

面向物流机器人的改进 Q-Learning 动态避障算法研究

王力^{1,2}, 赵全海¹, 黄石磊¹

(1. 河南中烟工业有限责任公司漯河卷烟厂, 河南 漯河 462005;

2. 中国人民解放军战略支援部队信息工程大学, 郑州 450000)

摘要: 为提升物流机器人 (AMR) 在复杂环境中的自主导航与避障能力, 改善传统 Q-Learning 算法在动态环境中的收敛速度慢、路径规划不够优化等问题; 研究引入模糊退火算法对 Q-Learning 算法进行路径节点和搜索路径优化, 删除多余节点和非必要转折; 并为平衡好 Q-Learning 算法的探索和利用问题, 提出以贪婪法优化搜索策略, 并借助改进动态窗口法对进行路径节点和平滑加速改进, 实现局部路径规划, 以提高改进 Q-Learning 算法在 AMR 动态避障中的搜索性能和效率; 结果表明, 改进 Q-Learning 算法能有效优化搜索路径, 能较好避开动态障碍物和静态障碍物, 与其他算法的距离差幅至少大于 1 m; 改进算法在局部路径中的避障轨迹更趋近于期望值, 最大搜索时间不超过 3 s, 优于其他算法, 且其不同场景下的避障路径长度和运动时间减少幅度均超过 10%, 避障成功率超过 90%; 研究方法能满足智慧仓储、智能制造等工程领域对物流机器人高效、安全作业的需求。

关键词: 物流机器人; Q-Learning 算法; DWA; 多目标规划; 障碍物; 避障

Improved Q-Learning Dynamic Obstacle Avoidance Algorithm for Logistics Robots

WANG Li^{1,2}, ZHAO Quanhai¹, HUANG Shilei¹

(1. Luohe Cigarette Factory, Henan Tobacco Industry Co., Ltd., Luohe 462005, China;

2. Information Engineering University, Strategic Support Force of the Chinese People's

Liberation Army, Zhengzhou 450000, China)

Abstract: To enhance the autonomous navigation and obstacle avoidance capabilities of logistics robots (AMRs) in complex environments, and to address the issues of slow convergence speed and not optimized path planning in traditional Q-Learning algorithms within dynamic settings, this paper introduces a fuzzy annealing algorithm to optimize path nodes and search paths in the Q-Learning algorithm, and eliminate redundant nodes and unnecessary transitions. To balance the exploration and exploitation in the Q-Learning algorithm, a greedy method is proposed to refine the search strategy. Additionally, an improved dynamic window method is employed to enhance path nodes and smooth acceleration, and implement effective local path planning, which significantly improves the search performance and efficiency of the enhanced Q-Learning algorithm in AMR dynamic obstacle avoidance scenarios. The results demonstrate that the improved Q-learning algorithm effectively optimizes the search path, adeptly avoiding both dynamic and static obstacles, with a distance advantage of at least 1 m over other algorithms. The obstacle avoidance trajectory of the improved algorithm in local path planning is closer to expected values, achieving a maximum search time of no more than 3 s, which outperforms other algorithms. Furthermore, the enhanced algorithm reduces the obstacle avoidance path length and motion time by more than 10% in various scenarios, with an obstacle avoidance success rate of over 90%. This algorithm effectively meets the needs for efficient and safe operations of logistics robots in engineering fields such as smart warehousing and intelligent manufacturing.

Keywords: logistics robots; Q-Learning algorithm; DWA; multi-objective planning; obstacle; obstacle avoidance

收稿日期: 2025-01-05; 修回日期: 2025-02-14。

基金项目: 中国航发产学研合作项 (HFZL2021CXY007); 航空发动机及燃气轮机基础科学中心项目 (P2022-B-V-002-001)。

作者简介: 王力 (1976-), 男, 硕士, 工程师。

通讯作者: 黄石磊 (1986-), 男, 硕士, 工程师。

引用格式: 王力, 赵全海, 黄石磊, 等. 面向物流机器人的改进 Q-Learning 动态避障算法研究[J]. 计算机测量与控制, 2025, 33(3):

267-274.

0 引言

数字化时代的发展以及电子商务的兴起使得快递物流得到了较多增长,但传统物流方式难以克服劳动力成本上升以及物流运输过程复杂等现实问题,仅借助人力资源来完成物品搬移及分拣等工作,不可避免存在运营成本高以及出错率高等问题,且在效率、准确性和灵活性方面逐渐显露出较大的局限性^[1]。因此在该发展背景下,物流机器人(AMR, autonomous mobile robot)概念应运而生,并逐渐成为快递物流领域不可或缺的一部分。AMR通过搭载多种传感器以及集成各种导航技术,自主高效且准确地完成货物的搬运、分拣、装载和卸载等任务,极大缩减处理时间,提高物流行业的工作效率^[2-3]。AMR凭借其自主性、灵活性和智能化的特点,能够适应多样化的物流场景,有效提高物流运作的效率和质量,但其在落地过程中存在诸多挑战,如技术成熟度、避障安全性以及操作灵敏性等^[4]。不同学者也尝试对AMR的应用进行了研究,文献[5]为提高非结构环境下多机器人编队的动态避障能力,以六足机器人为例,采用领航-跟随策略,设计分群一致性编队控制策略。结果表明,该方法能有效实现动态避障,应用效率提升效果明显。文献[6]考虑到机器人工作环境的动态随机性问题,提出以改进人工势场法结合静态与动态机制,预测并规避障碍物的方法。结果表明,所提方法较常规方法的避障效率和导航耗时均有不同程度的提高,提升幅度超过20%。文献[7]以改进深度Q网络实现机器人避障规划,即通过设计障碍学习规则,优化奖励机制,动态调整探索因子来实现。结果表明,该改进方法的避障成功率提升幅度远超过50%,缩短了路径长度。文献[8]将人工蜂群优化算法和递归神经网络相结合,并设计了一个混合控制器来应对多移动机器人在不同环境下的规划避障问题。仿真结果表明,该控制器的应用偏差较小,且较现有控制器的路径长度有明显提高,提升幅度大于10%。文献[9]提出高斯混合模型结合高斯混合回归的协作机器人路径规划方法以达到避障效果,并以果蝇优化算法提高计算效率。结果表明,该方法能有效提高规划路径平滑度和避障有效性。

