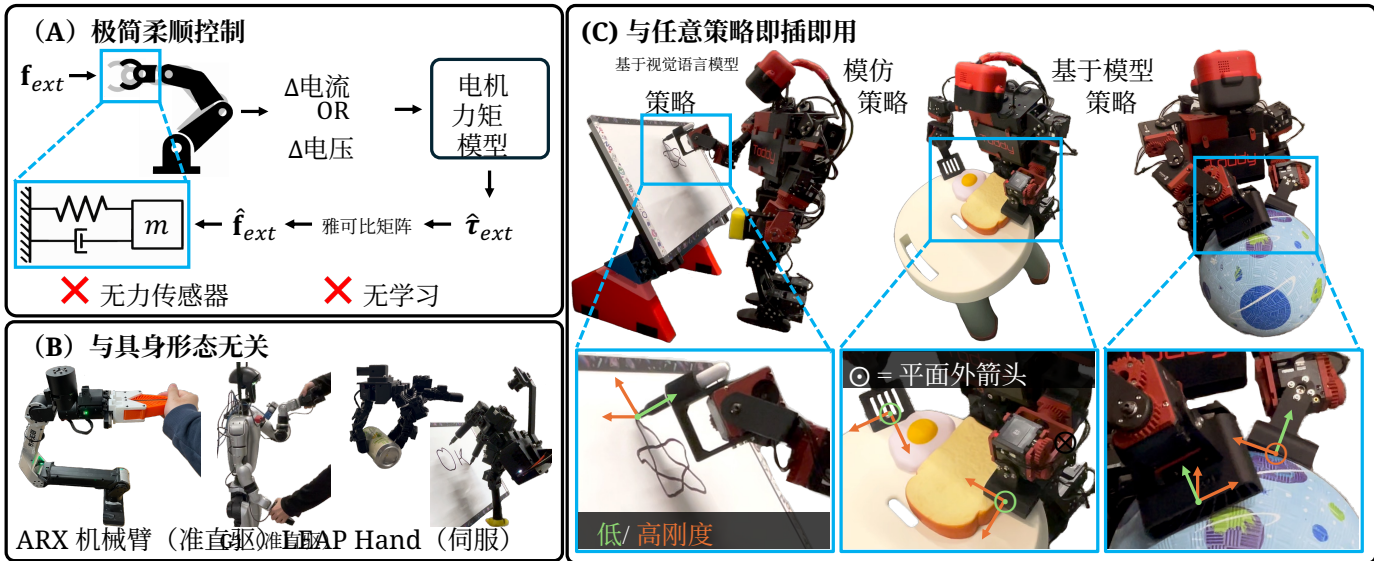
arXiv:2603.00913v1 [cs.RO] 1 Mar 2026

极简柔顺控制

Haochen Shi* **SongboHu*** **YifanHou Weizhuo Wang C.KarenLiu† ShuranSong†**

*共同贡献†共同指导 斯坦福大学

https://minimalist-compliance-control.github.io



ˆ 图1：**极简柔顺控制**（A）无需力传感器或学习，即可利用电机力矩模型和雅可比矩阵，直接根据电机电流或电压信号估算外部力旋量$\mathbf{f}_{ext}$。这些估算值驱动弹簧-质量-阻尼模型，以更新任务空间位置参考。（B）这种极简方法能够适用于不同本体，从机械臂、灵巧手到人形机器人；（C）还可与任意策略即插即用，例如基于视觉语言模型的策略、模仿策略和基于模型的策略，并适用于擦拭、绘画、舀取和手内操作等各种任务。

摘要——**柔顺控制对于安全的物理交互至关重要，但其应用受到力/力矩传感器等硬件要求的限制。近期的强化学习方法虽试图绕过这些限制，却往往面临仿真到现实差距、缺乏安全保证以及增加系统复杂性等问题。我们提出极简柔顺控制，仅利用现代伺服电机和准直驱电机中现成可用的电机电流或电压信号，即可实现柔顺行为，无需力传感器、电流控制或学习。我们根据执行器信号和雅可比矩阵估算外部力旋量，并将其纳入任务空间导纳控制器，在保留足够力测量精度的同时，实现稳定且响应灵敏的柔顺控制。我们的方法不受机器人本体限制，并可与各种高层规划器即插即用。我们在一台机械臂、一只灵巧手和两台人形机器人上，使用视觉语言模型、模仿学习和基于模型的规划，在多种丰富接触任务中验证了该方法。结果表明，该方法在不同本体和规划范式下均能实现稳健、安全且柔顺的交互。**

I. 引言

柔顺控制的广泛应用目前受到显著硬件障碍的阻碍。标准的导纳控制和阻抗控制通常依赖昂贵的力/力矩传感器，或具备精确力反馈能力的执行器而许多广泛使用的机器人平台并不具备这些条件 [11, 26, 21, 25, 43, 17, 37, 33, 44, 9]。为解决这一问题，近期研究提出利用强化学习（RL）来学习柔顺控制策略 [34, 45, 36, 23, 5, 15, 46, 24]。通常，这些方法会通过训练期间的奖励塑造，学习一个位置控制器来模仿柔顺行为。然而，基于强化学习的柔顺控制存在两个主要局限：首先，这些方法在实际部署时会遭遇仿真到现实的差距，例如位置跟踪和刚度精度方面的差异。因此，所学习的控制器往往缺乏安全保证，并可能在物理交互过程中产生意外的大幅力峰值 [45, 36, 23]。此外，基于强化学习的流程会增加框架的复杂性，也可能难以调参 [24, 34]。

为克服硬件障碍，同时避免不必要的复杂性，一个关键发现是，电机电流或 PWM（脉宽调制，本质上就是电压）本身就包含可用于估计外部力旋量的信息。事实上，20世纪90年代的早期研究就利用电机电流进行扭矩估计，但需要精心设计基于观测器的滤波器，以减轻显著的测量噪声并还原具有物理意义的外力 [27, 12]。相比之下，

现代执行器设计采用了显著改进的电流传感和高带宽电流控制回路，并已证明，根据电机电流估计出的扭矩与真实测量值十分接近 [28, 20, 39]。

另一项关键观察是，刚度实现不准确与外部力旋量估计不精确，会以不同方式影响柔顺性。柔顺控制并不要求力的幅值具有极高精度；相反，它主要依赖正确的力符号/方向，以及相关频带内合理的精度，从而确保对扰动做出稳定且及时的响应。只要保留这些特性，近似的力旋量估计就足以产生柔顺行为，而提高幅值精度主要是改善性能，而非安全性 [45]。相比之下，基于强化学习的方法无法明确保证力符号的一致性或频域行为。

因此，我们提出了面向配备现代伺服电机或准直接驱动（QDD）电机的机器人的极简柔顺控制（图 1）。我们的方法（1）对电机特性（例如扭矩常数）进行校准，（2）根据电机电流或 PWM 估计外部力旋量，以及（3）利用估计的力旋量实施任务空间导纳控制。这种方法具有三项优势，共同推动广泛应用于各类机器人系统的柔顺行为普及：

- **无力传感器：**我们仅通过现代执行器中普遍存在的电机电流或 PWM 信号来估计外力和力矩，从而不再依赖直接的力、力矩或触觉感知硬件，大幅降低系统成本和维护需求。值得注意的是，我们的方法甚至不需要电流传感器或电流控制电机；仅凭原始 PWM 信号就足够了。尽管方法简单，这种基于模型的估计仍能保持足够的力测量精度，从而实现稳定、灵敏的柔顺控制。相比之下，基于 RL 的方法无法对此提供明确保证，因而在现实世界接触过程中增加了发生不安全行为的风险。同时，我们的方法还具备其他优势，包括反应式行为，以及应对末端执行器力矩以外扰动的能力。

- **与具身形态无关：**我们的框架以任务空间导纳控制为基础，能够自然地适用于各种形态，从固定基座机械臂到灵巧手和浮动基座人形机器人。

- **与任意策略即插即用：**我们基于模型的方法与多种高层规划策略兼容，且只需进行极少的调参，适用范围涵盖视觉-语言模型（VLM）、模仿学习和基于模型的运动规划。

