

DOI: 10.20189/j.cnki.CN/61-1527/E.202501012

无人机航迹规划智能算法研究综述

鲜 勇, 郭 靖, 任乐亮, 李少朋
(火箭军工程大学, 陕西 西安 710025)

摘 要: 针对传统无人机航迹规划算法计算效率低、适应性差的问题, 从算法改进的角度出发, 系统地梳理了国内外在无人机航迹规划领域的研究进展。从全局优化能力、动态适应性、计算效率以及鲁棒性等方面, 阐述了无人机航迹规划智能算法具有的显著优势; 同时, 对航迹规划智能算法进行了详细分类和总结, 从智能优化算法与机器学习算法的几种典型算法改进策略出发, 对比分析了每种典型航迹规划智能算法的优缺点、应用场景以及改进方向; 最后, 结合当前无人机航迹规划所面临的多重挑战, 对无人机航迹规划智能算法未来的研究趋势和潜在应用方向进行了展望, 强调了航迹规划智能算法在提升算法性能、适应复杂环境、满足多样化任务需求以及智能算法融合等方面的重要性, 为无人机航迹规划智能算法的研究与发展提供了参考。

关键词: 无人机; 航迹规划; 智能优化算法; 机器学习算法; 强化学习; 神经网络

中图分类号: TP273

文献标志码: A

文章编号: 2097-2741(2025)01-106-16

Review of Intelligent Algorithms for Unmanned Aerial Vehicle Trajectory Planning

XIAN Yong, GUO Jing, REN Leliang, LI Shaopeng
(Rocket Force University of Engineering, Xi'an 710025, China)

ABSTRACT: To address the problems of low computational efficiency and poor adaptability of conventional unmanned aerial vehicle (UAV) trajectory planning algorithms, progress in domestic and international research in this field was reviewed from the perspective of algorithm improvement. First, the substantial advantages of intelligent algorithms for UAV trajectory planning were elaborated based on global optimization capability, dynamic adaptability, computational efficiency, and robustness. Subsequently, a detailed classification and summary of intelligent algorithms for trajectory planning were presented. Furthermore, the advantages, disadvantages, application scenarios, and improvement directions of each typical intelligent algorithm were comparatively analyzed according to typical algorithm improvement strategies of intelligent optimization and machine learning algorithms. Finally, combined with the multiple challenges currently encountered in UAV trajectory planning, future research trends and potential application directions of intelligent algorithms for UAV trajectory planning were discussed. Additionally, the importance of intelli-

收稿日期: 2024-08-28

第一作者: 鲜勇 (1972—), 男, 教授, 主要从事飞行力学与任务规划研究。E-mail: xy603xy@163.com

引用格式: 鲜勇, 郭靖, 任乐亮, 等. 无人机航迹规划智能算法研究综述[J]. 火箭军工程大学学报, 2025, 39 (1): 106-121. XIAN Yong, GUO Jing, REN Leliang, et al. Review of intelligent algorithms for unmanned aerial vehicle trajectory planning[J]. Journal of Rocket Force University of Engineering, 2025, 39 (1): 106-121.

gent algorithms for trajectory planning in improving algorithm performance, adapting to complex environment, meeting diverse task requirements, and integrating intelligent algorithms were emphasized. The review provides a reference for the research and development of intelligent algorithms for UAV trajectory planning.

KEYWORDS: unmanned aerial vehicle; trajectory planning; intelligent optimization algorithm; machine learning algorithm; reinforcement learning; neural network

随着科技的飞速发展, 无人机 (Unmanned Aerial Vehicle, UAV) 已经在多个领域展现出了巨大的应用潜力, 包括农业勘察^[1]、环境监测^[2]、灾害救援^[3]、物流配送^[4]以及军事侦察^[5]等。在这些应用场景中, 无人机的航迹规划显得尤为重要, 直接关系到任务的执行效率和无人机的飞行安全。传统的航迹规划方法往往依赖于预设的飞行路线和固定的决策逻辑, 难以适应复杂多变的环境和任务需求。因此, 开发一种能够智能地规划无人机航迹的算法成为当下研究的热点。

近年来, 随着人工智能技术的突破, 特别是强化学习^[6]和深度学习^[7]方法的广泛应用, 为无人机航迹规划带来了新的可能性。智能算法^[8]能够通过学习和优化, 使无人机在面对复杂环境时能够实时调整航迹, 以提高飞行效率、安全性和任务完成率。

本文通过介绍几种主流的智能算法, 分析其在无人机航迹规划中的应用效果以及这些算法在实际应用中可能面临的问题, 综合探讨无人机航迹规划智能算法的研究进展、技术挑战及其未来发展方向, 以期为其研究与发展提供借鉴。

1 无人机航迹规划智能算法概述

1.1 无人机航迹规划定义

无人机航迹规划是指无人机在执行特定任务时, 根据任务需求、环境约束和无人机自身性能, 通过算法和计算机技术, 为无人机设计出从起始点到目标点的一条最优轨迹。

无人机航迹规划的主要目的在于实现无人机的安全、高效和精准飞行。通过航迹规划, 可以确保无人机在复杂多变的环境中, 避开潜在的风险和障碍物, 选择最优的飞行路径, 从而保障无人机的飞行安全和任务的顺利完成。

1.2 无人机航迹规划智能算法分类

航迹规划智能算法主要分为两大类, 一类为智能优化算法, 另一类为机器学习算法, 如图1所

示。智能优化算法侧重于通过迭代搜索找到最优解, 而机器学习算法侧重于从数据中学习并做出预测或决策, 以适应复杂环境。

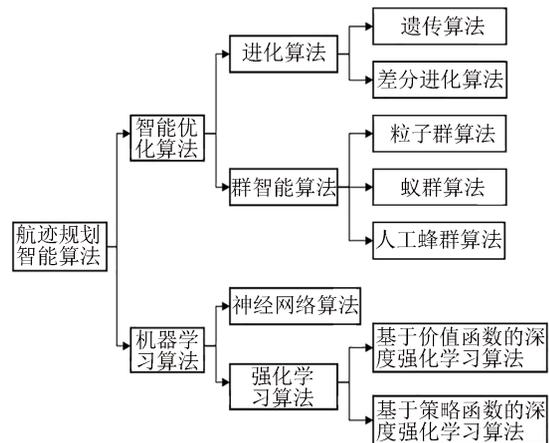


图1 无人机航迹规划智能算法分类

Fig. 1 Classification of intelligent algorithms for UAV trajectory planning

1.3 无人机航迹规划智能算法的优势

在无人机航迹规划中, 传统算法如 A*^[9]、D*^[10]、快速扩展随机树 (Rapidly-exploring Random Tree, RRT) 算法^[11]及人工势场法各具特点。A*算法因其搜索效率高, 能找到最短路径而广受欢迎, 但在高维空间或复杂环境中, 其计算量和存储空间需求急剧增加。D*算法作为 A*的变种, 更适用于动态环境, 能实现实时重规划, 但同样面临计算复杂度高的问题。RRT算法通过快速生成随机树探索空间, 适合解决高维和复杂约束条件下的路径规划问题, 其优势在于路径规划的快速性和对未知环境的适应性, 但生成的路径可能不是最优的, 且需要较多迭代次数。人工势场法通过构造虚拟的吸引和排斥势场引导无人机避障, 实现简单且实时性好, 但算法在局部极小值处易陷入死锁, 且不适用于处理复杂形状障碍物和动态实时路径规划问题。

综上所述, A*和D*算法主要用于静态或低动态环境下的无人机航迹规划, RRT算法则擅长处理高维和复杂环境下的路径规划问题, 而人工势

场法更适用于无人机实时避障。然而，它们均存在计算复杂度高、路径非最优、易陷入局部极小值等问题，因此在实际应用中需根据具体场景和任务需求进行选择或结合其他算法进行优化。

无人机航迹规划智能算法相比传统算法具有显著的优势，主要体现在以下几个方面。

1) 全局优化能力。智能算法，如遗传算法和蚁群算法，通过模拟生物进化或社会行为等自然过程，展现出强大的全局搜索和优化能力，能够在复杂的解空间中有效寻找全局最优解，从而规避局部最优的陷阱。相比之下，传统算法如 Dijkstra、A* 算法在处理复杂环境下的路径规划问题时，可能会受到局部最优解的限制，并且当飞行环境较为复杂时，其最优路径搜索时间会急剧增加，导致效率显著下降。

2) 动态适应性。智能算法具有出色的动态适应性，能够实时感知并响应环境的变化。例如，利用人工神经网络和强化学习技术，无人机可以在飞行中根据获取的实时数据灵活调整航迹，有效应对各种突发状况。而传统算法则多依赖于事先设定的环境模型进行路径规划，相比之下，在面对动态变化的环境时，其适应能力明显较弱。

3) 计算效率。虽然智能算法可能需要消耗更多的计算资源，但通常能更迅速地找到近似最优解，尤其在处理大规模问题时表现出色。与之相对，传统算法在处理复杂难题时，可能需要经历更多的迭代次数、耗费更长的计算时间，在计算效率上不及智能算法。

