

DOI:10.19344/j.cnki.issn1671-5276.2026.02.024

基于无人艇路径规划方法的海上救援技术研究

毛君赫^a, 李峰^a, 马裕清^b, 胡家庆^a

(中国人民解放军海军军医大学 a. 卫生勤务学系; b. 海军医学系, 上海 200433)

摘要: 为了提高海上救援的效率和安全性, 根据无人艇的性能优化动作空间, 采用 Q-Learning 算法学习适应复杂多变的环境, 选择模拟退火算法保证全局最优解的搜索。结果表明: 在较小规模海上环境中, 该方法的平均路径长度比传统的模拟退火算法短了 10.02%, 平均耗时减少了 1.41%, 且没有出现急转角。

关键词: 路径规划; USV; Q-Learning; 模拟退火算法; 海上救援

中图分类号: R197.32 **文献标志码:** B **文章编号:** 1671-5276(2026)02-0126-05

Maritime Rescue Technology Reserch Based on Unmanned Boat Path Planning Method

MAO Junhe^a, LI Feng^a, MA Yuqing^b, HU Jiaqing^a

(a. Department of Health Services;

b. Department of Naval Medicine, Naval Medical University of the People's Liberation Army, Shanghai 200433, China)

Abstract: In order to improve the efficiency and safety of maritime rescue, this study optimizes the action space based on the performance of unmanned boats, adopts Q-Learning algorithm to learn and adapt to complex and changing environments, and selects simulated annealing algorithm to ensure the search for the global optimal solution. The results show that the average path length of the adopted method in small-scale offshore environments is 10.02% shorter than traditional simulated annealing algorithms, and the average time is reduced by 1.41%, without sharp corners.

Keywords: path planning; USV; Q-Learning; simulated annealing algorithm; sea rescue

0 引言

海上救援是一项重要的任务, 其关乎人身安全和财产损失的最小化。当前的海上救援方式主要包括自救和互救两种方式, 在互救方面往往依靠专业的海上救援队伍、救援船只、直升机等^[1]。其中, 无人艇(unmanned surface vessel, USV)具有机动灵活、环境适应性强、活动区域广、使用成本低等优点, 这在海上救援方面有着显著的优势。王伟等^[2]为了减少 USV 雷达监测跟踪的误差, 构思了一种自适应阈值分割匹配算法, 结果表明该算法的距离误差小于 2%。路径规划是 USV 智能程度的重要体现, 通过合理的路径规划, USV 能够自主导航、自主避障和自主识别与探测, 从而极大地提升其自主性能。DVORAK 等^[3]针对 USV 的避障规划, 通过遗传算法(genetic algorithm, GA)为 USV 提供全局规划, 结果表明该设计保证了实用性。同时, 在风、浪、流的扰动等复杂海况下, 高

效的路径规划能够保证 USV 实现快速响应海上救援功能, 确保 USV 的安全、有效的执行任务。Q-Learning 在路径规划方面的应用是通过强化学习的方法, 来训练智能体学习如何在环境中找到最优路径。该算法便于理解, 计算简单, 鲁棒性强, 在复杂和多变的海洋环境中具有较强的鲁棒性, 能够适应环境的变化。周治国等^[4]为了提高 USV 的避障能力, 开发了一种结合长短期记忆网络的深度 Q 网络(deep Q network, DQN)算法, 加速收敛并保存训练信息, 结果表明该算法存在可行性, 但在超大维度的环境下, 该算法的计算开销急剧增加, 这会导致算法在计算资源有限的情况下难以实际应用, 并且降低学习速率。因此, 研究创新性地结合 USV 性能优化 Q-Learning 的动作空间, 并以模拟退火算法(simulated annealing, SA)自适应调整策略。研究旨在提高 Q-Learning 在复杂海洋环境中 USV 路径规划的效率和效果, 使其更适合实际应用, 提升海上救援的效率。这

第一作者简介: 毛君赫(2002—), 女, 北京人, 硕士研究生, 研究方向为军事后勤学, maomaozhang966@yeah.net。

通信作者简介: 胡家庆(1973—), 男, 上海人, 教授, 博导, 博士, 研究方向为军队卫勤保障、海上无人智能化搜救, jiaqinghu@126.com。

