

DOI:10.19651/j.cnki.emt.2416152

融合动态窗口法和 DDPG 的无人船避碰方法*

韩志强¹ 李立刚^{1,3} 贺则昊² 宗律¹ 戴永寿¹

(1. 中国石油大学(华东)海洋与空间信息学院 青岛 266580; 2. 中国石油大学(华东)控制科学与工程学院 青岛 266580;
3. 自然资源部海上丝路海洋资源环境组网观测技术创新中心 青岛 266580)

摘要: 针对复杂环境中障碍物运动状态多变导致无人船避碰难以兼顾安全性、经济性和平滑性的问题,提出了一种融合动态窗口法和 DDPG 的无人船避碰方法。首先,在传统碰撞危险度模型中增加最近会遇距离和最近会遇点舷角作为评价因素,使无人船的碰撞危险度评价更为合理;其次,设计了基于动态窗口法的局部引导方法,将动态窗口法搜索的无人船可达位置作为局部引导点,增加了引导奖励来增大靠近引导点的动作奖励,使 DDPG 算法在训练中获得更准确的更新方向;最后,开展了在多种障碍物环境中的无人船避碰测试,实验结果显示,相较于传统 DDPG 算法,本文方法规划的路径更合理、平滑、风险更低,同时收敛速度提高了约 37.5%,验证了所提方法的有效性。

关键词: 无人船;碰撞危险度;最近会遇点舷角;局部引导方法;DDPG

中图分类号: U664.82;TN966 **文献标识码:** A **国家标准学科分类代码:** 510.80

USV collision avoidance method combining DWA and DDPG algorithm

Han Zhiqiang¹ Li Ligang^{1,3} He Zehao² Zong Lyu¹ Dai Yongshou¹

(1. College of Oceanography and Space Informatics, Chinese University of Petroleum(East China), Qingdao 266580, China;

2. College of Control Science and Engineering, Chinese University of Petroleum(East China), Qingdao 266580, China;

3. Technology Innovation Center for Maritime Silk Road Marine Resources and Environment Networked Observation, Qingdao 266580, China)

Abstract: To address the challenge of ensuring safety, efficiency, and smoothness in collision avoidance decisions for unmanned surface vessel in complex environments with dynamically changing obstacles, we propose a collision avoidance method that combines the dynamic window approach and DDPG algorithm. Firstly, in the traditional collision risk model, the distance to closest point of approach and the bearing angle at the closest point of encounter are added as evaluation factors to make the risk evaluation of the unmanned surface vessel more reasonable. Next, we design a local guidance method based on dynamic window approach, the reachable position of unmanned surface vessel by dynamic window approach is used as the local guide point, and the guide reward is added to increase the action reward near the guide point, so that DDPG algorithm can obtain more accurate updating direction in training. Finally, the method is tested in various obstacle environments. Experimental results show that compared to the traditional DDPG algorithm, the proposed method generates more reasonable, smoother and less risky paths. Additionally, it improves convergence speed by approximately 37.5%, verifying the effectiveness of the proposed method.

Keywords: USV; collision risk index; CPA bulwark angle; local guidance method; DDPG

0 引言

无人船(unmanned surface vessel, USV)作为一种新型海上运载平台,因其具有工作时长、维护成本低、响应速度快的优点,在海上勘探、搜救、巡逻等领域具有广阔的应用前景^[1]。

为确保无人船在复杂水域中安全航行,张金泽等^[2]将

模糊控制与动态窗口法(dynamic window approach, DWA)结合,根据无人船航行环境动态调整评价函数权重,提高了避碰规划的效率。杨敏豪等^[3]结合了间隙分析法和碰撞预测改进了传统 DWA 算法,提高了算法对动态障碍物环境的适应能力。Huang 等^[4]提出了基于船舶领域模型的人工势场法,该方法在计算受力时考虑了两船长度差异的影响,并进一步结合船舶操纵特性对船舶避碰进行约束。丁志国

收稿日期:2024-05-31

* 基金项目:中国石油大学(华东)自主创新科研计划项目(22CX01004A)资助

等^[5]通过结合避碰规则和驾驶实践经验,细化了会遇场景和避让责任,并将其融入船舶危险度评价和避碰决策中,提高了避碰算法面对不同会遇态势时的适应能力。赵文涛等^[6]提出了一种改进的快速行进平方法,通过构建符合避碰规则的动态障碍势场并引入局部风势场防止船舶陷入航行死区,提高了算法在复杂环境下避碰的安全性和经济性。张荣翔等^[7]提出了一种基于种群更新成功率的海洋捕食者算法,利用种群更新成功率代替种群迭代数作为选择搜索策略的标准,有效减少了算法陷入局部最优的问题。Liu 等^[8]为了使避碰算法更适用于复杂的海洋环境,提出了一种结合蚁群算法(ant colony optimization, ACO)和聚类算法的无人船避碰规划方法。这种基于 ACO 和聚类的两种智能算法的动态路径规划算法,具有收敛时间短、计算负载均衡、综合性能指标好的特点。Kim 等^[9]提出了一种改进的速度障碍法算法,采用碰撞时间和碰撞距离作为影响因素并引入到评价函数中寻找最优速度。但以上传统方法进行避碰时,避碰决策的相关参数通常要根据环境不同进行手动调整,这会导致在动态环境中方法的自适应能力较差。

