arXiv:2603.01151v1 [cs.RO] 1 Mar 2026

D-REX：用于学习灵巧抓取的可微真实到仿真再到现实引擎

Haozhe Lou[1, 2*], **Mingtong Zhang**[1, 2*], **HaoranGeng**[3], **Hanyang Zhou**[2], **Sicheng He**[1,2], **ZhiyuanGao**[1,2] **Siheng Zhao**[1,2], **Jiageng Mao**[1,2], **Pieter Abbeel**[3], **Jitendra Malik**[3], **Daniel Seita**[2], **YueWang**[1,2]物理超智能（PSI）实验室，南加州大学[2]维特比工程学院，南加州大学[3]电气工程与计算机科学系，加利福尼亚大学伯克利分校

摘要

仿真为数据生成和策略学习提供了一个经济高效且灵活的平台，用以开发机器人系统。然而，弥合仿真与真实世界动力学之间的差距仍然是一项重大挑战，尤其是在物理参数识别方面。在本工作中，我们引入了一种真实到仿真再到真实引擎，该引擎利用高斯泼溅表示构建可微分引擎，从而能够从真实世界视觉观测和机器人控制信号中进行物体质量识别，并同时实现抓取策略学习。通过优化被操作物体的质量，我们的方法能够自动构建高保真且物理上合理的数字孪生。此外，我们提出了一种新颖方法，通过将可行的人类示范转化为仿真机器人示范，从有限数据中训练力感知抓取策略。通过全面的实验，我们证明该引擎在各种物体几何形状和质量值上的质量辨识中均取得了准确且鲁棒的性能。这些优化后的质量值促进了力感知策略学习，在物体抓取中实现了优越的高性能，有效缩小了仿真到现实差距。我们的代码和项目页面可访问drex.github.io.

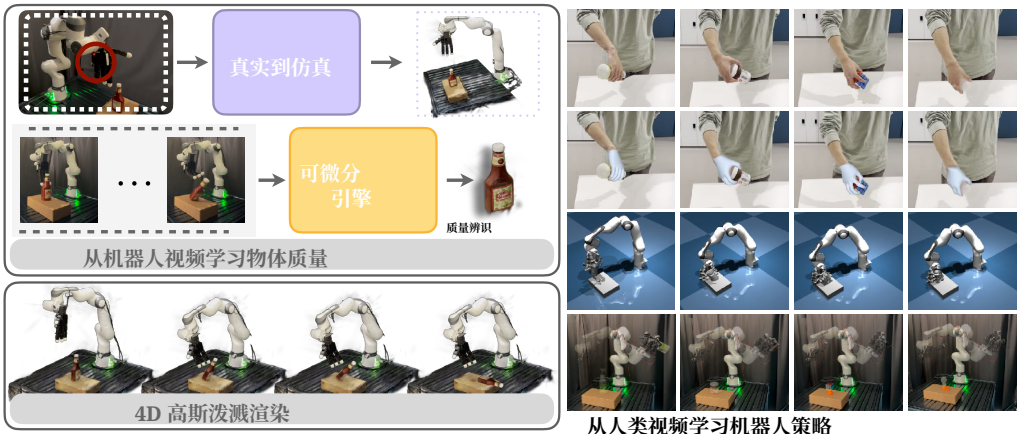


图 1：我们提出了 D-REX，这是一种可微的真实到仿真再到真实引擎，通过根据真实世界视觉观测和机器人交互数据辨识物体质量，实现 4D 照片级真实感渲染和物理仿真。D-REX 使用高斯泼溅表示重建物体几何形状，并利用可微物理引擎进行端到端质量辨识。随后，利用识别出的质量，从人类演示中进行力感知策略学习，从而支持灵巧抓取任务中的鲁棒抓取和仿真到现实迁移。

1 引言

仿真已成为机器人学的重要平台，提供了一种经济高效且可扩展的平台，减少了对深厚机器人学专业知识的依赖。通过可复用且可控的数据生成，仿真在加速策略学习方面取得了显著进展Akkaya

*共同贡献。

等人 (2019）； Hafner 等人 (2023）； Chen 等人(2021）； Agarwal 等人 (2023）； He 等人 (2024）； Ma 等人 (2023）。然而，尽管有这些优势，要复现真实世界的视觉真实感和复杂物理动力学仍是一项重大挑战。高保真物理仿真通常需要专业知识和复杂建模，这限制了基于仿真的方法在现实世界部署中的可扩展性和鲁棒性。

长期以来，研究者一直致力于弥合仿真到现实差距；当将在仿真中训练的模型迁移到真实世界配置中时，便会产生这一差距。它仍是机器人学中的一项根本性挑战。基于仿真的策略通常假设已准确了解并建模真实世界配置，包括其底层物理参数。然而，视觉观测所估计的几何形状和质量与其真实世界取值之间的差异，会扩大仿真到现实差距。为缓解这一差距，现有策略包括域随机化 Tobin 等人（2017）；Sadeghi 和 Levine（2016）；Peng 等人（2018a），通过改变仿真参数来增强鲁棒性；以及系统辨识 Hwangbo 等人（2019）；Tan 等人（2018）；Khalil 等人（2007），通过真实世界观测校准来完善仿真动力学。仿真保真度的提升 Ho 等人（2020）；Mittal 等人（2023）以及域适应 Bousmalis 等人（2018）；Ren 等人（2023）；Chen 等人（2023），进一步促进了机器人应用中模型从仿真向现实的迁移。作为这些工作的补充，真实到仿真框架致力于构建能够高精度复现真实世界几何结构和动力学的数字孪生 Chen 等人（2024）；Torne 等人（2024）；Jiang 等人（2022）。尽管如此，构建精确的数字孪生通常需要整合多种方法，例如几何重建和参数辨识。尽管取得了这些进展，当前的真实到仿真方法仍难以根据视觉观测实现精确建模。

这一挑战与系统辨识问题密切相关：需要根据视觉观测推断物理参数，以确保仿真环境能够忠实地反映真实世界动力学。即使能够获取完整的系统状态，从图像中估计物体属性和系统动力学仍然十分困难。尽管存在鲁棒的正向模拟器 Macklin 等人 (2014)，但其不可微分性限制了其在逆问题中的应用。有限差分等替代梯度方法被广泛采用 Cranmer 等人 (2020); Ramos 等人(2019); Wu 等人 (2017)，但在高维场景中扩展性较差。近期可微仿真领域的进展提升了学习效率。特别是，GradSim Jatavallabhula 等人 (2021a) 实现了从视觉观测到物体级物理参数的端到端求导。受此启发，我们的工作直接从视频中优化物体质量，从而实现以质量为条件的力感知抓取策略学习，并显著提升性能。

为应对这些挑战，本文提出 D-REX，这是一种可微分的真实到仿真再到真实框架，其仿真引擎基于可微仿真构建 Jatavallabhula 等人(2021a); Freeman 等人 (2021); Müller 等人 (2007); Macklin 等人 (2016) 以及高斯泼溅表示 Kerbl 等人 (2023)。这一可微分引擎能够通过机器人与物体交互中的视觉观测和机器人控制信号，识别物体质量。此外，我们提出一种用于灵巧操作的新型基于学习的方法：将人类演示转化为可在仿真中执行的机器人演示，然后利用所提出的方法同时优化抓取位置和力度。

我们的主要贡献包括：

- 一种真实到仿真再到真实框架，通过视觉观测和机器人控制信号，利用可微仿真实现端到端的物体质量辨识。
- 一种从人类演示数据中学习抓取策略的新方法，该策略以已辨识的物体质量为条件，并整合位置与力控制，以缩小仿真到现实差距，实现稳健、高性能的抓取。
- 通过实验证明，利用我们的可微框架准确辨识质量，并以此作为策略的条件，能够提升对具有挑战性的物体进行灵巧抓取的效果。

2 相关工作

2.1 面向机器人学的可微物理仿真

物理仿真的发展使机器人学能够高效地进行数据生成和策略训练 Liu 与 Negrut（2021）；Xu 等人 （2021；2022）。具体而言，可微物理仿真近年来取得了重大进展，因为它们为策略学习提供了高效的梯度。一种流行的方法是开发具有自动可微程序-

设计 de Avila Belbute-Peres 等人（2018年）；Hu 等人（2019）；Xu 等人（2022）；Li 等人（2025a）。另一条研究路线侧重于学习神经网络以近似真实世界动力学 Li 等人（2019）；Pfaff 等人（2020）；Xian 等人（2021），这些网络本质上可微，适用于规划与控制优化中的应用。在机器人应用方面，已研究了多种下游任务：流体操控 Xian 等人（2023）、软体操控 Huang 等人（2021）、布料操控 Peng 等人（2024）；Yu 等人（2023），以及软体机器人形态与控制策略的协同优化 Bhatia 等人（2021）；Li 等人（2023a）；Li 等人（2025b）。值得注意的是，Jatavallabhula 等（2021a）提出利用可微多物理场仿真从像素进行系统辨识。我们的框架提出了一种新视角，即通过可微方法将从视觉观测获得的梯度反向传播用于系统辨识，并使我们能够基于真实到仿真结果进行灵巧操作策略学习。

2.2 真实到仿真再到真实的迁移

真实到仿真使我们能够在仿真中复现真实世界资产及其动力学，从而增强数据驱动洞见、优化效果和机器人能力。通过捕捉自然统计特性、动态行为和运动学结构，该方法支持稳健决策、高效模型训练以及复杂场景评估。近期的多项研究体现了这些趋势。Jiang 等人（2022）创建了用于仿真的关节物体交互式数字孪生。Chen 等人（2024）；Mandi 等人（2024）根据图像生成关节仿真，而 Sundaresan 等人（2022）则利用点云调整可变形物体的参数。神经辐射场也已应用于操作和运动等机器人任务（Kerr 等人，2022；Rashid 等人，2023；Zhou 等人，2023；Byravan 等人，2023；Wang 等人，2023），但通常无法准确体现物理真实性。更近期的研究（Abou-Chakra 等人，2024；Zhang 等人，2024a；Kerr 等人，2024；Jiang 等人，2025年；Zhobro 等人，2025年；Abou-Chakra 等人，2025年；Yang 等人，2025年；Xie 等人，2023）利用高斯泼溅，根据真实世界视觉输入构建数字孪生。Pfaff等人（2025）和 Khalil 等人（2007）通过关节扭矩传感辨识机器人及负载参数。与先前研究（Zhu 等人，2025年；Turpin 等人，2022）不同，后者往往因表示方式或仿真引擎的局限，未能与可微真实到仿真再到真实框架集成；我们的方法则支持在仿真中准确识别物体质量并学习策略，以便直接部署于现实世界。

2.3 从人类视频中学习机器人操作

人类演示视频为机器人操作提供了可扩展且语义丰富的数据来源。然而，由于具身形态和感知方式存在差异，将人类动作映射为机器人控制仍然具有挑战性。以往工作通过利用预训练视觉表征 Nair 等人（2022）；Radosavovic 等人（2023）；Ma 等人（2022），或提取可供性 Bahl 等人（2023）和以物体为中心的光流 Xu 等人（2024）等中间线索来弥补这一差距，但这些方法难以实现精细且灵巧的操作。其他研究则聚焦于三维人体运动估计，以实现技能迁移 Shaw 等人（2023a）；Patel 等人（2022）；Peng 等人（2018b）；Lum 等人（2025），但这类方法往往受限于人类与机器人具身形态差距。近期方法尝试放宽这一限制：Guzey 等人（2024）基于物体跟踪构建奖励函数，Singh 等人（2024）则从野外视频中生成三维手物体轨迹，以进行重定向。尽管这些方法前景可观，但在从人类视频中进行可泛化策略学习方面仍面临困难。与之相比，我们的框架将人类示范迁移为可在仿真中执行的机器人示范，从而实现可扩展的策略学习，并根据物体质量进行条件化，以提升适应性和性能。

3 问题陈述

我们聚焦于现实到仿真再到现实任务，旨在构建一个能够高度还原真实世界几何结构、物理属性和外观的仿真环境。我们假设可以获取几种类型的 RGB 视频：以场景为中心的视频序列，记为 $\mathcal{I}_s$，用于捕捉静态环境；以物体为中心的视频，记为 $\mathcal{I}_o$，提供被操作物体的多个视角，以支持准确的视觉与几何重建；以及人类演示视频 $\{\mathcal{I}_t\}_{t=1}^{T}$，用于展示任务执行过程。我们还从真实世界的机器人运行轨迹中提取物体轨迹，记为 $\{\mathbf{s}_t^{\text{real}}\}_{t=1}^{T}$，并在我们的框架中模拟相应的轨迹 $\{\mathbf{s}_t^{\text{sim}}\}_{t=1}^{T}$。

场景和物体视频（$\mathcal{I}_s$、 $\mathcal{I}_o$）通过视觉与几何重建用于初始化真实到仿真过程。真实环境和仿真环境中的策略展开轨迹通过可微分引擎实现质量辨识，从而得到优化后的物体质量 m，并将其作为物理参数纳入仿真器。同时，人类演示视频 $\{\mathcal{I}_t\}_{t=1}^{T}$ 被转换

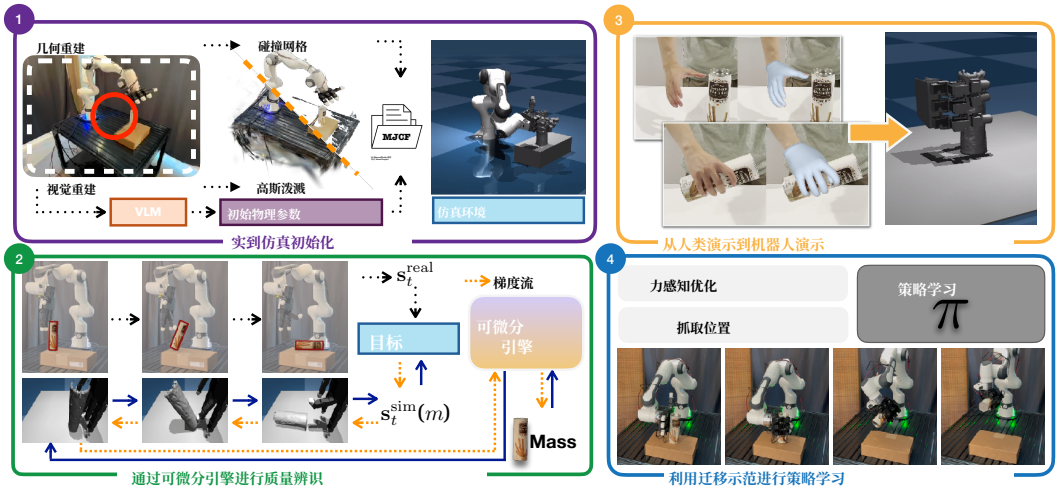


图 2：**我们方法的概述**。我们的方法由四个组件构成：(1) 真实到仿真，(2) 质量辨识，(3) 从人类演示中学习，以及 (4) 策略学习。我们首先采集场景视频和人类演示数据。随后，在仿真和真实世界中执行机器人动作，通过我们的可微物理引擎辨识物体质量。最后，利用演示数据和辨识出的质量训练操作策略。

使用我们提出的方法转换为机器人可执行的轨迹 $\{\mathbf{A}_t\}_{t=1}^T$，并用于训练具备力感知能力的操作策略 π。

4 方法

我们提出了一种真实到仿真再到真实框架，该框架利用可微分引擎，根据视觉观测和机器人控制信号构建准确的仿真环境，并通过系统辨识确定物体质量，从而实现稳健的策略学习和仿真到现实迁移。该框架基于 MuJoCo（Todorov 等人（2012）），这是一款通用物理仿真器，并结合了可微分引擎 Brax（Freeman 等人（2021））和 GradSim（Jatavallabhula 等（2021a））。该框架分为图 2 所示的四个步骤：真实到仿真（第 4.1 节）、质量辨识（第 4.2 节）、从人类演示中学习（第 4.3 节）以及策略学习（第 4.4 节）。首先，根据 RGB 视频 $\mathcal{I}_s$ 和 $\mathcal{I}_o$ 重建场景和物体，捕捉静态环境和目标物体。输出的仿真环境以 MJCF 格式表示为 $\mathcal{S} = \{\mathcal{K}, \boldsymbol{\theta}\}$，其中 $\mathcal{K}$ 表示碰撞网格，$\boldsymbol{\theta}$ 表示物理参数。接下来，该框架在仿真环境和真实世界中执行一致的机器人动作，收集 $\{\mathbf{s}_t^{\text{real}}\}_{t=1}^T$ 和 $\{\mathbf{s}_t^{\text{sim}}\}_{t=1}^T$，以辨识物体质量 m。第三，将人类演示 $\{\mathcal{I}_t\}_{t=1}^T$ 转换为机器人可执行的轨迹 $\{\mathbf{A}_t\}_{t=1}^T$。最后，利用这些轨迹在仿真环境中训练策略，然后将其直接部署到真实世界中。

4.1 视觉与几何重建

我们的框架首先重建操作环境中关键元素的高保真视觉模型和几何模型，包括 $\mathcal{I}_s$物体、机械臂、灵巧手和工作空间。这种重建可确保碰撞几何和视觉外观都得到准确表示。为将这些模型整合到可微仿真中，我们采用高斯泼溅表示Kerbl 等人(2023)；Huang 等人(2021)，该表示可实现照片级真实感渲染，并能按照Lou 等人(2024)的方法生成用于碰撞检测的高质量网格。具体而言，我们处理由移动设备采集的视频，训练两组高斯基元：一组用于碰撞几何，另一组用于视觉外观。具体而言，带有表面法线估计的二维高斯泼溅Ye 等人(2024a)可为仿真提供准确的几何形状，而三维高斯泼溅则确保高保真渲染。该过程会产生两个互补的输出：碰撞网格 $\mathcal{K}$ 和高斯粒子 $\mathcal{P}$。更多细节请参见附录。

4.2 基于机器人-物体交互的物理参数识别

准确辨识物理参数 $\boldsymbol{\theta}$ 对于构建物理上合理的仿真至关重要。我们首先使用视觉-语言模型 Hurst 等人(2024生成初始 MJCF 表示 $\mathcal{S}$，依据环境图像和提示词Zhang 等人(2024b。尽管这种方法提供了合理的结构先验，但由于缺乏可观测的物理线索，仅依据视觉输入推断出的参数往往与真实值存在偏差Asenov 等人(2020)。

为此，我们专注于辨识物体质量。它是动力学中的关键参数，并且能够可靠地测量。准确的质量辨识可以提高仿真保真度，并促进稳健的策略学习。我们选择一项基于虚拟支点假设的平面推动任务，以减少摩擦效应，并优化物体质量 m，从而最大限度地减小仿真轨迹与真实世界轨迹之间的差异Jatavallabhula 等人(2021a)：

$$\min_{m>0} \mathcal{L}_{\text{traj}}(m) := \sum_{t=1}^{T} \left\| \mathbf{s}_t^{\text{sim}}(m) - \mathbf{s}_t^{\text{real}} \right\|_2^2, \tag{1}$$

其中$\mathbf{s} = [\mathbf{p}, \mathbf{q}]^\top \in \mathbb{R}^7$ 表示物体的六自由度位姿，由位置$\mathbf{p} \in \mathbb{R}^3$ 和以单位四元数表示的方向$\mathbf{q} \in \mathbb{R}^4$。$\mathbf{s}_t^{\text{real}}$ 由 FoundationPose Wen 等人(2024) 在真实世界中估计，而 $\mathbf{s}_t^{\text{sim}}(\mathbf{m})$ 则通过在仿真中执行相同动作获得。

