

数据驱动下的机器人轴孔装配研究综述

王战玺, 张邦海, 李景德, 罗子彦, 郑晨

(西北工业大学机电学院, 西安 710072)

摘要: 轴孔装配作为制造业核心工艺,其智能化对产品精度与可靠性至关重要。传统人工算法存在模型依赖强、自适应不足等问题,而数据驱动方法通过从数据中隐式学习响应模式,展现出良好的泛化性能。本文系统综述了基于数据驱动的轴孔装配技术,从环境感知、装配控制及任务课程设计梳理进展。分析了智能体在轴孔装配中的感知、任务理解与控制以及课程设计对任务的影响,并针对现有研究在动态响应、鲁棒性及任务理解等方面的瓶颈,展望未来研究需聚焦以任务需求为引导的低成本感知、先验知识+数据驱动的决策框架、未知任务空间的性能评估,以及探索人机共融下的任务理解与控制。

关键词: 轴孔装配;数据驱动;任务理解;多模态感知;课程设计

中图分类号: TP24

文献标志码: A

文章编号: 1005-2615(2025)03-0397-15

Data-Driven Robotic Shaft-Hole Assembly: A Review

WANG Zhanxi, ZHANG Banghai, LI Jingde, LUO Ziyan, ZHENG Chen

(School of Mechanical Engineering, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: Shaft-hole assembly, as a core manufacturing process, critically determines product precision and reliability through intelligent advancement. Traditional manual algorithms suffer from strong model dependency and insufficient adaptability, whereas data-driven methods exhibit promising generalization capabilities by implicitly learning response patterns from operational data. This paper systematically reviews data-driven shaft-hole assembly technologies, organizing advancements across three dimensions: Environmental perception, assembly control, and task curriculum design. It analyzes agent perception mechanisms, task comprehension-control interactions, and curriculum design impacts in shaft-hole assembly. Addressing current bottlenecks in dynamic response, robustness, and task understanding, the study proposes future research directions: Task-oriented low-cost perception systems, prior knowledge-enhanced data-driven decision frameworks, performance evaluation in unknown task spaces, and task comprehension-control strategies under human-robot collaboration.

Keywords: peg-in-hole assembly; data-driven; task comprehension; multimodal perception; curriculum design

装配是根据工艺要求将多个零部件接合为新部件或产品的过程,通常处于整个产品生产流程的末端,因此装配工序不仅影响整个产品的精度、寿命和可靠性等关键指标,失败时还可能会导致零部件之前所有生产过程的失效。在所有的装配工艺中,小到钢丝螺套的安装^[1]、螺栓的安装^[2],大到航

空航天工业^[3]、汽车生产^[4]的装配,其中有超过40%的装配工作都可以被视为轴孔装配作业,该类装配包含了装配工作中的大多数典型特征^[5]。

轴孔装配过程通常被划分为3个阶段,包括到达、对齐和插入^[6]。到达和对齐阶段要求机器人可以安全地到达孔的正上方,并将轴和孔的位

基金项目: 航空科学基金(2024Z072053001);陕西省重点研发计划项目(2023ZDLGY-40);陕西省重点研发计划项目:(2022ZDLGY03-06)。

收稿日期: 2025-03-30; **修订日期:** 2025-05-20

通信作者: 郑晨,男,博士,副教授, E-mail: chen.zheng@nwpu.edu.cn。

引用格式: 王战玺,张邦海,李景德,等. 数据驱动下的机器人轴孔装配研究综述[J]. 南京航空航天大学学报(自然科学版), 2025, 57(3): 397-411. WANG Zhanxi, ZHANG Banghai, LI Jingde, et al. Data-driven robotic shaft-hole assembly: A review[J]. Journal of Nanjing University of Aeronautics & Astronautics(Natural Science Edition), 2025, 57(3): 397-411.

姿进行对齐,插孔阶段则需要机器人高效且低接触力地将轴插入孔内^[7]。轴孔装配方法可分为基于传统的人工算法和基于数据驱动的方法两类。到达和对齐阶段可被分解为机器人路径规划,即机器人末端需要从任意起始点到达目标位姿,传统的人工算法中,路径规划算法包括经典方法^[8]、启发式方法^[9]、进化算法^[10]、混合方法^[11]以及机器学习方法^[12]。插孔阶段属于接触式任务,且插入过程中轴和孔的接触状态变化复杂,控制算法的两个核心性能指标为完成任务所需时间和任务过程中轴孔接触力的最大值,因此装配控制算法的核心目的在于控制接触力,控制接触力的方法有被动式和主动式两种^[13],被动式方法通过设计柔顺性的末端降低接触力,传统的主动柔顺控制算法通常将机器人系统看作质量-弹簧-阻尼二阶线性系统,通过调整机器人与外界交互过程中的阻抗参数来降低接触力/力矩^[14]。具体方法包括阻抗控制、导纳控制、力/位混合控制,以及可消除稳态误差的自适应阻抗控制。总体而言,传统的人工算法中,路径规划算法具有一定的适应性,但当存在多个障碍物、轴孔形状不规则、装配空间狭窄、视觉遮挡等情况时,人工算法实现难度大,难

以保证路径的最优性和高效性^[15]。装配控制算法则面临依赖精确的数学模型,在装配任务或条件发生变化时缺乏自主决策能力,自适应能力有限。相较而言,基于数据驱动的控制方式具有适应性强、鲁棒性好、可自我改进装配策略、无需精确建模等优点,可将一个实用策略泛化到从未见过的任务或场景中。同时也需注意该方式在数据量、计算资源以及实时性方面的挑战,以及模型决策过程的可解释性差等特点。

图 1 中的柱状图为 Web of Science 中机器学习相关论文的收录数量,并展示了机器学习的发展历程和每次关键节点,从图 1 中的趋势可知,机器学习的相关研究论文呈指数式递增。随着深度网络在视觉识别、翻译和语言等领域中得到快速发展,基于数据驱动的机器人控制技术也迎来井喷式发展。基于数据样本训练智能体,由智能体理解人类的意图,再进行机器人控制的方法已经表现出良好的泛化性能^[16]。本文将系统地分析轴孔装配基本原理、总结智能体训练流程,并围绕智能体现今突出的样本效率问题,从环境感知、装配控制以及任务认知方面讨论智能体进行轴孔装配的研究现状和发展趋势。

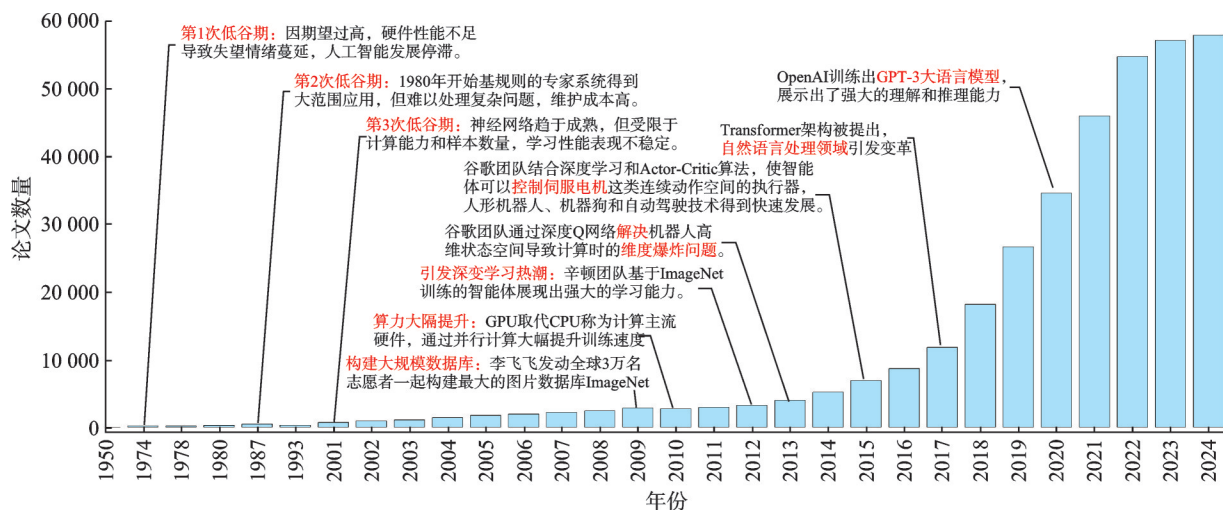


图 1 Web of Science 关于机器学习的论文数量

Fig.1 Publications on machine learning in Web of Science

1 数据驱动的轴孔装配

1.1 轴孔装配

1.1.1 装配阶段分析

机器人轴孔装配过程到达、对齐、插入分别如图 2(a~c)所示。首先,机器人需要夹持轴从起始位置运动到待插入孔的正上方,这一阶段需要克服环境不确定性和轴孔相对位置观测误差,轴孔相对位置观测为机器人提供该阶段的目标位置,通常会因为工业现场的光照变化,孔特征遮挡、加工误差

等因素出现观测误差^[17],使目标位置出现偏差,这些环境不确定性给机器人在起始位置和目标位置之间搜寻出一条安全且高效的运动轨迹增加难度。其次,机器人需要将轴的位姿和孔的位姿进行初步对齐^[7],通常这一阶段采取非接触对齐以防止对齐过程中对轴或孔造成损伤,其挑战主要在于如何获取高精度的轴孔相对位姿^[18],以消除两者的位姿偏差。最后,在插孔阶段,在机器人控制轴与孔保持低接触力/力矩的情况下,快速完成插入,该

