

DOI:10.19651/j.cnki.emt.2519240

基于机械臂模仿学习的高效轨迹优化策略^{*}羊清宇¹ 袁亮^{1,2} 吕凯¹

(1.新疆大学智能制造现代产业学院 乌鲁木齐 830046; 2.上海交通大学南加州大学文化创意产业学院 上海 200240)

摘要: 模仿学习为机械臂在非结构化场景中完成复杂任务提供了强大支持。然而,许多先进的方法仍受输入数据冗余的影响,导致训练效率低下,同时在复杂任务中面临轨迹预测精度受限的问题。为此,提出一种基于关键点提取的示教轨迹优化方法(KPT-O)。通过关键点筛选减少机械臂学习的范围,同时优化关键点分布以提升轨迹预测精度。为了验证其性能,将KPT-O在先进框架下进行训练,并在HelloWorld和RoboTasks数据集上与当前先进方法进行比较。实验结果表明,KPT-O在显著缩短训练时间的同时,也获得了更高的轨迹预测精度。此外,在真实机器人平台上评估了该方法的性能,证明其在涉及位置和方向变化的现实世界机械臂任务中的有效性。

关键词: 模仿学习;机械臂;轨迹优化;关键点

中图分类号: TP242; TP18; TN7 **文献标识码:** A **国家标准学科分类代码:** 540.50

Efficient trajectory optimization strategy for robotic arms via imitation learning

Yang Qingyu¹ Yuan Liang^{1,2} Lyu Kai¹

(1. School of Mechanical Engineering, Xinjiang University, Urumqi 830046, China;

2. USC-SJTU Institute of Cultural and Creative Industry, Shanghai 200240, China)

Abstract: Imitation learning offers a powerful approach for enabling robotic arms to perform complex tasks in unstructured environments. However, many state-of-the-art methods are hindered by redundant input data, leading to inefficient training and limited trajectory prediction accuracy in complex tasks. To address these issues, this paper proposes KPT-O, a method for optimizing demonstrated trajectories by extracting keypoints. The method filters keypoints to streamline the learning data and optimizes their distribution to enhance prediction accuracy. To validate its performance, KPT-O was trained within a state-of-the-art framework and compared against leading methods on the HelloWorld and RoboTasks datasets. The results demonstrate that KPT-O not only significantly reduces training time but also achieves superior trajectory prediction accuracy. Furthermore, evaluations on a physical robot platform confirm the method's effectiveness in real-world robotic arm tasks involving changes in both position and orientation.

Keywords: imitation learning; robotic arm; trajectory optimization; keypoints

0 引言

随着机器人技术的快速发展,机械臂正从传统的流水线环境迈向灵活且非结构化的现实场景^[1-2]。这一转变要求机械臂能够自主学习经验或技能,以适应新任务。模仿学习(imitation learning, IL)作为一种强大的范式,为机械臂赋予运动技能提供了有效途径^[3]。其核心思想是从少量示教轨迹中提取运动特征,并将其泛化至新情境,使机械臂具备较强的适应能力^[4-5]。

目前,超网络^[6](hypernetwork, HN)与神经微分

程^[7](neural ordinary differential equations, NODE)结合的方法在模仿学习领域展现出优异性能,该框架不仅能够捕捉轨迹的细粒度变化,还能实现连续时间的轨迹预测,在高维复杂任务中展现出良好的泛化能力^[6]。然而,该方法直接训练原始示教数据,在复杂任务和长时间步轨迹中易产生复合误差,导致其在实际应用中仍面临计算开销大和轨迹预测精度受限的两大挑战^[8-9]。

在模仿学习领域,轨迹数据的高效表示与预测精度是两大核心议题^[10-12]。在轨迹压缩方面,Douglas-Peucker^[13]通过递归地去除轨迹中对整体形状影响较小的点,实现了

收稿日期:2025-07-01

* 基金项目:国家自然科学基金(52275003)、“天山英才”培养计划(2023TSYCLJ0052)项目资助

高效的轨迹简化。DiffTORI (differentiable trajectory optimization for deep reinforcement and imitation learning)^[14]通过可微分轨迹优化过程,最大化任务性能,解决了先前模型中存在的“目标不匹配”问题,在多任务中表现出色。在提升预测精度方面,为解决复合误差(compounding errors)难题,HYDRA (hybrid robot actions)^[15]算法采用混合高层和低层动作,使其能够动态适应复杂环境并减少误差积。Perceiver-Actor^[16]算法采用Transformer^[17]模型捕捉轨迹中的复杂关系,在处理高维输入和捕捉非线性轨迹特征方面具有优势。尽管现有方法在特定任务中表现良好,但在处理高维、复杂的机器人任务时,仍普遍面临着压缩率与预测精度难以兼顾、计算开销大以及泛化能力不足的挑战。

针对上述问题,本文作者提出了一种基于关键点提取的示教轨迹优化方法(KPT-O),通过关键点筛选减少机械臂学习的范围。该方法引入全局压缩策略和自适应误差优化策略对原始示教轨迹进行预处理,优化后的关键点序列替代原始示教轨迹,作为输入在HN+NODE框架下进行训练和轨迹预测。KPT-O旨在通过更少但更具代表性的轨迹点学习动态模型参数,在实现轨迹全局压缩的同时保留局部精细特征,从而提升轨迹学习的效率和精度。

1 策略算法框架

1.1 轨迹压缩目标

在机械臂模仿学习中,示教轨迹通常包含大量的冗余数据,例如缓慢移动或静止的阶段,而关键轨迹信息往往由少量重要状态构成。直接使用完整轨迹进行训练不仅增加了计算开销,还可能导致策略在无关紧要的轨迹点上过拟合^[18]。因此,目标变成了通过轨迹压缩的方法,从示教轨迹 $\tau = \{x_0, x_1, \dots, x_T\}$ 中提取一个最优的关键路点集 $W = \{w_0, w_1, \dots, w_L\}$,以简化机械臂的学习范围,如图1。其中 $x_t \in R^d$ 代表机器人在时间步 t 处的状态, $L \ll T$,在保证轨迹误差不超过设定阈值的情况下,最大程度地减少数据规模,同时保留轨迹的全局结构和关键运动信息。

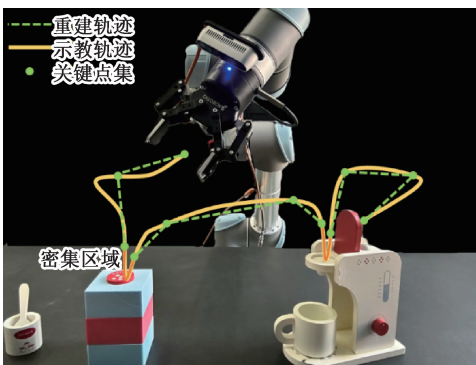


图1 轨迹压缩简图

Fig.1 Schematic diagram of trajectory compression

为了衡量轨迹压缩的质量,定义轨迹的重建误差(reconstruction loss):

$$L(\hat{\tau}, \tau) = \max_{x \in \tau} \min_{\hat{x} \in \hat{\tau}} l(x, \hat{x}) \quad (1)$$