4) 鲁棒性。智能算法展现出较强的鲁棒性，能够在一定程度上容忍数据中的噪声和不确定性，保持稳定的性能。相比之下，传统算法对输入数据的准确性有着更高的要求，在处理包含噪声或不确定性的数据时，性能可能会受到较大影响，表现出相对较弱的处理能力。

2 无人机航迹规划智能算法介绍

航迹规划作为无人机自主飞行和任务执行中的关键环节，其算法的选择和设计直接关系到无人机的飞行效率和任务完成质量。为深入理解各类航迹规划算法的原理和特性，针对当前主流的航迹规划智能算法进行详细的分析和讨论。

2.1 智能优化算法

智能优化算法主要是通过模拟自然进化过程

或物理现象来搜索最优路径，以提高无人机的飞行效率和安全性。智能优化算法主要分为 2 类，一类为进化算法，包括遗传算法和差分进化算法；另一类是群智能算法，包括粒子群算法、蚁群算法、人工蜂群算法等。

2.1.1 进化算法

1) 遗传算法

1975 年，美国密歇根大学的约翰·霍兰德教授首次系统地提出了遗传算法 (Genetic Algorithm, GA) 的概念^[12]。该算法的运作机制模仿了生物进化过程中的基因重组和自然选择机制，通过模拟基因的遗传、交叉和变异等操作，逐步演化出最优的解决方案。遗传算法通常用于解决复杂的优化问题，特别适用于搜索空间较大或无法使用传统方法进行求解的问题，算法流程如图 2 所示。

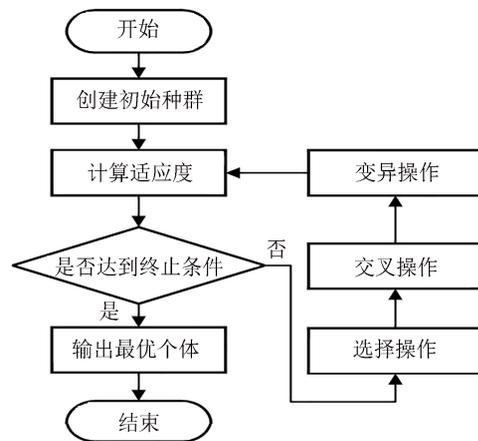


图 2 遗传算法流程

Fig. 2 Flow chart of genetic algorithm

遗传算法的收敛速率和效率也受到其参数配置和起始种群选择的影响。为了提高性能，一般可以根据具体问题设置适当的参数和操作，或结合其他优化算法进行改进。

① 自适应参数调整。王豪等^[13]提出了一种改进的自适应遗传算法，其关键在于将路径长度和拐点数量同时纳入适应度评估体系，并对轮盘赌选择法进行优化。通过设计自适应的交叉算子和变异算子，实现了更有效的个体进化，利于逼近全局最优解。该方法不但缩短了平均最短路径长度，而且减少了路径的拐点数量和算法收敛所需的迭代次数。

② 混合智能优化方法。程泽新等^[14]针对无人机航迹规划过程中后期收敛速度减缓以及容易陷入局部最优的问题，提出在遗传算法中引入差分进化变异策略，旨在增加算法变异的多样性，从

而增强全局搜索能力。同时, 通过将遗传算法与模拟退火算法相结合, 有效地避免了算法陷入局部最优解, 并显著加快了收敛速度。这种结合策略在提高规划效率的同时, 也有效抑制了算法的早熟现象。

③编码和操作策略改进。李文广等^[15]提出了一种多无人机协同侦察航迹规划算法, 通过在交叉和变异操作中引入适应度更新策略, 不仅提高了航迹规划的精度, 还显著提升了规划的效率, 使得多无人机协同侦察任务能够被更加高效、准确地完成。

④适应度函数设计。贾广芝^[16]为了解决遗传算法常见的早熟收敛问题, 对适应度函数进行了改进, 设计了一种能够随着进化迭代次数动态调整的非线性适应度函数。这种改进显著提升了算法的收敛速度, 而且还与稀疏A*算法相结合, 能很好地适应三维在线航迹规划。

从算法改进的角度出发, 自适应参数调整能够确保算法在进化过程中更有效地生成优质个体; 混合智能优化方法则通过结合其他算法的优点, 增强全局搜索能力并避免陷入局部最优; 编码和操作策略改进以及适应度函数设计则分别从操作层面和函数设计层面提升了算法的精度和效率。未来, 随着技术的不断进步, 遗传算法有望在更多领域发挥重要作用, 同时, 如何进一步提高其收敛速度、优化效率以及增强全局搜索能力, 将是研究者们关注的重点方向。

2) 差分进化算法

差分进化算法^[17] (Differential Evolution, DE) 作为一种高效的全局优化算法, 起源于对生物进化理论中自然选择和基因杂交的模拟。该算法由Storn和Price在1995年提出, 旨在解决实数编码下的复杂非线性优化问题。差分进化算法通过随机初始化种群, 经历变异、交叉、选择等迭代步骤优化个体, 收敛至最优解, 算法流程如图3所示。但是在无人机航迹规划时, 算法在后期可能难以保持进化种群的多样性, 面临陷入局部最优的风险。

为了深入理解并运用差分进化算法, 学者们对此进行了大量的研究工作, 探讨了算法的改进策略、性能评估及其在不同领域中的应用。

①结合其他算法。严炜等^[18]结合差分进化算法的全局搜索能力和量子退火算法的快速收敛性, 通过模拟量子退火过程并结合差分进化的变异、

交叉和选择策略, 有效解决了传统航迹规划算法在复杂环境和多约束条件下难以找到高质量飞行路径的问题, 显著提高了无人机航迹规划的效率和质量。

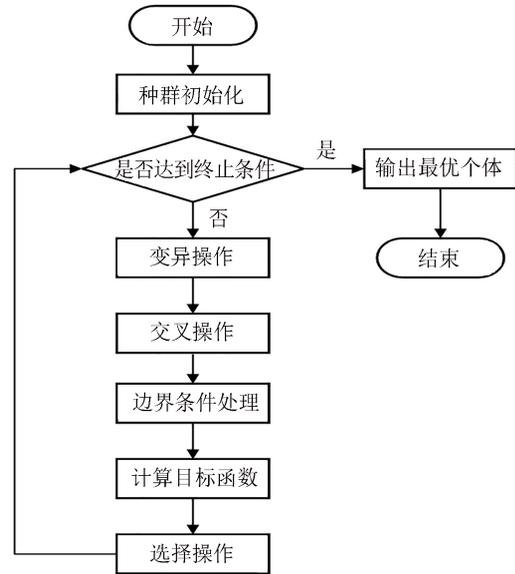


图3 差分进化算法流程

Fig. 3 Flow chart of differential evolution algorithm

②策略与参数调整。张宗豪等^[19]通过引入改进的变异、交叉和选择策略, 对算法进行了优化, 有效解决了传统算法在三维空间中路径搜索复杂度高、计算量大、易陷入局部最优解的问题。

③种群结构改进。Chai等^[20]提出了一种多策略融合差分进化算法, 通过划分子种群并应用不同的变异策略, 平衡了算法全局和局部搜索的能力; 同时, 引入参数自适应策略动态调整关键参数, 成功解决了无人机航迹规划中的多目标约束问题。

表1为典型进化算法在无人机航迹规划应用中的对比与改进。进化算法在无人机航迹规划领域中展现出显著优势, 能有效处理复杂优化问题, 通过模拟自然进化机制实现全局搜索。然而, 算法性能亦受多方面因素影响: 一方面, 算法性能高度依赖于参数配置、适应度函数设计及种群多样性管理, 不当设置易导致算法早熟收敛或计算效率低下; 另一方面, 算法改进策略虽多, 但需针对具体问题量身定制, 适用性不强。因此, 未来研究需更注重算法的自适应性、混合策略的创新性及种群结构的优化, 以平衡全局与局部搜索能力, 提高算法在不同应用场景下的鲁棒性和计算效率。

表 1 典型进化算法总结

Tab. 1 Summary of typical evolutionary algorithms

算法	适用场景	算法优点	算法缺点	算法改进方向
遗传算法	复杂、多变环境 离线规划	1. 全局搜索能力强 2. 易于并行处理 3. 无需梯度信息, 适用于复杂问题	1. 收敛速度慢 2. 早熟收敛(局部最优) 3. 参数选择敏感	1. 引入精英策略, 保留优秀个体 2. 改进交叉、变异操作, 提高搜索效率 3. 设计自适应参数调整策略
差分进化算法	连续、可导环境 实时或离线规划	1. 原理简单, 易于实现 2. 收敛速度快 3. 鲁棒性强, 适应性强	1. 对高维问题性能下降 2. 局部搜索能力较弱 3. 参数选择影响性能	1. 引入局部搜索策略, 增强局部优化能力 2. 结合其他算法处理高维问题 3. 设计参数自适应策略, 减少参数选择的影响