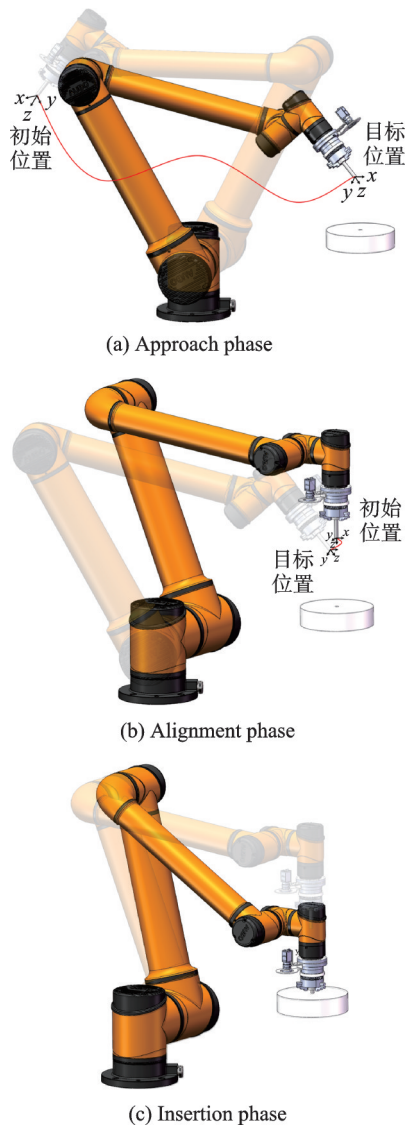


图2 机器人轴孔装配

Fig.2 Robotic shaft-hole assembly

阶段是轴孔装配的核心,由于插入过程中,机器人需同时观测轴孔的相对位姿状态和接触力/力矩,因此难以仅靠单一的视觉或力觉模态完成插入任务^[19-20]。

1.1.2 插孔阶段分析

轴孔装配的插孔阶段过程中,机器人需要减少轴孔接触力/力矩,因此本文针对单轴孔装配的插孔阶段进行接触模型分析。在插孔阶段,影响装配精度的因素较多,包括待装配轴孔的位姿偏差、配合公差、加工误差和粗糙度等,因此需简化插孔阶段,便于对轴孔接触状态进一步分析。由于单轴孔具有三维对称性,因此可将三维接触模型简化为二维平面模型,如图3所示,并建立以下假设条件:(1)轴孔为刚性物体,且横向和扭转刚度互相独立;(2)接触分析时整个插孔过程将被视为准静态平衡,忽略惯性力影响;(3) yoz 平面和 xoz 平面具有对称性;(4)插孔过程仅机器人端工件发生弹性形

变。插孔阶段可细分为单点接触、两点接触和完全接触^[21]。

力主动柔顺控制中,柔顺中心的位置选择将会直接影响力柔顺控制精度和效果,因该位置的受力情况是控制系统的关键输入量。通常该位置选择机器人末端六维力传感器端面的圆心点,该点的受力/力矩可由传感器直接获得。柔顺控制时,机器人主要调整柔顺中心相对于孔中心线的位置和轴孔中心线夹角来减少轴孔的接触力/力矩。

如图3(a)所示,轴孔的接触点处存在摩擦力 f 和支持力 F_N ,其中 $f = \mu F_N$, μ 为摩擦力因数。则对于单点接触阶段,末端柔顺中心 P 在 yoz 面上的受力可以表示为

$$\begin{cases} F_z = f \\ F_y = F_N \\ T_x = F_N(l - l_s) - \mu F_N r_P \end{cases} \quad (1)$$

式中: l 为柔顺中心与轴底沿轴中心线的距离, r_P 为轴的半径。 xoz 面受力分析类似,此处不再赘述。

随着插孔的继续,插孔过程将由单点接触过渡到两点接触,如图3(b)所示,两个接触点均出现摩擦力和支持力。此时,柔顺中心点 P 与孔中心轴线之间的距离 d 和轴孔中心线夹角 α 之间的约束关系可以表示为

$$\begin{cases} 2r_s = (l_s + 2r_P) \sin \alpha \\ d = l_s \tan \alpha \end{cases} \quad (2)$$

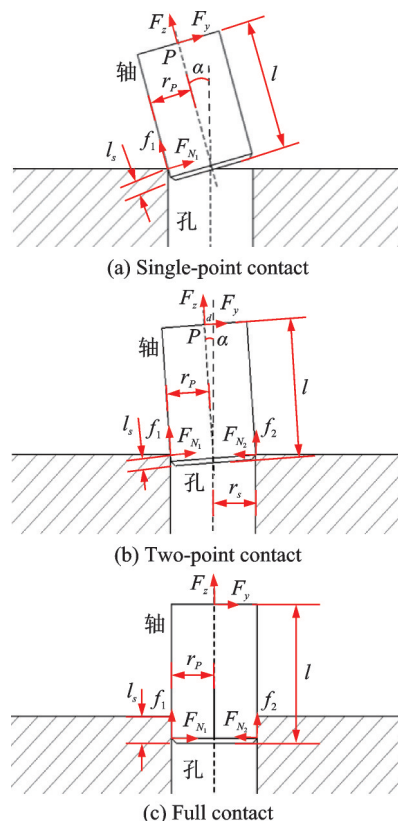


图3 插孔阶段受力分析

Fig.3 Force analysis during peg-in-hole insertion phase

在两点接触中,轴孔之间最易卡死,为避免卡死,各参数之间需遵循如下约束关系

$$\sin \alpha < \frac{r_s - r_p}{\mu r_p} \quad (3)$$

两点接触中柔顺中心在 yoz 面的受力可表示为

$$\begin{cases} F_z = f_1 + f_2 + F_{N_2} \sin \alpha \\ F_y = F_{N_1} - F_{N_2} \cos \alpha + f_2 \sin \alpha \\ T_x = F_{N_1}(l - l_s) - l(F_{N_2} \cos \alpha - f_2 \sin \alpha) + \\ \quad r_p(F_{N_2} \sin \alpha + f_2 \cos \alpha) - f_1 r_p \end{cases} \quad (4)$$

完全接触阶段如图 3(c) 所示,此时轴孔的接触形式由点接触变为面接触,此时轴孔的中心线重合,轴孔之间无力矩作用,只有摩擦力 $f_1 + f_2$ 和对向的支持力 $F_{N_1} - F_{N_2}$ 。但需要注意机器人自身进度限制导致的接触力影响:若采用规划机器人末端路径点,求逆解的方式控制机器人进行轴孔装配时,在插孔最后阶段可能会出现高频颤振导致的控制系统失稳。这是因为当机器人走直线时,从宏观角度来看机器人的末端轨迹是一条直线,但是在运动最小分辨率时,从微观角度观测来看,机器人末端轨迹是一个微小的圆弧。这个现象在其他非接触任务或接触力要求不高的时候无关紧要,但在轴孔的插孔阶段后期,当轴深入孔内,轴孔间隙极小或过盈装配时,机器人的轴向运动就会导致轴孔之间的高频振动^[22],进而使控制系统失稳。

1.2 基于数据驱动的运动控制

传统的轴孔装配方法通常基于运动学和接触动力学解析模型,分别对各阶段的任务需求构建控制逻辑。基于数据驱动的机器人轴孔装配方法不同于传统的显式物理建模,而是由智能体从大量交互数据中隐式学习不同状态的响应模式,以完成人类所需的任务意图。这种方式可显著降低任务过程对先验知识的依赖。通常,在训练完成后,即使是构建者本人,也难以完全理解智能体如何实现从输入的当前环境状态到输出下一状态所需动作。

2012 年辛顿团队基于李飞飞构建的 ImageNet 数据库,使用深度卷积神经网络(Convolutional neural network, CNN)突破传统视觉识别方法的性能瓶颈后,基于数据驱动的智能体应用迎来了快速发展。在训练智能体进行机器人运动控制的机器学习方法中,有监督学习^[23]、无监督学习^[24]、强化学习^[25]和模仿学习^[26]等各种学习方式。其中,机器人运动控制中强化学习被广泛应用。2013 年谷歌的 DeepMind 团队提出深度 Q 网络(Deep Q-network, DQN)^[27],解决了 Q 学习无法处理高维状态空间的问题,2014 年将策略梯度与深度学习结合,提出了确定性策略梯度(Deterministic policy gradient, DPG)算法^[28],使智能体可直接控制伺服

电机这类具有连续动作空间的执行器。随后,在基于值函数的 DQN 算法上,出现了双深度 Q 网络(Double deep Q-network, DDQN)、Dueling DQN^[29]、连续动作深度 Q 网络(Continuous deep Q-network, CDQN)^[30]等,在基于策略梯度的 DPG 算法基础上,出现了近端策略优化(Proximal policy optimization, PPO)^[31]、任域策略优化(Trust region policy optimization, TRPO)^[32]、双延迟深度确定性策略梯度(Twin delayed deep deterministic policy gradient, TD3)^[33]、柔性演员-评论家算法(Soft actor-critic, SAC)^[34]等训练方法。基于这些方法,基于数据驱动的机器人、汽车等控制方法迎来了井喷式发展。

然而,在训练智能体进行轴孔装配这类较复杂的任务时,还存在以下问题:(1)智能体难以高精度观测轴孔的相对位姿,以及轴孔接触力/力矩的大小及对应位置;(2)智能体难以高效地在各种场合中快速理解人类的装配任务意图,并从交互样本中学习完成所需任务所需的装配技能;(3)智能体在面对复杂装配任务时,训练无法收敛或样本效率低下。这 3 个问题分别对应 20 个院所 2025 年联名发表的智能体长篇综述^[35]中智能体感知、认知及样本效率部分。本文针对以上问题,从智能体对轴孔状态的感知、任务理解与控制以及任务课程设计 3 个方面对相关方法及进展进行综述,并探讨其发展趋势。