其中, $\hat{\tau} = f(W)$ 为通过插值函数 f 还原的轨迹, $l(x, \hat{x})$ 为轨迹点之间的欧几里得距离:

$$l(x, \hat{x}) = \|x - \hat{x}\|_2 \quad (2)$$

该误差衡量了原始轨迹与压缩轨迹之间的最大偏差,希望通过优化 W 使得:

$$\min_w |W| \text{ s.t. } L(f(W), \tau) \leq \eta \quad (3)$$

其中, η 是允许的最大重建误差,它的取值直接决定了轨迹压缩率与重建精度之间的权衡。一个过小的值会保留过多细节,导致压缩率低;一个过大的值则会丢失关键特征,影响轨迹精度。在本文中, η 的选择是基于对示教轨迹数据特征的分析。实验证明,当 η 在 $[0.2, 0.4]$ 区间内时,模型能够在保持较高压缩率的同时,取得良好的预测精度。因此,在后续所有实验中,为HelloWorld和RoboTasks数据集统一设置 $\eta = 0.3$ 。

为实现这一目标,提出了全局压缩策略和自适应误差优化策略,分别从数据压缩和轨迹精度优化的角度进行改进。

1.2 全局压缩策略

全局压缩策略的核心目标是在尽可能减少数据规模的前提下,保留轨迹的全局结构。相比于传统的轨迹压缩方法,如均匀采样、滑动窗口筛选等,该方法在于引入全局误差控制和均匀优化,确保压缩后的轨迹在误差范围内保持全局特性,同时保证轨迹点的合理分布。

1) 基于全局误差约束的关键点选取

为实现轨迹的全局压缩,采用了一种基于全局误差约束的迭代优化方法。该方法首先选取轨迹的关键端点作为初始路点集,并基于全局误差约束进行逐步优化,以确保轨迹的整体形态不失真。具体而言,算法首先从示教轨迹 $\tau = \{x_0, x_1, \dots, x_T\}$ 中提取起点 x_0 和终点 x_T 作为初始路点集 $W = \{x_0, x_T\}$ 。随后,在基于贪心策略的每次迭代中,计算轨迹中当前误差最大的点 w^* ,如图2所示,即该点到当前路点集 W 所重建轨迹 $\hat{\tau}$ 的最小距离最大的位置:

$$w^* = \operatorname{argmax}_{x \in \tau} \min_{w \in W} l(x, w) \quad (4)$$

其中, $l(x, w)$ 是点 x 到当前关键点集合 W 的最小欧几里得距离:

$$l(x, w) = \|x - w\|_2 \quad (5)$$

直观上,这意味着找到当前压缩轨迹偏差最大的位置。如果该误差超过设定的阈值 η ,则将 w^* 加入关键点集合 $W = W \cup \{w^*\}$,此过程不断重复,直到所有轨迹点的误差均满足 η 约束:

$$L(f(W), \tau) \leq \eta \quad (6)$$

2) 轨迹均匀优化

在初步轨迹压缩完成后,计算轨迹的累积长度,并在轨

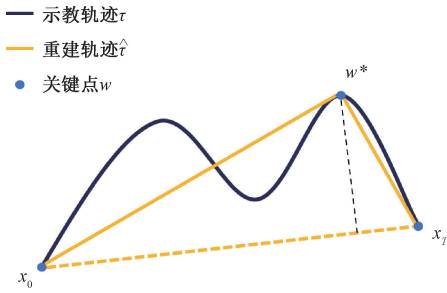


图 2 基于全局压缩的路点选取可视化

Fig.2 Visualization of waypoint selection based on global compression

迹上重新均匀分布路点,确保压缩后的轨迹保持全局特性:

$$d_{avg} = \frac{\sum_{i=1}^{|W|-1} l(\omega_i - \omega_{i-1})}{|W| - 1} \quad (7)$$

通过重新采样,使得所有相邻路点的间隔尽可能接近 d_{avg} ,从而提高轨迹的均匀性,并减少由于压缩导致的局部过拟合问题,如图 3 所示。

1.3 自适应误差优化策略

自适应误差优化策略旨在针对轨迹的局部区域动态调整关键点密度,确保轨迹的局部误差最小化。不同于全局压缩策略,该方法优先考虑轨迹的局部非线性特征,在复杂

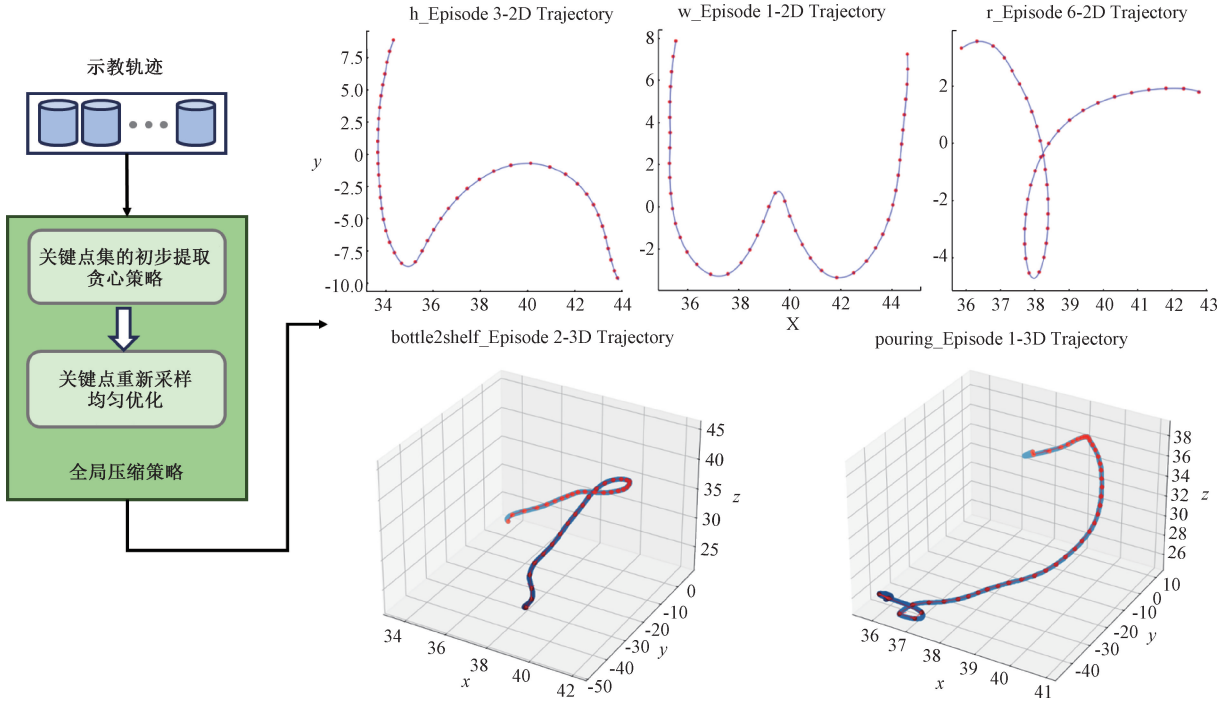


图 3 全局压缩策略架构图

Fig.3 Architecture diagram of global compression strategy

轨迹段增加路点密度,而在平滑段减少路点,使得在关键区域保留更多信息。

1) 局部误差驱动的路点分配

该方法首先基于现有关键路点集 W 采用线性插值方法对轨迹进行重建。给定两个相邻关键路点 ω_i 和 ω_{i+1} , 以及它们对应的时间索引 t_i 和 t_{i+1} , 轨迹点 x_i 的重建轨迹点 \hat{x}_i 由线性插值计算:

$$\hat{x}_i = \frac{t_{i+1} - t}{t_{i+1} - t_i} \omega_i + \frac{t - t_i}{t_{i+1} - t_i} \omega_{i+1} \quad (8)$$

在重建轨迹的基础上,定义局部误差 $E(x_i)$ 以衡量原始轨迹点 x_i 与重建轨迹点 \hat{x}_i 之间的偏差:

$$E(x_i) = \|x_i - \hat{x}_i\|_2 \quad (9)$$

这里引入局部误差阈值 η_{local} , 用于优化轨迹中非线性程度高的片段,确保局部细节不被丢失。为捕捉细微变

化,它应比全局阈值 η 更为严格。故设定 $\eta_{local} = \alpha \cdot \eta$, 其中 $0 < \alpha < 1$, 该设置确保了局部优化的约束始终在全局约束的范围之内。实验验证,当 $\alpha = 0.5$ 时,算法能够在复杂区域增加足够的关键点,同时避免在平滑区域引入不必要的点。

当 $E(x_i)$ 超过 η_{local} 时,说明当前关键点集 W 无法充分拟合该区域的轨迹变化,需动态增加新的关键点以优化轨迹重建质量。因此,在当前位置 x_i 处插入新的路点:

$$W = W \cup \{x_i\}, E(x_i) > \eta_{local} \quad (10)$$

这一机制确保了关键点在复杂轨迹区域的高密度分布,同时在平滑区域保持低密度分布,实现了对轨迹形态的自适应优化。基于局部误差的路点选取在图 4 中可可视化。

2) 动态规划优化关键点序列

在通过局部误差驱动初步增加路点集后,采用动态规

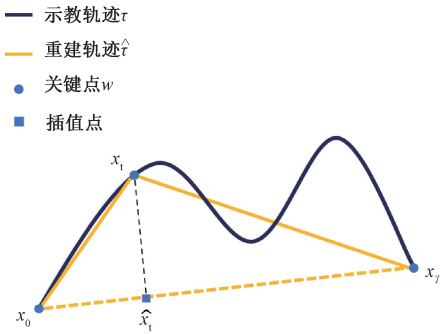


图4 基于局部误差的路点选取可视化

Fig. 4 Visualization of waypoint selection based on local errors

划(dynamic programming, DP)来进一步提炼,以找到满足误差约束 η 的最短关键点子序列。将此问题建模为一个在轨迹点上的最短路径寻优问题:

$$\min |W|, L(f(W), \tau) \leq \eta \quad (11)$$

具体而言,将其分解为一系列最优子问题。定义 $DP[i]$ 为从轨迹起点 x_0 到轨迹点 x_i 的最优关键点子序列的点的数量,目标是求解 $DP[T]$ 。

$DP[i]$ 的值可以通过尝试将 x_i 作为当前子序列的最后一个关键点,并向前寻找一个最优的前置关键点 x_j 来确定,其中 $0 \leq j < i$ 。如果轨迹段 $\tau_{j \rightarrow i}$ 仅用两个端点 $\{x_j, x_i\}$ 进行重建,且其误差 $L(f(\{x_j, x_i\}), \tau_{j \rightarrow i})$ 不超过阈值 η , 那么 x_j 就是一个有效的前置关键点。

因此, $DP[i]$ 的计算方式为,在所有有效的 j 中,找到那个能使 $DP[j] + 1$ 最小的值:

$$DP[i] = \min_{0 \leq j < i, L(f(\{x_j, x_i\}), \tau_{j \rightarrow i}) \leq \eta} \{DP[j]\} + 1 \quad (12)$$

通过从 $i = 1$ 到 T 的顺序迭代计算,最终得到 $DP[T]$, 即最优解包含的关键点数量。为了回溯得到最优的关键点集合 W , 需要在计算过程中记录下每一个 $DP[i]$ 是由哪个前置索引 j 转移而来的。

这种动态规划方法确保了能在多项式时间内找到满足误差约束的全局最优最简关键点序列,从而有效降低轨迹冗余,为后续的模仿学习任务提供高质量的输入。具体的算法流程在算法1中详细描述。

该方法能够有效降低局部误差,使轨迹重建更加精准,特别是在复杂轨迹段能够大幅提高预测精度,如图5所示。

1.4 基于 KPT-O 的轨迹压缩

在传统的 HN+NODE 算法框架下,示教轨迹直接用于训练,尽管该方法能够动态生成参数并进行时间序列建模,但仍然面临计算开销大、训练效率低以及复杂轨迹区域精度受限的问题。因此,全局压缩策略与自适应误差优化策略可以作为 HN+NODE 的前处理模块,如图6所示,有效减少计算负担,并提升预测精度。

综上,将全局压缩策略与自适应误差优化策略相结合。具体而言,首先利用全局压缩策略对原始示教轨迹进行初步压缩,从而去除大部分冗余信息,有效减少了数

算法1:动态规划算法

输入: $\tau = \{x_0, x_1, \dots, x_T\}, \eta$

输出: $W_{optimal}$

```

1: 初始化:  $DP[0, \dots, T], Prev[0, \dots, T], DP[0] \leftarrow 1$ 
2: for  $i \leftarrow 1$  to  $T$ 
3:      $DP[i] \leftarrow \infty$ 
4:  $Prev[i] \leftarrow -1$ 
5: end for
6: for  $i \leftarrow 1$  to  $T$ 
7:     for  $j \leftarrow 0$  to  $i-1$ 
8:          $\tau_{j \rightarrow i} \leftarrow \{x_j, x_{j+1}, \dots, x_i\}$ 
9:          $W_{segment} \leftarrow \{x_j, x_i\}$ 
10:         $error \leftarrow L(f(W_{segment}), \tau_{j \rightarrow i})$ 
11:        if  $error \leq \eta$  then
12:            if  $DP[j] + 1 < DP[i]$  then
13:                 $DP[i] \leftarrow DP[j] + 1$ 
14:                 $Prev[i] \leftarrow j$ 
15:            end if
16:        end if
17:    end for
18: end for
19:  $W_{optimal} \leftarrow \emptyset$ 
20:  $k \leftarrow T$ 
21: while  $k \neq -1$ 
22:      $W_{optimal}, insert\_at\_front(x_k)$ 
23:      $k \leftarrow Prev[k]$ 
24: end while
25: return  $W_{optimal}$ 