2.1.2 群智能算法

1) 粒子群算法

粒子群算法^[21] (Particle Swarm Optimization, PSO) 是一种基于群体智能的优化算法, 其灵感来源于对鸟群、鱼群等社会性动物群体行为的模拟。在粒子群算法中, 每个粒子具有位置、速度和适应度值 3 个属性。位置表示解在搜索空间中的位置, 速度决定粒子移动的方向和距离, 而适应度值则根据问题的目标函数计算得出, 用于评估粒子的优劣。算法通过迭代的方式更新粒子的位置和速度, 使粒子逐渐向最优解区域移动。图 4 为粒子群算法的流程图。

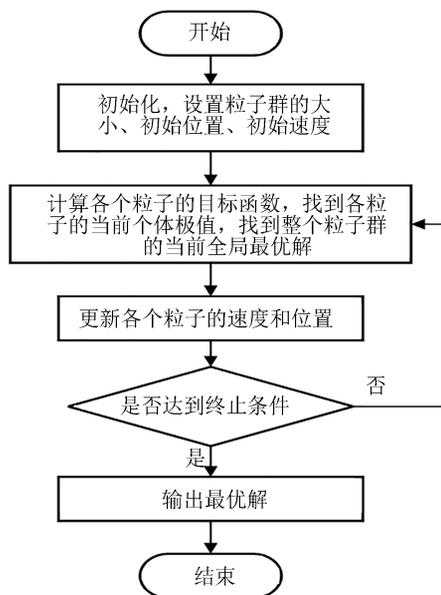


图 4 粒子群算法流程

Fig. 4 Flow chart of particle swarm optimization algorithm

粒子群算法具有许多优点, 如原理简单、易于实现、收敛速度快等。然而, 粒子群算法也存在一些缺点, 如缺乏速度的动态调节、容易陷入

局部最优等。为了克服这些缺点, 研究者们提出了许多改进的粒子群算法。

①策略更新改进。甯洋等^[22]为提升粒子群算法在航迹规划中的优化效果, 提出了一种融合压缩策略和变异策略的粒子群算法 (CMPSO), 通过自适应调节和增强种群多样性, 提高了跳出局部最优的能力, 该算法能获得更短路径, 耗时更少, 代价更低。

②混合算法策略。王飞等^[23]提出了一种改进的粒子群优化算法, 通过栅格法建立环境模型, 针对传统粒子群算法在求解此类问题时的局限性, 引入了 Singer 映射以改进粒子的初始分布, 同时采用线性调整加速因子和最大速度、新的粒子位置更新策略以及动态调整惯性权值等方法, 降低了计算代价, 提高了效率。

③参数自适应调整。黄晋等^[24]为解决无人机在复杂环境下进行航迹规划时搜索精度不足以及易陷入局部最优解的问题, 提出了引入分段自适应惯性权重和自适应指数学习因子, 有效平衡了算法的全局搜索和局部精细搜索能力。同时, 通过自适应调整系数来优化粒子的位置更新, 提升了所求路径的质量, 使其更为高效和精确。

目前, 研究者们已提出多种改进的粒子群算法, 包括策略更新改进、混合算法策略和参数自适应调整等。这些改进方法通过融合新策略、引入其他算法优点及自适应调整参数等方式, 提高了算法的全局搜索能力, 增强了种群多样性, 降低了计算代价。未来, 随着技术的不断进步, 粒子群算法有望在更多领域得到应用, 并且将更加注重算法的智能化、自适应化及与其他优化方法的融合, 以进一步提升其优化效果和应用范围。

2) 蚁群算法

蚁群算法^[25] (Ant Colony Optimization, ACO)

是一种模拟自然界中蚂蚁觅食行为的优化算法。该算法的核心思想在于模拟蚂蚁在寻找食物源过程中释放信息素的机制, 以及后续蚂蚁根据这些信息素浓度选择路径的行为。通过这一过程, 蚁群能够发现从巢穴到食物源的最短路径。算法流程如图5所示。

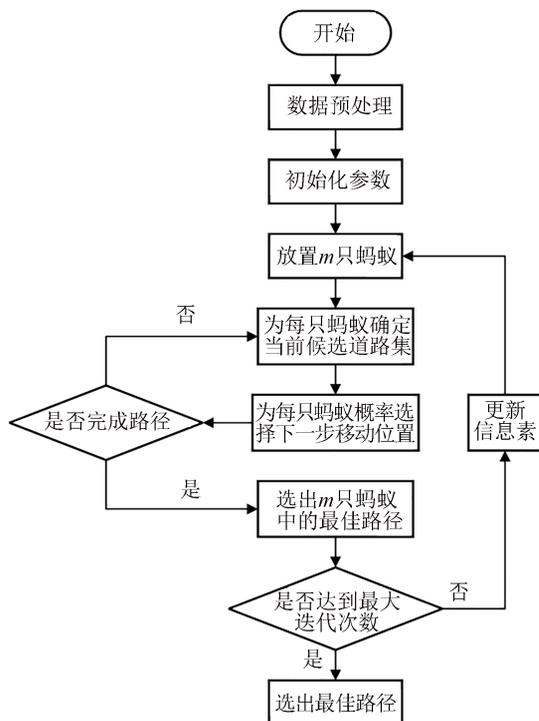


图5 蚁群算法流程

Fig. 5 Flow chart of ant colony optimization algorithm

蚁群算法具备全局搜索能力和适应性, 适用于解决复杂组合优化问题, 但存在收敛速度较慢、易遇滞留现象、过早收敛至局部最优等问题, 许多学者对其进行了改进。

①混合算法改进。田薇^[26]提出一种基于蚁群和D*算法的混合式无人机航迹规划算法。首先利用蚁群算法进行全局规划, 再结合D*算法对局部路径进行实时优化和重规划, 确保无人机在飞行过程中能够实时响应环境变化, 避免与障碍物发生碰撞, 并找到更加安全、高效的飞行路径。孙海文等^[27]提出了一种结合A*算法与蚁群算法的混合优化策略, 充分利用A*算法在稀疏或低密度环境区域进行高效航路规划的优势。在遇到复杂或高密度环境区域时, 引入蚁群算法来增强其局部规划能力和搜索效率, 实现了更全面、更灵活的航迹规划效果。

②动态自适应调整。邬琦等^[28]提出了一种动态自适应蚁群算法, 该算法通过实施动态自适应

航迹点选择策略, 并对信息素更新规则和挥发系数进行动态自适应调整, 显著提升了蚁群算法的求解效率。此外, 还基于战场已知的威胁源信息, 生成了Voronoi加权图, 并将其与所提的动态自适应蚁群算法相结合, 以在规划空间中求解出最优航迹。

③改进信息素更新机制。柳文林等^[29]提出了一种蚁群算法的改进策略, 通过动态调整信息素挥发系数和引入精英策略来优化信息素更新, 加速算法的收敛并提高解的质量, 增强了算法的有效性和实用性。

针对蚁群算法收敛速度慢、易陷入局部最优的问题, 研究者们不断探索改进策略。混合算法改进成为主流方向之一, 通过结合其他算法的优势, 来提升蚁群算法在全局与局部规划上的效率与灵活性。动态自适应调整策略则通过动态更新信息素规则和挥发系数, 增强了算法的适应性和求解速度。此外, 改进信息素更新机制也是关键, 通过精英策略和动态调整挥发系数, 加速算法收敛并提高解的质量。未来, 蚁群算法的发展将继续聚焦于提升收敛速度、增强全局搜索能力, 并探索更多与其他优化算法的融合应用, 以应对更复杂的实际问题。

3) 人工蜂群算法

人工蜂群 (Artificial Bee Colony, ABC) 算法^[30]是一种模拟自然界中蜜蜂群体觅食行为的优化算法。该算法基于蜜蜂群体的智能搜索和决策机制, 通过蜜蜂个体之间的信息交流和相互协作, 从而寻找问题的最优解, 实现了在复杂问题空间中的高效搜索和优化。人工蜂群的算法流程如图6所示。

人工蜂群算法简单灵活、全局搜索能力强、鲁棒性好, 适用于复杂航迹规划, 但随机性大、易陷入局部最优、收敛速度和精度受参数影响较大。针对上述缺点, 许多学者针对人工蜂群算法进行了改进。

①策略改进。朱金磊等^[31]通过全域采样随机初始化确保初始解集的完整性, 并在选择概率计算中引入开采次数因子以提高潜在较优解的选中概率。此外, 算法结合余弦函数的变化特点, 实现了全局最优个体引导下的自适应局部开发, 从而提升了局部开发的精度, 能够有效解决复杂灾害场景下的航迹规划问题。

