面向机电产品的深度强化学习多目标协同优化设计

傅亮星
广州恒龙建设工程有限公司　广东广州　510700

摘　要：机电产品的优化工作需要在性能、成本与可靠性等多个目标之间达成平衡，传统的优化手段常常会在目标冲突问题以及搜索效率低下的问题上遇到困难。深度强化学习同时具备强大的特征提取能力与序列决策能力，这为多目标的协同优化提供了一条新的途径。本文以机电产品在整个生命周期中的需求为指引，引入深度强化学习框架来构建优化模型。通过剖析设计过程中各目标的耦合关系，设计出与之适配的状态空间、动作空间以及奖励机制，进而实现各目标的动态平衡。此项研究打破了传统方法的限制，为机电产品设计提供了高效且精准的解决方案，对于提升产品的核心竞争力有着重要的意义。

关键词：机电产品；深度强化学习；多目标优化；协同设计；优化模型

引言

在智能制造的推动作用下，机电产品朝着高精度、多功能的方向发展，其设计工作需要兼顾结构、性能等多个目标。像遗传算法、粒子群优化这类传统的方法，在处理多目标耦合问题时，容易陷入局部最优的情况，探索效率也比较低。深度强化学习将深度学习的感知能力和强化学习的决策能力相融合，能够在复杂的环境中自主学习到最优策略，这为解决上述难题提供了机会。目前，深度强化学习在机器人控制领域已经取得了明显的成效，但在机电产品设计中的应用还不够深入，还面临着目标量化等方面的问题。所以，开展相关的研究既具有理论价值，又具有工程意义，本文将构建有针对性的优化框架来为设计实践提供支持。

1 机电产品多目标协同优化设计的核心需求与挑战

1.1 多目标的内涵与耦合关系

机电产品设计的核心目标涉及性能、经济性、可靠性等多个方面。性能目标一般包含运行精度、负载能力以及响应速度等，这些直接决定了产品的使用价值；经济性目标包括研发成本、制造成本以及运维成本等，是企业控制投入的关键所在；可靠性目标则与产品的使用寿命和故障发生率相关，会对用户体验和市场口碑产生影响。各个目标之间存在着明显的竞争与协同关系，比如进行轻量化设计能够降低材料成本和能耗，但可能会使结构强度降低；提高加工精度可以提升产品的性能，却会增加制造工艺的难度和成本。

1.2 传统方法的应用局限

遗传算法、模拟退火、粒子群优化等传统的智能算法，在机电产品优化设计领域已经有了较为广泛的应用。例如在机械结构参数优化中，遗传算法通过模拟生物进化的过程来寻找最优解，在简单结构的优化中取得了一定的效果。不过，随着机电产品复杂程度的提高，这些方法的局限性越来越突出。当面对包含数十甚至上百个设计变量的高维优化问题时，传统算法的搜索空间会急剧扩大，使得优化效率大幅降低，难以在合理的时间内获得令人满意的结果。同时，在处理强度约束、几何约束等强约束条件时，容易出现可行解丢失或者搜索方向偏离的情况。

1.3 智能制造背景下的新需求

随着智能制造与工业 4.0 的推进，对机电产品优化设计产生了一系列新的要求。个性化定制成为市场的主要发展趋向之一，不同用户对于产品功能、性能以及外观的需求存在较大差别，这就需要优化设计方法具备充足的灵活性，能够迅速对优化目标和参数进行调整，以契合多样化的设计需求。全生命周期设计理念的推广，让优化设计不再仅仅局限于研发环节，还需要兼顾生产、装配、运维、回收等后续阶段。比如在设计阶段需要考虑零部件的可加工特性，降低生产的难度和成本；同时要便于后期的维护，减少运维方面的费用。

2 深度强化学习在多目标优化中应用的技术根基

2.1 核心架构与运行机制

深度强化学习的核心架构由智能体、环境、状态、动

作和奖励这五大要素组成。智能体是做出决策的主体，承担着在与环境的交互过程中不断学习优化策略的任务；环境是智能体所处的外部环境，在机电产品优化场景中可对应设计空间和约束条件；状态是环境的具体体现，反映出当前设计方案的参数和性能；动作是智能体为改变状态而采取的操作，也就是对设计参数进行调整；奖励是环境对动作的反馈，用于评判设计方案的好坏。典型的深度强化学习算法包含深度 Q 网络、策略梯度等。深度 Q 网络借助神经网络来近似 Q 值函数，实现对动作价值的评估和优化；策略梯度则直接对策略函数进行参数更新，以寻找最优的决策策略。其运行机制本质上是智能体在环境中通过“尝试错误”不断积累经验，运用深度学习技术来处理高维状态信息。

