

文章编号: 1671-251X(2025)06-0081-07

DOI: 10.13272/j.issn.1671-251x.2025040020

基于内在动机强化学习算法的煤矿井下 运输机器人自主避障

赵克宝¹, 李灵锋¹, 陈茁¹, 韩骏¹, 尹瑞^{2,3}

(1. 河北建材职业技术学院 机电工程系, 河北 秦皇岛 066004;

2. 中煤张家口煤矿机械有限责任公司, 河北 张家口 076250;

3. 河北省高端智能矿山装备技术创新中心, 河北 张家口 076250)

摘要: 现有的机器人避障方法多依赖于预设规则或外部奖励信号, 难以适应煤矿井下复杂多变的环境。为实现煤矿井下运输机器人自主高效避障, 提出了一种基于内在动机强化学习(IM-RL)算法的机器人自主避障方法。煤矿井下运输机器人通过视觉传感器感知外界环境信息, 利用基于好奇心的内在动机取向函数计算判别外界环境属性的内部奖赏值, 利用外部动机奖励函数计算其动作属性的外部奖赏值, 结合内在动机取向函数的奖励权重和外部动机奖励函数的奖励权重, 计算运输机器人执行动作前后状态的综合奖赏值, 形成强化学习算法奖励机制, 通过深度置信网络对其状态进行训练和学习, 激励运输机器人主动探索未知环境, 同时利用自身记忆机制存储知识和经验, 通过不断学习训练实现自主避障。在静态环境、动态环境和煤矿井下实际环境中分别进行运输机器人自主避障实验, 结果表明: 基于IM-RL算法的机器人自主避障路径和搜索时间较短, 具有较强的泛化性和鲁棒性。

关键词: 内在动机; 强化学习; 运输机器人; 自主避障; 路径规划

中图分类号: TD67

文献标志码: A

Autonomous obstacle avoidance of underground coal mine transport robots based on intrinsic motivation reinforcement learning algorithm

ZHAO Kebao¹, LI Lingfeng¹, CHEN Zhuo¹, HAN Jun¹, YIN Rui^{2,3}

(1. Department of Mechanical and Electrical Engineering, Hebei Polytechnic of Building Materials, Qinhuangdao 066004, China; 2. China Coal Zhangjiakou Coal Mining Machinery Co., Ltd., Zhangjiakou 076250, China; 3. Hebei Province High-end Intelligent Mine Equipment Technology Innovation Center, Zhangjiakou 076250, China)

Abstract: Existing robot obstacle avoidance methods mostly rely on preset rules or external reward signals, making it difficult to adapt to the complex and variable underground environment in coal mines. To achieve autonomous and efficient obstacle avoidance for underground coal mine transport robots, an autonomous obstacle avoidance method for underground coal mine transport robot based on Intrinsic Motivation Reinforcement Learning (IM-RL) algorithm was proposed. The underground coal mine transport robot perceived external environmental information through visual sensors, calculated internal reward values for identifying external environmental attributes using a curiosity-driven intrinsic motivation orientation function, and computed external reward values for its action attributes using an external motivation reward function. By combining the reward

收稿日期: 2025-04-09; 修回日期: 2025-06-24; 责任编辑: 盛男。

基金项目: 国家重点研发计划项目(2017YFF0210606); 河北省高等学校科学研究计划项目(ZD2022018, ZC2024136)。

作者简介: 赵克宝(1977—), 男, 河北涿州人, 副教授, 硕士, 主要研究方向为计算机科学, E-mail: 99283920@qq.com。通信作者: 韩骏(1982—), 男, 河北抚宁人, 副教授, 高级工程师, 硕士, 研究方向为智能控制技术, E-mail: 384042235@qq.com。

引用格式: 赵克宝, 李灵锋, 陈茁, 等. 基于内在动机强化学习算法的煤矿井下运输机器人自主避障[J]. 工矿自动化, 2025, 51(6): 81-87.

ZHAO Kebao, LI Lingfeng, CHEN Zhuo, et al. Autonomous obstacle avoidance of underground coal mine transport robots based on intrinsic motivation reinforcement learning algorithm[J]. Journal of Mine Automation, 2025, 51(6): 81-87.



扫码移动阅读

weights of the intrinsic motivation orientation function and the external motivation reward function, it calculated a comprehensive reward value based on the robot's state before and after performing an action, forming the reward mechanism of the reinforcement learning algorithm. The robot's state was trained through a deep belief network, which encouraged the transport robot to actively explore unknown environments. Meanwhile, it used its own memory mechanism to store knowledge and experience, achieving autonomous obstacle avoidance through continuous learning and training. Autonomous obstacle avoidance experiments for the transport robot were conducted in static environments, dynamic environments, and actual underground coal mine environments. The results showed that robots using the IM-RL algorithm achieved the short obstacle avoidance paths and search times, demonstrating strong generalization and robustness.