总之，我们提出了一种极简柔顺控制框架，无需力传感器或学习即可实现柔顺交互，并将这一能力扩展到配备现代伺服和QDD电机的广泛机器人系统。我们在四个不同平台——机械臂、灵巧手和两台人形机器人——上，通过全面实验验证了该方法，涵盖丰富接触任务，例如擦拭、绘画、舀取和手内操作。

II. 相关工作

A. 机器人与力-扭矩传感器

许多广泛使用的机器人平台都缺少末端执行器力/扭矩传感器或关节扭矩传感器。这一局限存在于各种不同形态的机器人中，包括机械臂（如 ARX、ViperX 和 ALOHA [1]）、灵巧手（如 LEAP [29], ORCA [10], 和 DexHand [40]）以及人形机器人（如 Unitree G1、Berkeley Humanoid Lite [7], 和 ToddlerBot [30, 18, 38]）。安装外部六轴力/扭矩传感器虽是一种直接的解决方案，但 ATI Mini-45 等业界标准产品通常体积庞大、耐用性有限，而且价格高昂 [31, 3, 8]。尽管光学 [41], 电阻式 [19], 和磁性 [2]等替代触觉技术已经问世，但它们往往受限于耐久性、信号漂移或容易受到干扰。此外，这些技术通常需要复杂的校准，且主要用于感知法向力和粗略估算剪切力。另一个关键观察是，这些机器人系统均配备了现代伺服电机和准直驱电机，因此我们尝试利用这些平台上可直接获取的电机电流或 PWM 信号来估算外部力旋量。

B. 基于模型的柔顺控制

经典的基于模型的策略，例如任务空间导纳控制和阻抗控制 [11, 26, 21], ，通过虚拟弹簧–质量–阻尼器动力学，为调节接触力提供了有理论依据的框架。然而，这些方法通常需要六轴力/扭矩传感器，或精确的关节扭矩传感与控制。为降低对硬件的要求，先前的研究探索了利用关节级扰动观测器进行无传感器力估计，通过将外部诱导力矩与内部建模的动力学分离，实现无需专用力传感器的力或运动控制 [27, 12]。然而，由于传感精度有限，且20世纪90年代的边缘计算能力不足，当时的电流测量往往噪声极大，需要大量滤波，最终限制了其实际应用价值。近年来，随着电流传感、高带宽电流控制和适当校准技术的改进，研究者已能够直接根据电机电流准确估计扭矩，这主要适用于传动失真较小的系统，例如直驱和QDD电机 [28, 20,39]。我们的工作不再局限于低传动、高扭矩透明度的执行器，并表明，即使不进行电流/扭矩控制，采用高减速比传动（齿轮传动比超过200:1）的伺服电机也能可靠地用于扭矩估计及后续的导纳控制。

C. 基于强化学习的柔顺控制

作为应对硬件限制的另一种接近方式，近期研究探索了将强化学习（RL）用于柔顺控制，旨在通过神经网络从本体感知信号中隐式推断外部力旋量 [24, 36, 45]。这些方法通常分为两类：特定任务策略，它们使用围绕特定任务扩充的数据集进行训练，

运动片段以吸收扰动 [24], ; 以及任务无关策略，它们旨在通过模仿弹簧–质量–阻尼器动力学来实现跨任务泛化 [45, 36, 23]。然而，这些策略的黑箱特性往往会导致仿真到现实的差异，进而缺乏安全保证，且容易出现危险的力峰值。在这项工作中，我们通过显式的力旋量估计和经典的基于模型的控制，实现可靠的柔顺控制，无需仿真到现实迁移，也不依赖黑箱策略。

III. 方法

我们在第 III-A、III-B、III-C、III-D节介绍低层柔顺控制器，并在第 III-E、III-F、III-G节介绍高层规划策略。每种规划策略都可以独立地与我们的柔顺控制器配合使用。图 2总结了每种策略的输入和输出。

A. 弹簧–质量–阻尼模型

在我们的柔顺控制器中，每个末端执行器都被建模为一个能够响应指令运动和外部力旋量的弹簧–质量–阻尼器系统。其动力学由下式描述:

$$m\ddot{\mathbf{x}} = \mathbf{K}_p(\mathbf{x}_{\text{des}} - \mathbf{x}) + \mathbf{K}_d(\dot{\mathbf{x}}_{\text{des}} - \dot{\mathbf{x}}) + \mathbf{f}_{\text{cmd}} + \mathbf{f}_{\text{ext}}, \tag{1}$$

其中，$\mathbf{x} \in \mathbb{R}^6$ 表示末端执行器位姿，其方向由旋转向量表示； m为有效质量（为简化起见设为 $\mathbf{1}$）； $\mathbf{K}_p$ 和$\mathbf{K}_d$ 分别为刚度矩阵和阻尼矩阵； $\mathbf{x}_{\text{des}}$ 和 $\dot{\mathbf{x}}_{\text{des}}$ 分别为期望位姿和速度； $\mathbf{f}_{\text{cmd}}$ 和 $\mathbf{f}_{\text{ext}}$ 分别为指令力旋量和外部力旋量。

B. 电机扭矩估计

大多数伺服电机和准直驱电机无法直接测量电机力矩。因此，我们根据电机电流或可用于推断电机电流的脉宽调制（PWM）信号估算电机力矩。我们按照 [13, 22] 中描述的标准电机模型对电机进行系统辨识。

对于配备电流传感器的电机，我们直接从传感器读取电机电流 I_{w}。对于未配备电流传感器的电机，我们使用以下关系估算电机电流:

$$V_{\text{PWM}} = \text{PWM} \times V_{\text{bus}}, \tag{2}$$

$$V_{\text{emf}} = \frac{\dot{q}}{K_{\text{v}}}, \tag{3}$$

$$I_{\text{w}} = \frac{V_{\text{PWM}} - V_{\text{emf}}}{R_{\text{w}}}, \tag{4}$$

其中， I_{w} 为电机绕组电流， V_{PWM} 为PWM电压， V_{emf} 为反电动势， R_{w} 为绕组电阻，PWM $\in [0, 1]$ 为占空比， V_{bus} 为恒定母线电压，$\dot{q}$ 为电机速度， K_{v} 为速度常数。

根据公式 (3)，我们通过在空载条件下让电机以不同速度运转来校准 K_{v} ，其中$V_{\text{PWM}} \approx V_{\text{emf}}$。我们使用 R_{w} ，通过外部电流传感器测量 I_{w} 在不同 PWM 占空比下的值，然后根据公式R_{w} 求解 (4)。电机扭矩常数 K_{t} 通过电流与力矩之间的关系进行校准:

$$K_{\text{t}} = \frac{\tau_{\text{load}}}{\eta I_{\text{w}}}, \tag{5}$$

其中， τ_{load} 是测得的负载转矩， $\eta \in (0, 1]$ 是计入功率转换损耗后的电机效率， I_{w}是电机绕组电流。若制造商提供了扭矩常数 K_{t}，我们便使用该值；否则，我们会在不同负载和速度条件下校准 K_{t}。这意味着，对于配备电流传感器且制造商提供了 K_{t}的电机，无需进行额外的系统辨识。

为考虑正向驱动和反向驱动中的非对称功率传输，我们采用依赖方向的增益:

$$d = \begin{cases} \text{sign}(\tau_{\text{w}}\dot{q}), & |\dot{q}| > \epsilon_{\text{vel}}, \\ d_{\text{prev}}, & |\dot{q}| \le \epsilon_{\text{vel}}, \end{cases} \tag{6}$$

$$\tau_{\text{load}} = \begin{cases} \eta K_{\text{t}} I_{\text{w}}, & d > 0 \text{ (forward drive)}, \\ \eta^{-1} K_{\text{t}} I_{\text{w}}, & d \le 0 \text{ (backward drive)}, \end{cases} \tag{7}$$

其中， d 是根据功率流 $\tau_{\text{w}}\dot{q}$ 的符号确定的正向/反向驱动状态， ϵ_{vel} 是用于消抖的速度阈值，而 τ_{load} 是关节处的输出力矩。 d_{prev}初始化为 $\mathbf{1}$。在正向驱动（$d > 0$）时，电机加速负载，效率 η 用于计入功率损耗。在反向驱动（$d \le 0$）时，外力反向驱动电机，逆效率 η^{-1} 用于计入传动过程中的制动放大。系统辨识使用在 [30] 中描述的开源电机测试平台所采集的测量数据进行。