2.2 多目标场景的改进路径

传统的单目标强化学习算法在多目标优化场景中存在显著的适配问题。单目标算法以最大化单一奖励为目标，而在多目标优化中，多个目标往往相互抵触，无法通过单一奖励准确地衡量设计方案的综合性能，直接应用容易导致优化结果偏向某一个目标，忽视其他重要需求。帕累托最优理论为多目标强化学习提供了重要的支持，基于该理论构建优化目标，旨在寻找非支配解集合，即不存在任何一个解在所有目标上都优于另一个解的情况。奖励机制的设计是多目标强化学习的核心环节之一，权重自适应调整策略可以根据不同设计阶段的目标优先级，动态地改变各目标奖励的权重；多奖励融合则通过一定的数学方法将多个目标的奖励整合为综合奖励信号，既保留各目标的独立信息，又能实现对整体性能的评价。

2.3 与机电设计的适配性

深度强化学习和机电产品设计流程之间存在内在的对应联系，能够较好地匹配设计需求。机电产品设计一般遵循“参数设置 – 性能剖析 – 方案修改”的循环流程，这与深度强化学习“状态察觉 – 动作实施 – 奖励回馈”的交互过程十分契合。设计流程里的参数设置相当于智能体的动作，性能剖析的结果构成环境回馈的状态与奖励，方案修改则是智能体依据奖励进行的策略更新。机电产品设计所积累的大量历史数据和仿真数据，为强化学习模型的训练提供了充分的数据源。这些数据包含不同设计参数的组合以及相应的性能指标，可用于构建训练样本，增强模型的学习效果和泛化能力。

3 面向机电产品的深度强化学习多目标协同优化模型构建

3.1 状态与动作空间定义

状态空间和动作空间的合理界定，是搭建深度强化学习优化模型的根基，会直接对模型的优化效果和收敛速度产生影响。状态空间的构建需要紧密围绕机电产品的设计需求，全面包含反映产品性能的关键参数。拿机械结构件设计来说，状态空间应涵盖结构尺寸参数、材料属性参数以及对应的强度、刚度、重量等性能指标，形成能够完整呈现设计方案的高维状态向量。动作空间的界定则根据设计变量的可调范围和特性，采用离散和连续相结合的方式。对于螺栓型号、材料类型等离散变量，直接以类别形式来界定动作；对于尺寸参数等连续变量，把其取值范围划分成合理区间，通过连续数值的调整来实现动作表达。

3.2 奖励函数设计与优化

奖励函数是指引深度强化学习智能体学习方向的核心，其设计水平直接决定优化结果是否符合实际需求。多目标导向的奖励函数需要依据各设计目标的优先级和量化标准，搭建多维度的奖励体系。首先明确各目标的量化指标，比如将性能目标量化为精度误差、响应时间等，成本目标量化为单位制造成本等，然后根据目标的重要性分配初始权重，形成综合奖励函数。为解决固定权重难以适应设计过程变化的问题，采用自适应权重策略。

3.3 训练策略与收敛保障

优化模型的训练策略设计，是保证模型稳定收敛并获得最优解的关键。经验回放和目标网络双机制的引入，能够有效提高模型训练的稳定性。经验回放机制把智能体与环境交互产生的经验数据存储在回放缓冲区，训练时随机抽取样本进行学习，打破数据之间的时序相关性，防止模型训练出现震荡。目标网络与主网络结构相同，用于计算目标 Q 值，其参数定期从主网络复制更新，减缓目标值的波动，使训练过程更加平稳。采用分段训练策略，分阶段优化模型的探索和利用能力。

4 机电产品多目标协同优化的求解流程与实现路径

4.1 数据预处理及特征工程

数据作为深度强化学习模型训练的根基，其质量高低对模型学习成效影响显著。机电产品设计数据拥有广泛来源，像历史设计方案、仿真试验数据、实际运行监测数据等均包含在内。首先要开展全面的数据采集工作，保证数据能够涵

盖不同类型、不同工况下的产品设计信息。数据预处理环节意义重大，需要对采集到的数据开展清洗操作，去除掉因测量误差、设备故障等因素产生的异常数据；针对缺失值，要选用均值填充、插值法等恰当方式进行填补，以此保障数据的完整性。