Key words: intrinsic motivation; reinforcement learning; transport robot; autonomous obstacle avoidance; path planning

0 引言

煤矿井下运输机器人通常用于矸石、煤炭、设备和材料的运输。然而,运输机器人需要频繁穿梭于狭窄且障碍物繁多的巷道和综采工作面中,多数运输机器人不具备智能自主避障功能,躲避障碍物需依赖人工操作^[1],增加了工作人员劳动强度,若操作不当会导致翻车事故,存在一定的安全隐患。因此,研究具有自主避障功能的井下运输机器人对提高煤矿生产运输效率具有重要意义。

近年来,众多学者对机器人避障技术进行了研究。曹现刚等^[2]采用人工势场法改进快速随机扩展树算法,在此基础上提出了一种煤矸分拣机器人避障轨迹规划方法,设计了一种环境敏感型目标偏置策略来提升机器人路径局部特性。金将等^[3]针对蚁群算法应用于路径规划时存在的盲目搜索、路径平滑性差等缺点,提出了一种改进的蚁群算法,以提高搜索效果。张彪等^[4]提出了一种改进的动态寻优蚁群算法,实时评估和调整机器人路径,实现即时动态避障功能。王欣等^[5]建立机器人动力学模型,提出了一种判断转弯方向和转弯角度的策略,使得救灾机器人能够在井下实现自主避障。机器学习和人工智能的飞速发展,加速了其在煤矿机器人中的应用^[6]。巩固等^[7]提出了基于运动视频的光流场目标视觉识别方法,可识别煤矿井下环境目标。李芳威等^[8]通过时间弹性带(Time Elastic Band, TEB)算法进行局部路径规划。张立亚等^[9]基于红外摄像头与激光雷达多源传感融合技术,提出了一种基于子图像分割映射点云空间的机器人避障算法,具有较高的避障效率。宋秦中等^[10]提出了一种基于卷积神经网络(Convolutional Neural Network, CNN)的井下无人驾驶无轨胶轮车运动避障控制算法。郭爱军等^[11]提出了动态环境下的避障方法,实现了井下矿

车的动态自主避障和安全行驶。张可琨等^[12]利用障碍物评价子函数改进动态窗口(Dynamic Window Approach, DWA)算法,有效提升了搬运机器人动态避障能力。Yang Hongxia 等^[13]提出了一种改进 A*算法与增强型 DWA 算法相结合的机器人避障方案,对搜索点选择策略和代价函数进行优化,提升了机器人避障效率。Xu Zhenyang 等^[14]提出了一种基于改进 A*算法与 DWA 算法的移动机器人自主导航算法,采用 DWA 算法将移动机器人的驱动直接限制在速度空间内,同时与环境动态交互以实时调整机器人运动轨迹,实现了基于全局最优的未知障碍物规避。彭继国等^[15]提出了一种基于模糊控制的井下机器人智能视觉避障方法,通过模糊控制算法计算机器人的转向角度和加速度,实现井下机器人智能避障。王利民等^[16]提出了一种融合改进 A*算法与 DWA 算法的煤矿机器人路径规划方法,将改进 A*算法规划出的路径节点依次作为局部路径规划 DWA 的局部目标点,从而缩短路径长度,并通过调整 DWA 代价函数中的权值比例来提升避障性能。鲁志等^[17]提出了一种基于改进 A*算法与改进 DWA 算法的机器人动态避障方法,该方法在 A*算法中引入全局障碍物占比,在 DWA 算法中加入目标点代价子函数,从而实现移动机器人的动态避障。

现有的机器人避障方法多依赖于预设规则或外部奖励信号,难以适应井下复杂多变的动态环境。本文提出一种基于内在动机强化学习(Intrinsic Motivation Reinforcement Learning, IM-RL)算法的煤矿井下运输机器人自主避障方法。煤矿井下运输机器人通过视觉传感器感知外界环境信息,利用基于好奇心^[18-19]的内在动机取向函数^[20]计算其判断外界环境物体属性的内部奖赏值,利用外部动机奖励函数计算其动作属性的外部奖赏值,结合内在动机取向函数的奖励权重和外部动机奖励函数的奖励权重