动力学建模。为模拟物体运动，我们采用牛顿-欧拉机制的标准刚体形式。设$\mathbf{u}_t = [\mathbf{v}_t, \boldsymbol{\omega}_t]^\top$ 表示物体在时间步 t的速度，其中$\mathbf{v}_t$ 和 $\boldsymbol{\omega}_t$ 分别表示线速度分量和角速度分量。我们将控制方程表示为二阶常微分方程（ODE）Chen 等人(2019)：

$$\mathbf{M}(\mathbf{s}_t, \mathbf{u}_t, m, \boldsymbol{\theta})\, \dot{\mathbf{u}}_t \;=\; \mathbf{f}(\mathbf{s}_t, \mathbf{u}_t, \boldsymbol{\theta}), \tag{2}$$

其中$\mathbf{M}$ 是质量惯性矩阵Baraff (1992)，而$\mathbf{f}$ 汇总外力和接触力方程3、重力和力矩。我们采用以刚度和阻尼为参数的柔顺惩罚式接触模型$(k_e, k_d) \in \boldsymbol{\theta}$，该模型施加的法向力与穿透深度和接触速度成正比Todorov 等人(2012);Erez 等人(2015：

$$\mathbf{f}_n(\mathbf{s}, \mathbf{u}_t, \boldsymbol{\theta}) \;=\; -\,\mathbf{n}\left(k_e\, C(\mathbf{s}) + k_d\, \dot{C}(\mathbf{u})\right), \tag{3}$$

其中 $\mathbf{f}_n$ 为接触力，$\mathbf{n}$ 为接触法向，$C(\mathbf{s})$ 为穿透深度，且 $\dot{C}(\mathbf{u})$是 $C(\mathbf{s})$的导数。该接触模型是可微的，并可方便地集成到我们的仿真框架中。实践中，我们采用离散时间更新来实现动力学 Erez 等人 (2015)：

$$\mathbf{s}_{t+1}^{\text{sim}} = G\big(\mathbf{s}_t^{\text{sim}}, \mathbf{u}_t,\, m,\, \mathbf{f}_t\big), \quad t = 0, \ldots, T-1. \tag{4}$$

$\mathbf{f}_t$ 是时间步 t处的外力，包括执行器冲量、重力和物体与地面的接触。**可微物理**。为优化方程1，我们计算模拟轨迹相对于物体质量 m的梯度。遵循 Jatavallabhula 等人(2021a)的离散伴随方法，我们采用半隐式欧拉积分方案，以保证接触动力学下的稳定性。我们将 MjX/Brax Freeman 等人 (2021)的运动学与刚体动力学方程 2 以及接触模型方程 3相结合，形成可微计算图 Hu 等人 (2020)。

半隐式欧拉建模。更新函数 $G(\cdot)$ 在方程 4 中采用半隐式欧拉积分方案实现：

$$G\Big(\big[\mathbf{s}_t,\, \mathbf{u}_t\big], m,\, \boldsymbol{\theta}\Big) = \begin{bmatrix} \mathbf{s}_t \;+\; \Delta t\, \mathbf{u}_{t+1} \\ \mathbf{u}_{t+1} \end{bmatrix} = \begin{bmatrix} \mathbf{s}_t \;+\; \Delta t\Big(\mathbf{u}_t + \Delta t\, \mathbf{M}^{-1}(\mathbf{s}_t, \mathbf{u}_t, m, \boldsymbol{\theta})\, \mathbf{f}\big(\mathbf{s}_t,\, \mathbf{u}_t\big)\Big) \\ \mathbf{u}_t + \Delta t\, \mathbf{M}^{-1}(\mathbf{s}_t, \mathbf{u}_t, m, \boldsymbol{\theta})\, \mathbf{f}\big(\mathbf{s}_t,\, \mathbf{u}_t\big) \end{bmatrix} \tag{5}$$

其中 Δt 是积分时间步，$f(\cdot)$同时包含外力与接触力 .

可微真实到仿真再到真实优化。我们从初始条件$\mathbf{s}_0^{\text{sim}}$开始对该系统进行仿真，并通过方程4迭代更新状态。为量化仿真轨迹与真实世界轨迹之间的差异，我们将仿真状态$\mathbf{s}_t^{\text{sim}}$与对应的真实世界状态$\mathbf{s}_t^{\text{real}}$之间的轨迹损失定义为：

$$\mathcal{L}_{\text{traj}}(m) = \sum_{t=0}^{T} \left\| \mathbf{s}_t^{\text{sim}} - \mathbf{s}_t^{\text{real}} \right\|_2^2. \tag{6}$$

该目标促使以质量为参数的模拟轨迹 m随着时间推移，尽可能贴近观测到的现实世界动力学。梯度 $\nabla_m \mathcal{L}_{\text{traj}}(m)$ 通过自动微分计算，并采用 PyTorch Paszke 等人 (2019) 中实现的反向传播，具体如下：

$$\frac{\partial \mathcal{L}_{\text{traj}}}{\partial m} = \sum_{t=1}^{T} \frac{\partial \mathcal{L}_{\text{traj}}}{\partial \mathbf{s}_t^{\text{sim}}} \cdot \frac{\partial \mathbf{s}_t^{\text{sim}}}{\partial \mathbf{M}_t} \cdot \frac{\partial \mathbf{M}_t}{\partial m}. \tag{7}$$

不同于 GradSim 等系统辨识方法 Jatavallabhula 等人 (2021a)，这些方法依赖于人工指定的外力，而我们的方法通过直接利用仿真和真实世界中一致的机器人控制信号来建模施加于物体的外力，从而支持端到端优化。这使真实世界与模拟轨迹紧密耦合，使我们能够通过机器人与物体交互捕捉接触动力学。

重要的是，我们的方法不需要物体真实质量或接触点。物体几何和位姿通过第4.1节获得，而执行器信号和机器人与物体交互则由MJCF运动学模型推导得出。这些作为我们可微框架的输入，用于精确的质量优化。更多建模细节和实验见附录。

4.3 将人类示范迁移为机器人演示

在精确建模场景和物体之后，下一步是采集真实世界的人类示范，并将其迁移为用于策略学习的机器人演示。尽管直接从人类示范中学习很直观，但人手与机器人手之间的显著差异使抓取交互迁移变得复杂，尤其是由于物体几何形状和质量各不相同。

我们的方法旨在将从RGB 视频序列中采集的人类示范 $\{\mathcal{I}_t\}_{t=1}^{T}$转换为仿真中可执行的机器人演示。每个视频帧 $\mathcal{I}_t$ 均使用HaMeR Pavlakos 等人（2024）和MCC-HO Wu 等人（2024a）进行处理，以重建人手和被操作物体的详细关节模型。在每个时间步 t，这些方法输出：

$$\mathbf{h}_t \in \mathrm{SE}(3) \times \mathbb{R}^{J_h}, \quad \mathbf{o}_t \in \mathrm{SE}(3), \tag{8}$$

其中 $\mathbf{h}_t$ 编码六自由度腕部位姿和手指关节角度（J_h），而 $\mathbf{o}_t$ 描述物体的六自由度位姿。随后，我们使用 Dex-Retargeting Qin 等人 (2023) 将这些人手-物体位姿 $\mathbf{h}_t$,$\mathbf{o}_t$ 映射到具有 J_r 个自由度的机器人手上。由此生成机器人动作：

$$\mathbf{A}_t = \mathcal{R}(\mathbf{h}_t, \mathbf{o}_t) \in \mathbb{R}^{J_r}, \tag{9}$$

其中 $\mathbf{A}_t$ 表示机器人执行器的目标关节角。鉴于我们假设人类演示与机器人操作中的物体几何形状保持一致，所得动作集 $\mathbf{A}_t$ 可直接作为策略学习的数据来源。

4.4 利用迁移的机器人演示数据进行策略学习

我们使用第 $\{\mathbf{A}_t\}_{t=1}^{T}$ 节中所述的机器人演示数据初始化策略学习 4.3。每段演示数据都将重建物体的碰撞网格顶点 $\mathcal{K}$ 作为输入，对应到相应的机器人抓取位姿。这些观测-动作对直接用于监督操作策略 π_ϕ的训练，该策略将以物体为中心的观测映射为灵巧抓取配置。

抓取位置学习。 为捕捉物体的几何形状和位姿，我们对其碰撞网格的顶点进行编码 $\mathcal{K}$ 使用位置编码 Tancik 等人 (2020)，构成策略的输入。该策略以重建后的碰撞网格 $\mathcal{K}$ 和识别质量 m为条件。具体而言，π_ϕ 是一个多头神经网络，可预测灵巧手的关节位置$\hat{}$ $\hat{}\mathbf{A}$、接触相关奖励 $\hat{}\mathbf{r}$ 以及质量相关控制力 $\mathbf{f}$。

$$\pi_\phi(\mathbf{o}) = \begin{bmatrix} \hat{\mathbf{A}} \\ \hat{\mathbf{r}} \\ \hat{\mathbf{f}} \end{bmatrix} \in \mathbb{R}^{19}, \quad \hat{\mathbf{f}} = \frac{m \cdot g}{n_{\text{active}}}. \tag{10}$$

其中，$\hat{}\mathbf{A} \in \mathbb{R}^{16}$ 表示预测关节位置，$\hat{}\mathbf{r} \in \mathbb{R}^{2}$ 表示接触约束，$\hat{}\mathbf{f} \in \mathbb{R}$ 表示抓取力约束。变量 n_{active} 表示机器人手与物体之间活动接触点的数量。该网络采用带有 ReLU 激活的全连接层，后接一个池化层。更多细节请参见附录。

力感知优化设计。 在训练开始阶段，我们定义了一个二维接触约束$\hat{}\mathbf{r}$：其中一项用于鼓励在滚动执行过程中持续保持接触，另一项则确保在结束时保持物体。策略会根据通过我们的仿真引擎与仿真资产实现的手物交互，动态影响该约束 S，详见第4.1 节和4.2节。当整个时间范围内的有效接触点数超过阈值 $N_{\min}$ 时，该约束保持较高值 H：

$$\forall t \in [t_0, t_0 + H] : n_{\text{active}}(t) \geq N_{\min}, \quad \mathbb{I}_{\text{in_hand}}(t) = \begin{cases} 1, & \text{if } n_{\text{active}}(t) \geq 1, \\ 0, & \text{otherwise}, \end{cases} \tag{11}$$

随后，我们使用10中的基于力的约束重新训练操作策略，从而实现能够响应物体质量变化的自适应力控制Vassiliadis 等人(2021)；Zhang 等人(2025年)。这提升了从演示数据中学习到的抓取姿态的鲁棒性，确保其在不同动力学条件下保持稳定。

传统的基于位置的策略会复现人类演示中的抓取位姿 Patel 等人(2022); Lum 等人 (2025年); Chen 等人 (2025a); Wan 等人 (2023); Wei 等人 (2025年); Shaw 等人 (2022)，但忽略了未观测到的力，尤其是抵消重力的力。在不同物体质量下施加均匀的力往往会导致不稳定。为解决这一问题，我们提出了一种混合控制框架，将位置与力控制相结合，并使用以优化后的质量为条件的预测模块 m。该框架联合优化策略参数 ϕ 和抓取力，从而实现更稳健且符合物理规律的操作。

5 实验

我们的实验旨在从以下几个关键方面评估系统的性能：(1) 评估利用可微分引擎在不同物体几何形状、密度和类别下进行质量辨识的有效性与稳健性。(2) 分析引入物体质量对策略学习性能的影响，并评估基于力的控制的可行性。(3) 评估从迁移的机器人演示数据中学习抓取策略，以及直接进行仿真到现实部署的有效性。

5.1 通过物体推动进行质量辨识

我们通过物体推动实验评估质量辨识方法，分别在真实世界和仿真中施加相同的动作。所得轨迹用于通过我们的可微分引擎优化物体质量，详情见第4.2节。我们在两种设置下评估性能：（1）几何形状各异的物体；（2）几何形状相同但内部密度不同的复制品。我们的方法在这两种情况下都能准确恢复质量，展现出对不同形状的泛化能力，以及对细微物理差异的敏感度。用于质量辨识的物体如图3所示。

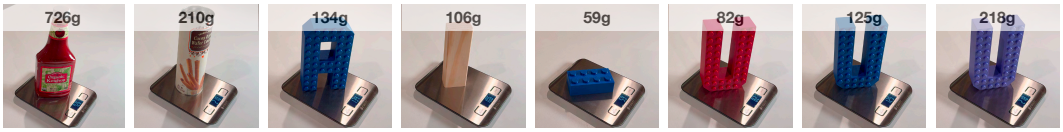


图3：**用于质量辨识的物体**。我们针对多种物体几何形状，以及几何形状相同但密度不同的物体，开展质量辨识实验。我们的方法在这两种设置下都能准确估计质量，展现出对形状和密度变化的鲁棒性。

为评估方法在形状、尺寸和质量尺度方面的鲁棒性，我们选择了多种物体进行质量辨识。这一测试检验了整个流程对不同接触几何形状的泛化能力。如表1所示，百分位误差介于 4.8% 至12.0%之间，表明无需针对特定物体进行调优即可实现准确的质量优化。

为单独考察质量的影响，我们制作了三个几何形状相同但内部密度不同的复制品 $\boldsymbol{\rho}$，方法是采用不同的3D打印填充率。由于形状保持不变，任何辨识误差都反映了对质量的敏感性。如表 2 所示，质量辨识准确，偏差低于 13 克，这证实了无需依赖几何形状即可估计物理参数的有效性。

如图 4所示，采用优化后的质量进行仿真，所得物体动力学与真实情况高度吻合；而采用错误的较轻质量进行仿真，则会产生显著偏差。这表明，准确的质量辨识能够提升仿真展开轨迹的物理真实性和视觉质量。

物体	字母U	字母 A	Lego	多米诺骨牌	饼干	番茄酱
推断质量（g）	500	500	300	500	500	1000
识别质量（g）	110	145	53	117	200	667
真实质量（g）	125	134	59	106	210	726
百分位误差（%）	12.0	9.0	8.6	9.3	4.8	8.1

表1：不同形状、尺寸和质量尺度的多种物体的质量辨识。推断质量由视觉语言模型（VLM）按照第4.2节所述方法获得。

密度	ρ_1	ρ_2	ρ_3
识别质量（克）	95	129	207
真实质量（克）	82	125	218

表 2：对于几何形状相同但密度不同的物体进行质量辨识。

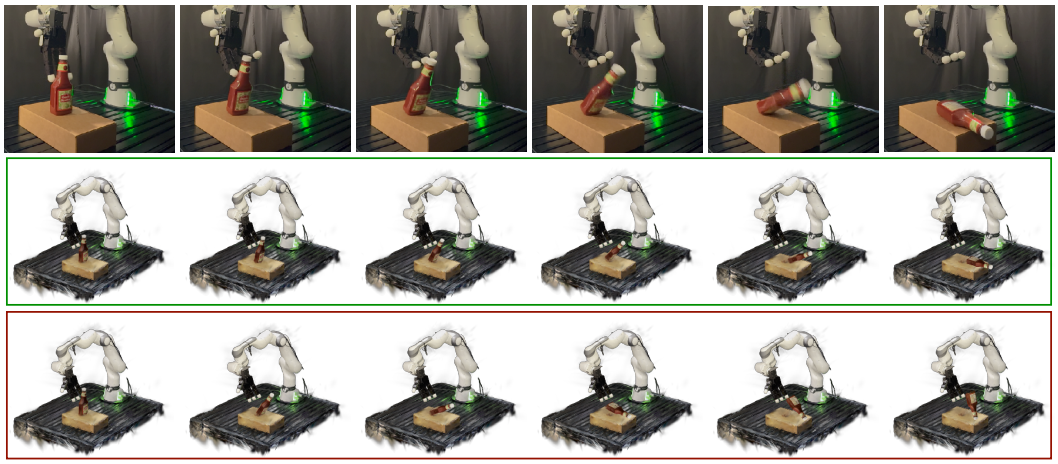


图4：**质量识别的定量结果**。我们展示了真实世界中的物体推动过程（上），以及使用高斯泼溅渲染的物体轨迹：采用优化后的质量进行仿真（中），以及采用更轻质量进行仿真（下），三者使用相同的机器人动作。优化后的质量能够紧密再现真实世界动力学，从而以高视觉保真度缩小仿真与现实差距。

训练 \ 评估	ρ_1	ρ_2	ρ_3
ρ_1	**75%**	30%	15%
ρ_2	40%	**80%**	30%
ρ_3	15%	40%	**95%**

表3：对使用不同物体密度训练、并在不同质量条件下评估的抓取策略进行交叉评估。每个单元格显示抓取成功率。只有训练质量与评估质量相匹配时，策略才能表现良好。

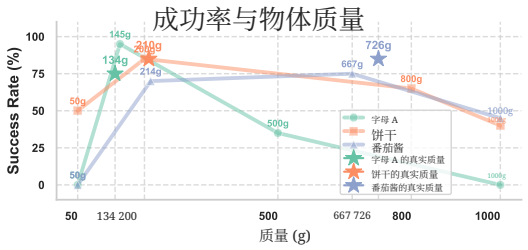


图5：三种具有不同质量值的物体的抓取成功率。

这些实验表明，我们的可微真实到仿真再到真实框架能够准确地辨识质量，适用于物体间差异和物体内部密度变化这两种情况。

5.2 通过抓取实现基于力的控制的有效性

在我们的抓取实验中，我们评估了纳入以物体质量为条件的基于力的约束如何影响仿真到真实性能。该设置突出了对质量感知力控制的需求，并展示了准确的质量辨识对策略成功率的影响。

我们首先在三个几何形状和演示数据完全相同、但质量不同的物体上评估抓取策略。每种策略都针对特定的物体质量进行训练，以评估质量感知力控制的影响。如图 6所示，只有训练质量与评估质量匹配时，策略才能表现良好：中等质量策略能成功抓取中型物体，但会因施力不足或施力过度而无法抓取较重和较轻的物体。质量不匹配同样会导致另外两种策略出现不稳定抓取。表 3 证实了这一趋势：在训练质量下，成功率最高（80%），而在质量不匹配的情况下，成功率分别降至 40% 和 30%。这些结果凸显了准确进行质量条件化对于实现稳健、可靠抓取的重要性。

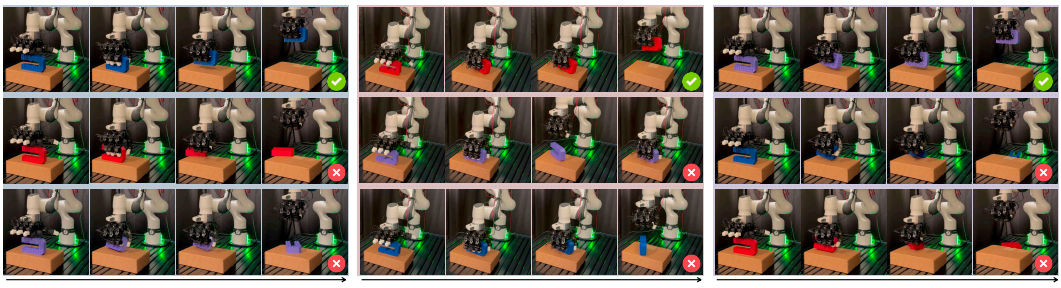


图 6：**定性结果**。从左到右：分别是在中型、轻型和重型物体上训练的策略。只有与物体质量匹配的策略才能实现稳定抓取，而质量不匹配的策略则会因作用力过大或不足而失败，导致较轻的物体弹开，较重的物体发生滑移。

