基于轨迹预测分歧的机械臂精细操作自适应动作分块方法

郑姗[1]，杨子仪[1]，刘宜坤[1]，周翰[2]，王刚[3]，吴清华[3]，沈甜雨[1]，李志伟[1]
（1. 北京化工大学，北京 100029；
2. 北京邮电大学，北京 100876；
3. 北京具身智能科技有限公司，北京 100085）

摘　要：针对具身操作任务中固定动作分块策略难以同时兼顾稳定阶段动作连续性与关键阶段执行响应能力的问题，提出一种基于预测分歧的自适应动作分块调度方法。在保持模型结构与训练过程不变的条件下，通过在推理阶段动态调整动作缓存刷新时写入并连续执行的动作步数，构建“固定长度预测-自适应长度执行”的推理机制。基于相邻预测块在时间重叠区间内的差异，定义轨迹级预测分歧指标，并结合时间平滑与分段阈值规则，实现执行块长的在线切换。实验结果表明，所提方法在任务成功率和轨迹误差方面取得最优或接近最优结果，并在执行平滑性方面显著优于固定小块长策略，从而在执行精度、关键阶段执行响应能力与轨迹稳定性之间实现了更优综合平衡；同时，该方法在视觉扰动与几何扰动条件下表现出较好的鲁棒性。进一步的消融实验与超参数分析表明，轨迹级预测分歧、时间平滑和分段阈值调度共同提升了执行块长调度的稳定性与阶段适应性。该方法能够较好地表征任务阶段变化，并提升不同阶段下的执行适应性。

关键词：具身智能；机械臂操作；自适应动作分块
中图分类号：TP39
文献标志码：A
doi: 10.11959/j.issn.2096-6652.ZNKX26051

Adaptive action chunking for fine manipulation via trajectory prediction discrepancy

Zheng Shan[1], Yang Ziyi[1], Liu Yikun[1], Zhou Han[2], Wang Gang[3], Wu Qinghua[3], Shen Tianyu[1], Li Zhiwei[1]
1. Beijing University of Chemical Technology, Beijing 100029, China
2. Beijing University of Posts and Telecommunications, Beijing 100876, China
3. Beijing Embodied Intelligence Technology Co., Ltd., Beijing 100085, China

Abstract: To address the difficulty of fixed action chunking strategies in simultaneously maintaining action continuity during stable phases and execution responsiveness during critical phases in embodied manipulation tasks, an adaptive action chunking scheduling method based on prediction discrepancy was proposed. Without modifying the model architecture or training procedure, the proposed method established an inference mechanism of “fixed-length prediction and adaptive-length execution” by dynamically adjusting the number of action steps written into the action cache and executed consecutively during cache refresh at inference time. A trajectory-level prediction discrepancy metric was defined based on the differences between adjacent predicted chunks over their overlapping temporal intervals, and online switching of the execution chunk length was achieved by combining temporal smoothing with piecewise threshold rules. Experimental results on a real robotic arm platform demonstrate that the proposed method achieves the best or near-best performance in task success rate and trajectory error, while maintaining substantially better execution smoothness than the fixed small-chunk strategy, thereby providing a better overall balance among execution accuracy, closed-loop responsiveness, and tra-

收稿日期：2026-04-30；**修回日期：**2026-07-06
通信作者：李志伟，lizw@buct.edu.cn
基金项目：北京市自然科学基金资助项目（No.4252048）
Foundation Item: Beijing Natural Science Foundation (No.4252048)

jectory stability. The method also exhibits good robustness under both visual and geometric disturbances. Further ablation experiments and hyperparameter analysis indicate that trajectory-level prediction discrepancy, temporal smoothing, and piecewise threshold scheduling collectively enhance the stability and phase adaptability of execution chunk-length scheduling. The results demonstrate that this method effectively captures task phase transitions and improves execution adaptability across different phases.

Key words: embodied intelligence, robotic manipulation, adaptive action chunking

0　引言

工业精细操作正由固定工位、单一工序向开放环境、动态扰动和长时程多阶段任务演进。以装配、插接、压装与对齐为代表的典型工况中，机器人不仅需要在自由空间完成快速到位与避障，还需要在接触发生后持续应对摩擦不确定、装配间隙极小以及姿态微偏差易累积等问题。Jiang 等[1]指出工业装配通常涵盖规划、执行与评估等完整流程；Abbas 等[2]进一步指出，在更复杂的协同操作场景中，系统还需处理机械臂之间几何约束、力分配、自碰撞及协同配合等耦合因素。因此，在不过度增加系统建模、训练与部署复杂度的前提下，实现兼具高效率、稳定性与鲁棒性的机械臂精细操作控制，仍是当前机器人落地应用中的关键问题。

模仿学习（imitation learning）因能够直接从专家示教中学习状态到动作的映射关系、降低显式动力学建模负担，已成为机器人控制的重要技术路径。Argall 等[3]系统总结了机器人示教学习的基本范式，Osa 等[4]从算法视角归纳了模仿学习的主要技术路线，张超等[5]进一步综述了其在机器人任务中的研究进展。在具体方法上，行为克隆通过监督学习拟合专家动作，但易受到分布偏移影响[3-5]；Ross 等[6]提出 DAgger，通过交互式数据聚合缓解复合误差；Ho 等[7]提出 GAIL，从分布匹配角度学习更稳健的策略。随着 Transformer 与生成式策略的发展，模仿学习逐步由单步动作回归演进为长时序行为建模。Brohan 等[8]提出 RT-1，展示了大规模数据与高容量模型在真实机器人控制中的可扩展性；Chi 等[9]提出 Diffusion Policy，在接触任务中提升了动作多模态表达能力与轨迹平滑性。在精细操作场景中，Zhao 等[10]提出 ACT 算法通过动作分块与 CVAE-Transformer 结构预测短时间窗动作序列，显著缩短有效决策视界，并在低成本双臂平台上验证了模仿学习完成精细任务的可行性。然而，上述方法在长时程任务中的误差累积问题及对阶段切换的敏感性仍未得到根本解决。

近年来，围绕动作序列策略的推理阶段优化，已有研究开始关注动作块执行过程中的闭环性与反应性问题。Diffusion Policy 采用基于动作序列预测的滚动执行机制，提高了视觉运动策略在接触操作中的轨迹连续性与多模态表达能力[9]；ACT 通过动作分块降低长时程模仿学习中的有效决策视界，并结合时间集成提升执行平滑性[10]。在此基础上，InterACT 和 Bi-ACT 等方法进一步面向双臂协同操作扩展动作块生成结构，增强了动作序列建模与双臂依赖关系表达能力[11-12]。近期，Liu 等[13]从测试时解码角度分析了动作分块对策略反应性的影响，并通过闭环重采样机制改善动作块执行过程中的时序一致性与环境响应能力。这些研究表明，动作块策略的性能不仅取决于动作序列生成质量，也受到推理阶段执行方式的显著影响。

除模仿学习外，工业精细装配与协同操作领域亦形成了较为系统的研究体系。在装配控制方面，Hogan[14]提出的阻抗控制通过调节机器人与环境之间的动态关系实现接触柔顺性，Raibert 等[15]提出的混合力/位置控制通过对不同自由度实施力控与位控实现力位协调。为提高对复杂工况的适应能力，研究者进一步将强化学习与多模态感知方法引入接触任务。Zhao 等[16]将离线元强化学习应用于工业插入任务以提升策略泛化能力；Nair 等[17]提出离线到在线微调框架，实现对新型连接器插入任务的自适应优化；Jin 等[18]融合视觉与力觉信息并结合课程学习提升接触任务稳定性；Wang 等[19]在半结构化环境中结合残差学习与示教实现机器人装配。在协同操作方面，Smith 等[20]综述了双臂操作在建模、规划与控制中的关键问题；Ren 等[21]提出通用双臂协同操作框架以提升系统灵巧性与泛化能力；Kim 等[22]将 Transformer 引入双臂模仿学习以增强高维时序建模能力；Lee 等[11]提出 InterACT 建模双臂依赖关系以改进动作块生成；Buamanee 等[12]提出 Bi-ACT，将动作分块与双向控制结合以增强接

触任务中的协调性与适应能力。总体来看，现有研究从经典控制到数据驱动方法已形成较为完整的技术体系，其中接触密集型装配与协同操作场景是验证复杂耦合控制能力的典型任务。

尽管已有研究从动作块生成结构、测试时重采样以及双臂协同建模等角度提升了模仿学习策略的执行性能，但面向长时程、多阶段机械臂精细操作的推理阶段执行粒度调度仍有进一步研究空间。现有 ACT 及其扩展方法多采用固定执行块长[10]，测试时闭环重采样方法主要通过多候选动作块采样与解码准则改善动作选择一致性[13]，而不确定性估计或动态 horizon 方法通常需要额外置信度建模、采样过程或规划视界调整。对于真实机械臂精细操作任务，系统在自由空间接近、稳定搬运、接触对齐和精密插入等阶段呈现明显的时间尺度差异：较大的执行块长有利于降低动作缓存刷新频率和动作子段切换频率，从而保持稳定阶段的动作连续性，但在关键接触阶段可能使最新观测对应的预测结果较晚作用于实际执行，导致块内误差累积；较小的执行块长能够使动作缓存更频繁地刷新，使最新预测结果更快作用于实际执行，从而增强局部纠偏能力，但频繁的动作子段切换也可能降低执行平滑性并引入控制波动。因此，该问题并非简单选择某一固定块长即可解决，而是需要在不增加训练成本与模型复杂度的前提下，根据当前任务阶段动态调整推理阶段的执行粒度。

针对上述问题，本文提出一种基于轨迹预测分歧的机械臂精细操作自适应动作分块方法，在保持现有动作块生成网络结构、训练目标与固定预测长度不变的前提下，仅在推理执行环路中引入轻量级调度机制。具体而言，基线策略在每个控制时刻仍生成固定长度动作块，调度器利用相邻预测动作块在时间轴上的重叠关系，对同一未来区间构造多次预测，并以重叠区间内的轨迹级预测分歧作为策略自一致性信号。当系统处于自由空间接近或平稳搬运等稳定阶段时，预测一致性较高，调度器选择较大的执行块长以降低动作缓存刷新频率、保持动作连续性；当进入接触、对齐或插入等高不确定阶段时，预测分歧增大，调度器切换为较小的执行块长，使最新观测对应的预测结果更快写入动作缓存并作用于实际执行，从而增强纠偏能力。为保证调度稳定性，进一步引入时间平滑机制，以抑制瞬时预测波动导致的频繁切换。本文的主要贡献如下。