计算其综合奖赏值,将此奖赏值作为强化学习算法^[21]奖励机制,通过深度置信网络对运输机器人状态进行训练和学习,激励运输机器人主动探索未知环境,使其向自身的记忆机制中存储知识和经验,通过不断学习和训练实现井下运输机器人自主避障。

1 IM-RL 算法

IM-RL 算法利用一个七元组模型 $\langle A, S, V(i), H, R(s_i, a_i), Q(s_i, a_i), p(s_i, a_i) \rangle$ 进行序列化描述,其中各元素具体含义如下。

1) A : 运输机器人动作集合。 $A=\{a_i|i=1, 2, \dots, m_a\}$, 其中 a_i 为运输机器人第 i 个动作,以当前机器人碰到或避开障碍物的动作、接近或远离目标点的动作等来描述, m_a 为运输机器人所有动作的数量。

2) S : 运输机器人状态集合。 $S=\{s_i|i=1, 2, \dots, m_s\}$, 其中 s_i 为运输机器人第 i 个动作对应的第 i 个状态,以当前机器人与障碍物的距离、当前机器人与目标点的距离等来描述, m_s 为运输机器人所有状态的数量。

3) $V(i)$: 评价函数。

$$V(i) = \max U^*(s_i, a_i) \quad (1)$$

式中 $U^*(s_i, a_i)$ 为运输机器人在状态 s_i 时执行动作 a_i 后对动作的评价。

4) H : 内在动机取向函数。取向函数用于引导运输机器人在环境中的学习方向,运输机器人根据环境中前后位置的取向函数变化对不稳定空间中的内部奖赏值进行更新。一般情况下,当运输机器人靠近障碍物时,内部奖赏值较低;当运输机器人靠近目标点时,内部奖赏值较高。

$$H = \frac{1 - \exp(-\lambda V(i))}{1 + \exp(-\lambda V(i))} \quad (2)$$

式中 λ 为内在动机取向函数的修饰参数, $0 < \lambda \leq 1$ 。

5) $R(s_i, a_i)$: 综合奖赏值,即运输机器人在状态 s_i 时执行动作 a_i 转移到下一状态 s_{i+1} 后所得到的综合奖赏值。

$$R(s_i, a_i) = \xi H + \eta r_{\text{ex}} = \xi \frac{1 - \exp(-\lambda V(i))}{1 + \exp(-\lambda V(i))} + \eta r_{\text{ex}} \quad (3)$$

式中: ξ 为内在动机取向函数的奖励权重; η 为外部动机奖励函数的奖励权重, $\xi + \eta = 1$; r_{ex} 为外部动机奖励函数。

6) $Q(s_i, a_i)$: 值函数。

将外部动机奖励函数 r_{ex} 分为 2 个部分:

$$r_{\text{ex}} = r_{\text{ex-obs}} + r_{\text{ex-goal}} \quad (4)$$

式中: $r_{\text{ex-obs}}$ 为运输机器人相对障碍物的运动方向、距离和位置所产生的奖励 ($r_{\text{ex-obs}} = -0.5$, 接近障碍物; $r_{\text{ex-obs}} = +0.5$, 远离障碍物; $r_{\text{ex-obs}} = 0$, 其他运动); $r_{\text{ex-goal}}$ 为运输机器人相对目标点的运动方向、距离和位置所产生的奖励 ($r_{\text{ex-goal}} = -0.5$, 远离目标点; $r_{\text{ex-goal}} = +0.5$, 接近目标点; $r_{\text{ex-goal}} = 0$, 其他运动)。

则 IM-RL 算法的值函数迭代公式为

$$Q(s_{i+1}, a_{i+1}) = Q(s_i, a_i) + \kappa \left[\xi \frac{1 - \exp(-\lambda V(i))}{1 + \exp(-\lambda V(i))} + \eta r_{\text{ex}} + \gamma Q(s_{i+1}, a_{i+1}) - Q(s_i, a_i) \right] \quad (5)$$