2 轴孔状态感知

状态感知作为轴孔装配关键技术,是实现自动化和智能化的必要前提^[36-37]。在当今制造业智能化升级的浪潮下,轴孔装配日渐呈现出高度复杂、高度精密以及高度柔性的“三高”特性。装配环境中的光照变化、零件的微小形变、装配位置的细微偏差等多种不确定性因素交织,单一模态的感知手段^[38-39]已难以满足精准装配的需求。具体而言,动态光照造成的视觉特征漂移、工件微观形变引发的接触力突变、机器人末端执行器的累积定位误差、以及装配夹具的微振动干扰等多物理场耦合作用,使得视觉、力觉这类单一模态传感器易陷入“感知盲区”,无法全面、准确地获取装配环境的信息^[40]。故而必须开展多模态融合感知,通过整合视觉、力觉、本体(关节编码器)等多种感知信息,提升机器人对装配环境和任务状态的全面理解与判断能力。

在进行多模态特征感知装配环境的研究中,根据融合方式,可将多模态信息融合方法分为以下 3 类:特征级融合^[41]、决策级融合^[42]和模态级融合。特征级融合通过提取各模态的高维特征向量(如视

觉的CNN特征、触觉的力反馈时序特征、语言的词嵌入向量),在特征空间进行对齐与拼接。如图4所示,伯克利大学的Jones等^[43]提出了一种通过将自然语言作为跨模态的桥梁,将视觉、触觉和声音等多模态数据融合到策略模型中,从而允许策略在零样本情况下进行多模态推理,尤其该方法可融合包括听觉、视觉、本体感知和触觉等更多模态的信息,但是该方法依赖于对各模态信息进行语言描述的准确度,且难以处理高度复杂和动态环境变化,

在一定程度上限制了该方法的应用范围。以自然语言为桥梁进行多模态特征融合的过程中,无需大量特定模态数据,就能将不同模态与预训练通用策略连接,适合零样本和组合式推理,但自然语言表达的方式多样而且模糊,容易描述与理解不一致,从而导致模态转换后出现信息丢失或扭曲等现象。此外,还有以视觉特征^[44]、音频特征^[45]、知识图谱^[46]和空间向量^[47]等作为桥梁进行多模态信息特征级的交互和融合。

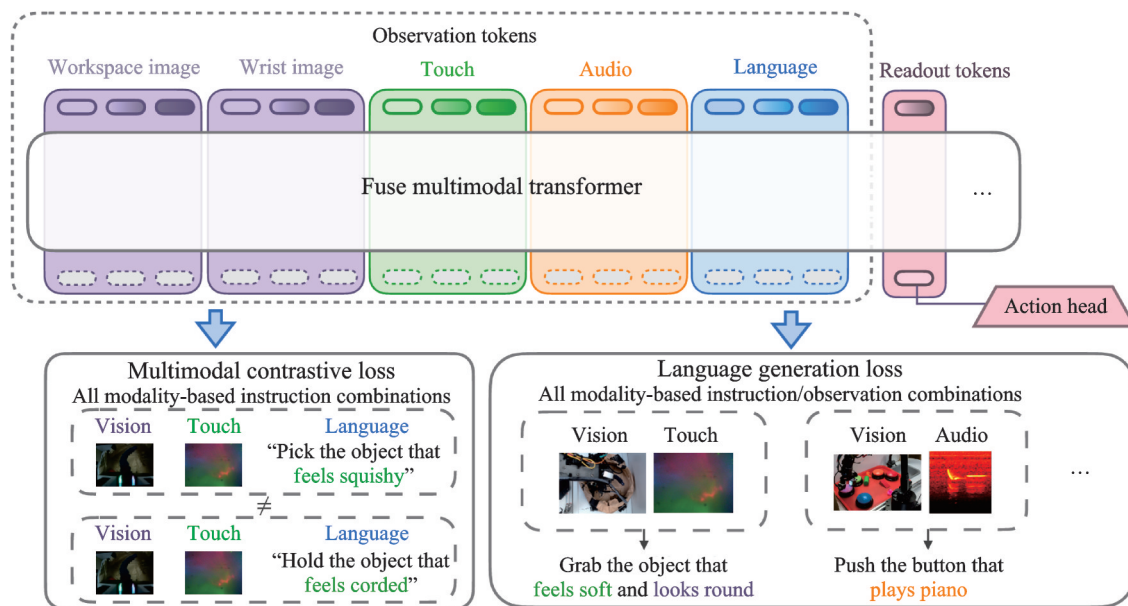


图4 多模态数据融合微调架构图^[43]

Fig.4 Multimodal data fusion fine-tuning architecture diagram^[43]

中山大学胡海峰团队Mai等^[48]为考虑特征向量不同部分的差异,提出分层融合策略,受卷积神经网络启发,采用滑动窗口便利对齐的多模态特征向量,对窗口内的特征进行局部融合,从特征维度对多模态信息进行针对性融合。然而该研究在融合过程中将各模态信息同等对待,而在轴孔装配各阶段中,不同模态信息对智能体的重要程度不同。华北电力大学Cao等^[49]由特征提取子网(Feature extraction subnetwork, FES)提取特征后,利用多级特征聚合模块跳跃连接深度融合特征,使各级特征图包含更多细节和语义信息。大连理工大学的Liu等^[50]针对多模态融合中目标模糊和纹理细节丢失等问题,构建了一种层次聚合融合架构,从特征级和决策级融合的角度提取和优化融合特征。然而该研究主要在红外、图像融合等场景进行试验,在更复杂、存在背景干扰的多模态融合场景的有效性有待验证。

特征级模态融合需解决特征尺度的差异(如视觉特征的高维和力觉的低维)和时序对齐问题(如力觉的高频采样和视觉的低维采样)等问题。

决策级融合中,先将多模态数据进行滤波、特征提取、对齐等一些必要的预处理后,直接输入给策略模型,让智能体自行感知轴孔状态,这种融合方式保留了各模态的独立性,在策略模型的输入端进行功能性互补。例如,博洛尼亚大学Gregorio等^[51]仅使用视、力觉传感器互相配合地在狭窄或遮挡的环境中进行装配作业:视觉系统通过CNN对环境进行三维重建,使用力觉传感器量化接触力进行轨迹调整。如图5所示,上海交通大学Zhang等^[52]分别训练了视觉策略和力控制策略,在不同的装配阶段执行不同的策略组合来提高样本效率,

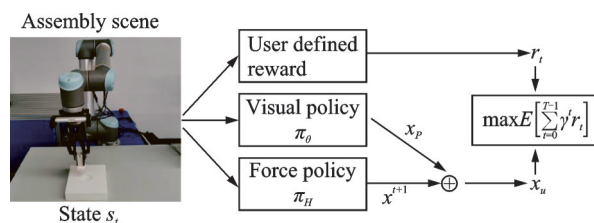


图5 基于视觉、力觉策略的装配^[52]

Fig.5 Assembly based on visual and force perception strategies^[52]

然而这种端到端的融合方式仅对多模态信息进行了功能性互补,难以获取视、力觉特征层面的融合信息,例如由力接触信息获得视觉被遮挡部分的位姿信息,且该方法需要策略模型对多模态信息进行融合利用,这将大幅增加策略模型的负担。决策级融合方式需设计多模态协调机制,如动态权重分配等,以避免模态冲突(视觉遮挡情况下过度依赖力觉导致振荡)。

因此面对需要多模态融合信息的任务中,将模态信息输入给策略模型前,先进行模态融合,获取智能体装配控制所需信息的模态级融合方法已经被广泛应用^[53]。这种融合方式介于特征级和决策级融合之间,通过表示学习实现模态交互。如图 6 所示,斯坦福大学李飞飞团队 Jin 等^[54]提出一种基于自监督学习的多模态表示学习模型,该模型可通过变分推理方法预测视觉信息、接触力信息以及机器人末端姿态等自监督目标,从而将状态估计与控制策略学习解耦,显著提高了策略模型训练的样本效率,但是该融合过程需要使用多个编码器和解码器末端姿态预测、时间对齐预测、接触力预测等,这可能导致其需要庞大的计算资源和数据进行训练。文献[55]提出一种多模态双重化技术,对视觉

和本体感知信息进行融合以感知轴孔位姿。具体而言,该方法针对两种模态特征设计了特定编码器:视觉编码器由 CNN 网络构成,通过 3 帧连续图像堆叠捕捉动态信息;本体感知数据则通过多层感知机(Multilayer perceptron, MLP)编码为低维向量。随后,通过双重化技术将视觉与本体感知的多模态表示分别输入 Actor 和 Critic 模型以优化性能。然而,该方法在稀疏奖励环境下的表现欠佳,尤其在复杂装配场景中,因缺乏高频真实接触力反馈,限制了策略模型的学习效率和探索能力。在使用模态级融合方式进行模态间交互时,需平衡模态间的信息冗余(如轴孔装配中视觉和本体感知均包含位姿信息)和计算效率。

目前轴孔状态感知研究已经取得部分进展,各融合方式横向对比如表 1 所示。其中,在输入前进行特征融合的方式被广泛证明可有效提高智能体感知能力,但在面对复杂装配场景中,智能体不仅需要对齐、互补融合后的多模态特征,还需要智能体在任务过程中关注的环境感知信息,例如,在进行曲面多装配的对齐阶段,智能体更加关注目标周围孔群的位姿,而非其他上百个相邻孔位姿或其他信息。