针对传统避碰算法存在的问题,许多研究者将深度学习和强化学习技术应用到避碰规划领域,通过较强的自学能力来提高避碰算法对不同环境的适应能力。李鹏程等^[10]提出了在传统深度 Q 网络(deep Q-network, DQN)算法的基础上,融合了航海优先级和优先级经验回放策略,改进了算法学习过程中优秀经验的利用频率。Dong 等^[11]提出了一种基于贪心算法和深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法的避碰方法,该方法根据训练进程动态调整探索因子,更好地解决了探索和利用失衡的问题,提高了算法的学习效率。Li 等^[12]利用人工势场法改进 DQN 算法的奖励函数,使无人船训练获得更全面的反馈。张永梅等^[13]提出基于好奇心驱动的 DDPG 算法,利用好奇心模块生成预测状态和实际状态的误差奖励,一定程度上解决了奖励稀疏的问题。Wang 等^[14]将 DDPG 算法与免疫算法结合,通过对经验池中的样本进行免疫优化,提高了优秀样本的利用率。邓修明等^[15]提出了一种改进的 DQN 算法,该方法利用子调节贪婪因子调整探索与利用的概率,并根据人工势场法设计了相应的奖励函数,解决了算法训练中的探索利用失衡的问题,提高了训练效率。Duan 等^[16]提出了一种基于自适应模糊奖励的 DQN 算法,通过建立模糊逻辑动态调整奖励标准,使评价结果更符合实际的会遇态势,提高了无人船在不同场景下避碰的适应能力。Li 等^[17]将双重延迟深度确定性策略梯度(twin delayed deep deterministic policy gradient, TD3)算法与优先级经验回放策略结合,并采用动态延迟更新方式减少估计误差对策略网络的影响,该方法能够有效提高算法的训练效率和避碰成功率。以上方法虽在一定程度上缓解了强化学习训练效率低、学习效果差的问题,但仍

未考虑动作空间连续时会存在不同规划奖励相近,算法难以在训练中分辨优秀动作经验的问题,这会导致算法学习效率低且往往难以学习到最优规划策略。

针对以上问题,本文提出了一种融合 DDPG 和 DWA 算法的无人船避碰方法。首先,为更合理地评价无人船的危险程度,引入最近会遇距离和最近会遇点的舷角改进传统碰撞危险度评价模型。其次,基于 DWA 算法设计了局部引导方法,通过增大动作之间的奖励区分度,使算法在训练中能更有效地区分优秀动作并确定避碰学习的方向。最后在多种障碍物环境下对无人船进行训练并开展了实船实验,验证了改进方法的有效性。

1 碰撞危险度模型构建

在传统避碰中,采取避碰措施多以无人船与障碍物的距离为依据,这虽然可以使无人船与障碍物保持一定的距离,但也降低了无人船对航行危险的分析能力,不适用于多障碍物环境。为更合理地评价无人船的碰撞危险,本文在传统碰撞危险度模型的基础上,添加了最近会遇距离(distance to closest point of approach, DCPA)和最近会遇点舷角作为危险度评价因素,并结合海上避碰经验,从避碰时机选取、避碰措施幅度等方面完善传统评价标准。各评价因素的评价函数设计如下所示。

1.1 基于 DCPA 的碰撞危险度

$$R_{DCPA} = \frac{d_2 - DCPA}{d_2} \quad (1)$$

式中: d_2 表示无人船的观测范围。DCPA 是评价障碍物碰撞危险的主要指标之一,是无人船与障碍物在最紧迫时刻的相遇距离,能帮助无人船更合理、有效地评价障碍物对无人船产生的危险。

1.2 基于相对距离 D 的碰撞危险度

$$R_D = \frac{d_2 - D}{d_2} \quad (2)$$

相较于 DCPA 而言,无人船与障碍物的距离则是反映当前时刻的紧迫情况,该指标能使无人船注重与自身距离更近的障碍物带来的碰撞危险。

1.3 基于最近会遇点舷角 β 的碰撞危险度

无人船在面对不同会遇态势时,其避碰难度也是不同的。根据海上避碰经验,当面对正前方障碍物时,无人船往往需要采取更早的避碰时机,并做出较多的转向调整以规避危险。当障碍物从后方和侧方接近时,则仅需要小幅度的调整船速和航向即可避免大多数碰撞的发生,因此,仅依靠距离信息难以准确分析相障碍物的实际碰撞危险。为使评价模型能更有效地判断碰撞危险,本文引入最近会遇点舷角作为评价指标,将无人船与障碍物最紧迫时刻的相对方位划分为多个区域^[18],如图 1 所示。