在复杂多变的物流环境中,AMR需要实时感知周围的障碍物,并迅速做出准确的避障决策,以避免碰撞和保障货物的安全运输。自动化操作、多任务处理、精细作业以及较低的成本优势使得AMR配送效率的提升效果明显。但其也不可避免面临着感知决策能力仍需提升、续航能力、协同作业以及维修费用等挑战。如文献[10]考虑到多载位仓储机器人路径规划任务冲突问题,提出借助多智能体强化学习任务监管器实现冲突消解方

案。文献[11]在考虑困难指数和运行时间因素约束下,提出设计AMR移动路径选择函数以缩短运行时间。传感器技术的限制、算法复杂度挑战、避障实时性以及环境适应性等都限制了AMR的进一步发展,且随着科技创新的不断发展,多机器人系统正逐步成为研究的新领域。如何协调多个机器人的避障行为,避免相互之间的碰撞和干扰,是一个亟待解决的问题。其中Q-Learning算法鼓励最优动作价值函数逐渐接近期望值,在不依赖环境条件的同时,借助交互学习来获取最优策略。Q-Learning算法更适用于复杂多变的应用环境,且其在动态避障领域具有一定的应用潜力。因此研究以Q-Learning算法为基础,针对避障环境及多机器人协调创新地提出Q-Learning算法改进方法,并在改进Q-Learning算法全局规划基础上引入动态窗口法(DWA, dynamic window approach)进行进一步改进,以实现AMR动态避障中的全局和局部规划,旨在为AMR在物流领域的广泛应用提供更加可靠和高效的技术支持,推动物流行业的智能化发展。

AMR的动态避障性能直接关系到其在复杂环境中的运行效率和安全。研究从全局避障和局部避障两个方面来进行设计。全局避障主要关注的是在整个工作环境中,为AMR规划一条从起点到终点的无碰撞路径,涉及到对整个环境障碍物位置的动态预测和全局路径规划。局部避障则侧重于AMR在行进过程中遇到突发障碍物时的即时反应和路径调整,要求其计算出绕过障碍物的最佳局部路径,确保AMR能够安全地继续其任务。

1 改进 Q-Learning 算法

Q-Learning算法不依赖于数学模型的构建,通过不断与环境进行交互和学习,获取最大收益下的动作值,体现在物流AMR中,则可根据场景环境中的障碍物位置、速度和运动方向等信息,计算出应采取的最优行动,如转向、加速等^[12]。但动态避障所要求的环境变化较为复杂,致使Q-Learning算法在收敛速度和稳定性上存在一些局限。Q-Learning算法具有较强的适应性和自学习能力,但其在复杂的动态环境中,可能存在Q值表更新过程缓慢,Q值表的存储和计算成本也可能会随着状态和动作空间的增大而急剧增加^[13-14]。因此研究尝试借助模拟退火算法(SA, simulated annealing)对Q-Learning算法进行迭代优化,包括对路径节点和搜索路径的优化。最后借助节点优化算法删除Q-Learning算法中未经过障碍物节点连接线的多余节点和非必要转折点,直至最后得到最优的规划路径。SA算法主要是借助循环迭代来判断是否接受新结果,直至满足终止条件^[15]。SA算法通过引入随机因素和概率接受机制,可

以有效帮助 Q-Learning 算法跳出局部最优解, 探索全局最优解, 找到全局最优路径。具体来看, 借助 SA 算法优化 Q-Learning 算法时, 首先需要将 Q 值表进行初始化, 并设置初始温度和降温速率。之后依旧当前状态和贪婪策略选择动作, 其数学表达见式 (1):

$$a = \begin{cases} \text{Random action}, \epsilon \\ \text{argmax}_a Q(s, a'), 1 - \epsilon \end{cases} \quad (1)$$

式中, a 表示在状态 s 下选择的动作, 表示动作值函数 $Q(s, a')$, ϵ 为选择随机动作的概率, $1 - \epsilon$ 为选择最优动作的概率。执行动作并观察新状态和即时奖励, 更新 Q 值。改进算法的 AMR 全局路径规划流程具体如图 1 所示。

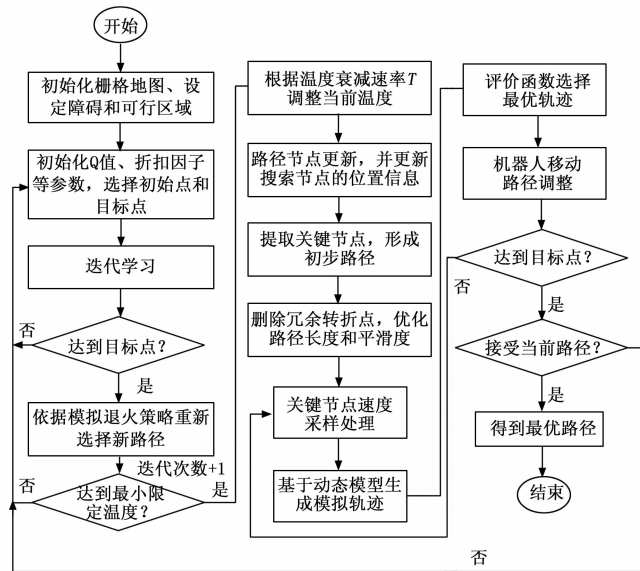


图 1 改进 Q-Learning 算法的全局路径规划

图 1 中, 首先对栅格地图和各参数进行初始化处理, 其中 Q 值更新的数学表达见式 (2):

$$Q'(s, a) = (1 - \alpha) * Q(s, a) + \kappa * [R(s, a) + \sigma * \max Q'(s, a)] \quad (2)$$