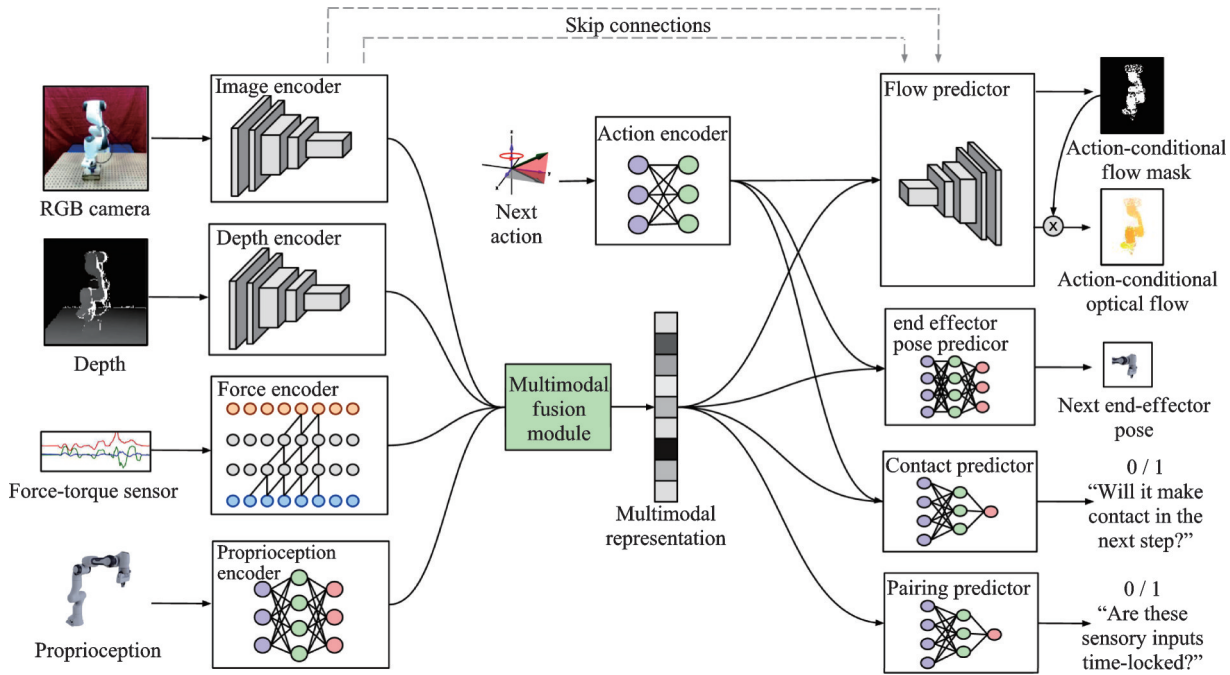


图 6 基于变分模型的多模态表示^[54]

Fig.6 Multimodal representation based on variational models^[54]

表 1 多模态融合方式横向对比

Table 1 Horizontal comparison of multimodal fusion methods

维度	特征级融合	决策级融合	模态级融合
融合阶段	特征提取后,策略输入前	策略推理过程中	表示学习过程中
信息整合深度	特征结合(高维空间)	高层决策协作(功能互补)	中层语义关联(表示学习)
适用场景	多模态强关联任务	动态环境下的模态互补	复杂耦合任务
计算复杂度	高(需跨模态对齐)	低(独立模态处理)	中(需联合优化编码器)

3 数据驱动的装配控制

机器人的装配控制可分为传统的控制算法和基于数据驱动的装配策略以及混合控制^[21]。其中,传统控制算法可分为无模型控制方法和基于模型的控制方法^[56],无模型的方法通过接触力反馈进行实时调整实现柔顺装配,控制时需要调试控制参数,耗时且繁琐。基于模型的控制方法在控制时,利用精确的接触动力学模型进行轨迹实时修正,面对复杂环境时难以建立精确的动力学模型。传统的装配控制算法难以适应机器人与工件之间的刚柔耦合效应,且在非结构化环境中,难以兼顾任务的高效性和安全性,任务泛化性较差。以数据

驱动的装配策略通过隐式建模、自主学习表现出良好的自主决策性能,系统性解决了传统方法中的模型依赖、自适应不足和泛化受限等问题。

现有以数据驱动的装配控制方法中,总体可分为3类:进行探索学习的探索式学习的装配策略^[57-59]、从先验知识(如专家演示路径、数据集等)中获得装配技能的策略^[43,60],以及将前两者结合的控制策略^[61]。在探索式学习策略中,华中科技大学丁汉团队Zhao等^[62]提出一种强化-阻抗策略,如图7所示,通过最大熵强化学习框架训练装配策略,策略接收轴孔状态信息,输出动作指令,由阻抗控制器进行执行,从而实现高精度装配任务。

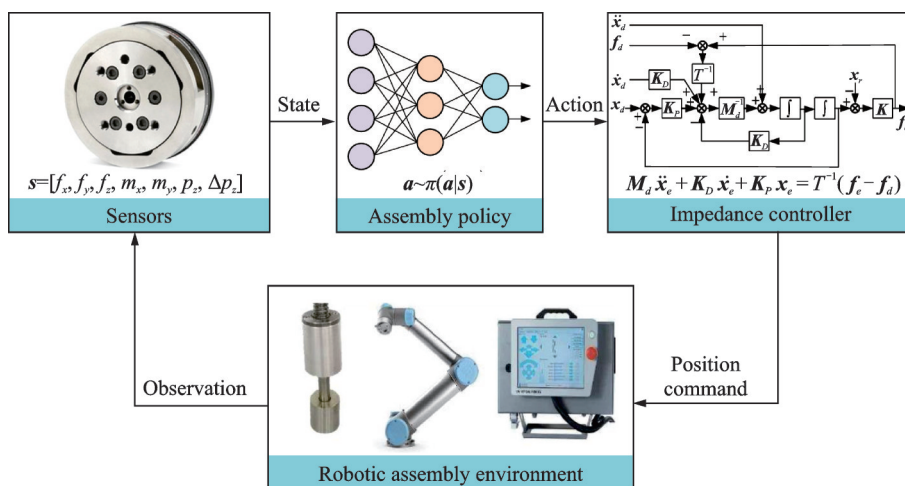


图7 基于阻抗-强化的轴孔装配策略^[62]

Fig.7 Shaft-hole assembly strategy based on impedance-reinforcement^[62]

清华大学吴丹团队Gai等^[63]通过结合基于模型参考的阻抗控制与迁移强化学习,提出了一种模型混合合规控制框架,使智能体更适合处理装配任务中复杂几何特征下的非线性力-位映射,利用迁移学习将已有的任务知识迁移到新任务中,提高了对不同几何特征的适应性,如图7所示。

意大利技术研究院的Zhang等^[64]通过将安全强化学习与变阻抗控制结合,提出一种安全强化学习框架,训练后的策略不仅负责生成动作,还可对阻抗参数进行调整,从而在任务效率和安全性之间实现平衡,解决了在轴孔装配任务中的安全探索问题,但该方法在复杂装配任务中的性能有待进一步验证。

如图8所示,弗吉尼亚大学的Zhu等^[65]提出一种结合任务导向安全场模型和DQN的残差强化学习方法,可利用传统控制器处理轴孔装配到达阶段的避障、导航等任务,由强化策略在装配过程中动态调整参数来减少动作空间与训练复杂度,但是该方法依赖于对环境可完全观测的假设,未考虑到传感器被遮挡或者噪声、预测误差对策略鲁棒性的

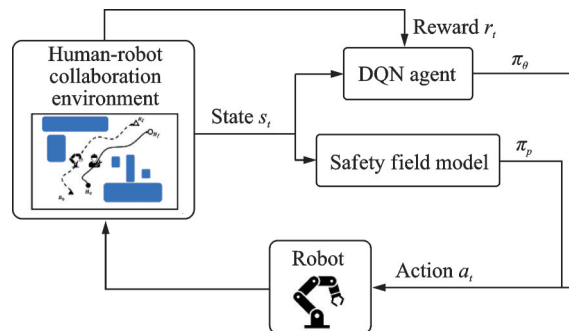
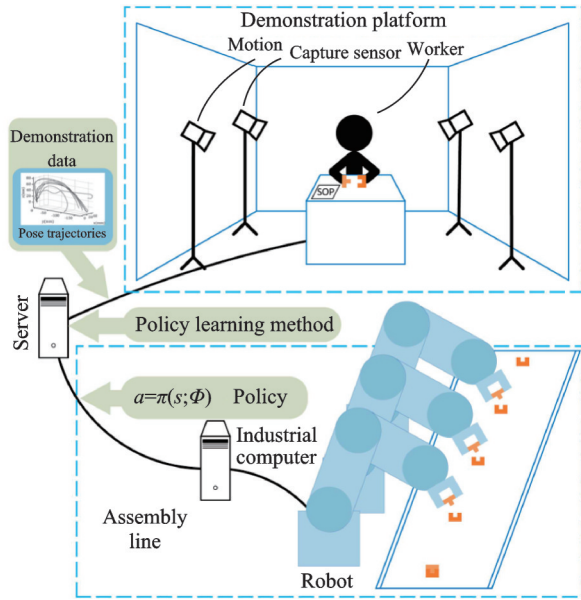


图8 到达阶段探索式控制策略^[65]

Fig.8 Exploratory control strategies in the arrival stage^[65]

