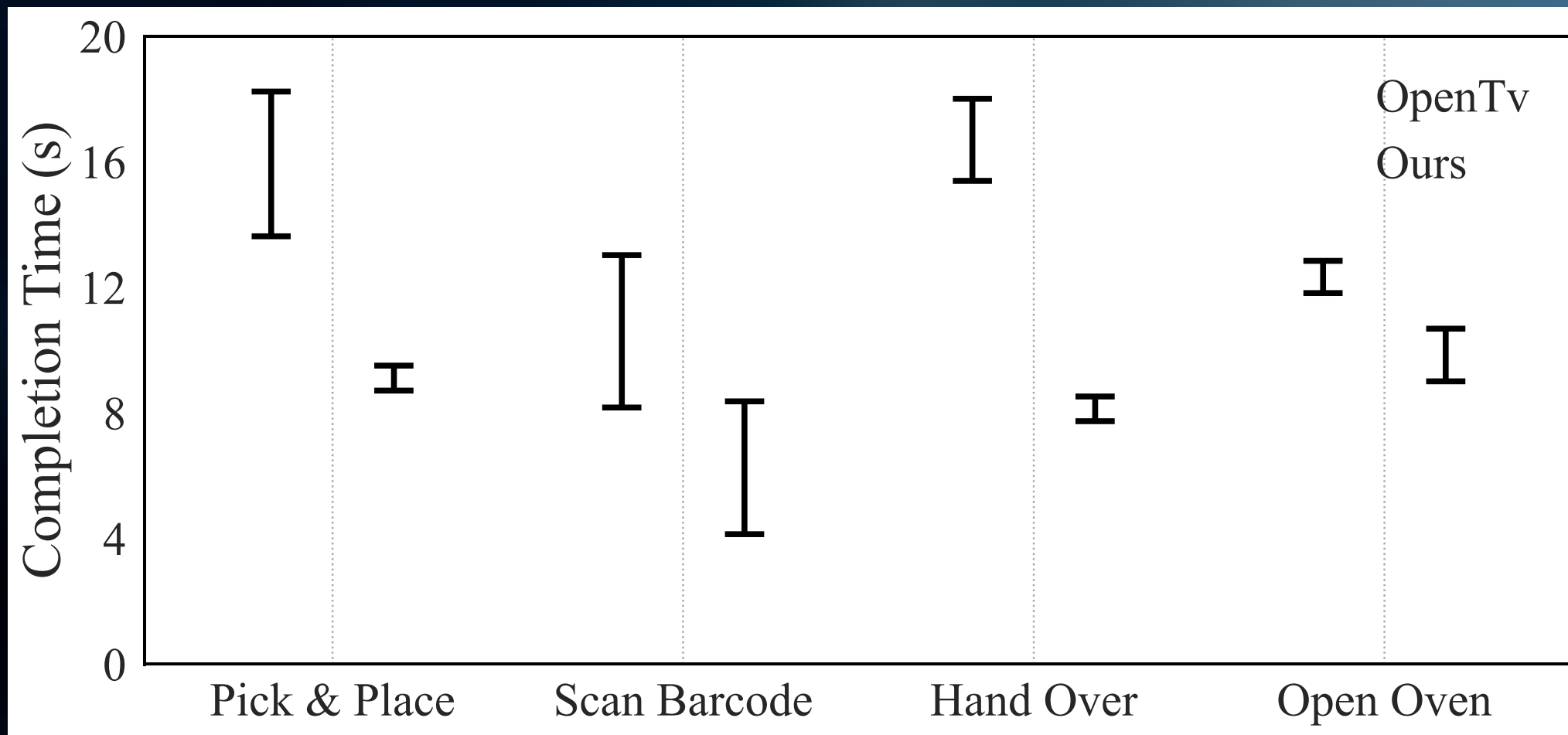


3、脚踏板:

控制机器人运动命令
实现单一操作者的全身遥操

HOMIE: 人型机器人移动操作驾驶舱

遥操作效果对比: 任务用时约是OpenTelevision一半



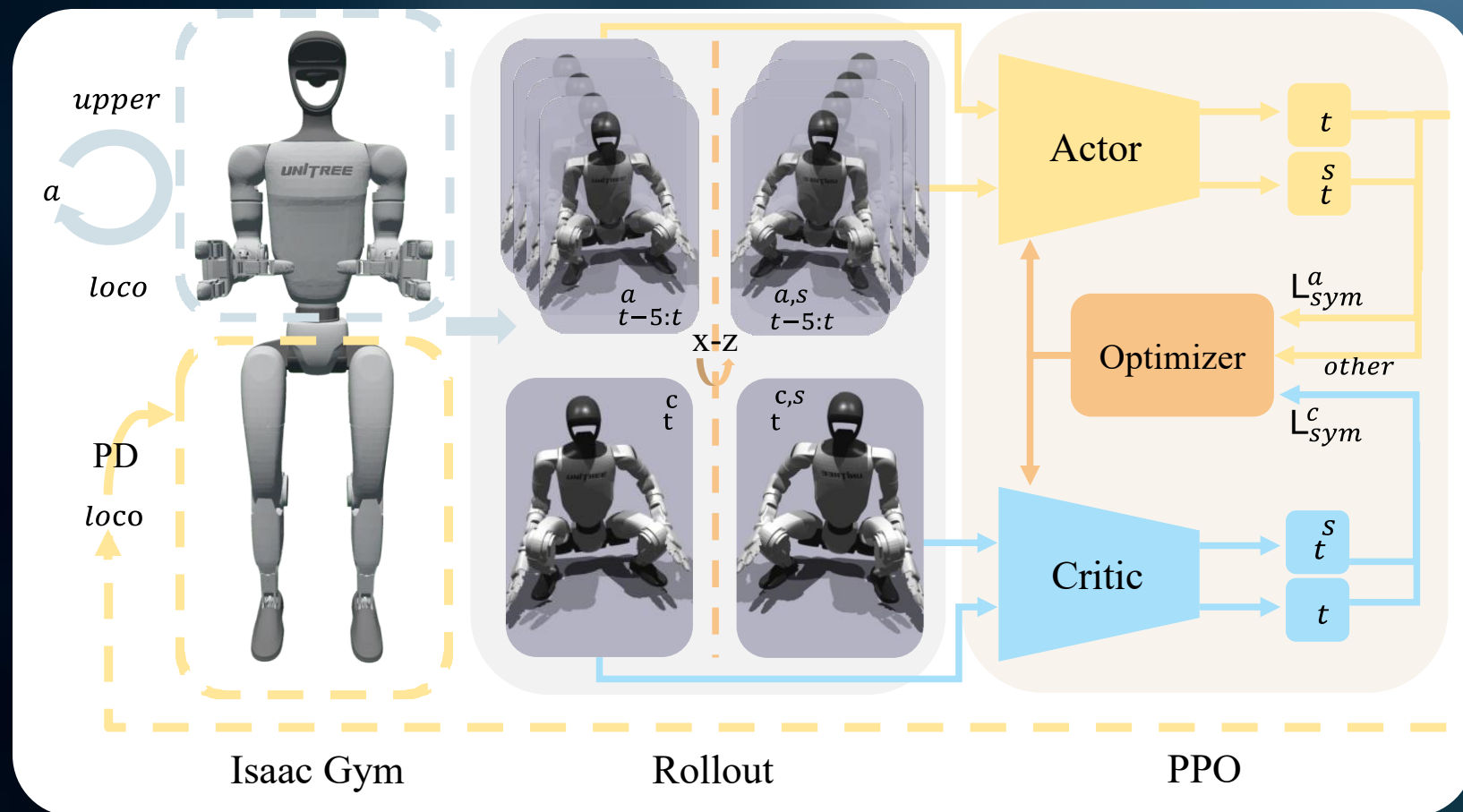
HOMIE: 人型机器人移动操作驾驶舱

“小脑” —— 首个支持任意上身姿势稳定移动下蹲的算法

1、课程学习：逐步增大上身运动范围，课程式学习逐渐适应任意上身姿势（无需使用动捕数据）

2、启发式奖励：膝盖相关启发式奖励函数（低了曲膝，高了反之），引导机器人更好学会下蹲到任意高度（首创稳健下蹲）

3、对称性利用：利用机器人结构的对称性进行数据增强并约束网络输出的对称性（运动更对称）

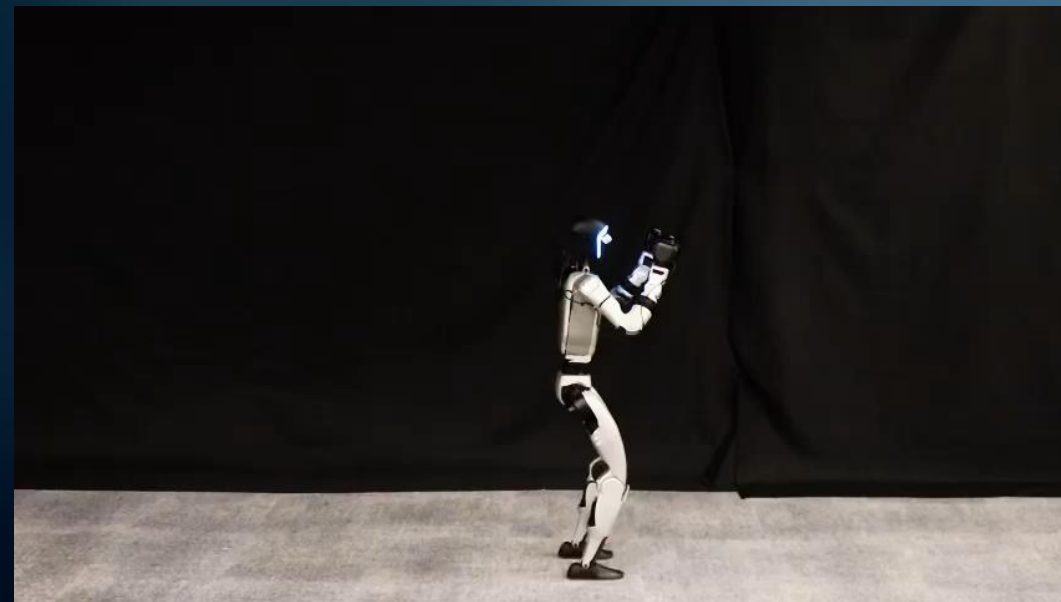


HOMIE: 人型机器人移动操作驾驶舱

“小脑” —— 首个支持任意上身姿势稳定移动下蹲的算法



稳健下蹲



任意上身姿势

HOMIE: 人型机器人移动操作驾驶舱



总结

- 人形机器人的基础运动能力在强化学习（Reward Engineering）的帮助下有非常好的进步
- 人形机器人的移动操作还非常遥远

- 场景泛化：Perception – 传感器方案，表征方式，融合方式
- 技能泛化：Unified Model – Motion Prior? Scaling?

Limits

- **Basic humanoid control** benefits from RL(Reward Engineering)