②混合算法改进。王庆海等^[32]为应对无人机

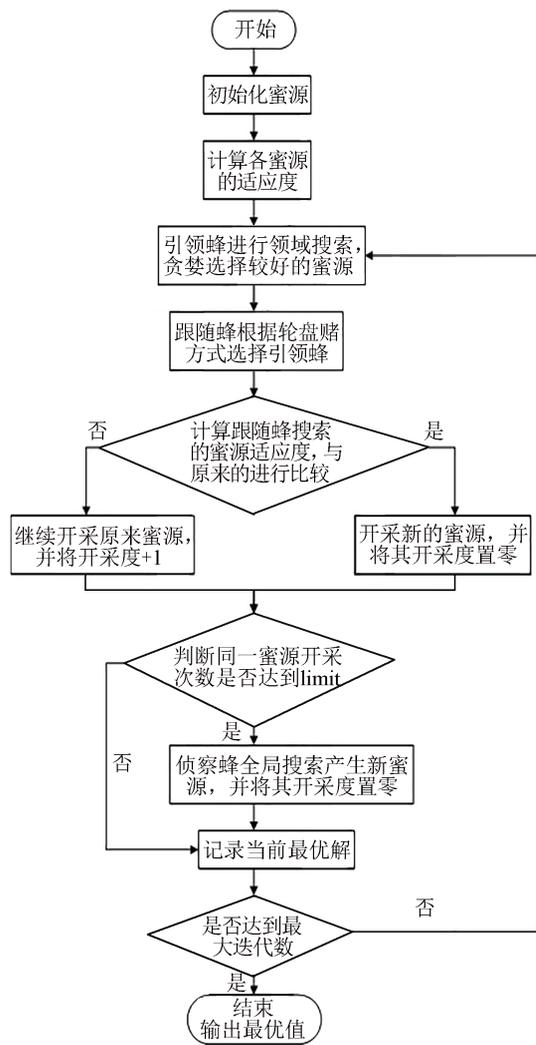


图 6 人工蜂群算法流程

Fig. 6 Flow chart of artificial bee colony algorithm

复杂任务, 提出了一种改进的人工蜂群算法, 通过优化初始化蜜源、融入模拟退火算法的 Me-

tropolis 准则和复合形法, 提升了无人机航迹规划的鲁棒性、收敛速度和精度。

人工蜂群智能算法的优势在于能够利用群体智慧实现全局搜索, 且具有较强的适应性和鲁棒性。但算法性能受参数设置、策略更新及信息素管理机制等的影响显著, 不当配置容易导致收敛速度慢和陷入局部最优等问题。未来研究需更注重算法的自适应性、混合策略的创新性以及参数优化的智能化, 以平衡全局与局部搜索能力, 提高算法在不同应用场景下的效率和精度。表 2 为典型群智能算法在无人机航迹规划应用中的对比与改进。

2.2 机器学习算法

机器学习算法旨在通过学习和分析历史飞行数据, 使无人机能够自主地进行航迹规划。在提高无人机航迹规划效率和精度的同时, 使无人机能够自适应地根据环境变化做出决策。机器学习算法还能够实现无人机航迹规划的自主学习和优化, 提高航迹规划算法的适应性和鲁棒性。

机器学习算法用于航迹规划主要有三大类, 一类是神经网络算法, 另一类是强化学习算法, 第三类是深度强化学习算法。

2.2.1 神经网络算法

神经网络 (Neural Networks, NNs) 作为一类智能计算模型, 其设计灵感来源于生物神经系统。这些网络由大量简单且功能相似的神经元相互连接构成, 其中的信息处理机制^[33]依赖于神经元之间结构化的信息传递方式。鉴于神经网络具有出色的拟合复杂非线性函数的能力, 神经网络

表 2 典型群智能算法总结

Tab. 2 Summary of typical swarm intelligence algorithms

算法	适用场景	算法优点	算法缺点	算法改进方向
粒子群算法	未知或半已知环境、实时规划	1. 通用性强, 不依赖问题信息 2. 原理简单, 易于实现 3. 收敛速度快 4. 适用于动态环境, 实时性强	1. 局部最优问题 2. 对复杂地形问题处理不佳 3. 参数选择困难	1. 引入速度动态调整机制 2. 结合地形信息优化搜索策略 3. 设计自适应参数调整策略
蚁群算法	静态或缓慢变化环境、离线规划	1. 正反馈机制, 易于寻找全局最优解 2. 分布式并行计算, 鲁棒性强 3. 适用于组合优化问题	1. 收敛速度慢 2. 局部最优问题 3. 对动态环境适应性差	1. 引入启发式信息, 提高收敛速度 2. 引入局部搜索策略, 避免局部最优 3. 结合动态环境信息, 增强实时性
人工蜂群算法	复杂多变环境、实时规划	1. 全局寻优能力强 2. 适用范围广, 适应性强 3. 适用于多模态函数优化	1. 收敛速度慢 2. 精度相对较低 3. 对参数敏感	1. 引入更高效的搜索策略 2. 结合地形信息提高搜索精度 3. 设计自适应参数调整策略

算法展现出优异的泛化能力、非线性映射能力和学习能力, 因此在无人机自主航迹规划领域具有显著的研究价值。

在无人机自主航迹规划方面, 基于神经网络的方法可划分为两大类。第一类方法直接利用神经网络生成飞行航线, 并结合其他优化算法对航线进行精细调整。李波等^[34]将无监督学习型神经网络应用于无人机全区域侦察航迹规划领域。将无人机的威胁信息、搜索需求和实时位置信息作为无监督学习型神经网络的输入特征, 旨在直接控制无人机的运动, 其输出直接对应无人机的引擎推力调整。另外结合遗传算法, 通过构建适应度函数来评估不同飞行航线的优劣。不仅提高了无人机侦察的智能化水平, 也显著提升了侦察效率和适应性。

另一类方法则是利用神经网络计算目标函数, 进而指导其他算法生成飞行航线。陈侠等^[35]首先运用动态步长策略来初始化飞行航迹, 然后借助神经网络模型构建了一个威胁能量函数。通过计算航点与潜在威胁边缘之间的距离来评估航点的威胁程度, 并输出相应的威胁能量值。随后, 利用梯度下降算法来构建飞行航线的运动方程, 确保航线能够迅速而有效地避开潜在的威胁。该算法提高了航线规划的效率和准确性, 显著缩短了计算时间。在实际应用中, 这类方法常与智能优化算法相结合, 以实现更加高效和准确的无人机航迹规划^[36-37]。

深度神经网络是包含有多个隐含层的神经网络, 通过学习飞行样本的内在规律进行航迹规划。Liu等^[38]提出了一种利用残差卷积神经网络进行无人机实时航迹规划的创新方法。该方法通过将无人机的当前位置、目标位置以及周围环境信息作为输入数据训练网络, 以识别和生成相应的环境状态和无人机航向。经过离线训练, 该网络模型展现出了优异的实时航迹规划性能。Wu等^[39]提出了一种端到端的神经网络架构, 将复杂的三维航迹规划任务转化为一个神经网络的分类问题。该研究引入了动作分解与合成的理念, 利用二维卷积神经网络来执行值迭代, 有效解决了三维空间中的航迹规划问题。

2.2.2 强化学习算法

近年来, 随着人工智能的兴起和计算机性能的提升, 强化学习的无人机航迹规划算法成为了研究热点。强化学习是一种不需要先验信息, 可

以直接和环境不断交互获取反馈信息来优化自身决策的人工智能优化算法, 其模型如图7所示。相对于传统算法, 强化学习算法的泛化性更好, 对动态变化的环境具有更强的适应能力, 且能更好地满足在线航迹规划问题的实时性要求。强化学习在训练过程中会自动获取样本数据, 在智能体与环境的不断交互过程中, 根据环境反馈的奖励值评估动作的好坏, 算法向获得奖励最大的方向进行学习, 最大奖励对应的策略即为最优策略^[40]。传统的强化学习算法主要包括动态规划法、蒙特卡洛法和时间差分法等^[41]。

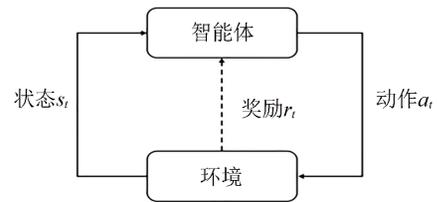


图7 强化学习模型

Fig. 7 Reinforcement learning model

强化学习不需要依赖大量预先标记样本数据, 在飞行样本相对较少的无人机自主航迹规划问题中更具优势。因此, 强化学习为无人机在有限数据条件下的航迹规划提供了一种有效的解决方案。

段建民等^[42]针对Q学习(Q-Learning)算法在航迹规划应用中存在收敛速度慢、效率不高的问题, 提出通过利用环境势能值来初始化Q值, 从而引导机器人在学习过程中快速收敛, 不仅提升了学习效率, 还增强了算法的适用性。

宋丽君等^[43]提出一种改进的Q-Learning算法, 通过引入概率突变性探索因子和设计深度学习因子, 再融合遗传算法, 实现了探索与利用之间的有效平衡。这种算法显著提高了学习效率, 有效避免了陷入局部最优解, 并且降低了在动态地图探索过程中的重复率, 为复杂环境中的学习提供了更为高效和稳定的解决方案。