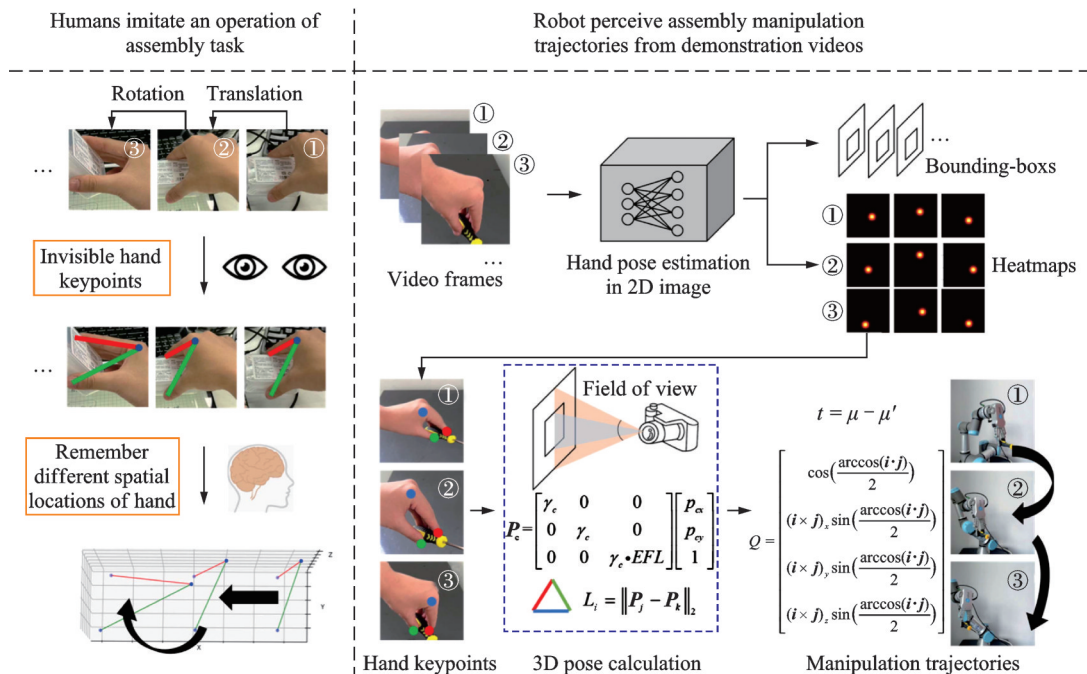
影响。因此,探索式的学习方法适用于轴孔装配的插孔阶段,可对环境变化进行动态响应,实现高精度轴孔装配任务。

利用先验知识获得装配策略的方法中,典型代表为利用模仿学习从专家演示路径中获取装配技能。如图9所示,哈尔滨工业大学人工智能研究院Hu等^[66]针对装配的3个阶段分别训练不同的装配控制策略,利用光学动作捕捉进行专家示教,摆脱传统示教方法对机器人本体的依赖,并支持利用多

图9 基于专家示教的机器人装配系统^[66]Fig.9 Robot assembly system based on expert teaching^[66]

组演示数据学习轴孔装配技能,但该模型训练时间会随运动基元的数量呈指数级增长,无法支持实时策略更新,难以适应复杂高精度装配任务。同时,哈工大机电学院 Zha 等^[26]提出了一种基于几何特征表示的任务空间建模方法,将复杂的装配任务简

化为几何特征空间中的表示来减少任务表示维度,并创新性地将概率运动基元(Probabilistic movement primitives, ProMPs)与行为克隆(Behavioral cloning, BC)结合提出新的模仿学习框架来训练全局装配策略,其中,ProMPs 用于从少量演示样本中提取全局任务策略,而 BC 则负责训练神经网络,学习全局任务空间中的技能。然而论文中设计了动态时间规整、ProMPs 模型提取计算复杂性较高的步骤,在进行高精度装配时,可能面临计算资源的挑战,并且论文进行了时间阶段划分,通过高斯模型估计当前状态所属的时间阶段,这种方法虽然简单,但在复杂任务中可能难以精确反映任务状态的时间变化,导致动态响应能力不足。如图 10 所示,北京理工大学 Deng 等^[67]仅通过单目相机,利用神经网络估计 2D 手部姿态,将其映射到真实 3D 装配空间,实现智能体从演示视频中就可以学习到轴孔装配技能。由此可见,通过模仿学习训练装配控制策略时,其优秀的任务理解能力可大幅缩短训练时间并提高样本效率。然而,该方法面临环境动态响应能力不足的挑战,尤其是在复杂装配任务中,智能体难以同时实现高效任务理解和实时高精度控制^[68-69]。

图10 机器人从原始视频中学习装配技能^[67]Fig.10 Robot learning assembly skills from original video^[67]

将探索策略与从先验知识中学习的策略相结合的混合控制方法中,中国科学院徐德团队 Ma 等^[56]提出了一种如图 11 所示的基于模仿学习和强化学习的精密装配插入控制策略,在训练过程中先利用高斯混合模型从多个演示动作中

学习潜在的共同特征,建立从状态到动作的映射模型,并用高斯混合回归将该模型泛化为初始策略,然后引入残差强化学习,在插入环境中进行探索来优化初始策略,从而提高智能体训练的样本效率。

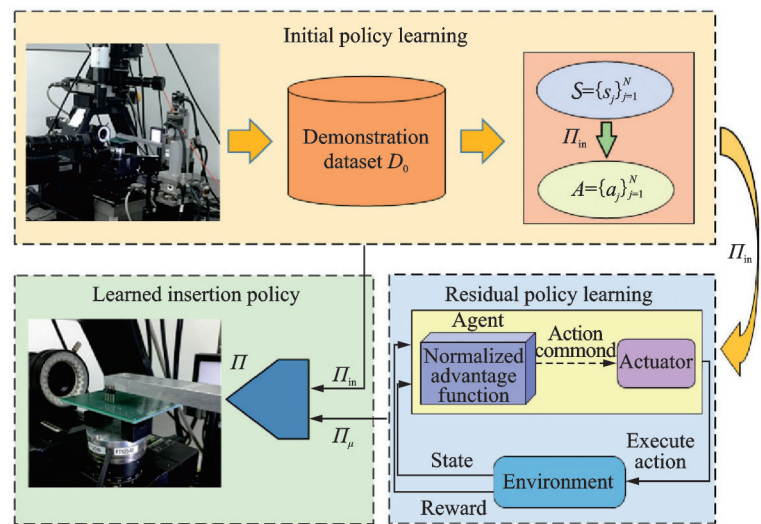


图 11 精确装配策略学习框架^[56]

Fig.11 Precise assembly strategy learning framework^[56]

然而该研究虽提及在插入环境变化时需重新学习插入策略,但缺乏对复杂环境变化的深入研究,例如在实际的高精度装配中,温度、湿度以及装配工件的微小振动等都可能会引起控制失稳,因此在实际应用中可能面临挑战。如图 12 所示,爱丁堡大学 Davchev 等^[70]联合 Facebook、谷歌提出了一种结合动态运动基元(Dynamic movement primi-

tives,DMP)和强化学习的残差学习框架,先通过 DMP 从单次演示中提取基础策略,再利用强化学习任务空间的非线性残差来校正策略,该过程虽然引入了方向残差校正,采用四元数空间的角度一轴表示法,避免了奇异性问题,但方向残差学习仍面临非欧几里得空间优化的挑战,研究表明智能体既可以快速理解任务,也可以进行精细化的操作,然而论文未深入分析方向校正的收敛性,也未对比其他方向表示(如旋转矩阵)的有效性。由此可见,在复杂轴孔装配任务中,利用先验知识进行快速的任務理解,并由探索策略进行优化相较于单独使用两种方法具有很大的优势。

因此国内外研究学者在数据驱动的装配控制方面已经取得了部分进展,通过模仿示教、强化探索以及两种方法结合尝试提高智能体的任务理解能力和高精度轴孔装配能力,利用 3 种方式进行装配技能学习的优缺点及应用场景如表 2 所示。尽管如此,现有策略对复杂装配任务的任务性能仍存在不足,尤其是在强非线性、高动态特性的装配场景中,不同工艺要求的重组将面临任务理解与装配控制之间的深层次耦合,这将加剧智能体在进行任务理解和控制时的难度。

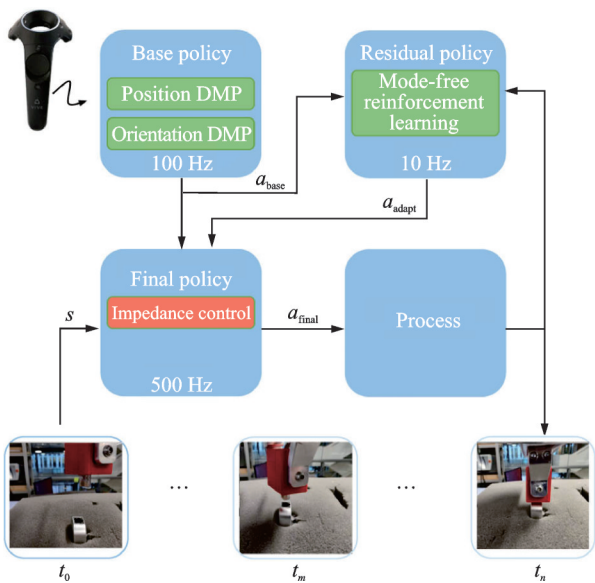


图 12 基于 DMP 和残差策略的装配框架^[70]

Fig.12 Assembly framework based on DMP and residual strategy^[70]

表 2 装配控制学习方式横向对比

Table 2 Horizontal comparison of learning methods for assembly control				
维度	核心特点	优点	缺点	应用场景
探索式学习	通过试错自主探索	自主性强、泛化能力好、具有良好的创新性	试错成本高、收敛速度慢、有较强的数据依赖性	容错率高、完全未知环境
从先验知识中学习	利用已有数据学习技能	学习效率、性能稳定、训练风险低	数据依赖性强、泛化性能有限、创新性不足	标准化装配场景
混合学习策略	已有数据上自主探索	数据效率高、鲁棒性强	融合机制较复杂、难以协调	可容错的半未知场景