式中: κ 为学习因子; γ 为折扣因子。

7) $p(s_i, a_i)$: 运输机器人在状态 s_i 时执行动作 a_i 后的适应率。

2 基于 IM-RL 算法的煤矿井下运输机器人自主避障流程

基于 IM-RL 算法的煤矿井下运输机器人自主避障流程如图 1 所示,具体步骤如下。

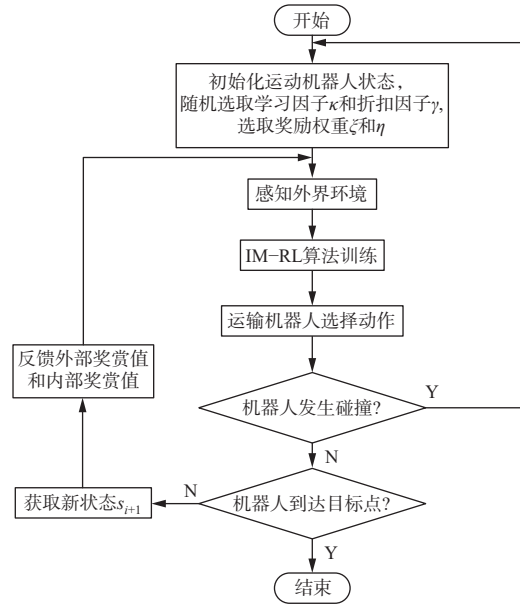


图 1 基于 IM-RL 算法的井下运输机器人自主避障流程

Fig. 1 Autonomous obstacle avoidance process of underground transport robot based on IM-RL algorithm

Step 1: 初始化当前运输机器人状态, 随机选取学习因子 κ 和折扣因子 γ , 同时根据运输机器人工作环境选择恰当的内在动机取向函数的奖励权重 ξ 和外部动机奖励函数的奖励权重 η 。

Step 2: 运输机器人通过视觉传感器感知外界环境信息, 随机获取运输机器人当前状态信息。

Step 3: 根据强化学习马尔可夫决策过程 (Markov Decision Process, MDP) 迭代公式计算当前可能执行

的动作的奖赏值,运输机器人根据当前奖赏值选择合适的动作 a_i 。

Step 4: 执行动作 a_i 后,运输机器人得到全新状态 s_{i+1} ,并随着下一状态的学习,通过内在动机取向函数计算其判断外界环境物体属性的内部奖赏值,利用外部动机奖励函数计算其动作属性的外部奖赏值。若机器人在训练过程中发生碰撞,则返回一定范围的初始状态重新开始训练。

Step 5: 根据不同的奖励权重、内部奖赏值和外部奖赏值计算综合奖赏值,形成强化学习算法奖励机制,通过外界环境反馈到深度置信网络进行下一步训练。

Step 6: 重复 Step 2 至 Step 5,直到完成自主学习。

3 机器人自主避障仿真实验

3.1 静态环境下机器人自主避障仿真实验

3.1.1 实验环境和评价指标

创建尺寸为 $36\text{ m} \times 36\text{ m}$ 的仿真实验环境地图,如图 2 所示,图中黑色区域表示实验环境中不可触及的障碍物区域,白色方格区域表示煤矿运输机器人可自由行走的无障碍区域。实验硬件配置为 13th Gen Intel(R) Core(TM) i9-13900K 3.00 GHz 处理器、12 GiB 的 NVIDIA RTX A2000 GPU、Windows10 操作系统,开发环境为 Matlab。实验采用运输机器人避障路径距离和到达目标点的搜索时间对 IM-RL 算法进行性能评估。

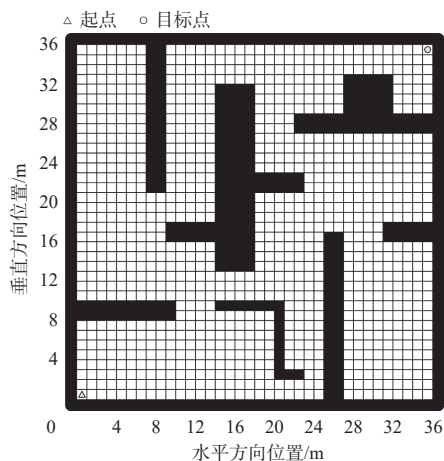


图 2 实验环境地图

Fig. 2 Experimental environment map

3.1.2 不同奖励权重下机器人静态避障仿真实验

为获得 IM-RL 算法的内在动机取向函数的奖励权重 ζ 和外部动机奖励函数的奖励权重 η ,进行机器人静态避障仿真实验,避障路径如图 3 所示,数据见表 1。可看出,当 $\zeta=0.95, \eta=0.05$ 时,外部奖励权重较

低,机器人识别外部环境障碍物和目标点的能力相对较弱,在搜寻目标位置时出现识别准确度低的情况;当 $\zeta=0.85, \eta=0.15$ 时,外部奖励权重较高,机器人识别外部环境障碍物和目标点的能力相对较强,导致机器人远距离躲避障碍物,进而出现冗余路径。当 $\zeta=0.9, \eta=0.1$ 时,机器人避障路径最优。