其次，我们评估以自动识别的质量为条件的策略，能否达到以物体质量为条件训练的策略的性能。如图 5所示，成功率始终在真实质量或识别质量处达到峰值。值得注意的是，使用识别质量的策略通常能达到甚至超过使用真实值的策略，同时显著优于以任意质量为条件的策略。这些结果彰显了我们所采用的质量辨识方法的有效性：无需获取真实质量值，即可实现稳健的力感知抓取。

这些实验表明，准确识别质量对于有效进行力感知抓取至关重要。使用物体质量训练的策略始终优于使用不匹配质量训练的策略；而以自动识别的质量为条件的策略，其表现与采用真实值的策略相当。综上，这些结果验证了我们的质量识别框架是一种实用且可靠的解决方案，无需先验的物体质量信息即可实现鲁棒抓取。

5.3 桌面物体抓取实验

我们在多种物体上将我们的抓取策略与两种基线方法进行比较：(1) DexGraspNet 2.0 Zhang 等人 (2024c)，该方法基于大规模仿真数据集训练；(2) Human2Sim2RobotLum 等人 (2025年），这是一种近期提出的真实到仿真再到真实方法，可从人类演示的 RGBD 视频中学习灵巧操作策略。所有方法均使用由我们的真实到仿真框架生成的碰撞网格 $\mathcal{K}$ 作为输入。如图 7所示，在几何形状和质量各异的八种物体上，我们的方法始终优于基线方法，取得了较高的成功率，且方差显著更低。随着物体质量增加，基线方法的性能会下降，而我们的力感知策略在各种物体特征范围内均能保持稳定、可靠的抓取。

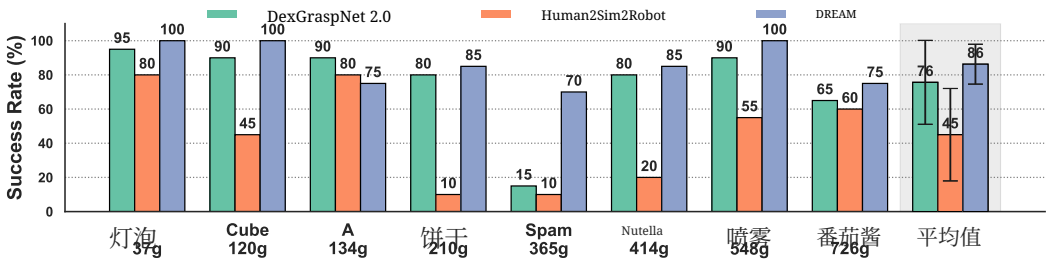


图 7：**抓取策略的定量结果**。八种几何形状和质量值各异的物体的抓取成功率，以及各方法的平均值和标准差。

图8 展示了我们的力感知抓取策略在一系列物体上的定性结果。顶排呈现了到达预抓取姿态的运动过程，底排则展示了最终的抓取后构型。这些示例展现了该策略在物体几何形状和质量值各异的情况下，始终实现稳定、牢固抓取的能力。

物体	# 图像	COLMAP+3DGS+2DGS 运行时间	网格顶点
Lego	306	30 分钟	47,266
番茄酱	337	35 分钟	13,487
多米诺骨牌	320	30 分钟	35,445
饼干	309	30 分钟	25,682
A	306	30 分钟	67,895
U	327	33 分钟	15,322

表 4：代表性物体的离线重建统计，包括输入图像数量、总运行时间以及提取的网格的顶点数量。

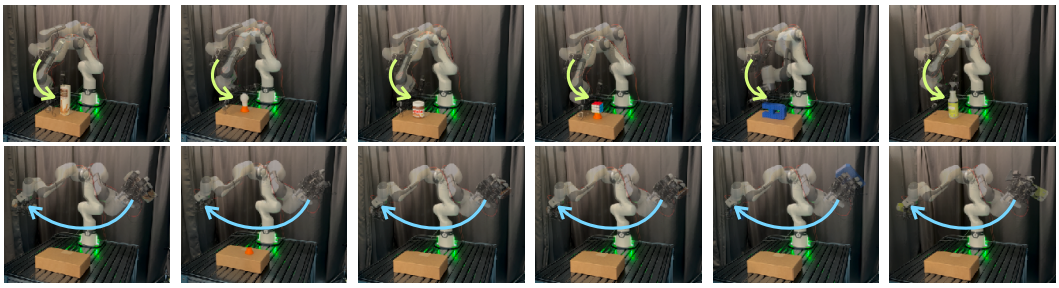


图8：**OurPolicy 的定性结果**。我们评估了力感知抓取策略在各种物体上的表现。第一行展示了接近预抓取姿态的过程，第二行展示了两个抓取后位置，说明该策略能够实现稳定、安全的抓取。

5.4 运行时间分析

接下来，我们分析真实到仿真重建和质量辨识流程的计算成本。对于每个物体，我们采集约 300–340 张 RGB图像。端到端离线重建——通过 COLMAP 进行运动恢复结构 Pan 等人 (2024);Schönberger 和 Frahm(2016年)，随后进行 3D 高斯泼溅和 2DGS——在我们的设备上通常每个物体需要 30–35 分钟。表 4列出了具有代表性的统计数据。尽管网格复杂度因物体而异，从 $\sim 1.3 \times 10^4$到 $\sim 6.8 \times 10^4$ 个顶点，但不同物体的重建时间仍处于较窄范围内。
基于可微物理的质量辨识所需时间随网格顶点数量增加而增长。在我们的硬件上，每次训练迭代约需 1.43–1.68 秒，通常在约 200 个训练轮次内即可收敛。实际操作中，对于表 4 中列出的大多数物体，这相当于 5–20 分钟。所有重建和质量辨识步骤均离线执行，因此不会影响学习得到的抓取策略的实时执行。

6 结论

D-REX 是一个真实到仿真再到真实框架，它利用可微仿真，根据视觉观测和机器人控制信号创建视觉逼真且物理属性准确的数字孪生，从而实现稳健的灵巧抓取策略。通过机器人与物体交互辨识物体质量，该框架能够泛化到形状和密度各异的物体。此外，将以质量为条件的力感知控制融入模仿学习，可增强策略的稳健性和适应性，从而为可扩展的数据生成和可泛化策略的开发带来良好前景，代表着迈向稳健现实世界机器人系统的重要一步。

7 可复现性声明

我们依赖若干开源基础模型。复现实验所需的实现细节见附录与补充材料。对于我们的主要贡献——从视频中学习质量以及力–位置混合策略——我们将代码库和数据集包含在补充材料及我们的网站中。

8 伦理声明

本研究遵守 ICLR 道德准则；所有作者均已阅读并确认该准则，作为投稿流程的一部分。若研究涉及人工标注员或用户数据，我们已获得机构伦理审查委员会的批准或记录在案的豁免以及知情同意，仅使用经许可收集的数据，实施去标识化并限制访问；否则，我们依据相应许可使用公开数据集。我们评估了潜在风险——包括隐私、安全、偏见／歧视、双重用途或滥用以及环境影响——并记录了局限性、失效模式和缓解措施，以尽量减少伤害。我们亦披露在构思、编码或写作中对大语言模型（LLM）的任何使用，并对全部内容承担全部责任；大语言模型并非作者。

9 致谢

南加州大学物理超智能实验室感谢丰田研究院、杜比、Google DeepMind、Capital One、英伟达、博世、美国国家科学基金会和高通的慷慨支持。Yue Wang 还获得了鲍威尔研究奖的资助。

本研究得益于与 Stephen Tu 教授、Quankai Gao 博士和 Yuming Gu 博士的讨论。

参考文献

Ilge Akkaya, Marcin Andrychowicz, Maciek Chociej, Mateusz Litwin, Bob McGrew, Arthur Petron, Alex Paino, Matthias Plappert, Glenn Powell, Raphael Ribas 等。使用机器人手解决魔方。*arXiv* 预印本 *arXiv:1910.07113*，2019。

Danijar Hafner, Jurgis Pasukonis, Jimmy Ba 和 Timothy Lillicrap。通过世界模型掌握多样化领域。*arXiv* 预印本 *arXiv:2301.04104*，2023。

Tao Chen、Jie Xu 和 Pulkit Agrawal。通用手内物体重新定向系统，2021。网址 https://arxiv.org/abs/2111.03043。

Ananye Agarwal、Ashish Kumar、Jitendra Malik 和 Deepak Pathak。利用自我中心视觉在具有挑战性的地形中实现足式运动。在机器人学习会议，第 403–415 页。PMLR，2023。

Tairan He、Chong Zhang、Wenli Xiao、Guanqi He、Changliu Liu 和 Guanya Shi。敏捷而安全：学习无碰撞的高速足式运动。发表于*Robotics: Science and Systems (RSS)*，2024年。

Yecheng Jason Ma、William Liang、Guanzhi Wang、De-An Huang、Osbert Bastani、Dinesh Jayaraman、Yuke Zhu、Linxi Fan 和 Anima Anandkumar。Eureka：通过编写大语言模型代码实现人类水平的奖励设计。*arXiv*预印本 *arXiv: Arxiv-2310.12931*，2023年。

Josh Tobin、Rachel Fong、Alex Ray、Jonas Schneider、Wojciech Zaremba 和 Pieter Abbeel。用于将深度神经网络从仿真迁移至真实世界的域随机化。发表于*2017*年*IEEE/RSJ*智能机器人与系统国际会议（*IROS*），第23–30页。IEEE，2017年。

Fereshteh Sadeghi和Sergey Levine。Cad2rl：无需任何真实图像的真实单图像飞行。*arXiv*预印本 *arXiv:1611.04201*，2016 年。

Xue Bin Peng、Marcin Andrychowicz、Wojciech Zaremba和Pieter Abbeel。基于动力学随机化的机器人控制仿真到现实迁移。载于*2018*年*IEEE*机器人与自动化国际会议（*ICRA*），第3803–3810页。IEEE，2018年a。

Jemin Hwangbo、Joonho Lee、Alexey Dosovitskiy、Dario Bellicoso、Vassilios Tsounis、Vladlen Koltun 和 Marco Hutter。为足式机器人学习敏捷、动态的运动技能。《*Science Robotics*》，4(26):eaau5872，2019。

Jie Tan、Tingnan Zhang、Erwin Coumans、Atil Iscen、Yunfei Bai、Danijar Hafner、Steven Bohez 和 Vincent Vanhoucke。仿真到现实：为四足机器人学习敏捷运动。*arXiv* 预印本 *arXiv:1804.10332*，2018年。

Wisama Khalil、Maxime Gautier 和 Philippe Lemoine。工业机械臂负载惯性参数的辨识。载于*2007年 IEEE*机器人与自动化国际会议论文集，第 4943–4948 页，2007 年。doi：10.1109/ROBOT.2007.364241。

Daniel Ho、Kanishka Rao、Zhuo Xu、Eric Jang、Mohi Khansari 和 Yunfei Bai。Retinagan：一种面向对象感知的仿真到现实迁移方法，2020 年。网址https://arxiv.org/abs/2011.03148。

Mayank Mittal、Calvin Yu、Qinxi Yu、Jingzhou Liu、Nikita Rudin、David Hoeller、Jia Lin Yuan、Ritvik Singh、Yunrong Guo、Hammad Mazhar、Ajay Mandlekar、Buck Babich、Gavriel State、Marco Hutter 和 Animesh Garg。Orbit：用于交互式机器人学习环境的统一仿真框架。*IEEE*机器人与自动化快报，8(6)：3740–3747，2023 年。doi：10.1109/LRA.2023.3270034。

Konstantinos Bousmalis、Alex Irpan、Paul Wohlhart、Yunfei Bai、Matthew Kelcey、Mrinal Kalakrishnan、Laura Downs、Julian Ibarz、Peter Pastor、Kurt Konolige 等。利用仿真和域适应提高深度机器人抓取的效率。发表于*2018年IEEE*机器人与自动化国际会议（*ICRA*），第 4243–4250 页。IEEE，2018年。

Allen Z Ren、Hongkai Dai、Benjamin Burchfiel 和 Anirudha Majumdar。Adaptsim：用于仿真到现实迁移的任务驱动仿真适配。*arXiv* 预印本 *arXiv:2302.04903*，2023。

Tao Chen、Megha Tippur、Siyang Wu、Vikash Kumar、Edward Adelson 和 Pulkit Agrawal。视觉灵巧性：基于深度信息的手内灵巧操作。发表于*ICML* 学习、控制与动力系统新前沿研讨会，2023。

Zoey Chen、Aaron Walsman、Marius Memmel、Kaichun Mo、Alex Fang、Karthikeya Vemuri、Alan Wu、Dieter Fox 和 Abhishek Gupta。Urdformer：一种从真实世界图像构建多关节仿真环境的流程。*arXiv* 预印本 *arXiv:2405.11656*，2024。

Marcel Torne、Anthony Simeonov、Zechu Li、April Chan、Tao Chen、Abhishek Gupta 和 P ulkit Agrawal。通过仿真调和现实：一种用于稳健操作的真实到仿真再到真实方法。*arXiv* 预印本 *arXiv:2403.03949*，2024。

Zhenyu Jiang、Cheng-Chun Hsu 和 Yuke Zhu。Ditto：通过交互构建关节物体的数字孪生。发表于计算机视觉与模式识别会议（*CVPR*），2022。

Miles Macklin、Matthias Müller、Nuttapong Chentanez 和 Tae-Yong Kim。面向实时应用的统一粒子物理。*ACM* 图形学汇刊，33(4)，2014年7月。ISSN 0730-0301。doi：10.1145/ 2601097.2601152。网址 https://doi.org/10.1145/2601097.2601152。

Kyle Cranmer、Johann Brehmer 和 Gilles Louppe。基于仿真的推断前沿。美国国家科学院院刊，117(48):30055–30062，2020。

Fabio Ramos、Rafael Carvalhaes Possas 和 Dieter Fox。Bayessim：通过概率推断为机器人仿真器进行自适应域随机化。*arXiv* 预印本 *arXiv:1906.01728*，2019。

Jiajun Wu、Joshua B Tenenbaum 和 Pushmeet Kohli。神经场景反渲染。载于*IEEE* 计算机视觉与模式识别会议 *(CVPR)*，2017。

Krishna Murthy Jatavallabhula、Miles Macklin、Florian Golemo、Vikram Voleti、Linda Petrini、Martin Weiss、Breandan Considine、Jerome Parent-Levesque、Kevin Xie、Kenny Erleben、Liam Paull、Florian Shkurti、Derek Nowrouzezahrai 和 Sanja Fidler。GradSim：用于系统辨识和视觉运动控制的可微仿真，2021a。网址 https://arxiv.org/abs/2104.02646。

C. Daniel Freeman、Erik Frey、Anton Raichuk、Sertan Girgin、Igor Mordatch 和 Olivier Bachem。Brax——用于大规模刚体仿真的可微物理引擎。2021年。网址 http://github.com/google/brax。

Matthias Müller、Bruno Heidelberger、Marcus Hennix 和 John Ratcliff。《基于位置的动力学》。《视觉传播与图像表示》，18(2):109–118，2007年。ISSN 1047-3203。doi: https://doi.org/10.1016/j.jvcir.2007.01.005。网址 https://www.sciencedirect.com/science/article/pii/S1047320307000065。Miles Macklin、Matthias Müller 和 Nuttapong Chentanez。XPBD：柔顺约束动力学的基于位置的仿真。载于第9届国际游戏运动会议论文集，MIG '16，第49–54页，美国纽约州纽约市，2016年。美国计算机协会。ISBN 9781450345927。doi: 10.1145/2994258.2994272。网址 https://doi.org/10.1145/2994258.2994272。Bernhard Kerbl、Georgios Kopanas、Thomas Leimkühler 和 George Drettakis。用于实时辐射场渲染的三维高斯泼溅。《ACM图形学汇刊》，42(4):139–1，2023年。C. Karen Liu 和 Dan Negrut。基于物理的模拟器在机器人学中的作用。《控制、机器人与自主系统年度评论》，4（第4卷，2021年）：35–58，2021年。ISSN 2573-5144。doi: https://doi.org/10.1146/annurev-control-072220-093055。网址 h
t
t
ps://www.annualreviews.org/content/journals/10.1146/annurev-control-072220-093055。Jie Xu、Tao Chen、Lara Zlokapa、Michael Foshey、Wojciech Matusik、Shinjiro Sueda 和 Pulkit Agrawal。用于接触感知机器人设计的端到端可微框架。arXiv预印本 arXiv:2107.07501，2021年。Jie Xu、Viktor Makoviychuk、Yashraj Narang、Fabio Ramos、Wojciech Matusik、Animesh Garg 和 Miles Macklin。通过并行可微仿真加速策略学习。arXiv预印本 arXiv:2204.07137，2022年。Filipe de Avila Belbute-Peres、Kevin Smith、Kelsey Allen、Josh Tenenbaum 和 J. Zico Kolter。用于学习与控制的端到端可微物理。载于 S. Bengio、H. Wallach、H. Larochelle、K. Grauman、N. Cesa-Bianchi 和 R. Garnett（编），《神经信息处理系统进展》，第31卷。Curran Associates, Inc.，2018年。网址 https://proceedings.neurips.cc/paper_files/paper/2018/file/842424a1d0595b76ec4fa03c46e8d755-Paper.pdf。Yuanming Hu、Luke Anderson、Tzu-Mao Li、Qi Sun、Nathan Carr、Jonathan Ragan-Kelley 和 Frédo Durand。DiffTaichi：用于物理仿真的可微编程。arXiv预印本 arXiv:1910.00935，2019年。Wenxuan Li、Hang Zhao、Zhiyuan Yu、Yu Du、Qin Zou、Ruizhen Hu 和 Kai Xu。PIN-WM：学习用于非抓取式操作的物理信息世界模型，2025a。网址 https://arxiv.org/abs/2504.16693。Yunzhu Li、Jiajun Wu、Russ Tedrake、Joshua B Tenenbaum 和 Antonio Torralba。学习粒子动力学，以操作刚体、可变形物体和流体。载于国际学习表征会议，2019年。Tobias Pfaff、Meire Fortunato、Alvaro Sanchez-Gonzalez 和 Peter Battaglia。使用图网络学习基于网格的仿真。载于国际学习表征会议，2020年。Zhou Xian、Shamit Lal、Hsiao-Yu Tung、Emmanouil Antonios Platanios 和 Katerina Fragkiadaki。超动力学：利用超网络元学习物体与智能体动力学。arXiv预印本 arXiv:2103.09439，2021年。Zhou Xian、Bo Zhu、Zhenjia Xu、Hsiao-Yu Tung、Antonio Torralba、Katerina Fragkiadaki 和 Chuang Gan。FluidLab：用于复杂流体操作基准测试的可微环境。载于国际学习表征会议，2023年。Zhiao Huang、Yuanming Hu、Tao Du、Siyuan Zhou、Hao Su、Joshua B Tenenbaum 和 Chuang Gan。PlasticineLab：具有可微物理的软体操作基准。arXiv预印本 arXiv:2104.03311，2021年。Weikun Peng、Jun Lv、Yuwei Zeng、Haonan Chen、Siheng Zhao、Jichen Sun、Cewu Lu 和 Lin Shao。TieBot：通过真实到仿真再到真实方法，从视觉演示中学习打领带。arXiv预印本 arXiv:2407.03245，2024年。