在其他因素相同的情况下,基于最近会遇点舷角的碰撞危险度如表 1 所示。

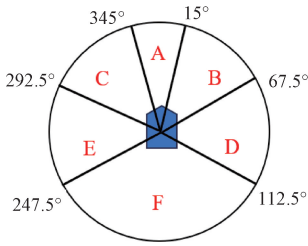


图 1 最近会遇点区域划分图
Fig. 1 Regional division of CPA

表 1 基于最近会遇点舷角的危险度评价

Table 1 Risk assessment of CPA bulwark angle

方位区域	A	BC	DE	F
碰撞危险度	1	0.8	0.6	0.5

1.4 碰撞危险度评价模型

在无人船碰撞危险度评价中,相较于距离 D,DCPA 和 CPA 的舷角能更直观地反映出无人船不采取避碰措施的后果,因此在评价中给予其评价函数更大的权重。此外,当无人船运动状态改变时,DCPA 和 CPA 舷角均会发生较为明显的变化,为更有效地体现会遇态势的危险变化,本文将两者的评价结果进行组合。

综上所述,构建无人船的碰撞危险度评价模型为:

$$CRI = 0.7R_{DCPA} \cdot R_{\beta} + 0.3R_D \quad (3)$$

若无人船未通过与障碍物的最近会遇点,则计算该障碍物对无人船产生的危险。该模型能够快速反映碰撞危险度变化的同时,通过相对距离的评价函数反映出无人船与障碍物的实际紧迫程度。

2 融合 DWA 和 DDPG 算法的无人船避碰方法

2.1 问题分析和算法基本原理

在传统 DDPG 算法的避碰规划中,为更好地区分不同决策产生的影响,奖励函数通常设计为连续函数。这种方式会导致不同规划获得的奖励相近,使算法在训练过程中难以区分最优规划经验并确定合适的更新方向。

为增大不同规划的奖励区分度,使算法训练获得更为准确的更新方向,本文设计了一种基于 DWA 算法的局部引导方法。首先,利用 DWA 算法搜索无人船在不同状况下可达的合适位置作为局部引导点;接着,根据无人船的航速和航行时间设置引导区域的范围,判断无人船是否进入引导区域;最后,通过增加引导奖励函数,对 DDPG 算法做出的规划进行评价。该方法通过增大到达引导区域内的规划奖励,使算法能更有效地确定避碰学习的主要方向,进而提高算法的训练效率和收敛效果。DWA-DDPG 算法基本原理如图 2 所示。

2.2 DWA-DDPG 算法

DWA-DDPG 的设计包括四部分:状态空间、动作空间、基于 DWA 的局部引导方法以及奖励函数。

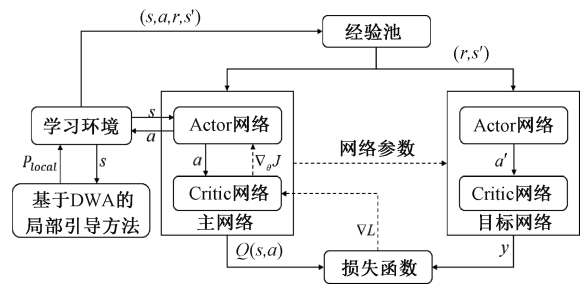


图 2 DWA-DDPG 算法基本原理图

Fig. 2 DWA-DDPG algorithm training flowchart

1) 状态空间设计

状态空间 s 是无人船进行决策所需状态信息的集合,包括无人船自身的状态信息 s_u 、障碍物状态信息 s_e 和目标点的状态信息 s_g 。其中 s_u 包括无人船的位置 P_u 、速度 V_u 、航向 ψ_u 和偏航角 δ_u 。即:

$$s_u = (P_u, V_u, \psi_u, \delta_u) \quad (4)$$

s_e 是周边障碍物信息的集合,由障碍物的位置 P_{e-i} 、航向 ψ_{e-i} 以及障碍物与无人船的相对速度 V_{e-i} 组成,其中 i 表示第 i 个障碍物。

$$s_e = ((P_{e-1}, V_{e-1}, \psi_{e-1}), \dots, (P_{e-i}, V_{e-i}, \psi_{e-i})) \quad (5)$$

s_g 包括目标点的坐标 P_g :

$$s_g = (P_g) \quad (6)$$

综上,算法的状态空间设计为:

$$s = s_u \cup s_e \cup s_g \quad (7)$$

2) 动作空间设计

动作空间 a 包含了无人船在规划中需要控制的运动参数。本文设定无人船的速度为恒定值,选择对无人船的转向幅度 τ 进行控制。

$$a = (\tau) \quad (8)$$

为了确保执行动作的可行性,规定无人船在每个状态下转向幅度的范围为 $[-10^\circ, 10^\circ]$ 。

3) 基于 DWA 的局部引导方法

动态窗口法的原理是根据无人船运动特性,在速度限制范围内采样多组航行速度和角速度组合,并利用一定的评价标准选择最合适的速度组合^[19]。

本文中将动作空间的范围作为采样范围,首先由 DWA 算法计算出合适的转向动作并计算无人船执行动作后到达的位置,将该位置作为局部引导点;接着计算无人船执行 DDPG 算法后到达的位置,计算公式如下:

$$x_{local} = x_u + V_u \cdot \cos(\psi_u + \Delta\psi) \cdot \Delta t \quad (9)$$

$$y_{local} = y_u + V_u \cdot \sin(\psi_u + \Delta\psi) \cdot \Delta t \quad (10)$$

$$x_{u'} = x_u + V_u \cdot \cos(\psi_u + \tau) \cdot \Delta t \quad (11)$$

$$y_{u'} = y_u + V_u \cdot \sin(\psi_u + \tau) \cdot \Delta t \quad (12)$$

式中: x_u 和 y_u 为无人船的当前位置, $\Delta\psi$ 为 DWA 算法做出的转向决策, Δt 为执行决策后无人船航行的时间, $x_{u'}$ 和 $y_{u'}$ 为无人船执行 DDPG 算法决策后到达的实际位置, τ 为 DDPG 算法做出的决策, x_{local} 和 y_{local} 为局部引导点的坐标。

最后根据无人船航速和航行时间动态确定引导范围,判断无人船执行 DDPG 算法的决策后是否进入引导范围内,若进入则给予显著的正向引导奖励,否则给予惩罚。

$$d_{local} = \sqrt{(x_{u'} - x_{local})^2 + (y_{u'} - y_{local})^2} \quad (13)$$

$$d_{guide} = 0.3 \cdot V_u \cdot \Delta t \quad (14)$$

式中: d_{local} 为无人船实际位置与局部引导点的距离, d_{guide} 为引导范围。

4) 奖励函数的改进设计

奖励函数是强化学习算法至关重要的一部分,是影响无人船避碰训练效果的关键因素,反映了执行动作后对环境的影响,奖励函数的设计会直接影响算法的学习效率和学习效果。

本文将奖励分为终止奖励和过程奖励。终止奖励包括无人船到达目标点附近或与障碍物发生避碰获得的奖励与惩罚。考虑到到达目标点和避开障碍物是无人船的首要任务,在目标点附近应设计显著的正向奖励以吸引无人船到达,发生避碰则给予该路径较大的惩罚。终止奖励计算公式如下:

$$R_{term} = \begin{cases} 500, & D_{ug} \leq \epsilon_r \\ -1\ 000, & D_{uo} \leq \epsilon_o \\ 0, & \text{其他} \end{cases} \quad (15)$$

其中, D_{ug} 是无人船与目标点的距离, ϵ_r 是到达目标点附近的距离阈值, D_{uo} 是无人船与障碍物的距离, ϵ_o 是发生碰撞的距离阈值。

过程奖励包括偏航奖励、避碰奖励和基于局部引导点的引导奖励。其中偏航奖励是由无人船偏航程度决定的,促使无人船向目标点前进,该奖励计算公式如下:

$$R_{head} = \begin{cases} 5 - \delta_u, & \delta_u \leq 5 \\ -20, & \text{其他} \end{cases} \quad (16)$$

避碰奖励由无人船前后时刻的碰撞危险度变化决定的,若规划后危险度变小则给予奖励,变大则施以惩罚,该奖励计算公式如下:

$$R_{collision} = (CRI_{t-1} - CRI_t) \cdot 20 \quad (17)$$

局部引导方法中的引导奖励由无人船与局部引导点的位置关系决定,该奖励旨在增大引导点附近规划的奖励,为算法参数更新提供更准确、合理的梯度方向,既促使了无人船快速学习 DWA 算法的规划方式,也使其在此基础上进一步探索更为合理的规划策略。

$$R_{DWA} = \begin{cases} 20, & d_{local} \leq d_{guide} \\ -20, & \text{其他} \end{cases} \quad (18)$$

综上所述,无人船在训练过程中的奖励函数为:

$$R = R_{term} + R_{head} + R_{collision} + R_{DWA} \quad (19)$$

3 实验分析

3.1 DDPG 模型参数

本文基于 Python 语言,在多种障碍物环境下进行了避碰模型的训练。模型关键参数如表 2 所示。

表 2 模型参数设置

Table 2 Model parameter settings

序号	参数	数值
1	学习率	0.001
2	隐藏层神经元个数	128
3	经验池容量	100 000
4	样本抽样个数	256
5	衰减系数	0.9
6	软更新率	0.01