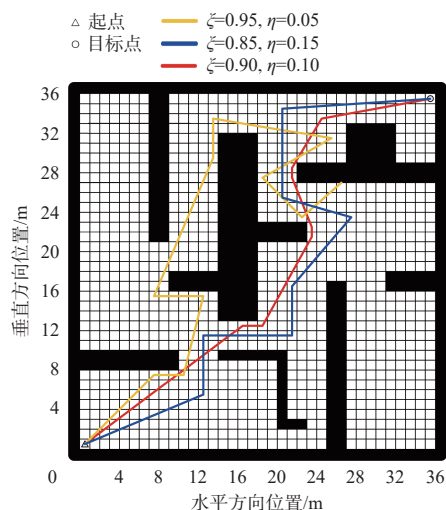


图 3 不同奖励权重下机器人静态避障路径

Fig. 3 Static obstacle avoidance paths of robot under different reward weights

表 1 不同奖励权重下机器人静态避障仿真数据

Table 1 Simulation data of robot static obstacle avoidance under different reward weights

奖励权重	路径距离/m	搜索时间/s
$\zeta=0.95, \eta=0.05$	76.59	—
$\zeta=0.85, \eta=0.15$	73.53	6.56
$\zeta=0.90, \eta=0.10$	56.70	5.68

3.1.3 不同算法下机器人静态避障仿真实验

IM-RL 算法、CNN 算法、混合 A*算法和改进 A*-DWA 算法下机器人静态避障仿真结果见表 2。可看出在静态环境中,与 CNN 算法、混合 A*算法和改进 A*-DWA 算法相比,IM-RL 算法的避障路径距离分别缩短了 4.26%, 3.16% 和 1.73%, 搜索时间分别缩短了 80.70%, 63.33% 和 46.57%, IM-RL 算法的效率更高。

表 2 不同算法下机器人静态避障仿真实验数据

Table 2 Simulation experiment data of robot static obstacle avoidance under different algorithms

算法	路径距离/m	搜索时间/s
CNN	59.22	29.42
混合 A*	58.55	15.49
改进 A*-DWA	57.70	10.63
IM-RL	56.70	5.68

3.2 动态环境下机器人自主避障仿真实验

3.2.1 不同奖励权重下机器人动态避障仿真实验

在静态障碍物仿真环境的基础上增加 $0.4\text{ m} \times 0.4\text{ m}$ 的正方形作为动态未知障碍物,其他参数不变。IM-RL 算法的内在动机取向函数的奖励权重 ζ 和外部动机奖励函数的奖励权重 η 取值不同时,机器人动态避障路径如图 4 所示,数据见表 3。可看出,当 $\zeta=0.95, \eta=0.05$ 时,外部奖励权重较低,机器人识别外部环境障碍物和目标点的能力相对较弱,在搜寻目标位置时出现识别准确度低的情况;当 $\zeta=0.85, \eta=0.15$ 时,外部奖励权重较高,机器人识别外部环境障碍物和目标点的能力相对较强,导致机器人远距离躲避障碍物,进而出现冗余路径。当 $\zeta=0.9, \eta=0.1$ 时,机器人避障路径最优。

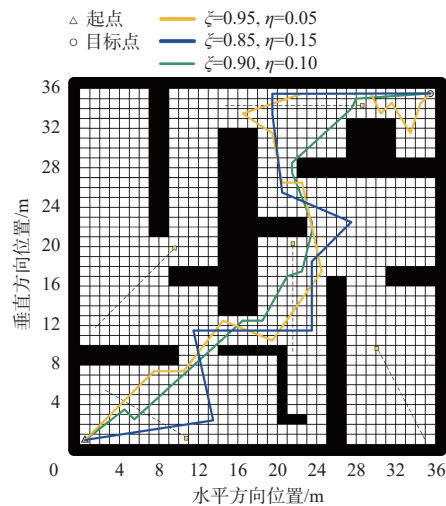


图 4 不同奖励权重下机器人动态避障路径

Fig. 4 Dynamic obstacle avoidance paths of robot under different reward weights

表 3 不同奖励权重下机器人动态避障仿真实验数据
Table 3 Simulation experiment data of robot dynamic obstacle avoidance under different reward weights

奖励权重	路径距离/m	搜索时间/s
$\zeta=0.95, \eta=0.05$	76.44	13.43
$\zeta=0.85, \eta=0.15$	82.21	12.39
$\zeta=0.90, \eta=0.10$	58.28	9.12