为分离外力矩，我们通过从测量力矩中减去 τ_{grav} 来补偿重力:

$$\tau_{\text{ext}} = -(r \cdot \tau_{\text{load}} - \tau_{\text{grav}}), \tag{8}$$

其中， r 是传动比，τ_{grav} 是重力诱导力矩。我们假设系统处于准静态交互工况，在此工况下，关节加速度和科里奥利效应等惯性项和速度相关项，相较于重力和外部施加的力矩可以忽略不计。在此假设下， τ_{ext} 可可靠地估计外部诱导力矩。根据经验，对于大多数需要柔顺性的任务，此假设均成立：交互主要由低频准静态接触主导，而非高动态运动。

C. 外部力旋量估计

外部力旋量根据电机力矩估计，采用一种灵活的公式，既支持全旋量恢复，也支持选择性轴向对齐估计。启用全旋量估计时，我们通过求解正则化最小二乘问题来恢复接触力，并可选择是否恢复力矩:

$$\hat{\mathbf{f}}_{\text{ext}}^{p} = \arg\min_{\mathbf{f}} \left\| J_p^{\top}\mathbf{f} - \boldsymbol{\tau}_{\text{ext}} \right\|^2 + \lambda\|\mathbf{f}\|^2, \tag{9}$$

其中，$\mathbf{f}^{p}{}_{\text{ext}} \in \mathbb{R}^3$ 是$\mathbf{f}_{\text{ext}}$的平移分量，$J_p \in \mathbb{R}^{3\times n}$ 是接触点的平移雅可比矩阵， λ 是正则化系数。力矩分量可使用旋转雅可比矩阵 J_r以类似方式恢复。

为了改善数值条件和噪声鲁棒性，我们可选择仅估计物理上相关的分量——即沿接触法向的力。对于一个

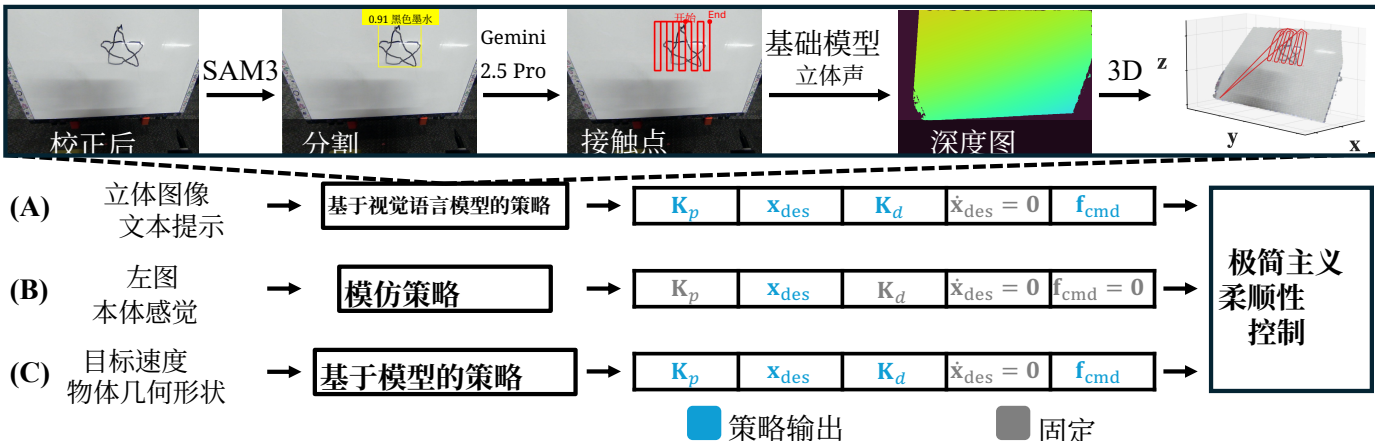


图2：**策略输入与输出**。我们表明，极简柔顺控制可与一系列受益于柔顺交互的高层策略即插即用。我们展示了（A）基于视觉语言模型的策略、（B）模仿策略和（C）基于模型的策略的输入与输出。数学符号对应于式（1）中的符号。（A）和（C）仅预测刚度$\mathbf{K}_p$和力指令$\mathbf{f}_{\text{cmd}}$的方向，而其幅值保持固定。对于这三者，阻尼矩阵均按临界阻尼设置为$\mathbf{K}_d = 2\mathbf{K}_p^{1/2}$，并假设为单位惯性矩阵，其中$(\cdot)^{1/2}$表示矩阵平方根。我们还可视化了VLM流水线，用于预测三维接触点和法向量，并对其进行插值以生成$\mathbf{x}_{\text{des}}$。

所选单位轴 $\hat{\boldsymbol{u}} \in \mathbb{R}^3$，我们将相应的轴对齐力估计为

$$\hat{\mathbf{f}}_{\text{ext}}^{p} = \frac{(\hat{\boldsymbol{u}}^T J_p)\,\boldsymbol{\tau}_{\text{ext}}}{(\hat{\boldsymbol{u}}^T J_p)(\hat{\boldsymbol{u}}^T J_p)^T + \lambda}\,\hat{\boldsymbol{u}}. \tag{10}$$

该表述可简化为一组独立的一维正则化最小二乘问题，从而避免病态的雅可比矩阵求逆，同时在所关注的方向上保持柔顺行为。沿选定轴的力矩分量以相同方式利用 J_r进行估计。

D. 基于逆运动学的导纳控制

给定估计得到的外部力旋量$\mathbf{f}_{\text{ext}}$，通过积分弹簧–质量–阻尼器动力学来实现导纳控制，从而更新任务空间运动参考。在每个控制步 k，根据式（1）计算参考加速度 $\ddot{\mathbf{x}}_{\text{ref}}^{k}$，并使用半隐式欧拉法推进系统：

$$\dot{\mathbf{x}}_{\text{ref}}^{k+1} = \dot{\mathbf{x}}_{\text{ref}}^{k} + \Delta t\,\ddot{\mathbf{x}}_{\text{ref}}^{k}, \tag{11}$$

$$\mathbf{x}_{\text{ref}}^{k+1} = \mathbf{x}_{\text{ref}}^{k} + \Delta t\,\dot{\mathbf{x}}_{\text{ref}}^{k+1}. \tag{12}$$

其中，k 表示离散控制步，Δt 表示控制周期。通过逆运动学（IK）求解器 [42], 将得到的任务空间参考映射到关节空间，从而生成由关节位置控制器跟踪的关节位置目标。

E. 基于*VLM*的规划

第一种规划方法使用视觉基础模型和视觉-语言模型预测三维接触点及法向量，并将其插值为稠密轨迹。

像素空间中的推理。我们的流程首先在校正后的RGB 图像上使用 SAM3 [4] 进行开放词汇分割，以定位目标物体。对于擦拭任务，在掩模边界框内对候选像素进行网格采样；对于绘画任务，则根据预定义的末端执行器工作空间矩形生成候选像素。随后，向视觉语言模型（VLM）提供校正后的图像和候选像素坐标列表，提示其为每个末端执行器输出有序的航点序列。

提升至三维空间。使用 Foundation Stereo 估计深度 [35]，并利用相机内参将其提升至三维空间。随后，使用头部位姿和相机外参，将生成的三维点云从相机坐标系变换到世界坐标系。接着，我们以50 Hz插值生成平滑轨迹，其中包含接近与撤离段、暂停，以及独立的自由空间和接触速度限制。沿估计的表面法向量施加接触力$\mathbf{f}_{\text{cmd}}$；根据表面法向量以及固定的切向和法向值计算刚度矩阵$\mathbf{K}_p$，并在假设临界阻尼的前提下推导阻尼矩阵$\mathbf{K}_d$。该流程如图2所示。

F. 模仿学习

数据采集。我们采用了与 [1]类似的遥操作系统。主臂通过 MuJoCo 中的正向运动学计算末端执行器位姿 [32]，并将其流式传输给从臂；从臂收到的位姿用于设定期望的末端执行器目标$\mathbf{x}_{\text{des}}$。随后，从臂在物理交互过程中使用低级柔顺控制器跟踪该目标。