式中, $Q'(s, a)$ 表示状态 s 和动作 a 更新后的 Q 值, R 为奖励值, $\max Q'(s, a)$ 为新状态下的最大 Q 值, κ 为学习率, σ 为折扣因子。SA 算法计算当前解的能量和新解的能量, 并根据概率和约束条件, 实现对温度的更新, 得到更新后的路径节点。初始化处理后, 依据原有 Q-Learning 算法进行迭代处理, 若达到目标点则借助 SA 算法进行路径选择, 并对最优路径进行节点优化。最后对关键节点进行速度采样处理, 并基于评价函数选择最优轨迹, 得到最后的最优路径。

2 全局动态避障

Q-Learning 算法在规划过程中, 因环境的复杂性和不确定性可能导致算法的学习过程较为漫长, 迭代路径存在前期起伏较大等问题, 一定程度会影响到其避障效

果的稳定性^[16]。因此, 为了提高 Q-Learning 算法的动态避障能力, 研究在 Q-Learning 算法路径规划优化基础上进行动态避障。首先进一步扩大搜索的区域范围, 将传统搜索方向的 4/8 方向扩展为 16 方向, 细化规划步长, 并依据步长状态设定不同的奖励函数, 划定区域边界。由于 Q-Learning 算法在搜索时间与学习速度上存在矛盾, 其主要体现在智能体需要在探索新策略和利用已知策略之间找到一个平衡点。如果智能体过于保守 (即过于注重利用), 那么它可能会错过发现更优策略的机会, 从而导致学习速度变慢; 而如果智能体过于冒险 (即过于注重探索), 那么它可能会花费大量时间在那些看似无意义的状态上, 从而导致搜索时间增加^[17]。因此为平衡好 Q-Learning 算法的探索和利用问题, 研究借助 ϵ -greedy 贪婪法进行改进。借助概率的突变型来优化探索利用选择, 探索因子的引入可在充分借助先验知识的同时改进搜索策略^[18]。式 (3) 为常规贪婪法的数学表达:

$$\pi(a | s) = \begin{cases} 1 - \epsilon + \frac{\epsilon}{|A(s)|} \\ \frac{\epsilon}{|A(s)|}, other \end{cases} \quad (3)$$

式 (3) 中的第一列约束条件为 $a = \text{arg} a \in A \text{max} Q(s, a)$, 探索因子的范围为 $[0, 1]$, 式 (3) 容易影响探索效率。因此, 研究依据探索步长的有效状态对搜索概率进行适当增减, 提高搜索效率, 得到调整后的探索因子, 见式 (4):

$$\epsilon = \frac{\epsilon}{(1 + \lg_{10} N \Delta x / 2x)} \quad (4)$$

式中, ϵ 表示探索概率, 为 Δx 阶段补偿探索的有效回报状态, N 为步长总数量。图 2 为全局避障设计思路。

图 2 中, ϵ -greedy 贪婪的引入优化 Q-Learning 算法的探索因子, 提高搜索性能, 并在更新过程中实现全局路径的动态避障。

3 局部动态避障设计

全局规划避障能为机器人提供了一条相对最优的路径来想着目标前进, 但在实际路径过程中, 不可避免会出现各种突发情况, 因此要求机器人需依据环境感知信息来应用突发情况和实现动态避障^[19-20]。电机在驱动机器人时会有定义好的最大加速度和减速度, 其主要与电机特性、机械结构以及重量等因素有关。研究设计机器人的运动学模型, 示意图如图 3 所示。

假定两时刻下的机器人均作直线运动, 则可得到式 (5):

$$\begin{cases} x(t+1) = x(t) + l(t+1) \cos \theta(t) \Delta t \\ y(t+1) = y(t) + l(t+1) \sin \theta(t) \Delta t \\ \theta(t+1) = \theta(t) + \tau(t+1) \Delta t \end{cases} \quad (5)$$