3.2.2 不同算法下机器人动态避障仿真实验

IM-RL 算法、CNN 算法、混合 A*算法和改进 A*-DWA 算法下机器人动态避障路径如图 5 所示,数据见表 4。可看出在存在未知障碍物的动态环境中,与 CNN 算法、混合 A*算法和改进 A*-DWA 算法相比,IM-RL 算法的避障路径距离分别缩短了 23.32%, 18.99% 和 12.98%, 搜索时间分别缩短了 75.06%, 72.51% 和 70.83%; IM-RL 算法下机器人能

实时、有效避开动态障碍物,表明 IM-RL 算法具有较高的鲁棒性。

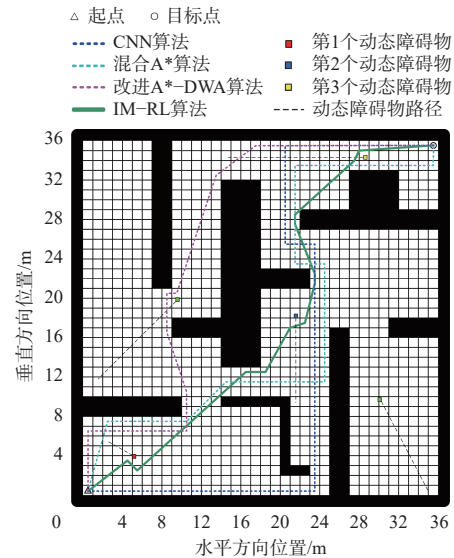


图 5 不同算法下机器人动态避障路径

Fig. 5 Dynamic obstacle avoidance paths of robot under different algorithms

表 4 不同算法下机器人动态避障仿真实验数据
Table 4 Simulation experiment data of robot dynamic obstacle avoidance under different algorithms

算法	路径距离/m	搜索时间/s
CNN	76.00	36.57
混合A*	71.94	33.18
改进A*-DWA	66.97	31.26
IM-RL	58.28	9.12

4 煤矿井下环境中机器人自主避障实验

由于煤矿井下环境黑暗,在机器人前端安装 2 个 LED 补光灯进行照明。在煤矿井下环境中,IM-RL 算法、CNN 算法、混合 A*算法和改进 A*-DWA 算法下机器人避障路径如图 6 所示,数据见表 5。可看出在煤矿井下环境中出现未知障碍物时,IM-RL 算法和改进 A*-DWA 算法能够实现实时避障;而 CNN 算法和混合 A*算法会出现与障碍物碰撞的情况;改进 A*-DWA 算法虽然能够实时避开障碍物顺利达到目标点,但其路径较长,影响运输效率;与 CNN 算法、混合 A*算法和改进 A*-DWA 算法相比,IM-RL 算法的避障路径距离分别缩短了 9.41%, 8.17% 和 14.49%, 搜索时间分别缩短了 79.65%, 78.58% 和 76.43%; IM-RL 算法下机器人能实时、有效避开未知障碍物,表明 IM-RL 算法泛化性能好。

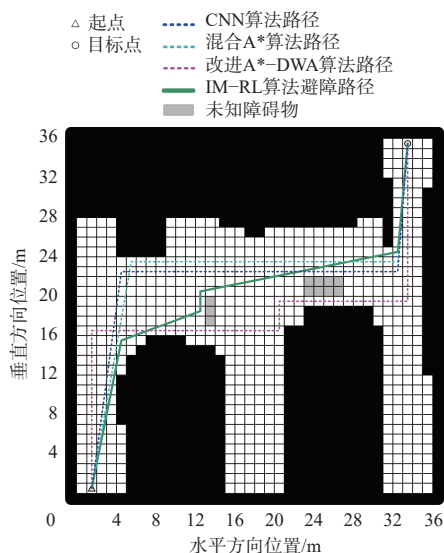


图 6 煤矿井下环境中不同算法下机器人避障路径

Fig. 6 Obstacle avoidance paths of robot under different algorithms in coal mine underground environment

表 5 煤矿井下环境中不同算法下机器人避障实验数据

Table 5 Experimental data of robot obstacle avoidance under different algorithms in coal mine underground environment

算法	路径距离/m	搜索时间/s
CNN	63.24	57.36
混合A*	62.39	54.47
改进A*-DWA	67.00	49.52
IM-RL	57.29	11.67

5 结论

1) 运输机器人利用视觉传感器感知外界环境信息,利用基于好奇心的内在动机取向函数计算运输机器人判断外界环境物体属性的内部奖赏值,结合运输机器人动作属性的外部奖赏值和奖励权重,计算运输机器人的综合奖赏值,激励运输机器人主动探索未知环境,使其向自身的记忆机制中存储知识和经验,通过不断的学习训练实现井下运输机器人自主避障。