（1）提出一种基于轨迹预测分歧的推理阶段自适应动作分块调度方法，在不改变训练目标与模型结构的前提下，实现执行块长的动态选择。

（2）构建由轨迹预测分歧度量、时间平滑与分段阈值调度组成的推理端调度流程，在不增加模型参数与离线训练成本的条件下，为现有动作块策略提供一种轻量化扩展方式。

（3）围绕单机械臂与双机械臂两类代表性工业精细操作任务开展实验验证，并进一步通过消融实验与超参数敏感性分析考察核心模块和关键参数对调度性能的影响。实验结果表明，本文方法能够兼顾任务成功率、关键阶段响应能力与整体执行稳定性。

1 动作块生成模型

1.1 系统输入与动作块定义

本文采用“观测到动作块”的模仿学习建模范式。在离散时间步 t，系统输入由多视角视觉观测与机器人本体状态共同构成，记为：

$$o_t=\{I_t,q_t\} \tag{1}$$

其中，I_t 表示多视角 RGB 视觉观测，$q_t\in\mathbb{R}^{d_q}$ 表示机器人当前本体状态（如机械臂关节位置与夹爪状态），并在输入网络前进行归一化处理，以便与视觉特征在统一表示空间中融合。

不同于传统行为克隆方法仅预测当前时刻动作 a_t，本文将策略输出定义为未来有限时域内的动作序列，即动作块（action chunk）：

$$a_{t:t+H-1}=\{a_t,a_{t+1},\cdots,a_{t+H-1}\},\quad a_t\in\mathbb{R}^{d_a} \tag{2}$$

其中，H 为动作块长度，d_a 为动作维度。本文在模型训练与预测分歧计算中采用统一的动作标准化表示；在实际执行阶段，输出动作再映射回机器人底层控制器所需的原始控制量空间。策略目标可形式化为条件生成问题，即：

$$\pi_\theta\left(a_{t:t+H-1}\middle|o_t\right) \tag{3}$$

该建模方式使模型能够在单次前向推理中联合预测未来多个控制步骤，从而显式建模局部时间窗口内的动作相关性，并缩短长时程任务中的有效决策视界[9-10,23]。

1.2 基于 Transformer 的动作块生成策略

本文采用以 Transformer 为核心的动作块生成模型作为基线策略。其基本流程包括：对多模态观

测进行编码、建立观测与未来动作块之间的条件映射，并生成长度为H的动作序列预测[24-28]。

具体而言，首先将多视角图像I_t通过视觉编码器映射为特征表示f_I，并与状态向量$\boldsymbol{q}_t$融合得到统一观测特征：

$$\boldsymbol{f}_t=\phi\left(\boldsymbol{I}_t,\boldsymbol{q}_t\right) \tag{4}$$

在此基础上，Transformer编码器对$\boldsymbol{f}_t$进行上下文建模，解码器则以该条件特征为输入，自回归或并行地生成未来动作块$a_{t:t+H-1}$。

为增强模型对人类示教中多模态性与不确定性的表达能力，本文采用条件变分自编码器（CVAE）结构进行扩展，引入潜变量z表征隐式策略分布：

$$\pi_\theta\left(a_{t:t+H-1}\middle|o_t,z\right) \tag{5}$$

训练阶段通过最大化重构似然并约束潜变量分布进行优化；测试阶段采用先验分布均值进行确定性推理。该设计继承了ACT框架的核心思想，在保持动作块建模优势的同时，提高了对复杂示教数据的拟合能力[29]。

在推理阶段，模型以滚动窗口的方式在每个控制时刻生成固定长度动作块，并通过时间集成（temporal ensembling）对历史预测中指向同一未来时刻的动作进行加权融合。设当前控制时刻为t，历史预测队列中第r次预测给出的未来时刻τ的动作记为$\hat{a}_\tau^{(r)}$，则参与融合的预测集合定义为$\mathcal{R}(\tau)=\left\{r\middle|r\leqslant t,\tau\in\left[r,r+H-1\right]\right\}$，融合动作定义为：

$$\tilde{a}_\tau=\frac{\sum\limits_{r\in\mathcal{R}(\tau)}w(t-r)\hat{a}_\tau^{(r)}}{\sum\limits_{r\in\mathcal{R}(\tau)}w(t-r)} \tag{6}$$

该机制在保持固定预测长度的同时，提升同一执行时刻动作指令的连续性与稳定性[9-10]。这一统一预测框架为后续引入自适应动作缓存长度调度提供了基础。

1.3　固定动作分块策略的局限

尽管动作块建模能够有效缓解长时程模仿学习中的误差累积问题，但现有ACT及其扩展方法在推理阶段通常采用固定执行长度（fixed chunk size）。这一设定隐含假设任务在整个执行过程中具有一致的时间尺度与控制需求，而该假设在工业精细操作场景中往往不成立[6,28]。

工业操作任务通常具有显著的阶段异质性。例如，在自由空间接近或稳定搬运阶段，系统状态变化平缓，较大执行块长有助于降低动作缓存刷新频率和动作子段切换频率，从而保持动作连续性；而在接触对齐、精密插入等关键阶段，环境约束增强且误差敏感性显著提高，系统更依赖最新观测对实际执行动作的及时修正。已有研究表明，接触式操作对几何配合、力位协同及微小误差极为敏感[30-32]，而双臂协同任务进一步引入跨臂耦合与协调约束，使这一问题更加突出[28,32]。

在此背景下，固定动作分块策略不可避免地面临动作连续性与关键阶段响应能力之间的内在矛盾：较大执行块长虽有利于保持动作连续性，但在接触、对齐和插入等关键阶段，过长的动作缓存可能延迟最新观测对实际执行动作的影响，从而导致块内误差累积；较小执行块长能够使动作缓存更频繁地刷新，使最新预测结果更快作用于实际执行，但频繁的动作子段切换也可能降低执行平滑性并引入控制波动。

更为关键的是，该问题并不源于某一固定块长选择不当，而在于固定分块机制本身缺乏对任务阶段变化的自适应能力。换言之，现有方法能够生成高质量动作块，但缺乏在推理阶段根据当前状态动态调整执行节奏的能力。因此，有必要在不改变训练过程与模型结构的前提下，引入面向推理阶段的自适应调度机制，使系统能够在稳定阶段保持动作连续性，并在关键阶段增强执行动作的刷新响应能力。

2　基于预测分歧的自适应动作分块调度方法

为解决固定动作分块难以适应多阶段工业任务的问题，本文提出一种基于预测分歧的自适应动作分块调度方法。该方法不改变基线模型的训练目标、网络结构及固定预测长度，仅在推理阶段根据当前状态是否进入不稳定区域，自适应调整实际动作缓存刷新时的执行块长。基于轨迹预测分歧的自适应动作分块方法总体框架如图1所示。其核心思想是：利用连续两次动作块预测在时间重叠区间内形成的多次预测集合，构造轨迹级预测分歧指标，以此刻画模型对未来局部轨迹的自一致性；当预测分歧较小时，采用较大的执行块长以降低动作缓存刷新频率并保持动作连续性；当预测分歧增大时，缩短执行块长，使最新观测对应的预测结果更快写入动作缓存并作用于实际执行，从而增强局部纠偏能力[9-10]。

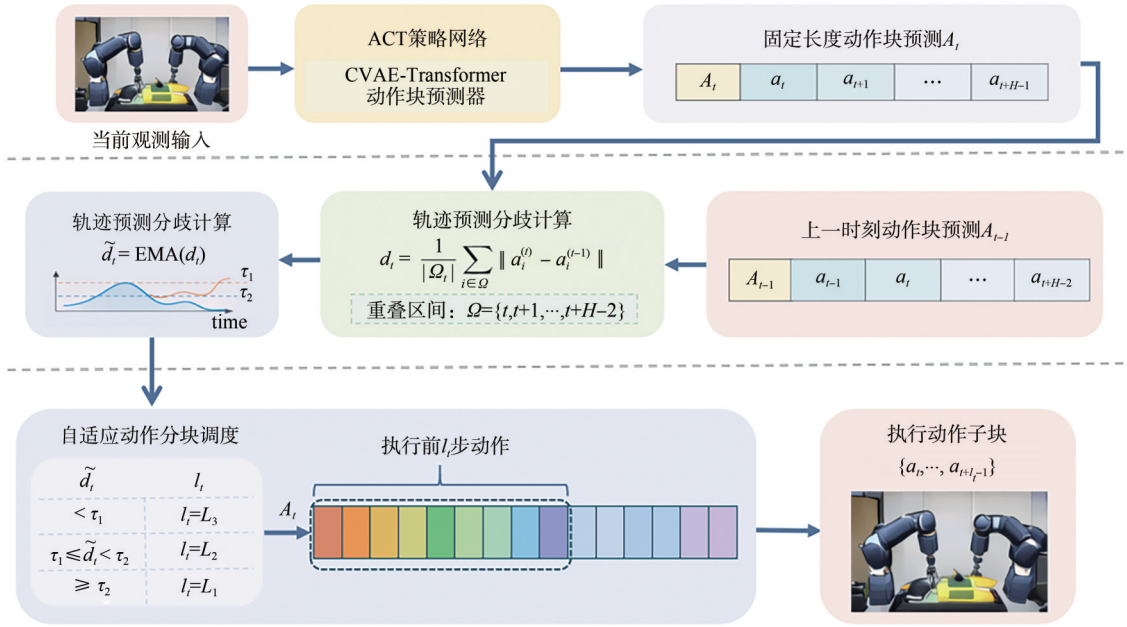


图1 基于轨迹预测分歧的自适应动作分块方法总体框架

2.1 轨迹预测分歧度量

在滚动预测框架下，策略在每个时刻t都会输出一个长度为H的未来动作块预测，记为：

$$\hat{A}^{(t)}=\left\{\hat{a}_t^{(t)},\hat{a}_{t+1}^{(t)},\cdots,\hat{a}_{t+H-1}^{(t)}\right\} \tag{7}$$

其中，$\hat{a}_{t+j}^{(t)}$表示在时刻t基于当前观测所预测的、将于未来时刻$t+j$执行的动作。

由于系统在下一时刻$t+1$仍会基于最新观测重新生成未来动作块，因此相邻两次预测在时间轴上天然存在重叠区间。对于同一未来执行时刻，系统实际上获得了来自不同时刻观测条件下的多次预测。若系统处于状态变化平缓、模型预测稳定的阶段，则这些预测应具有较高一致性；反之，若环境约束发生突变，或系统进入接触、对准、插入等高敏感阶段，则相邻预测块在重叠区间上的差异通常会明显增大。

