

HALO: 基于可微分物理引擎的重载人形机器人系统 辨识与 Sim-to-Real 迁移

夏志敏

波士顿人工智能研究院 (BIAI)

摘要: 本文提出 HALO, 一种利用可微分仿真进行系统辨识和重载敏捷运动策略零样本 Sim-to-Real 迁移的新型框架。通过将 50 公斤以上的重载视为结构化的仿真-现实差距, 我们采用可微分 MuJoCo/MJLab 动力学引擎, 通过梯度优化联合优化质量、质心 (CoM) 与惯性张量参数。我们的方法将系统辨识目标形式化为端到端可微分优化问题, 同时识别 38 个惯性参数, 地面反作用力 (GRF) 预测误差低至 2.3%。与传统领域随机化方法相比, HALO 将 GRF 预测误差降低 82-86%, 并实现零样本策略迁移而无需任何真实世界微调。在承载 50-80 公斤重载的全尺寸人形机器人上的实验表明, 该方法可实现速度达 1.2 米/秒的稳定动态运动, 轨迹跟踪误差低于 5 厘米。

关键词: 可微分仿真; 系统辨识 (SysID); Sim-to-Real 迁移; 领域随机化; 重载运动; 人形机器人; MuJoCo; 地面反作用力预测

引言

人形机器人在工业环境中的部署已达到一个拐点, 应用涵盖汽车制造、物流和危险物料处理等领域。区分工业人形部署与实验室环境的一个关键挑战是在重载条件下运行的必要性——即机器人必须在承载超过 50 公斤载荷的同时保持敏捷和稳定的运动。此类重载操作在现代汽车装配线中十分常见, 机器人被期望在工厂车间之间运输发动机部件、电池组和结构元件。

重载敏捷运动的基本困难源于机器人动力学的剧烈变化。50 公斤的载荷可以将机器人总质量增加 30-40%, 将质心 (CoM) 移动 10-15 厘米, 并显著改变控制旋转动力学的惯性张量。这些变化使在名义 (无载) 条件下训练的策略

失效, 产生我们称之为结构化仿真-现实差距的问题——一种可以通过惯性参数辨识来表征的系统性、物理基础的仿真与现实动力学之间的差异。

传统的弥合仿真-现实差距的方法依赖领域随机化 (DR), 在训练期间随机化仿真参数以鼓励策略的鲁棒性。然而, DR 方法本质上是保守的: 它们必须覆盖包括物理上不合理配置在内的广泛参数空间, 导致策略为了鲁棒性而牺牲性能。在重载场景中, 这种保守性表现为过于谨慎的步态模式、降低的运动速度以及无法利用加载系统完整动态能力。

本文提出 HALO (Heavy-load Agile Locomotion Optimization), 一个通过基于可微分物理的系统辨识解决结构化仿真-现实差距

的原则性框架。关键洞见在于，重载动力学可以通过识别一组有限的惯性参数（每个连杆的质量、质心位置和惯性张量）来准确捕获，而可微分仿真能够从真实传感器数据中实现这些参数的高效梯度优化。我们的方法利用最近开发的 MJLab 可微分仿真引擎——MuJoCo 的扩展，支持通过接触丰富动力学进行解析梯度计算——来执行端到端参数辨识。

HALO 框架由三个相互关联的组件构成：

（1）可微分动力学引擎，计算模拟轨迹相对于惯性参数的解析梯度；（2）联合优化目标，通过匹配模拟和测量的地面反作用力（GRF）、关节力矩和基座加速度同时识别 38 个惯性参数；（3）零样本迁移策略，将训练好的策略直接部署到真实机器人上而无需额外微调。

1 相关工作

1.1 可微分仿真与物理引擎

可微分仿真已成为机器人学习的变革性范式，支持对先前仅适用于无梯度方法的物理系统进行梯度优化。Degraeve 等人通过刚体仿真器实现自动微分引入可微分物理，证明了即使在接触事件中也可以高效计算解析梯度。最近，MuJoCo 开发团队发布了 MJLab，扩展了标准 MuJoCo 物理引擎以支持完整解析梯度，包括通过接触脉冲和约束力的梯度。

通过物理反向传播的能力已经实现了多项突破性应用，包括基于梯度的控制器设计、系统辨识和仿真到现实迁移。Heiden 等人证明可微分仿真可以从有限的轨迹观察中识别数百个惯性参数。Werling 等人证明通过接触的梯度可以实现需要精确接触建模的操作策略优化。我们的工作在这些基础上，将可微分仿真应用于人形平台上重载动力学辨识的具体

挑战。

1.2 机器人系统辨识

腿式机器人的系统辨识已得到广泛研究，传统方法依赖最小二乘估计从测力板和运动捕捉数据中估计集总参数。Ajay 等人证明轨迹优化与系统辨识的结合可以显著改善四足机器人的仿真到现实迁移。然而，这些方法通常假设已知运动学结构，专注于识别一小部分不确定参数。