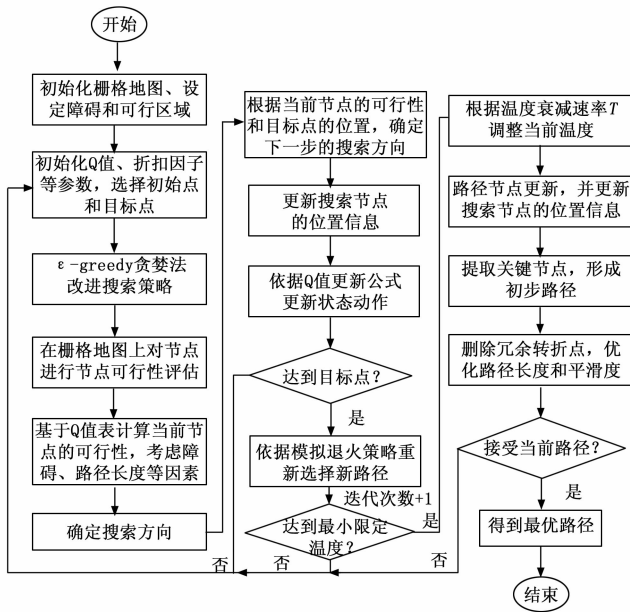


图 2 基于改进 Q-Learning 算法的全局动态避障设计思路

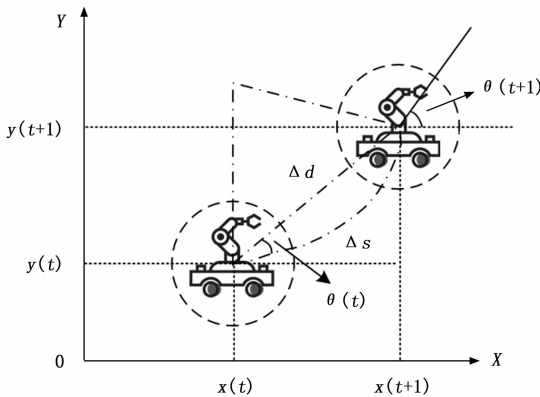


图 3 机器人的运动学模型

式中, t 表示时间, l 为线速度, τ 为角速度^[21]。机器人在遇到障碍物时, 需保证机器人在距离障碍物行驶轨迹最短长度 2 倍时做出减速行为。考虑机器人电机限制可有效降低碰撞风险, 提高机器人的安全性与可靠性^[22]。因此研究限定机器人速度范围, 并以轨迹评定的评价函数 $P(v, w)$ 来选取最优轨迹, 其可描述为机器人目标运动方向的夹角偏差 $H(v, w)$ 与最小避障距离 $\text{dist}(v, w)$ 和轨迹速度特性函数 $Ve(v, w)$ 之间的关系, 具体见式 (6):

$$P(w) = \sigma(\alpha \times H(w) + \beta \times \text{dist}(w) + \gamma Ve(w)) \quad (6)$$

式中, α, β, γ 为运动分配权重, β 为轨迹平滑因子, w 为机器人移动速度。

基于物流机器人的动力学模型, 研究引入 DWA 方法来弥补 Q-Learning 算法在局部规划上的缺陷。DWA

可以将机器人运动线速度和角速度作为分析对象, 通过预测时间轨迹和评价函数来选取最优轨迹^[23-24]。考虑到传统 DWA 算法容易陷入局部最小区域以及路径引导节点过多的问题, 研究对其进行节点引导和平滑路段加速改进, 得到改进 DWA 算法。具体来看, 对改进的 Q-Learning 算法的路径节点进行筛选, 删除相邻路段线段角度差小于 20° 的路径节点, 以剩下节点作为 DWA 的临时目标点。假定路径节点集合为 $J = [j_1, j_2, j_3, \dots, j_n]$, 节点 i 的方向向量可表示为 $d_i = \frac{j_{i+1} - j_i}{\|j_{i+1} - j_i\|}$, 节点的方向向量角度可表示为式 (7):

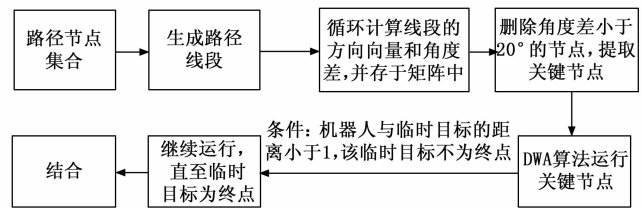
$$\theta_i = \text{atan2}(d_i^{(2)}, d_i^{(1)}) \quad (7)$$

式中, $d_i^{(1)}, d_i^{(2)}$ 表示点方向向量的第一和第二分量。相邻路段的线段角度差可表示为式 (8):

$$\Delta\theta_i = \theta_{i+1} - \theta_i \quad (8)$$

引入全局路径关键节点为临时目标, 判断机器人与临时目标点之间的距离来依次选择, 直至达到最后的目标点。之后研究基于改进 Q-Learning 对全局平滑路段进行改进, 以提高 DWA 算法的搜索速度, 减少旁侧障碍物对 AMR 速度的限制。通过引入航向障碍距离, 评估机器人在当前航向上的障碍物威胁程度, 并将其与制动距离进行比较, 评价函数的距离仍作为评分使用^[25]。得到 AMR 动态避障规划的方法, 如图 4 所示。

节点优化:



平滑加速:

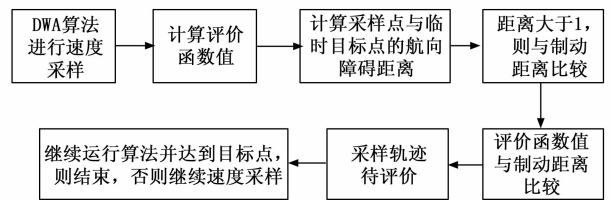


图 4 AMR 动态避障规划方法

图 4 中, 对关键节点的引导可删除部分平直路径的节点, 选取转折处的节点。依据机器人与临时目标点的距离计算来选取下一临时目标点。在平滑加速部分, 航向障碍距离的判断更符合前向轨迹预测结果。研究借助 Q-Learning 算法全局路径规划优势来实现 AMR 动态避障, 并引入改进 DWA 进行进一步改进, 实现局部环境精细规划, 能够有效提高 AMR 移动中的灵活性。

4 实验与分析

研究借助 RGB-D 深度相机 (1080P)、STM32F40-7VET6 主控芯片、Jetson Nano 主控板搭载机器人, 并在机器人操作系统 (Robot Operating System, ROS) 上进行程序编写, 以 Ubuntu18.04 为自主导航系统。同时在物流机器人分析过程中, 研究借助 ROS 操作系统来实现软硬件的通讯, 其导航系统包括机械本体、硬件控制部分、传感器采集部分和软件操作层。其机械本体包括底盘、机器人机架及防撞装置, 硬件控制模块为 STM32 控制器, 软件操作系统会发送控制指令趋势电机转动, 并保证机器人按照设定内容到达目的地。机器人运动过程中被采集的信息会经由编码器传递给上层控制器, 主控板会依据传感器信息进行控制。图 5 为机器人控制平台。

