

融合多目标优化的强化学习无人机路径规划策略设计

徐焕时 尹子豪

新疆理工职业大学

DOI:10.12238/er.v8i12.6624

[摘要] 无人机路径规划是确保无人机任务成功执行的关键技术之一。现有的路径规划方法在面对复杂动态环境时存在局限性，难以兼顾多目标优化需求。本文提出了一种基于多目标优化的强化学习路径规划策略，结合强化学习的自适应特性和多目标优化方法，旨在解决传统算法在动态障碍物、能耗最小化和飞行效率等多重需求下的不足。通过实验验证，所提出的策略能够在复杂环境中实现快速、准确的路径规划，显著提高飞行效率，并有效降低能量消耗。该方法为无人机在复杂环境中的自主飞行提供了更具潜力的解决方案。

[关键词] 无人机路径规划；多目标优化；强化学习；自适应性；飞行效率

中图分类号：V279 文献标识码：A

Design of a Reinforcement Learning-Based Path Planning Strategy for UAVs with Multi-Objective Optimization

Huanshi Xu, Zihao Yin

Xinjiang Vocational University of Technology

Abstract: Unmanned Aerial Vehicle (UAV) path planning is a critical technology for ensuring the successful execution of UAV missions. Existing path planning methods exhibit limitations in complex dynamic environments, struggling to address the demands of multi-objective optimization. This paper proposes a novel path planning strategy based on reinforcement learning with multi-objective optimization. By integrating the adaptive characteristics of reinforcement learning with multi-objective optimization techniques, the proposed approach addresses the shortcomings of traditional algorithms in handling dynamic obstacles, minimizing energy consumption, and maximizing flight efficiency. Experimental results demonstrate that the proposed strategy achieves rapid and accurate path planning in complex environments, significantly enhancing flight efficiency while effectively reducing energy consumption. This method offers a promising solution for autonomous UAV navigation in challenging environments.

Keywords: UAV path planning; multi-objective optimization; reinforcement learning; adaptability; flight efficiency

引言

无人机路径规划是无人机自主飞行和任务执行的关键技术之一。在复杂的动态环境中，传统的路径规划方法往往无法满足实时性和多目标优化的需求，特别是在面对动态障碍物和复杂任务时，现有算法表现出一定的局限性。为了应对这一挑战，强化学习与多目标优化方法的结合为路径规划提供了新的解决思路。强化学习能够根据环境反馈不断优化决策，具备较强的自适应性，而多目标优化则能够同时兼顾飞行效率、能耗和避障等多个目标。本文提出了一种基于多目标强化学习的无人机路径规划策略，旨在提升无人机在复杂环境中的规划能力，确保任务的高效和安全执行。

1 无人机路径规划的现有挑战与不足

1.1 传统路径规划算法的局限性

现有的无人机路径规划方法在复杂动态环境中往往面临明显的局限性。在静态环境下，许多经典算法能够较好地实现路径规划，但当遇到动态障碍物或复杂的避障需求时，它们的计算量迅速增加，导致效率降低。随着飞行环境的变化，原本预设的路径可能需要不断调整，增加了计算的负担和延迟。尤其是在需要多目标优化的情况下，现有方法往往只能聚焦于单一目标，如最短路径或最小能耗，无法同时处理多个优化目标，从而限制了路径规划的效果。某些优化算法可能陷入局部最优解，导致无法找到最优的全局路径，进一步降低了其在复杂环境中的适应性。面对动态变化的飞行任务，迫切需要新的算法来提升路径规划的实时性和多目标优化能力。

1.2 多目标优化在无人机路径中的应用问题

多目标优化方法在无人机路径规划中应用逐渐增多，主要是因为无人机任务通常需要同时满足多个目标，如飞行时间、能耗和避障能力等。传统的路径规划方法往往将这些目标单独考虑，而多目标优化则能够在多个目标之间找到一种折中解。然而，实际应用中，多目标优化面临一些挑战。如何合理地设计多目标的权重和优化函数，使得各个目标的权衡能够根据任务的不同需求动态调整，仍然是一个难题。路径规划过程中的计算复杂度较高，尤其是在高维度、多障碍物的环境中，如何有效地减少计算量，提高算法的实时性和适应性，依然是一个亟待解决的问题^[1]。多目标优化的最终路径解是否能满足所有目标的需求，并且在复杂环境下仍保持稳定，这也是该方法应用中的关键挑战。针对多目标优化在无人机路径中的应用，亟需进行进一步算法改进和优化。

