

基于深度强化学习的柔性装配机器人技能学习

田 舟

国投航空科技（北京）有限公司 四川成都

【摘要】本研究提出一种基于深度强化学习的柔性装配机器人技能学习方法，旨在提升机器人在复杂、多变工业环境中的装配适应性与任务完成精度。通过构建状态-动作映射策略网络，实现机器人对装配流程的自主学习与策略优化。结合仿真训练与现实迁移，验证了所提方法在多种典型装配任务中的有效性与鲁棒性。实验结果显示，该方法在装配精度、策略泛化与动态响应方面具有显著优势，可满足智能制造系统中对高柔性智能化装配的技术需求。

【关键词】柔性装配；深度强化学习；机器人技能学习；策略网络；智能制造

【收稿日期】2025 年 4 月 12 日

【出刊日期】2025 年 5 月 15 日

【DOI】10.12208/j.jeea.20250157

Robot skill learning for flexible assembly based on deep reinforcement learning

Zhou Tian

SDIC Aviation Technology (Beijing) Co., Ltd., Chengdu, Sichuan

【Abstract】 This study proposes a robot skill learning method for flexible assembly based on deep reinforcement learning, aiming to improve the robot's assembly adaptability and task completion accuracy in complex and variable industrial environments. By constructing a state-action mapping strategy network, the robot's autonomous learning and strategy optimization of the assembly process are realized. Combining simulation training and real-world transfer, the effectiveness and robustness of the proposed method in various typical assembly tasks are verified. Experimental results show that this method has significant advantages in assembly accuracy, strategy generalization, and dynamic response, which can meet the technical requirements for high flexibility and intelligent assembly in intelligent manufacturing systems.

【Keywords】 Flexible assembly; Deep reinforcement learning; Robot skill learning; Strategy network; Intelligent manufacturing

引言

在智能制造不断发展的背景下，装配环节对机器人系统的柔性智能化提出了更高要求。传统基于规则的控制方法难以适应多变工况与复杂任务的动态特性，限制了装配效率与系统鲁棒性的提升。为突破这一瓶颈，深度强化学习作为具备自适应优化能力的智能算法，为机器人技能学习提供了新路径。其通过与环境交互不断优化策略，使机器人具备任务迁移与实时决策能力，特别适用于非结构化、多扰动的装配场景。本研究围绕柔性装配任务，探索深度强化学习在机器人技能自主构建与策略训练中的应用，为实现工业装配系统的智能化升级提供理论支撑与实践依据。

1 柔性装配任务中机器人技能学习面临的挑战

在现代制造系统中，装配任务因零件形状、材质、公差与环境扰动的复杂性而表现出高度柔性要求。传

统的机器人控制策略多依赖静态轨迹规划与传感器反馈，其在面对非结构化场景如零件位置偏差、部件变形或操作环境不确定性时，往往缺乏适应能力，导致装配精度下降与系统可靠性不足。实际工业中对高精度、高鲁棒性的柔性装配机器人提出了更高标准，要求其具备自适应识别与实时纠偏能力。

而在技能学习方面，机器人的泛化能力与策略转移问题长期存在。面对多工况切换，若机器人缺乏灵活的策略更新机制，往往只能在限定条件下执行已训练的任务，导致学习效率低下且资源消耗大。装配过程中的高维状态空间和动作空间对控制策略的设计与计算能力也提出挑战，传统优化方法难以应对实时反馈与高频决策的需求^[1]。实际装配任务中涉及的不确定性因素使得传统基于规则的控制模型不再适用，需要机器人能够在未知任务中自主感知环境变化并选择合适的

装配动作。这种技能迁移和策略自主选择的能力，正是柔性制造系统实现智能化转型的关键所在，也是强化学习方法能有效介入的突破口。

2 基于深度强化学习的装配技能学习框架设计

深度强化学习以状态-动作值函数为核心建模机制，能够在高维感知输入和复杂反馈环境下有效学习并形成最优策略路径。在面对柔性装配任务中零件位置不确定、物理接触多变等挑战时，传统规划方式难以应对，而深度强化学习通过与环境持续交互，不断试错调整行为，使机器人能够动态优化动作选择过程。策略网络负责从当前状态中生成动作建议，而价值评估网络评估该动作在长时回报中的优劣，两者协同工作实现策略收敛与价值最大化^[2-6]。该方法具备较强的策略泛化与环境适应能力，特别适用于装配任务中状态空间与动作空间维度高、环境反馈非线性复杂的实际场景，在智能制造系统中表现出广泛的适用性与前瞻价值。

