

## 基于强化学习的机械臂示教编程自适应优化

李家忠

辽宁省营口市营口中裕耐火材料有限公司 辽宁营口

**【摘要】**机械臂示教编程在工业自动化领域广泛应用，但传统方式存在灵活性差、效率低等问题。为实现更高效智能的示教编程，强化学习技术被引入。通过构建包含状态空间、动作空间和奖励函数的强化学习模型，对机械臂示教编程过程进行自适应优化。模型以机械臂关节角度、末端执行器位姿等作为状态输入，以关节运动指令为动作输出，基于任务完成情况和能耗等设定奖励机制。经仿真与实验验证，该优化方法显著提升机械臂示教编程效率，降低操作复杂性，为机械臂在复杂任务场景下的应用提供有效技术支持。

**【关键词】**强化学习；机械臂；示教编程；自适应优化；工业自动化

**【收稿日期】**2025 年 3 月 10 日

**【出刊日期】**2025 年 4 月 18 日

**【DOI】**10.12208/j.ijme.20250021

### Adaptive optimization of teaching programming for robotic arms based on reinforcement learning

Jiazhong Li

Liaoning Yingkou Zhongyu Refractory Material Co., Ltd, Yingkou, Liaoning

**【Abstract】** Teaching programming for robotic arms is widely used in industrial automation, but traditional methods suffer from issues such as poor flexibility and low efficiency. To achieve more efficient and intelligent teaching programming, reinforcement learning technology has been introduced. By constructing a reinforcement learning model that includes the state space, action space, and reward function, the teaching programming process for robotic arms can be optimized adaptively. The model uses the joint angles of the robotic arm and the end effector's position as state inputs, and the joint motion commands as action outputs. A reward mechanism is established based on the task completion and energy consumption. Simulation and experimental results show that this optimization method significantly enhances the efficiency of teaching programming for robotic arms, reduces operational complexity, and provides effective technical support for the application of robotic arms in complex task scenarios.

**【Keywords】** Reinforcement learning; Robotic arm; Teaching programming; Adaptive optimization; Industrial automation

#### 引言

随着工业自动化程度不断提高，机械臂应用愈发广泛。示教编程作为机械臂任务执行的基础，其编程效率和准确性直接影响生产质量与效率。传统示教编程方式依赖人工重复操作，难以适应复杂多变的任务需求和动态环境。强化学习具备自主学习与决策能力，能够通过与环境交互不断优化策略。将强化学习应用于机械臂示教编程自适应优化，可有效提升机械臂智能化水平，具有重要的研究价值和应用前景。

#### 1 优化问题剖析

在智能制造中，机械臂作为核心执行单元，其编程复杂性随着生产需求升级而增加。复杂任务要求

机械臂高精度动作规划，如精密电子装配中元器件的精确抓取与放置，以及多自由度协调以避免干涉。传统编程方法难以处理复杂轨迹，导致动作问题。动态环境的不确定性，如 AGV 物流车的随机移动，使得传统程序难以实时适应，可能造成生产中断。机械臂的物理特性，如冗余自由度和动力学约束，也增加了编程难度，需要工程师仔细筛选关节角度并考虑力矩限制。

从生产效率角度看，人工编程的局限性愈发明显。传统示教编程依赖技术人员通过示教器手动记录每个动作节点，这一过程耗时耗力。在 3C 产品生产线中，每当产品设计迭代，机械臂需要重新编程调

试,平均耗时长达数天。这种低效的编程模式不仅拉长了产品研发周期,还增加了试错成本。随着个性化定制需求的兴起,小批量、多品种的生产模式逐渐成为主流,传统编程方式难以满足频繁调整生产工艺的需求<sup>[1]</sup>。家具制造企业面对客户多样化的定制订单,机械臂需要频繁切换钻孔、打磨等工艺程序,传统编程的低效率严重制约了产能释放。人工编程的经验依赖性强,不同工程师的编程水平差异会导致机械臂运行效果参差不齐,难以形成标准化的生产流程。