策略部署。我们使用扩散策略 [6]，该策略采用条件一维 U-Net，在视觉特征的条件约束下预测动作序列上的扩散噪声；视觉特征由 ResNet-18 编码器提取 [14]。在每个控制步，策略将观测到的末端执行器位姿$\mathbf{x}$ 作为输入，并输出期望的末端执行器目标 $\mathbf{x}_{\text{des}}$，随后将其传递给我们的柔顺控制器，以生成$\mathbf{x}_{\text{ref}}$ 供执行。

G. 基于模型的规划

混合力-速度规划。当接触位置已知时，可以使用条件优化的混合伺服控制，通过闭式解计算出稳健的柔顺性曲线

表 I：**与基线方法的定量比较**。 我们报告以毫米为单位的位置跟踪误差和以弧度为单位的姿态跟踪误差，定义为$\mathbf{x}$ 与 $\mathbf{x}_{\text{des}}$之间的偏差。我们还报告人形机器人的根部俯仰角，作为接触力的替代指标。由于接触点与质心之间的距离近似固定，力臂几乎恒定，因此根部俯仰角与施加力相关。

	位置误差（毫米）	姿态误差（弧度）	根部俯仰角（弧度）
UniFP [45]	57.8 ± 30.2	0.147 ± 0.119	0.068 ± 0.020
FACET [36]	22.4 ± 11.0	0.151 ± 0.087	**0.018** ± **0.008**
我们的方法（不含） $\hat{\mathbf{f}}_{\text{ext}}$	22.5 ± 9.6	0.082 ± 0.040	0.060 ± 0.038
Ours	**15.9** ± **5.1**	**0.048** ± **0.043**	0.029 ± 0.012

（OCHS） [16]。在每个步骤中，OCHS接收期望的系统速度和接触信息，以计算混合力-速度控制。力控制动作维持期望接触的安全接合，而速度控制动作则确保物体/机器人的期望速度。为了在我们的柔顺控制器上实现计算得到的混合力-速度动作，我们在速度控制方向上设置高刚度，同时为每个末端执行器在力控制方向上设置低刚度和力偏置。对于具有期望笛卡尔速度 $\boldsymbol{v}$ 的末端执行器，其刚度矩阵计算如下

$$\mathbf{K} = k_{\text{low}}\mathbf{I} + (k_{\text{high}} - k_{\text{low}})\,\frac{\boldsymbol{v}\boldsymbol{v}^{\top}}{\|\boldsymbol{v}\|^{2}}. \tag{13}$$

为简化实现，将OCHS产生的速度分量随时间积分，以生成笛卡尔位置指令；该指令与力和刚度指令一起由我们的柔顺控制器联合执行。采用这一形式，机器人能够执行富接触操作任务，例如手内旋转，同时保持稳健接触行为。

IV. 实验

A. 硬件设置

我们在四种形态各异的机器人平台上验证了我们的柔顺控制方法：配备准直驱电机的ARXX5机械臂；配备准直驱电机的Unitree G1；配备Dynamixel舵机的ToddlerBot [30]；以及配备Dynamixel舵机的LEAP Hand [29] 。这些平台均未配备力/力矩传感器，其中一些甚至缺乏电流传感和控制功能，因此非常适合仅基于电机电流或PWM反馈以及位置控制来评估我们的方法。

B. 外部力旋量估计

在本实验中，我们在运行柔顺控制器的同时，将 ATI Mini45力/力矩传感器抵在末端执行器上，沿每个笛卡尔轴施加推力，并估计——

估计由此产生的三轴接触力。估计力与传感器测量值高度吻合，在自由空间中接近于零，并能准确追踪接触事件。定量结果见图 3。尽管未显式建模摩擦，

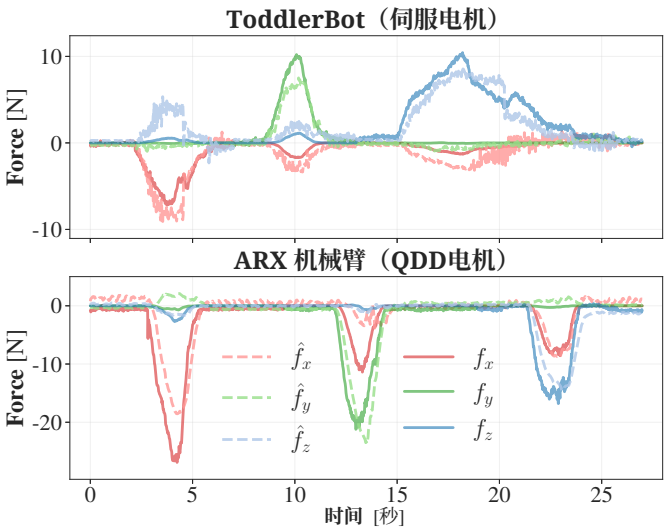


图3：**与力传感器读数的比较**。虚线ˆ（f）表示估计力，实线（f）表示 ATI Mini45传感器的真实测量值。配备伺服电机（传动比为 >200:1）的 ToddlerBot 的平均绝对误差为 0.69 ± 0.73 牛顿，配备准直驱电机（传动比为 ≈ 10:1）的 ARX 机械臂的平均绝对误差为 1.05 ± 1.60 牛顿。

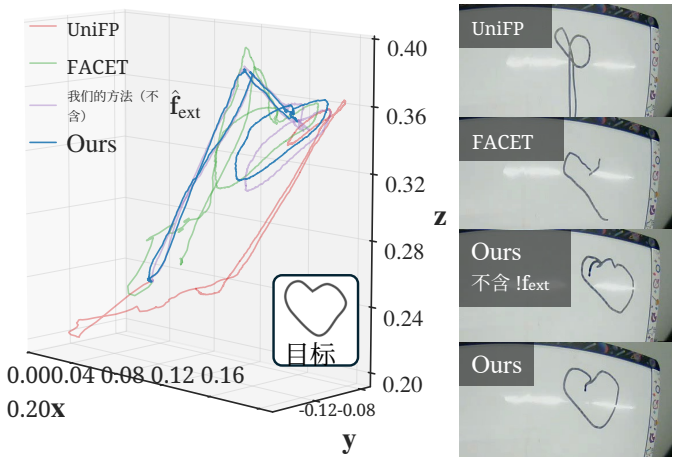


图4：**与基线方法的定性比较**。在此实验中，ToddlerBot 根据视觉语言模型生成的目标轨迹在白板上画心形。所有方法均遵循相同的指令，该指令指定了期望的末端执行器位置$\mathbf{x}_{\text{des}}$、速度 $\dot{\mathbf{x}}_{\text{des}}$、刚度$\mathbf{K}_p$、阻尼$\mathbf{K}_d$以及指令力旋量$\mathbf{f}_{\text{cmd}}$。我们展示了 UniFP 的三维末端执行器轨迹和真实世界执行结果 [45], ˆ FACET [36],我们的方法（不含$\mathbf{f}_{\text{ext}}$）以及我们的完整方法。

考虑驱动状态 d （正向/反向）足以实现精确估计，变速箱效率 η 则隐式涵盖了高减速比带来的摩擦和传动损耗。为降低噪声，我们分别为 ToddlerBot 和 ARX 机械臂应用指数移动平均（EMA）滤波器，取值为 $\alpha = 0.9$ 和 $\alpha = 0.1$ 。较大的 α会使估计响应更灵敏，但噪声也更大，凸显了响应性与平滑之间的权衡。