考虑到本文动作向量同时包含机械臂关节目标和夹爪开合状态，不同动作维度的量纲和数值范围可能存在差异。为避免大尺度维度在欧氏范数计算中占据主导，本文在计算轨迹预测分歧时采用与策略训练阶段一致的动作标准化表示。具体而言，动作向量在进入分歧计算前按照训练示范数据统计得到的逐维均值和标准差进行标准化；该统计量仅由训练示范数据计算得到，并在正式测试过程中保持固定，不根据测试结果更新。

基于上述观察，本文定义时刻t的轨迹预测分歧为相邻两次预测块在重叠区间上的多步动作差异平均值。基于重叠动作分块的时序分歧计算方法如图2所示，相邻时刻生成的预测动作块在时间轴上存在重叠区间，本文基于该区间内对应动作的差异定义轨迹预测分歧指标。设Ω_t表示由$\hat{A}^{(t-1)}$与$\hat{A}^{(t)}$形成的重叠时间索引集合，则轨迹预测分歧定义为：

$$D_t=\frac{1}{|\Omega_t|}\sum_{\tau\in\Omega_t}\left\|\hat{a}_\tau^{(t)}-\hat{a}_\tau^{(t-1)}\right\|_2 \tag{8}$$

其中，$|\Omega_t|$表示重叠区间长度；$\|\cdot\|_2$为欧氏范数，式（8）中的动作向量均指进入分歧计算前经过逐维标准化后的动作预测值，因此该欧氏范数是在标准化动作空间中计算，用于衡量同一未来执行时刻下两次动作预测之间的差异。

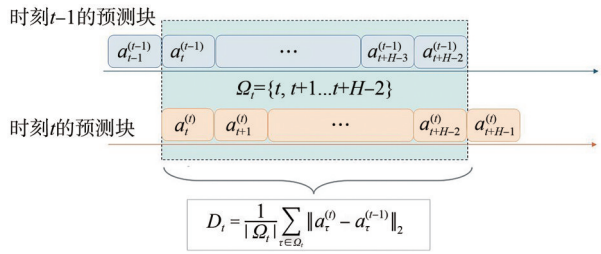


图2 基于重叠动作分块的时序分歧计算方法

式（8）刻画了在观测更新后模型对未来局部轨迹的重估程度。与仅比较单一步动作差异的方法相比，轨迹级预测分歧综合利用了重叠区间内的多

步信息，能够在一定程度上抑制瞬时噪声带来的随机波动，更稳定地反映系统当前所处阶段的预测一致性变化。因此，该指标更适合作为推理阶段执行节奏切换的判据。

2.2　分歧时间平滑

在实际推理过程中，轨迹预测分歧D_t不仅受到任务阶段变化的影响，还会受到视觉观测噪声、状态估计误差以及模型瞬时波动的干扰。若直接依据原始分歧值进行调度，则在阈值附近容易引起执行块长的频繁切换，从而带来不必要的控制抖动与规划开销。为提高调度器的鲁棒性，本文对分歧信号进行时间平滑处理。

本文采用指数滑动平均（exponential moving average，EMA）方式构造平滑分歧指标，即：

$$\bar{D}_t=\alpha D_t+(1-\alpha)\bar{D}_{t-1} \tag{9}$$

其中，$\alpha\in[0,1)$为平滑系数，$\bar{D}_t$表示时刻t的平滑分歧值。本文令初始平滑分歧$\bar{D}_0=0$，并在后续控制时刻按式（9）递推更新。

平滑系数α用于调节当前分歧观测与历史平滑分歧之间的权重关系。较大的α提高调度器对当前分歧变化的响应速度，较小的α则增强历史信息对调度结果的约束作用，有助于抑制短时噪声引起的块长频繁切换。本文在正式测试前通过独立预实验确定平滑系数，并在所有测试实验中固定使用$\alpha=0.3$；平滑系数的不同取值会影响调度器的响应速度与切换稳定性，本文在第3.5节进一步通过敏感性分析验证其对执行性能的影响。该设计能够在保留阶段变化整体趋势的同时，抑制局部随机扰动导致的误触发。平滑后的分歧信号为后续调度规则提供了更稳定的输入，使执行块长的更新更多反映真实任务阶段变化，而非短时噪声波动。

2.3　自适应动作分块调度规则

在获得平滑分歧$\bar{D}_t$后，本文采用分段阈值规则在动作缓存刷新时确定当前执行块长，设候选执行块长集合为：

$$K=\{k_{\min},k_{\mathrm{mid}},k_{\max}\} \tag{10}$$

并设置两个阈值θ_1与θ_2，其中$\theta_1<\theta_2$，则调度规则定义为：

$$k_t=\begin{cases}k_{\max}, & \bar{D}_t<\theta_1\\ k_{\mathrm{mid}}, & \theta_1\le\bar{D}_t<\theta_2\\ k_{\min}, & \bar{D}_t\ge\theta_2\end{cases} \tag{11}$$

式（11）的物理含义如下：当$\bar{D}_t$较小时，相邻预测块在重叠区间内保持较高一致性，说明系统处于相对稳定阶段，此时采用较大的执行块长可以降低动作缓存刷新频率和动作子段切换频率，从而保持动作连续性；当$\bar{D}_t$处于中等范围时，系统可能正处于由稳定推进向精细操作过渡的阶段，此时采用中等块长有助于在动作连续性与执行动作刷新响应能力之间取得折中；当$\widehat{D}_t$明显增大时，表明模型对未来轨迹的重估程度上升，系统更可能处于接触、对准或插入等高敏感阶段，此时应切换为较短的执行块长，使最新观测对应的预测结果更快写入动作缓存并作用于实际执行，从而增强局部误差修正能力[9,30]。

需要强调的是，本文方法改变的是推理阶段的执行缓存长度，即实际执行块长，而非基线策略的“预测长度”。基线模型在每个控制时刻仍输出长度固定为H的动作块，调度器仅在动作缓存刷新时决定写入并连续执行的动作步数k_t，其余未立即写入当前缓存的预测结果仍保存在预测历史队列中，用于后续重叠区间上的预测分歧计算和时间集成。因此，本文方法在模型结构、训练过程、固定预测长度和策略预测频率上与固定分块基线保持一致，其性能改进来源于推理阶段执行缓存长度调度机制的引入，而非模型容量、训练复杂度或策略查询频率的差异。在本文实验中，候选执行块长集合设置为$\{10,50,100\}$。其中，k=10对应较短动作缓存，主要用于接触、对齐和插入等误差敏感阶段，使最新观测对应的预测结果能够更快作用于实际执行；k=50对应中等执行粒度，用于阶段过渡或中等不确定性状态；k=100对应较长动作缓存，主要用于自由空间接近和平稳搬运等状态变化较缓阶段，以减少动作子段切换并保持动作连续性。该候选集合覆盖了从快速纠偏到连续执行的典型执行时间尺度。

由于轨迹预测分歧D_t在标准化动作空间中计算，阈值θ_1和θ_2也对应标准化分歧尺度下的无量纲阈值。本文采用阈值$\theta_1=0.15$、$\theta_2=0.30$，这两个阈值通过正式评估前的独立预实验确定。具体而言，本文在不同候选阈值组合下运行策略，综合比较分歧曲线与任务阶段变化的一致性，以及任务成功率、轨迹误差、执行平滑性和块长切换次数等指标，最终选取能够较稳定地区分低分歧稳定阶段、中等分歧过渡阶段和高分歧关键阶段的阈值组合。正式测试过程中不再根据测试结果调整阈值。

2.4 推理执行流程

结合上述轨迹预测分歧度量、分歧时间平滑与自适应动作分块调度规则，本文在推理阶段构建“固定长度预测–连续分歧监测–缓存边界调度”的在线闭环执行流程。该流程不改变基线动作块生成网络结构和训练方式，而是将策略预测与动作执行解耦：策略网络在每个控制时刻均生成固定长度为H的预测动作块，用于更新预测分歧信号并维护时间集成历史；机器人实际执行的动作则来自当前动作缓存。只有当动作缓存为空或执行完毕时，调度器才根据当前平滑分歧更新下一段实际执行块长。

具体而言，在每个控制时刻t，系统首先采集当前观测o_t，并输入动作块生成策略π_θ，得到长度为H的未来动作块预测$\hat{A}^{(t)}$。若上一时刻预测块$\hat{A}^{(t-1)}$存在，则系统在两者的重叠区间上计算轨迹预测分歧D_t，并进一步通过式（9）更新平滑分歧$\bar{D}_t$。需要强调的是，预测分歧的更新与执行块长的生效并非逐时刻等价：前者在每个控制时刻连续进行，用于形成稳定的阶段变化信号；后者仅在当前动作缓存为空或者执行完毕时触发。具体地，当动作缓存尚未执行完毕时，系统继续执行缓存中的下一步动作，同时持续更新预测历史队列和当前平滑分歧$\bar{D}_t$；当缓存为空或者执行完毕时，调度器读取当前时刻的$\bar{D}_t$，根据式（11）确定当前执行块长k_t，随后，系统基于预测历史队列对未来$t, t+1, \cdots, t+k_t-1$这k_t个控制步的动作进行时间集成，生成新的动作缓存。最后系统继续按控制频率逐步执行缓存中的动作，从而形成“连续分歧监测–缓存边界调度”的闭环推理过程。

为更清晰地描述本文方法的推理过程，完整流程如算法1所示。

算法1 基于预测分歧的自适应动作分块推理流程

（1）初始化动作块生成策略π_θ

（2）初始化分歧平滑变量$\bar{D}_0$

（3）初始化上一时刻预测动作块$\hat{A}^{(0)}=\varnothing$，预测历史队列$P=\varnothing$

（4）设定固定预测长度H、候选执行块长集合$\mathcal{K}$

（5）设定分歧平滑系数α以及阈值θ_1, θ_2

（6）初始化动作缓存$B=\varnothing$，缓存指针$j=1$

（7）对每个控制时刻$t=1,2,\cdots$重复具体步骤

（8）采集当前观测o_t

（9）利用策略网络生成当前预测动作块$\hat{A}^{(t)}$，加入预测历史队列P

（10）若$\hat{A}^{(t-1)}\neq\varnothing$，则

（11）计算当前预测块与上一预测块的重叠区间Ω_t

（12）根据式(8)计算轨迹预测分歧D_t

（13）根据式(9)更新平滑分歧$\bar{D}_t$

（14）否则，令$D_t=0$，$\bar{D}_t=\bar{D}_{t-1}$

（15）结束若

（16）若动作缓存B为空，或当前缓存已执行完毕，则

（17）根据式（11）和当前$\bar{D}_t$选择当前执行块长k_t

（18）对未来执行时刻$\tau=t, t+1, \cdots, t+k_t-1$，从预测历史队列中选取所有覆盖该时刻的预测结果，经式（6）时间集成后写入动作缓存B