4.2 优化求解流程

运用深度强化学习进行机电产品多目标协同优化求解，需依照清晰的流程步骤实施。首先要搭建机电产品设计的仿真环境，该环境应具备准确模拟设计参数与产品性能间映射关系的能力，并实现与强化学习模型的实时交互。仿真环境需包含参数输入接口、性能计算模块以及结果输出接口，当智能体输入设计参数后，环境能够迅速计算并反馈相应的性能指标与奖励值。随后要确定模型训练的参数设置，其中包括学习率、批次大小、探索率等超参数的初始值，这些参数需要结合模型特点与设计问题进行合理配置。

4.3 优化结果评估与筛选

优化结果的评估和筛选是保障设计方案具备实用性的重要环节，需要构建包含多维度的评估指标体系。该体系应涵盖产品性能指标，例如精度、强度、效率等；经济性指标，如研发成本、制造成本、生命周期成本等；可靠性指标，像故障率、使用寿命、维护周期等。借助这些指标对优化得到的设计方案进行全面评判，防止出现因单一指标评价导致方案具有片面性的情况。采用帕累托前沿分析方法对多目标优化结果进行处理，筛选出非支配最优解集合。

5 深度强化学习优化方法的泛化性与扩展性研究

5.1 模型迁移策略

机电产品存在多种类型，涵盖机床、工程机械、电子产品等不同类别，其设计兼具共性与个性化差异。共性主要体现为均需考虑性能、成本等基础目标，且优化流程具有相似特征；个性化差异则表现为具体设计参数、约束条件以及目标优先级的不同。基于迁移学习的模型参数复用方式，可有效提升深度强化学习模型的泛化能力。针对新类型机电产品的优化需求，可将训练成熟的模型参数作为初始数值，无需重新开展从零开始的训练工作。

5.2 模型自适应更新机制

机电产品的设计需求并非固定不变，在设计过程中可能会因市场需求改变、技术条件提升等因素而发生调整。建立设计需求变化的识别与量化手段，是实现模型自适应更新的前提条件。通过对设计目标权重、约束条件阈值等关键参数的变化进行监测，识别需求变更类型，并将其量化为具体的参数调整指标。针对需求变化情况，提出模型参数的增量更新策略，避免对模型进行全量的重新训练。依据需求变更的范围与影响程度，确定需要更新的网络模块与参数范围，仅对相关部分进行调整，以实现模型的快速响应。

5.3 与其他技术的融合路径

深度强化学习并非独立的优化技术，与其他优化技术进行融合能够进一步提高其优化能力。深度强化学习与传统优化算法的混合优化模式，是重要的融合方向之一。可采用“传统算法初始化 + 强化学习精优化”的策略，利用遗传算法等传统方法快速搜索到较优的解空间区域，再通过深度强化学习在该区域内进行精准搜索，以兼顾优化效率与精度。数字孪生技术的发展为机电产品优化提供了新的支持，将深度强化学习与数字孪生技术相融合，可构建与物理实体完全对应的虚拟模型。

6 结论

本文围绕面向机电产品的深度强化学习多目标协同优化设计开展研究，系统分析了机电产品多目标优化的核心需求以及传统方法存在的局限性，阐述了深度强化学习应用于多目标优化的技术基础与适配性。在此基础上构建了有针对性的深度强化学习多目标协同优化模型，明确了状态空间、动作空间及奖励函数的设计方法，提出了稳定高效的模型训练策略。设计了完整的优化求解流程，包含数据预处理、模型训练、结果评估与筛选等关键环节，并对模型泛化性与扩展性的提升路径进行了探讨。

参考文献：

[1] 张洪浩 . 如何做好旧机电产品进口检验 [J]. 中国海关 ,2025,(11):50.

[2] 周天朋 , 徐明鸽 , 淡丽艳 . 机电产品与火工品组合系统可靠性综合评估方法研究 [J]. 环境技术 ,2025,43(10):21–26.

[3] 王明宇 , 李涛 , 杨晨 , 等 . 基于孪生数据模型的机电产品生命周期环境影响评价方法 [J]. 大连理工大学学报 ,2025,65(05):501–510.

[4] 钟莹莹 . 依托数字经济推动中国机电产品出口的对策研究 [J]. 老字号品牌营销 ,2025,(15):64–66.

[5] 秦闯 , 辛凯 , 高天奇 , 等 . 基于机器视觉的机电产品质量检测技术研究 [J]. 产品可靠性报告 ,2025,(06):65–66.

作者简介： 傅亮星 (1985.09–), 男 , 汉 , 江西人 , 大学本科 , 研究方向：机电工程