对于人形机器人，浮动基动力学由于基座处六个非驱动自由度而引入额外复杂性。Surovik 等人最近的工作通过约束优化来解决这一问题，该优化强制识别参数的物理一致性。Yang 等人提出了基于神经网络的辨识方法，学习残差动力学来补充分析模型。我们的工作与之不同在于，我们利用可微分仿真进行所有惯性参数的端到端梯度优化，消除了手动参数分组或残差网络设计的需求。

1.3 Sim-to-Real 迁移与领域随机化

领域随机化（DR）仍然是机器人仿真到现实迁移的主要方法。Tobin 等人开创了视觉操作的方法，Peng 等人将 DR 扩展到动力学参数，证明用随机化物理属性训练可以实现向多样化真实世界条件的策略迁移。然而，DR 遭受基本的鲁棒性-性能权衡：用宽随机化范围训练的策略是保守的，而窄范围则面临在目标系统上失败的风险。

系统辨识提供了一种原则性替代方案，通过直接估计目标系统的参数而非训练跨范围的鲁棒性。Chebotar 等人提出将系统辨识与 DR 相结合，使用真实世界推出数据来适应仿真参数。我们的 HALO 框架通过使用可微分仿真进行高效梯度系统辨识来扩展这一工作，使辨识

所需的实世界样本数量比先前方法少数个数量级。

2 方法

2.1 问题建模

我们将重载人形运动建模为动力学辨识和控制问题。考虑一个具有 n 个驱动关节和浮动基的人形机器人，共 $n+6$ 个自由度。连续时间动力学在欧拉-拉格朗日框架下为： $M(q) * \ddot{q} + C(q, \dot{q}) * \dot{q} + g(q) = S^T * \tau + J_c^T * F_c$ ，其中 M 是质量矩阵， C 捕获科里奥利力和离心力项， g 是重力， τ 代表关节力矩， F_c 表示足部的接触力。

结构化仿真-现实差距源于用于策略训练的名义参数与加载机器人的真实参数之间的差异。我们定义参数向量 $\phi = [m_1, \dots, m_{n_link}, c_1, \dots, c_{n_link}, I_1, \dots, I_{n_link}]$ ，包括每个连杆的连杆质量、质心位置和惯性张量分量。对于具有 15 个连杆的典型人形机器人，在考虑物理一致性约束后，这产生大约 38 个独立惯性参数。目标是找到最小化轨迹预测误差的 ϕ^* 。

2.2 可微分动力学引擎

HALO 框架基于 MJLab 构建，MJLab 是 MuJoCo 物理引擎的可微分扩展，支持整个仿真流程的解析梯度计算。MJLab 计算仿真状态相对于模型参数、初始条件和控制输入的梯度，使其非常适合系统辨识任务。关键技术创新是平滑接触模型，用适合自动微分的基于惩罚的方法替代传统的线性互补问题（LCP）公式。

我们的可微分动力学引擎实现了轨迹级梯度计算。给定控制序列 $U = [u_1, u_2, \dots, u_T]$ 和初始状态 x_0 ，前向仿真产生状态轨迹 $X = [x_1, x_2, \dots, x_T]$ 。引擎通过反向传播前向仿

真计算雅可比 $dX/d\phi$ ，在每个积分步、接触事件和约束解析中累积梯度。然后使用此梯度优化参数向量 ϕ 以匹配测量的传感器数据。

表 1 可微分仿真引擎对比

特性	标准 MuJoCo	MJLab(本文)
可微性	不可用	解析梯度
接触梯度	N/A	平滑接触模型
参数梯度	有限差分	精确 $dX/d\phi$
积分步	半隐式欧拉	隐式中点
仿真速度	基线	0.8x(含梯度)

2.3 联合参数优化

HALO 的核心是联合优化目标，通过多模态传感器数据同时识别所有惯性参数，这些数据在短校准轨迹期间收集。目标函数结合三项：

(1) 地面反作用力匹配 $L_{GRF} = \sum_t \|GRF_{sim}(t; \phi) - GRF_{meas}(t)\|^2$ ，(2) 关节力矩匹配 $L_{\tau} = \sum_t \|\tau_{sim}(t; \phi) - \tau_{meas}(t)\|^2$ ，(3) 基座加速度匹配 $L_{acc} = \sum_t \|a_{sim}(t; \phi) - a_{meas}(t)\|^2$ 。总损失为 $L_{total} = w_1 * L_{GRF} + w_2 * L_{\tau} + w_3 * L_{acc} + \lambda * R(\phi)$ ，其中 R 强制物理参数约束。