Xinyuan Yu、Siheng Zhao、Siyuan Luo、Gang Yang 和 Lin Shao。Diffclothai：具有无交叉摩擦接触以及与关节刚体可微双向耦合的可微布料仿真。载于*2023 IEEE/RSJ*智能机器人与系统国际会议（*IROS*），第400–407页，2023。doi: 10.1109/IROS55552.2023.10341573。

Jagdeep Bhatia、Holly Jackson、Yunsheng Tian、Jie Xu 和 Wojciech Matusik。Evolution gym：用于演化软体机器人的大规模基准。《神经信息处理系统进展》，34:2201–2214，2021。

Sizhe Li、Zhiao Huang、Tao Chen、Tao Du、Hao Su、Joshua B. Tenenbaum 和 Chuang Gan。Dexdeform：结合人类示范与可微物理的灵巧可变形物体操作，2023a。网址 https://arxiv.org/abs/2304.03223。

Sizhe Lester Li、Annan Zhang、Boyuan Chen、Hanna Matusik、Chao Liu、Daniela Rus 和 Vincent Sitzmann。利用深度网络推断雅可比场以控制多种机器人。《自然》，643(8070):89–95，2025年6月。ISSN 1476-4687。doi: 10.1038/s41586-025-09170-0。网址 http://dx.doi.org/10.1038/s41586-025-09170-0。

Zhao Mandi、Yijia Weng、Dominik Bauer 和 Shuran Song。Real2code：通过代码生成重建关节物体。*arXiv*预印本 *arXiv:2406.08474*，2024。

Priya Sundaresan、Rika Antonova 和 Jeannette Bohgl。Diffcloud：利用点云、可微仿真以及可变形物体渲染实现真实到仿真。在*2022 IEEE/RSJ*智能机器人与系统国际会议（*IROS*）上，第10828–10835页。IEEE，2022。

Justin Kerr、Letian Fu、Huang Huang、Yahav Avigal、Matthew Tancik、JeffreyIchnowski、Angjoo Kanazawa 和 Ken Goldberg。Evo-neRF：演进 NeRF，用于对透明物体进行序列机器人抓取。载于 第*6*届机器人学习年度会议，2022。网址 https://openreview.net/forum?id=Bxr45keYrf。

Adam Rashid、Satvik Sharma、Chung Min Kim、Justin Kerr、Lawrence Yunliang Chen、Angjoo Kanazawa 和 Ken Goldberg。用于零样本任务导向抓取的语言嵌入辐射场。载于 第7届机器人学习年度会议，2023。网址https://openreview.net/forum?id=k-Fg8JDQmc。

Allan Zhou、Moo Jin Kim、Lirui Wang、Pete Florence 和 Chelsea Finn。掌中 NeRF：通过新视角合成进行机器人学的校正增强。载于 *IEEE/CVF*计算机视觉与模式识别会议论文集，第17907–17917页，2023年。

Arunkumar Byravan、Jan Humplik、Leonard Hasenclever、Arthur Brussee、Francesco Nori、Tuomas Haarnoja、Ben Moran、Steven Bohez、Fereshteh Sadeghi、Bojan Vujatovic 等。Nerf2real：使用神经辐射场将视觉引导的双足运动技能从仿真迁移到现实。载于*2023 IEEE*机器人与自动化国际会议（*ICRA*），第 9362–9369 页。IEEE，2023。

Luobin Wang、Runlin Guo、Quan Vuong、Yuzhe Qin、Hao Su 和 Henrik Christensen。利用神经表面重建实现稳健物体抓取的现实到仿真到现实方法，2023。网址https://arxiv.org/abs/2210.02685。

Jad Abou-Chakra、Krishan Rana、Feras Dayoub 和 Niko Suenderhauf。具身高斯泼溅：一种可实时校正的机器人学世界模型。载于第 *8* 届机器人学习年度会议，2024。网址 https://openreview.net/forum?id=AEq0onGrN2。

Mingtong Zhang、Kaifeng Zhang 和 Yunzhu Li。用于基于图的神经动力学建模的动态三维高斯跟踪。*arXiv* 预印本 *arXiv:2410.18912*，2024a。

Justin Kerr、Chung Min Kim、Mingxuan Wu、Brent Yi、Qianqian Wang、Ken Goldberg 和 Angjoo Kanazawa。机器人看，机器人做：利用单目四维重建模仿关节物体操作。收录于第 *8* 届机器人学习年度会议，2024。网址 https://openreview.net/forum?id=2LLu3gavF1。

Hanxiao Jiang、Hao-Yu Hsu、Kaifeng Zhang、Hsin-Ni Yu、Shenlong Wang 与 Yunzhu Li。Phystwin：基于视频的可变形物体物理信息重建与仿真。*arXiv* 预印本 *arXiv:2503.17973*，2025年。

Mikel Zhobro、Andreas René Geist 与 Georg Martius。从 RGB 视频学习三维高斯仿真器。*arXiv* 预印本 *arXiv:2503.24009*，2025年。Jad Abou-Chakra、Lingfeng Sun、Krishan Rana、Brandon May、Karl Schmeckpeper、Maria Vittoria Minniti 与 Laura Herlant。Real-is-Sim：借助动态数字孪生弥合仿真到现实差距，以进行真实世界机器人策略评估。*arXiv* 预印本 *arXiv:2504.03597*，2025年。Sizhe Yang、Wenye Yu、Jia Zeng、Jun Lv、Kerui Ren、Cewu Lu、Dahua Lin 与 Jiangmiao Pang。基于高斯泼溅的新型演示生成实现鲁棒的单次操作，2025年。网址 https://arxiv.org/abs/2504.13175。Tianyi Xie、Zeshun Zong、Yuxing Qiu、Xuan Li、Yutao Feng、Yin Yang 与 Chenfanfu Jiang。PhysGaussian：面向生成动力学的物理集成三维高斯。*arXiv* 预印本 *arXiv:2311.12198*，2023。Nicholas Pfaff、Evelyn Fu、Jeremy Binagia、Phillip Isola 与 Russ Tedrake。可扩展 Real2Sim：通过机器人抓取与放置装置实现物理感知资产生成，2025年。网址 https://arxiv.org/abs/2503.00370。Yifan Zhu、Tianyi Xiang、Aaron Dollar 与 Zherong Pan。通过端到端可微仿真与渲染实现的单次真实到仿真，2025年。网址 https://arxiv.org/abs/2412.00259。Dylan Turpin、Liquan Wang、Eric Heiden、Yun-Chun Chen、Miles Macklin、Stavros Tsogkas、Sven Dickinson 与 Animesh Garg。Grasp'd：面向多指手的可微富接触抓取合成，2022。网址 https://arxiv.org/abs/2208.12250。Suraj Nair、Aravind Rajeswaran、Vikash Kumar、Chelsea Finn 与 Abhinav Gupta。R3M：面向机器人操作的通用视觉表征。*arXiv* 预印本 *arXiv:2203.12601*，2022。Ilija Radosavovic、Tete Xiao、Stephen James、Pieter Abbeel、Jitendra Malik 与 Trevor Darrell。基于掩码视觉预训练的真实世界机器人学习。载于机器人学习会议，第416–426页。PMLR，2023。Yecheng Jason Ma、Shagun Sodhani、Dinesh Jayaraman、Osbert Bastani、Vikash Kumar 与 Amy Zhang。VIP：通过值隐式预训练迈向通用视觉奖励与表示。*arXiv* 预印本 *arXiv:2210.00030*，2022。Shikhar Bahl、Russell Mendonca、Lili Chen、Unnat J ain 与 Deepak Pathak。来自人类视频的可供性作为机器人学的通用表示。*CVPR*，2023。Mengda Xu、Zhenjia Xu、Yinghao Xu、Cheng Chi、Gordon Wetzstein、Manuela Veloso 与 Shuran Song。以流作为跨域操作接口。*arXiv* 预印本 *arXiv:2407.15208*，2024。Kenneth Shaw、Shikhar Bahl 与 Deepak Pathak。VideoDex：从互联网视频学习灵巧性。载于机器人学习会议，第654–665页。PMLR，2023a。Austin Patel、Andrew Wang、Ilija Radosavovic 与 Jitendra Malik。从互联网视频学习模仿物体交互。*arXiv:2211.13225*，2022。Xue Bin Peng、Angjoo Kanazawa、Jitendra Malik、Pieter Abbeel 与 Sergey Levine。SFV：从视频中强化学习物理技能。*ACM*图形学汇刊（*TOG*），37(6):1–14，2018年b。Tyler Ga Wei Lum、Olivia Y Lee、C Karen Liu 与 Jeannette Bohg。利用一次人类演示，通过仿真到现实强化学习跨越人机具身差距。*arXiv* 预印本 *arXiv:2504.12609*，2025年。Irmak Guzey、Yinlong Dai、Georgy Savva、Raunaq Bhirangi 与 Lerrel Pinto。通过面向对象奖励弥合从人到机器人的灵巧性差距。*arXiv* 预印本 *arXiv:2410.23289*，2024。Himanshu Gaurav Singh、Antonio Loquercio、Carmelo Sferrazza、Jane Wu、Haozhi Qi、Pieter Abbeel 与 Jitendra Malik。基于视频的手物交互预训练。*arXiv* 预印本 *arXiv:2409.08273*，2024。

Emanuel Todorov、Tom Erez 和 Yuval Tassa。MuJoCo：用于基于模型的控制的物理引擎。载于*2012*年*IEEE/RSJ*智能机器人与系统国际会议，第5026–5033页。IEEE，2012。doi：10.1109/IROS.2012.6386109。

Haozhe Lou、Yurong Liu、Yike Pan、Yiran Geng、Jianteng Chen、Wenlong Ma、Chenglong Li、Lin Wang、Hengzhen Feng、Lu Shi、Liyi Luo 和 Yongliang Shi。Robo-GS：一种采用混合表示的、适用于机械臂的物理一致时空模型，2024。网址 https://arxiv.org/abs/2408.14873。

Chongjie Ye、Lingteng Qiu、Xiaodong Gu、Qi Zuo、Yushuang Wu、Zilong Dong、Liefeng Bo、Yuliang Xiu 和 Xiaoguang Han。StableNormal：降低扩散方差，实现稳定且清晰的法线。*ACM*图形学汇刊 *(TOG)*，2024a。

Aaron Hurst、AdamLerer、AdamP Goucher、Adam Perelman、Aditya Ramesh、Aidan Clark、AJ Ostrow、Akila Welihinda、Alan Hayes、Alec Radford 等。GPT-4o 系统卡。*arXiv* 预印本 *arXiv:2410.21276*，2024。

Tianyuan Zhang、Hong-Xing Yu、Rundi Wu、Brandon Y. Feng、Changxi Zheng、Noah Snavely、Jiajun Wu 和 William T. Freeman。PhysDreamer：通过视频生成与三维物体进行基于物理的交互。载于欧洲计算机视觉会议。施普林格，2024b。

Martin Asenov、Michael Burke、Daniel Angelov、Todor Davchev、Kartic Subr 和 Subramanian Ramamoorthy。Vid2param：从视频中建模动力学参数，2020。网址 https://arxiv.org/abs/1907.06422。

Bowen Wen、Wei Yang、Jan Kautz和Stan Birchfield。FoundationPose：新物体的统一六维位姿估计与跟踪。收录于*CVPR*，2024。

Ricky T. Q. Chen、Yulia Rubanova、Jesse Bettencourt和David Duvenaud。神经常微分方程，2019。网址https://arxiv.org/abs/1806.07366。

David Baraff。刚体仿真。*SIGGRAPH 1992*课程讲义，19，1992。

Tom Erez、Yuval Tassa和Emanuel Todorov。基于模型的机器人技术的仿真工具：Bullet、Havok、MuJoCo、常微分方程与PhysX的比较。收录于*2015*年*IEEE*机器人与自动化国际会议（*ICRA*），第4397–4404页，2015。doi: 10.1109/ICRA.2015.7139807。

Yuanming Hu、Luke Anderson、Tzu-Mao Li、Qi Sun、Nathan Carr、Jonathan Ragan-Kelley和Frédo Durand。DiffTaichi：面向物理仿真的可微编程，2020。网址https://arxiv.org/abs/1910.00935。

Adam Paszke、Sam Gross、Francisco Massa、Adam Lerer、James Bradbury、Gregory Chanan、Trevor Killeen、Zeming Lin、Natalia Gimelshein、Luca Antiga、Alban Desmaison、Andreas Köpf、Edward Yang、Zach DeVito、Martin Raison、Alykhan Tejani、Sasank Chilamkurthy、Benoit Steiner、Lu Fang、Junjie Bai 和 Soumith Chintala。《PyTorch：一种命令式风格的高性能深度学习库》，2019。网址 https://arxiv.org/abs/1912.01703。

Georgios Pavlakos、Dandan Shan、Ilija Radosavovic、Angjoo Kanazawa、David Fouhey 和 Jitendra Malik。使用 Transformer 在三维空间中重建手部。载于 *CVPR*，2024。

Jane Wu、Georgios Pavlakos、Georgia Gkioxari 和 Jitendra Malik。重建手持物体的三维模型。*arXiv* 预印本 *arXiv:2404.06507*，2024a。

Yuzhe Qin、Wei Yang、Binghao Huang、Karl Van Wyk、Hao Su、Xiaolong Wang、Yu-Wei Chao 和 Dieter Fox。Anyteleop：一种通用的基于视觉的灵巧机器人臂手遥操作系统。载于机器人学：科学与系统，2023 年。

Matthew Tancik、Pratul P. Srinivasan、Ben Mildenhall、Sara Fridovich-Keil、Nithin Raghavan、Utkarsh Singhal、Ravi Ramamoorthi、Jonathan T. Barron 和 Ren Ng。傅里叶特征使网络能够在低维域中学习高频函数。神经信息处理系统大会，2020 年。

Pierre Vassiliadis、Gerard Derosiere、Cecile Dubuc、Aegryan Lete、Frederic Crevecoeur、Friedhelm C. Hummel 和 Julie Duque。奖励促进基于强化的运动学习。*iScience*，24(7): 102821，2021年7月。doi: 10.1016/j.isci.2021.102821。

Hui Zhang、Zijian Wu、Linyi Huang、Sammy Christen 和 Jie Song。Robustdexgrasp：对一般物体进行稳健的灵巧抓取，2025年。网址 https://arxiv.org/abs/2504.05287。

Hongyi Chen、Yunchao Yao、Yufei Ye、Zhixuan Xu、Homanga Bharadhwaj、Jiashun Wang、Shubham Tulsiani、Zackory Erickson和Jeffrey Ichnowski。Web2grasp：从手物交互的网络图像中学习功能性抓取，2025年a。网址https://arxiv.org/abs/2505.05517。

Weikang Wan、Haoran Geng、Yun Liu、Zikang Shan、Yaodong Yang、Li Yi和He Wang。Unidexgrasp++：通过几何感知课程与迭代式通才-专才学习改进灵巧抓取策略学习，2023。网址 https://arxiv.org/abs/2304.00464。

Zhenyu Wei、Zhixuan Xu、Jingxiang Guo、Yiwen Hou、Chongkai Gao、Zhehao Cai、Jiayu Luo和Lin Shao。 $\mathcal{D}(\mathcal{R},\mathcal{O})$ grasp：一种面向跨具身形态灵巧抓取的机器人与物体交互统一表示，2025年。网址 https://arxiv.org/abs/2410.01702。

Kenneth Shaw、Shikhar Bahl 和 Deepak Pathak。《VideoDex：从互联网视频中学习灵巧性》，2022年。网址https://arxiv.org/abs/2212.04498。

Jialiang Zhang、Haoran Liu、Danshi Li、Xinqiang Yu、Haoran Geng、Yufei Ding、Jiayi Chen 和 He Wang。《DexGraspNet 2.0：在大规模合成杂乱场景中学习生成式灵巧抓取》，2024c。网址 https://arxiv.org/abs/2410.23004。

Linfei Pan、Dániel Baráth、Marc Pollefeys 和 Johannes L. Schönberger。《重新审视全局运动恢复结构》，2024年。网址https://arxiv.org/abs/2407.20219。Johannes L. Schönberger 和 Jan-Michael Frahm。《重新审视运动恢复结构》。载于*2016*年 *IEEE*计算机视觉与模式识别会议（*CVPR*），第 4104–4113 页，2016年。doi：10.1109/CVPR.2016.445。

Krishna Murthy Jatavallabhula、Miles Macklin、Florian Golemo、Vikram Voleti、Linda Petrini、Martin Weiss、Breandan Considine、Jerome Parent-Levesque、Kevin Xie、Kenny Erleben、Liam Paull、Florian Shkurti、Derek Nowrouzezahrai 和 Sanja Fidler。GradSim：用于系统辨识和视觉运动控制的可微仿真。国际学习表征会议（*ICLR*），2021b。网址https://openreview.net/forum?id=c_E8kFWfhp0。

Clement Fuji-Tsang、Masha Shugrina、Jean-Francois Lafleche、Charles Loop、Towaki Takikawa、Jiehan Wang、Wenzheng Chen、Sanja Fidler、Jason Gorski、Rev Lebaredian、Jianing Li、Michael Li、Krishna Murthy Jatavallabhula、Artem Rozantsev、Frank Shen、Edward Smith、Gavriel State 和 Tommy Xiang。Kaolin：一个用于加速三维深度学习研究的 PyTorch 库。
https://github.com/NVIDIAGameWorks/kaolin，2019年。Xuan Li、Yi-Ling Qiao、Peter Yichen Chen、Krishna Murthy Jatavallabhula、Ming Lin、Chenfanfu Jiang 和 Chuang Gan。PAC-NeRF：用于几何无关系统辨识的物理增强连续体神经辐射场，2023b。网址https://arxiv.org/abs/2303.05512。Peter I Corke。一种简单且系统化的 Denavit–Hartenberg 参数赋值方法。*IEEE*机器人学汇刊，23(3):590–594，2007年。Matthias Müller、Bruno Heidelberger、Marcus Hennix 和 John Ratcliff。基于位置的动力学。《视觉传播与图像表示》，18(2):109–118，2007年。Chenfanfu Jiang、Craig Schroeder、Joseph Teran、Alexey Stomakhin 和 Andrew Selle。用于模拟连续体材料的物质点法。载于*ACM SIGGRAPH 2016*课程，SIGGRAPH '16，美国纽约州纽约市，2016年。美国计算机协会。ISBN 9781450342896。doi: 10.1145/2897826.2927348。网址 https://doi.org/10.1145/2897826.2927348。Jan Bender、Matthias Müller 和 Miles Macklin。基于位置的动力学综述，2017年。载于欧洲计算机图形学协会论文集：教程，EG '17，德国戈斯拉尔，2017年。Eurographics协会。doi: 10.2312/egt.20171034。网址 https://doi.org/10.2312/egt.20171034。