2 强化学习在无人机路径规划中的应用与优势

2.1 强化学习的基本原理与优势

强化学习是一种基于试错法的机器学习方法，它通过智能体与环境的交互，不断调整其行为策略，以最大化累积的奖励。在无人机路径规划中，强化学习通过将路径选择问题转化为状态—动作的决策过程，使得无人机能够在动态环境中自主选择最佳路径。其核心优势在于能够通过不断的反馈来优化其路径选择策略，从而应对复杂和变化的飞行环境^[2]。与传统的规划方法不同，强化学习无需事先知道整个环境的具体信息，只需依赖于实时的状态观测和环境反馈，通过探索和学习不断优化决策过程。强化学习具有较强的自适应能力，能够在面对复杂障碍、突发事件或飞行环境变化时，迅速调整飞行策略，使得其在动态环境中的应用具有明显优势。

2.2 强化学习与多目标优化的融合方法

强化学习与多目标优化的融合可以有效解决传统路径规划算法中单一目标优化的问题，提升无人机在复杂环境中的任务执行能力。在该融合方法中，强化学习不仅能够通过奖励机制自主学习路径规划策略，还可以在多目标优化框架下进行多任务协同优化。具体而言，每个目标（如飞行时间、能耗、避障等）可以设定一个相应的奖励函数，强化学习智能体通过同时优化多个目标，逐步收敛到一个最优的路径规划方案。融合方法还可以利用多目标的评价机制，在面对不同飞行环境时，通过动态调整权重来适应不同任务需求。在能源紧张的情况下，系统可以自动增加对能耗的优化权重，从而使无人机选择最节能的路径；在避障要求较高的情况下，系统则倾向于选择避障能力更强的路径。通过这种方法，强化学习可以在多目标优化中充分发挥其优势，实现路径规划的多维度最优化。

2.3 基于多目标优化的强化学习路径规划策略设计

2.3.1 目标设定与算法模型

设计基于多目标优化的强化学习路径规划策略时，首先

需要明确任务的多个优化目标，这些目标通常包括最短飞行时间、最低能耗、最强避障能力等。在此基础上，可以构建一个多目标强化学习模型，其中每个目标对应一个不同的奖励函数。通过设定合适的目标函数，可以确保无人机在飞行过程中同时考虑多个优化目标，进而实现综合性能的提升。采用深度强化学习（Deep Reinforcement Learning, DRL）方法来进行路径规划，可以利用深度神经网络来近似状态值函数和策略函数，使得模型能够处理更复杂的环境信息和更大规模的状态空间。这种基于深度强化学习的模型能够更好地捕捉复杂环境中的模式，从而为无人机路径规划提供更为精准的决策支持。

2.3.2 多目标优化策略的实施与改进

在多目标优化策略的实施过程中，平衡各个目标之间的冲突是路径规划中的核心问题。在强化学习框架下，如何处理多个优化目标的相互影响是实现高效路径规划的关键。一个有效的解决方案是采用加权法和法，通过为每个目标分配适当的权重，将各个目标的奖励函数合并成一个综合目标函数。在实际应用中，这些权重可以根据任务需求进行动态调整，以确保在不同的环境和飞行任务中，路径规划能够平衡飞行时间、能量消耗、避障能力等多个目标的优化^[3]。多策略学习方法也能够同一时刻同时优化不同目标，从而减少目标之间的竞争和冲突，提高整体性能。为了进一步提升算法的效率，可以结合模糊控制算法或自适应优化技术，对奖励函数进行实时动态调整。这些方法可以帮助模型在面对动态环境变化时，快速收敛并提高规划精度。通过优化后的多目标强化学习模型，能够在保证路径规划效率的同时满足多维度的任务需求，提高无人机的综合性能。

3 强化学习路径规划策略的实验验证与性能分析

3.1 实验环境与测试方法

为了验证多目标强化学习路径规划策略的有效性，实验通常采用多种模拟环境进行测试。这些环境设计的目的是尽可能贴近实际应用场景，考察路径规划策略在不同条件下的表现^[4]。在动态环境中，通过添加随机障碍物和复杂天气变化来模拟不确定性和实时障碍。这些变化要求无人机在飞行中不断做出反应，调整路径以保证飞行安全与任务完成。实验环境通常包括不同地形类型，如城市街区、山区、森林等，模拟无人机可能面临的多样化地理和飞行条件。对于不同的飞行任务，实验设计了包括高速避障、能量消耗最小化、任务时效性等多项任务要求，并将其作为路径规划的评价标准。测试方法覆盖了路径规划多个维度，包括时效性、飞行精度、避障能力、能耗等方面的综合评估。为了全面衡量新策略的优势，实验还将这些指标与传统路径规划算法进行对比，进行多维度的性能测试与分析。