2.2.3 深度强化学习算法

在航迹规划问题中, 随着环境复杂性和对智能体自主决策能力要求的提升, 传统的强化学习算法逐渐显露出其局限性。传统的Q-Learning、Sarsa等算法在处理复杂环境和大规模状态空间时, 往往难以有效地学习并找到最优策略。因此, 为了处理更加复杂的航迹规划任务, 研究者们将深度学习强大的函数逼近能力与强化学习的决策机制相结合, 发展出了深度强化学习算法。深度

强化学习算法通过深度学习网络来近似表示值函数或策略函数,能够处理大规模的状态空间和动作空间问题,使得智能体能够在复杂环境中自主进行航迹规划,并表现出更高的学习效率和更优的决策性能。

基于深度强化学习的无人机航迹规划方法按照策略更新方式分为基于价值函数的深度强化学习方法和基于策略的深度强化学习方法两类。

1) 基于价值函数的深度强化学习方法

基于价值函数的深度强化学习方法(Value-based Deep Reinforcement Learning)是将深度学习与强化学习中的价值函数估计相结合,通过深度神经网络来近似状态值函数或动作值函数,从而用于处理高维状态空间和连续动作空间问题。

深度 Q 网络^[44](Deep Q-Network, DQN)是一种结合深度学习与 Q-Learning 的强化学习算法,旨在解决传统 Q-Learning 在处理大规模或连续状态空间时遇到的维度灾难问题。深度 Q 网络通过使用深度神经网络(Deep Neural Networks, DNN)来近似表示 Q 值函数,使得算法能够在高维状态空间中有效地学习并找到最优策略,为复杂环境下的决策学习开辟了新的途径。Li 等^[45]提出了一种改进的深度 Q 网络航迹规划算法。通过引入多种优化策略,包括改进的 Q 值函数近似、优化动作选择机制以及动态调整学习率等,解决了传统深度 Q 网络算法在复杂环境航迹规划中面临的学习效率低、收敛速度慢以及难以适应动态障碍物等问题。Gu 等^[46]提出了一种名为 DMDQN(Dueling Munchausen Deep Q Network)的算法,该算法结合了 Dueling 网络和 Munchausen 强化学习技术,提高了深度 Q 网络算法在航迹规划任务中的性能和计算稳定性。通过引入 2 个独立的神经网络流分别估计状态值和优势值,并结合 Munchausen 正则化来减少过估计和不稳定性,实现了更高效和鲁棒的航迹规划。Zhang 等^[47]提出了一种基于长短期记忆(Long Short-Term Memory, LSTM)网络的无人机实时航迹规划算法,该算法融合了递归神经网络(Recursive Neural Network, RNN)和深度强化学习的记忆能力,利用长短期记忆网络作为 Q 值网络,实现了决策的记忆性,避免了单步决策的局限性。同时,通过设计分层奖惩函数,优化了无人机的航迹规划,在复杂环境中表现出更高的鲁棒性和准确性。

基于价值函数的深度强化学习方法在离散动

作空间的航迹规划任务中表现良好,然而在处理连续动作空间时面临挑战,并且无法直接应用于需要随机策略的情境^[48]。

2) 基于策略的深度强化学习方法

在处理航迹规划这类涉及连续动作空间和长期策略规划的问题时,面对复杂、动态的环境,基于价值函数的方法在动作选择上存在局限性。基于策略的深度强化学习方法(Policy-based Deep Reinforcement Learning)通过直接优化策略参数,能够学习到适应复杂环境、考虑长期收益的策略,从而更好地应对连续动作选择、动态环境变化等挑战,实现更稳定、更高效的航迹规划。

根据策略输出动作的性质,这些方法可分为随机性策略算法和确定性策略算法两类。随机性策略在给定状态下输出动作的概率分布,通过随机选择来平衡探索和利用;确定性策略直接输出确定的动作值,适用于需要精确控制的场合。这两类策略的代表算法包括 Actor-Critic、信赖域策略优化(Trust Region Policy Optimization, TR-PO)、近端策略优化(Proximal Policy Optimization, PPO)等随机性策略算法,以及深度确定性策略梯度(Deep Deterministic Policy Gradient, DDPG)、双延迟深度确定性策略梯度(Twin Delayed DDPG, TD3)等确定性策略算法。

①随机策略深度强化学习方法。随机策略深度强化学习算法是智能体在给定状态下根据概率分布随机选择动作,通过与环境交互和奖励反馈来优化策略性能的算法。

在随机性策略算法中,最具代表性的是 Actor-Critic 方法及其各种变种。Actor-Critic 方法^[49]是一种深度强化学习算法,融合了策略梯度与价值函数的思想。在这种方法中,Actor 网络专注于学习策略,通过生成动作的概率分布来指导智能体与环境进行互动;Critic 网络用于学习价值函数,评估 Actor 生成动作的质量,为 Actor 提供反馈。Actor 和 Critic 之间通过共享环境反馈的奖励信号进行交互,Actor 根据 Critic 的评估调整其策略,Critic 则根据 Actor 的动作和环境的奖励来更新其价值函数估计。Actor-Critic 算法框架如图 8 所示。

信赖域策略优化算法^[50]通过引入信赖域约束限制策略更新的幅度,确保策略更新稳定且奖励增加。在无人机航迹规划问题中,信赖域策略优化算法解决了因策略更新步长过大导致的性能波

动问题, 提升了训练的稳定性和收敛速度。然而, 信赖域策略优化算法在处理高维状态和动作空间时仍旧面临计算复杂度高、收敛速度受限等问题。

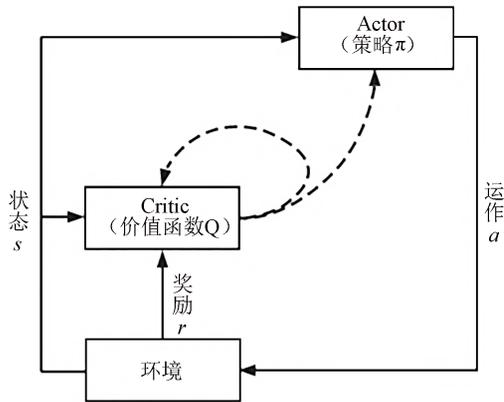


图 8 Actor-Critic 算法框架

Fig. 8 Framework of Actor-Critic algorithm

近端策略优化算法^[51]通过截断策略更新的比率来限制新策略和旧策略的差异, 从而在保障稳定性的同时实现高效的策略更新。在无人机航迹规划问题中, 该算法解决了信赖域策略优化算法因KL散度约束导致的计算复杂度高的问题, 使得训练更加高效。然而, 尽管该算法提高了训练效率和稳定性, 但在处理高维、复杂航迹规划问题时, 仍可能面临样本效率不高和陷入局部最优解的问题。Xiao等^[52]提出了一种基于Beta策略的改进分布式采样近端策略优化算法, 该算法通过引入Beta策略来对近端策略优化算法的更新过程进行改进。Beta策略能够根据历史数据动态调整采样分布, 使得算法能够更快速地适应环境变化, 并找到更优的航迹规划方案。此外, 分布式采样设计使得算法能够并行处理大量数据, 显著提高了训练速度, 降低了计算资源的消耗, 进一步提升了全局航迹规划的性能和效率。Zhao等^[53]提出了一种基于深度强化学习的双无人机协同侦察任务规划方法, 该方法通过整体建模考虑任务特定属性和参数, 将侦察任务描述为马尔可夫决策过程, 并设计了4个带权重的奖励函数以增强任务理解, 然后使用近端策略优化算法, 结合离线和在线规划进行仿真。结果表明, 本文所提方法可适应未知连续高维环境变化, 规划出满足要求的路径。

为了解决近端策略优化算法因大量采样而导致学习效率下降的问题, Haarnoja等^[54]提出了SAC (Soft Actor-Critic) 算法。SAC是一种基于

最大熵框架的异策略随机策略梯度算法, 旨在通过改进策略学习的方式来提高学习效率。通过结合策略梯度学习和价值函数估计, 引入熵最大化来鼓励智能体探索更多的动作, 从而增强算法的鲁棒性和适应性, 能够在高维状态空间和动作空间中学习有效的策略。Zhang等^[55]提出了一种在动态室内环境中的SAC-LSTM航迹规划方法。该方法结合了序列到序列架构和长短时记忆网络, 实现了对动态室内环境中高效、准确、适应性强、实时性好的航迹规划。Zhou等^[56]提出了一种改进SAC算法, 该方法通过在Actor网络中引入自注意力机制来处理三维环境空间中的高维复杂状态信息, 并将人工势场法与SAC算法中的奖励函数相结合, 解决了在线轨迹规划中的奖励稀疏问题。

②确定性策略深度强化学习方法。确定性策略深度强化学习算法是一种结合了深度学习和确定性策略梯度的强化学习算法, 在处理连续动作空间的问题时表现出色。这种算法的核心思想是通过深度神经网络来逼近策略和值函数, 同时采用确定性策略来直接输出动作, 从而在高维连续动作空间中提供更精确和高效的动作选择。此外, 确定性策略深度强化学习算法还引入了目标网络和经验回放机制, 以增强学习过程的稳定性和数据利用率。