（19）重置缓存指针$j=1$

（20）结束若

（21）执行当前动作缓存中的第j步动作

（22）更新缓存指针$j\leftarrow j+1$

（23）更新上一时刻预测块：$\hat{A}^{(t-1)}\leftarrow\hat{A}^{(t)}$

（24）进入下一控制时刻

（25）结束循环

从算法1的结构上看，该流程体现了“高频分歧估计、缓存边界调度”的设计思想。系统在每个控制时刻均生成动作块预测并更新平滑分歧，从而持续跟踪模型对未来局部轨迹的一致性变化；但执行块长仅在当前动作缓存为空或执行完毕时，根据最新平滑分歧$\bar{D}_t$重新选择。因而，外层循环始终以控制时刻为单位递增，执行块长表示当前动作缓存连续执行的步数。该设计既保留了逐时刻预测分歧监测对阶段变化的敏感性，又避免了瞬时分歧波动导致的逐步块长重置。

2.5 计算开销分析

本文方法的轻量化主要体现在不改变基线策略网络结构、模型参数和离线训练流程。本文自适应动作分块调度机制不引入额外策略网络、价值网络、不确定性估计模块或多模型集成。

在推理阶段，为保证轨迹预测分歧的连续估计，并保证与固定块长基线之间的公平比较，本文所有方法均在每个控制时刻生成固定长度为H的预测动作块，并采用相同的时间集成机制。因此，

Adaptive-k 与 Fixed-k=10、Fixed-k=50、Fixed-k=100 在策略网络查询频率上保持一致，性能差异主要来自动作缓存刷新时执行块长的选择方式，而非策略网络调用频率差异。

相较于固定块长基线，本文方法新增的后处理计算主要包括相邻预测块重叠区间内的标准化动作向量差分、欧氏范数计算、均值计算、EMA 平滑以及分段阈值判断。上述操作均作用于策略网络已经输出的低维动作序列，计算量随预测长度 H 和动作维度 d_a 线性增长。与基于 Transformer 的策略网络前向推理相比，该部分属于动作空间上的轻量级后处理计算。

因此，本文方法不增加模型参数和离线训练成本，也不引入额外神经网络模块。其额外开销主要来自动作序列后处理和缓存调度判断。性能改进主要来源于推理阶段执行缓存长度的自适应调度，而非模型容量、训练规模或策略查询频率差异。

3　实验

围绕本文提出的基于预测分歧的自适应动作分块调度方法开展实验验证，重点评估该方法在不增加模型结构复杂度与训练成本的前提下，是否能够实现：①提升任务成功率与系统鲁棒性；②增强关键接触阶段执行动作的刷新响应能力；③在稳定阶段保持较好的执行连续性；④通过轨迹级预测分歧、时间平滑与分段阈值调度的协同作用，提升执行块长调度的稳定性。

3.1　实验平台与任务设置

本文实验基于机械臂模仿学习与遥操作一体化平台构建。该平台由示教端双臂遥操作控制器、执行端双臂执行机械臂、平行夹爪、多视角视觉感知系统以及控制与算力平台组成，具备示教数据采集、离线策略训练以及模型部署后的在线推理与实机执行能力。

系统整体采用示教端–执行端（leader-follower）架构：在示教阶段，示教端通过人机交互产生参考控制信号；执行端在执行示教动作的同时，同步记录关节状态与多视角观测数据，用于后续模仿学习训练。在策略部署阶段，执行端基于当前观测进行模型推理，并输出控制指令驱动机械臂执行，从而形成完整的感知–决策–执行闭环。

视觉系统采用多路 RGB-D 相机（顶视、底视及双腕视角）对操作空间进行覆盖，为单臂和双臂精细操作任务提供多视角观测信息。平台遥操作采样频率为 100 Hz，策略执行频率约为 50 Hz，能够满足接触密集型任务对实时性与控制精度的要求。为验证本文方法在不同阶段特性任务中的表现，选取两类具有代表性的工业机械臂操作任务。

（1）Pick & Place（螺丝放置任务）

该任务采用单臂执行，以自由空间运动与弱接触操作为主，系统状态变化平缓，稳定推进阶段占比较高。此类任务对快速缓存刷新和局部纠偏的依赖较弱，更强调执行过程的连续性与稳定性。因此，该任务主要用于评估在稳定阶段条件下，自适应分块策略相较于固定分块策略在任务完成能力、动作连续性和执行稳定性方面的表现。

（2）Insert/Align（双臂轴承插入任务）

该任务为典型的双臂协作精细操作任务，包含抓取、传递及对准插入等多个阶段。前期阶段以大范围运动为主，更依赖动作连续性；而对准与插入阶段处于接触约束条件下，对位姿误差高度敏感，对执行动作的刷新响应能力与快速纠偏能力提出更高要求。该任务用于评估自适应分块策略在多阶段任务中的动态调度能力，以及其在动作连续性、关键阶段响应能力与控制精度之间的平衡性能。

实验平台的主要硬件参数见表 1，实验平台与典型任务场景示意如图 3 所示。

表 1　实验平台硬件参数

参数	数值
DOF	14（双臂，每臂 7DOF）
机械臂系统	睿尔曼 RM71-D（控制端）+RM75-B-V（执行端）
末端执行器	平行夹爪（行程 70~80 mm，夹持力 40~140 N）
视觉系统	Intel RealSense D435i ×4RGB-D 相机（双目深度+RGB+IMU）
图像分辨率	最高 1 280 像素×720 像素
控制频率	100 Hz（遥操作采样），执行频率 50 Hz
观测模态	关节状态+多视角图像
计算平台	RTX 4090

3.2　数据采集与轨迹可视化

3.2.1　数据来源与采样设置

本文实验数据基于真实机械臂遥操作示范构建。系统采用示教端–执行端（leader-follower）架构，由示教端（leader）进行示教控制，执行端（follower）在执行过程中同步记录机器人本体状态与多模态观测信息，并可基于当前观测进行策略推理与动作执行。平台底层遥操作与控制采样频率为 100 Hz。

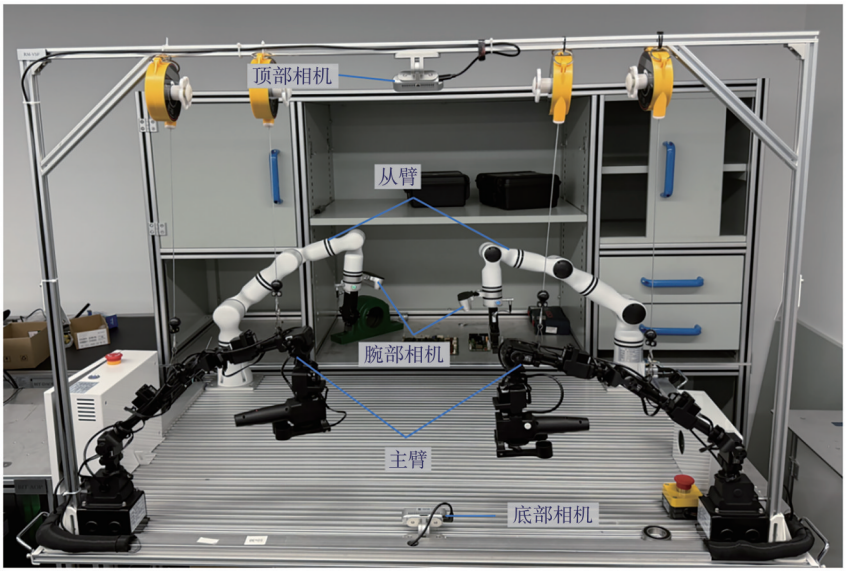


图3　实验平台与典型任务场景示意

为与策略执行频率保持一致，本文在数据预处理阶段将示范数据重采样至50 Hz，并在策略训练与部署阶段均采用约50 Hz的时间步进行建模与执行，各模态数据在时间维度上保持严格对齐。

在每一时间步t，系统记录如下信息。

（1）机械臂关节位置及夹爪状态，作为机器人本体状态。

（2）来自顶视、底视及左右腕部视角的4路RGB-D图像，作为视觉观测输入。

针对不同任务场景，本文分别采集约200条用于训练的成功示范轨迹，单条轨迹长度约400~700步（8~14 s），覆盖自由空间运动与接触对齐等不同阶段的操作过程。数据集统计信息见表2，本文在单臂Pick & Place任务中仍采用双臂关节与夹爪状态构成16维状态/动作表示，其中非操作臂在任务执行过程中保持静止，其对应维度作为恒定状态记录。

表2　数据集统计信息

任务	示范条数	轨迹长度/步	采样频率/Hz	相机数量/个	控制维度
Pick & Place	200	400~700	50	4	16
Insert / Align	200	400~700	50	4	16

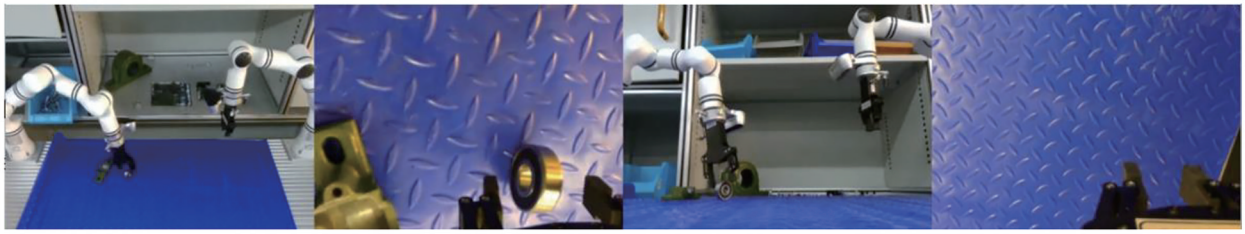
本文实验未采用图像分类等离线监督任务中常见的固定训练集、验证集和测试集划分方式。每个任务采集的成功示范轨迹用于模仿学习策略训练，模型性能不通过离线轨迹回放评估，而是在真实机器人平台上进行独立闭环执行测试。平滑系数与分段阈值通过正式测试前的独立预实验确定，预实验结果不计入主实验、鲁棒性实验、消融实验和超参数敏感性分析的统计结果。正式测试阶段，各方法均部署到真实机器人平台进行多次独立执行，测试结果仅用于性能统计，不再用于模型训练或调度参数调整。通过统计任务成功率、轨迹误差及执行平滑性等指标，评估策略在闭环控制条件下的实际性能。