```

据规模。随后在压缩后的关键点序列上应用自适应误差优化策略,通过动态调整局部路点密度,进一步提升预测精度。最终,优化后的关键点序列代替原始示教轨迹,作为任务嵌入向量输入超网络(HN),HN根据不同向量输出相应参数,NODE作为目标网络使用这些参数处理当前任务的输入,并生成最终的预测轨迹。这样一来,HN能够基于更少但更具代表性的轨迹点学习动态模型参数,而NODE则利用优化后的轨迹信息进行时间序列建模。该方法在理论上实现了数据压缩与预测精度之间的最佳平衡,更加适用于高维复杂任务下的轨迹学习问题。

2 数据集与评估指标

2.1 数据集

1) HelloWorld 数据集

HelloWorld 数据集包含7个任务,分别为小写字母 h、

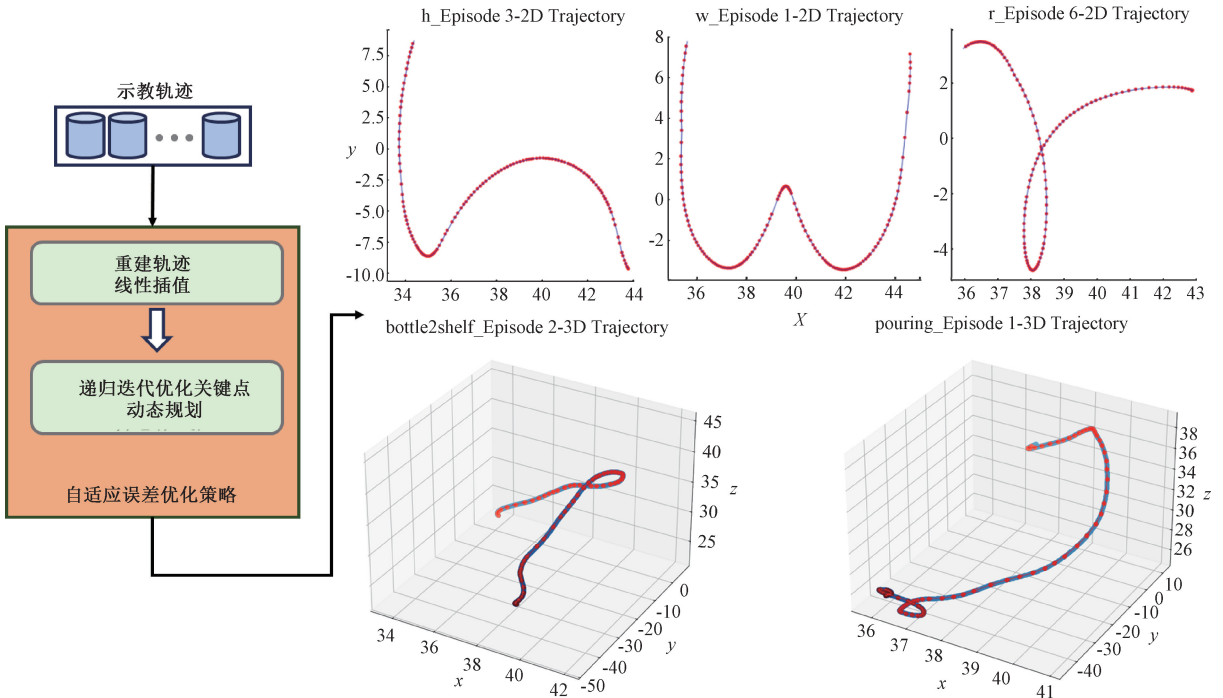


图 5 自适应误差优化策略架构图

Fig. 5 Architecture diagram of adaptive error optimization strategy

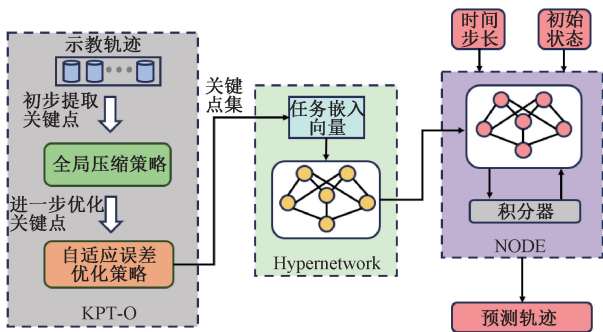


图 6 KPT-O 整体框架图

Fig. 6 KPT-O overall framework diagram

e、l、o、w、r 和 d 的二维轨迹。该数据集通过人类用户的运动示教记录末端执行器的轨迹。每个任务包含 8 条演示轨迹，每条轨迹由 1 000 个二维点序列构成。HelloWorld 数据集的特点是轨迹结构复杂，包含大量回环，这对预测精度和模型的泛化能力提出了较高的要求。

2) RoboTasks 数据集

RoboTasks 数据集记录了机械臂执行四个操作任务的三维轨迹：(1) 开箱任务：打开盒盖；(2) 放瓶任务：将竖直状态的瓶子放置在水平位置的架子上；(3) 叠盘任务：将竖直状态的盘子放到高台上，同时避开障碍物；(4) 倒豆任务：将咖啡豆从杯中倒入容器中。每个任务包含 9 条演示轨迹，每条轨迹由 1 000 个时间步组成。与 HelloWorld 数据集相比，RoboTasks 任务复杂度更高，轨迹维度更大，同时演示轨迹之间存在更显著的变异性。

3) 真实场景任务设置

为了验证所提出方法在真实场景中的有效性，设计了两个不同的实际任务：(1) 抓取任务：机械臂抓取咖啡豆并将其放入咖啡机，随后合上盖子制作咖啡。(2) 非抓取任务：机械臂通过末端执行器按压获取消毒液。每个任务通过人类用户的运动示教记录轨迹，并基于这些数据评估所提出方法在真实操作环境中的表现。

2.2 评估指标

为了衡量不同方法的计算效率，以不同算法框架的训练时间作为评估指标，并比较各方法在不同任务上的训练时长。此外，使用 DTW (dynamic time warping) 误差、Frechet Distance 误差和 Swept Area 误差来量化生成轨迹与真实轨迹之间的相似程度，误差越小表示轨迹预测精度越高。最后，根据 DTW 误差设置阈值，将误差低于阈值的预测轨迹视为准确，并统计准确预测轨迹的比例，计算模型的总体准确率 (ACC)。