Trevor Standley、Ozan Sener、Dawn Chen 和 Silvio Savarese。image2mass：根据图像估算物体质量。载于 Sergey Levine、Vincent Vanhoucke 和 Ken Goldberg 编，第一届机器人学习年度会议论文集，第78卷，机器学习研究论文集，第324–333页。PMLR，2017年11月13–15日。网址https://proceedings.mlr.press/v78/standley17a.html。

Chongjie Ye、Yinyu Nie、Jiahao Chang、Yuantao Chen、Yihao Zhi 和 Xiaoguang Han。Gaustudio：用于 3D 高斯泼溅及其他任务的模块化框架。*arXiv* 预印本 *arXiv:2403.19632*，2024b。

Binbin Huang、Zehao Yu、Anpei Chen、Andreas Geiger 和 Shenghua Gao。用于几何精确辐射场的 2D 高斯泼溅。载于 *ACM SIGGRAPH 2024* 会议论文集，第 1–11 页，2024。

Andy Zeng、Shuran Song、Matthias Nießner、Matthew Fisher、Jianxiong Xiao 和 Thomas Funkhouser。3DMatch：从 RGB-D 重建中学习局部几何描述符。载于*CVPR*，2017。

Peter Yichen Chen、Chao Liu、Pingchuan Ma、John Eastman、Daniela Rus、Dylan Randle、Yuri Ivanov 和 Wojciech Matusik。通过可微机器人-物体交互利用机器人本体感知学习物体属性，2025b。网址 https://arxiv.org/abs/2410.03920。

Yiren Lu、Justin Fu、George Tucker、Xinlei Pan、Eli Bronstein、Rebecca Roelofs、Benjamin Sapp、Brandyn White、Aleksandra Faust、Shimon Whiteson、Dragomir Anguelov 和 Sergey Levine。《模仿还不够：利用强化学习提升模仿在复杂驾驶场景中的鲁棒性》，2023。网址 https://arxiv.org/abs/2212.11419。

Eli Bronstein、Mark Palatucci、Dominik Notz、Brandyn White、Alex Kuefler、Yiren Lu、Supratik Paul、Payam Nikdel、Paul Mougin、Hongge Chen、Justin Fu、Austin Abrams、Punit Shah、Evan Racah、Benjamin Frenkel、Shimon Whiteson 和 Dragomir Anguelov。《用于自动驾驶规划的分层模型式模仿学习》，2022。网址 https://arxiv.org/abs/2210.09539。

Balakumar Sundaralingam、Siva Kumar Sastry Hari、Adam Fishman、Caelan Garrett、Karl Van Wyk、ValtsBlukis、Alexander Millane、Helen Oleynikova、Ankur Handa、Fabio Ramos 等人。curobo：并行化的无碰撞最小加加速度机器人运动生成。*arXiv* 预印本 *arXiv:2310.17274*，2023。

Kenneth Shaw、Ananye Agarwal 和 Deepak Pathak。LEAP Hand：用于机器人学习的低成本、高效且仿人型手。*Robotics: Science and Systems (RSS)*，2023b。

Guanjun Wu、Taoran Yi、Jiemin Fang、Lingxi Xie、Xiaopeng Zhang、Wei Wei、Wenyu Liu、Qi Tian 和 Xinggang Wang。用于实时动态场景渲染的4D高斯泼溅。载于*IEEE/CVF*计算机视觉与模式识别会议（*CVPR*）论文集，第20310–20320页，2024b年6月。

Haoqi Yuan、Bohan Zhou、Yuhui Fu 和 Zongqing Lu。利用强化学习进行跨具身灵巧抓取。*arXiv*预印本 *arXiv:2410.02479*，2024年。

Zhenjia Xu、Jiajun Wu、Andy Zeng、Joshua B. Tenenbaum 和 Shuran Song。Densephysnet：通过多步动态交互学习稠密物理对象表征，2019年。网址https://arxiv.org/abs/1906.03853。

Vickie Ye、Ruilong Li、Justin Kerr、Matias Turkulainen、Brent Yi、Zhuoyang Pan、Otto Seiskari、Jianbo Ye、Jeffrey Hu、Matthew Tancik 和 Angjoo Kanazawa。gsplat：一个用于高斯泼溅的开源库。 *arXiv* 预印本 *arXiv:2409.06765*，2024c。网址 https://arxiv.org/abs/2409.06765。

OpenAI：Marcin Andrychowicz、Bowen Baker、Maciek Chociej、RafalJozefowicz、BobMcGrew、Jakub Pachocki、Arthur Petron、Matthias Plappert、Glenn Powell、Alex Ray 等人。灵巧手内操作学习。《国际机器人研究杂志》，39(1)：3–20，2020。

A 附录

A.1 预备知识

本文旨在仅凭视觉观测和机器人控制信号，准确重建人手抓取的物理过程，而无需获取真实物理参数。我们的方法基于两个关键组件。第一个是可微分的粒子物理仿真引擎 Freeman 等人（2021），它能够对物体质量等物理属性进行基于梯度的优化。第二个是基于高斯泼溅的真实到仿真重建流程Lou 等人（2024），可让我们根据视频输入构建逼真且空间一致的三维场景。结合这两个组件，我们构建了一条完全可微分的流程，连接真实世界感知与物理仿真，支持对动态手物交互进行准确建模，并实现仿真中的稳健策略学习。

MuJoCo 等机器人仿真引擎 Todorov 等人 (2012）、Isaac Sim Mittal 等人 (2023）和 GradSim Jatavallabhula 等人 (2021b）；Fuji-Tsang 等人 (2019）从根本上都建立在力学的拉格朗日表述之上 Li 等人 (2023b），该表述通过追踪一组固定的粒子或参考点在空间和时间中的变化，对物理系统的演化进行建模。这种方法假设仿真环境具有一致且预先定义的结构，通常使用 MJCF 或 URDF 等格式进行描述。这些配置规定了系统组件的数量和排列方式，例如关节、连杆和驱动元件，并且这些内容在整个仿真过程中保持不变。在每个离散时间步，系统会根据反映仿真引擎中所嵌入物理原理的动力学和运动学方程，更新每个物体的状态。因此，物体位姿、速度和接触交互的演化由引擎内部的数值求解器和积分方案控制。这种结构化且融入物理属性的表示，对于准确建模机器人抓取场景中的力传递、接触行为和运动至关重要。

可微物理。 我们的引擎有一项基本假设：静态场景重建完成后，在系统辨识和策略训练阶段，环境的物理配置始终保持不变。也就是说，仿真环境或真实场景中都不会新增物体或机器人，也不会移除已有组件。因此，观测阶段捕获的所有实体状态均保持一致，并直接用于在仿真和真实世界执行中部署控制信号。这确保了仿真展开的保真度以及不同域之间动力学的一致性。

我们的系统架构由两类基本方程支配。第一类涉及运动学方程 Corke (2007)，用于对机械臂和机器人手的关节式运动建模，并考虑关节角、速度和末端执行器轨迹。这些方程支撑着机器人以受控方式伸向并操纵物体的能力。第二类由动力学方程 Jatavallabhula 等人 (2021b)构成，用于描述物体、机器人手与支撑面（例如桌面）之间的交互。这些动力学方程描述接触过程中产生的力和力矩，从而能够准确仿真物体的响应。

为了仿真和优化物体行为，我们采用由 MJX（Brax 基于 JAX 的后端）和 GradSim 组成的双引擎架构。在可微物理引擎中进行空间表示时，我们将物体的网格顶点作为基本粒子。这些顶点充当几何和物理描述符，可对物体与手部以及物体与环境的交互进行精细建模。

MJX 用于对机器人运动学建模，并在仿真展开过程中提取详细的接触信息。它能提供与机器人手交互所产生的精确接触点、表面法向量和力向量。这些信息对于为系统辨识和后续策略学习设定准确的边界条件至关重要。

与此同时，GradSim Jatavallabhula 等人 (2021b) 提供了一个基于 PyTorch 的引擎，用于对物体动力学进行基于梯度的仿真。它对重力、惯性力和外部扰动（例如机器人施加的推力或与地面的碰撞）的影响进行建模，从而实现梯度随时间平滑流动。这一方案有助于高效地优化质量参数，并支持同时涉及感知和控制的端到端训练流程。

我们方案中的一个关键假设是，仿真中物体、地面和机器人手之间的相对位姿与真实世界中的情况十分接近。这种对齐对于确保仿真中的接触事件反映真实世界的条件至关重要，从而能够对物理交互进行高保真建模。为此，我们使用从 FoundationPose 等视觉跟踪管线获得的估计姿态来对齐物体的位置，确保坐标系保持一致。
尽管我们的引擎采用了基于位置的动力学（PBD）Müller 等人（2007）以提升稳定性和效率，但我们针对性地进行了改进，以增强碰撞检测和接触求解。具体而言，我们改进了宽相位碰撞检测算法，使其能够更好地处理高分辨率网格和非凸几何体。这对于准确建模具有精细表面细节的复杂物体，以及确保密集接触交互过程中的梯度传播稳健性至关重要。

结合 MJX 在运动学建模和接触提取方面的优势，以及 GradSim 在基于梯度的物理仿真方面的能力，我们的引擎实现了端到端质量识别和力感知策略训练。这些能力为在现实环境中实现准确且具有泛化能力的机器人抓取奠定了基础，并通过基于物理的学习弥合仿真到现实差距。

A.1.1 基于粒子的物理仿真

基于粒子的物理仿真广泛应用于计算物理和图形学，用于模拟物体的动态行为Jiang 等人（2016年）。不同于依赖连续体积或多边形网格的传统方法，基于粒子的方法将物体离散为大量独立粒子，每个粒子都具有质量等物理属性 m_i、位置$\mathbf{x}_i$和速度$\mathbf{v}_i$，以及包括弹性、摩擦和阻尼在内的材料属性。这种离散表示能够高效、逼真地模拟复杂行为，尤其适用于涉及可变形或碎裂物体、流体和颗粒材料的场景。

粒子系统的质心 **COM** 可通过以下公式计算：

$$\mathbf{COM} = \frac{\sum_i m_i \mathbf{x}_i}{\sum_i m_i}. \tag{12}$$

惯性张量 **I**描述物体对转动加速度的抵抗，其相对于质心的计算公式为：

$$\mathbf{I} = \sum_i m_i \left[\|\mathbf{r}_i\|^2 \mathbf{E} - \mathbf{r}_i \mathbf{r}_i^\top \right], \quad \text{where} \quad \mathbf{r}_i = \mathbf{x}_i - \mathbf{C} \tag{13}$$

其中 **E**表示单位矩阵。

基于位置的动力学（PBD）。基于位置的动力学（PBD）因其稳定性、简洁性以及处理约束驱动动力学的高效性，成为实时交互式物理仿真中广泛采用的一种范式。Bender 等人(2017）。与通过显式积分力和力矩来计算运动的传统力驱动方法不同，PBD 通过迭代投影粒子位置，使其满足一组预先定义的几何和物理约束，从而确保物理一致性。这种基于投影的形式能够自然地支持较大的仿真时间步，因此特别适用于机器人抓取和交互式环境等高速应用。

预测步骤（隐式积分）。仿真首先使用半隐式欧拉积分预测粒子状态，这种方法具有良好的数值稳定性，并能减少刚性接触过程中的振荡。对于每个粒子 i，其平移运动计算如下：

$$\mathbf{v}_i^{t+\Delta t} = \mathbf{v}_i^t + \Delta t \frac{\mathbf{f}_i^t}{m_i}, \tag{14}$$

$$\mathbf{x}_i^{t+\Delta t} = \mathbf{x}_i^t + \Delta t \, \mathbf{v}_i^{t+\Delta t}, \tag{15}$$

其中，$\mathbf{f}_i^t$表示外力（例如重力或接触冲量），m_i 表示粒子质量，Δt 表示仿真时间步。对于刚体部件，角运动的预测方式如下：

$$\boldsymbol{\omega}_i^{t+\Delta t} = \boldsymbol{\omega}_i^t + \Delta t \, \mathbf{I}_i^{-1} \left(\boldsymbol{\tau}_i^t - \boldsymbol{\omega}_i^t \times (\mathbf{I}_i \boldsymbol{\omega}_i^t) \right), \tag{16}$$

$$\mathbf{q}_i^{t+\Delta t} = \mathbf{q}_i^t + \frac{\Delta t}{2} \tilde{\boldsymbol{\omega}}_i^{t+\Delta t} \mathbf{q}_i^t, \tag{17}$$

其中，$\mathbf{I}_i$ 表示惯性张量，$\boldsymbol{\tau}_i^t$ 表示外力矩，$\mathbf{q}_i^t$ 表示以单位四元数表示的方向。此处，$\tilde{\boldsymbol{\omega}}_i = [0, \boldsymbol{\omega}_i^\top]^\top$ 将角速度嵌入四元数代数中。

约束投影步骤。 获得预测状态后，通过迭代校正来施加位置约束。每个约束 $C(\mathbf{x}_i, \mathbf{q}_i) \geq 0$ 代表一项物理要求（例如避免相互穿透、保持固定距离和体积），并通过基于梯度的位置校正方案求解。为满足约束，位置更新计算如下：

$$\Delta \mathbf{x}_i = -\lambda \frac{1}{m_i} \nabla_{\mathbf{x}_i} C(\mathbf{x}_i), \quad \text{with} \quad \lambda = \frac{C(\mathbf{x}_i)}{\sum_j \frac{1}{m_j} \|\nabla_{\mathbf{x}_j} C(\mathbf{x}_j)\|^2}, \tag{18}$$

其中，拉格朗日乘子 λ 确保以符合物理规律的方式施加约束。采用迭代式高斯-赛德尔或雅可比求解器，使系统收敛至满足约束的有效构型。

速度更新步骤。 施加约束后，更新粒子速度以反映校正后的位置：

$$\mathbf{v}_i^{t+\Delta t} \leftarrow \frac{\mathbf{x}_i^{t+\Delta t} - \mathbf{x}_i^t}{\Delta t}. \tag{19}$$

这确保了位置校正与后续动力学的一致性，在保持动量的同时保留 PBD 的稳定性优势。

讨论。 基于粒子的表述能够实现精细的空间分辨率，并直接抓取几何属性；这对于仿真高自由度机械手与刚性、可变形或复杂形状物体之间的交互尤其有益。此外，约束的隐式处理避免了许多与刚性力模型相关的数值不稳定性，因此 PBD 非常适用于可微仿真场景，在这些场景中，鲁棒性和梯度流至关重要 Standley 等人 (2017)。

A.1.2 高斯泼溅

高斯泼溅已成为机器人真实到仿真流程中的强大技术，可用于以高几何保真度和照片级真实细节捕捉场景、物体和背景。它能够根据单目视频输入灵活高效地建模复杂环境，从而实现精准的空间重建与渲染。在我们的引擎中，我们采用了Lou 等人（2024）提出的真实到仿真流程，该流程可将真实世界扫描视频转换为可用于仿真的资产。借助高斯泼溅，我们能够高效地将重建的物体网格与仿真环境对齐，实现无缝集成。

为进一步提升几何一致性，我们采用了Ye 等人（2024b；a）提出的稳定法线约束，该约束可确保重建点的表面法向量保持一致。这一约束对于保留精细的表面细节并减少噪声尤为重要，尤其是在几何形状复杂或纹理精细的场景中。

总之，这一过程使我们能够为可微分物理建模恢复两个关键组成部分：（1）物体的三维几何形状；（2）物体相对于机械臂的相对位姿。这两者对于准确的系统辨识和仿真对齐都至关重要。

A.2 实现细节

A.2.1 真实到仿真重建的实现

首先，我们采用基于粒子的高斯泼溅方法，构建目标环境在视觉和几何上都十分精确的数字孪生 Kerbl 等人 (2023);Huang 等人 (2024)。通过移动设备拍摄的以环境为中心的（$\mathcal{I}_s$）视频流，我们借助运动恢复结构（SfM）获得经过校准的相机轨迹 Pan 等人(2024); Schönberger 和 Frahm (2016年)。随后，整个流程会训练两组互不相交的高斯基元集合，各自追求不同的目标。

1）体渲染集合。我们维护一组三维高斯

$$\mathcal{P}^{\text{rend}} = \left\{(x_i, y_i, z_i, r_i, g_i, b_i, o_i, s_i, \Sigma_i)\right\}_{i=1}^{N_{\text{rend}}},$$

其中 $(x_i,\ y_i,\ z_i) \in \mathbb{R}^3$ 是第 i 个高斯的中心，$(r_i,\ g_i,\ b_i) \in [0,1]^3$ 是其 RGB 颜色，$o_i \in [0,1]$ 是用于阿尔法混合的不透明度系数，$\Sigma_i \in \mathbb{R}^{3\times3}$ 是指定各向异性范围的对称正定协方差，s_i 表示高斯的语义与实例 ID，N_{rend} 表示此类基元的总数。这些粒子仅针对光度保真度进行优化，从而支持可微体积泼溅，并实现实时新视角合成。

2）表面重建集。几何形状由一组独立的二维表面对齐高斯近似

$$\mathcal{P}^{\text{surf}} = \left\{(x_j, y_j, z_j, \mathbf{t}_{u,j}, \mathbf{t}_{v,j}, s_{u,j}, s_{v,j})\right\}_{j=1}^{N_{\text{surf}}},$$

其中，$(x_j, y_j, z_j) \in \mathbb{R}^3$ 表示圆盘中心，$\mathbf{t}_{u,j}, \mathbf{t}_{v,j} \in \mathbb{R}^3$ 为正交归一的切向量，而 $s_{u,j}$,$s_{v,j} > 0$ 则设定沿这些方向的标准差。向外的表面法线为

$$\mathbf{n}_j = \mathbf{t}_{u,j} \times \mathbf{t}_{v,j}.$$

该集合通过深度失真和法线一致性项进行训练，以确保几何精度，且不受光度损失的影响。

训练完成后，$\mathcal{P}^{\text{surf}}$ 中的表面高斯被栅格化为多视角深度图，融合为截断有符号距离场，再通过移动立方体转换为三角网格。表面法线由 Ye 等人 (2024a) 估计，得到最终的碰撞网格 $\mathcal{M}$。由于 $\mathcal{P}^{\text{rend}}$ 和 $\mathcal{P}^{\text{surf}}$ 不共享参数且采用互不重叠的损失函数，外观的改进不会损害几何保真度。

A.2.2 使用重建的高斯和网格表示构建 MJCF 模型

MuJoCo XML 控制格式（MJCF）编码了关键仿真组件，包括物体的运动学结构、PID 控制增益、刚度参数、碰撞几何体 $\mathcal{K}$ 以及表面点云 $\mathcal{P}^{\text{surf}}$ Zeng 等人（2017），以及驱动关节的规格。为了根据重建得到的高斯泼溅和网格表示构建完整的 MJCF 模型，我们首先将静态环境嵌入为不可移动的背景，并在仿真环境中将重建的物体定义为带有自由关节的刚体。

随后，我们按照 Lou 等人（2024）中描述的变换过程，使重建得到的高斯坐标系和手性与 MuJoCo 的约定保持一致。为确保仿真的真实性，我们提取真实场景中物体与机械臂之间的相对位姿，并将该变换作为仿真中自由关节物体的初始配置。在整合所有相关的位置和控制信息后，我们使用视觉-语言模型（VLM）推断物理参数的初始估计值，包括物体质量；这些参数对于后续的仿真保真度至关重要。