3.2 无人船避碰算法对比实验

避碰实验对比了基于距离信息指导避碰的 DDPG 模型、基于碰撞危险度模型指导避碰的 DDPG、DWA-DDPG、TD3^[20] 以及 DQN 模型。障碍物场景设计为静态障碍物、动态障碍物和密集障碍物,地图中蓝色、绿色和紫色轨迹分别表示无人船、障碍物 1、障碍物 2 的航行路径,对应颜色的星形符号表示无人船与各障碍物航行的目标点,无人船速度设置为 0.7 m/s,设置无人船安全距离为 2.5 m,起点为(0,0),初始航向为 45°(以纵坐标为 0°),无人船的观测距离为 20 m,对障碍物采取避碰的碰撞危险度阈值为 0.6,障碍物参数如表 3 所示。

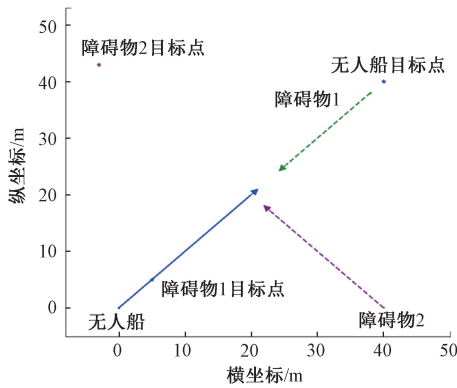
表 3 障碍物初始参数

Table 3 Initial parameters of obstacles

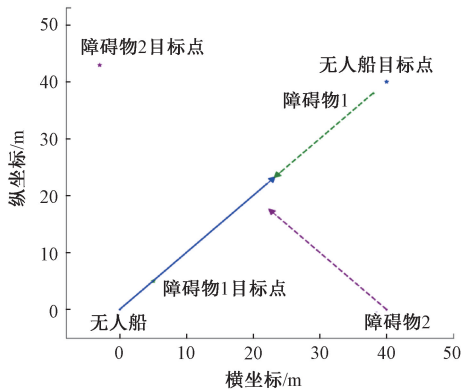
障碍物	起点	航向/ (°)	航速/ (m·s ⁻¹)	最近会遇 点航角/(°)
1	(38,38)	225	0.5	0.0
2	(40,0)	-45	0.6	130.6

首先利用基于距离信息的传统危险度评价方法和本文提出的评价方法对无人船航行过程中的碰撞危险度进行了评价。如图 3(c)、(d)所示,传统方法仅距障碍物较近才认为有较大的危险,这会导致无人船采取避碰措施时已距障碍物较近,不利于避碰的安全性。相较于传统方法,本文方法对无人船的碰撞危险变化更敏感,评价结果更合理。在后续避碰实验中能确保无人船安全的同时,使无人船在不同场景下规划的路径分别缩短了 10.14%、9.78%、12.50%。

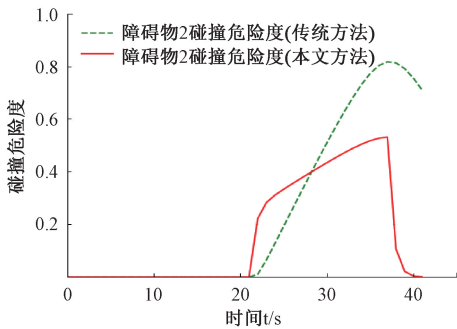
图 4 和表 4 展示了各强化学习算法在 3 种避碰测试环境中的表现,在图 4(a)中,DQN 算法规划的路径较长且与障碍物距离较近。DDPG 和 TD3 算法规划的路径较短且与 DWA-DDPG 算法的规划结果相近,但仍存在路径不平滑的情况。在图 4(b)中,各算法均能规划出较短的路线到达目标点,但 DQN 算法规划的路径平滑性较差。DDPG 和 TD3 算法一定程度上提高了路径的平滑性,但 DDPG 算法在规划初期做出的航向调整不合理,并非最优规划,TD3 算法也存在航向调整较早的问题,还需要进一步寻找合适



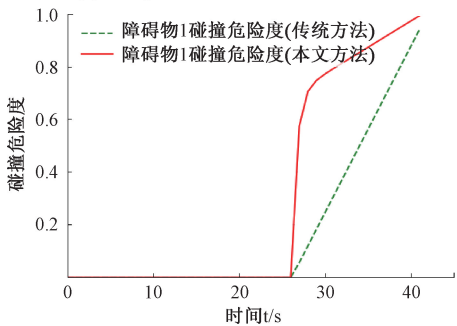
(a) 无人船与障碍物2的最近会遇点
(a) The CPA between USV and obstacle 2