3.2.2　数据表示与可视化

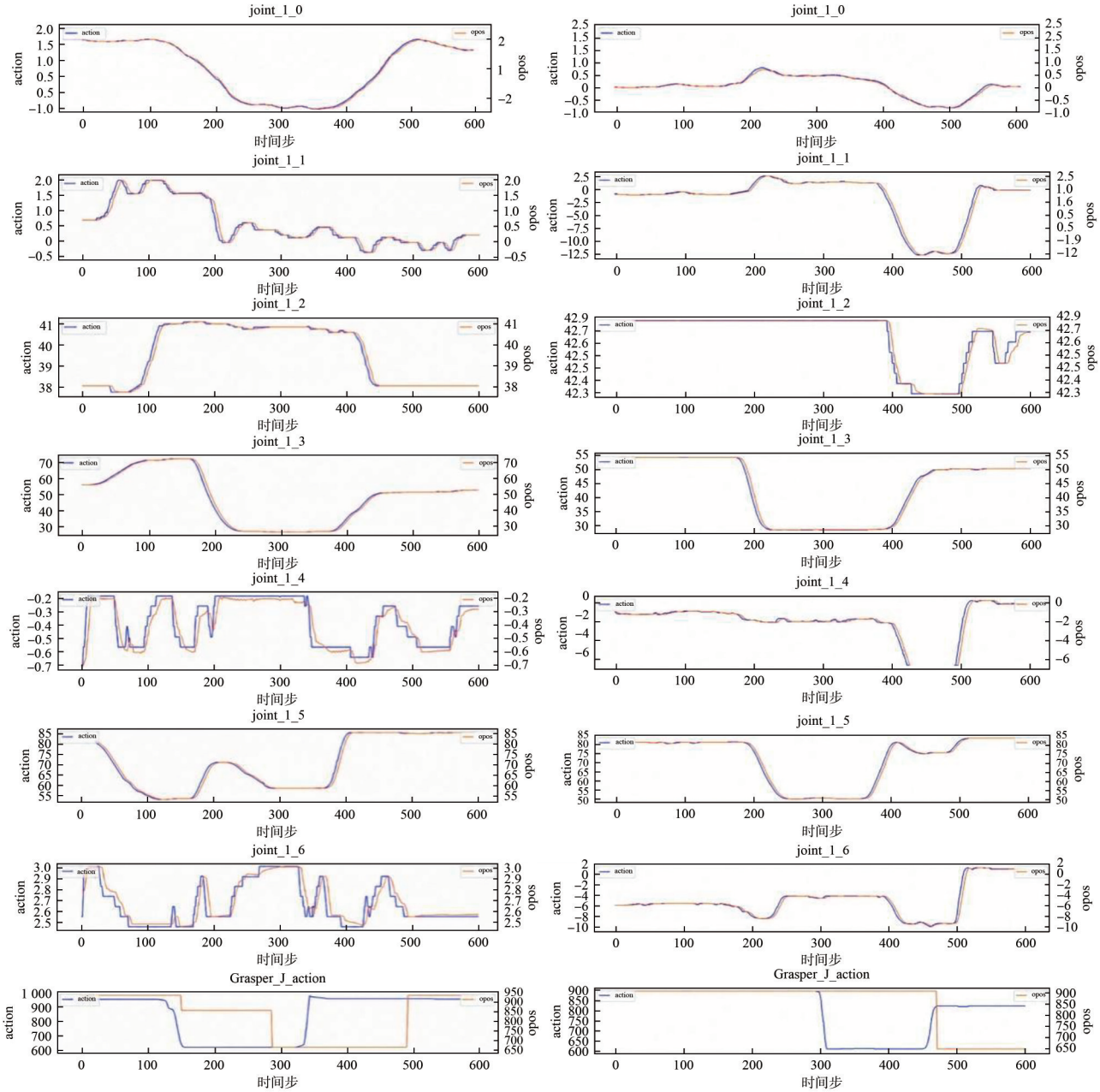
本文采用HDF5格式对多模态示范数据进行统一存储与管理。每条示范轨迹对应一个独立数据条目，包含时间严格对齐的机器人本体状态与多视角视觉观测序列。其中，本体状态包括机械臂关节位置及夹爪开合状态；视觉观测由顶视、底视及左右腕部视角的4路RGB-D图像序列构成，各模态数据在时间维度上保持同步一致。

为对示范数据的结构特性与质量进行分析，本文从动作表示与视觉观测两个层面对轨迹进行可视化建模。

（1）视觉层面：对四视角图像序列进行关键帧提取与拼接展示，从多视角呈现示范过程中环境变化、遮挡情况及关键接触时刻的视觉特征。多视角观测关键帧拼接结果如图4（a）所示，多视角观测能够覆盖机械臂操作空间的主要区域，并在关键阶段提供完整的视觉信息，有助于评估观测数据的完整性与一致性。



（a）多视角观测关键帧拼接结果



（b）机械臂关节与夹爪状态时序变化曲线

图4　机械臂操作示范数据可视化结果

（2）动作层面：提取机械臂关节位置及夹爪状态，构建随时间变化的多维时序曲线（共16维），用于表征示范过程中动作的动态变化特征。机械臂关节与夹爪状态时序变化曲线如图4（b）所示，关节与夹爪状态在时间维度上整体呈现出连续且平滑的变化趋势，无明显异常跳变，表明数据采集过

程具有良好的稳定性。同时，通过对主从两侧关节状态的对比，可以观察到较高的一致性，验证了遥操作执行过程中的跟随准确性。

上述可视化结果主要用于对示范数据质量进行验证，包括轨迹连续性、动作变化合理性及多视角观测一致性等方面的检查，为后续策略训练与实验评估提供可靠的数据基础。

3.3 实验设计

3.3.1 模型结构与训练设置

本文基线策略参考ACT框架构建，采用CVAE-Transformer动作块生成模型。模型以多视角视觉观测和机器人本体状态为输入，并行预测未来固定长度H的动作序列。训练阶段采用CVAE目标函数优化动作重构损失与KL正则项；推理阶段丢弃CVAE encoder，仅保留策略decoder，并将潜变量设为先验分布均值$Z=0$进行确定性动作块预测。所有对比方法共享相同的模型结构、训练数据、固定预测长度、动作表示和时间集成机制，仅在推理阶段动作缓存刷新时的执行块长选择方式上不同。实验模型结构与训练参数设置见表3。

表3 实验模型结构与训练参数设置

项目	设置
基线策略模型	CVAE-Transformer
视觉编码器	ResNet18
动作块生成方式	Transformer decoder并行预测未来H步动作序列
固定预测长度H/步	100
CVAE encoder	Transformer encoder
CVAE decoder/policy	ResNet18 + Transformer encoder + Transformer decoder
encoder layers	4
decoder layers	7
推理阶段潜变量	Z=0
hidden dimension	512
Feedforward dimension	3 200
attention heads	8
Reconstruction loss	L1 loss
KL权重β	10
dropout	0.1
learning rate	1×10^{-5}
batch size	8

3.3.2 对比算法与基线模型

为保证实验对比的公平性与结论的可靠性，本文对训练阶段和推理阶段的基础设置进行严格控制。所有方法均基于相同的示范数据、模型结构及训练策略进行训练，并在推理阶段采用相同的固定预测长度H=100、相同的策略预测频率和相同的时间集成机制。不同方法仅在动作缓存刷新时的执行块长选择方式上存在差异，从而将性能差异主要归因于执行缓存长度调度策略本身，而非模型能力差异。

在该统一设置下，构建如下对比方法。

（1）Fixed-k=100：采用较大的固定执行块长，该方法在每次动作缓存刷新时固定写入100个控制步的动作，因此动作缓存刷新频率较低，动作子段切换较少，有利于保持稳定阶段的执行连续性；但在接触或对齐等关键阶段，最新观测对应的预测结果需要等待当前缓存执行完后才能写入新的动作缓存，因而对局部状态变化的响应相对较慢。

（2）Fixed-k=50：采用中等固定执行块长，该方法在动作连续性和执行动作刷新响应能力之间形成折中，用作中等执行粒度的固定块长基线。

（3）Fixed-k=10：采用较小的固定执行分块，该方法动作缓存刷新频率较高，能够使最新观测对应的预测结果更快作用于实际执行，具备较强的局部纠偏能力；但频繁的动作子段切换可能引入动作波动，降低执行平滑性。

（4）本文方法（Adaptive-k）：本文基于预测分歧的自适应动作分块方法。在候选集合$\{10, 50, 100\}$上进行在线调度，根据平滑预测分歧动态选择动作缓存刷新时的执行块长，从而在不同阶段实现动作连续性、关键阶段响应能力与轨迹稳定性之间的自适应平衡。

平台底层控制与遥操作采样频率为100 Hz，本文在数据预处理阶段将训练数据重采样至50 Hz，以与策略执行频率保持一致。执行块长的选取基于系统控制频率（50 Hz）与任务时间尺度进行设定。其中，k=10、50、100分别对应约0.2 s、1 s、2 s的执行时间窗口，覆盖了从快速刷新到连续执行的典型范围。

需要说明的是，所有方法在训练阶段均基于固定预测长度为H的动作块模型进行学习。在推理阶段，所有方法均以相同策略频率生成固定长度动作块，并采用相同时间集成机制。自适应策略仅在动作缓存刷新时根据平滑预测分歧选择执行块长，不访问未来信息，也不改变策略网络结构。与Fixed-k基线相比，其额外计算主要来自预测分歧计算、

EMA 平滑和阈值判断等动作序列后处理操作。

为保证实验结果的统计可靠性，本文参考机器人模仿学习相关工作中的评估方式，对不同策略进行多次独立执行测试。对于主实验和鲁棒性实验，每种方法在对应任务或扰动条件下累计获得 100 次独立执行结果，任务成功率按成功执行次数占总执行次数的比例统计；轨迹误差、执行平滑性等连续指标取多次实验结果的平均值。对于新增消融实验和超参数敏感性分析，采用相同统计方式，以降低单次执行偶然性对实验结论的影响。