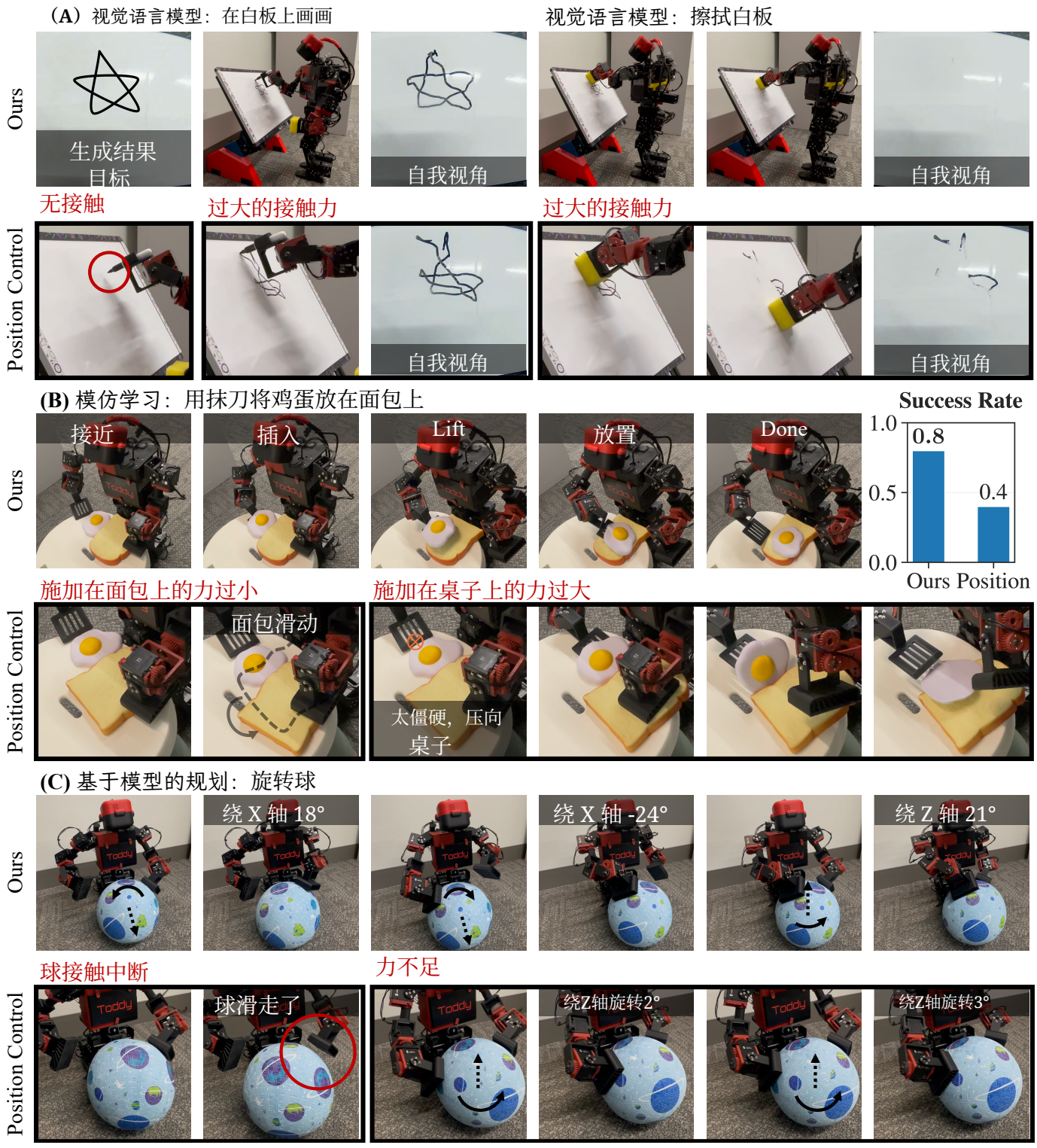


图5：**ToddlerBot实验结果**。我们展示了极简柔顺控制可应用于浮动基座人形机器人，并能作为即插即用模块无缝集成到各种高层规划器中。任务包括：使用基于视觉语言模型的策略在白板上绘画和擦拭；使用抹刀和模仿策略将鸡蛋放在面包上；以及使用基于模型的策略旋转球。奇数行（我们的方法）展示了该方法的成功执行效果，偶数行（位置控制）则展示了基线方法具有代表性的失败案例。我们的控制器能够维持适当的接触力和稳定的交互，而位置控制基线方法经常失去接触、施力不足或施力过大，导致任务失败。施力不足通常无法建立或维持接触，而施力过大会增加切向摩擦，导致更大的切向跟踪误差。

经指数移动平均滤波器处理。ToddlerBot 的 z 轴力在约 $t = 5$s 时出现瞬时尖峰，这是由于雅可比矩阵暂时病态，导致 x 轴和 z 轴的估计相互耦合；在实际应用中，这对任务性能的影响可以忽略不计。

C. 柔顺控制

基线方法实现。为评估我们的柔顺控制器与基于学习的方法，我们将其与两种基于强化学习的基线方法 UniFP [45] 和 FACET [36]进行比较。在这两种情况下，我们都保留原有的强化学习策略公式和实现，同时仅进行少量实现层面的调整，以确保不同方法之间的比较公平。UniFP 基线方法直接使用作者发布的代码库和策略架构，而 FACET 基线方法则经过重新实现，以匹配相同的基础架构。

在训练期间，末端执行器位姿、外部和指令交互力旋量、刚度及阻尼均在机器人的可达工作空间内进行均匀采样。采样范围确保运动学可行性，并与评估期间遇到的轨迹保持一致（图4），从而避免不可行或分布外指令。我们还调整了奖励权重，以强调接触条件下的稳定、灵敏行为。更多详情请参见附录 A。

评估结果。如图4所示，心形目标轨迹由基于视觉语言模型的策略生成，并由所有方法执行。UniFP和FACET难以在与白板保持适当接触力的同时准确跟踪位置。我们方法中的外力旋量估计对于调节接触力也至关重要。表I报告了位置和姿态跟踪误差以及与接触力相关的人形机器人根部俯仰角的定量结果。FACET的根部俯仰角最小，因为它未能对白板施加足够的接触力。

附加特性。我们的控制器每步运行约12ms，其中约 10ms 用于逆运动学求解器 [42]。这些耗时是在Jetson Orin NX上测得的。此外，由于电机电流会对机器人任意位置的扰动作出响应，控制器能够对末端执行器接触之外的情况作出反应。定性结果见补充视频。

D. ToddlerBot结果

我们在ToddlerBot人形机器人平台上评估了三种高层规划器，以证明我们的柔顺控制器能够跨规划范式无缝集成。如图5所示，我们考虑三项任务：使用基于视觉语言模型的策略在白板上进行绘画和擦拭、使用模仿策略借助抹刀将鸡蛋放在面包上，以及使用基于模型的策略旋转球。这些任务需要适当的接触力，从而促使采用柔顺控制。每项任务都包含一个位置控制基线，该基线会因接触丢失、力不足（未能建立或维持接触）或力过大（增加切向摩擦和跟踪误差）而失败。我们观察到，准静态假设并未降低这些任务中的性能。

基于视觉语言模型的策略。对于绘画和擦拭任务，我们在 ToddlerBot 的左手上安装海绵，右手上安装一支笔。给定文本提示（例如“星星”）和白板的初始立体观测后，视觉基础模型和视觉语言模型会预测一条三维接触轨迹，由我们的柔顺控制器在保持接触力的同时执行。绘画完成后，机器人会获取新的观测，以推断用于擦拭的接触点。柔顺控制器通过柔顺交互，补偿有时不够准确的三维接触点和法向量预测。相比之下，位置控制基线要么无法建立接触，要么施加过大的力，导致切向跟踪效果不佳。端到端流程耗时 3–6 秒，其中约 90% 的延迟来自视觉语言模型 API 调用。定性结果见图5。更多细节见附录B。

模仿策略。对于鸡蛋放置任务，我们在 ToddlerBot 的右手上安装了一把抹刀。为训练扩散策略，我们在柔顺控制器启用的情况下收集了 200 次示范，并记录了 $\mathbf{x}_{\text{des}}$ 和 $\mathbf{x}_{\text{ref}}$。为确保公平比较，启用柔顺控制的策略使用 $\mathbf{x}_{\text{des}}$ 作为动作输出，而位置控制基线方法使用 $\mathbf{x}_{\text{ref}}$，即通过基于弹簧–质量–阻尼器模型的柔顺调整后得到的轨迹。由于推理时轨迹并非严格符合分布，柔顺控制器会根据估计的外部力旋量作出响应，从而修正局部误差，而纯位置控制器则无法做到这一点。基线方法常常无法将面包稳定地抵住挡块，或是用抹刀施加过大的力，导致切向跟踪效果不佳。我们的方法成功率为 $16/20 = 80\%$，而位置控制基线方法为 $8/20 = 40\%$。