由于每个任务包含多个真实演示，需要首先计算所有演示对之间的 DTW 误差。对于 HelloWorld 数据集，选取其最大误差值，并乘以 3 作为阈值。该乘法因子是基于经验设定的，以允许一定的误差范围，使得视觉上与演示轨迹形状一致的预测轨迹被认为是准确的。最终得到 HelloWorld 数据集的 DTW 阈值为 1 821。对于 RoboTasks 数据集，同样计算所有演示对之间的 DTW 误差，由于真实的示教轨迹已经非常多样化，所以直接取其最大值作为误差阈值，该阈值设定为 7 131。

3 实验结果及分析

3.1 HelloWorld 数据集

1) 仿真二维任务定量分析

在 HelloWorld 数据集中,依次训练模型完成 7 个任务,并分别评估基线模型(HN+NODE)和全局压缩策略、自适应误差优化策略及 KPT-O 在基线框架下训练后的表现。图 7 显示了不同模型完成 7 个任务时分别所需的训练时间(越低越

好),X 轴显示当前任务。条状表示的是每个任务 5 个独立种子的平均训练时长。实验结果表明,全局压缩策略通过去除轨迹中大量冗余信息,大幅减少了训练数据量,其训练时间相较基线模型缩短约 43%。相比之下,自适应误差优化策略虽然能够提高轨迹精度,但在一些包含回环结构的任务中,因需要更多迭代优化,训练时间甚至可能超过基线模型。KPT-O 在全局压缩的基础上优化了关键点分布,平均训练时间较基线减少 33%,但略高于全局压缩策略。

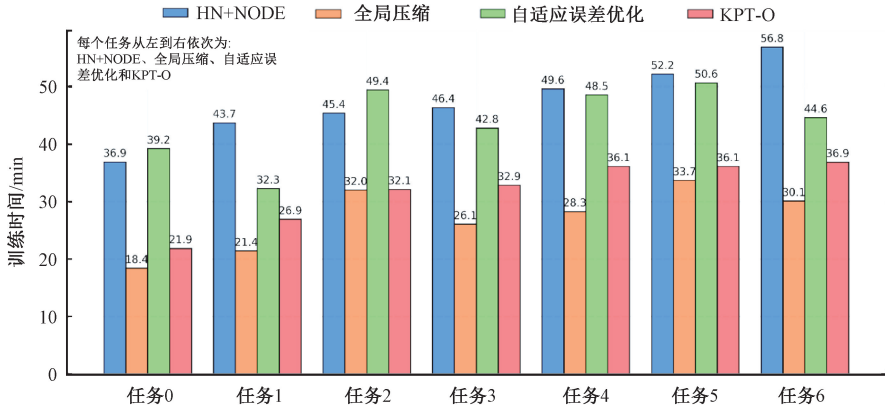


图 7 HelloWorld 数据集训练时间
Fig. 7 Training time of HelloWorld dataset

在轨迹预测精度方面,使用 DTW 误差、Frechet Distance 误差和 Swept Area 误差衡量生成轨迹的质量。图 8(上方)展示了 HelloWorld 数据集在所有任务的误差分布情况(越低越好),图 8(下方)是缩放表现良好策略的误差分布情况。X 轴显示当前任务。在学习一个任务后,所有当前和以前的任务都将被评估。线条显示每个任务 5 个独立种子的中位数,阴影区域表示 5 个独立种子误差的下四分位数和上四分位数。从图中可以观察到,全局压

缩策略的 DTW 误差与基线方法相近,这表明该策略在大幅减少数据规模的同时,仍能将其精度损失控制在可接受的范围内。然而,由于其采用统一的压缩方式,忽略了轨迹局部非线性变化,在某些复杂任务上的预测误差甚至高于基线框架。相比之下,自适应误差优化策略和 KPT-O 的 DTW 误差均显著低于基线方法。其中,自适应误差优化策略通过动态调整关键点密度,确保局部误差最小化,其 Frechet Distance 误差和 Swept Area 误差均显著优于其他

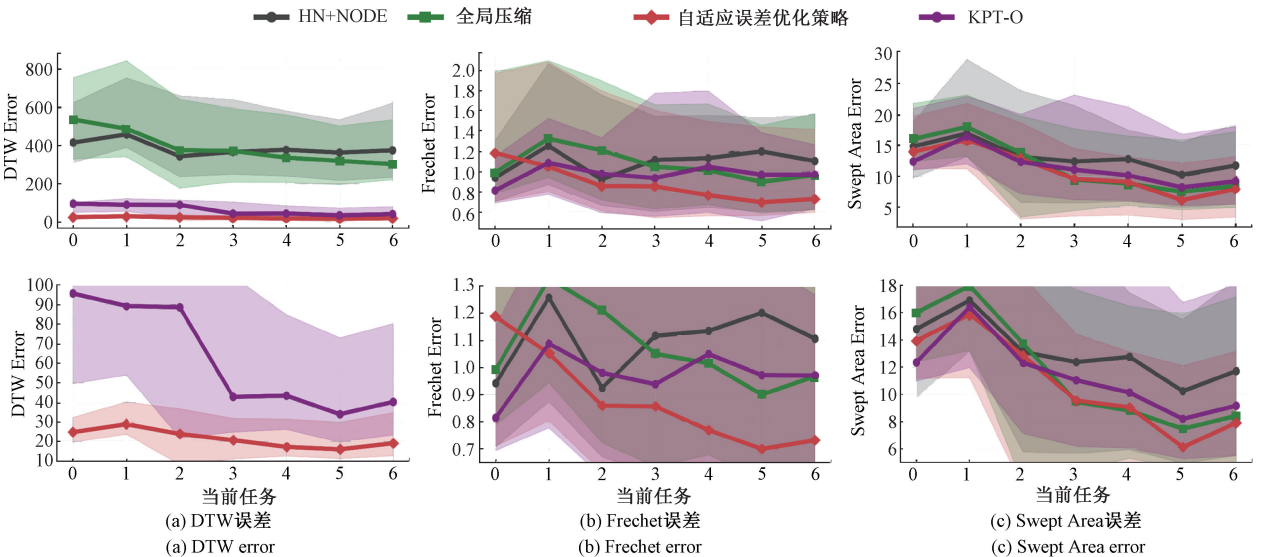


图 8 HelloWorld 数据集预测轨迹的误差

Fig. 8 Error of trajectory prediction in HelloWorld dataset

框架,表明在该策略下生成的轨迹在整体形状和局部特征上更接近真实轨迹。KPT-O 在全局特征与局部精细度之间达到了良好的平衡,其误差略高于自适应误差优化策略,但依然远优于全局压缩策略和基线框架,展现出更稳定的性能。