所得的 MJCF 模型具有精确对齐的坐标、初始位姿和几何形状，为后续的系统辨识以及基于物理的策略学习奠定了坚实基础。它还支持高保真渲染和精确的真实到仿真的迁移。

A.2.3 质量辨识的实现

本节阐述我们的质量辨识引擎的两个关键方面：（1）质量-惯性建模策略，以及（2）为支持在具有多样物理属性和几何变化的物体上进行质量学习所必需的自适应参数集合。

质量-惯性建模。在常规设定中，物体的真实质量通常均匀分布在其组成粒子上，如公式 12 所定义。然而，该策略在真实到仿真再到真实优化引擎中常常导致数值不稳定性和梯度爆炸，尤其是在处理包含超过 50,000 个顶点但质量相对较低的高分辨率物体时 Chen 等人 (2025b）。在此类条件下，所得平均粒子质量可能低于 10^{-6} kg，从而引入显著的数值误差。

为缓解这一问题，我们将物体的全部质量分配给每个粒子。重力均匀作用于所有粒子，而外力则根据采样顶点的数量按比例缩放。这种公式通过避免每个粒子的质量值过小，保持了数值稳定性。

此外，由于重建物体的顶点数量各不相同，我们会自适应地选取一部分活动顶点，这些顶点位于物体与机器人手指之间的接触表面上。这进一步提高了仿真保真度，并确保重点关注相关的物理交互。

为确保真实世界观测与仿真环境之间的一致性，我们对 FoundationPose 跟踪系统与以 MJCF 格式定义的 MuJoCo 仿真进行显式同步，包括帧率、时间范围（开始和结束时间）以及空间居中。

接触建模、显式梯度表示、自适应学习参数。我们从在仿真和真实世界环境中执行的机器人动作展开中提取接触点及相应的力。在仿真中，完成真实到仿真重建后，将物体放置在与其真实世界配置一致的相对位置。为确保基于位置的动力学（PBD）引擎中的接触建模保持稳定，我们将物体初始化在略高于地面的位置（例如，$[0.05, 0.05, \text{高度}_2 + 0.01]$），以避免过早发生地面接触并维持仿真稳定性。

真实世界物体轨迹、机器人控制信号与仿真展开之间的精确时间同步，对于可靠的质量辨识至关重要。我们利用 FoundationPose Wen 等人 (2024) 获取精确的物体位姿估计，并据此对齐仿真时间线，以确保观测运动与仿真运动之间的一致性。

为了进行显式梯度计算，我们按照 Jatavallabhula 等人 (2021a) 提出的公式实现半隐式积分方案，从而能够通过接触事件和物体动力学进行可微反向传播。

自适应学习策略 为了适应质量尺度各异的物体，我们采用自适应学习策略。初始时，每个顶点的粒子质量统一设为约 0.002 kg，但为了确保稳定收敛，必须根据物体的总体质量进行调整。对于较重的物体，例如番茄酱瓶（0.8 kg），训练需要更高的学习率和更长的训练周期，通常最多需要 2000 个训练轮次才能收敛。相比之下，中型物体（0.1 kg）通常使用适中的学习率，在 100 个训练轮次内即可高效收敛。轻型物体（0.05 kg）则受益于学习率衰减，同样能在 100 个训练轮次内收敛。

成功进行质量学习还取决于几个关键因素。施加冲量的持续时间由机器人手指与物体之间的活动接触间隔决定，会直接影响估计出的动力学。我们从 FoundationPose 中选取活动跟踪帧，以标记从运动开始到静止的关键转变。此外，我们应用规范重居中向量，使仿真空间中的物体位置对齐，从而减少相机视角差异带来的变化。最后，根据物体的顶点数量和活动接触区域按比例调整估计的接触面积，以便准确建模手物交互 Lu 等人 (2023)；Bronstein 等人 (2022)。

关于 z 轴误差的讨论，来自 FoundationPose。 实际上，FoundationPose 产生的 z 轴误差不可忽视。我们发现，在某些情况下，必须对该误差进行人工后处理，才能使估计位姿更好地与真实到仿真的数字资产对齐。

A.2.4 半隐式欧拉与显式欧拉的消融实验

在本补充章节中，我们扩展了参数辨识公式，并基于常微分方程（ODE）对目标进行了更正式的描述，其中包括边界条件和初始条件、采样及收敛性。

我们设置中的物体动力学由牛顿-欧拉方程建模

$$m\ddot{x}(t) = f_{\text{contact}}(x(t), u(t)) + f_{\text{grav}}, \tag{20}$$

where $x(t)$ 表示物体位姿（或其质心位置），$u(t)$ 表示机器人的 hand actions，且 m 是未知的物体质量。

初始条件和几何约束由以下因素确定：(i) 通过高斯泼溅重建的物体几何和位姿（初始位姿/形状），(ii) 机器人手构型以及

已知的驱动指令（控制输入 $u(t)$），以及 (iii) 同样由重建提供的地平面和环境。

在观测轨迹上沿途的离散时间步 $\{t_k\}_{k=1}^{T}$ 处评估动力学残差，并根据高斯泼溅重建的网格顶点计算接触力。FoundationPose 提供时间 $\{t_k\}_{k=1}^{T}$ 处的真实轨迹 $\{\hat{x}_k\}_{k=1}^{T}$。我们的可微仿真器在这些采样点上强制满足牛顿-欧拉方程，因此，通过最小化适定常微分方程组的残差，推动学习得到的动力学和质量参数的收敛。

在用于质量辨识的特定"向下推"交互中，令 $z(t)$表示物体沿推动轴的位置，$u(t)$表示机器人施加的已知接触力：

$$m\ddot{z}(t) = u(t) - mg. \tag{21}$$

对时间积分两次，得到

$$z(t;m) = z_0 + v_0 t + \frac{1}{m}\int_0^t \int_0^\tau u(s)\, ds\, d\tau - \frac{1}{2} g t^2, \tag{22}$$

可写为

$$z(t;m) = \alpha(t) + \frac{1}{m}\,\beta(t), \tag{23}$$

其中，$\alpha(t)$和 $\beta(t)$仅取决于已知量（初始条件和测得的机器人作用力）。因此，在这一简化接触模型下，轨迹与质量倒数呈仿射关系 $1/m$。对时间离散化可得

$$z_k(m) = \alpha_k + \frac{1}{m}\,\beta_k, \quad k = 1, \ldots, T. \tag{24}$$

我们使用的轨迹损失为，

$$\mathcal{L}_{\text{traj}}(m) = \sum_{k=1}^{T} \left\| z_k(m) - \hat{z}_k \right\|^2, \tag{25}$$

随后便成为关于参数 $\theta = 1/m$ 的标准最小二乘问题。在这种参数化下，$\mathcal{L}_{\text{traj}}(\theta)$ 是关于 θ 的二次函数，因此是凸函数。在实际操作中，我们的学习过程通过可微仿真器中的向量–雅可比积进行，但其底层结构等价于针对质量求解一个条件良好的最小二乘辨识问题。

现在，我们说明在可微仿真器中使用半隐式欧拉格式与使用显式欧拉格式所产生的影响。

For 对于离散时间动力学，我们采用半隐式（辛）欧拉格式。对于时间步长 p
Δt, 更新公式为

$$v_{k+1} = v_k + \Delta t\, m^{-1} f(x_k, u_k), \tag{26}$$

$$x_{k+1} = x_k + \Delta t\, v_{k+1}, \tag{27}$$

whe其中 $f(x_k,\ u_k)$ 包含重力和接触力。相比之下，显式欧拉格式将使用 e

$$v_{k+1} = v_k + \Delta t\, m^{-1} f(x_k, u_k), \tag{28}$$

$$x_{k+1} = x_k + \Delta t\, v_k. \tag{29}$$

这一半隐式更新规则与基于位置的动力学（PBD）方法中常用的积分器密切相关，并且已知其对刚性系统具有更大的稳定域。

在接触丰富动力学中，尤其是对于较小的质量或刚性接触力，显式欧拉的稳定域要小得多，并且往往在 x_k中产生大幅振荡更新。这在我们的设定中尤其成问题，因为我们将质量初始化为一个相对较小的值（保守下界），这会放大有效加速度 $m^{-1}f$，并在使用显式欧拉时导致非常大的位置更新和梯度爆炸。相比之下，半隐式欧拉先更新速度，然后使用更新后的速度推进位置，经验上可得到更稳定、更准确的轨迹。

我们在基于推动的质量识别任务中，通过实验比较了这两种格式。表5 报告了每个物体的识别质量（单位：克）以及每次迭代的平均优化时间（单位：秒）：

物体	真实值（g）	半隐式（g）	显式（g）	半隐式（s）	显式（s）
Lego	59	51	34	1.38	1.21
番茄酱	726	667	685	1.43	1.19
饼干	210	200	189	1.41	1.18
多米诺骨牌	106	117	135	1.36	1.17
U	125	110	98	1.39	1.20
A	134	145	120	1.42	1.22

表5：比较半隐式欧拉格式和显式欧拉格式在质量辨识中的消融实验。我们报告了真实质量、使用各积分器得到的识别质量，以及每次迭代的平均优化时间（单位：秒）。半隐式欧拉能以适度的运行时间开销获得更准确的质量估计。

算法 1 力感知策略训练 输入：物体网格及质量集合： $\{(\mathbf{K}_i, \mathbf{M}_i)\}_{i=1}^{N}$ **输出：学习得到的动作和力：** $\{(\text{动作}_i, \text{力}_i)\}_{i=1}^{N}$ **1：对于每个演示 (**$\mathbf{K}_i$**,**$\mathbf{M}_i$**)，执行以下操作：2：使用 HaMeR Pavlakos 等人（2024）和 MCC-HO Wu 等人（2024a）提取人手姿态和物体位姿。3：将人手姿态及对应的末端执行器位姿重定向到机器人手上。4：位置编码：使用位置编码对顶点进行编码，以获得特征表示。5：数据集构建：准备训练批次，其中包含编码后的顶点、物体质量** $\mathbf{M}_i$ **和真实动作。加载由 Real2Sim 生成的相应 MJCF 文件。6：第一阶段训练（监督式）：将力和接触预测头的真实标签设为 1，以训练策略网络并优化初始抓取预测。7：第二阶段训练（基于仿真的细化）：使用由 Real2Sim 生成的 MJCF 文件，在 MuJoCo 仿真器中执行预测动作。根据仿真结果计算力奖励和接触奖励，并通过反向传播细化模型。8：现实世界部署：使用重建的物体网格，将抓取策略部署到真实机器人系统上，并通过力控制执行预测动作。9：结束循环**

总体而言，显式欧拉法每次迭代的速度略快，但准确度始终较低，与真实质量的偏差更大，优化过程中也更容易出现数值不稳定性。半隐式欧拉法在稳定性与准确度之间取得了更好的平衡，这对于在我们的 Real2Sim 流程中可靠地进行可微质量辨识至关重要。

A.2.5 D-REX 抓取策略的实现

表 6 详细介绍了我们在 *GraspMLP* 中使用的神经网络架构，算法 1 和算法 2 则描述了训练流程。对于标准物体，每个物体的抓取策略使用约 200 条演示数据进行训练。对于几何或动态复杂度较高的物体，我们将数据集扩大至最多包含 5000 条演示数据，以确保充分覆盖策略学习所需的方差，从而提升策略的鲁棒性。根据实证结果，我们发现，结合轻型策略网络、对人手与物体交互的精确建模以及精准的物理信息约束，能够针对每种物体实现可靠且高性能的抓取行为。

A.2.6 基于力的抓取策略的实现细节

对于抓取位置策略，我们采用两阶段训练流程，将人类演示与基于仿真的细化相结合（更多细节请参见附录 A.2.4）。我们的抓取策略记为*GraspMLP*，首先使用人类数据进行训练，随后在仿真中进行细化。

人类演示数据的输入表示与监督。 我们首先收集人类抓取演示，并使用

算法 2 两阶段训练流程

1: **初始化：** 模型参数 θ、优化器、数据加载器 $\mathcal{D}$、环境 $\mathcal{E}$、损失函数：均方误差损失 ($\mathcal{L}_{\text{MSE}}$)、二元交叉熵损失 ($\mathcal{L}_{\text{BCE}}$)。
2: **阶段 1：监督预训练**
3: **对于** 训练轮次 $= 1, \ldots, E_1$ **执行**
4: **对于**批次 $(x, a, r, f) \sim \mathcal{D}$ **执行**
5: 计算预测值：$(\hat{a}, \hat{r}, \hat{f}) \leftarrow$ 模型$(x;\ \theta)$
6: 计算损失：
7: $\mathcal{L}_a \leftarrow \mathcal{L}_{\text{MSE}}(\hat{a}, a)$
8: $\mathcal{L}_r \leftarrow \mathcal{L}_{\text{BCE}}(\hat{r},\ r)$
9: $\mathcal{L}_f \leftarrow \mathcal{L}_{\text{MSE}}(\hat{f}, f)$
10: 反向传播总损失：$\mathcal{L} = \mathcal{L}_a + \mathcal{L}_r + \mathcal{L}_f$
11: 更新参数 θ
12: **结束循环**
13: **结束循环**
14: **阶段 2：环境交互**
15: **对于** 训练轮次 $= 1, \ldots, E_2$ **执行**
16: **对于**批次 $x \sim \mathcal{D}$ **执行**
17: 预测动作和奖励：$(\hat{a}, \hat{r}, \hat{f}) \leftarrow$ 模型$(x;\ \theta)$
18: 在环境 $\mathcal{E}$ 中执行 $\hat{a}$，并观察奖励 r_{env} 和基于接触的力 f_{env}
19: 计算缩放后的真实力：$f_{\text{env}} = \text{clip}(\frac{m \cdot g \cdot \text{num_contacts}}{f_{\max}},\ 0,\ 1)$
20: 计算损失：
21: $\mathcal{L}_r \leftarrow \mathcal{L}_{\text{BCE}}(\hat{r}, r_{\text{env}})$
22: $\mathcal{L}_f \leftarrow \mathcal{L}_{\text{MSE}}(\hat{f}, f_{\text{env}})$
23: 反向传播加权损失：$\mathcal{L} = 0.8\mathcal{L}_r + 0.3\mathcal{L}_f$
24: 更新参数 θ
25: **结束 for**
26: **结束 for**

HaMeR Pavlakos 等人 (2024) 和 MCCHOWu 等人 (2024a) 重建人手和物体的位姿。这些人手轨迹被重定向至目标机器人手，从而生成关节空间参考动作。对于每个物体，我们使用 Real2Sim 流程重建的网格及其识别质量。我们对网格顶点进行均匀采样，对其三维坐标应用位置编码，并将这些特征与物体质量拼接，形成输入 GraspMLP 的逐顶点输入。

网络架构。 GraspMLP 由一个共享 MLP 主干网络组成，该主干网络处理以质量为条件的逐顶点特征，随后接三个任务专用头：(i) 一个用于关节空间抓取动作的 16 维头，(ii) 一个预测基于接触的奖励/可行性分数的 2 维头，以及 (iii) 一个输出 $[0, 1]$ 中归一化抓取力的 1 维头。完整架构（层宽度与激活函数）总结于正文的表 4。

阶段 1：监督预训练。 在第一阶段，我们在人类示范上训练 GraspMLP，以回归关节动作并预测接触/力标签。我们对关节动作和归一化力输出使用 ℓ_2（MSE）损失，并对接触/可行性头使用二元交叉熵（BCE）损失（算法 2，第 3–13 行）。在此阶段，所有已演示的抓取均被视为成功，接触/力标签直接源自人类数据。

阶段 2：基于仿真的细化。 在第二阶段，我们在仿真中细化策略。我们使用 Real2Sim 生成的 MJCF 模型以及已识别的物体质量，在 MuJoCo 中滚动执行 GraspMLP（算法 1，第 6–7 行；附录中的算法 2，第 15–26 行）。根据这些策略展开，我们计算：（i）反映抓取成功与稳定性的基于接触的奖励，以及（ii）基于物体质量、重力和有效接触点数量的缩放后的力目标。随后，我们以 BCE 损失（用于奖励/可行性信号）与均方误差损失（用于力目标）的加权组合更新 GraspMLP，从而在保持监督学习框架的同时，有效地执行强化学习式微调。

部署与力控制。 在部署时，策略以观测到的物体点云（格式与训练顶点相同）以及已识别的质量作为输入。在真实硬件上，我们通过将预测的归一化抓取力映射到执行器电流，实现基于力的控制器。对于采用直驱无刷电机的 Allegro 和 LEAP 手部，电机电流到指尖力的映射近似为线性，因此电流控制可作为控制抓取力的可靠代理。

对不同物体尺度的适用性。尽管我们的实验聚焦于标准桌面尺度的物体，但我们的方法并非天生受限于这些尺寸。质量识别和抓取模块都基于位姿轨迹和重建几何运行，因此原则上也可应用于更小的物体。然而在实践中，非常小的物体受到两个因素的制约：（i）感知流程仍须提供准确的六自由度位姿和足够精细的几何信息；（ii）机器人硬件（包括末端执行器设计、手指尺寸和感测分辨率）必须能够可靠地抓取和操作这类物体。在我们当前的设置中，对于微小物体而言，主要瓶颈在于感知和硬件方面的限制，而非 Real2Sim 或质量识别方法本身的局限。这些挑战是大多数基于位姿的操作方法所共有的。因此，在本研究中，我们聚焦于现有六自由度位姿估计基准和机器人抓取设置中常见的物体尺寸，因为在这些条件下，感知和硬件都较为可靠。

A.2.7 计算细节与耗时

默认情况下，我们的抓取策略使用每个物体包含 200 至 300 个演示位姿的数据集进行训练，在一块 NVIDIA RTX 4090 GPU 上，每个物体的训练时长约为 2 分钟。对于更复杂或方差更大的物体，如果需要提高数据覆盖，我们会将训练数据集扩充至最多包含 5000 条演示数据。在这种情况下，由于批次大小增加，每个物体的训练时间会延长至约 20 分钟。
推理效率很高。策略训练并部署后，只需输入重建后的 URDF 或 MJCF 表示，其中包含物体的几何形状、位姿和物理属性。给定此类输入，策略可在约 0.5 秒内为每个位姿预测一个稳定抓取配置。这种低延迟推理速度使该系统适用于实时场景和机器人端应用，尤其适用于需要快速适应动态物体放置或朝向的场景。

总体而言，我们的引擎在训练成本与部署效率之间实现了良好的权衡，具备可扩展的训练能力，且推理时的运行开销较低。

组件	操作	输出维度	细节
输入	位置编码	$N \times 3$	N 物体顶点（XYZ）
线性层 1	全连接	256	输入： $3 \to 256$
激活函数 1	ReLU	256	非线性
线性层 2	全连接	256	$256 \to 256$
激活函数 2	ReLU	256	非线性
线性层 3	全连接	256	$256 \to 256$
激活函数 3	ReLU	256	非线性
动作头	线性层	16	关节动作输出
奖励头	线性层 +Sigmoid	2	接触约束预测
力头	线性层 +Sigmoid	1	抓取力预测

表6：所提出的GraspMLP网络架构。输入由逐顶点三维坐标组成。共享主干网络将输入映射到潜在特征空间，随后解码为用于预测关节动作、基于接触的奖励信号和抓取力的独立头部。

A.2.8 物体抓取基线方法的实现

Human2Sim2Robot 基线。 在 Human2Sim2Robot 引擎中，Lum 等人 (2025年)，我们假设从人类演示视频中提取的抓取末端执行器位姿