深度确定性策略梯度算法^[57]是确定性策略深度强化学习领域的代表性算法。该算法基于确定性策略的思想, 构建了4个神经网络来分别学习策略函数和价值函数。为了打破数据之间的相关性, 深度确定性策略梯度算法引入独立的目标网络结构以及经验回放机制, 以提高学习过程的稳定性和效率。在无人机航迹规划问题中, Wu等^[58]提出了一种多批判延迟深度确定性策略梯度算法, 通过引入多批判网络和延迟学习来减少高估问题, 同时添加噪声以增强鲁棒性, 该算法收敛更快、效果更优、稳定性更强, 能使无人机更有效地学习应对复杂环境。深度确定性策略梯度算法能够在连续动作空间中学习最优的确定性策略, 避免了传统方法中对于离散动作空间的依赖, 提高了航迹规划的灵活性和准确性。

双延迟深度确定性策略梯度算法^[59]是深度确定性策略梯度算法的改进版本, 通过引入双重Critic网络 (即“Twin”网络) 来解决深度确定性策略梯度算法中的价值函数过估计问题, 同时采用延迟更新的策略来减少策略网络更新的频率,

从而提高了算法的稳定性和性能。Luo 等^[60]提出了基于优先经验回放的平均双延迟深度确定性策略梯度算法用于无人机的航迹规划。该算法结合深度强化学习中的双延迟深度确定性策略梯度算法和优先经验回放机制,通过引入平均策略的概念,对多个时间步的策略进行平均,提高了策略的稳定性和鲁棒性。

表 3 为典型机器学习算法在无人机航迹规划应用中的对比与改进方向。

3 航迹规划智能算法对比分析

智能优化算法与机器学习算法均可应用于无人机的航迹规划问题,智能优化算法擅长进行全局搜索,而学习能力较弱;机器学习算法在学习能力上表现突出,但全局搜索能力有限。通过分析智能算法的研究进展和比较这 2 种算法的实际应用场景,可以得出以下结论。

1) 智能优化算法以其结构的简洁性和较强的

表 3 机器学习算法总结

Tab. 3 Summary of machine learning algorithms

算法类型	算法	适用场景	算法优点	算法缺点	算法改进方向
神经网络算法	深度神经网络	复杂环境、离线或实时规划	1. 非线性映射能力强 2. 可通过学习优化路径	1. 需要大量训练数据 2. 对模型参数敏感	1. 设计更有效的训练策略 2. 结合其他算法减少数据需求 3. 引入自适应参数调整机制
强化学习算法	Q 学习	静态或简单动态环境	1. 不需要完整的环境模型 2. 离线策略学习,可以学习更广泛的策略	1. 数据效率低下,需要大量探索 2. 对大规模或连续状态空间不够高效	1. 结合深度学习(如深度 Q 网络)提高数据的利用率 2. 使用优先经验回放等技术提高学习效率
	Sarsa	已知环境、实时决策	1. 在线策略学习,实时性强 2. 适用于动态环境,基于实际经验进行更新	1. 可能会陷入局部最优解 2. 对大规模或连续状态空间不够高效	1. 使用更先进的探索策略以平衡探索和利用 2. 结合模型预测以提高环境适应能力
基于随机性策略的深度强化学习方法	深度 Q 网络	动态环境、实时规划	1. 可处理连续状态空间 2. 强大的自学习能力	1. 样本效率低 2. 收敛速度慢	1. 引入优先经验回放 2. 结合其他算法提高收敛速度 3. 设计更有效的探索策略
	近端策略优化算法	实时规划	1. 适用于连续动作空间 2. 能够直接优化策略	1. 稳定性差 2. 对超参数敏感	1. 引入正则化技术提高稳定性 2. 设计自适应参数调整策略 3. 结合其他算法改善性能
	SAC 算法	动态环境、实时规划	1. 探索能力和鲁棒性强 2. 适用于连续动作空间 3. 速度快,易收敛	1. 对超参数敏感 2. 在高维空间中计算复杂	1. 设计自适应超参数调整策略 2. 引入近似方法减少计算量 3. 结合其他算法提高性能
基于确定性策略的深度强化学习方法	深度确定性策略梯度	实时规划	1. 高效利用数据 2. 适用于连续动作空间	1. 对模型误差敏感 2. 难以处理部分可观察环境	1. 引入模型不确定性估计 2. 结合其他算法处理部分可观察性 3. 设计鲁棒性更强的探索策略
	双延迟深度确定性策略梯度算法	实时规划	1. 双重网络结构,减少 Q 值估计偏差 2. 目标策略平滑正则化,提高价值评估准确性 3. 延迟更新策略,提高稳定性	1. 对模型参数和结构敏感 2. 难以处理极端环境	1. 设计自适应参数调整机制 2. 结合其他算法处理极端环境 3. 引入不确定性估计以应对模型误差

可实现性, 特别适合在资源受限的地面控制站中实施。这些算法遵循启发式原则, 易于实现, 并且具有良好的扩展性, 能够适应多变的飞行环境和任务需求。但是在面对大规模解空间时可能会遇到局部最优解的困境, 且收敛速度可能较慢。这导致算法在搜索全面性和收敛效率之间难以平衡, 有时可能无法保证航迹规划的质量, 或导致规划过程耗时较长, 难以应用于需要实时响应的复杂环境中。

2) 机器学习算法, 特别是强化学习方法, 以其强大的非线性问题处理能力, 在无人机航迹规划中展现出巨大潜力。算法通过分析历史航线数据, 学习环境特征与决策之间的内在联系, 形成高效的航迹规划策略。机器学习算法在环境变化时能够快速适应, 无需从头开始搜索, 更适合实时航迹规划。但受限于航线数据的可用性, 且在训练初期可能会进行大量的探索, 降低了规划过程的效率。如果训练数据不足或模型超参数未得到恰当调整, 也可能影响算法性能和最终规划出的航线质量。

4 结论

随着无人机任务多样化和飞行环境复杂化, 航迹规划算法需进一步优化以适应复杂多变的战场环境, 从而增加航迹规划效率和任务执行的成功率。因此, 关于无人机航迹规划智能算法的研究仍然有很大发展空间。

1) 环境模型复杂性。随着无人机在复杂战场环境中的应用日益广泛, 实时三维航迹规划显得尤为重要。二维预先规划已无法满足无人机的飞行安全和任务需求, 而传统方法在处理复杂问题时存在局限性。因此, 未来研究需聚焦于构建贴近真实战场的实时三维环境模型, 并着力提升智能算法的实时性和路径规划效率。同时, 为应对动态环境中的不确定性和未知因素, 航迹规划算法还需具备快速数据处理、高效计算策略以及强大的动态适应性和鲁棒性, 以确保无人机能实时响应环境变化并稳定地重新规划航迹。

2) 目标约束多样性。在复杂战场环境中, 无人机航迹规划面临多重挑战。设计智能算法的目标函数是关键, 现有算法考虑要素不足, 如仅仅只考虑了长度代价、威胁代价、平滑代价目标函数, 难以确保航迹最优性和可靠性。未来目标函

数需综合考虑任务需求、飞行成本及其他各类约束, 如无人机动力学和终端航向角约束等。同时, 航迹规划还需满足飞行高度、速度、能耗等多重约束, 以优化飞行性能和任务效果。未来研究应着重于构建全面合理的目标函数, 平衡不同约束, 避免局部最优解, 从而求得最优航迹规划方案。

3) 算法融合与智能化提升。随着人工智能技术的飞速发展, 未来航迹规划算法将更加注重多种智能算法的融合与创新。当前航迹规划算法虽然在一定程度上能够实现自动化和智能化, 但在处理极端复杂环境、高动态变化场景时仍显不足。因此, 未来的发展方向应聚焦于如何将深度学习、强化学习、遗传算法等先进人工智能技术更有效地融合到航迹规划中, 以提升算法的智能化水平和高度决策能力。此外, 随着算法复杂度的增加, 如何保证算法的实时性和稳定性也成为一大挑战, 需要研究者在算法设计和优化上付出更多努力。

4) 人机交互与自主决策协同。随着无人机航迹规划技术的不断进步, 未来航迹规划算法将更加注重人机交互与自主决策的协同。当前, 虽然航迹规划算法已经取得了一定的自动化水平, 但在实际应用中, 仍需要人工干预和决策, 特别是在复杂多变的战场环境中。因此, 如何实现人机交互与自主决策的有效协同, 成为未来航迹规划算法发展的重要方向。此外, 还需要解决人机交互与自主决策之间的冲突和矛盾, 确保两者在协同过程中能够相互补充、相互促进, 而不是相互干扰或制约。

综上所述, 无人机航迹规划算法的发展与挑战紧密相连, 需要考虑多个方面因素进行深入研究。展望未来, 随着人工智能、大数据和云计算等技术的不断发展, 新的算法和模型将被引入到无人机航迹规划领域, 为无人机提供更智能、更高效的航迹规划解决方案, 推动其在更多领域的应用和发展。