基于模型的策略。在球体旋转任务中，我们指定目标角速度，并遵循一套固定的旋转指令序列，其中规定了旋转轴以及使用哪只手。对于水平轴旋转，机器人使用单手（左手或右手）；而垂直轴旋转则由双手完成。规划器还将物体的几何形状和尺度作为输入。我们的方法平均实现了 21° 的旋转，而基线方法通常会打滑，或仅将球旋转 2°–3°。旋转角度通过多视角观测进行估计：拟合球体轮廓、匹配表面特征、将这些特征提升至三维球面点，并求解最佳拟合旋转。

E. LEAP Hand 结果

我们进一步在 LEAP Hand 上评估控制器，以展示其在灵巧操作中的有效性。我们考虑两项任务：使用基于视觉语言模型的策略进行白板绘制和擦拭，以及使用基于模型的策略对各种物体进行手内重新定向（图6）。这两项任务都要求精确调节接触力。尽管 LEAP Hand 的机械设计提供了一定的被动柔顺性，位置控制器可在一定程度上加以利用，但我们证明，采用主动柔顺控制对于实现可靠性能仍至关重要。

基于视觉语言模型的策略。我们采用了与 Toddler-Bot 实验类似的设置，但 LEAP Hand 可以同时用多根手指绘画。第三人称视角立体相机观察白板。笔安装在食指和无名指上，海绵安装在中指上。给定提示词“LetterO and LetterK”，基于视觉语言模型的策略会预测



图 6：**LEAP Hand 结果**。我们展示了极简柔顺控制可泛化至不同本体，包括灵巧操作。我们评估了两种高层策略：一种基于视觉语言模型的策略，用于在白板上进行多指绘画和擦拭；另一种是基于模型的规划，用于对各种物体进行手内重新定向。这两项任务都要求精确调节接触力。此外，每项任务都包含一个位置控制基线，但该基线会因接触丢失、力不足或力过大而失效。

三维轨迹，并由柔顺控制器可靠地执行。相比之下，位置控制基线要么无法建立接触，要么因施力过大而难以准确跟踪。

基于模型的策略。在手内重新定向任务中，我们指定目标线速度或角速度，并执行预定义的指令序列。规划器接收物体的几何形状和尺度作为输入。我们的方法平均实现了 3.6 厘米的平移和 27° 的旋转，并能持续进行重新定向，超过三个平移和旋转指令周期。相比之下，位置控制基线方法通常会在最初几条指令内就让物体掉落。

五、结论

总之，极简柔顺控制无需力传感器或学习，仅利用易于获取的执行器信号便能实现可靠的柔顺性，使许多配备现代伺服电机和QDD电机的机器人系统都能轻松采用这一能力，并降低开展丰富接触操作的门槛。我们在机械臂、灵巧手和两台人形机器人上，通过多种丰富接触任务验证了该方法，证明它能够跨不同本体泛化，并

与视觉-语言、模仿和基于模型的策略相结合，从而比位置控制基线实现更稳定的接触。

局限性。 我们当前的接近方法未考虑不可反驱或自锁执行器。具有明显静摩擦或齿隙的电机也未被明确建模，需要额外的系统辨识。若对关节加速度、科里奥利项和热动力学等其他效应进行建模，可能会提高准确性，但这会带来准确性与复杂性之间的权衡。在本研究中，我们优先采用极简主义设计，以确保在丰富接触任务中实现可靠性能所需的准确性。

致谢

作者谨向 Toddy Team 致以诚挚谢意，感谢其提供硬件支持，并感谢 Stanford TML 和 REALab 的成员所进行的有益讨论。本研究部分得到了 NSF 奖项 # 2143601、#2037101、#2132519、#2153854、斯隆研究奖以及斯坦福大学以人为本人工智能研究院的支持。本文所述观点和结论仅代表作者，并不应被视为必然代表资助方明示或暗示的官方政策。

参考文献

[1]Jorge Aldaco、Travis Armstrong、Robert Baruch、Jeff Bingham、Sanky Chan、Debidatta Dwibedi、Chelsea Finn、Pete Florence、Spencer Goodrich、Wayne Gramlich、Alexander Herzog、Jonathan Hoech、Thinh Nguyen、Ian Storz、Baruch Tabanpour、Jonathan Tompson、Ayzaan Wahid、Ted Wahrburg、Sichun Xu、Sergey Yaroshenko和 Tony Z Zhao。ALOHA 2：一种用于双臂遥操作的增强型低成本硬件，2024年。[2]Raunaq Bhirangi、Venkatesh Pattabiraman、Enes Erciyes、Yifeng Cao、Tess Hellebrekers和Lerrel Pinto。AnySkin：面向机器人触觉的即插即用皮肤传感。载于*2025*年*IEEE*机器人与自动化国际会议（*ICRA*），第16563–16570页，2025年5月。doi: 10.1109/ICRA55743.2025.11128638。[3] Max Yiye Cao、Stephen Laws和Ferdinando Rodriguez y Baena。面向机器人应用的六轴力/扭矩传感器：综述。《*IEEE*传感器期刊》，21(24): 27238–27251，2021年12月。ISSN 1558-1748。doi: 10.1109/JSEN.2021.3123638。[4] Nicolas Carion、Laura Gustafson、Yuan-Ting Hu、Shoubhik Debnath、Ronghang Hu、Didac Suris、Chaitanya Ryali、Kalyan Vasudev Alwala、Haitham Khedr、Andrew Huang、Jie Lei、Tengyu Ma、Baishan Guo、Arpit Kalla、Markus Marks、Joseph Greer、Meng Wang、Peize Sun、Roman R¨dle、Triantafyllos Afouras、Effrosyni Mavroudi、Katherine Xu、Tsung-Han Wu、Yu Zhou、Liliane Momeni、Rishi Hazra、Shuangrui Ding、Sagar Vaze、Francois Porcher、Feng Li、Siyuan Li、Aishwarya Kamath、Ho Kei Cheng、Piotr Doll´ar、Nikhila Ravi、Kate Saenko、Pengchuan Zhang和 Christoph Feichtenhofer。

SAM 3：基于概念的分割一切，2025年11月。[5]Sirui Chen、Zi-ang Cao、Zhengyi Luo、Fernando Casta˜neda、Chenran Li、Tingwu Wang、Ye Yuan、Linxi "Jim" Fan、C. Karen Liu 和 Yuke Zhu。CHIP：通过事后扰动实现人形机器人控制的自适应柔顺性，2025年12月。[6] Cheng Chi、Siyuan Feng、Yilun Du、Zhenjia Xu、Eric Cousineau、Benjamin Burchfiel 和 Shuran Song。扩散策略：通过动作扩散实现视觉运动策略学习。载于*Robotics: Science and Systems XIX*。Robotics: Science and Systems Foundation，2023年7月。ISBN 978- 0-9923747-9-2。doi: 10.15607/RSS.2023.XIX.026。[7]Yufeng Chi、Qiayuan Liao、Junfeng Long、Xiaoyu Huang、Sophia Shao、Borivoje Nikolic、Zhongyu Li 和 Koushil Sreenath。展示 Berkeley Humanoid Lite：一款开源、易于使用且可定制的三维打印人形机器人，2025年4月。[8] Hojung Choi、Jun En Low、Tae Myung Huh、Gabriela A. Uribe、Seongheon Hong、Kenneth A. W. Hoffman、Julia Di、Tony G. Chen、Andrew A. Stanley 和 Mark R. Cutkosky。CoinFT：一种用于机器人应用的硬币大小电容式六轴力/力矩传感器。https://arxiv.org/abs/2503.19225v2，2025年3月。[9] Hojung Choi、Yifan Hou、Chuer Pan、Seongheon Hong、Austin Patel、Xiaomeng Xu、Mark R. Cutkosky 和 Shuran Song。借助 UMI-FT 在真实环境中进行柔顺操作，2026年1月。[10]Clemens C. Christoph、Maximilian Eberlein、Filippos Katsimalis、Arturo Roberti、Aristotelis Sympetheros、Michel R. Vogt、Davide Liconti、Chenyu Yang、Barn- abas Gavin Cangan、Ronan J. Hinchet 和 Robert K. Katzschmann。ORCA：一款开源、可靠、经济高效的仿人机器人手，可持续进行灵巧任务学习，2025年9月。[11]J.J. Craig 和 M.H. Raibert。机械臂混合位置/力控制的系统方法。载于*COMPSAC 79. Proceedings. Computer SoftwareandThe IEEEComputer Society's ThirdInternational Ap-plications Conference,1979.*，第446–451页，1979年11月。doi：10.1109/CMPSAC.1979.762539。[12] K.S. Eom、I.H. Suh、W.K. Chung 和 S.-R. Oh。基于扰动观测器、无需力传感器的机器人机械臂力控制。载于*Proceedings. 1998IEEE Inter-national Conferenceon Robotics and Automation (Cat.No.98CH36146)*，第4卷，第3012–3017卷第4页，1998年5月。doi：10.1109/ROBOT.1998.680888。[13]Jos´e Carlos Gamazo-Real、Ernesto V´azquez-S´anchez 和 Jaime G´omez-Gil。采用无传感器技术的无刷直流电机位置与速度控制及应用趋势。*Sensors(Basel, Switzerland)*，10(7):6901–6947，2010年7月。ISSN 1424-8220。doi：10.3390/s100706901。[14] Kaiming He、Xiangyu Zhang、Shaoqing Ren 和 Jian Sun。用于图像识别的深度残差学习。载于*2016 IEEEConferenceonComputerVisionandPattern*