(b) 无人船与障碍物1的最近会遇点
(b) The CPA between USV and obstacle 1



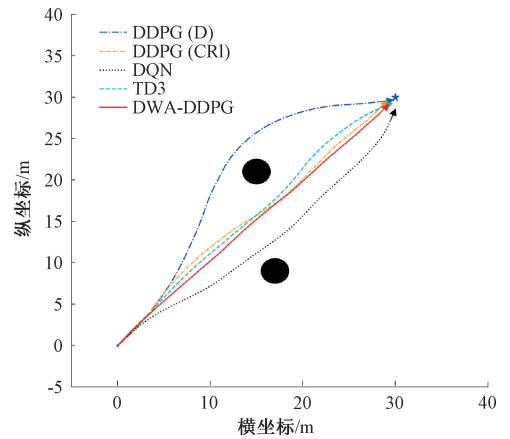
(c) 无人船与障碍物2的碰撞危险度变化
(c) Change in CRI between USV and obstacle 2



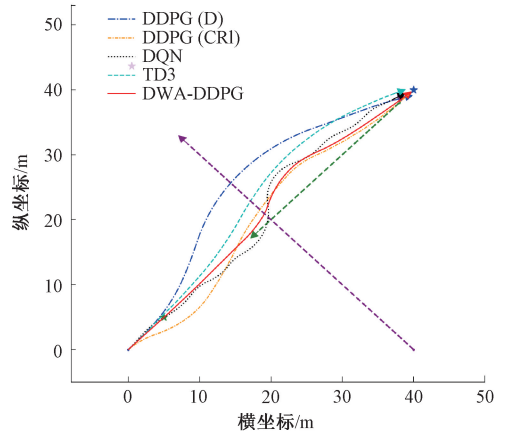
(d) 无人船与障碍物1的碰撞危险度变化
(d) Change in CRI between USV and obstacle 1

图 3 无人船航行轨迹和碰撞危险度变化

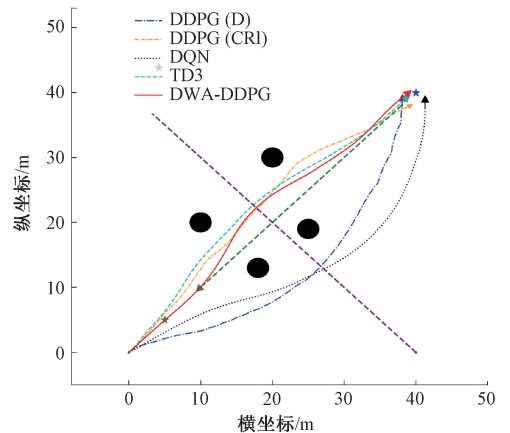
Fig. 3 Change in navigation trajectory and CRI of USV



(a) 静态障碍物场景下无人船的避碰路径
(a) Avoidance path of USV in static obstacle scenario



(b) 动态障碍物场景下无人船的避碰路径
(b) Avoidance path of USV in dynamic obstacle scenario



(c) 密集障碍物场景下无人船的避碰路径
(c) Avoidance path of USV in dense obstacle scenario

图 4 无人船避碰结果

Fig. 4 Collision avoidance results of USV

的避碰时机。相比之下,DWA-DDPG 算法规划的路径长度更短且采取避碰的时机更优。在图 4(c)中,由于 DQN 算法受限于动作空间的离散性,难以在多障碍物区域中采取合适的转向规划,导致无人船从障碍物区域外侧绕行,降低了路径规划的经济性。DDPG 和 TD3 算法虽然能规划

表 4 各算法避碰路径数据对比

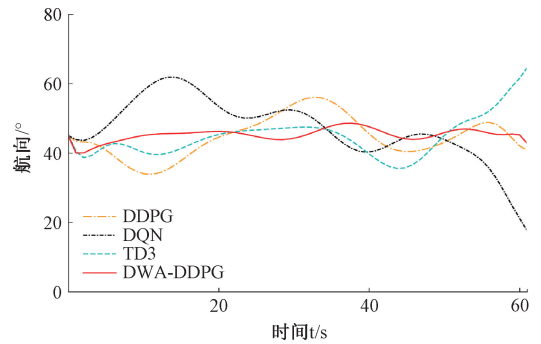
Table 4 Comparison of avoidance path data for various algorithms

场景	方法	路径	与障碍物
		长度/m	最近距离/m
静态障碍	DDPG(D)	48.3	4.24
	DDPG(CRI)	43.4	3.86
	DQN	45.5	3.62
	TD3	43.4	3.73
	DWA-DDPG	42.0	4.18
动态障碍	DDPG(D)	64.4	7.61
	DDPG(CRI)	58.1	3.89
	DQN	60.2	4.11
	TD3	61.6	5.23
	DWA-DDPG	57.4	4.61
密集障碍	DDPG(D)	67.2	5.25
	DDPG(CRI)	58.8	3.97
	DQN	70.7	5.16
	TD3	60.2	4.07
	DWA-DDPG	58.1	4.34