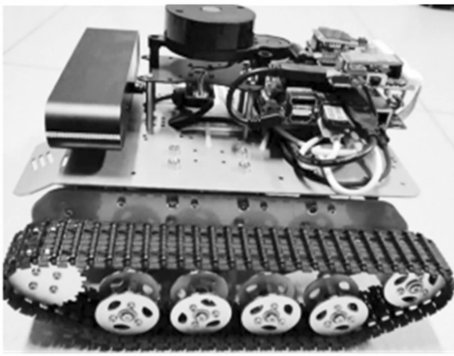


图 5 机器人控制平台

物流机器人的搜索路径控制和动态避障是实现其自主导航和高效作业的关键技术, 其中搜索路径主要是在已知或未知环境中规划并优化行驶路径的过程, 需考虑路径的长度、时间、安全性等多个因素, 选择最优路径。还需在局部规划位置根据障碍物的位置、速度等信息, 动态调整机器人的行驶方向^[26-27]。根据规划的路径和避障算法的结果, 可实时调整机器人的运动状态, 确保其能在完成指定路径规划的同时, 有效避开障碍物。研究借助 MATLAB 2019a 软件进行仿真分析, 设定初始温度为 80 °C, 温度迭代次数为 8, 衰减系数为 0.45, 学习了和折扣因子分别为 0.8 和 0.9。首先对研究提出的改进 Q-Learning 算法与传统 Q-Learning 算法进行搜索路径结果比较, 结果如图 6 所示。

图 6 (a) 结果表明, 相同起点和终点情况下, 改进 Q-Learning 算法的搜索路径经过的节点数和转折次数都明显少于传统 Q-Learning 算法, 其最短路径小于 10 m, 最短路径提升幅度为 3.24%。图 6 (b) 结果表明, 改进 Q-Learning 算法的迭代曲线避免了传统 Q-Learning 算法前期起伏过大情况, 达到稳定值所需的迭代次数更小, 为 7 次。而传统 Q-Learning 算法的迭代

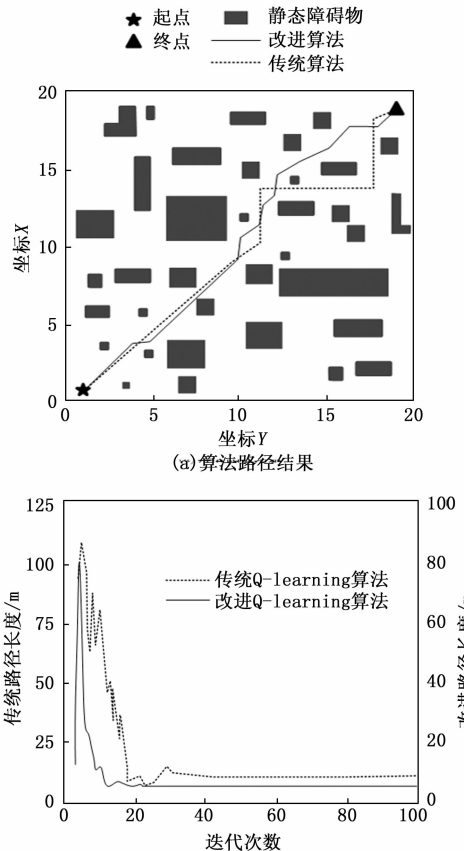


图 6 Q-Learning 算法改进前后的搜索路径和迭代结果

曲线表现出明显的节点波动情况, 在 20 次后趋于收敛。之后将研究提出的改进 Q-Learning 算法与改进动态窗口算法 (IDWA, improving dynamic window approach)、融合强化学习算法 (RL-RSA, reinforcement learning-reptile search algorithm)、改进深度 Q 网络算法 (IDQN, improving deep Q-Network) 和改进麻雀搜索算法 (ISSA, improving sparrow search algorithm) 进行比较, 搜索路径结果如图 7 所示。

图 7 (a) 结果表明, 改进 Q-Learning 算法的规划路径长度整体较短, 且存在较少的转折节点, 路径平滑度远高于其他比较算法。相比之下, IDWA 算法和 RL-RSA 算法的路径长度相对较长, 其转折点均多于改进 Q-Learning, 且表现出较为明显的路径随机性与不稳定性, 其更加偏向于栅格地图的周边, 在搜索过程中易陷入局部最优。IDQN 算法与研究算法的搜索结果相对较为贴合, 但其路径规划上的路径长度与平滑度仍有待提升, 在非密集障碍情况下的避障路径转弯角较多。而 ISAA 算法则因搜索策略的局限性, 导致部分节点的路径曲折且冗长, 规划路径最长。在图 7 (b) 迭代结果中, 改进 Q-Learning 算法的搜索效率与路径优化能力更是表现较为突出, 最小路径长度小于 20 m, 达到收敛的次数 (<20 次), 明显少于其他算法。其余比较算