此外，考虑到现有模仿学习方法在实际部署中通常采用固定长度的动作执行策略，本文以不同执行块长的固定策略作为主要对比基线，以系统性分析执行分块策略对任务性能的影响。

3.3.3　评价指标

为评估不同动作分块策略在任务完成能力、关键阶段执行精度和动作连续性方面的差异，本文采用任务成功率、轨迹误差和执行平滑性 3 个指标进行定量评价。其中，任务成功率（success rate，SR）用于衡量策略完成任务的整体可靠性；轨迹误差（trajectory error）用于衡量放置、对齐和插入等关键阶段的执行精度；执行平滑性（smoothness）用于衡量相邻控制步之间动作指令的连续性。

（1）任务成功率

任务成功率用于衡量策略完成任务的整体可靠性，定义为：

$$\mathrm{SR}=\frac{N_{\mathrm{success}}}{N_{\mathrm{total}}} \tag{12}$$

其中，N_{success} 表示成功完成任务的执行次数，N_{total} 表示总执行次数。本文在主实验和鲁棒性实验中均对每种方法进行 100 次独立执行测试，任务成功率按成功次数占总执行次数的比例统计。

针对不同任务，其成功判据定义如下。

①Pick & Place：目标物体被放置到指定目标区域内，释放后未发生明显滑移或掉落，并在目标区域内保持稳定。

②Insert/Align：目标部件完成对准插入，插入深度达到任务要求，且执行过程中未出现明显卡滞、脱落或姿态失稳。

（2）轨迹误差

轨迹误差用于衡量策略在闭环执行过程中的控制精度，反映不同动作分块策略在任务执行过程中产生的误差累积情况。轨迹误差定义为该次执行过程中任务误差的平均值：

$$e_{\mathrm{traj}}=\frac{1}{T}\sum_{t=1}^{T}e(t) \tag{13}$$

其中，T 表示一次独立闭环执行过程中参与统计的有效控制步数，$e(t)$ 表示时刻 t 的任务相关误差。针对不同任务，误差定义如下。

①Pick & Place：目标物体中心相对于目标放置区域中心的位置偏差。

②Insert/Align：末端执行器或目标部件相对于目标装配位置的对齐偏差或插入误差。

对于每种方法，本文分别计算多次独立执行实验中的轨迹误差，并取其平均值作为最终统计结果。为保证不同方法之间的可比性，所有对比方法均采用相同的误差计算方式、相同的有效统计范围和相同的统计单位。轨迹误差的统计单位为 m。该指标数值越小，表示执行轨迹与任务目标之间的偏差越小，策略的执行精度越高。

（3）执行平滑性

执行平滑性用于刻画策略在闭环执行过程中输出动作指令的连续性与稳定性，主要反映不同动作分块策略是否会在动作缓存刷新和动作子段切换过程中引入指令突变或控制抖动。本文比较的是不同执行块长调度策略对实际执行过程的影响，因此平滑性基于机器人执行过程中实际发送至底层控制器的动作指令序列计算。本文采用相邻控制步动作指令变化量的平方范数均值作为执行平滑性指标，计算式为：

$$S=\frac{1}{T-1}\sum_{t=1}^{T-1}\left\| a_{t+1}^{\mathrm{cmd}}-a_{t}^{\mathrm{cmd}} \right\|_2^2 \tag{14}$$

其中，T 表示一次独立闭环执行过程中的有效控制步数，a_t^{cmd} 表示时刻 t 实际发送至底层控制器的动作指令向量。该动作指令由时间集成后的动作缓存产生，与机器人实际执行过程一一对应。对于每次独立执行实验，本文根据完整执行过程中的动作指令序列计算平滑性，并对多次实验结果取平均值作为最终统计结果。

该指标数值越小，表示相邻控制步之间的动作指令变化越平缓，执行过程越稳定；数值越大，则说明动作指令变化越剧烈，执行过程中更容易出现抖动或不连续现象。为保证不同方法之间的可比性，所有方法均采用相同的动作表示、相同的控制频率和相同的统计方式来计算执行平滑性。

3.4 实验结果与分析

3.4.1 主实验结果与分析

在无外部扰动条件下，各方法在两类任务上的实验结果见表4。由表4可以看出，不同固定分块策略在执行精度与轨迹稳定性之间呈现出明显的权衡关系，而本文自适应分块方法在两类任务中均表现出更优的综合性能。

表4 主实验结果

任务	模型	成功率↑	轨迹误差↓	执行平滑性↓
Pick & Place	Fixed-k=100	74%	0.061	0.82
	Fixed-k=50	83%	0.038	1.36
	Fixed-k=10	89%	0.021	2.91
	本文（Adaptive-k）	93%	0.021	1.08
Insert/Align	Fixed-k=100	41%	0.078	0.95
	Fixed-k=50	58%	0.051	1.74
	Fixed-k=10	71%	0.030	3.48
	本文（Adaptive-k）	82%	0.027	1.41

Fixed-k=100由于采用较大的执行块长，在两类任务中均获得较低的平滑性（Pick & Place：0.82；Insert/Align：0.95），表明较低的动作缓存刷新频率和较少的动作子段切换有利于维持动作连续性。然而，缓存长度较大，最新观测对应的预测结果需要等待当前缓存执行完后才能写入新的动作缓存，因此在接触、对齐等关键阶段响应较慢，导致成功率与轨迹误差表现较差（例如在Insert/Align任务中成功率仅为41%，轨迹误差达到0.078）。

相比之下，Fixed-k=10通过更频繁地刷新动作缓存，使最新观测对应的预测结果更快作用于实际执行，因此在两类任务中均有效降低了轨迹误差并提升了任务成功率。但由于动作子段切换频率较高，其执行过程更容易出现动作波动，使平滑性指标明显增大（分别为2.91和3.48）。

Fixed-k=50在性能上介于上述两者之间，在一定程度上实现了精度与稳定性的折中，但仍难以兼顾不同任务阶段的动态需求。

相比之下，本文Adaptive-k方法在两类任务中均取得最高成功率（Pick & Place：93.0%；Insert/Align：82%），同时在轨迹误差上达到或接近最优水平（分别为0.021和0.027）。在平滑性方面，其表现虽略低于Fixed-k=100，但显著优于Fixed-k=10（分别为1.08和1.41），表明该方法在保证执行精度的同时，有效抑制了高频抖动现象。

从任务层面分析，Pick & Place任务以稳定推进为主，各方法之间性能差异相对有限；而在Insert/Align任务中，由于涉及接触与精细对齐，固定分块策略的局限性更加显著，Adaptive-k方法在成功率与轨迹误差方面均表现出更明显的优势。该方法能够根据任务阶段动态调整执行块长，在关键阶段缩短缓存长度以提高执行动作的刷新响应能力，在稳定阶段增大缓存长度以保持动作连续性，从而在任务成功率、轨迹误差和执行平滑性之间实现更优平衡。

本文方法在不改变模型结构与训练过程的前提下，通过推理阶段的自适应分块机制，显著提升了策略执行性能，验证了其在多阶段操作任务中的有效性与适用性。

3.4.2 鲁棒性实验与性能分析

为评估本文方法在工业场景不确定性条件下的稳定性与适应能力，本文在Pick & Place螺丝放置任务上进一步开展鲁棒性实验，并在实际执行过程中引入两类典型扰动：视觉扰动与几何扰动。其中，视觉扰动包括短时遮挡与光照变化，用于模拟感知不稳定性；几何扰动包括目标物体或初始位姿偏移，以及抓取、搬运或放置阶段的轻微外力干预，用于模拟实际操作中的位姿误差与环境交互干扰。

为便于分析，本文从可恢复性与恢复过程质量两个维度对系统鲁棒性进行评价。其中，可恢复性以任务成功率衡量；恢复过程质量通过轨迹误差与执行平滑性指标进行表征。

Pick & Place任务上的鲁棒性实验结果见表5。由表5可以看出，在引入扰动后，各方法的成功率均出现不同程度下降，同时轨迹误差增大、平滑性降低，表明扰动对任务执行过程产生显著影响。相比之下，本文Adaptive-k方法在所有扰动条件下均取得最高成功率，并在轨迹误差与平滑性之间保持较优平衡。

在视觉扰动条件下（遮挡与光照变化），Fixed-k=100由于动作缓存长度较大，最新观测对应的预测结果较晚作用于实际执行，对感知偏差的响应能力有限，其成功率分别下降至61%和66%，轨迹误差分别为0.078和0.071。Fixed-k=10虽在误差控制方面表现较优（0.036和0.032），但平滑性明显下降（3.05和2.98），表明频繁的动作缓存刷新和动作子段切换易引入动作波动。相比之下，Adaptive-k方法在两种视觉扰动下均取得最高成功率（86%和

表5　Pick & Place任务上的鲁棒性实验结果

扰动类型	扰动设置	方法	成功率	轨迹误差	执行平滑性
视觉扰动	遮挡(腕部/顶视)	Fixed-k=100	61%	0.078	0.88
		Fixed-k=50	72%	0.052	1.42
		Fixed-k=10	78%	0.036	3.05
		本文Adaptive-k	86%	0.031	1.26
	光照变化	Fixed-k=100	66%	0.071	0.85
		Fixed-k=50	76%	0.047	1.39
		Fixed-k=10	82%	0.032	2.98
		本文Adaptive-k	89%	0.028	1.21
几何扰动	初始位姿偏移	Fixed-k=100	48%	0.089	0.93
		Fixed-k=50	63%	0.061	1.68
		Fixed-k=10	74%	0.041	3.32
		本文Adaptive-k	85%	0.034	1.47
	接触外力扰动	Fixed-k=100	42%	0.095	0.97
		Fixed-k=50	57%	0.067	1.82
		Fixed-k=10	69%	0.046	3.55
		本文Adaptive-k	81%	0.038	1.58

89%），同时保持较低轨迹误差（0.031和0.028）及较好的平滑性（1.26和1.21），体现出较强的感知适应能力。

在几何扰动条件下（初始位姿偏移与接触外力干预），各方法性能差异进一步扩大。Fixed-k=100由于缺乏有效的实时纠偏能力，其成功率分别降至48%和42%，轨迹误差分别达到0.089和0.095；Fixed-k=10虽能够降低轨迹误差（0.041和0.046）并提高成功率（74%和69%），但其平滑性显著下降（3.32和3.55），表现出明显振荡特性。相比之下，Adaptive-k方法在两类几何扰动下均取得最高成功率（85%和81%），并在轨迹误差（0.034和0.038）与平滑性（1.47和1.58）之间保持较优折中，表明其在强扰动条件下具有更稳定的执行表现。