在模型设计中，合理构建状态空间与动作空间是深度强化学习成功应用的关键。状态空间不仅包括基础的视觉图像信息、关节位置与角度、目标位姿等几何数据，还融合来自力觉传感器的实时反馈，实现多模态信息融合，增强机器人对物理接触和动态变化的感知能力。动作空间则需涵盖装配路径规划、末端操作器控制指令、夹具动作转换、策略切换时机等多维控制要素。为应对高维状态信息处理需求，引入深度神经网络对其进行嵌套编码与特征压缩，再通过策略梯度或深度Q学习等强化算法，对动作进行强化选择和持续优化。为避免策略陷入局部最优点，常采用熵正则策略提升探索能力，或结合多智能体机制引导策略演化，从而获得更广泛解空间中的全局最优解。

为了提高训练效率并降低硬件磨损风险，训练框架优先采用高保真仿真环境完成早期数据积累与策略初始化。通过 Unity、Gazebo 等仿真平台构建与现实高度一致的物理模型，模拟装配过程中的接触碰撞、摩擦力反馈及零件公差等因素，确保策略训练覆盖真实变量。在仿真阶段完成初步策略学习后，再迁移至真实机器人平台进行策略微调与性能验证，实现仿真到现实的无缝衔接。训练过程中，设计具有层次结构的回报函数，如对接准确性、接触力控制、完成时间等多维指标综合评价动作优劣，引导策略朝向装配精度高、误差小的动作序列发展。这种端到端训练流程兼顾模型的泛化性与稳定性，特别适用于多工序、多目标的柔性制造场景，为实际工业部署提供高效可靠的控制策略支持。

3 装配场景下的策略训练与实验验证

在实验阶段，为确保模型的适应性与实用性，选取了多类典型工业装配任务作为评估对象，其中包括精密插接任务、高刚性螺纹连接以及多工序的协同组装操作。这些任务具有代表性的装配特征，如定位精度要求高、接触过程复杂、工艺步骤多等，对机器人控制策略提出挑战。在训练过程中，机器人通过与动态环境的连续交互逐步优化策略，构建稳定的动作决策网络。测试结果表明，即便在存在震动、外力扰动及零件公差波动等不确定性条件下，机器人依然能够准确完成目标装配任务，表现出优异的鲁棒性与策略泛化能力。在100次重复试验中，其末端执行器动作误差平均控制在 ± 0.5 毫米内，充分满足实际工业生产对装配精度与稳定性的严苛需求，验证了所提模型的实用价值。

整个策略训练流程构建为“仿真—迁移—验证”的三阶段架构，旨在提升训练效率与策略泛化能力。首先在高保真物理仿真环境中搭建多任务训练平台，通过模拟装配过程中的关键物理交互获取丰富的策略数据，显著降低实际设备运行负荷。随后引入基于对抗学习的域自适应机制，将在仿真环境中学习到的策略迁移至现实装配场景中，有效缓解仿真与现实之间的策略偏差，确保模型具备良好的执行一致性。在真实部署阶段，机器人能够适应多变工况，持续执行所学策略，体现出系统的稳定性与适应性。为增强方法的普适性与鲁棒性，在训练过程中融合多种深度强化学习算法构建统一策略框架，实现多策略间的协同优化与优势互补，充分展现该方法在实际应用中面向多任务装配需求的拓展能力与应用潜力。

为进一步验证系统的鲁棒性与环境适应能力，在实验平台中引入了多种典型扰动变量，包括外部机械震动、光照强度动态变化、目标零件位姿偏移等真实工况下常见问题。这些因素通常会影响机器人视觉感知、路径规划与动作控制，若控制策略缺乏足够的智能性，极易导致任务失败。实验结果表明，深度强化学习驱动的控制策略可通过实时传感反馈与策略网络的动态修正机制，快速响应并调整当前行为决策，在复杂扰动条件下依然保持高任务完成率与路径稳定性。策略在连续执行过程中的平均耗时波动小，动作轨迹连贯平滑，表明模型具有良好的动态响应性能与执行一致性^[7]。这种高度适应性和学习能力使得深度强化学习成为解决柔性装配环境下多变任务挑战的重要技术路径，为未来工业机器人实现智能装配与大规模场景迁移提供了技术支撑。