既准确又适合机器人执行。这些抓取姿态通常从手物交互序列中获取，并直接重定向至 LeapHand，所使用的是 Leap Hand 仓库提供的官方重定向实现，从而保留原始抓取意图的空间保真度。

为确保基线比较公平且一致，我们将 Human2Sim2Robot 中使用的原始演示资产和物体网格替换为我们通过真实到仿真重建的网格；这些网格具备第 Real2sim 节所述的光度几何保真度和物理真实性。利用这些资产训练抓取策略直至收敛，这通常需要约 20,000 个训练轮次，才能使奖励信号和行为趋于稳定。

在部署时，我们假设相对末端执行器位姿在 Franka 机械臂和 CuRobo 运动规划框架下是可行的。也就是说，我们预期从人类示范中迁移的抓取位姿无需在真实世界试验期间进行额外的重新规划或修正即可执行。尽管这一假设与原始基线设定一致，但在具有挑战性的物体构型下，可能会限制系统的鲁棒性。

需要注意的是，Human2Sim2Robot 使用的原始控制器 fabric（包括闭环视觉伺服和抓取调整机制）并未公开。因此，我们的重新实现仅关注静态推理：给定固定的 RGBD 帧和已知的物体位姿，系统会预测一个单步抓取动作，不提供在线反馈或纠正性重新规划。我们在评估中考虑了这一限制，以确保比较公平。

DexGraspNet 2.0 基线。 我们采用 DexGraspNet 2.0 提出的两阶段抓取流程 Zhang 等人 (2024c)，该流程通过运动规划将抓取位姿生成与执行分离。然而，我们的方法并非直接根据合成训练数据回归相对平移和旋转，而是通过 MCC-HO 推断这些抓取参数。MCC-HO 是一个预训练模型，可从视频中提取有意义的抓取特征；这些视频记录了真实的人手与物体交互。这些交互以通过 Real-to-Sim 流程重建的几何形状为基础，其中，由点云得到的物体顶点被直接用于估计三维空间中的可行抓取位姿。

生成抓取位姿后，我们使用 CuRobo 进行轨迹规划和执行。规划的轨迹受到机械臂运动学与动力学限制的约束，从而确保推断出的抓取位姿能够安全、可行地部署于现实世界。

为与 DexGraspNet 2.0 进行公平比较（该方法假设所有测试场景中的物体质量均固定为约 0.1 千克），我们将评估范围限制在质量相近的物体上，以匹配其策略的训练条件。然而，不同于这一固定质量假设，我们的方法会利用通过可微真实到仿真再到真实流程识别出的质量，显式优化抓取策略。这使得力感知抓取成为可能，因为识别出的质量可用于改进力预测并提升抓取稳定性。

通过利用人类演示和精确的物理建模，与仅依赖仿真训练数据和启发式质量假设的方法相比，我们的方法在不同物体形状和动态属性下具有更强的泛化能力，同时提升了真实感和适应性。

物体跟踪与运动规划。 我们采用 FoundationPose Wen 等人(2024) 在抓取过程中实时估计物体的六自由度位姿。这一稳健的视觉跟踪系统能够提供时间上连贯的姿态预测，从而支持机器人末端执行器进行动态的碰撞感知轨迹规划。这些物体姿态估计为在杂乱或动态环境中构建抓取轨迹奠定了基础。

在可靠地跟踪物体位姿后，我们引入由 MCC-HO 生成的腕部姿态预测 Wu 等人 (2024a)，这是一种预训练模型，旨在根据人类演示视频重建手—物体交互轨迹。从这些交互中提取的腕部姿态代表了源自人类演示的可行抓取构型。结合真实到仿真得到的物体位姿，它们共同确定了抓取执行所需的目标末端执行器姿态。

为生成无碰撞运动规划，我们使用 CuRobo Sundaralingam 等人(2023) 提出一个受约束的逆运动学（IK）优化问题。具体而言，我们寻找机器人关节构型 $\mathbf{A}^*$，使机器人正向运动学（FK）的输出与……之间的距离最小

目标末端执行器位姿 $\mathbf{X}_{\text{ee}}^{\text{des}}$，同时保持在机器人的无碰撞构型空间 $\mathcal{Q}_{\text{free}}$ 内：

$$\mathbf{A}^* = \text{参数} \min_{\mathbf{A} \in \mathcal{Q}_{\text{free}}} \left\| \text{正向运动学}(\mathbf{A}) - \mathbf{X}_{\text{ee}}^{\text{des}} \right\|_p. \tag{30}$$

此处，$\mathbf{X}_{\text{ee}}^{\text{des}}$通过将物体位姿（利用高斯泼溅和摄影测量重建）与人类演示中的腕部位姿对齐得到，从而形成一个有依据且具有物理意义的抓取目标。范数 $\|\cdot\|_p$（通常为 L_2）用于衡量规划位姿与期望位姿在三维特殊欧几里得群中的空间误差。

这使得抓取执行在物理上合理且与任务相关，并能利用现实世界感知、人类演示先验和可微仿真，完成仿真到现实闭环。

A.2.9 真实世界实验

在我们的实验设置中，场景由五个主要组件组成：固定桌面、固定背景、目标物体、机械臂和机器人手。在每次实验的整个过程中，桌面和背景均保持静止不变，从而提供一致的空间语境。机械臂和机器人手均为全驱动并受到精确控制，所有关节运动均被准确跟踪，以确保可复现性和可靠的系统行为。
目标物体完全是被动的，既不被驱动，也不被直接控制。其运动仅来自与机器人手的物理交互，例如抓取或推动过程中由接触引起的力。这种以物体为中心的动态行为构成了我们系统辨识和策略学习任务的基础。

为进行视觉跟踪，我们采用第三人称视角的Intel RealSense D435i RGB-D相机，其位置可覆盖整个抓取工作空间。为随时间估计6自由度物体位姿，我们使用FoundationPoseWen 等人(2024)，这是一种实时物体位姿估计引擎，即使在遮挡或杂乱环境下也能确保稳健且帧间一致的预测值。

为重建实验场景的几何细节——包括桌子、物体和机器人——我们使用基于智能手机的摄影测量来补充深度相机数据。我们用手机拍摄一段简短的单目视频，再运用多视图立体视觉技术生成环境的稠密三维重建。通过这一过程，我们能够构建高分辨率的物体网格和经过空间对齐的场景表示，随后用于初始化仿真环境并进行现实到仿真的迁移。

精确跟踪与高保真几何重建相结合，为基于真实物理的仿真、物理参数识别以及稳健的现实世界策略部署奠定了基础。

A.2.10 硬件设置

在实验引擎中，我们采用两种不同的机器人手，以满足系统辨识和灵巧抓取的不同需求：Allegro Hand 和 LEAP Hand。两者均配备 16 个可独立驱动的自由度（DoF）。选择这两种平台，是为了在各项实验任务中兼顾机械精度和扭矩能力。

Allegro 和 LEAP Hand 都是直驱式机械手，我们已测试并验证了该方法能够提高其成功率。对于 Inspire 这类连杆驱动式机械手，我们尚不清楚基于力的控制是否有效。

Allegro Hand 是一款广泛应用于灵巧抓取研究的 16 自由度仿人机器人手。它采用内置布线和紧凑的机械结构，可最大限度地减少物理交互过程中的外部干扰。其低矮外形设计和简洁的关节布局简化了运动学与动力学建模，使其非常适合物理参数识别任务，例如物体质量估计。外部线缆的减少有助于实现更稳定的接触建模，并在可微物理优化过程中获得更清晰的梯度流。

LEAP Hand Shaw 等人（2023b）是一款高扭矩、经济高效的机器人手，设计时充分考虑了模块化和现实应用。它由 3D 打印组件与现成执行器组合构成，便于定制、维修和实验。手指长度、关节刚度和指间距等关键机械属性均可调整，以适应特定的抓取场景或物体几何形状。LEAP Hand 采用新颖的腱驱动运动学结构，可实现高度灵巧、类似人手的关节运动。每个关节都能产生超过人手的力矩，同时保持最高约 8 弧度每秒的真实运动速度。

LEAP Hand 的一项核心设计原则，是尽可能提高执行器质量在手部总质量中所占的比例，从而增强握持力，同时保持紧凑的外形。这一设计使其能够处理需要强大且自适应力控制的重型或形状不规则的物体。值得注意的是，LEAP Hand 集成了电流和扭矩限制机制，因此既能实现强力抓取，也能进行轻柔抓取。这些特性使其特别适合执行真实场景抓取任务，因为这类任务要求力控制既稳健又具备柔顺性。

在实验中，我们通过调整 LEAP Hand 执行器的电流上限来调节其施加的抓取力；电流上限与施加的关节力矩呈线性相关。这种控制方案能够根据物体质量和表面属性精确调节接触力，而这对于力感知策略学习中的仿真到现实泛化至关重要。

借助 Allegro 和 LEAP 手部各自的优势，我们的引擎兼具精确的物理建模能力和高性能真实场景抓取能力，从而促进端到端的现实到仿真再到现实学习与部署。

使用不同手部的理由 我们在质量辨识实验中采用**Allegro Hand**，因为它采用紧凑且一体化的机械设计，能够最大限度地减少外部干扰。其内置布线和低扭矩驱动有助于实现稳定、无噪声的接触动力学，因此非常适合需要准确梯度传播和精确系统辨识的任务。在利用可微物理通过机器人与物体交互估计物体质量时，这些特性尤为重要，因为机械噪声或不一致的接触会显著降低优化性能。Allegro Hand 稳定的运动学特性和低惯量结构进一步提升了真实到仿真辨识阶段物体动力学建模的保真度。

我们使用 **LEAP Hand** Shaw 等人 (2023b) 进行抓取任务，因为它具备较强的扭矩能力，以及模块化、类人的运动学结构。LEAP Hand 采用腱驱动，并配备强劲的电机，能够产生远高于 Allegro Hand 的力，因此能够可靠地抓取形状、重量和柔顺性各异的物体。这对于评估现实世界中的策略部署尤为重要，因为在此过程中，鲁棒性和抓取稳定性至关重要。其设计优先考虑力量和灵巧性，因此适合在物理条件逼真的环境中执行力感知策略。该手部还采用电流控制驱动，能够精确调节抓取力；我们利用这一特性使策略适应不同的物体质量。
然而，LEAP Hand 包含外露的线路和腱线布置，会引入机械噪声并增加建模复杂性，尤其是在质量辨识等敏感的参数估计阶段。这些结构因素可能干扰精确的接触建模，并导致可微仿真过程中的力反馈不一致。

通过解耦两只手的职责——使用 Allegro Hand 进行精确的物理参数估计，并使用 LEAP Hand 实现鲁棒抓取——我们能够优化现实到仿真到现实引擎的各个阶段。这种职责分离使我们的引擎能够兼顾准确性与实用性，支持在多种抓取场景下进行高保真建模和现实世界部署。

A.2.11 数据集采集与实验部署

为支持精确的真实到仿真建模，我们使用第三人称 RGB-D 相机（Intel RealSense D435i）或 iPhone 等扫描设备，为每个场景采集约 300 张 RGB图像。这些图像用于高保真三维重建，既能捕捉物体几何形状，也能

捕捉环境背景。每个场景的完整重建过程通常需要约 30 分钟，并生成高斯泼溅表示。

随后，我们将重建的视觉资产转换为可用于仿真的 MJCF。此转换将物体几何编码为碰撞网格，指定物体运动学，并初始化物理参数，以便在 MuJoCo 等仿真环境中使用。我们还提取物体与机器人基座之间的相对位姿，这对于仿真部署期间的对齐至关重要。

我们的实验环境由一台 7 自由度机械臂（Franka Emika Panda）、一只灵巧机器人手（根据任务选择 Allegro 或 LEAP）以及一张放置物体的固定桌面组成。在数据采集和评估期间，机器人系统在与物体交互时执行预定义的控制轨迹或学习得到的策略。同时，FoundationPose Wen 等人(2024) 使用第三视角 RGB-D 视频输入，实时跟踪 6 自由度物体位姿。这确保了现实世界运动与相应仿真轨迹之间的精确对齐。
所有采集的传感器数据——包括 RGB 帧、深度图、机器人关节状态和物体位姿——都会同步并记录下来，以供后续仿真、策略训练和评估使用。该结构化数据集为质量辨识和抓取策略学习奠定基础，使不同实验间能够实现一致的真实到仿真再到真实的迁移。

A.3 消融实验

A.3.1 规模化性能

我们研究了人类演示数据的数量如何影响机器人对一种具有挑战性的物体——紧凑型高密度螺丝刀——的抓取性能。由于该物体接触面积小且转动惯量大，操控难度尤其高，因此是评估学习引擎可扩展性的理想基准。如图 9a所示，我们使用以不同数量演示数据训练的策略，在 20 次真实世界试验中评估抓取成功情况。这些演示数据经过自动筛选，并通过 Real-to-Sim 流程转换为可由机器人执行的抓取位姿。

训练数据集的规模从 255 个到 6,386 个抓取位姿不等，这些位姿提取自 1 至 40 段独特的人类视频示范。结果显示，演示数据的数量与抓取成功率呈明显正相关：只有少量示例时，策略难以泛化，且经常无法稳定物体。然而，随着演示数据数量增加，策略得以充分接触多样的物体构型和交互模式，从而实现更稳健、更一致的抓取。图9b 展示了不同数据规模下学习到的抓取位姿的定性可视化。数据量极少时，策略生成的抓取效果欠佳或不稳定，且往往与物体的几何形状或平衡点不匹配。随着数据集增大，学习到的位姿逐渐更符合物理稳定且类似人类的策略。这些结果凸显了数据集规模对于训练力感知抓取策略的重要性，并表明我们的系统能够有效利用人类视频示范，提升灵巧抓取性能。

A.3.2 使用可微优化在多样物体上实现鲁棒质量识别

我们在图11中展示了可微质量优化过程的三个代表性示例，并说明其在一组多样物体上的收敛性行为：饼干、乐高和番茄酱。在所有情况下，优化均从刻意低估的2g初始质量开始——大约100×、 350×和 30× 分别小于饼干、乐高和番茄酱的真实质量。

对于具有中等质量和接触动力学的饼干物体，尽管初始差距较大，优化仍能平滑收敛到正确值。这表明我们的引擎在轻度质量偏差下具有稳健性。对于接触面较小且惯性较低的乐高物体，较大的初始误差会引发早期过冲。尽管如此，基于梯度的优化器仍能够恢复，并在稳定的迭代次数内将系统引导至正确的质量值。番茄酱瓶因其高质量和复杂几何形状而成为最具挑战性的情形。初始质量与真实质量之间的显著不匹配导致

方法	优化运行时间	峰值信噪比 ↑	结构相似性指数 ↑	学习感知图像块相似度 ↓
4D 高斯泼溅 Wu 等人(2024b)	30 分钟	17.14	0.632	0.411
D-REX（3D 高斯泼溅 +优化轨迹）	20 分钟	**23.54**	**0.783**	**0.237**

表 7：真实推挤番茄酱序列上的真实到仿真重建质量。我们的 D-REX 流水线相较于原始 4D 高斯泼溅提升了光度学指标，同时减少了优化时间。

较高的初始损失。然而，通过采用自适应学习率并增加训练轮数，系统成功收敛到准确的质量估计。

这些示例共同凸显了我们的可微分引擎的灵活性与有效性。无论物体的尺度或动态属性如何，我们的方法都能从较差的初始化可靠地优化质量估计，从而实现力感知策略学习所必需的基于物理的仿真。

A.3.3 真实到仿真重建质量

我们首先评估用于质量辨识和策略学习的重建 4D 场景的质量。对于第 5.1 节中的真实推挤番茄酱序列，我们沿着已执行的机器人轨迹渲染重建场景，并与记录的相机帧进行比较。该序列的视频以及交互式 3DGS查看器均可在我们的匿名项目网站上查看（“通过物体推挤进行质量识别”和“交互式3D工作空间”部分）。重建结果高度还原了观察到的物体运动和接触交互，表明真实到仿真映射在视觉上与真实世界数据高度吻合。

为定量评估真实到仿真保真度，我们将 D-REX Real2Sim流水线（3DGS+ 优化轨迹）与 4D高斯泼溅（4DGS）Wu 等人(2024b) 在相同的推挤番茄酱序列上进行比较。按照标准的 D-NeRF 弹跳球协议，我们使用峰值信噪比、结构相似性指数和学习感知图像块相似度评估机器人轨迹上的渲染质量，并报告优化运行时间。结果汇总于表7。

D-REX 在全部三项光度学指标上均明显优于原始 4D 高斯泼溅，同时所需的优化时间更少。我们将这些提升归因于直接对 6自由度物体位姿进行监督，并在可微物理下将被操作物体建模为单一刚体，而非独立优化每一个高斯。

A.3.4 范围、局限性以及通向通用策略的路径。

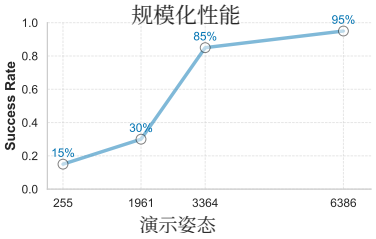
这些结果凸显了我们引擎的灵活性，同时阐明了其适用范围。目前，泛化能力取决于（i）准确的网格重建与质量辨识，以及（ii）接触条件能被我们的刚体模拟器良好近似的任务设置。当前策略仍是物体特定的；不过，以估计质量为条件提供了一种即插即用的信号，可与专为类别级或多物体训练设计的架构相结合（例如，Zhang 等人 （2024c）），从而在具备合适演示数据时获得更通用的策略。将学习建立在人类示范和有针对性的参数辨识之上，可降低对人工设计奖励和大规模机器人采集数据集的依赖，从而实现跨任务的数据高效迁移。

表 8：从较大的电动螺丝刀到较小的电动螺丝刀的跨物体泛化能力。策略基于五个人类示范进行训练，并以物体质量和重建网格为条件。

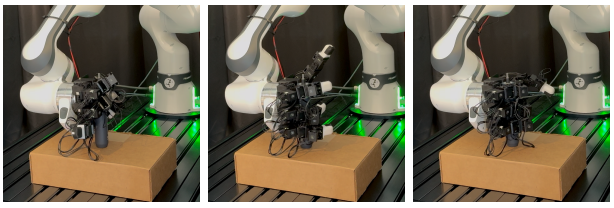
训练物体	**测试物体**	**成功率**
$10 \times 3 \times 3$ 厘米，600 克	$7 \times 2 \times 2$ 厘米，500 克	90% $\rightarrow$ 70%

A.3.5 质量感知学习与域随机化强化学习

我们将 D-REX（基于人类演示数据训练，并以准确推断的物体质量为条件）与 CrossDex Yuan 等人(2024)，一种使用域随机化（DR）的强化学习基线进行比较。CrossDex 在训练期间将质量随机设定在 0.5–1.5kg 范围内，并报告了仿真中的成功率为 89%。



(a)**规模化性能**。 成功率会随着演示数据数量的增加而提高。抓取成功率随数据集规模增大而提升，表明策略能够利用更多演示数据来实现稳健性能。



(b) **不同演示数据数量下抓取姿态的可视化**。 如图所示，使用 1 至 10 条演示数据训练时，策略会产生不稳定抓取；使用 20 条演示数据训练时，则能生成稳定抓取。