更深层次的挑战在于,传统编程缺乏自适应能力。在半导体制造的洁净车间,温湿度、气压等环境参数的细微变化都可能影响机械臂的运行精度,但预先编写的程序无法感知这些变化并作出调整。在食品分拣场景中,不同批次物料的尺寸、重量存在差异,传统编程机械臂难以灵活调整抓取力度,容易造成物料损坏。这种“刚性”编程模式无法应对生产过程中的不确定性,使得机械臂在复杂工况下的应用受限<sup>[2]</sup>。随着工业4.0的推进,智能制造对生产系统的柔性化、智能化提出更高要求,传统示教编程方式已难以满足实时决策、动态优化的需求,迫切需要引入新的技术手段,实现机械臂编程从“人工驱动”向“智能自主”的转变。

## 2 强化学习模型构建

构建适用于机械臂示教编程的强化学习模型,本质上是为机械臂赋予自主学习与决策的“大脑”。在状态空间的构建过程中,需全面捕捉机械臂运行的核心特征。机械臂各关节角度是反映其姿态的基础参数,不同的关节组合决定了末端执行器的空间位置;角速度参数则体现了机械臂的运动趋势,能够帮助模型预测下一时刻的状态变化。而末端执行器的位置和姿态信息至关重要,例如在焊接任务中,焊枪的空间位姿直接影响焊缝质量。环境感知数据也应纳入状态空间,如激光雷达检测到的障碍物距离、视觉传感器识别的目标位置等,这些信息使机械臂能够“感知”周围环境,为决策提供依据。通过整合这些多维度数据,状态空间能够完整描述机械臂与环境的交互状态,为学习过程奠定基础。

动作空间的设计需兼顾机械臂的运动特性与控制需求。关节角度变化量是最直接的控制指令,通过细微调整各关节角度,机械臂可以实现末端执行器的精确定位;速度控制指令则用于调节机械臂的运动节奏,在快速移动阶段采用高速度以提升效率,在

接近目标时降低速度确保精度。还可引入加速度控制指令,优化机械臂的启停过程,减少惯性冲击<sup>[3]</sup>。对于具有多个自由度的机械臂,动作空间的维度较高,如何合理设计动作指令的组合方式是关键。在冗余自由度机械臂中,可设计优先策略,在保证任务执行的前提下,优先选择能耗低、关节磨损小的动作组合。为了确保机械臂运行的安全性,需对动作空间进行约束,如限制关节角度变化的最大步长,避免机械臂因动作突变引发碰撞。

强化学习模型的核心在于奖励函数的设计,它指导机械臂学习最佳策略。任务完成度是关键,如装配任务中零件正确安装和搬运任务中物料准确送达应获得高奖励。运动时间短和能耗低也是奖励的重要指标,以促进高效节能的运动策略。安全方面,碰撞或关节超限应受到负奖励,强化安全意识<sup>[4]</sup>。奖励函数需具备动态调整能力,初期放宽奖励以鼓励探索,后期提高奖励以引导收敛至最优策略。模型通过平衡奖励因子,学习到综合优化策略,兼顾效率、安全与质量。

## 3 优化方案实施

强化学习模型优化机械臂编程,包含训练和应用两个阶段,旨在提升机械臂智能决策。训练阶段利用模拟环境作为“数字训练场”,创建高度仿真的虚拟场景,模拟复杂工况。机械臂在模拟环境中根据状态选择动作,观察奖励信号,通过“试错”学习有效行为。在模拟焊接中,机械臂学习调整焊枪姿态以优化焊接效果。经验回放机制存储历史经验,防止学习陷入局部最优,提升学习效率和策略泛化。

实际应用阶段是将训练成果转化为生产力的关键环节。机械臂通过传感器实时感知环境状态,如视觉传感器识别工件位置、力传感器检测接触力大小,这些感知数据作为输入传递给训练好的强化学习模型。模型基于当前状态快速计算并输出最优动作指令,控制机械臂执行任务<sup>[5]</sup>。在物流分拣场景中,机械臂能够根据实时识别的包裹尺寸和位置,动态调整抓取姿态和力度,实现高效分拣。值得注意的是,实际环境往往比模拟环境更复杂,存在传感器噪声、模型误差等干扰因素。优化方案设计了动态更新机制,机械臂在执行任务过程中,会持续收集新的经验数据,并定期对模型进行微调。