综上，固定分块策略在扰动条件下难以同时兼顾误差控制与动作稳定性：较大执行块长虽有利于提高轨迹平滑性，但动作缓存长度较大，最新观测对应的预测结果较晚作用于实际执行，纠偏能力不足；较小执行块长虽提升了响应能力，但频繁的动作缓存刷新和动作子段切换易引入动作波动。本文Adaptive-k方法能够根据任务状态变化动态调整执行块长，在扰动阶段提高纠偏能力，在稳定阶段维持轨迹平滑性，从而在复杂环境条件下实现更优的综合性能。

3.4.3　关键阶段分析

机械臂操作任务执行过程示意如图5所示，图5（a）为单臂抓取放置任务，图5（b）为双臂抓取−传递−对准−插入任务。



（a）单臂抓取放置任务



（b）双臂抓取−传递−对准−插入任务

图5　机械臂操作任务执行过程示意

不同任务中执行块长的阶段性变化对比如图6所示，其中图6（a）对应单臂抓取放置任务，图6（b）对应抓取–传递–对准–插入任务。由图6可以看出，k在任务执行过程中呈现出显著的阶段相关性：在自由运动阶段维持较大取值，在接触、传递及对齐等关键操作阶段迅速收缩，并在关键阶段结束后恢复至较大水平。

该变化规律在两类任务中具有一致性，但在图6（b）所示任务中，小分块区间持续时间更长。这是由于该任务包含双臂传递与精对齐插入两个关键阶段，对控制精度与执行动作刷新响应能力的要求更高，因此需要在更长时间范围内维持较小执行块长以保证操作稳定性。

从机理上分析，执行块长的动态调度反映了执行粒度、动作连续性与局部纠偏能力之间的内在权衡关系。在自由运动阶段，系统状态变化相对平缓且可预测性较强，采用较大的执行块长能够有效降低动作缓存刷新频率和动作子段切换频率，从而保持动作连续性；而在接触与对齐等误差敏感阶段，减小执行块长能够提高动作缓存刷新频率，使最新观测对应的预测结果更快写入动作缓存并作用于实际执行，从而增强局部调整能力并降低误差累积风险。

在关键操作阶段，不同策略之间的性能差异更为显著。固定大分块策略动作缓存较长，最新预测结果难以及时作用于实际执行，易导致误差累积；固定小分块策略虽提高了响应能力，但由于动作子段切换频繁，容易引入控制振荡并降低执行稳定性。相比之下，本文方法能够根据任务阶段自适应调整执行块长，并结合时间集成机制对相邻预测结果进行平滑处理，使系统在保证关键阶段刷新响应能力的同时维持连续稳定的控制行为，从而在接触与插入过程中实现更可靠的对齐与执行。

3.5　消融实验与超参数分析

为进一步验证本文方法各组成模块的有效性，在 Insert/Align 双臂轴承插入任务上开展消融实验与超参数分析。该任务包含抓取、传递、对准和插入等多个连续阶段，既涉及自由空间中的稳定运动，也涉及接触约束下的精细对齐与插入过程，对执行块长调度的阶段适应性和关键阶段执行响应能力要求较高。因此，相较于 Pick & Place 任务，Insert/Align 任务更适合用于分析轨迹预测分歧、时间平滑和分段阈值调度机制对系统性能的影响。

本文方法主要由轨迹级预测分歧度量、指数滑动平均时间平滑以及分段阈值调度 3 个部分组成。为分析各模块的独立贡献，构建如下消融版本。

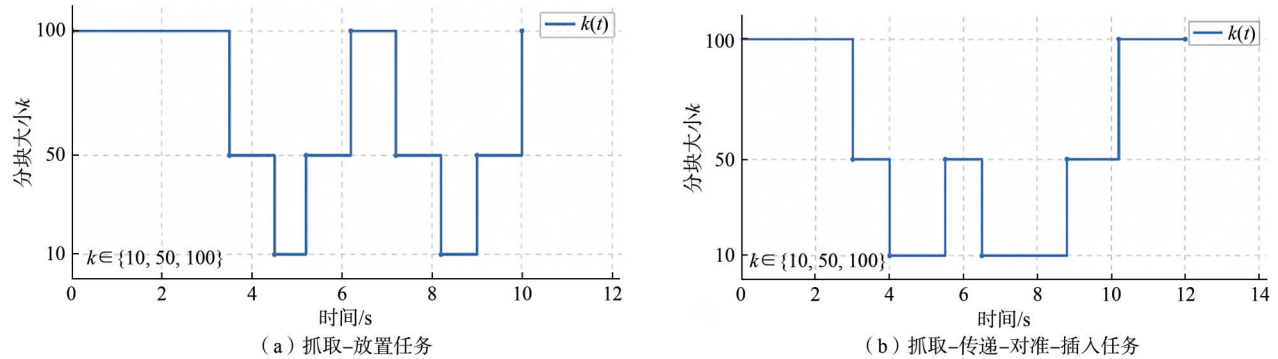
（1）w/o trajectory-level discrepancy：去除轨迹级预测分歧度量，仅比较相邻预测块在重叠区间内首个对应动作的差异。该设置用于验证多步重叠区间信息相较于单步动作差异在任务阶段判别中的作用。

（2）w/o EMA smoothing：去除指数滑动平均平滑，直接使用原始预测分歧进行阈值判断和块长调度。该设置用于验证时间平滑在抑制瞬时预测波动、减少频繁切换以及提升执行稳定性方面的作用。

（3）w/o threshold scheduling：去除分段阈值调度机制，不再根据低、中、高分歧区间分别选择不同执行块长，而采用固定中等执行块长 k=50。该设置用于验证分段阈值调度在不同任务阶段执行粒度分配中的必要性。

（4）Full Adaptive-k：本文完整方法，同时保留轨迹级预测分歧、指数滑动平均平滑和分段阈值调度机制，并在候选执行块长集合{10, 50, 100}中进行在线选择。

Insert/Align 任务上的消融实验结果见表 6。除



（a）抓取–放置任务　　（b）抓取–传递–对准–插入任务

图6　不同任务中执行块长的阶段性变化对比

任务成功率、轨迹误差和执行平滑性外，本文进一步统计块长切换次数，以衡量调度过程的稳定性。较高的块长切换次数通常表明调度器对短时噪声较为敏感，容易在阈值附近发生频繁跳变，从而引入额外控制抖动；而过低的切换次数则可能说明调度器对任务阶段变化的响应不足。因此，块长切换次数能够从执行节奏稳定性的角度补充评价自适应调度机制的有效性。

表6　Insert/Align任务上的消融实验结果

方法	成功率	轨迹误差	执行平滑性	块长切换次数
w/o trajectory-level discrepancy	74%	0.034	1.76	13.8
w/o EMA smoothing	76%	0.032	2.23	21.5
w/o threshold scheduling	58%	0.051	1.74	0
Full Adaptive-k	82%	0.027	1.41	11.6

由表6可以看出，完整Adaptive-k方法在任务成功率、轨迹误差、执行平滑性及块长切换次数等指标之间取得了较优的综合性能。将轨迹级预测分歧替换为单步动作差异后，任务成功率由82%降至74%，轨迹误差由0.027增至0.034。该结果表明，单步动作差异易受局部预测噪声或个别动作维度突变的影响，对任务阶段变化的表征稳定性不足。相比之下，轨迹级预测分歧能够综合利用重叠区间内的多步动作信息，更稳定地刻画模型对未来局部轨迹的自一致性变化，因此更适合作为推理阶段执行块长调度的判据。

去除EMA时间平滑后，方法仍可依据预测分歧进行自适应调度，但执行平滑性由1.41下降至2.23，块长切换次数由11.6增加至21.5。该现象说明，在真实机器人执行过程中，视觉观测噪声、状态估计误差以及模型瞬时预测波动均可能导致原始分歧信号在阈值附近产生短时振荡，从而引发执行块长的频繁切换。引入指数滑动平均后，局部随机波动得到有效抑制，调度器能够更多地依据任务阶段变化的整体趋势进行决策，从而提升执行过程的连续性与稳定性。

去除分段阈值调度后，方法退化为固定中等执行块长$k=50$，因此其结果与主实验中的Fixed-k=50一致，块长切换次数为0。与完整Adaptive-k相比，该设置成功率由82%降至58%，轨迹误差由0.027增至0.051，说明固定中等执行粒度难以同时适应自由运动、双臂传递、精密对准和接触插入等不同阶段的控制需求。仅采用固定中等分块时，在稳定推进阶段无法充分利用较长缓存带来的动作连续性，在关键接触阶段又无法像短缓存一样快速刷新实际执行动作，因此容易出现阶段适应不足。

为进一步分析平滑系数对方法性能的影响，本文对指数滑动平均中的平滑系数α进行敏感性分析，结果见表7。平滑系数α用于调节当前分歧观测与历史平滑分歧之间的权重关系。当α取值较小时，平滑分歧更多依赖历史信息，有助于减少块长频繁切换，但可能降低调度器对任务阶段突变的响应速度；当α取值较大时，调度器对当前分歧变化更为敏感，但也更容易受到瞬时噪声影响，导致执行块长频繁变化并降低动作连续性。

表7　平滑系数敏感性分析

平滑系数α	成功率	轨迹误差	执行平滑性	块长切换次数
0.1	76%	0.033	1.35	7.4
0.3	82%	0.027	1.41	11.6
0.5	80%	0.029	1.62	15.3
0.7	77%	0.032	1.98	19.8

由表7可以看出，平滑系数α对调度器的响应特性和执行稳定性具有显著影响。当$\alpha=0.1$时，平滑分歧更多依赖历史信息，因而块长切换次数较少，执行轨迹也相对平滑。然而，过强的历史平滑会降低调度器对当前状态变化的响应速度，使系统在进入接触、对准等误差敏感阶段时可能无法及时切换至较小执行块长，从而导致任务成功率和轨迹误差指标下降。当$\alpha=0.7$时，调度器对当前分歧变化更加敏感，能够更快响应局部状态变化，但同时也更容易受到视觉噪声、状态估计误差和模型瞬时预测波动的影响，导致块长切换次数显著增加，并降低执行平滑性。相比之下，$\alpha=0.3$在阶段响应速度与调度稳定性之间取得了较优折中，使方法在任务成功率、轨迹误差、执行平滑性和块长切换次数等指标上表现更为均衡。因此，本文在后续实验中采用$\alpha=0.3$作为默认设置。