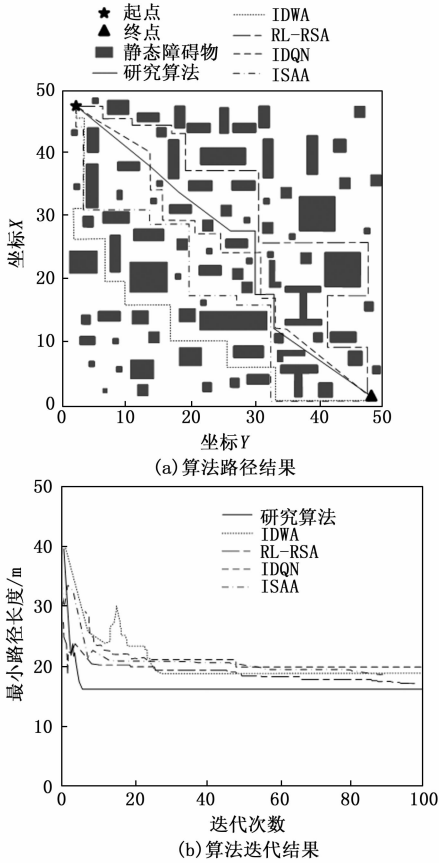


图 7 不同算法的搜索路径结果

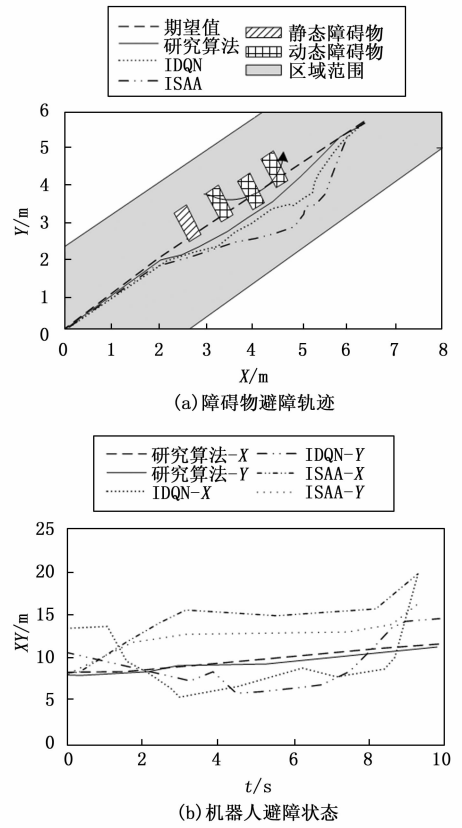


图 8 机器人避障结果

法在迭代后期所达到的最小路径长度较为先进，但在迭代次数小于 20 次之前，IDWA 算法的最短路径长度表现最差。改进 Q-Learning 算法凭借其高效的策略迭代与状态值函数估计，能有效对路径节点进行规划。

之后设计静态和动态两种障碍物，对研究提出的改进 DWA 融合设计思路进行检验，并与 IDQN 算法和 ISSA 算法进行比较。设置机器人的最大线速度和角速度分别为 1.0 m/s 和 0.349 rad/s，对应加速度为 0.2 m/s² 和 0.872 rad/s²，轨迹评价函数中的对应运动分配权重系数分别为 0.05、0.2 和 0.1。航向权重系数为 0.1，障碍物距离权重为 0.05。评估机器人的避障性能，结果如图 8 所示。

图 8 (a) 为机器人的避障轨迹，从中可以看出的是，研究算法的避障轨迹更趋近于期望值，且能较好避开动态障碍物和静态障碍物，整体轨迹曲线更为平滑。而 IDQN 算法和 ISSA 算法在避障动态障碍物时，所产生的轨迹路线较长，且存在明显的节点起伏情况，与研究算法至少存在 1 m 的最大距离差幅。图 8 (b) 为机器人的避障状态，IDQN 算法和 ISAA 算法的避障状态曲线表现出较大的斜率，且在两个方向上的曲线差幅较为明显，而研究所提算法的曲线斜率更平缓，避障过程更为稳定。上述结果表明，AMR 在改进 DWA 方法下

能稳定动态避障。之后对研究提出的改进 DWA 与 IDQN 算法进行规划性能比较，结果如图 9 所示。

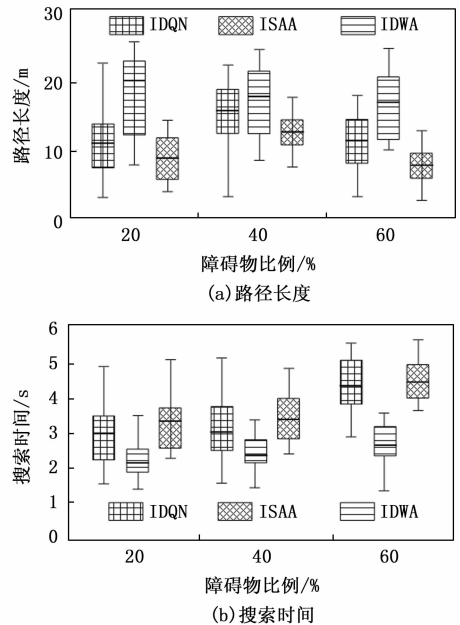


图 9 随机场景下的路径规划结果

图 9 (a) 结果表明，在不同障碍物比例下，改进 DWA 算法在 20%、40% 和 60% 障碍物下的路径长度平

均值分别为 20 m、17 m 和 16 m。而 IDQN 算法和 ISSA 算法的路径长度最大值不超过 15 m, 其中 ISSA 算法的路径长度受障碍物比例影响较为明显。图 9 (b) 结果中, 改进 DWA 算法在障碍物下的最大搜索时间不超过 3 s, 最小平均值更是达到了 2 s。其次表现较好的 IDQN 算法, 最后是 ISSA 算法。在 60% 障碍物下, IDQN 算法和 ISSA 算法的搜索时间均超过了 4 s, 避障规划性能有待提升。之后设置直线状态和曲线状态两种环境, 对物流机器人的动态避障效果分析, 结果如图 10 所示。