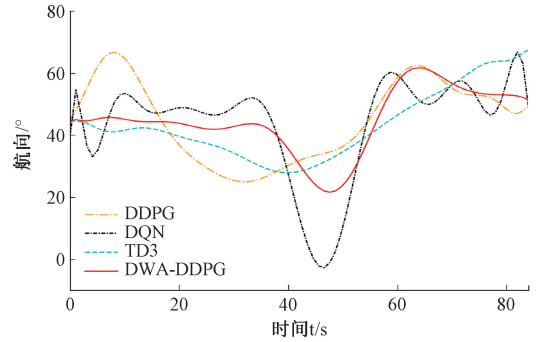
出较短的路径,但在路径平滑性和避碰时机的选取上仍存在不足,DWA-DDPG 算法在此场景中则仍能够较好地平衡安全性、经济性和平滑性的问题。综上所述,DWA-DDPG 算法在 3 个避碰测试场景中的表现均优于其他算法。

图 5 展示了各算法在规划过程中无人船的航向变化情况。由图 5(a)~(c)的无人船航向变化可知,DQN 算法对无人船航向调整的幅度较大,且在调整过程中存在较大的波动,导致路径的平滑性降低。相比之下,DDPG 和 TD3 算法能够在较小范围内调整无人船航向,但在规划过程中仍然存在航向波动的问题,增加了无人船航向调整的频率,不利于实际控制。与其他算法相对,DWA-DDPG 显著减少了无人船避碰规划过程中航向的波动,使得航向变化更加平稳,进一步提高了规划路径的平滑性。

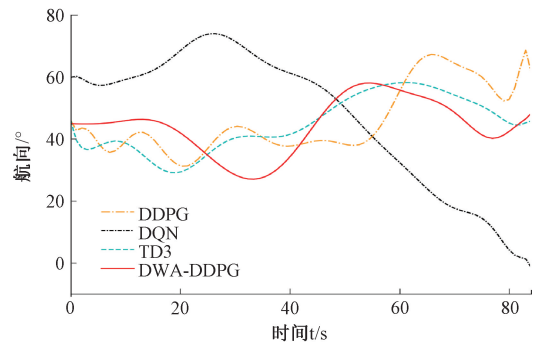
最后对比了各深度强化学习算法避碰训练的奖励值变化曲线。根据图 6 中训练的奖励曲线可知,DDPG 和 DQN 算法在约 2 400 回合时才逐渐到达目标点,在 3 000 回合后才能逐渐达到收敛,且在训练过程中的曲线波动较大。TD3 算法与 DWA-DDPG 算法在约 1 500 回合时便学习到了可以达到目标点的合适路径,学习效率相较于 DDPG 和 DQN 算法提高了约 37.5%。此外,与 TD3 算法相比,DWA-DDPG 算法能更准确地确定学习方向,TD3 算法由于受到奖励相近的经验影响,在训练后期仍在寻找更合适的策略,导致算法存在较为剧烈的震荡。该实验证明了 DWA-DDPG 算法能够更稳定且更高效地学习到合适的规划策略。



(a) 静态障碍物场景下无人船的航向变化
(a) Heading change of USV in static obstacle scenario



(b) 动态障碍物场景下无人船避碰的航向变化
(b) Heading change of USV in dynamic obstacle scenario



(c) 密集障碍物场景下无人船避碰的航向变化
(c) Heading change of USV in dense obstacle scenario

图 5 无人船航向变化对比图

Fig. 5 Comparison of USV heading change

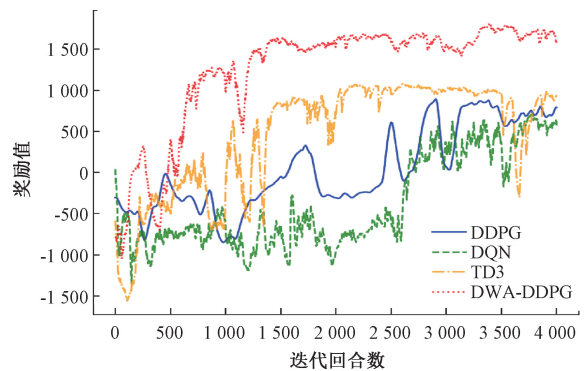


图 6 各算法训练的奖励值曲线图

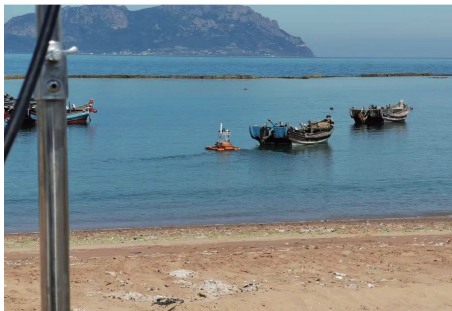
Fig. 6 Reward value curves for various algorithms