综上，消融实验与超参数分析表明，本文方法的性能提升来源于轨迹级预测分歧、EMA平滑和分段阈值调度的共同作用。轨迹级预测分歧为任务阶段变化提供了更稳定的自一致性判据，EMA平滑增强了调度过程对噪声和瞬时预测波动的鲁棒性，分段阈值调度则实现了不同任务阶段下执行粒度的动

态分配。轨迹级预测分歧、EMA平滑和分段阈值调度会引入一定的动作序列后处理开销，但不增加额外模型参数、离线训练成本或策略网络查询频率。因此，该方法能够在保持统一推理频率的条件下，通过自适应缓存长度调度，提升固定动作分块策略在长时程、多阶段精细操作任务中的执行性能。

4 结束语

本文围绕长时程、多阶段具身操作中的执行粒度调度问题开展研究。实验分析表明，固定执行块长难以同时适应不同任务阶段的控制需求，执行块长的选择本质上需要在动作连续性与局部响应能力之间进行动态权衡。稳定运动阶段更适合采用较大的执行粒度，以减少动作子段切换并保持执行连续性；而在接触、传递、对齐和插入等误差敏感阶段，较小的执行粒度能够使最新观测对应的预测结果更快作用于实际执行，从而增强局部调整能力。

进一步分析表明，轨迹预测分歧能够反映模型对未来局部轨迹预测的一致性变化，并可作为任务阶段变化及执行粒度调整的有效判据。相比单步动作差异，利用重叠区间内多步信息构造的轨迹级分歧具有更好的稳定性；时间平滑能够抑制瞬时预测波动引起的频繁切换；分段阈值调度则使不同分歧水平能够对应不同执行粒度。消融实验与参数分析进一步说明，上述机制共同决定了调度过程的阶段适应性与执行稳定性。

从整体上看，本文研究在保持策略模型及训练过程不变的情况下，对推理阶段的动作执行方式进行合理设计，同样可以改善长时程机械臂操作中的执行表现。这为现有动作块策略提供了一种从“动作生成”之外进一步优化“动作执行”的思路，也为多阶段具身操作中的动态执行节奏设计提供了参考。

未来工作将进一步引入力觉等多模态信息，与现有视觉观测形成互补，以提高接触阶段的状态感知与阶段判别能力；同时，可研究连续化的执行块长调节机制以及更加自适应的阈值确定方法，降低人工参数设定对不同任务和场景的依赖。

参考文献：

[1] Jiang Y Z, Huang Z Z, Yang B, et al. A review of robotic assembly strategies for the full operation procedure: planning, execution and evaluation[J]. Robotics and Computer-Integrated Manufacturing, 2022, 78: 102366.

[2] Abbas M, Narayan J, Dwivedy S K. A systematic review on cooperative dual-arm manipulators: modeling, planning, control, and vision strategies[J]. International Journal of Intelligent Robotics and Applications, 2023, 7(4): 683-707.

[3] Argall B D, Chernova S, Veloso M, et al. A survey of robot learning from demonstration[J]. Robotics and Autonomous Systems, 2009, 57(5): 469-483.

[4] Osa T, Pajarinen J, Neumann G, et al. An algorithmic perspective on imitation learning[J]. Foundations and Trends® in Robotics, 2018, 7(1/2): 1-179.

[5] 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 C, Bai W S, Du X, et al. Survey of imitation learning: tradition and new advances[J]. Journal of Image and Graphics, 2023, 28(6): 1585-1607.

[6] Ross S, Gordon G, Bagnell D. A reduction of imitation learning and structured prediction to no-regret online learning[C]//Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, 2011-04-11/13, Fort Lauderdale, USA. New York: PMLR, 2011: 627-635.

[7] Ho J, Ermon S. Generative adversarial imitation learning[C]//Proceedings of the 30th International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2016: 4565-4573.

[8] Brohan A, Brown N, Carbajal J, et al. RT-1: robotics transformer for real-world control at scale[C]//Proceedings of the Robotics: Science and Systems XIX. Robotics: Science and Systems Foundation, 2023: 1-31.

[9] Chi C, Xu Z J, Feng S Y, et al. Diffusion policy: visuomotor policy learning via action diffusion[J]. The International Journal of Robotics Research, 2025, 44(10/11): 1684-1704.

[10] Zhao T, Kumar V, Levine S, et al. Learning fine-grained bimanual manipulation with low-cost hardware[C]//Proceedings of the Robotics: Science and Systems XIX. Robotics: Science and Systems Foundation, 2023: 1-18.

[11] Lee A C-W, Chuang I, Chen L-Y, et al. InterACT: inter-dependency aware action chunking with hierarchical attention transformers for bimanual manipulation[C]//Proceedings of the 8th Conference on Robot Learning. New York: PMLR, 2025: 1730-1743.

[12] Buamanee T, Kobayashi M, Uranishi Y, et al. Bi-ACT: bilateral control-based imitation learning via action chunking with transformer[C]//Proceedings of the 2024 IEEE International Conference on Advanced Intelligent Mechatronics (AIM). Piscataway: IEEE Press, 2024: 410-415.

[13] Liu Y, Hamid J I, Xie A, et al. Bidirectional decoding: improving action chunking via guided test-time sampling[C]//Proceedings of the 13th International Conference on Learning Representations (ICLR 2025). Vancouver: ICLR, 2025: 72787-72820.

[14] Hogan N. Impedance control: an approach to manipulation: part I: theory[J]. Journal of Dynamic Systems, Measurement, and Control, 1985, 107(1): 1-7.

[15] Raibert M H, Craig J J. Hybrid position/force control of manipulators[J]. Journal of Dynamic Systems, Measurement, and Control, 1981, 103(2): 126-133.

[16] Zhao T Z, Luo J L, Sushkov O, et al. Offline meta-reinforcement learning for industrial insertion[C]//Proceedings of the 2022 International Conference on Robotics and Automation (ICRA). Piscataway: IEEE Press, 2022: 6386-6393.

[17] Nair A, Zhu B, Narayanan G, et al. Learning on the job: self-rewarding offline-to-online finetuning for industrial insertion of novel connectors from vision[C]//Proceedings of the 2023 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE Press, 2023: 7154-7161.

[18] Jin P P, Lin Y J, Song Y X, et al. Vision-force-fused curriculum learning

for robotic contact-rich assembly tasks[J]. Frontiers in Neurorobotics, 2023, 17: 1280773.
[19] Wang C, Su C P, Sun B Z, et al. Extended residual learning with one-shot imitation learning for robotic assembly in semi-structured environment[J]. Frontiers in Neurorobotics, 2024, 18: 1355170.
[20] Smith C, Karayiannidis Y, Nalpantidis L, et al. Dual arm manipulation: a survey[J]. Robotics and Autonomous Systems, 2012, 60(10): 1340-1353.
[21] Ren Y, Zhou Z H, Xu Z W, et al. Enabling versatility and dexterity of the dual-arm manipulators: a general framework toward universal cooperative manipulation[J]. IEEE Transactions on Robotics, 2024, 40: 2024-2045.
[22] Kim H, Ohmura Y, Kuniyoshi Y. Transformer-based deep imitation learning for dual-arm robot manipulation[C]//Proceedings of the 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Piscataway: IEEE Press, 2021: 8965-8972.
[23] Ravichandar H, Polydoros A S, Chernova S, et al. Recent advances in robot learning from demonstration[J]. Annual Review of Control, Robotics, and Autonomous Systems, 2020, 3: 297-330.
[24] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]// Advances in Neural Information Processing Systems 30. Red Hook: Curran Associates, Inc., 2017: 5998-6008.
[25] Zeng A, Florence P, Tompson J, et al. Transporter networks: rearranging the visual world for robotic manipulation[C]//Proceedings of the 2020 Conference on Robot Learning. New York: PMLR, 2021: 726-747.
[26] Shridhar M, Manuelli L, Fox D. CLIPort: what and where pathways for robotic manipulation[C]//Proceedings of the 5th Conference on Robot Learning. New York: PMLR, 2022: 894-906.
[27] SHRIDHAR M, MANUELLI L, FOX D. Perceiver-actor: a multi-task transformer for robotic manipulation[C]//Proceedings of The 6th Conference on Robot Learning. New York: PMLR, 2023: 785-799.
[28] Belkhale S, Cui Y, Sadigh D. Hydra: hybrid robot actions for imitation learning[C]//Proceedings of The 7th Conference on Robot Learning. New York: PMLR, 2023: 2113-2133.
[29] Sohn K, Yan X C, Lee H. Learning structured output representation using deep conditional generative models[C]//Proceedings of the 29th International Conference on Neural Information Processing Systems - Volume 2. New York: ACM Press, 2015: 3483-3491.
[30] Chen Y Z, Merwe M V D, Sipos A, et al. Visuo-tactile transformers for manipulation[C]//Proceedings of The 6th Conference on Robot Learning. New York: PMLR, 2023: 2026-2040.
[31] Whitney D E. Quasi-static assembly of compliantly supported rigid parts[J]. Journal of Dynamic Systems, Measurement, and Control, 1982, 104(1): 65-77.
[32] Grannen J, Wu Y, Vu B, et al. Stabilize to act: learning to coordinate for bimanual manipulation[C]//Proceedings of the 7th Conference on Robot Learning. New York: PMLR, 2023: 563-576.

[作者简介]



郑姗（2004– ），女，北京化工大学信息科学与技术学院本科生，主要研究方向为机器人与人工智能。



杨子仪（2005– ），男，北京化工大学信息科学与技术学院本科生，主要研究方向为机器人与人工智能。



刘宜坤（1997– ），女，北京化工大学信息科学与技术学院博士生，主要研究方向为具身智能感知与决策。



周翰（1983– ），男，北京邮电大学电子工程学院博士生，主要研究方向为具身智能感知与操作。



王刚（2001– ），男，北京具身智能科技有限公司技术总监，主要研究方向为具身智能与机器人。



吴清华（2001– ），男，北京具身智能科技有限公司首席运营官，主要研究方向为机器人与人工智能。



沈甜雨（1996– ），女，博士，北京化工大学信息科学与技术学院副教授，主要研究方向为机器视觉、智能感知与机器人系统。



李志伟（1990– ），男，博士，北京化工大学信息科学与技术学院副教授，主要研究方向为具身智能感知、操作与学习。