4 任务课程设计

训练初期若智能体直接接触高难度装配任务,易导致收敛困难、样本效率低下等问题^[71]。因此受到人类学习过程遵循由简到难规律的启发,Bengio等^[72]在2009年提出课程学习的概念。最初,课程学习以先简单后难的顺序给智能体输入训练数据,作为一个灵活的即插即用子模块,而随着课程学习效果的显现,这种学习方法逐步被广泛应用到视觉识别、语言处理,以及强化学习的任务中。

清华大学朱文武团队 Wang等^[73]指出:课程学习使强化学习智能体能够解决在无课程学习情况下无法完成的复杂目标任务。在具有高维、时变和强约束特性的装配场景中,这将增加该装配任务的难度,如图13所示,课程学习机制被广泛引入智能体训练中,用以引导智能体进行由简到难的训练过程,保证智能体训练的收敛性和样本效率。

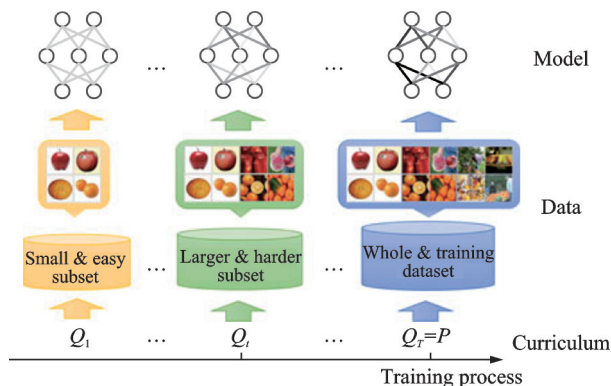


图13 课程学习过程^[73]

Fig.13 Curriculum learning process^[73]

课程学习过程的本质是适当调整智能体的训练难度,让难度适合于当前的智能体,其调整的作用对象有数据、模型、任务,课程学习中,以任务为作用对象在调整任务难度时有4类方法^[74]:平衡课程学习(Balanced curriculum learning, BCL)^[75]、渐进式课程学习(Progressive curriculum learning,

PCL)^[76]、师生课程学习^[77]、自步课程学习(Self-paced curriculum learning, SPCL)^[78-79]。其中:(1)平衡课程学习更多的关注状态和动作空间的多样性,在训练过程中,由开始的单独场景训练,逐步过渡到更加多样化的装配场景,可提高其泛化能力;(2)渐进式课程学习以关卡式的方式训练智能体,将任务按难度进行排序,智能体在完成简单任务后才能进入稍微困难的任务中;(3)师生课程学习则更像是在进行技能迁移,需要先预训练一个可以很好完成复杂装配任务的策略模型作为教师模型,在对教师模型进行动作预测时,从不同的角度实时给出建议。

自步课程学习主要侧重于预测出当前智能体在一个新任务分布上的性能表现^[80],最理想的情况是每次课程设计时都逼近智能体的性能边界,在保证智能体可以在新任务分布上完成任务的同时,以最大速度向目标复杂任务靠拢。为了尽可能逼近性能边界,课程设计时需要围绕一些特定的替代目标进行生成,以更好地引导智能体学习,例如在课程设计时最大化智能体学习进度^[81]、中间任务难度^[82]、遗憾值^[83]以及Q函数之间的差异^[84]等。如图14所示,伦敦大学 Jiang等^[85]使用固定大小的回放缓冲区实现高遗憾值任务的近似分布,同时鼓励频繁回放缓冲区内的任务,以更准确地估计遗憾值。法国国家信息与自动化研究所 Portelas等^[81]使用高斯混合模型来逼近可以实现高学习进度的任务分布。由于装配任务中智能体在初始阶段由于环境噪声等干扰因素不易精准控制,此时遗憾值和Q函数的计算对计算资源和数据量要求高,而中间任务难度相对更容易衡量和调整,并且可以更好地平衡智能体探索新策略和利用已有策略之间的关系,从而更稳定地提升性能^[86],因此部分学者采用在初始任务分布与目标任务分布之间进行插值的方式进行课程设计。

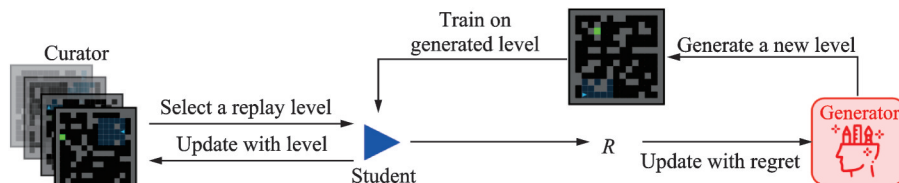


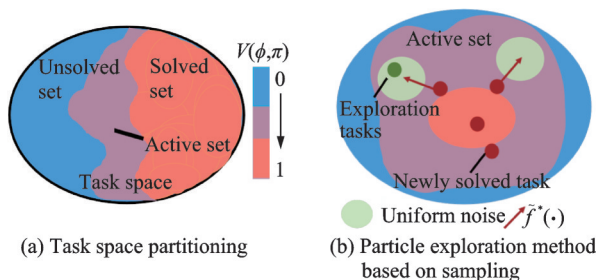
图14 师生式课程学习^[85]

Fig.14 Teacher-student curriculum learning^[85]

通过插值进行课程设计时,关键在于衡量插值的中间任务分布与目标任务分布之间的距离。德国达姆施塔特工业大学 Klink等^[78]将任务插值解释为一种概率分布生成过程,用KL(Kullback-Leibler)

散度衡量中间任务分布与目标任务分布的距离。

如图15所示,清华大学 Chen等^[87]提出一种变分自动课程学习,基于变分梯度下降生成新的任务分布,然而这些任务插值的过程中不可避免地需要

图15 基于变分自动课程的任务插值^[87]Fig.15 Task interpolation via variational automatic curriculum learning^[87]

依赖高斯分布^[77,88],虽然该方法可以通过近似来简化更新规则,简化计算过程,但其依赖于对任务分布的假设(例如均匀分布),而在如曲面多轴孔装配这类复杂装配任务中,任务难度的多因素耦合将导致任务并非呈规则的分布,也就无法充分考虑任务分布之间的相似性。对此,亚马逊(德国)的Klink等^[80]提出通过 Wasserstein 距离替代 KL 散度,确保任务分布插值时的连续性,避免参数化假设(如高斯分布)的限制,并通过显式任务相似度建模,严格性能约束,使其在稀疏奖励、高维任务空间中表现突出。

通过任务课程设计提高智能体在复杂任务中的性能表现,在一定程度上提高了训练的样本效率,提高了智能体在面对复杂轴孔装配时训练收敛的速度和可能性,但当各类任务参数对任务难度的影响规律不明时,课程设计机制仍难以准确预测智能体在中间任务分布中的性能表现,这将可能使插入的中间任务分布对智能体来说难度过大或过小,降低其训练效率。

5 结 论

轴孔装配作为制造业核心工艺,其智能化实现依赖于数据驱动技术对传统方法的突破。本文系统梳理了轴孔装配的阶段特性与关键挑战,从状态感知、装配控制及任务课程设计3方面总结了数据驱动方法的研究进展:多模态融合感知通过特征级、决策级和模态级融合提升复杂环境下的状态理解能力。装配控制中强化学习、模仿学习及混合策略展现出对非线性力-位映射的处理优势。任务课程设计通过渐进式、自步式等方法引导智能体由简入难学习,可有效提高样本效率与训练收敛性。现有研究虽在精度、效率上取得进步,但在强不确定性、高动态耦合的复杂装配场景中,智能体的实时动态响应、多物理场耦合下的鲁棒性以及任务理解-运动控制耦合机制仍是当前研究的瓶颈。

未来研究可从以下方向展开:

(1) 深化多模态融合的层次化建模,结合注意力机制动态聚焦关键感知信息,解决复杂光照、遮挡等干扰下的轴孔位姿精确估计问题,以低计算成本为智能体提供任务执行过程中智能体所关注的装配状态细节。

(2) 构建“先验知识引导+数据驱动优化”的混合控制框架,将接触动力学模型、几何约束等领域知识嵌入深度神经网络,提升智能体在过盈配合、非对称轴孔等严苛条件下的决策可靠性与动作精细化程度。

(3) 发展自适应课程学习理论,基于智能体实时评估动态生成任务序列,结合元学习实现不同装配工艺的快速迁移,突破任务参数耦合导致的难度预测瓶颈。

(4) 探索人机共融场景下的交互式学习,利用人类示教的语言指令、手势引导等自然交互数据,训练智能体理解隐含装配规则,实现从“完成任务”到“理解意图”的认知升级。此外,引入因果推理与可解释 AI 技术有望提升智能体决策透明度,推动工业场景下的可靠应用。随着具身智能与仿真技术的发展,数据驱动的轴孔装配将向全自主、高柔性的下一代智能制造系统演进。

此外,传统的装配控制算法在部分场景中仍然具有不可替代的作用,如在高度结构化且稳定的装配场景下,传统算法无须训练就能以低成本达到很高的装配精度;同时,在计算资源受限、对实时性要求高、装配工艺成熟且具有高可靠性要求等场景中,仍须依赖传统算法进行控制。

参考文献:

- [1] 赵富,张鹏,杨羲昊,等. 钢丝螺套自动安装系统设计及应用[J]. 机电产品开发与创新, 2023, 36(5): 73-75.
ZHAO Fu, ZHANG Peng, YANG Xihao, et al. Design and application of automatic installation system for steel wire bushings[J]. Development & Innovation of Machinery & Electrical Products, 2023, 36(5): 73-75.
- [2] LIU Z, CHEN J, CHEN C, et al. Target reconstruction and process parameter decision-making for bolt intelligent assembly based on robot and multi-camera[J]. Expert Systems with Applications, 2024, 256(1): 124871.
- [3] JIN J, HU J, LI C, et al. A digital twin system of reconfigurable tooling for monitoring and evaluating in aerospace assembly[J]. Journal of Manufacturing Systems, 2023, 68(1): 56-71.
- [4] HAMNER B, KOTERBA S, SHI J, et al. An auton-