参考文献

- [1] 武锦龙, 吴虹麒, 李浩, 等. 基于改进DeepLabV3+的荞麦苗期无人机遥感图像分割研究[J/OL]. 农业机械学报, 2024: 1-12. <https://link.cnki.net/urlid/11.1964.S.20240328.0844.002>.
WU J L, WU H Q, LI H, et al. Segmentation of buckwheat by UAV based on improved lightweight DeepLabV3+ at seedling stage[J/OL]. Tra-

- nsactions of the Chinese Society for Agricultural Machinery, 2024: 1-12. <https://link.cnki.net/urlid/11.1964.S.20240328.0844.002>.
- [2] 陈炳耀, 信誉, 路方婷, 等. 长江江豚监测现状及展望[J]. 中国环境监测, 2023, 39(2): 1-10.
CHEN B Y, XIN Y, LU F T, et al. Monitoring status and prospects of Yangtze finless porpoise [J]. Environmental Monitoring in China, 2023, 39 (2): 1-10.
- [3] 白文超, 班明飞, 宋梦, 等. 电动汽车-无人机联合救援系统协调调度模型[J]. 上海交通大学学报, 2024, 58(9): 1443-1453.
BAI W C, BAN M F, SONG M, et al. Coordinated scheduling model for electric vehicle-unmanned aircraft joint rescue system[J]. Journal of Shanghai Jiao Tong University, 2024, 58 (9) : 1443-1453.
- [4] 王飞, 杨清平. 面向多无人机物流配送的双层任务规划方法[J/OL]. 北京航空航天大学学报, 2024: 1-14. <https://doi.org/10.13700/j.bh.1001-5965.2023.0719>.
WANG F, YANG Q P. Two-layer task planning method for multi-UAV logistics distribution [J/OL]. Journal of Beijing University of Aeronautics and Astronautics, 2024: 1 - 14. <https://doi.org/10.13700/j.bh.1001-5965.2023.0719>.
- [5] 杨孝鹏, 李伟春, 孙锐. 基于国家安全情报的海岸警卫队态势感知研究[J]. 飞航导弹, 2020(6): 87-91.
- [6] 唐恒, 孙伟, 吕磊, 等. 融合动态奖励策略的无人机编队路径规划方法[J]. 系统工程与电子技术, 2024, 46(10): 3506-3518.
TANG H, SUN W, LU L, et al. UAV formation path planning approach incorporating dynamic reward strategy[J]. Systems Engineering and Electronics, 2024, 46(10): 3506-3518.
- [7] 王伟伦, 尤明, 孙磊, 等. 未知环境下无人机集群智能协同探索路径规划[J]. 工程科学学报, 2024, 46(7): 1197-1206.
WANG W L, YOU M, SUN L, et al. Intelligent collaborative exploration path planning for UAV swarm in an unknown environment[J]. Chinese Journal of Engineering, 2024, 46(7): 1197-1206.
- [8] 李明阳, 许可儿, 宋志强, 等. 多智能体强化学习算法研究综述[J]. 计算机科学与探索, 2024, 18(8): 1979-1997.
LI M Y, XU K E, SONG Z Q, et al. Review of research on multi-agent reinforcement learning algorithms[J]. Journal of Frontiers of Computer Science and Technology, 2024, 18(8): 1979-1997.
- [9] 郑锴, 郑献民, 殷少锋, 等. 基于改进A*算法的无人机任务分配和航迹规划优化方法[J]. 电光与控制, 2022, 29(10): 7-11, 101.
ZHENG K, ZHENG X M, YIN S F, et al. UAV mission assignment and path planning optimization based on improved A* algorithm[J]. Electronics Optics & Control, 2022, 29 (10) : 7-11, 101.
- [10] 汪小帅, 朱其新, 朱永红, 改进D*算法下的无人机三维路径规划[J]. 西安工程大学学报, 2023, 37(3): 83-91.
WANG X S, ZHU Q X, ZHU Y H. Three-dimensional path planning for unmanned aerial vehicle based on an improved D* algorithm[J]. Journal of Xi'an Polytechnic University, 2023, 37(3) : 83-91.
- [11] 郝峰, 张栋, 唐硕, 等. 基于改进RRT算法的巡飞弹快速航迹规划方法[J]. 飞行力学, 2019, 37(3): 58-63.
HAO F, ZHANG D, TANG S, et al. A rapid route planning method of loitering munitions based on the improved RRT algorithm[J]. Flight Mechanics, 2019, 37(3): 58-63.
- [12] 常洪江. 遗传算法综述[J]. 电脑学习, 2010(3): 115-116.
- [13] 王豪, 赵学军, 袁修久. 基于改进自适应遗传算法的机器人路径规划[J]. 电光与控制, 2022, 29(5): 72-76.
WANG H, ZHAO X J, YUAN X J. Robot path planning based on improved adaptive genetic algorithm[J]. Electronics Optics & Control, 2022, 29 (5) : 72-76.
- [14] 程泽新, 李东生, 高杨. 一种改进遗传算法的无人机航迹规划[J]. 计算机仿真, 2019, 36(12): 31-35.
CHENG Z X, LI D S, GAO Y. GASA drone path planning to improve mutation strategy[J]. Computer Simulation, 2019, 36(12): 31-35.
- [15] 李文广, 胡永江, 庞强伟, 等. 基于改进遗传算法的多无人机协同侦察航迹规划[J]. 中国惯性技术学报, 2020, 28(2): 248-255.
LI W G, HU Y J, PANG Q W, et al. Track planning of multi-UAV cooperative reconnaissance based on improved genetic algorithm[J]. Journal of Chinese Inertial Technology, 2020, 28 (2) : 248-255.
- [16] 贾广芝. 基于遗传算法和稀疏A*算法的无人机三维航迹规划研究[D]. 南京: 南京邮电大学, 2017:

- 25-27.
- JIA G Z. Research on three-dimensional path planning of UAV based on genetic algorithm and sparse A* algorithm[D]. Nanjing: Nanjing University of Posts and Telecommunications, 2017: 25-27.
- [17] 张荣, 王彬. 基于改进差分进化算法的无人机任务规划研究[J]. 电子制作, 2022, 30(19): 49-54.
- [18] 严炜, 龙长江, 李善军. 基于差分量子退火算法的农用无人机路径规划方法[J]. 华中农业大学学报, 2020, 39(1): 180-186.
- YAN W, LONG C J, LI S J. A path planning method for agricultural UAV based on DEQA algorithm[J]. Journal of Huazhong Agricultural University, 2020, 39(1): 180-186.
- [19] 张宗豪, 徐斌, 胡铮. 应用改进差分进化算法的三维路径规划[J]. 电光与控制, 2022, 29(6): 6-10, 15.
- ZHANG Z H, XU B, HU Z. Application of modified differential evolution algorithm in 3D path planning[J]. Electronics Optics & Control, 2022, 29(6): 6-10, 15.
- [20] CHAI X Z, ZHENG Z S, XIAO J M, et al. Multi-strategy fusion differential evolution algorithm for UAV path planning in complex environment [J]. Aerospace Science and Technology, 2022, 121: 107287.
- [21] 朱润泽, 赵静, 蒋国平, 等. 基于改进粒子群算法的无人机三维路径规划[J/OL]. 南京邮电大学学报(自然科学版), 2024: 1-10. <https://link.cnki.net/urlid/32.1772.tn.20240621.1643.016>.
- ZHU R Z, ZHAO J, JIANG G P, et al. UAV 3D path planning based on improved particle swarm optimization algorithm[J/OL]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2024: 1-10. <https://link.cnki.net/urlid/32.1772.tn.20240621.1643.016>.
- [22] 甯洋, 郑波, 龙足腾, 等. 基于CMPSO算法的无人机复杂三维路径规划[J]. 电光与控制, 2024, 31(4): 35-42.
- NING Y, ZHENG B, LONG Z T, et al. Complex 3D path planning for UAVs based on CMPSO algorithm[J]. Electronics Optics & Control, 2024, 31(4): 35-42.
- [23] 王飞, 杨清平. 基于改进粒子群算法的城市物流无人机路径规划[J]. 科学技术与工程, 2023, 23(30): 13187-13194.
- WANG F, YANG Q P. Route planning of urban logistics unmanned aerial vehicle based on improved particle swarm optimization algorithm[J]. Science Technology and Engineering, 2023, 23(30): 13187-13194.
- [24] 黄晋, 李云飞, 王圣淳, 等. 基于改进PSO算法的无人机城域三维路径规划[J]. 电光与控制, 2024, 31(2): 41-45, 76.
- HUANG J, LI Y F, WANG S C, et al. 3D path planning of UAVs in urban environments based on improved PSO algorithm[J]. Electronics Optics & Control, 2024, 31(2): 41-45, 76.
- [25] 王辉, 胡晓阳. 基于蚁群算法的无人机航迹规划研究[J]. 科技资讯, 2020, 18(10): 29-30.
- [26] 田薇. 基于改进蚁群-Dstar混合算法的无人机航迹规划算法研究[D]. 哈尔滨: 哈尔滨工业大学, 2018: 29-30.
- TIAN W. Research on UAV path planning algorithm based on improved ant colony-Dstar hybrid algorithm[D]. Harbin: Harbin Institute of Technology, 2018: 29-30.
- [27] 孙海文, 肖玉杰, 王生玉. 基于改进型A*-蚁群混合算法的USV航迹规划[J]. 舰船科学技术, 2021, 43(8): 83-87.
- SUN H W, XIAO Y J, WANG S Y. USV route planning based on improved A*-ant hybrid algorithm[J]. Ship Science and Technology, 2021, 43(8): 83-87.
- [28] 邬琦, 潘广贞, 杨江涛. 基于Voronoi图和动态自适应蚁群算法的UAV航迹规划[J]. 计算机测量与控制, 2014, 22(9): 3037-3040.
- WU Q, PAN G Z, YANG J T. UAV planning based of UAV based on Voronoi diagram and dynamic adaptive ant colony algorithm[J]. Computer Measurement & Control, 2014, 22(9): 3037-3040.
- [29] 柳文林, 潘子双, 赵红超. 基于改进蚁群算法的无人机2维航路规划[J]. 兵工自动化, 2022, 41(11): 68-72, 88.
- LIU W L, PAN Z S, ZHAO H C. 2D path planning for UAV based on improved ant colony algorithm[J]. Ordnance Industry Automation, 2022, 41(11): 68-72, 88.
- [30] 高晨阳, 鱼小军, 闫妍. 蜂群算法研究综述[J]. 信息与电脑, 2021(22): 63-65.
- GAO C Y, YU X J, YAN Y. A review of bee colony optimization[J]. China Computer & Communication, 2021(22): 63-65.
- [31] 朱金磊, 袁晓兵, 裴俊. 基于改进人工蜂群算法的灾害场景下路径规划[J]. 中国科学院大学学报, 2023, 40(3): 397-405.
- ZHU J L, YUAN X B, PEI J. Path planning in