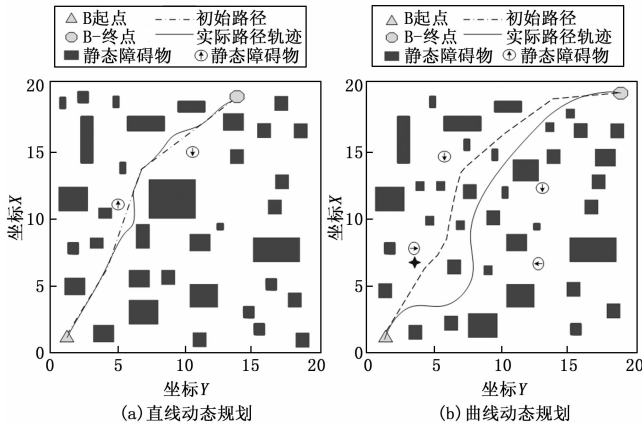


图 10 机器人避障性能

图 10 (a) 为直线动态障碍情况, 就初始路径来看, 机器人在规划路径过程中因避障目的可能存在碰撞风险。但经路径规划后, 机器人的实际运动轨迹发生了改变, 当检测到碰撞隐患后, 机器人会实现避障处理, 规划路线出现了平滑阶段。图 10 (b) 为曲线动态障碍情况, 其中机器人在避障过程中, 初始路径可能存在明显的相撞风险。借助这种融合算法来进行协调机器人时。而在实际路径中, 机器人重新规划路径, 并在避障情况下尽可能减少路径长度, 继续前进。上述结果表明, 研究所提出的避障策略能有效改善机器人行进状况, 实现动态避障。之后对研究提出的动态避障算法在不同场景下的运动规划结果进行检验, 包括时间花费情况、避障成功率等, 具体可见表 1。表 1 中的减少率是相较于传统 Q-Learning 算法而言的。

表 1 不同场景下的 AMR 应用性能

场景(起始点位置)	时间/s	路径长度/%	运行时间/%	避障成功率/%	搜索时间/%	平均安全距离/%
水平	4.4716	-21.25	14.50	94.31	-10.34	8.12
对角	4.4895	-18.34	-10.27	95.16	-8.67	5.64
斜角	5.9610	-11.33	-22.37	96.43	-7.12	6.58

表 1 结果表明, 在 3 种场景中, 机器人的路径长度和运动时间减少幅度均超过 10%, 避障成功率达到了 94.31%、95.16% 和 98.43%。且机器人在水平、对角

和斜角场景下的最大搜索时间减少幅度达到 10.34%, 平均安全距离均有不同程度的提升。

5 结束语

物流运输智能化时代的到来使得具有自动化性能的移动机器人越来越被应用于物流运输领域, 研究基于当前物流机器人的应用问题, 提出改进 Q-Learning 动态避障算法以提高其应用性能。对研究所提方法进行仿真分析, 结果表明, 相同起点和终点情况下, 改进 Q-Learning 算法的搜索路径经过的节点数和转折次数都明显少于传统 Q-Learning 算法, 其最短路径小于 10 m, 最短路径提升幅度为 3.24%。且改进 Q-Learning 算法存在较少的转折节点, 路径平滑度远高于其他比较算法, 而 IDWA 算法和 RL-RSA 算法的路径长度相对较长, 表现出较为明显的路径随机性与不稳定性。在局部规划结果中, 研究算法的避障轨迹更趋近于期望值, 且能较好避开动态障碍物和静态障碍物, 整体轨迹曲线更为平滑。而 IDQN 算法和 ISSA 算法在避障动态障碍物时, 存在明显的节点起伏情况, 与研究算法至少存在 1 m 的最大距离差幅。在局部规划结果中, 改进 DWA 算法的避障轨迹更趋近于期望值, 且能较好避开动态障碍物和静态障碍物, 而 IDQN 算法和 ISSA 算法存在明显的节点起伏情况, 与研究算法至少存在 1 m 的最大距离差幅。在不同障碍物比例下, 改进 DWA 算法在 20%、40% 和 60% 障碍物下的路径长度平均值分别为 20 m、17 m 和 16 m。而 IDQN 算法和 ISSA 算法的路径长度最大值不超过 15 m。且改进 DWA 算法在障碍物下的最大搜索时间不超过 3 s, 最小平均值更是达到了 2 s, 优于其他算法。在直线状态和曲线状态两种环境下, 物流机器人的实际运动轨迹发生了改变, 能实现避障处理, 规划路线出现了平滑阶段。且其在水平、对角和斜角场景下的最大搜索时间减少幅度达到 10.34%, 路径长度和运动时间减少幅度均超过 10%, 避障成功率达到了 94.31%、95.16% 和 98.43%, 平均安全距离均有不同程度的提升。

研究所提方法优化了路径规划长度, 在不同避障场景中展现出较好的适应性和鲁棒性, 有效降低了冲突碰撞风险。研究设计的贪婪改进方法能对搜索因子进行调整, 这能一定程度上减少误判或漏判情况。其在今后尝试增加从多传感器冗余检测、优先级避障、速度空间搜索等方面进行改进以提高物流机器人的安全性和可靠性。值得注意的是, 研究所提出算法在极端密集或高度动态环境中的表现有待进一步优化, 未来可考虑借助集成先进传感器技术和轮廓模型预测控制等先进算法来实现多模态感知。同时还可借助多机器人系统架构、领航者一跟随者策略等来解决通信协作和编队避障问题, 进

一步提升物流机器人在不同条件下的智能决策能力。

参考文献:

- [1] 陈 肯, 赵 慧, 蒋 林, 等. 动态障碍物环境下的扫地机器人路径规划 [J]. 传感器与微系统, 2024, 43 (7): 153 - 156.
- [2] 王豪杰, 马向华, 代婉玉, 等. 改进 DWA 算法的移动机器人避障研究 [J]. 计算机工程与应用, 2023, 59 (6): 326 - 332.
- [3] 侯嘉瑞, 万 熠, 梁西昌, 等. 面向未知环境的机器人动态路径规划算法研究 [J]. 单片机与嵌入式系统应用, 2022, 22 (1): 29 - 32, 38.
- [4] 吴 琼, 袁 杰, 马圣山, 等. 基于安全点引导混合算法的启发式机器人动态路径规划 [J]. 计算机集成制造系统, 2023, 29 (10): 3284 - 3295.
- [5] 庄红超, 王 柠, 董凯伦, 等. 非完整约束大负重比六足机器人多机动态协同编队避障控制策略 [J]. 机械工程学报, 2024, 60 (1): 284 - 295.
- [6] 章一鸣, 姚文广, 陈海进. 动态环境下自主机器人的双机制切向避障 [J]. 浙江大学学报 (工学版), 2024, 58 (4): 779 - 789.
- [7] 田箫源, 董秀成. 基于改进 DQN 的移动机器人避障路径规划 [J]. 中国惯性技术学报, 2024, 32 (4): 406 - 416.
- [8] KUMAR S, PARHI D R, MUNI M K. Path planning and obstacle avoidance of multi-robotic system in static and dynamic environments [J]. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 2023, 237 (9): 1376 - 1390.
- [9] HU Y, WANG Y, HU K, LI W. Adaptive obstacle avoidance in path planning of collaborative robots for dynamic manufacturing. Journal of Intelligent Manufacturing, 2023, 34 (2): 789 - 807.
- [10] 黄 捷, 詹维捷, 潘聪捷, 等. 面向智慧工厂的多仓储机器人路径规划仿真教学研究 [J]. 实验技术与管理, 2024 (11).
- [11] 冯仁宇, 夏 凯. 多重约束下智慧仓储搬运机器人移动路径选择研究 [J]. 现代电子技术, 2023, 46 (22): 1 - 4.
- [12] LI Z, LI J, WANG W. Path planning and obstacle avoidance control for autonomous multi-axis distributed vehicle based on dynamic constraints [J]. IEEE Transactions on Vehicular Technology, 2022, 72 (4): 4342 - 4356.
- [13] 邵浩然, 陈建松. RL-RSA 算法的移动机器人自主避障技术 [J]. 机械科学与技术, 2024, 43 (8): 1411 - 1417.
- [14] 田箫源, 董秀成. 基于改进 DQN 的移动机器人避障路径规划 [J]. 中国惯性技术学报, 2024, 32 (4): 406 - 416.
- [15] 朱金坛. 融合神经网络及麻雀算法的机器人避障研究 [J]. 计算机测量与控制, 2023, 31 (4): 258 - 263.
- [16] GAO X, ZHANG Q, LI M, et al. Mobile Robot Combination Autonomous Behavior Strategy to Inspect Hazardous Gases in Relatively Narrow Man-Machine Environment [J]. Chinese Journal of Mechanical Engineering, 2022, 35 (1): 1 - 13.
- [17] ZHAO K, NING L. Hybrid navigation method for multiple robots facing dynamic obstacles [J]. Tsinghua Science and Technology, 2022, 27 (6): 894 - 901.
- [18] LI S, HAN K, LI X, et al. Hybrid trajectory replanning-based dynamic obstacle avoidance for physical human-robot interaction [J]. Journal of Intelligent & Robotic Systems, 2021, 103 (3): 1 - 14.
- [19] JUAREZ-LORA A, RODRIGUEZ-ANGELES A. Bio-Inspired Autonomous Navigation and Formation Controller for Differential Mobile Robots [J]. Entropy, 2023, 25 (4): 1 - 16.
- [20] NAH M C, LACHNER J, HOGAN N. Robot control based on motor primitives: A comparison of two approaches [J]. The International Journal of Robotics Research, 2024, 43 (12): 1959 - 1991.
- [21] ZHENG K, WU F, CHEN X. Laser-based people detection and obstacle avoidance for a hospital transport robot [J]. Sensors, 2021, 21 (3): 1 - 24.
- [22] FAN Y, LI X, LIU X, et al. Multi-robot consensus formation based on virtual spring obstacle avoidance [J]. Mechanical Sciences, 2024, 15 (1): 195 - 207.
- [23] HU Y, WANG Y, HU K, et al. Adaptive obstacle avoidance in path planning of collaborative robots for dynamic manufacturing [J]. Journal of Intelligent Manufacturing, 2023, 34 (2): 789 - 807.
- [24] 陈 楠, 乔立春. 基于势场蚁群算法的仓储搬运机器人避障控制方法 [J]. 计算机测量与控制, 2024, 32 (8): 168 - 173.
- [25] KUMAR S, PARHI D R, MUNI M K. Path planning and obstacle avoidance of multi-robotic system in static and dynamic environments [J]. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 2023, 237 (9): 1376 - 1390.
- [26] TURHANLAR E E, EKREN B Y, LERHER T. Autonomous mobile robot travel under deadlock and collision prevention algorithms by agent-based modelling in warehouses [J]. International Journal of Logistics Research and Applications, 2024, 27 (8): 1322 - 1341.
- [27] TANG Y, QI S, ZHU L, et al. Obstacle avoidance motion in mobile robotics [J]. Journal of System Simulation, 2024, 36 (1): 1 - 26.