2) 静态环境和动态环境下运输机器人自主避障仿真实验结果表明,当 IM-RL 算法的内在动机取向函数的奖励权重 $\zeta=0.90$ 、外部动机奖励函数的奖励权重 $\eta=0.10$ 时,运输机器人避障路径和搜索时间均达到最优。

3) 在静态环境、存在未知障碍物动态环境和煤矿井下环境中,IM-RL 算法具有较高的鲁棒性、泛化性和实时性,自主避障效果优于 CNN 算法、混合 A* 算法和改进 A*-DWA 算法。

参考文献(References):

- [1] 杨春雨, 张鑫. 煤矿机器人环境感知与路径规划关键技术[J]. 煤炭学报, 2022, 47(7): 2844-2872.
YANG Chunyu, ZHANG Xin. Key technologies of coal mine robots for environment perception and path planning[J]. Journal of China Coal Society, 2022, 47(7): 2844-2872.
- [2] 曹现刚, 藏家松, 吴旭东, 等. 基于 AE-RRT* 的煤矸分拣机器人避障轨迹规划方法[J/OL]. 煤炭学报: 1-12 [2025-03-27]. <https://link.cnki.net/doi/10.13225/j.cnki.jccs.2024.1195>.
CAO Xiangang, ZANG Jiasong, WU Xudong, et al. Obstacle avoidance trajectory planning method for coal gangue sorting robot based on AE-RRT*[J/OL]. Journal of China Coal Society: 1-12 [2025-03-27]. <https://link.cnki.net/doi/10.13225/j.cnki.jccs.2024.1195>.
- [3] 金将, 王小平, 臧铁钢, 等. 基于改进蚁群算法的机器人避障路径规划[J]. 计算机工程与设计, 2025, 46(4): 950-958.
JIN Jiang, WANG Xiaoping, ZANG Tiegang, et al. Robot obstacle avoidance path planning based on improved ant colony algorithm[J]. Computer Engineering and Design, 2025, 46(4): 950-958.
- [4] 张彪, 李永强. 基于动态寻优蚁群算法的移动机器人路径规划[J]. 仪器仪表学报, 2025, 46(3): 74-85.
ZHANG Biao, LI Yongqiang. Path planning of mobile robot based on the dynamic optimization ant colony algorithm[J]. Chinese Journal of Scientific Instrument, 2025, 46(3): 74-85.
- [5] 王欣, 邓玉娇, 常俊林. 矿井救灾机器人运动学分析及避障策略研究[J]. 煤矿机械, 2013, 34(2): 69-71.
WANG Xin, DENG Yujiao, CHANG Junlin. Kinematics analysis and obstacle avoidance strategy research of mine rescue robot[J]. Coal Mine Machinery, 2013, 34(2): 69-71.
- [6] 张辰, 范永, 李贻斌, 等. 人工智能在煤矿机器人中的应用[J]. 中国煤炭, 2021, 47(1): 93-98.
ZHANG Chen, FAN Yong, LI Yibin, et al. Application of artificial intelligence in coal mine robots[J]. China Coal, 2021, 47(1): 93-98.
- [7] 巩固, 朱华. 基于目标识别与避障的煤矿救援机器人自主行走[J]. 南京理工大学学报, 2022, 46(1): 32-39.
GONG Gu, ZHU Hua. Autonomous walking of coal mine rescue robot based on target recognition and obstacle avoidance[J]. Journal of Nanjing University of Science and Technology, 2022, 46(1): 32-39.
- [8] 李芳威, 鲍久圣, 王陈, 等. 基于 LD 改进 Cartographer 建图算法的无人驾驶无轨胶轮车井下 SLAM 自主导航方法及试验[J]. 煤炭学报, 2024, 49(增刊 2): 1271-1284.
LI Fangwei, BAO Jiusheng, WANG Chen, et al. Unmanned trackless rubber wheeler based on LD