此外,在设定的 DTW 阈值(1 821)下,自适应误差优化策略和 KPT-O 的轨迹准确率(ACC)均达到 100%,显著高于基线框架的 97%和全局压缩策略的 95%,如表 1 所示。这一结果表明,在二维任务中,自适应误差优化策略和 KPT-O 均能有效提升预测精度,且 KPT-O 兼顾了数据压缩效率,适用于复杂二维任务。

2) 仿真二维任务定性分析

图 9 展示了在所有 HelloWorld 数据集任务上顺序训练之后,基线框架、KPT-O 和自适应误差优化策略针对过去所有任务预测轨迹的实例。从图中可以观察到,HN 在

表 1 HelloWorld 数据集轨迹预测准确率指标

Table 1 Trajectory prediction accuracy indicators for HelloWorld dataset

算法框架	准确率(ACC)
HN+NODE	0.97
全局压缩策略	0.95
自适应误差优化策略	1.00
KPT-O	1.00

学习具有回环结构的字母 e、o 和 d 时,其预测轨迹误差较大,存在明显偏差。而 KPT-O 在此类复杂轨迹上表现优异,能够有效缓解该问题,使得预测轨迹更接近真实轨迹。此外,自适应误差优化策略在精细化局部关键点分布方面更具优势,其生成的轨迹精度最高,与原始演示轨迹的匹配度最佳。

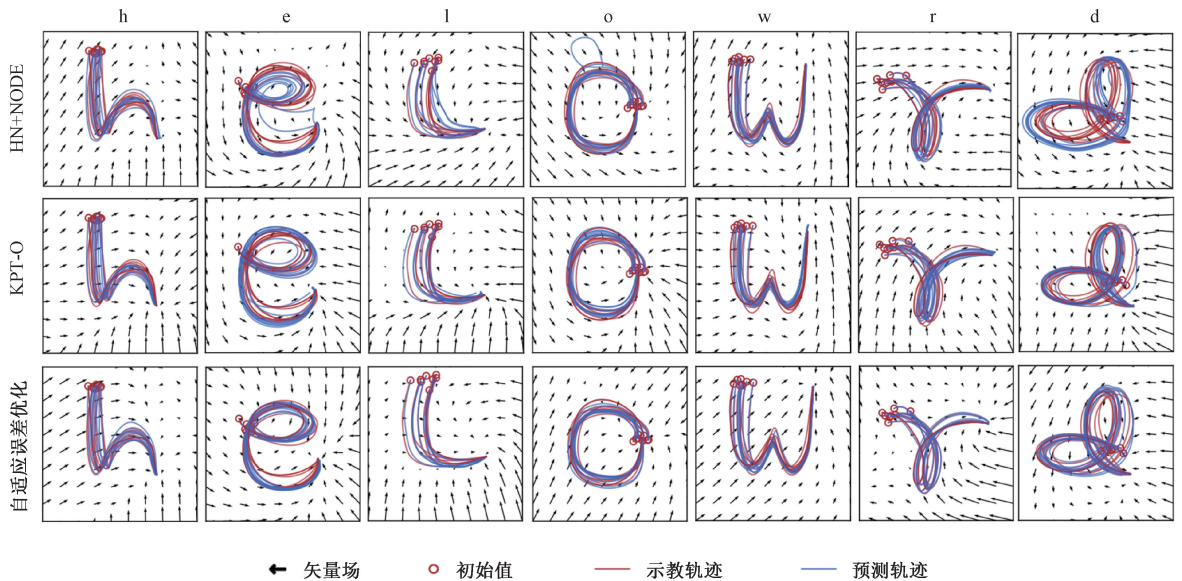


图 9 HelloWorld 数据集预测轨迹的可视化

Fig. 9 Visualization of trajectory prediction in HelloWorld

3.2 RoboTasks 数据集

1) 仿真三维任务定量分析

在 RoboTasks 数据集中,依次训练模型完成 4 个任务,并评估各方法在高维复杂轨迹任务中的表现。图 10 展示了不同模型完成 4 个任务所需的训练时间,X 轴显示当前任务,条状表示的是每个任务 5 个独立种子的平均训练时长(越低越好)。与 HelloWorld 结果类似,全局压缩策略显著减少了轨迹冗余点数,使训练效率最高,平均训练时间比基线框架减少 52%。KPT-O 进一步优化了关键点分布,训练时间较基线减少 39%,略高于全局压缩策略,但优于自适应误差优化策略。

在预测精度方面,图 11 和表 2 显示,自适应误差优化策略在所有精度指标上表现最佳,其 ACC 达到 99%,

Frechet Distance 误差和 Swept Area 误差也显著优于其他方法,表明该策略能够精确捕捉复杂任务中的局部特征。然而,该方法的训练效率较低,难以满足对实时性要求较高的任务需求。KPT-O 在全局和局部特征捕捉上取得良好平衡,各项误差略高于自适应误差优化策略,但远优于基线框架和全局压缩策略,ACC 达到 97%,明显优于基线框架的 88%和全局压缩策略的 83%。这一结果进一步验证了 KPT-O 在应对具有更高的复杂性和维度的任务中的整体性能优势。

2) 仿真三维任务定性分析

图 12 展示了在所有 RoboTasks 数据集任务上顺序训练后,基线框架、KPT-O 和自适应误差优化策略针对过去所有任务预测轨迹的实例。从图中可以看出,HN 在学习