有助于推动 USV 在海上救援中的实际应用,为应对未来的海上救援工作提供了有力的技术支持。

1 基于 Q-Learning 和 SA 的 USV 路径规划

1.1 结合 USV 性能优化 Q-Learning 动作空间

Q-Learning 不依赖于环境的先验知识,USV 即使对环境不了解也能进行学习和规划。在复杂多变的海洋环境中,USV 无需预先知道环境的详细信息,便可以通过自我探索来找到最佳路径^[5]。由于海域环境特殊,具有一定复杂性,在此环境下,可根据 USV 回转特性优化算法的动作空间,以便保证获取的路径来满足 USV 的操纵特性。USV 的动作空间优化及其航行示意图,如图 1 所示。

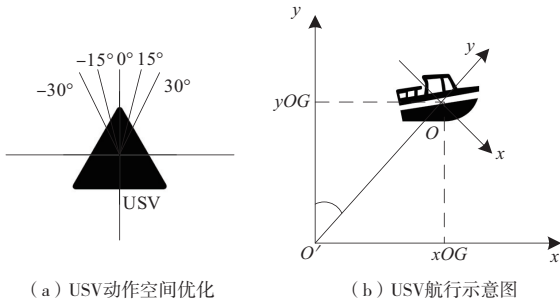


图 1 USV 的动作空间优化及其航行示意图

图 1(a)中,优化 USV 的航向角度,设置其动作空间为 $[-30^\circ, -15^\circ, 0^\circ, 15^\circ, 30^\circ]$ 。这样可以更好地满足 USV 在实际航行中的操纵特性,保证航行过程中的安全性和节能性,提高路径规划的效率和质量。这一优化方法能够有效克服传统栅格法的局限性,使得 USV 在复杂多变的环境中能够更加稳定和高效地进行路径规划^[6-7]。图 1(b)中,根据优化之后的动作空间,重新定义和建模 Q-Learning 的状态空间和状态-动作 Q 值表格。通过 USV 的航向和位置信息设计状态空间,USV 所在的海洋范围为 $xO'y$, xOG 、 yOG 分别为 USV 质心 G 在 x 轴、y 轴的投影。设定 USV 的初始位置为 $(x(O), y(O))$,航行速度 v 保持不变, t 时刻其位置为

$$\begin{cases} x(t) = x(O) + \int v \cdot \sin\varphi_0 dt \\ y(t) = y(O) + \int v \cdot \cos\varphi_0 dt \end{cases} \quad (1)$$

式中 φ_0 为 USV 航向。因此 USV 在航行时航向和其所在位置坐标有关,以此为基础,搭建 Q 学习

的状态空间。尽管 Q-Learning 在路径规划中有诸多优势,但在海上救援 USV 应用中也存在一些问题。假设 USV 可能的动作较大,或者环境的可能状态非常复杂,Q-Learning 的学习过程会变慢。这是因为在大的动作和复杂的状态空间中,Q 值表格的维度会显著增加^[8]。因此,研究结合 USV 的性能对 Q-Learning 算法进行优化,通过合理的状态和动作空间简化,可以减少学习过程中的计算开销和时间。为了保证 USV 学习重点关注区域,提升学习速度,研究需要先查询 USV 学习时的状态,若该状态不存在,需要添加后更新 Q 值,反之则查询后跳转^[9]。 ϵ -贪婪策略的探索率 ϵ 为固定不变的,行动选择过程的随机策略 $\pi(s)$ 为

$$\pi(s) = \begin{cases} \operatorname{argmax} Q(s, a), & \sigma > \epsilon \\ a \in A(s), & \sigma \leq \epsilon \end{cases} \quad (2)$$

式中: Q 为值函数; s 为状态; a 为动作; $A(s)$ 为动作的集合; σ 为随机数,当随机数大于 ϵ ,操作根据当前状态下的 Q 值最佳行为决定,若小于 ϵ ,操作行为随机。