- omous mobile manipulator for assembly tasks[J]. *Autonomous Robots*, 2010, 28(1): 131-149.
- [5] FERREIRA P, DOLTSINIS S, LOHSE N. Symbiotic assembly systems—A new paradigm[J]. *Procedia Cirp*, 2014, 17(1): 26-31.
- [6] 李明富, 王飞鸿, 朱凌枫, 等. 接触丰富的自动装配方法研究进展: 机器人装配接触力控制技能学习[J/OL]. *机械工程学报*, 1-19[2025-05-29]. <http://kns.cnki.net/kcms/detail/11.2187.th.20241212.1303.039.html>.
- LI Mingfu, WANG Feihong, ZHU Lingfeng, et al. Research progress on contact-rich automated assembly methods: Learning robotic assembly contact force control skills[J/OL]. *Journal of Mechanical Engineering*, 1-19[2025-05-29]. <http://kns.cnki.net/kcms/detail/11.2187.th.20241212.1303.039.html>.
- [7] GE M, JIN H, XU T, et al. Attitude alignment based on hole attitude estimation in robot assembly[J]. *IEEE Transactions on Instrumentation and Measurement*, 2025, 74(1): 7505111.
- [8] TUMULA S, RAMADEVI Y, PADMALATHA E, et al. An opportunistic energy-efficient dynamic self-configuration clustering algorithm in WSN-based IoT networks[J]. *International Journal of Communication Systems*, 2024, 37(1): e5633.
- [9] EKINCI S, IZCI D, EKER E, et al. Hunger games pattern search with elite opposite-based solution for solving complex engineering design problems[J]. *Evolving Systems*, 2024, 15(3): 939-964.
- [10] GHASEMI M, ZARE M, ZAHEDI A, et al. Geyser inspired algorithm: A new geological-inspired meta-heuristic for real-parameter and constrained engineering optimization[J]. *Journal of Bionic Engineering*, 2024, 21(1): 374-408.
- [11] YANG L, FU L, LI P, et al. An effective dynamic path planning approach for mobile robots based on ant colony fusion dynamic windows[J]. *Machines*, 2022, 10(1): 50.
- [12] ABDALLAH M A B, KHEMILI I, AIFAOU N, et al. Computer-aided design tool for typical flexible mechanisms synthesis by means of evolutionary algorithms[J]. *Robotica*, 2024, 42(4): 1172-1211.
- [13] 王战玺, 张益铭, 张邦海, 等. 双机器人协作轴孔装配力柔顺控制技术研究[J]. *机械工程学报*, 2024, 60(3): 28-33.
- WANG Zhanxi, ZHANG Yiming, ZHANG Banghai, et al. Influence of excitation amplitude and load on the characteristics of a quasi-zero stiffness isolator[J]. *Journal of Mechanical Engineering*, 2024, 60(3): 28-33.
- [14] GE D, ZHAO H, LI D, et al. A dual-arm robotic cooperative framework for multiple peg-in-hole assembly of large objects[J]. *Robotics and Computer-Integrated Manufacturing*, 2025, 95(1): 102991.
- [15] HERNÁNDEZ-ALVARADO R, GARCÍA-VALDOVINOS L G, SALGADO-JIMÉNEZ T, et al. Neural network-based self-tuning PID control for underwater vehicles[J]. *Sensors*, 2016, 16(9): 1429.
- [16] KIM Y L, AHN K H, SONG J B. Reinforcement learning based on movement primitives for contact tasks[J]. *Robotics and Computer-Integrated Manufacturing*, 2020, 62(1): 101863.
- [17] LI S, YE L, YU H, et al. JamTac: A tactile jamming gripper for searching and grasping in low-visibility environments[J]. *Soft Robotics*, 2023, 10(5): 988-1000.
- [18] 陈平, 李灿, 雷学军. 3D 视觉结合图像检测与导纳控制的圆轴孔零件机器人装配[J]. *控制与决策*, 2023, 38(4): 963-970.
- CHEN Ping, LI Can, LEI Xuejun. Robotic assembly of cylindrical shaft and hole parts based on 3D vision, image detection and admittance control[J]. *Control and Decision*, 2023, 38(4): 963-970.
- [19] CHOI J W, FANG T H, YOO W S, et al. Sensor data fusion using perception net for a precise assembly task[J]. *IEEE/ASME Transactions on Mechatronics*, 2003, 8(4): 513-516.
- [20] JIANG J, YAO L, HUANG Z, et al. The state of the art of search strategies in robotic assembly[J]. *Journal of Industrial Information Integration*, 2022, 26(1): 100259.
- [21] 宁睿, 刘志晨, 刘毅, 等. 强化学习引导变导纳控制的机械臂精密轴孔装配模糊位姿估计与精确调整[J]. *仪器仪表学报*, 2025, 45(11): 170-177.
- NING Rui, LIU Zhichen, LIU Yi, et al. Reinforcement learning guided variable admittance control for fuzzy pose estimation and precise adjustment of robotic precise peg-in-hole assembly[J]. *Chinese Journal of Scientific Instrument*, 2025, 45(11): 170-177.
- [22] YOON I, NA M, SONG J B. Assembly of low-stiffness parts through admittance control with adaptive stiffness[J]. *Robotics and Computer-Integrated Manufacturing*, 2024, 86(1): 102678.
- [23] SCHIAPPA M C, RAWAT Y S, SHAH M. Self-supervised learning for videos: A survey[J]. *ACM Computing Surveys*, 2023, 55(13): 1-37.
- [24] BRÖKER F, HOLT L L, ROADS B D, et al. Demystifying unsupervised learning: How it helps and hurts[J]. *Trends in Cognitive Sciences*, 2024, 28(11): 974-986.
- [25] WANG H, ZHOU X, KANG Y, et al. Multi-agent reinforcement learning with weak ties[J]. *Information*

- Fusion, 2025, 118(1): 102942.
- [26] ZANG Y, WANG P, ZHA F, et al. Peg-in-hole assembly skill imitation learning method based on ProMPs under task geometric representation[J]. Frontiers in Neurorobotics, 2023, 17(1): 1320251.
- [27] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Playing atari with deep reinforcement learning[J]. arXiv preprint arXiv:1312.5602, 2013.
- [28] SILVER D, LEVER G, HEESS N, et al. Deterministic policy gradient algorithms[C]//Proceedings of International Conference on Machine Learning. [S.l.]: PMLR, 2014: 387-395.
- [29] WANG Z, SCHAUL T, HESSEL M, et al. Dueling network architectures for deep reinforcement learning [C]//Proceedings of International Conference on Machine Learning. [S.l.]: PMLR, 2016: 1995-2003.
- [30] LILLICRAP T P, HUNT J J, PRITZEL A, et al. Continuous control with deep reinforcement learning [J]. arXiv preprint arXiv:1509.02971, 2015.
- [31] SCHULMAN J, WOLSKI F, DHARIWAL P, et al. Proximal policy optimization algorithms[J]. arXiv preprint arXiv:1707.06347, 2017.
- [32] SCHULMAN J, LEVINE S, ABBEEL P, et al. Trust region policy optimization[C]//Proceedings of International Conference on Machine Learning. [S.l.]: PMLR, 2015: 1889-1897.
- [33] FUJIMOTO S, HOOF H, MEGER D. Addressing function approximation error in actor-critic methods [C]//Proceedings of International Conference on Machine Learning. [S.l.]: PMLR, 2018: 1587-1596.
- [34] HAARNOJA T, ZHOU A, ABBEEL P, et al. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor[C]//Proceedings of International Conference on Machine Learning. [S.l.]: PMLR, 2018: 1861-1870.
- [35] LIU B, LI X, ZHANG J, et al. Advances and challenges in foundation agents: From brain-inspired intelligence to evolutionary, collaborative, and safe systems[J]. arXiv preprint arXiv:2504.01990, 2025.
- [36] CUI J, ZHANG Y, LU Y, et al. Towards perceptive assembly: Real-time full-field deformation twin perception method for large components under parametric load[J]. Journal of Manufacturing Systems, 2025, 79(1): 101-116.
- [37] JIANG M, GUO Y, HUANG S, et al. Generating the assembly instructions of helicopter subassemblies using the hierarchical pruning strategy and large language model[J]. Journal of Industrial Information Integration, 2024, 42(1): 100723.
- [38] LV J, ZHANG R, LI X, et al. A multimodality scene graph generation approach for robust human-robot collaborative assembly visual relationship representation [J]. IEEE Transactions on Industrial Informatics, 2023, 20(3): 3242-3251.
- [39] ZHAO F, ZHANG C, GENG B. Deep multimodal data fusion[J]. ACM Computing Surveys, 2024, 56(9): 1-36.
- [40] WANG C, LIN Z, LIU B, et al. Task attention-based multimodal fusion and curriculum residual learning for context generalization in robotic assembly[J]. Applied Intelligence, 2024, 54(6): 4713-4735.
- [41] ZHANG R, NIE F, LI X, et al. Feature selection with multi-view data: A survey[J]. Information Fusion, 2019, 50(1): 158-167.
- [42] WEI Y, WU D, TERPENNY J. Decision-level data fusion in quality control and predictive maintenance [J]. IEEE Transactions on Automation Science and Engineering, 2020, 18(1): 184-194.
- [43] JONES J, MEES O, SFERRAZZA C, et al. Beyond sight: Finetuning generalist robot policies with heterogeneous sensors via language grounding[J]. arXiv preprint arXiv:2501.04693, 2025.
- [44] XU X, CHEN Z, YIN F. Deep mutual information multi-view representation for visual recognition[J]. Applied Intelligence, 2022, 52(13): 14888-14904.
- [45] SANKALA S. Multi-feature integration for speaker embedding extraction[C]//Proceedings of ICASSP 2022—2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). [S.l.]: IEEE, 2022: 7957-7961.
- [46] CHEN Z, WAN Y, LIU Y, et al. A knowledge graph-supported information fusion approach for multifaceted conceptual modelling[J]. Information Fusion, 2024, 101(1): 101985.
- [47] SENEVIRATNE G, WEERAKOON K, ELNOOR M, et al. CROSS-GAIT: Cross-attention-based multimodal representation fusion for parametric gait adaptation in complex terrains[J]. arXiv preprint arXiv:2409.17262, 2024.
- [48] MAI S, HU H, XING S. Divide, conquer and combine: Hierarchical feature fusion network with local and global perspectives for multimodal affective computing[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, Italy: Association for Computational Linguistics, 2019: 481-492.
- [49] CAO J, YANG G, YANG X. A pixel-level segmentation convolutional neural network based on deep feature fusion for surface defect detection[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 70(1): 1-12.
- [50] LIU R, LIU Z, LIU J, et al. Searching a hierarchical-