图 9：随着演示数据数量增加，我们的力感知抓取策略的规模化性能。

为单独考察质量的影响，我们在一组 *Symbol Y* 物体上进行评估。这些物体的几何形状完全相同，但质量不同： 117 g、 206 g和 324 g (即 0。117、 0。206和 0。324kg）。值得注意的是，这三种测试质量值均低于 CrossDex 的训练范围。如表 9所示，CrossDex 在最重的变体上表现良好，在中型变体上表现尚可，但在最轻的物体上表现不佳。这表明域随机化（DR）部署于分布外（OOD）质量值时的敏感性，尤其是在远离训练支持范围的情况下。
相比之下，D-REX 利用低成本的人类演示来推断物体质量，并据此调整策略，无需额外进行随机化训练即可实现有针对性的适应。尽管相较于 DR 基线的训练范围存在分布外（OOD）偏移，但在这三种质量条件下，抓取成功率始终很高。

表 9：单一物体几何形状（*SymbolY*）不同质量变体下的真实世界抓取成功率；每种条件进行 10 次试验。CrossDex 使用 $[0.5, 1.5]$kg 范围内的质量随机化进行训练；所有测试质量均低于该范围。数值越高越好。

方法	**117 g（轻型）**	**206 g（中型）**	**324 g（重型）**
CrossDex	4/10	7/10	9/10
Ours	9/10	10/10	9/10

这些结果表明了两个互补的结论：（i）域随机化在训练支持范围内或附近能够取得出色表现，但面对更大的分布外质量变化（例如非常轻的物体）时，性能会下降；（ii）显式的质量感知条件化提供了一种简单且数据高效的机制，无需进行广泛的随机化，即可实现跨质量变化的鲁棒迁移。域随机化与质量感知学习并非互斥，但我们的研究结果表明，准确的参数辨识是实现现实世界泛化能力的有力手段，尤其是在部署条件超出域随机化所覆盖范围时。

A.4 学习质量以外的物理参数及处理易碎物体

我们重点估计物体质量，因为它具有明确的真实值，易于通过实验验证，并且会直接、明显地影响抓取性能（例如，欠驱动导致的抓取失败）。原则上，学习其他物理参数——例如摩擦、刚度或阻尼——也是可行的，但考虑到质量易于测量、在不同设置下稳定，且与抓取动力学直接相关，我们优先研究质量。此外，可以在多种物体上系统地改变质量，从而有助于在受控条件下评估跨几何形状和密度的泛化能力。

先前研究已探索在仿真中学习更丰富的物理属性集合（例如，GradSim Jataval- labhula 等人（2021b）），以及从视觉观测中学习稠密物体属性（例如，Xu 等人（2019））。然而，要在真实世界中可靠地验证此类参数，仍然面临更大的挑战

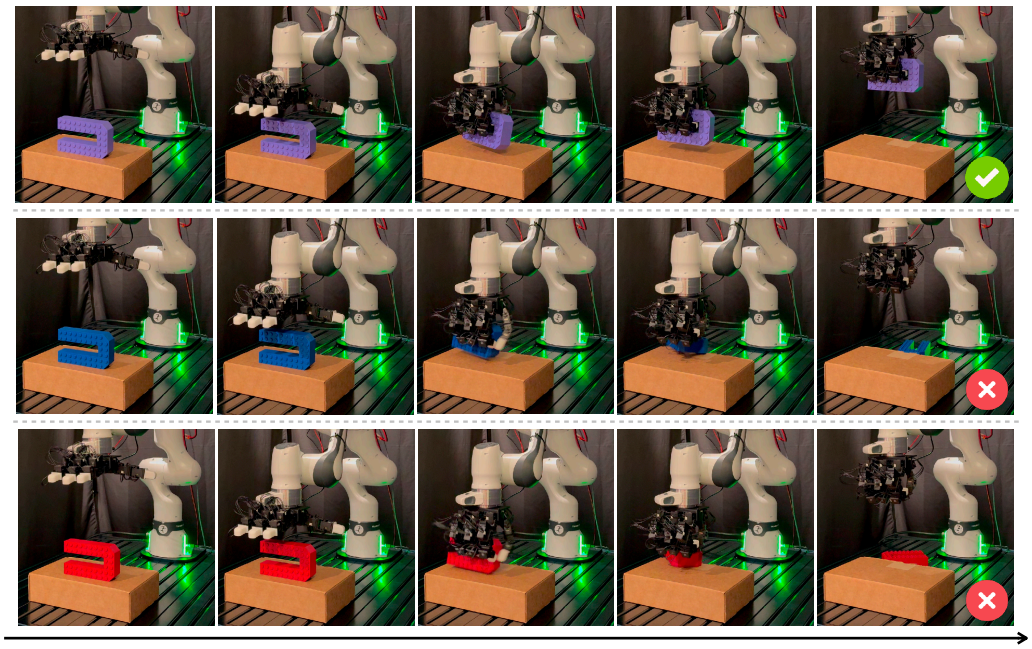


图 10：**在重型物体上训练的物体抓取定量结果**。只有质量匹配的策略才能实现稳定抓取，而不匹配的策略会因作用力过大或不足而失败，导致较轻的物体弹开，较重的物体发生滑移。

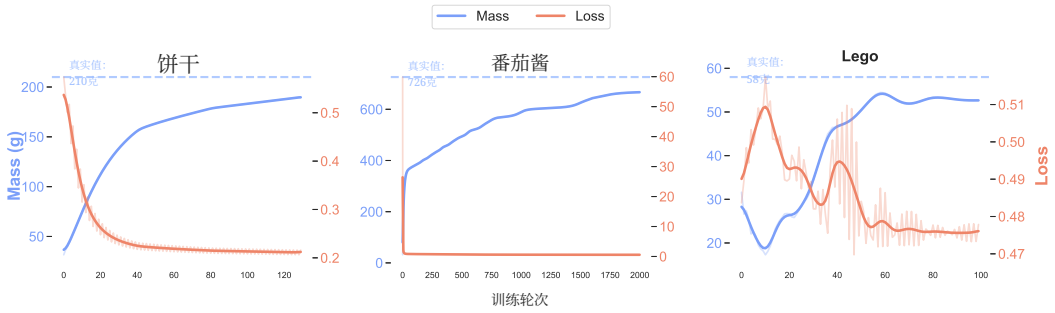


图11：**质量-损失曲线**。我们展示了系统用于质量辨识的三个示例。蓝色曲线表示估计质量，均能可靠地收敛到真实值，展现了我们方法的准确性。

这是因为接触依赖性、时空变化以及对表面条件的敏感性。因此，将参数学习扩展到质量以外的范围，不在本研究的讨论范畴之内。

在当前的 D-REX 引擎中，我们假设刚体动力学，并将物体质量作为唯一可学习的物理参数。这一设计有两个目的：(i) 隔离质量在灵巧抓取中的因果作用；(ii) 证明 D-REX 能够直接从数据中恢复这一关键量。策略以针对同一物体的演示数据为条件，因此是针对特定物体的，而非完全通用的。若要实现更广泛的泛化能力，可以将质量估计模块用作插件：先估计物体的质量，再根据推断出的数值微调（或对其进行条件化）任务策略。

对于无法承受推搡的易碎物体，我们不会尝试学习额外的接触参数，而是采用力度较小的抓取策略来降低接触不确定性，例如，在遵循预定义定向轨迹的同时滚动物体或使其重新定向（通过四元数 球面线性插值）Xu 等人 (2019)，从而限制冲击相互作用，而无需显式估计摩擦特性。

最后，在完成真实到仿真对齐步骤后，执行过程不会受到物理场景中外部人员的干预。这一假设与我们仿真器所依据的拉格朗日刚体动力学保持一致，并确保将观察到的性能差异明确归因于学习得到的质量参数，而非不受控的外部修正。

A.5 关于 D-REX 的泛化能力

正如我们在局限性中所述，当前策略是针对特定物体的：训练和评估假定物体的质量能够被一致地识别，并在仿真与真实系统之间传递。尽管如此，我们在该假设下考察 D-REX 展现跨物体泛化与跨任务泛化的程度。

类别内跨物体迁移。为评估类别内的迁移，我们针对一把更大的电动螺丝刀采集了二十段人类演示片段，并训练了一个以其重建网格和质量为条件的策略。在测试时，无需任何微调，我们将网格和质量替换为同一类别中一把更小螺丝刀的对应值，并执行该策略进行了 10 次试验。如表 8 所示，该策略保持了稳定表现，成功率仅略有下降，表明 D-REX 能够在类别内对几何形状和质量的适度变化进行泛化。定性策略展开见匿名项目网站。

超越抓取：关节物体与精细操作任务。为探究更广泛的适用性，我们在更复杂的抓取场景中评估了 D-REX，包括关节物体交互（例如操作订书机）和精细操作任务（例如操控计算机鼠标）。每项任务均使用 5–10 个人类演示数据，并通过相同的真实到仿真流程进行处理；但对于关节物体的数字资产创建，我们手动进行分割和重新组装，并将质量辨识纳入训练。我们发现，只要重建的仿真能够捕捉到显著的关节结构和任务相关几何形状，策略就能可靠地迁移到这些场景中。这表明 D-REX 并不局限于抓取，还能支持更广泛的物体交互任务。

A.6 基线性能下降

基线性能下降的主要原因是缺乏显式的力（或阻抗）控制，同时训练质量分布范围狭窄。*Human2Sim2Robot*和 *DexGraspNet 2.0* Zhang 等人(2024c) 均完全在仿真中训练，物体质量集中在 $\approx$0.1 千克左右。当部署于明显更重的物品上时（例如 *Spam*、番茄酱、*Nutella*），这些物品超出了训练分布范围，控制器施加的抓取力基本固定或仅能微弱自适应，不足以防止滑移——也就是说，欠驱动导致抓取失败。

对于需要抵抗重力的摩擦抓取，每个接触点所需的法向力会随物体重量增加，并与有效摩擦系数成反比。在简化的平行夹爪情形下，若有两个对称的接触点，

$$F_n \gtrsim \frac{mg}{2\mu}\gamma,$$

其中，m 为物体质量，g 为重力加速度，μ 为有效（取决于任务的）摩擦系数，$\gamma \geq 1$ 涵盖了力旋量分布、接触几何形状和安全裕度。如果 m超出训练期间遇到的范围，或在部署时被低估，固定的位置控制策略就无法相应地调节 F_n，从而违反该不等式并导致滑移。

低摩擦表面（例如保鲜膜或光滑金属）会通过降低 μ进一步提高力需求；当施加的力本已接近临界值时，失败风险也会随之加剧。相反，较轻的物体对轻微的位置或力误差更加宽容，即使控制不够理想，也可能仍能保持稳定。我们也观察到一些较重物体成功抓取的例外情况，这归功于偶然形成的有利几何形状：例如，喷雾瓶喷头可能会意外地在手指间形成部分形封闭（或笼困），从而在一定程度上弥补摩擦支撑不足。

我们的方法通过质量感知力调节增强策略：根据机器人动作和视频估计物体质量，并在测试时根据推断质量调整抓取力（或阻抗设定值）。这种有针对性的适应可在较重

或其他具有挑战性的物体上恢复足够的接触力，减少滑移并提高成功率。更广泛而言，这些发现凸显了质量条件控制对于在各种真实世界物体和表面条件下实现稳健、可泛化的灵巧抓取的必要性。

A.7 构建精准数字孪生的理由

构建精准的数字孪生并应用域随机化，是弥合仿真与现实差距的两种互补策略；具体任务和部署环境不同，各自的优势也有所不同。精准的数字孪生旨在忠实复现现实世界的物理和视觉保真度等，从而实现：1）在真实动力学条件下进行精准的策略评估和基准测试；2）开展系统辨识，尤其适用于高接触任务，或质量、摩擦等敏感物理参数的辨识；3）对物理属性或控制策略进行基于梯度的优化，这需要可微且真实的仿真反馈。

我们的方法不止于视觉或几何层面的数字孪生，还融入了可微系统辨识，以捕捉底层物理规律——这一直是机器人学和图形学领域的一项长期挑战。这使参数适配更加准确、高效，进而提升真实感和策略迁移能力。我们认为，数字孪生与域随机化是互补工具；高保真建模有助于在高接触或动态场景中进行有依据的适配，因为在这类场景中，仅靠随机化可能会忽略关键约束。

A.8 质量辨识与基于力的策略学习之间的关系

我们有意将质量辨识 与策略学习 解耦，以分离质量在仿真到现实迁移中的因果作用，并实现准确评估。具体而言，我们根据少量人类演示和机器人抓取 $\mathcal{D}$ 估计一个标量质量

$$\hat{m} \;=\; \arg\min_{m} \; \mathcal{L}_{\mathrm{id}}(m; \mathcal{D}),$$

然后训练一个显式以该估计值为条件的控制策略，

$$u \;=\; \pi_\theta(x, \hat{m}),$$

其中， x 表示机器人/物体状态， u 表示控制指令。这种两阶段设计避免了参数估计与控制优化之间的混杂，因此能够将下游性能变化明确归因于$\hat{}m$的准确性。

以$\hat{}m$ 为条件的策略能够显式调节力或阻抗设定值，以及前馈重力项（例如， $u \supset g(q; \hat{}m)$）。实际应用中，我们根据$\hat{}m$调整抓取力目标值和柔顺性参数，从而在较重物体上恢复足够的接触力，同时避免对较轻物体施加不必要的高力。由此，无需大量重新训练，即可在较宽的质量范围内提升鲁棒性。

A.9 系统子模块与局限性

先前的可微现实到仿真方法（例如，Chen 等人（2025年b）通常依赖丰富的机器人本体感知（例如，电机扭矩感测）和严格的硬件校准来进行系统辨识。这些要求限制了方法在仪器设备完善的实验室之外的可部署性，与我们的应用场景也大不相同。相比之下，我们采用视觉驱动的辨识方法，仅使用从外部观测到的信号；我们发现，这种方法在实践中更易于使用，也更具可扩展性。因此，我们通过真实物理测量值（例如质量）而非仅在仿真中使用的指标进行评估，从而直接衡量其对现实世界的保真度。据我们所知，很少有现实到仿真引擎能够提供在部署时仍切实可行的端到端可微性；而能够做到这一点的引擎，往往需要在非结构化环境中难以满足的假设。

我们的流程利用 *FoundationPose* Wen 等人 (2024) 作为一种稳健的六自由度位姿估计器。这些位姿是辨识和控制的主要观测信号，无需机载力矩传感。我们将这些估计结果与可微物理引擎（MJX）相结合，该引擎使用 Real2Sim 生成的 MJCF 资产，提供几何形状、惯性属性和接触模型。

GradSim Jatavallabhula 等人 (2021b) 旨在利用仿真生成的渲染图像观测进行系统辨识，并结合基于状态的损失和光度损失。我们的设定有两个根本区别：

真实世界中的部分观测。我们直接处理真实视频，以及由 FoundationPose 估计的六自由度物体位姿Wenet 等人(2024）。不同于Jatavallabhula 等人(2021b）的方法假设可以获取完整的仿真器状态和梯度，我们通过在可微仿真器中最小化随时间变化的状态空间轨迹误差，从部分且含噪的观测中估计物理参数（例如质量）。

在我们的设置中，光度监督并不切实可行。采用Jatavallabhula 等人(2021b）的方法需要对光照进行精心控制、使用标定相机，通常甚至还需要属性已知的 3D打印物体，才能获得可靠的光度损失。我们尝试使用 4D高斯泼溅生成渲染图像以进行光度对齐，但在我们的场景中，优化过程不稳定且精度不高，这进一步凸显了纯粹依赖基于图像的损失进行物理部署的局限性。

物理约束辨识。在我们的引擎中，可微仿真器充当遵循物理定律的数值求解器。给定已知的机器人输入（例如指令关节轨迹）以及从 Real2Sim 获取的准确初始/边界条件，我们将质量估计表述为一个约束优化问题：寻找能够最佳复现观测到的 FoundationPose 轨迹的参数值。

基于这些原因，我们不将Jatavallabhula 等人（2021b）视为评估中的竞争基线。相反，我们在内部复用其可微渲染机制，同时由 MJX 提供机器人手的物理运动学和动力学。

模块化、鲁棒性与可扩展性。我们的系统有意采用模块化设计：(i) 位姿估计（FoundationPose Wen 等人（2024））；(ii) 资产生成和场景重建（Real2Sim）Lou 等人（2024）；Ye 等人（2024b；c）；(iii) 可微物理（MJX、GradSim）Jatavallabhula 等人（2021b）；Todorov 等人（2012）；以及 (iv) 策略学习。类似的多组件设计在机器人学和视觉系统研究中十分常见Pfaff 等人（2025年）；Andrychowicz 等人（2020），因为它们便于针对性改进、随着更好的工具出现而替换子模块，并复用经过充分验证的组件。我们并未将这些模块视为黑箱；相反，我们依据实证可靠性来选择它们，并通过合理性检查和数据筛选将其集成起来。

数据策略与实用性。 我们主要依靠人类演示视频作为策略学习的监督信号。这类视频成本低廉且易于获取（例如，来自公共互联网平台和现有数据集），与机器人执行的演示或强化学习相比，大幅减轻了数据收集负担；后两者通常需要人工设计奖励和漫长的训练周期。个别子模块偶尔出现故障（例如，短暂的位姿估计误差），通常只会导致少量低质量演示数据被筛除；通过规模化仍能保持整体有效性。可扩展的监督方式与可微分的真实到仿真辨识相结合，为实现稳健的仿真到现实迁移提供了一条实用且可扩展的途径。

局限性。 D-REX 框架目前仅支持刚体动力学，并且仅将质量作为主要可学习参数。真实到仿真阶段结束后，为了在所假设的拉格朗日动力学框架下保持一致性，我们的仿真引擎要求真实世界操作场景中不发生人类交互。

A.10 大语言模型的使用

我们仅使用大语言模型对稿件进行语法润色和风格润色。概念构思、实验设计、实现或分析均未依赖这些模型。

A.11 符号列表

符号	描述
I_s	以场景为中心的 RGB 视频序列
I_o	以物体为中心的 RGB 视频序列
$\{I_t\}_{t=1}^{T}$	人类演示 RGB 视频序列
$\{s_t^{\text{real}}\}_{t=1}^{T}$	真实世界物体轨迹
$\{s_t^{\text{sim}}\}_{t=1}^{T}$	模拟物体轨迹
m	优化后的物体质量
π	力感知抓取策略
S	仿真环境表示（MJCF）
K	物体几何形状的碰撞网格
θ	物理仿真参数
P	用于呈现视觉外观的高斯泼溅粒子
s_t	时间步 t 时的物体状态（位置和方向）
u_t	时间步 t 时的物体速度
v_t	时间步 t 时的线速度分量
ω_t	时间步 t 的角速度分量
M	质量惯性矩阵
f	外力与接触力
f_n	接触力向量
k_e, k_d	接触刚度与阻尼参数
$G(\cdot)$	离散时间更新函数
Δt	仿真时间步
L_{traj}	轨迹损失函数
h_t	时间步 t 的人手姿态
o_t	时间步 t 的物体位姿
A_t	时间步 t 的机器人动作
π_ϕ	学习得到的抓取策略（由 ϕ 参数化）
$\hat{A}$	预测的机器人关节位置
$\hat{r}$	预测的接触约束
$\hat{f}$	预测的抓取力约束
n_{active}	机器人与物体之间的有效接触数
ρ	物体密度参数