1.2 基于 Q-Learning 和 SA 的路径规划方法

在 Q-Learning 的路径规划上,SA 能够有效跳出局部最优解,通过模拟物理退火过程,逐步降低温度,在搜索空间内进行广泛探索,从而找到全局最优路径。SA 在初期接受较差解的概率较高,随着温度降低,逐渐减少接受较差解的概率,保证在早期有足够的探索,后期能逐步收敛到全局最优解^[10-11]。在 ϵ -贪婪策略中,为了保证 ϵ 能够自适应调整,提高路径规划效果的稳定性,研究选择 SA 调整 ϵ ,以便能够更有效地探索环境,从而加快学习过程。迭代 k 次的温度衰减函数 T_k 为

$$T_k = \max[f(sq(Q_N - Q_{N-1})), v_s] \cdot T_{k-1} \quad (3)$$

式中: k 为迭代次数; sq 为 Q 值表格变化的平方和; Q_N 为迭代第 N 次的 Q 值表格; v_s 为温度衰减速率。其中归一化 sigmoid 函数 f 关系式为

$$f(s) = \frac{1}{1 + e^{-s}} \quad (4)$$

每次迭代时,为了确定 ϵ 的变化,研究选择 Metropolis 准则验证。该准则是 SA 算法中的一个重要部分,用于决定是否接受一个新的解。对于当前状态 s 和下一状态 s' ,如果 s' 的目标函数值优于 s ,即 $Q(s) \leq Q(s')$,则无条件接受新解,则 ϵ 的情况为

$$\epsilon = \begin{cases} 1, & Q(s) \leq Q(s') \\ \exp\left(-\frac{Q(s) - Q(s')}{T_k}\right), & \text{other} \end{cases} \quad (5)$$

式中 $Q(s)$ 、 $Q(s')$ 分别为当前状态 s 、下一状态 s' 的 Q 值。当 T_k 较高时,接受较差解的概率较高,允许更多的探索。随着 T_k 逐渐降低,接受较差解的概率减小,算法逐步收敛到最优解。在 SA 作用下自适应调整,可以保证 ε 随 Q 值表格的变化

情况进行动态调整,保证充分探索环境,并在后期保持算法的稳定性,确保利用和探索的平衡^[12-13]。优化 Q-Learning 的 USV 路径规划流程示意图,如图 2 所示。

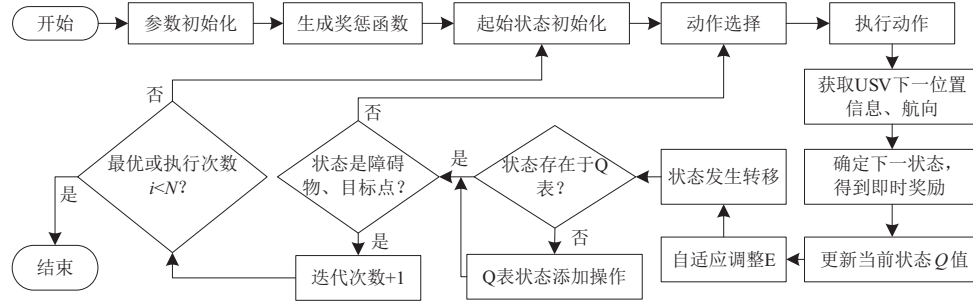


图 2 优化 Q-Learning 的 USV 路径规划流程示意图

图 2 中,确定仿真环境生成地图之后设置 USV 中相关参数,比如最大迭代次数 n 、学习步长 α 、奖励折扣 γ 、 T_k 、 v_s 等结合目标点信息和设计奖励函数。初始状态为 USV 的起点和航向,之后进行动作选择并执行,确定下一状态 s_{i+1} ,更新当前状态 Q 值,迭代表达式为

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha(r(t) + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t)) \quad (6)$$

式中 $Q(s_t, a_t)$ 为当前时刻状态-行为对应的值函数。调整结束后,状态发生转移 $s \leftarrow s_{t+1}$ 。判断状态 s ,若 USV 的 s 位于障碍物区域,需要增加一次迭代次数。当最优或执行次数 $i < N$,则输出最优路径策略 $\pi^*(s)$,其关系式为

$$\pi^*(s) = \operatorname{argmax}_a Q(s, a) \quad (7)$$

USV 在航行过程中的行为决策是否正确,且其是否有效地避开了障碍物,可以用奖惩函数进行评价。奖惩函数对 USV 的路径规划学习过程具有导向作用,帮助无人艇逐渐学习到最佳的路径规划策略。考虑 USV 与目标点和障碍物的距离以及速度大小,通过负反馈信号和速度约束,实现了对 USV 航行安全性的保证。通过这种方式,USV 能够逐渐学习到最优的路径规划策略,确保安全高效的航行。关于障碍物的总惩罚函数 R 为

$$R = R_1 + R_2 = -100 \quad (8)$$

式中 R_1 、 R_2 分别为海上矩形、圆形的障碍物惩罚函数。当 USV 做出一个符合目标的行动,比如接近目标点、成功避开障碍物等就会得到正奖励。反之,做出接触障碍物、偏离路线等不利于目标的行动,则会受到惩罚。

