

# 融合强化学习与群智能算法的移动机器人自主导航策略优化

王娜<sup>1</sup> 金小婷<sup>2</sup>

1. 昆明冶金高等专科学校; 2. 辽宁开放大学(辽宁装备制造职业技术学院)

**摘要:** 移动机器人在复杂动态环境中的自主导航仍面临路径规划不稳定、全局最优性不足与动态避障能力有限等问题。为此, 本文构建融合强化学习与群智能算法的协同优化框架, 通过粒子群算法实现策略网络参数初始化与搜索引导, 通过蚁群算法提供全局参考路径, 并以强化学习完成局部实时决策与动态避障。二者通过奖励设计与搜索空间反馈实现双向耦合。实验结果表明, 该方法在路径长度、规划效率、避障成功率与策略稳定性方面均显著优于单一算法, 尤其在动态障碍环境中展现出更强的鲁棒性与适应能力, 为移动机器人自主导航提供了一种高效可扩展的优化策略。

**关键词:** 移动机器人; 自主导航; 强化学习; 群智能算法; 路径规划

移动机器人自主导航能力是智能制造、仓储物流与公共服务等场景的重要基础, 其核心在于在复杂动态环境中完成路径规划、避障与实时决策。传统算法如A\*与Dijkstra在静态已知环境中表现较好, 但面对多障碍、动态变化与不确定性条件时, 容易出现规划僵化、路径质量下降等问题。近年来, 强化学习因具备自适应和在线学习能力, 被广泛用于导航策略优化, 但其训练过程依赖大量交互样本, 收敛速度慢, 且在稀疏奖励环境中易陷入局部最优。

群智能算法具有良好的全局搜索能力, 能够为复杂路径规划提供整体启发式优化, 但其迭代更新机制难以满足实时导航需求, 在高维连续空间中搜索效率也会明显下降。二者在优势与不足上形成互补: 强化学习擅长局部动态决策, 群智能算法擅长全局结构优化。因此, 将两类方法进行深度融合, 有望实现全局最优性与局部适应性的兼具, 从而提升机器人在复杂环境中的自主导航性能。基于此, 本文构建融合强化学习与群智能算法的协同优化框架, 通过全局参考路径与策略初始化引导强化学习快速收敛, 并利用强化学习的在线评价信息动态优化全局搜索空间, 以实现更高效、更稳定的自主导航策略。

## 1 方法

### 1.1 系统总体架构

为提高移动机器人在复杂环境中的自主导航性能, 本文构建了融合群智能算法与强化学习的分层式导航系统。系统由环境感知、状态表示、全局搜索、策略学习与运动控制五部分组成。环境感知模块基于

激光雷达与深度传感器获取障碍物信息; 状态表示模块通过栅格化处理与局部地图构建, 将多源数据转化为可学习的连续状态。全局搜索模块由群智能算法完成, 负责生成具有全局合理性的参考路径; 策略学习模块由强化学习承担, 通过与环境交互获得线速度和角速度等动作输出, 实现对动态障碍的实时应对; 运动控制模块将动作指令转换为可执行的控制量, 完成闭环导航。不同于传统“全局规划+局部避障”的串行架构, 本文提出的融合框架实现了群智能算法与强化学习的双向交互: 群智能算法为强化学习提供初始策略与全局参考, 而强化学习的价值评估又会反向影响群智能算法的搜索方向, 使系统逐步形成协同优化机制, 从而兼具全局路径质量与局部决策灵活性。

### 1.2 群智能算法与强化学习的协同优化机制

在全局优化层, 粒子群算法用于对强化学习策略网络进行参数初始化, 并在训练关键节点进行参数重置, 以缓解强化学习初期收敛不稳定、易陷入局部最优等问题。其适应度函数综合路径平滑度、碰撞次数与输出稳定性, 从高维参数空间中搜索出一组更优的初始权重, 提高训练起点质量。蚁群算法用于在栅格地图中生成参考路径, 通过信息素更新与启发式函数形成整体导航趋势, 但其输出不作为硬性约束, 而是以软引导方式融入强化学习的奖励结构, 使策略更易沿全局优路径方向收敛。在局部策略学习层, 本文采用兼具稳定性与连续控制能力的强化学习算法。状态空间包含机器人位姿、局部环境、障碍物位置以及参考路径偏移量等信息; 动作空间为速度控制指令。奖