优化方案的实施还需解决实际应用中的工程化问题。在硬件层面,需确保传感器与执行器的高精度

配合,例如采用高精度编码器实时反馈关节角度,利用伺服电机实现精确的速度和力矩控制。在软件层面,要优化模型的部署方式,将强化学习算法移植到嵌入式控制器中,满足机械臂实时决策的计算需求。为了保障系统的安全性,需建立多层防护机制,当模型输出的动作指令可能引发危险时,硬件限位和软件安全策略会立即介入,强制机械臂停止或调整动作<sup>[6]</sup>。人机协同也是优化方案的重要组成部分,操作人员可以通过可视化界面监控机械臂的运行状态,并在必要时进行人工干预,实现智能决策与人工经验的有机结合,进一步提升生产系统的可靠性和灵活性。

#### 4 验证与结果

通过仿真实验和实际场景测试,验证了基于强化学习的机械臂示教编程优化方案的有效性和实用性。仿真实验构建了包含多种工业场景的虚拟测试平台。在汽车零部件装配模拟中,机械臂需在动态障碍物中规划路径,精准抓取零件。强化学习优化的机械臂能实时调整轨迹,缩短任务时间。在曲面打磨仿真中,优化后的机械臂能自适应调整打磨姿态和力度,提升工件表面质量。这些实验表明强化学习在任务规划和环境适应方面的优势。

实际工业场景测试进一步验证了优化方案的工程实用性。在3C产品生产线的测试中,机械臂承担精密元器件的贴装任务。面对产品型号频繁切换、元器件规格多样的挑战,传统编程方式需要技术人员花费数小时重新调试程序;而采用强化学习优化的机械臂,通过对历史任务数据的学习,能够在几分钟内完成新任务的参数调整,快速适应生产变化。在实际运行过程中,机械臂不仅保持了较高的贴装精度,还将生产效率提升,同时降低了因编程错误导致的元器件损耗率<sup>[7]</sup>。在食品加工行业的包装生产线测试中,机械臂需要处理不同形状、重量的食品包装任务。强化学习优化的机械臂能够根据实时检测的物料特性,自动调整抓取力度和包装动作,有效避免了食品损坏,产品合格率提高。

通过仿真与实际测试的双重验证,基于强化学习的机械臂示教编程优化方案展现出强大的技术优势。与传统编程方法相比,该方案赋予了机械臂自主学习、动态决策的能力,使其能够在复杂多变的生产环境中灵活应对各种挑战。这种智能化的编程方式不仅提升了生产效率和产品质量,还降低了对人工编程经验的依赖,为智能制造的发展提供了新的技

术路径<sup>[8]</sup>。随着技术的不断完善,该优化方案有望在更多工业领域得到推广应用,推动机械臂从“自动化设备”向“智能作业伙伴”的转型,助力制造业向更高水平的智能化迈进。

#### 5 结语

强化学习为机械臂示教编程自适应优化提供了新途径,有效解决了传统编程方式的局限性。当前已在提升编程效率和任务适应性上取得良好成果。未来,随着强化学习算法不断发展和硬件性能提升,可进一步拓展优化应用场景,如在危险、高精度等极端环境下的应用。结合多模态感知技术,提升机械臂对复杂环境的感知与理解能力,推动机械臂智能化水平迈向新高度。

#### 参考文献

- [1] 姜龙云,高永强,王文成,等.带电抢修机器人液压柔性关节多机械臂协同控制[J].液压气动与密封,2024,44(10):49-54.
- [2] 刘瞰东,张馨月,林晨滢,等.基于分段动态运动基元的机械臂轨迹学习与避障方法[J].机器人,2024,46(03):275-283.
- [3] 王天瑞,陶平.双机械臂协作最优装配位置及多目标轨迹优化[J].机械传动,2024,48(01):39-45.
- [4] 倪浩君,温秀兰,顾云阳,等.基于LabVIEW的机械臂控制与图像处理示教平台设计[J].计量与测试技术,2023,50(11):17-19.
- [5] 金慧萍,兰天翔,刘腾,等.计算机视觉的桌面型机械臂实验平台设计[J].实验室研究与探索,2023,42(06):46-50.
- [6] 贺鑫鑫,雷同飞,王瑞龙.基于CAD的离线仿真编程焊接机器人生产线设计分析[J].机电产品开发与创新,2022,35(06):19-22.
- [7] 蒋浩,宁芳,张家奇.基于AHP-熵权TOPSIS法的轻载机械臂示教系统交互质量评价方法[J].机电产品开发与创新,2022,35(04):157-160+167.
- [8] 杨亮亮,窦岩杰,张晖.基于FreeRTOS和emWin图形库的工业机械臂示教器软件设计[J].软件工程,2022,25(06):40-44.

**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**