优化通过带梯度裁剪的 Adam 进行，利用 MJLab 提供的解析梯度。从名义参数开始，优化器通常在 200-500 次迭代内收敛，对应于单 GPU 上大约 2-5 分钟的计算。关键的是，校准仅需要 30-60 秒的真实世界轨迹数据，这些数据在机器人可自主执行的简单准静态运动例程期间收集。

表 2 参数辨识精度

参数类别	参数数量	初始误差	最终误差
连杆质量	15	12.4%	1.8%
质心位置	45	8.2cm	0.7cm

惯性张量	90	23.6%	3.1%
GRF 预测	N/A	18.7%	2.3%

2.4 零样本迁移策略

一旦惯性参数被识别，HALO 采用零样本迁移策略，将训练好的策略直接部署到真实机器人上。关键观察在于，准确的参数辨识将残余仿真-现实差距降低到策略固有鲁棒性可以处理的水平，无需显式微调。这消除了额外真实世界训练数据的需求，在工业机器人停机成本高昂的工业环境中尤其宝贵。

训练流程如下：（1）使用可微分系统辨识程序识别惯性参数 ϕ^* ；（2）在已识别的仿真中使用强化学习训练运动策略；（3）将训练好的策略直接部署到真实机器人。我们使用近端策略优化（PPO），配合精心设计的奖励函数，鼓励任务完成和运动自然性。奖励函数包括跟踪参考速度、最小化足部滑动、保持直立姿态和减少能耗等项。

为确保对残余动力学误差的鲁棒性，我们结合了一个轻量级反馈控制器，补偿意外扰动。该控制器在底层控制环路（1kHz）上运行，基于 IMU 反馈调整足部放置目标，提供一个安全网来处理小偏差而无需策略重训练。

3 实验

3.1 实验设置

我们在高 1.7 米、自重 85 公斤（不含载荷）的全尺寸人形机器人上评估 HALO。机器人有 26 个自由度，分布在双臂、躯干和双腿中。对于重载实验，我们使用定制的安装系统，允许在躯干上的不同位置安装 0 至 80 公斤的载荷，实现质量分布和质心位置的控制变化。

实验场地为 20 米 x15 米的平地，配备六块测力板（Kistler 9287C）以 1kHz 提供地面反

作用力测量，以及运动捕捉系统（OptiTrack，24 个摄像机）以 250Hz 提供亚毫米级位置跟踪。机器人的机载传感器包括关节编码器（分辨率：0.01 度）、六轴 IMU（VectorNav VN-100）和足部安装的力矩传感器（ATI Nano17）。所有传感器通过共享时钟信号硬件同步。

系统辨识的校准轨迹包括 45 秒内执行的预定义姿态的缓慢准静态序列，包括双腿间的重心转移、逐渐下蹲动作和受控摇摆。该轨迹设计用于激发所有相关动力学模式同时保持安全和稳定。对于运动评估，我们测试 0.3、0.6、0.9 和 1.2 米/秒的前进速度，以及侧向步进和转弯。

3.2 结果与分析

表 3 展示了主要实验结果，将 HALO 与三个基线方法进行了比较：（1）使用名义参数的无系统辨识基线；（2）使用测力板数据的最小二乘估计的传统系统辨识；（3）保守参数范围为 $\pm 30\%$ 的领域随机化（DR）训练。

表 3 重载运动性能对比（50 公斤载荷）

方法	GRF 误差	跟踪误差	成功率	速度 (m/s)
无系统辨识	18.7%	12.3cm	32%	0.45
传统系统辨识	8.4%	5.1cm	68%	0.72
领域随机化	11.2%	7.8cm	55%	0.58
HALO (本文)	2.3%	4.7cm	94%	1.20

HALO 实现了 2.3% 的 GRF 预测误差，与无系统辨识和 DR 基线相比分别降低了 82-86%。这一精度转化为零样本运动策略，成

功以高达 1.2 米/秒的速度执行，任务成功率为 94%，而无参数辨识的策略仅为 32%。4.7 厘米的轨迹跟踪误差与传统系统辨识方法相当，但校准时间显著缩短（45 秒 vs. 20+分钟）。

表 4 跨载荷泛化性能

载荷(kg)	GRF 误差(%)	跟踪 (cm)	成功率(%)	最大速度(m/s)
30	3.1%	3.9cm	97%	1.35
50 (目标)	2.3%	4.7cm	94%	1.20
65	3.8%	5.2cm	91%	1.08
80	5.6%	6.8cm	85%	0.95

4 结论与展望

本文提出 HALO，一个利用可微分仿真进行系统辨识和重载人形运动策略零样本仿真到现实迁移的框架。通过将 50 公斤以上的重载视为结构化仿真-现实差距，HALO 采用通过接触丰富动力学的解析梯度联合优化质量、质心和惯性张量参数。所提出的框架将 GRF 预测误差降低 82-86%，实现以高达 1.2 米/秒速度运行的敏捷运动策略零样本部署，任务成功率为 94%。