励函数由距离变化、避障、安全惩罚与路径偏移惩罚构成,可在每一步提供及时反馈,缓解稀疏奖励带来的学习困难。训练过程中,强化学习将高价值状态持续反馈给群智能算法,使后者在下一轮搜索中更聚焦潜在优区域,实现全局与局部的动态协同。

通过上述结构,群智能算法提供全局视角,强化学习提供局部实时性,两者的交互更新使导航策略在复杂场景中同时具备收敛速度快、路径质量高与动态适应能力强等优势。

## 2 结果

### 2.1 实验设置

实验环境包括仿真平台与真实机器人实验。仿真部分使用 Gazebo 构建三类地图:静态稀疏环境、静态密集障碍环境与动态多障碍环境。真实实验在机器人平台 Turtlebot3 上进行,通过 ROS 进行控制。实验比较算法包括:纯 RL、纯 PSO、RL+ACO 串行融合、RL+PSO 串行融合、本文提出的 RL-SIA 深度融合方法。评价指标包括平均路径长度、平均执行时间、避障成功率、策略稳定性(方差)与训练收敛速度。

数据采集通过重复执行任务 50 次获得统计结果。为排除偶然效果,所有算法均采用相同初始位置与环境条件。

### 2.2 仿真结果分析

在静态稀疏环境中,各算法均可完成导航,但性能差异明显。纯 RL 策略在训练初期表现较差,偶尔出现碰撞,收敛速度慢;纯 PSO 虽具备全局搜索能力,但由于实时性不足,路径执行时间较长;RL+ACO 串行融合可获得一定的路径平滑性,但动态避障能力不佳。相比之下,本文提出的融合方法在路径长度、控制平滑性与成功率方面表现最优。其路径长度相较纯 RL 下降约 23%,平均执行时间减少 18%,避障成功率保持在 98% 以上。

在静态密集障碍环境中,群智能算法的全局参考路径发挥显著作用,RL 能够沿参考路径进行局部调整。实验结果进一步显示,融合方法在复杂环境中能够保持较高的策略稳定性。与串行融合方法相比,其性能方差降低了约 35%,说明策略学习更稳定。

在动态环境实验中,随机移动的动态障碍对导航策略提出了更高要求。纯 PSO 和 ACO 无法应对实时避障,纯 RL 在训练过程中出现大量失效。本文方法通过强化学习的快速响应能力,结合群智能提供的全局策略引导,实现了对动态障碍的有效规避。实验中,融合方法的动态避障成功率达到 91%,显著高于纯 RL (63%)。

### 2.3 真实机器人实验

真实环境的复杂性远高于仿真环境,本研究选择实验室走廊、仓储货架区与多移动障碍场景进行测试。融合方法在三个场景中均表现稳定。路径执行过程中机器人轨迹与全局参考路径保持高度一致,局部调整灵活有效,未出现明显抖动或策略异常。机器人在空间狭窄区域能够依靠 RL 形成的局部策略进行平滑避障,ACO 提供的全局路径引导避免其进行无效探索,整体表现优于其他对照算法。

## 3 讨论

### 3.1 融合方法性能优势的原因分析

融合强化学习与群智能算法的方法在多个维度上提升了导航性能。其核心在于全局搜索与局部优化的互补特征,使策略既具备全局最优性,又具备实时适应能力。PSO 提供的参数初始化避免了 RL 训练初期的随机探索造成的震荡,使训练过程更快、更平滑;ACO 提供的全局参考路径使 RL 获得明确的方向引导,减少无效探索;RL 通过状态价值反馈调整 ACO 的信息素分布,使 ACO 在后续搜索中更加聚焦关键区域。通过双向耦合机制,RL 不再单纯依赖自身的梯度信息,而是获得外部智能体的搜索支持,从而扩大有效探索区域。