Recognition (CVPR)，第770–778页，拉斯维加斯，美国内华达州，2016年6月。电气电子工程师学会。ISBN 978-1-4673-8851-1。doi：10.1109/CVPR.2016.90。[15] Zewen He、Chenyuan Chen、Dilshod Azizov 和 Yoshihiko Nakamura。CoTaP：柔顺任务流程及其控制器的强化学习与柔顺性调节，2025年10月。[16] Yifan Hou 和 Matthew T. Mason。一种用于最优混合力-速度控制的高效闭式方法。载于*2021*年*IEEE*机器人与自动化国际会议*(ICRA)*，第11125–11131页，2021年5月。doi：10.1109/ICRA48506.2021.9561426。[17] Yifan Hou、Zeyi Liu、Cheng Chi、Eric Cousineau、Naveen Kuppuswamy、Siyuan Feng、Benjamin Burchfiel 和 Shuran Song。自适应柔顺策略：为扩散引导控制学习近似柔顺性。载于*2025*年*IEEE*机器人与自动化国际会议*(ICRA)*，第4829–4836页，2025年5月。doi：10.1109/ICRA55743.2025.11128452。[18]Kaizhe Hu、Haochen Shi、Yao He、Weizhuo Wang、C. Karen Liu 和 Shuran Song。机器人训练机器人：人形机器人的真实世界策略自动适应与学习，2025年8月。[19] Binghao Huang、Yixuan Wang、Xinyi Yang、Yiyue Luo 和 Yunzhu Li。3D-ViTac：利用视觉-触觉感知学习精细操作。载于第八届机器人学习年会，2024年9月。[20] Marco Hutter、Christian Gehring、Dominic Jud、Andreas Lauber、C. Dario Bellicoso、Vassilios Tsounis、Je min Hwangbo、Karen Bodie、Peter Fankhauser、Michael Bloesch、Remo Diethelm、Samuel Bachmann、Amir Melzer 和 Mark Hoepflinger。ANYmal——一种高度灵活、动态性能出色的四足机器人。载于*2016 IEEE/RSJInternationalConference on Intelligent Robots and Sys-tems (IROS)*，第38–44页，2016年10月。doi：10.1109/ IROS.2016.7758092。[21]O. Khatib。机器人机械臂运动与力控制的统一方法：操作空间公式。*IEEE*机器人与自动化期刊，3(1):43–53，1987年2月。ISSN 2374-8710。doi：10.1109/JRA.1987.1087068。[22] Ung Hee Lee、Tor Shepherd、Sangbae Kim、Avik De、Hao Su、Robert Gregg、Luke Mooney 和 Elliott Rouse。如何为轻量化机器人系统的设计建立无刷电动机模型，2023年9月。[23]Qingzhou Lu、Yao Feng、Baiyu Shi、Michael Piseno、Zhenan Bao 和 C. Karen Liu。GentleHumanoid：学习适用于高接触人机与物体交互的上半身柔顺性，2025年11月。[24]Gabriel B. Margolis、Michelle Wang、Nolan Fey 和 Pulkit Agrawal。SoftMimic：从示例中学习全身柔顺控制，2025年10月。[25] Roberto Mart´ín-Mart´ín、Michelle A. Lee、Rachel Gard- ner、Silvio Savarese、Jeannette Bohg 和 Animesh Garg。末端执行器空间中的变阻抗控制：一种用于接触环境中强化学习的动作空间

丰富接触任务。载于*2019*年*IEEE/RSJ*智能机器人与系统国际会议（*IROS*），第1010–1017页，2019年11月。doi：10.1109/IROS40897.2019. 8968201.[26] Matthew T. Mason。计算机控制机械臂的柔顺与力控制。*IEEE*《系统、人类与控制论汇刊》，11(6)：418–432，1981年6月。ISSN 2168-2909。doi：10.1109/TSMC.1981. 4308708. [27] T. Murakami、F. Yu 和 K. Ohnishi。多自由度机械臂的无转矩传感器控制。*IEEE*《工业电子学汇刊》，40(2)：259–265，1993年4月。ISSN 1557-9948。doi：10.1109/41.222648.[28]Sangok Seok、Albert Wang、Meng Yee Chuah、David Otten、Jeffrey Lang 和 Sangbae Kim。高效四足机器人的设计原则及其在 MIT猎豹机器人上的实现。载于*2013*年*IEEE*机器人与自动化国际会议，第3307–3312页，2013年5月。doi：10.1109/ICRA.2013.6631038.[29]Kenneth Shaw、Ananye Agarwal 和 Deepak Pathak。LEAP Hand：面向机器人学习的低成本、高效仿人手。载于*Robotics:Science andSystems XIX*。Robotics: Science and Systems Foundation，2023年7月。ISBN 978-0-9923747-9-2。doi：10.15607/ RSS.2023.XIX.089.[30] Haochen Shi、Weizhuo Wang、Shuran Song 和 C. Karen Liu。ToddlerBot：面向移动操作的开源、兼容机器学习的人形机器人平台，2025年2月。[31]Stefano Stassi、Valentina Cauda、Giancarlo Canavese 和 Candido Fabrizio Pirri。基于压阻复合材料的柔性触觉传感：综述。传感器（巴塞尔，瑞士），14(3)：5296–5332，2014年3月。ISSN 1424-8220。doi：10.3390/s140305296.[32] 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，2012年10月。doi：10.1109/IROS.2012.6386109.[33]Changhao Wang、Xiang Zhang、Zhian Kuang 和 Masayoshi Tomizuka。笛卡尔空间变阻抗控制的安全在线增益优化。载于*2022*年*IEEE*第*18*届自动化科学与工程国际会议（*CASE*），第751–757页，2022年8月。doi：10.1109/CASE49997.2022.9926697.[34] Lai Wei、Xuanbin Peng、Ri-Zhao Qiu、Tianshu Huang、Xuxin Cheng 和 Xiaolong Wang。HMC：学习用于接触丰富移动操作的异构元控制，2025年11月。[35] Bowen Wen、Matthew Trepte、Joseph Aribido、Jan Kautz、Orazio Gallo 和 Stan Birchfield。FoundationStereo：零样本立体匹配，2025年4月。[36] Botian Xu、Haoyang Weng、Qingzhou Lu、Yang Gao 和 Huazhe Xu。FACET：通过阻抗参考跟踪实现足式机器人的力自适应控制。载于第九届机器人学习年会，2025年9月

.