3.2 实验结果与策略性能分析

实验结果表明，基于多目标强化学习的路径规划策略在

多个评估指标上均表现出优于传统路径规划算法的性能。在动态环境下，强化学习策略能够在障碍物变化和任务需求波动的情况下，实时调整飞行路径。相比传统路径规划方法，强化学习能够更快速地识别新障碍并作出避让决策，保证飞行安全。在飞行时间和能量消耗方面，强化学习策略显示出更好的优化效果。通过对比分析发现，强化学习在确保飞行时间最短的同时还能够大幅降低不必要的能量消耗，从而提高了任务的整体效率。所设计的多目标强化学习策略在避障能力方面表现突出，能够有效地在复杂环境中避开障碍物，避免飞行事故的发生。通过这些实验数据分析，可以看出，该策略在动态和多变的飞行环境中具有较强的适应性和灵活性，充分证明了其在实际无人机路径规划中的可行性与优势。见表 1：

表 1 国内无人机路径规划策略比较测试结果

路径规划策略	飞行时间 (分钟)	能量消耗 (kWh)	避障成功率 (%)	任务时效性 (分钟)
传统算法	20.5	2.3	85	25
强化学习策略	18.3	1.9	95	22
改进遗传算法	19.1	2.0	90	24

数据来源：实验室模拟测试，数据采集自 2024 年 1 月—2024 年 3 月的无人机路径规划实验，测试环境包括城市、山区与复杂天气条件。

4 提升无人机路径规划系统适应性的技术探索

4.1 动态环境下路径规划的应对策略

在动态环境中，路径规划需要考虑飞行过程中不断变化的障碍物、天气、飞行区域及任务需求等多个因素。传统路径规划方法通常依赖于静态环境数据，难以适应这种快速变化的环境。提高路径规划系统的适应性显得尤为重要。强化学习在这种环境下具有独特的优势，因为它能够通过与环境交互持续学习，并根据新的感知数据快速调整路径决策。无人机可以使用雷达、视觉传感器等实时感知系统，收集当前飞行状态的信息，并将这些数据反馈到路径规划系统中。这些数据包括动态障碍物的位置变化、气象信息的更新以及飞行过程中其他潜在的危险因素。强化学习算法根据这些实时数据调整规划的路径，不断优化飞行路线。智能路径规划系统还可以结合云计算和边缘计算技术，将计算任务分散到多个处理单元，实现实时路径优化和决策支持。这种结合实时感知和分布式计算的策略，使得无人机能够更加高效、灵活地应对动态环境变化，提高飞行安全性和任务执行效率。

4.2 未来路径规划技术的发展方向

未来路径规划技术的核心目标是实现更高水平的智能化、自动化与自适应无人机操作。随着应用领域的不断拓展，路径规划系统的需求日益复杂多样，传统方法已难以应对新挑战。为提升性能与效率，多目标强化学习仍将发挥重要作用，同时新的算法与技术不断涌现。将强化学习与模糊逻辑或群体智能相结合，可进一步增强系统在复杂环境下的鲁棒性与适应性。模糊逻辑有助于在高不确定性条件下实现灵活且精确的决策；群体智能则通过模拟多智能体协作提升优化能力^[5]。

未来路径规划还将与大数据分析、物联网及云计算平台深度融合，充分利用其计算与信息整合优势，使无人机能够综合考虑实时天气、空域流量、任务优先级等多因素，实现全局最优路径规划。通过多技术协同，路径规划将突破单一任务限制，在复杂环境中灵活应对，为无人机高效运行提供系统化支持。

5 结语

本研究提出的基于多目标强化学习的无人机路径规划策略，成功解决了传统路径规划方法在复杂动态环境中的局限性。通过实验验证，所设计的策略在飞行精度、时效性、能耗和避障能力等多个方面均表现出优于传统算法的性能。强化学习的自适应特性使得无人机能够在变化的环境中快速做出决策，优化飞行路径，提高任务执行效率。未来，随着技术的不断发展，多目标强化学习将进一步增强无人机的自主性和灵活性，为更加复杂的任务环境提供有力支持，推动无人机路径规划技术的应用和发展。

[参考文献]

[1]孙学章,时晨光,吴志锋,等.基于改进三维 A*算法的多无人机隐身路径优化研究[J].战术导弹技术,2025(4):150-160.

[2]范弘悦.基于深度强化学习的无人机自主避障路径规划优化方法[J].中国新技术新产品,2025(15):8-10.

[3]洪港,唐飞,钟爽奎,等.旋翼无人机配网主动巡检路径规划方法[J].电子设计工程,2025,33(15):160-165.

[4]裴明阳,邵康顺,李林青,等.无人机应急救援路径规划建模方法综述与展望[J/OL].华南理工大学学报(自然科学版),1-16[2025-09-13].

[5]廖榆信,王伟,滕卫明,等.基于多目标约束的无人机光顺路径生成全局优化方法[J].浙江大学学报(工学版),2025,59(7):1481-1491.

作者简介：

徐焕时（1999.08—），男，汉族，吉林辽源，硕士，助教，新疆理工职业大学，研究方向为强化学习。

尹子豪（1999.08—），男，汉族，山西阳泉，硕士，助教，新疆理工职业大学，研究方向为深度学习、图像处理。