- **Generalizable humanoid loco-manipulation** is way too hard
- What limits the **upper bound**? Hardware & Perception & Scaling



感知

Unified Model?

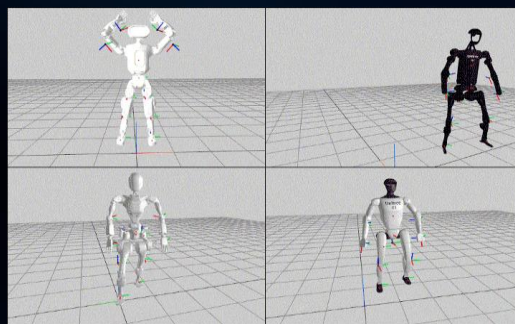
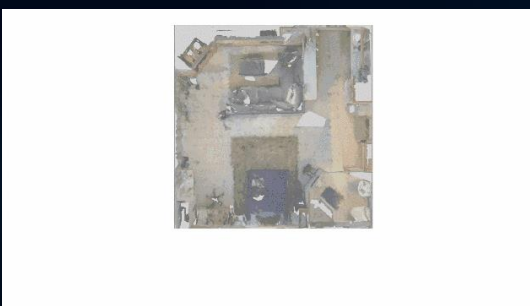
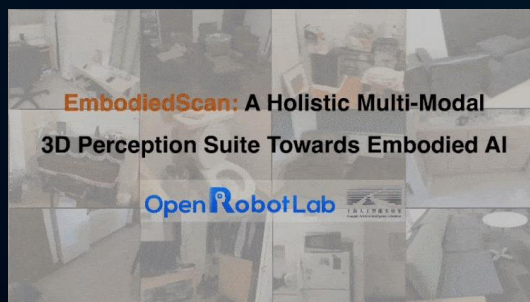


Motion Prior? & Data Scaling?

- **场景泛化**: Perception – 传感器方案, 表征方式, 融合方式
- **技能泛化**: Unified Model – Motion Prior? Scaling?

具身智能 + 人形机器人

数据基础



构建具身三维数据体系

城市级大规模场景，可交互物体
人体交互数据，多模态标注体系

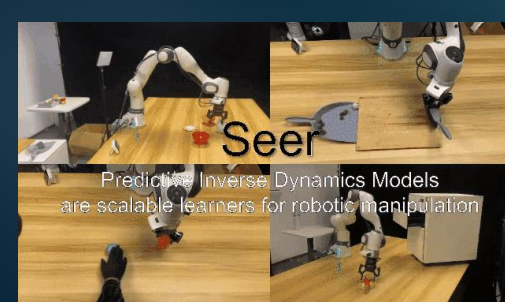
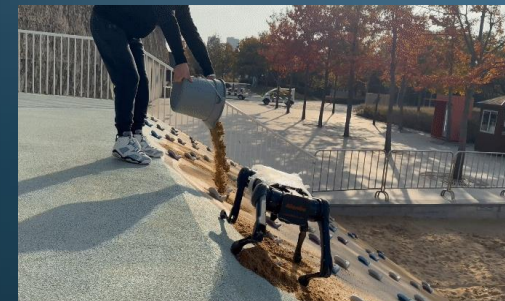
高效仿真



通用具身仿真平台与合成数据引擎

智能任务生成与评测
从交互中学习

虚实迁移



具身智能大模型

融合数据驱动与经典控制理论
实时交互，渐进式端到端

感谢聆听!