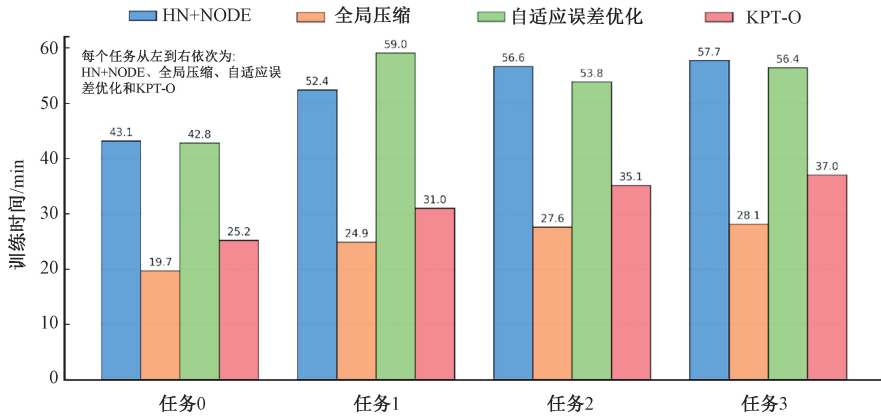


图 10 RoboTasks 数据集训练时间

Fig. 10 Training time of RoboTasks dataset

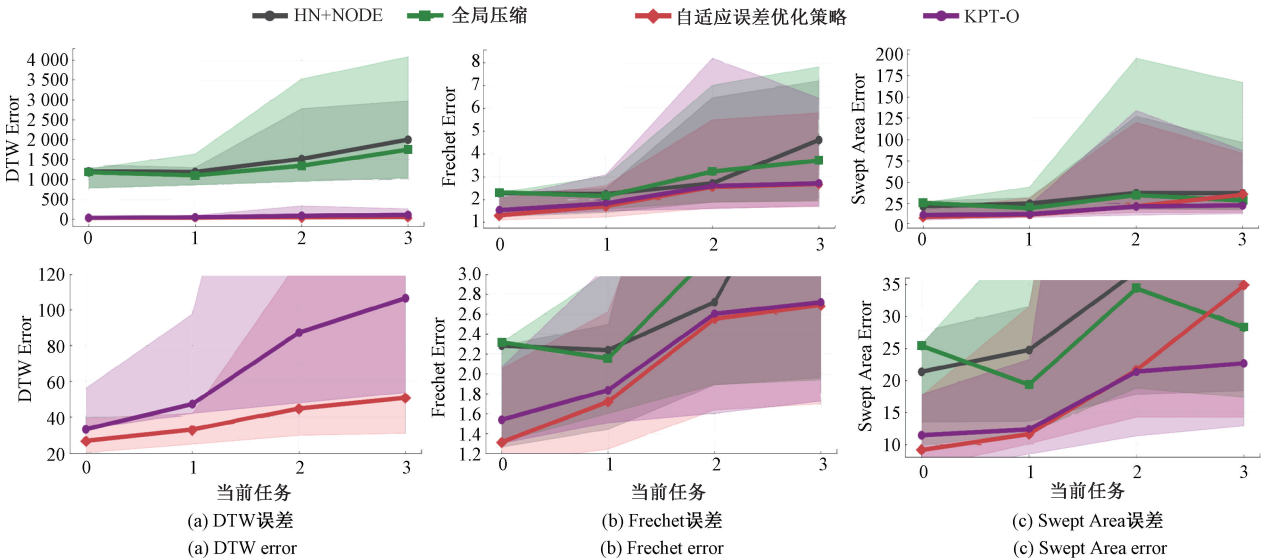


图 11 RoboTasks 数据集预测轨迹的误差

Fig. 11 Error of trajectory prediction in RoboTasks dataset

表 2 RoboTasks 数据集轨迹预测准确率指标

Table 2 Trajectory prediction accuracy indicators for

RoboTasks dataset

算法框架	准确率(ACC)
HN+NODE	0.88
全局压缩策略	0.83
自适应误差优化策略	0.99
KPT-O	0.97

高维度、高复杂度的轨迹时,预测误差明显增大,部分任务甚至出现轨迹失真。相比之下,KPT-O在全局压缩的基础上进一步优化了关键点分布,使得预测轨迹在整体形态和关键区域细节上均保持较高一致性,有效提升了轨迹重建质量。而自适应误差优化策略凭借对复杂区域的精细建模,使得预测轨迹更加精准,与真实轨迹几乎完全一致。

3.3 真机实验验证

在人机协作中,必须将模仿学习方法学习到的轨迹应用到机器人身上,以便机器人的运动能够实时适应情况。为了评估本文提出的方法在真实场景中的性能,特在真实的机器人平台 UR5 上进行了大量实验,实验场景如图 13 所示。该平台是一个六自由度的机械臂,平台末端配有两指平行夹爪。

1) 动态环境下的鲁棒性与泛化能力测试

该任务要求机械臂精准按压一个消毒液瓶的喷嘴。为了严格评估算法的泛化能力,在实验中引入了显著的环境变化。具体而言,在每次重复实验中,在 30 cm × 30 cm 的工作区域内随机改变消毒液瓶的初始位置,并引入 ±5 cm 的高度扰动。

该任务进行了 30 次重复实验,其中包含了不同的目标位置、高度和干扰物布局。实验结果表明,KPT-O 生成的轨迹能够实时适应这些变化,成功完成了 28 次任务,成

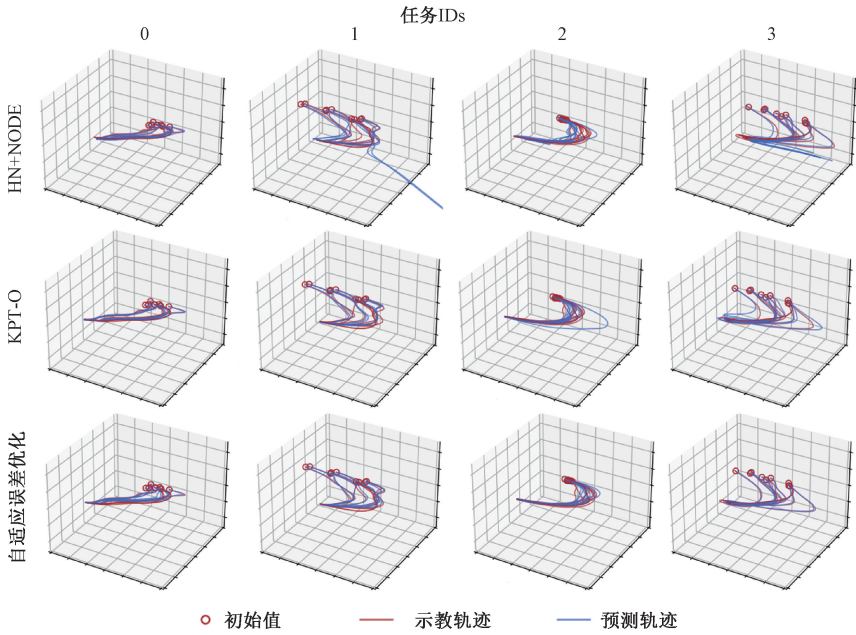


图 12 RoboTasks 数据集预测轨迹的可视化

Fig. 12 Visualization of trajectory prediction in RoboTasks dataset

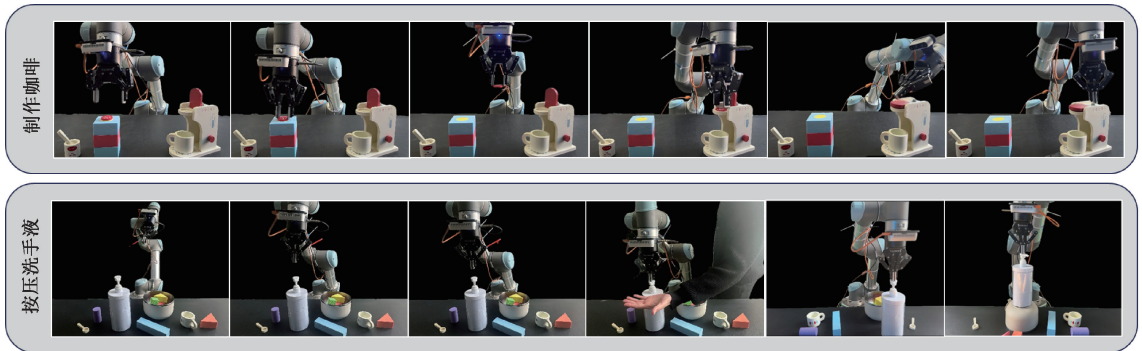


图 13 现实世界抓取与非抓取任务

Fig. 13 Real world grasping and non-grasping tasks

功率达到了 93.3%。这一量化结果有力地证明了该方法在动态、非结构化场景下的高度鲁棒性与可靠的泛化能力。

2) 复杂长序列操作任务的有效性验证

该任务旨在验证算法在执行复杂、多步骤、长序列操作时的有效性。任务流程为：(1)机械臂从桌面精准抓取一枚咖啡胶囊；(2)将其稳定运送至咖啡机入口并精确放置；(3)闭合咖啡机顶盖并按下启动按钮。此任务链条长、对多个不同阶段的定位精度要求高，是对算法综合性能的严峻考验。