### 3.2 融合机制的适用条件与局限性

尽管融合方法具有显著优势,但仍存在一定局限。首先,该方法对环境建模精度有一定要求,若环境地图存在较大误差,ACO 提供的参考路径可能偏离真实可行路径,从而降低 RL 训练效果。其次,融合算法结构相对复杂,训练时间长于单一算法;在机器人计算资源有限的场景中需要对两类算法的计算结构进行轻量化设计。此外,在高度动态的环境中,全局路径的参考价值可能下降,系统需设计更灵活的全局路径更新机制,以避免全局路径对 RL 形成误导。

### 3.3 对实时性与算力消耗的影响

融合算法的实时性取决于强化学习的推理速度与群智能算法的更新频率。由于机器人执行阶段只需调用强化学习策略网络进行实时推理,群智能算法主要在训练阶段运行,因此对实时性影响较小。实验结果表明,融合方法在推理速度与纯 RL 相当,不会导致明显延迟。然而,训练阶段的算力需求相对较高,需要 GPU 加速以保证训练效率。面向资源受限设备,可采用模型蒸馏、网络剪枝等技术减少算法体积。

### 3.4 未来研究方向

未来研究可从以下三个方向展开:第一,可将更

多群智能算法如灰狼优化、萤火虫算法等引入融合框架,构建更强的多算法协同系统;第二,可引入可解释强化学习与可解释群智能,使导航策略的决策过程透明化,以便应用于更高安全级别的场景;第三,可在真实复杂环境如户外无人配送、矿山勘探等场景进行验证,进一步提升算法的鲁棒性与泛化能力。

#### 4 结论

本研究围绕移动机器人在复杂动态环境中的自主导航问题,构建了融合强化学习与群智能算法的协同优化框架,通过全局搜索与局部学习的深度耦合,实现了导航策略在全局最优性、局部适应性与动态鲁棒性上的综合提升。研究表明,群智能算法在训练初期提供的参数初始化与全局参考路径显著改善了强化学习在稀疏奖励环境中的探索效率,使策略收敛更稳定;强化学习则通过持续的状态价值反馈优化群智能搜索空间,使全局路径逐步向高价值区域收敛,从而形成双向正反馈机制。在实际测试中,融合方法在路径长度、规划时间、避障成功率与策略稳定性等指标上均优于单一算法与串行融合方法,特别是在动态障碍场景中展现出更强的实时响应能力和策略鲁棒性。总体而言,该融合框架有效解决了传统群智能算法实时性不足与强化学习容易陷入局部最优的缺陷,为移动机器人在工业生产、智能物流与复杂作业场景中的

自主导航提供了一种可扩展、高性能的技术路径。未来可进一步扩展融合算法的可解释性、轻量化设计及跨场景迁移能力,以推动其在更大规模的真实应用中落地。

#### 参考文献:

- [1] 康亮,杜奕,尹丽华.一种面向异质多移动机器人的改进猫群算法[J].系统仿真学报,2024,36(08):1958-1968.
- [2] 董琳.基于群智能算法的移动机器人室内定位及路径规划方法研究[D].山东大学,2023.
- [3] 杜映峰,陈万米,范彬彬.群智能算法在路径规划中的研究及应用[J].电子测量技术,2016,39(11):65-70.
- [4] 江岩.基于群智能算法的移动机器人路径规划研究[D].安徽工程大学,2020.
- [5] 于诗茵.改进的鸡群优化算法及其应用研究[D].北方民族大学,2025.
- [6] 沈汝涵.基于群智能优化算法的管道辅助检测研究[D].安徽理工大学,2023.
- [7] 张晨.基于群智能算法的移动机器人导航系统设计及应用[D].上海工程技术大学,2016.
- [8] 莫定界,赵杰,凌港,等.混合粒子群与改进灰狼算法的移动机器人路径规划[J].软件导刊,2025,24(07):54-60.