- ly aggregated fusion architecture for fast multi-modality image fusion[C]//Proceedings of the 29th ACM International Conference on Multimedia. [S.l.]: ACM, 2021: 1600-1608.
- [51] GREGORIO D, ZANELLA R, PALLI G, et al. Integration of robotic vision and tactile sensing for wire-terminal insertion tasks[J]. IEEE Transactions on Automation Science and Engineering, 2018, 16(2): 585-598.
- [52] ZHANG Z, WANG Y, ZHANG Z, et al. A residual reinforcement learning method for robotic assembly using visual and force information[J]. Journal of Manufacturing Systems, 2024, 72(1): 245-262.
- [53] LEE M A, ZHU Y, ZACHARES P, et al. Making sense of vision and touch: Learning multimodal representations for contact-rich tasks[J]. IEEE Transactions on Robotics, 2020, 36(3): 582-596.
- [54] JIN P, LIN Y, SONG Y, et al. Vision-force-fused curriculum learning for robotic contact-rich assembly tasks[J]. Frontiers in Neurorobotics, 2023, 17(1): 1280773.
- [55] NOH S, LEE W, MYUNG H. Sample-efficient and occlusion-robust reinforcement learning for robotic manipulation via multimodal fusion dualization and representation normalization[J]. Neural Networks, 2025(1): 107202.
- [56] MA Y, XU D, QIN F. Efficient insertion control for precision assembly based on demonstration learning and reinforcement learning[J]. IEEE Transactions on Industrial Informatics, 2020, 17(7): 4492-4502.
- [57] 双丰, 卢王玉, 李少东, 等. 基于强化学习的机器人轴孔装配算法[J]. 机器人, 2023, 45(3): 321-332.
- SHUANG Feng, LU Wanyu, LI Shaodong, et al. Robotic peg-in-hole assembly algorithm based on reinforcement learning[J]. Robot, 2023, 45(3): 321-332.
- [58] 武黎明, 王明明, 罗建军, 等. 基于强化学习的空间机器人在轨装配柔顺控制方法[J]. 机器人, 2025, 47(2): 239-248, 258.
- WU Liming, WANG Mingming, LUO Jianjun, et al. Compliant control method for on-orbit assembly by space robots based on reinforcement learning[J]. Robot, 2025, 47(2): 239-248, 258.
- [59] CAO G, BAI J. Multi-agent deep reinforcement learning-based robotic arm assembly research[J]. PloS One, 2025, 20(2): 0311550.
- [60] JUNG D, LEE H, YOON S. Sample-efficient adversarial imitation learning[J]. Journal of Machine Learning Research, 2024, 25(31): 1-32.
- [61] WANG C, SU C, SUN B, et al. Extended residual learning with one-shot imitation learning for robotic assembly in semi-structured environment[J]. Frontiers in Neurorobotics, 2024, 18(1): 1355170.
- [62] ZHAO X, ZHAO H, CHEN P, et al. Model accelerated reinforcement learning for high precision robotic assembly[J]. International Journal of Intelligent Robotics and Applications, 2020, 4(2): 202-216.
- [63] GAI Y, WANG B, ZHANG J, et al. Robotic assembly control reconfiguration based on transfer reinforcement learning for objects with different geometric features[J]. Engineering Applications of Artificial Intelligence, 2024, 129(1): 107576.
- [64] ZHANG H, SOLAK G, LAHR G J G, et al. Srlvic: A variable stiffness-based safe reinforcement learning for contact-rich robotic tasks[J]. IEEE Robotics and Automation Letters, 2024, 9(6): 5631-5638.
- [65] ZHU C, YU T, CHANG Q. Task-oriented safety field for robot control in human-robot collaborative assembly based on residual learning[J]. Expert Systems with Applications, 2024, 238(1): 121946.
- [66] HU H, YAN H, YANG X, et al. Robot learning from demonstration for assembly with sequential assembly movement primitives[J]. IEEE/ASME Transactions on Mechatronics, 2023, 29(4): 2685-2696.
- [67] DENG X, LIU J, GONG H, et al. A human-robot collaboration method using a pose estimation network for robot learning of assembly manipulation trajectories from demonstration videos[J]. IEEE Transactions on Industrial Informatics, 2022, 19(5): 7160-7168.
- [68] HUA J, ZENG L, LI G, et al. Learning for a robot: Deep reinforcement learning, imitation learning, transfer learning[J]. Sensors, 2021, 21(4): 1278.
- [69] LI Y, WANG Y H, TAN X Y. Self-imitation guided goal-conditioned reinforcement learning[J]. Pattern Recognition, 2023, 144(1): 109845.
- [70] DAVCHEV T, LUCK K S, BURKE M, et al. Residual learning from demonstration: Adapting DMPS for contact-rich manipulation[J]. IEEE Robotics and Automation Letters, 2022, 7(2): 4488-4495.
- [71] LIU S, QING Y, XU S, et al. Curricular subgoals for inverse reinforcement learning[J]. IEEE Transactions on Intelligent Transportation Systems, 2025, 26(3): 3016-3027.
- [72] BENGIO Y, LOURADOUR J, COLLOBERT R, et al. Curriculum learning[C]//Proceedings of the 26th Annual International Conference on Machine Learning. [S.l.]: Association for Computing Machinery, 2009: 41-48.
- [73] WANG X, CHEN Y, ZHU W. A survey on curriculum learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(9): 4555-4576.
- [74] SOVIANY P, IONESCU R T, ROTA P, et al. Curriculum learning: A survey[J]. International Journal

- of Computer Vision, 2022, 130(6): 1526-1565.
- [75] HE C, SHAH A D, TANG Z, et al. Fedcv: A federated learning framework for diverse computer vision tasks[J]. arXiv preprint arXiv:2111.11066, 2021.
- [76] KARRAS T, AILA T, LAINE S, et al. Progressive growing of gans for improved quality, stability, and variation[J]. arXiv preprint arXiv:1710.10196, 2017.
- [77] HACHEN G, WEINSHALL D. On the power of Curriculum learning in training deep networks[C]//Proceedings of International Conference on Machine Learning. [S.l.]: PMLR, 2019: 2535-2544.
- [78] KLINK P, ABDULSAMAD H, BELOUSOV B, et al. A probabilistic interpretation of self-paced learning with applications to reinforcement learning[J]. Journal of Machine Learning Research, 2021, 22(182): 1-52.
- [79] EIMER T, BIEDENKAPP A, HUTTER F, et al. Self-paced context evaluation for contextual reinforcement learning[C]//Proceedings of International Conference on Machine Learning. [S.l.]: PMLR, 2021: 2948-2958.
- [80] KLINK P, D'ERAMO C, PETERS J, et al. On the benefit of optimal transport for curriculum reinforcement learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(11): 7191-7204.
- [81] PORTELAS R, COLAS C, HOFMANN K, et al. Teacher algorithms for curriculum learning of deep RL in continuously parameterized environments[C]//Proceedings of Conference on Robot Learning. [S.l.]: PMLR, 2020: 835-853.
- [82] FLORENSA C, HELD D, GENG X, et al. Automatic goal generation for reinforcement learning agents [C]//Proceedings of International Conference on Machine Learning. [S.l.]: PMLR, 2018: 1515-1528.
- [83] JIANG M, GREFFENSTETTE E, ROCKTÄSCHEL T. Prioritized level replay[C]//Proceedings of International Conference on Machine Learning. [S.l.]: PMLR, 2021: 4940-4950.
- [84] ZHANG Y, ABBEEL P, PINTO L. Automatic curriculum learning through value disagreement[J]. Advances in Neural Information Processing Systems, 2020, 33(1): 7648-7659.
- [85] JIANG M, DENNIS M, PARKER-HOLDER J, et al. Replay-guided adversarial environment design[J]. Advances in Neural Information Processing Systems, 2021, 34(1): 1884-1897.
- [86] FICKINGER A, COHEN S, RUSSELL S, et al. Cross-domain imitation learning via optimal transport [J]. arXiv preprint arXiv:2110.03684, 2021.
- [87] CHEN J, ZHANG Y, XU Y, et al. Variational automatic curriculum learning for sparse-reward cooperative multi-agent problems[J]. Advances in Neural Information Processing Systems, 2021, 34(1): 9681-9693.
- [88] KLINK P, D'ERAMO C, PETERS J R, et al. Self-paced deep reinforcement learning[J]. Advances in Neural Information Processing Systems, 2020, 33(1): 9216-9227.

(编辑:陈珺)