该任务执行了 25 次完整的流程尝试。实验结果显示，KPT-O 能够生成平滑且精准的全流程轨迹，成功完成了 21 次，成功率高达 84%。其中单次失败是由于初始抓取姿态的微小偏差导致。该实验的定量成功率表明，该方法不仅能处理简单的点到点运动，更能胜任需要细致操作

的复杂现实世界任务。

3) 真实场景实验总结

KPT-O 在真实场景中的实验结果表明，它不仅能在模拟任务中展现优异性能，同时也具备处理动态、复杂、长序列实际操作任务的潜力。

4 结 论

本文作者提出了一种基于关键点提取的示教轨迹优化方法(KPT-O)，以解决传统 HN+NODE 方法在训练时间和预测精度方面的瓶颈。我们在 HelloWorld、RoboTasks 公开数据集及真实机械臂操作任务上对我们的方法进行验证。实验结果显示，KPT-O 在保证全局轨迹压缩的同时优化局部关键点分布，预测精度和训练效率明显优于基线模型。此外，在真实机器人平台上的评估进一步验证了 KPT-O 的可行性，其生成的轨迹在复杂操作

任务中保持较高的稳定性和执行成功率。综上所述,KPT-O在提高模仿学习训练效率的同时,兼顾了轨迹预测的精度,为高效、精准的机械臂模仿学习提供了一种新的解决方案。

参考文献

- [1] 游达章,杨智杰,张业鹏.基于改进RRT-Connect算法的机械臂运动规划[J].电子测量技术,2023,46(8):112-119.
YOU D ZH, YANG ZH J, ZHANG Y P. Motion planning for robotic arm based on improved RRT-Connect algorithm [J]. Electronic Measurement Technology, 2023, 46(8):112-119.
- [2] WIJAYATHUNGA L, RASSAU A, CHAI D. Challenges and solutions for autonomous ground robot scene understanding and navigation in unstructured outdoor environments: A review [J]. Applied Sciences, 2023, 13(17): 9877.
- [3] 朱应钊,李嫒,胡颖茂.模仿学习在机器人领域的应用进展[J].广东通信技术,2020,40(10):85-90.
ZHU Y ZH, LI M, HU Y M. Application progress of imitation learning in the field of robotics [J]. Guangdong Communication Technology, 2020, 40(10): 85-90.
- [4] 张超,白文松,杜歆,等.模仿学习综述:传统与新进展[J].中国图象图形学报,2023,28(6):1585-1607.
ZHANG CH, BAI W S, DU X, et al. A survey of imitation learning: Traditional and new advances[J]. Journal of Image and Graphics, 2023, 28(6): 1585-1607.
- [5] FANG B, JIA S, GUO D, et al. Survey of imitation learning for robotic manipulation [J]. International Journal of Intelligent Robotics and Applications, 2019, 3(4): 362-369.
- [6] AUDDY S, HOLLENSTEIN J, SAVERIANO M, et al. Continual learning from demonstration of robotics skills[J]. Robotics and Autonomous Systems, 2023, 165: 104427.
- [7] CHEN R T Q, RUBANOVA Y, BETTENCOURT J, et al. Neural ordinary differential equations [J]. Advances in Neural Information Processing Systems, 2018, 31: 6571-6580.
- [8] 赵惠,李庆党,张明月.基于改进RRT算法的机械臂路径规划方法[J].电子测量技术,2021,44(16):45-49.
ZHAO H, LI Q D, ZHANG M Y. Path planning method for robotic arm based on improved RRT algorithm [J]. Electronic Measurement Technology, 2021, 44(16):45-49.
- [9] ZHANG H, ZHAO W. A memory-efficient neural ordinary differential equation framework based on high-level adjoint differentiation [J]. IEEE Transactions on Artificial Intelligence, 2022, 5(3): 1110-1120.
- [10] SPENCER J, CHOUDHURY S, VENKATRAMAN A, et al. Feedback in imitation learning: The three regimes of covariate shift [J]. ArXiv preprint arXiv:2102.02872, 2021.
- [11] 黄艳龙,徐德,谭民.机器人运动轨迹的模仿学习综述[J].自动化学报,2022,48(2):315-334.
HUANG Y L, XU D, TAN M. A survey of imitation learning for robot motion trajectory [J]. Acta Automatica Sinica, 2022, 48(2): 315-334.
- [12] 迟明善,姚玉峰,刘亚欣.模仿学习示教轨迹自动分割方法的研究进展[J].控制与决策,2019,34(7):1345-1354.
CHI M SH, YAO Y F, LIU Y X. Research progress on automatic segmentation methods for demonstration trajectories in imitation learning [J]. Control and Decision, 2019, 34(7): 1345-1354.
- [13] CHEN H, CHEN X. A trajectory ensemble-compression algorithm based on finite element method [J]. ISPRS International Journal of Geo-Information, 2021, 10(5): 334.
- [14] WAN W, WANG Z, WANG Y, et al. DiffTori: Differentiable trajectory optimization for deep reinforcement and imitation learning [J]. Advances in Neural Information Processing Systems, 2024, 37: 109430-109459.
- [15] BELKHALE S, CUI Y, SADIGH D. Hydra: Hybrid robot actions for imitation learning [C]. Conference on Robot Learning. PMLR, 2023: 2113-2133.
- [16] SHRIDHAR M, MANUELLI L, FOX D. Perceiver-actor: A multi-task transformer for robotic manipulation [C]. Conference on Robot Learning. PMLR, 2023: 785-799.
- [17] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [J]. Advances in Neural Information Processing Systems, 2017, 30: 5998-6008.
- [18] XIE M, LI AN Q, WYK K V, et al. Imitation learning via simultaneous optimization of policies and auxiliary trajectories [J]. ArXiv preprint arXiv:2105.03019, 2021.

作者简介

羊清宇,硕士研究生,主要研究方向为基机械臂的模仿学习。

E-mail:3268179580@qq.com

袁亮(通信作者),博士,教授,主要研究方向为智能机器人技术、机器视觉与图像处理、数字孪生及工业物联网。

E-mail:107552304354@stu.xju.edu.cn