- disaster scenarios based on improved artificial bee colony algorithm[J]. *Journal of University of Chinese Academy of Sciences*, 2023, 40(3): 397-405.
- [32] 王庆海, 刘广瑞, 郭珂甫. 基于改进人工蜂群算法的无人机航迹规划研究[J]. *机床与液压*, 2017, 45(21): 68-72.
- WANG Q H, LIU G R, GUO K F. Research on path planning of UAV based on improved ABC algorithm[J]. *Machine Tool and Hydraulics*, 2017, 45(21): 68-72.
- [33] 王东. 机器学习导论[M]. 北京: 清华大学出版社, 2021: 171-172.
- WANG D. Introduction to machine learning[M]. Beijing: Tsinghua University Press, 2021: 171-172.
- [34] 李波, 杨志鹏, 贾卓然, 等. 一种无监督学习型神经网络的无人机全区域侦察路径规划[J]. *西北工业大学学报*, 2021, 39(1): 77-84.
- LI B, YANG Z P, JIA Z R, et al. An unsupervised learning neural network for planning UAV full-area reconnaissance path[J]. *Journal of Northwestern Polytechnical University*, 2021, 39(1): 77-84.
- [35] 陈侠, 艾宇迪. 应用改进神经网络的无人机三维航迹规划[J]. *电光与控制*, 2018, 25(9): 7-11.
- CHEN X, AI Y D. Application of improved neural network in 3D path planning of UAVs[J]. *Electronics Optics & Control*, 2018, 25(9): 7-11.
- [36] 陈秋莲, 郑以君, 蒋环宇, 等. 基于神经网络改进粒子群算法的动态路径规划[J]. *华中科技大学学报(自然科学版)*, 2021, 49(2): 51-55.
- CHEN Q L, ZHENG Y J, JIANG H Y, et al. Improved particle swarm optimization algorithm based on neural network for dynamic path planning[J]. *Journal of Huazhong University of Science and Technology (Natural Science Edition)*, 2021, 49(2): 51-55.
- [37] DUAN H B, HUANG L Z. Imperialist competitive algorithm optimized artificial neural networks forUCAV global path planning[J]. *Neurocomputing*, 2014, 125: 166-171.
- [38] LIU Y, ZHENG Z, QIN F Y, et al. A residual convolutional neural network based Approach for real-time path planning[J]. *Knowledge-Based Systems*, 2022, 242: 108400.
- [39] WU K Y, ABOLFAZLI ESFAHANI M, YUAN S H, et al. TDDP-Net: Achieving three-dimensional path planning via a deep neural network architecture[J]. *Neurocomputing*, 2019, 357: 151-162.
- [40] SUTTON R S, BARTO A G. Reinforcement learning: An introduction[M]. Massachusetts: The MIT Press, 2018.
- [41] SUTTON R S. Learning to predict by the methods of temporal differences[J]. *Machine Learning*, 1988, 3(1): 9-44.
- [42] 段建民, 陈强龙. 利用先验知识的Q-Learning路径规划算法研究[J]. *电光与控制*, 2019, 26(9): 29-33.
- DUAN J M, CHEN Q L. Prior knowledge based Q-Learning path planning algorithm[J]. *Electronics Optics & Control*, 2019, 26(9): 29-33.
- [43] 宋丽君, 周紫瑜, 李云龙, 等. 改进Q-Learning的路径规划算法研究[J]. *小型微型计算机系统*, 2024, 45(4): 823-829.
- SONG L J, ZHOU Z Y, LI Y L, et al. Research on path planning algorithm based on improved Q-Learning algorithm[J]. *Journal of Chinese Computer Systems*, 2024, 45(4): 823-829.
- [44] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-level control through deep reinforcement learning[J]. *Nature*, 2015, 518: 529-533.
- [45] LI J X, CHEN Y T, ZHAO X N, et al. An improved DQN path planning algorithm[J]. *The Journal of Supercomputing*, 2022, 78(1): 616-639.
- [46] GU Y W, ZHU Z T, LYU J D, et al. DM-DQN: Dueling Munchausen deep Q network for robot path planning[J]. *Complex & Intelligent Systems*, 2023, 9(4): 4287-4300.
- [47] ZHANG J D, GUO Y K, ZHENG L H, et al. Real-time UAV path planning based on LSTM network[J]. *Journal of Systems Engineering and Electronics*, 2024, 35(2): 374-385.
- [48] 闫皎洁, 张镋石, 胡希平. 基于强化学习的路径规划技术综述[J]. *计算机工程*, 2021, 47(10): 16-25.
- YAN J J, ZHANG Q S, HU X P. Review of path planning techniques based on reinforcement learning[J]. *Computer Engineering*, 2021, 47(10): 16-25.
- [49] GRONDMAN I, BUSONI L, LOPES GAD, et al. A survey of actor-critic reinforcement learning: standard and natural policy gradients[J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 2012, 42(6): 1291-1307.
- [50] SCHULMAN J, LEVINE S, ABBEEL P, et al.

- Trust region policy optimization [C]//32nd International Conference on Machine Learning, 2015: 1889-1897.
- [51] SCHULMAN J, WOLSKI F, DHARIWAL P, et al. Proximal policy optimization algorithms[EB/OL]. (2017-08-28) [2024-08-17]. <https://doi.org/10.48550/arXiv.1707.06347>.
- [52] XIAO Q H, JIANG L, WANG M M, et al. An improved distributed sampling PPO algorithm based on beta policy for continuous global path planning scheme[J]. *Sensors (Basel)*, 2023, 23(13): 6101.
- [53] ZHAO X R, YANG R N, ZHANG Y, et al. Deep reinforcement learning for intelligent dual-UAV reconnaissance mission planning[J]. *Electronics*, 2022, 11(13): 2031.
- [54] HAARNOJA T, ZHOU A, ABBEEL P, et al. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor[EB/OL]. (2018-08-08) [2024-08-17]. <https://doi.org/10.48550/arXiv.1801.01290>.
- [55] ZHANG Y C, CHEN P Z. Path planning of a mobile robot for a dynamic indoor environment based on an SAC-LSTM algorithm[J]. *Sensors (Basel)*, 2023, 23(24): 9802.
- [56] ZHOU Y X, SHU J S, HAO H, et al. UAV 3D online track planning based on improved SAC algorithm[J]. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 2023, 46: 7-11.
- [57] MA Y, ZHU W B, BENTON M G, et al. Continuous control of a polymerization system with deep reinforcement learning[J]. *Journal of Process Control*, 2019, 75: 40-47.
- [58] WU R J, GU F Q, LIU H L, et al. UAV path planning based on multicritic-delayed deep deterministic policy gradient[J]. *Wireless Communications and Mobile Computing*, 2022: 9017079.
- [59] FUJIMOTO S, VAN HOOFF H, MEGER D. Addressing function approximation error in actor-critic methods[EB/OL]. (2018-10-22) [2024-08-20]. <https://doi.org/10.48550/arXiv.1802.09477>.
- [60] LUO X Q, WANG Q Y, GONG H F, et al. UAV path planning based on the average TD3 algorithm with prioritized experience replay[J]. *IEEE Access*, 2024, 12: 38017-38029.

(编辑: 刘建民)