- improved Cartographer mapping algorithm underground SLAM autonomous navigation method and test[J]. Journal of China Coal Society, 2024, 49(S2): 1271-1284.
- [9] 张立亚, 李晨鑫, 刘斌, 等. 基于子图像分割映射点云空间的机器人避障算法[J]. 煤炭科学技术, 2024, 52(增刊 2): 368-374.
- ZHANG Liya, LI Chenxin, LIU Bin, et al. Obstacle avoidance algorithm based on sub-image segmentation and mapping point cloud space[J]. Coal Science and Technology, 2024, 52(S2): 368-374.
- [10] 宋秦中, 胡华亮. 基于 CNN 算法的井下无人驾驶无胶轮车避障方法[J]. 金属矿山, 2023(10): 168-174.
- SONG Qinzhong, HU Hualiang. Obstacle avoidance method for underground unmanned trackless rubber-tired vehicle based on CNN algorithm[J]. Metal Mine, 2023(10): 168-174.
- [11] 郭爱军, 杨腾, 潘子宇. 动态环境下无人矿车速度规划与避障方法[J]. 矿业研究与开发, 2024, 44(7): 239-245.
- GUO Aijun, YANG Teng, PAN Ziyu. Speed planning and obstacle avoidance method for unmanned mining vehicle in dynamic environment[J]. Mining Research and Development, 2024, 44(7): 239-245.
- [12] 张可琨, 鲍久圣, 艾俊伟, 等. 基于改进 A* 与 DWA 算法的井下搬运机器人自主行走路径规划[J]. 煤炭科学技术, 2024, 52(11): 197-213.
- ZHANG Kekun, BAO Jiusheng, AI Junwei, et al. Autonomous walking path planning of underground handling robot based on improved A* and DWA algorithm[J]. Coal Science and Technology, 2024, 52(11): 197-213.
- [13] YANG Hongxia, TENG Xingqiang. Mobile robot path planning based on enhanced dynamic window approach and improved a algorithm[J]. Journal of Robotics, 2022. DOI: [10.1155/2022/2183229](https://doi.org/10.1155/2022/2183229).
- [14] XU Zhenyang, YUAN Wei. Mobile robot path planning based on fusion of improved A* algorithm and adaptive DWA algorithm[J]. Journal of Physics: Conference Series, 2022, 2330(1). DOI: [10.1088/1742-6596/2330/1/012003](https://doi.org/10.1088/1742-6596/2330/1/012003).
- [15] 彭继国, 张波, 孙凌飞, 等. 井下移动机器人智能视觉避障研究[J]. 工矿自动化, 2020, 46(9): 51-56, 63.
- PENG Jiguo, ZHANG Bo, SUN Lingfei, et al. Research on intelligent visual obstacle avoidance of underground mobile robot[J]. Industry and Mine Automation, 2020, 46(9): 51-56, 63.
- [16] 王利民, 孙瑞峰, 翟国栋, 等. 融合改进 A* 算法与动态窗口法的煤矿足式机器人路径规划[J]. 工矿自动化, 2024, 50(6): 112-119.
- WANG Limin, SUN Ruifeng, ZHAI Guodong, et al. Path planning of coal mine foot robot by integrating improved A* algorithm and dynamic window approach[J]. Journal of Mine Automation, 2024, 50(6): 112-119.
- [17] 鲁志, 刘莹煌, 张绪坤, 等. 融合 A* 与 DWA 算法的移动机器人动态避障研究[J]. 电子测量技术, 2025, 48(8): 34-45.
- LU Zhi, LIU Yinghuang, ZHANG Xukun, et al. Research on mobile robot dynamic obstacle avoidance by fusing A* and DWA algorithms[J]. Electronic Measurement Technology, 2025, 48(8): 34-45.
- [18] HARLOW H F. Learning and satiation of response in intrinsically motivated complex puzzle performance by monkeys[J]. Journal of Comparative and Physiological Psychology, 1950, 43(4): 289-294.
- [19] 李福进, 张俊琴, 任红格. 基于仿生学内在动机的 Q 学习算法移动机器人路径规划研究[J]. 现代电子技术, 2019, 42(17): 133-137.
- LI Fujin, ZHANG Junqin, REN Hongge. Research on mobile robot path planning by Q-learning algorithm based on bionics intrinsic motivation[J]. Modern Electronics Technique, 2019, 42(17): 133-137.
- [20] 阮晓钢, 张家辉, 黄静, 等. 一种结合内在动机理论的移动机器人环境认知模型[J]. 控制与决策, 2021, 36(9): 2211-2217.
- RUAN Xiaogang, ZHANG Jiahui, HUANG Jing, et al. An environment cognition model combined with intrinsic motivation for mobile robots[J]. Control and Decision, 2021, 36(9): 2211-2217.
- [21] 曾俊杰, 秦龙, 徐浩添, 等. 基于内在动机的深度强化学习探索方法综述[J]. 计算机研究与发展, 2023, 60(10): 2359-2382.
- ZENG Junjie, QIN Long, XU Haotian, et al. Exploration approaches in deep reinforcement learning based on intrinsic motivation: a review[J]. Journal of Computer Research and Development, 2023, 60(10): 2359-2382.