2 基于 Q-Learning 和 SA 算法的 USV 路径规划结果分析

研究选择的海洋环境仿真环境生成地图软件为 OpenSeaMap,其专注于航海地图的生成和共享。该软件能够提供全球范围内的航海地图,并可集成海洋环境数据。其中地图设置了两个规模的海洋环境,分别为 $200 \text{ km} \times 200 \text{ km}$ 和 $400 \text{ km} \times 400 \text{ km}$ 。实验设置学习步长为 0.8,初始 ε 为 1,奖励折扣因子为 0.9。SA 算法的控制参数设置中,温度下限值为 50,温度衰减速率为 0.85,初始温度为 2 000。为了验证所提出的算法的实用性和合理性,研究进行了仿真实验,保持相同的速度,对比 SA 算法在不同环境大小下从起始点到目标点的路径规划路线,如图 3 所示。

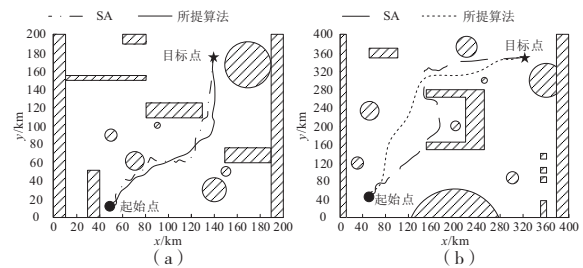


图 3 不同算法在不同海况下的路径规划路线

图 3(a) 中,对比 SA 算法,所提算法的路线满足所有的约束条件并且是最优的,平均长度为 434.6 km,平均耗时 3.488 s,并且未出现急转角的情况。SA 算法花费的时间更长,约为 3.538 s,平均路径长度为 483.0 km,急转角出现了 7 个。所提算法的平均路径长度比 SA 算法短了 10.02%,平均耗

时减少了 1.41%,同时优化动作空间的设计存在可行性,保证路径规划合理性。图 3(b)中,规模越大 SA 的急转角有 13 个,平均花费 11.98 s,平均路径长度为 118.4 km,而所提算法未出现急转角,平均花费 10.92 s,平均路径长度为 95.38 km。所提算法的平均路径长度比 SA 算法短了 19.44%,平均耗时减少了 8.85%。

为了更准确客观地评价提出的算法的有效性,实验选择在 200 km×200 km 规模的环境上进行,两种算法在相同的条件下分别进行了 10 次仿真实验。每次实验路径规划的最优长度对比,如图 4 所示。

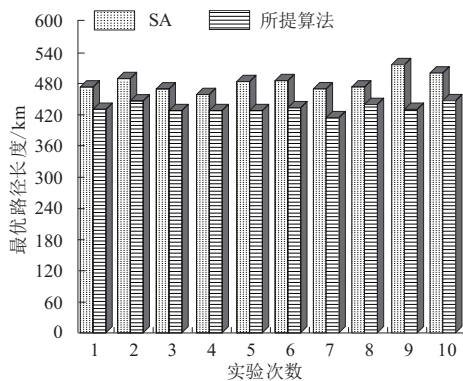


图 4 不同算法最优路径长度对比

图 4 中,10 次仿真实验结果提出的算法的最优路径长度均小于 SA 算法,其在路径规划性能上更具有优势。所提算法通过更有效的路径规划,使得 USV 能够在海上环境中更高效地确定最优路径,从而提高整体的执行效率和精确度。为了分析所提算法的合理性,研究与 3 种路径规划算法进行对比验证。对比算法选择 SA、改进 GA^[14] 和基于 K-means 的路径规划算法^[15]。将不同算法均各自训练 10 次,每一次的训练均包含 1 500 回合。不同算法的平均总步数和平均奖励对比,如图 5 所示。