[37] Xiaomeng Xu、Yifan Hou、Zeyi Liu 和 Shuran Song。合规残差 DAgger：通过人类纠正提升真实世界接触丰富操作。载于第三十九届神经信息处理系统年会，2025年10月。[38] Tae Hoon Yang、Haochen Shi、Jiacheng Hu、Zhicong Zhang、Daniel Jiang、Weizhuo Wang、Yao He、Zhen Wu、Yuming Chen、Yifan Hou、Monroe Kennedy、Shuran Song 和 C. Karen Liu。超越足部的运动，2026年1月。[39] Shuangyue Yu、Tzu-Hao Huang、Xiaolong Yang、Chunhai Jiao、Jianfu Yang、Yue Chen、Jingang Yi 和 Hao Su。用于轻量级髋部外骨骼的准直接驱动执行：高反驱性与高带宽。*IEEE/ASME*机电一体化学报，25(4)：1794–1802，2020年8月。ISSN 1941-014X。doi：10.1109/TMECH.2020.2995134.[40] Jianbo Yuan、Haohua Zhu、Jing Dai 和 Sheng Yi。具备本体感知柔顺控制的仿生腱驱动 DexHand 021的开发，2025年11月。[41] Wenzhen Yuan、Siyuan Dong 和 Edward H. Adelson。GelSight：用于估计几何形状和力的高分辨率机器人触觉传感器。*Sensors*，17(12)：2762，2017年12月。ISSN 1424-8220。doi：10.3390/s17122762.[42] Kevin Zakka。Mink：基于 MuJoCo 的 Python逆运动学，2025年5月。[43] Xiang Zhang、Liting Sun、Zhian Kuang 和 Masayoshi Tomizuka。通过逆强化学习为力相关任务学习变阻抗控制。*IEEE*机器人学与自动化快报，6(2)：2225–2232，2021年4月。ISSN 2377-3766。doi：10.1109/LRA.2021. 3061374.[44] Xiang Zhang、Changhao Wang、Lingfeng Sun、Zheng Wu、Xinghao Zhu 和 Masayoshi Tomizuka。通过在线导纳残差学习，高效实现接触丰富操作技能的仿真到现实迁移。载于第七届机器人学习年会，2023年8月。[45] Peiyuan Zhi、Peiyang Li、Jianqin Yin、Baoxiong Jia 和 Siyuan Huang。为足式移动操作中的位置与力控制学习统一策略。载于第九届机器人学习年会，2025年9月。[46] Xiang Zhou、Xinyu Zhang 和 Qingrui Zhang。HAC-LOCO：在持续外部扰动下为四足运动学习分层主动柔顺控制，2025年7月。

附录

A. 强化学习基线超参数

输入指令范围。训练期间，我们会在表 II所示范围内采样输入指令；FACET 和 UniFP 使用相同的指令结构和范围。

表 II：训练 UniFP 和 FACET 的指令范围。

指令	范围
左手	$\mathbf{x}_{\text{des}} \leftarrow \mathbf{0}$ $\mathbf{f}_{\text{ext}} \leftarrow \mathbf{0}$ $\mathbf{f}_{\text{cmd}} \leftarrow \mathbf{0}$
右手位置 $\mathbf{x}^{\text{pos}}$ des	$\mathbf{x} \in [0.00, 0.10]$ m $\mathbf{y} \in [0.00, 0.09]$ m $\mathbf{z} \in [0.00, 0.15]$ 米
右手方向 $\mathbf{x}^{\text{rot}}_{\text{des}}$	$\mathbf{r_x} \in [0.00, 0.10]$ 弧度 $\mathbf{r_y} \in [0.2, 0.34]$ 弧度 $\mathbf{r_z} \in [0.00, 0.00]$ 弧度
右手 $\mathbf{f}_{\text{ext}}$	$\mathbf{f}_{\text{ext},\ x} \in [0, 5]$ 牛顿 $\mathbf{f}_{\text{ext},y} \in [0, 3]$ 牛顿 $\mathbf{f}_{\text{ext},\ z} \in [0, 3]$ 牛顿
右手 $\mathbf{f}_{\text{cmd}}$	$\mathbf{f}_{\text{cmd},x} \in [0, 5]$ 牛顿 $\mathbf{f}_{\text{cmd}\ y} \in [-1.5, 1.5]$ 牛顿 $\mathbf{f}_{\text{cmd},z} \in [-2, 0]$ 牛顿
右手$\mathbf{K}_p$ 和$\mathbf{K}_d$	$\mathbf{K}_p^{\text{pos}},_x \in [120, 400]$ $\mathbf{K}_p^{\text{pos}},_y \in [350, 400]$ $\mathbf{K}_p^{\text{pos}},_{\ z} \in [350, 400]$ $\mathbf{K}_p^{\text{rot}} \in [20, 40]^3$ $\mathbf{K}_d = 2\mathbf{K}_p^{1/2}$

奖励。 表III 总结了用于训练 UniFP 和 FACET 的奖励项及其对应权重。大多数权重与原论文中的不同，并针对我们的机器人和任务进行了精心调优。

表 III：UniFP 和 FACET 的奖励项及权重。

Term	定义	UniFP	FACET
末端执行器位点位置	实验值$(-\sigma_{\text{pos}}\|\mathbf{x}^{\text{pos}} - \mathbf{x}^{\text{pos}}_{\text{参考值}}\|^2)$	10.0	6.0
末端执行器位点速度	实验值$(-\sigma_{\text{linvel}}\|\dot{\mathbf{x}}^{\text{pos}} - \dot{\mathbf{x}}^{\text{pos}}_{\text{参考值}}\|^2)$	1.0	0.6
末端执行器位点旋转	实验值$(-\sigma_{\text{rot}}\|\mathbf{x}^{\text{rot}} \ominus \mathbf{x}^{\text{rot}}_{\text{ref}}\|^2)$	3.0	4.0
电机位置	$\exp(-\|q - q_{\text{ref}}\|^2)$	4.0	0.5
能量	$-$平均值$(\|\tau\dot{q}\|^2)$	0.1	1.0
动作速率	$-$平均值$(\|a_t - a_{t\text{-}1}\|^2)$	0.08	2.0
碰撞	$-\sum \mathbf{1}(\|\mathbf{f}_{\text{contact}}\| > 0.1)$	3.0	3.0

B. 视觉语言模型实现细节

提示词。 我们使用结构化提示词来指定任务、目标物体，以及按末端执行器位置分组的候选接触点（像素坐标），并要求模型针对每个位置返回有序的 JSON 航点序列（仅包含关键帧；控制器随后对轨迹进行稠密化）。

图像： {图像}
任务： {任务描述}
目标物体： {物体标签}
候选接触点：
{候选线条}
动作要求：
为每个位置输出按顺序排列的航点列表（像素坐标）
以遍历目标区域。航点列表应定义一条单一的
连续路径，扫过标记区域且不离开
确保任何主要子区域都未遗漏。重点关注路径拓扑——覆盖顺序、
方向变化和子区域转换。我们稍后会加密路径；你的任务只是选出有意义的关键帧。
现在；你的任务只是选出有意义的关键帧。
JSON 输出格式：
{JSON 行}
JSON 规则：
返回有效的 JSON，使用双引号。
对于动态序列，必须严格保持各点的顺序。
请勿包含解释。
开始输出：

图 7：**视觉语言模型提示模板**。对于擦拭任务，任务描述为：“用橡皮擦擦掉白板上的黑色墨水”，默认的目标物体为“黑色墨水”。对于绘画任务，任务描述为：“用笔在白板上绘制 {物体标签} ”，其中目标物体对应于所提供的 {物体标签}。

候选生成。 对于擦拭，我们在校正后的左图像上使用 SAM3 分割目标物体“黑色墨水”，利用到中心距离滤波去除离群点，并在掩模边界框内生成候选像素的均匀网格。对于绘画，我们估计深度，使用相机内参和外参投影工作空间矩形，并在这些区域内采样候选像素网格。

浮动基座机器人。 对于浮动基座机器人，控制器会持续根据惯性测量单元估算躯干姿态。PD稳定器调节躯干俯仰角和偏航角，并在髋部和腰部发出纠正指令，使机器人在手臂与环境交互时保持直立姿态。在绘画或擦拭等任务中，期望偏航角会根据手部横向位移进行调整：手向侧方移动时，基座会按比例偏航，并受方向相关的限值约束，从而扩大手部的工作空间，同时保持平衡。除了静态站立平衡外，补充视频还展示了基于视觉语言模型的策略也可与运动策略结合，用于擦拭花瓶。