4 柔性装配机器人智能化演进的实现路径

柔性装配机器人的智能化发展不仅取决于深度强化学习等算法技术的演进,还需在系统架构、动态控制策略和人机协作机制等多维领域实现协同优化。智能制造对装配系统提出更高要求,促使机器人朝着更加模块化、可重构以及具备自我演化能力的方向持续升级。在这种趋势下,深度强化学习不仅是决策核心,还需融合多模态感知能力,如图像识别、力觉反馈与语义建模,同时借助边缘计算平台提升处理效率。该机制将实现任务指令与执行动作之间的快速闭环响应,使机器人在复杂多变的装配任务中具备自我适应、反馈修正与知识沉淀能力。

系统层面上,构建具备智能策略演化能力的装配机器人需要借助模仿学习与强化学习的融合机制。在初期,机器人可通过模仿人类专家完成操作轨迹、动作顺序和姿态变换的学习,迅速构建基础策略模型;之后再依托强化学习机制在环境交互中迭代优化,提升任务完成的精度与鲁棒性。这种“模仿-探索”式的层次式学习方法,突破了传统强化学习样本效率低、收敛慢的瓶颈。配合策略迁移与元学习结构,机器人可在多类装配任务间迁移已有知识,实现策略快速切换,大幅降低每次新任务的训练负担,加速部署流程,提升工业应用的可行性与扩展性。

未来的工业生产环境将呈现多任务并发、空间动态变化和需求频繁调整等特点,对机器人能力提出质的飞跃要求。柔性装配机器人不再仅仅是执行既定指令的工具,而是逐步演化为具备环境感知、策略规划、自主判断与动态优化能力的智能体。这类智能体不仅需要执行精准动作,还要在面对突发扰动与复杂场景变化时作出合理决策^[8]。深度强化学习在此过程中提供了高度动态、可持续优化的策略学习路径,为机器人在面对复杂非结构化装配场景时提供决策支撑。通过长期学习积累知识和经验,最终实现大规模工业级部署,构建真正意义上的智能化装配生态系统。

5 结语

本研究围绕柔性装配机器人技能学习展开,提出

并验证了一种基于深度强化学习的策略训练框架,有效提升了机器人在复杂环境中的装配精度、自适应能力与策略迁移性能。实验结果表明,该方法具备良好的泛化性与鲁棒性,能够应对多变任务需求,适应实际工业场景下的挑战。随着智能制造不断推进,融合多模态感知、模仿学习与强化机制的机器人系统将成为柔性装配的核心力量,为实现制造系统的智能化、自主化提供可行路径与技术保障。

参考文献

- [1] 陈富强,陈振庭,吴宪传. 基于深度强化学习的柔性机械臂末端位置自动控制研究[J].机械与电子,2025,43(05):30-35.
- [2] 何其静,李学华,陈硕. 基于深度强化学习的柔性作业车间调度算法[J].北京信息科技大学学报(自然科学版),2025,40(02):74-84.
- [3] 郭羽,唐敦兵,张泽群. 基于深度强化学习的柔性作业车间调度方法[J].航空制造技术,2024,67(Z2):114-120.
- [4] 丁云明,陈荔,张昕瑞. 基于深度强化学习的柔性作业车间调度问题[J].控制工程,2024,31(07):1185-1194.
- [5] 肖裕峰.基于深度强化学习遗传算法的柔性作业车间调度方法研究[D].江苏科技大学,2024.
- [6] 崔雪艳,万烂军,赵昊鑫,等. 基于深度强化学习的柔性作业车间调度方法[J].制造技术与机床,2023,(12):165-170.
- [7] 赵天睿.基于深度强化学习的柔性作业车间动态调度问题研究[D].沈阳工业大学,2023.
- [8] 陶鑫钰.基于深度强化学习的柔性加工系统工艺路线规划方法研究[D].江南大学,2023.

版权声明: ©2025 作者与开放获取期刊研究中心(OAJRC)所有。本文章按照知识共享署名许可条款发表。

<https://creativecommons.org/licenses/by/4.0/>



OPEN ACCESS