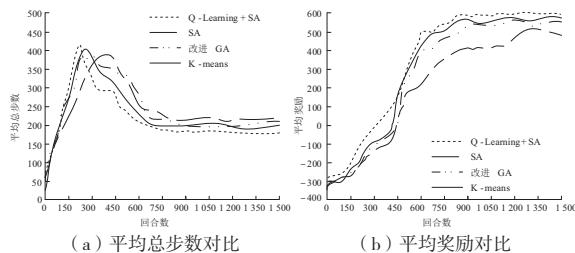


图 5 不同算法的训练曲线对比

图 5(a)中,所提算法的平均总步数在 220 回合数时达到最多,约为 400。之后平均总步数达到最多的分别为基于 K-means 的路径规划算法、

改进 GA 和 SA。在 600 回合数时,所提算法的平均总步数表现逐渐平稳,说明该算法能有效提升路径规划效率。此时 USV 正常到达目标点。图 5(b)中,大约 500 回合数之后,所提算法的平均奖励明显高于其余 3 种算法,在 600 回合数时所提算法的平均奖励约为 500。在 800 回合数时,所提算法的平均奖励变化开始趋于稳定,约为 600。该算法能够提供更好的路径规划结果和避障效果。所提算法在较少的回合数内即能达到稳定的平均奖励和总步数,从而显示出其高效性。

不同算法经过训练后,能够适应特定的环境和任务需求。在训练阶段,算法通过反复试验来优化其策略,使其在特定情境下能够有效完成路径规划任务。测试在与训练阶段相似的环境中进行,以确保测试结果的公平性和可比性。环境中包括起始点、目标点以及障碍物等,与训练环境相匹配。对比不同算法的奖励值、总路径长度和总步数指标,结果如表 1 所示。

表 1 不同算法的测试表现对比

参数	算法				平均值
	所提算法	SA	改进 GA	K-means	
奖励值	602.3	490.3	554.8	569.6	554.3
总路径长度/m	793.9	882.5	850.7	824.1	837.8
总步数	175.0	229.0	213.0	198.0	203.8

表 1 中,4 种算法的性能排名从高到低依次为所提算法、K-means、改进 GA、SA。对比 K-means 算法,所提算法的奖励值高了 5.7%,总路径长度短了 3.7%,总步数少了 11.6%。所提算法的奖励值比 SA 高出 22.8%,在路径规划的有效性和整体性能上的显著优势。其在总路径长度上明显比 SA 短了 10.0%,说明该算法能够找到更优的路径,使 USV 以更短的距离到达目标点,提高了路径规划的效率。其在总步数上比 SA 少了 23.6%,能够更快速地完成任任务,提高了整体的任务效率。整体来看所提算法在奖励值、总路径长度和总步数 3 个关键指标上都优于其他算法,表现出优良的综合性能,适合在复杂海域中实现高效、安全和可靠的路径规划。

3 结语

为了提高路径规划的准确性和稳定性,本文研究提出了一种结合 Q-Learning 和 SA 算法的 USV 路径规划方法。结果表明:在 400 km×400 km

大小的海上地图中该算法的平均路径长度比 SA 算法短了 19.44%, 平均耗时减少了 8.85%。该方法的平均总步数在 220 回合数时达到最多, 约为 400。600 回合数所提算法的平均总步数表现逐渐平稳, 说明该算法能有效提升路径规划效率。大约 500 回合数之后, 所提算法的平均奖励明显高于其余 3 种算法, 在 800 回合数时其平均奖励变化开始趋于稳定, 约为 600。该算法能够提供更好的路径规划结果和避障效果, 并且在较少的回合数内即能达到稳定的平均奖励和总步数, 显示出其高效性。该方法综合性能较好, 在复杂海域中可以更好地完成海上救援路径规划。但研究没有考虑风速、流速等环境因素的影响, 未来可以深入考虑这些因素的精细设计, 以便提高路径规划的决策质量。

参考文献:

- [1] 刘雅奇, 史誉洲, 曹石勇. 船舶航迹数据采集与可视化在海上搜救中的应用[J]. 舰船科学技术, 2023, 45(18): 166-169.
- [2] 王伟, 杜旭洋, 杨志伟, 等. 基于无人艇的导航雷达目标检测跟踪算法[J]. 系统工程与电子技术, 2024, 46(5): 1561-1572.
- [3] DVORAK M, DOLEZEL P, STURSA D, et al. Genetic algorithm-based task assignment for fleet of unmanned surface vehicles in dynamically changing environment[J]. Cybernetics and Systems, 2024, 55(6): 1314-1331.
- [4] 周治国, 余思雨, 于家宝, 等. 面向无人艇的 T-DQN 智能避障算法研究[J]. 自动化学报, 2023, 49(8): 1645-1655.
- [5] 贾薇, 王宁, 张富宇, 等. 基于多传感器融合 SLAM 的无人船靠泊技术[J]. 大连海事大学学报, 2023, 49(4): 65-75.
- [6] 徐茂竹, 李弘, 李亚光, 等. 单目视觉引导下的无人艇局部避障方法[J]. 重庆邮电大学学报(自然科学版), 2023, 35(4): 732-741.
- [7] 夏家伟, 朱旭芳, 罗亚松, 等. 基于深度强化学习的无人艇轨迹跟踪算法研究[J]. 华中科技大学学报(自然科学版), 2023, 51(5): 74-80.
- [8] 李子怡, 胡祥涛, 张勇乐, 等. 基于虚拟目标制导的自适应 Q 学习路径规划算法[J]. 计算机集成制造系统, 2024, 30(2): 553-568.
- [9] 黄岩松, 姚锡凡, 景轩, 等. 基于深度 Q 网络的多起点多终点 AGV 路径规划[J]. 计算机集成制造系统, 2023, 29(8): 2550-2562.
- [10] 洪涛清, 高雪芬. 改进 Q 学习下 3D 打印机器人路径生成方法[J]. 计算机仿真, 2023, 40(10): 417-421.
- [11] 王典, 周阳, 宋毅, 等. 基于 Q 学习的生物启发式目标导向导航路径规划模型[J]. 电子测量与仪器学报, 2023, 37(6): 68-76.
- [12] 凌琛. 考虑能耗的无人机路径规划及海洋捕食者算法求解[J]. 自动化技术与应用, 2024, 43(1): 5-8, 13.
- [13] 梁华尘, 周稳, 倪旭东, 等. 面向多目标联合电力巡检的车载无人机协同巡检路径规划方法[J]. 南方电网技术, 2023, 17(11): 138-147.
- [14] 郭兴华, 赵仓龙. 海上交通高峰航线船舶疏导路径规划仿真分析[J]. 舰船科学技术, 2023, 45(22): 206-209.
- [15] 张立华, 周寅飞, 贾帅东, 等. 一种有效顾及复杂海域避碰的路径规划方法[J]. 哈尔滨工程大学学报, 2023, 44(1): 56-64.

收稿日期: 2024-09-30

(上接第 125 页)

- [16] 朱冬生, 罗小平, 钱颂文. 微小流道(槽道)传热强化与节能[M]. 北京: 科学出版社, 2014.
- [17] 陶文铨. 数值传热学[M]. 2 版. 西安: 西安交通大学出版社, 2001.
- [18] 梁珍祥. 凹坑凸胞板式换热器性能的数值研究[D]. 郑州: 郑州大学, 2012.
- [19] 黄建伟. 类卵形肋柱对微通道流动传热特性的影响[D]. 吉林: 东北电力大学, 2021.
- [20] LI W M, LUO K, LI C, et al. A remarkable CHF of 345W/Cm² is achieved in a wicked-microchannel using HFE-7100[J]. International Journal of Heat and Mass Transfer, 2022, 187: 122527.
- [21] 田若男. 微冷却系统结构优化及流动换热特性研究[D]. 北京: 华北电力大学, 2021.
- [22] MOHIT M K, GUPTA R. Numerical investigation of the performance of rectangular micro-channel equipped with micro-pin-fin[J]. Case Studies in Thermal Engineering, 2022, 32: 101884.
- [23] 罗志宁, 张莹莹, 高磊, 等. 板壳式换热器壳程传热性能分析[J]. 机械制造与自动化, 2017, 46(2): 94-97, 129.

收稿日期: 2024-01-24