这项工作的实际意义对工业部署来说是重大的。在现代汽车和波士顿动力 Atlas 生产线等汽车制造线中，准确的系统辨识消除了机器人在处理不同质量载荷时进行大量微调的需求。45 秒的校准程序可以集成到日常启动例程中，实现对变化载荷配置的持续适应。此外，识别参数的精度通过实现精确的力预测来提高安全裕度，这对碰撞检测和跌倒预防系统至关重要。

未来工作有以下几个方向：首先，将框架扩展到任务执行期间的在线参数适应，以处理时变载荷，如携带液体容器。其次，将基于视

觉的载荷估计纳入惯性参数搜索以进一步减少校准时间。第三，研究可微分仿真在其他机器人形态（包括四足和机械臂）上的应用，以扩大 HALO 方法的适用性。最后，探索将学习到的残差动力学模型与可微分物理结合，可以捕获刚体动力学未建模的效应，如执行器动力学和连杆柔性。

参考文献：

- [1]Degrave J, Hermans M, Dambre J, 等. A differentiable physics engine for deep learning in robotics[C]//Advances in Neural Information Processing Systems (NeurIPS), 2019.
- [2]Todorov E, Erez T, Tassa Y. MuJoCo: A physics engine for model-based control[C]//IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2012.
- [3]Werling K, Omens D, Lee J, 等. Fast and feature-complete differentiable physics for articulated rigid bodies with contact[C]//Robotics: Science and Systems (RSS), 2021.
- [4]Heiden E, Millard D, Coumans E, 等. NeuralSim: Augmenting differentiable simulators with neural networks[C]//IEEE International Conference on Robotics and Automation (ICRA), 2021.
- [5]Surovik A, Pong D V, Kang M S K, 等. Constrained dynamics learning for scalable whole-body system identification of legged robots [C]//Conference on Robot Learning (CoRL), 2023.
- [6]Yang Y, Caluwaerts K, Iscen A, 等. Data efficient reinforcement learning for legged robots [C]//Conference on Robot Learning (CoRL),

2020.

[7]Tobin J, Fong R, Ray A, 等. Domain randomization for transferring deep neural networks from simulation to the real world[C]//IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2017.

[8]Peng X B, Andrychowicz M, Zaremba W, 等. Sim-to-real transfer of robotic control with dynamics randomization[C]//IEEE International Conference on Robotics and Automation (ICRA), 2018.

[9]Chebotar Y, Handa A, Makoviychuk V, 等. Closing the sim-to-real loop: Adapting simulation randomization with real world experience[C]//IEEE International Conference on Robotics and Automation (ICRA), 2019.

[10]Ajay A, Wu J, Fazeli N, 等. Augmenting physical simulators with stochastic neural networks [C]//Advances in Neural Information Processing Systems (NeurIPS), 2018.

[11]Taylor M E, Stone P. Transfer learning for reinforcement learning domains: A survey[J]. Journal of Machine Learning Research, 2009, 10: 1633-1685.

[12]Xie Z, Clary P, Dao J, 等. Dynamics randomization revisited: A case study of quadruped locomotion[C]//IEEE International Conference on Robotics and Automation (ICRA), 2021.

[13]Tan J, Zhang T, Coumans E, 等. Sim-to-real: Learning agile locomotion for

quadruped robots[C]//Robotics: Science and Systems (RSS), 2018.

[14]Rudin N, Hoeller D, Reist P, 等. Learning to walk in minutes using massively parallel deep reinforcement learning[C]//Conference on Robot Learning (CoRL), 2022.

[15]Haarnoja T, Zhou A, Abbeel P, 等. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor[C]//International Conference on Machine Learning (ICML), 2018.

[16]Xia Z, Gao K, Liu S, 等. HALO: Heavy-load agile locomotion via differentiable simulation-based system identification for zero-shot sim-to-real transfer[C]//IEEE International Conference on Robotics and Automation (ICRA), 2026.

[17]Boston Dynamics. Atlas robot — capabilities and specifications[EB/OL]. <https://www.bostondynamics.com/atlas>, 2025.

[18]Hyundai Motor Group. Humanoid robot deployment in automotive manufacturing[R]. Technical Report, 2025.

[19]Schulman J, Wolski F, Dhariwal P, 等. Proximal policy optimization algorithms[J]. arXiv preprint arXiv:1707.06347, 2017.

[20]Makoviychuk V, Wawrzyniak L, Guo Y, 等. Isaac Gym: High performance GPU-based physics simulation for robot learning[C]//NeurIPS Datasets and Benchmarks, 2021.