3.3 无人船避碰实验

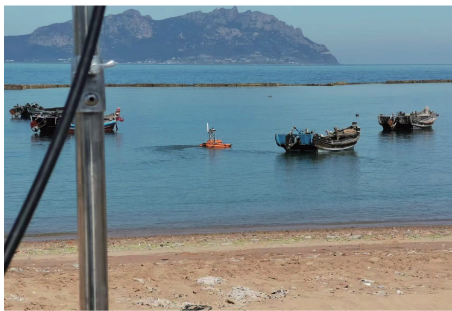
实验在古镇口海域开展,实验场地无明显海上风浪干扰,无人船避碰的障碍物为长约 5 m,宽约 2 m 的渔船。图 7 展示了无人船实际避碰效果。



(a) 无人船开始对障碍物避碰
(a) USV begins to avoid obstacles



(b) 无人船完成对障碍物的避碰
(b) USV completes collision avoidance of obstacles



(c) 无人船向目标点前进
(c) USV proceeds to the target point

图 7 无人船海上避碰结果

Fig. 7 Actual collision avoidance results of USV

根据无人船海上避碰的结果,本文避碰方法能够有效躲避障碍物,保障了无人船在海上航向的安全性,验证了方法实际应用的有效性。

4 结 论

本文提出一种融合 DWA 和 DDPG 算法的无人船避碰规划方法。该方法首先在传统危险度模型中增加了 DCPA 和最近会遇点舷角作为碰撞危险评价因素,提高了无人船对航行危险的分析能力。然后,设计了基于 DWA 算法的局部引导方法,通过为 DDPG 算法提供更为准确的避碰学

习方向,提高了算法的学习效率和效果。实验结果证明,本文方法在多种避碰测试环境中的规划表现均优于其他算法,且收敛速度更快。最后,开展了海上实船实验,验证了避碰方法的实际有效性,对无人船避碰技术的研究有一定的参考价值。下阶段工作将考虑在风浪流下如何保持无人船避碰的稳定性。

参考文献

- [1] 裴志远,戴永寿,李立刚,等. 无人船运动控制方法综述[J]. 海洋科学,2020,44(3):153-162.
PEI ZH Y, DAI Y SH, LI L G, et al. A review of motion control methods for unmanned ships [J]. Marine Science, 2020, 44(3): 153-162.
- [2] 张金泽,赵红,王宁,等. 密集障碍物下无人艇模糊双窗口 DWA 避障算法[J]. 中国舰船研究,2021,16(6): 10-18.
ZHANG J Z, ZHAO H, WANG N, et al. Fuzzy dual-window DWA obstacle avoidance algorithm for unmanned surface vehicles under dense obstacles[J]. Chinese Ship Research, 2021, 16(6): 10-18.
- [3] 杨敏豪,张国良,罗国攀,等. 融合间隙与 DWA 的移动机器人动态避障规划[J]. 国外电子测量技术,2023, 42(11):190-196.
YANG M H, ZHANG G L, LUO G P, et al. Dynamic obstacle avoidance planning for mobile robots by integrating gap and DWA[J]. Foreign Electronic Measurement Technology, 2023, 42(11): 190-196.
- [4] HUANG D J, LIU Z J, WANG X. Artificial potential field and ship maneuverability based collision avoidance path planning[C]. 2020 Chinese Automation Congress (CAC), IEEE, 2020: 2490-2495.
- [5] 丁志国,张新宇,王程博,等. 基于驾驶实践的无人船智能避碰决策方法[J]. 中国舰船研究,2021,16(1):96-104,113.
DING ZH G, ZHANG X Y, WANG CH B, et al. Intelligent collision avoidance decision-making method for unmanned ships based on driving practice [J]. Chinese Ship Research, 2021, 16(1): 96-104, 113.
- [6] 赵文涛,靖龙悦,孙朝阳. 基于改进快速行进平方方法的无人帆船动态避碰方法[J]. 中国舰船研究,2024,19: 1-14.
ZHAO W T, JING L Y, SUN CH Y. Dynamic collision avoidance method for unmanned sailboats based on an improved fast marching method [J]. Chinese Ship Research, 2024, 19: 1-14.
- [7] 张荣翔,戴永寿,李立刚,等. 基于漂角估计的无人船智能避碰方法[J]. 舰船电子工程,2023,43(12):65-70.
ZHANG R X, DAI Y SH, LI L G, et al. Intelligent collision avoidance method for unmanned ships based

- on drift angle estimation[J]. *Ship Electronic Engineering*, 2023, 43(12): 65-70.
- [8] LIU X Y, LI Y, ZHANG J, et al. Self-adaptive dynamic obstacle avoidance and path planning for USV under complex maritime environment[J]. *Ieee Access*, 2019, 7: 114945-114954.
- [9] KIM M, OH J H. Study on optimal velocity selection using velocity obstacle (OVVO) in dynamic and crowded environment[J]. *Autonomous Robots*, 2016, 40: 1459-1470.
- [10] 李鹏程,周远国,杨国卿. 基于深度Q网络的海上环境智能路径规划[J]. *电子测量技术*, 2024, 47(5): 77-84.
- LI P CH, ZHOU Y G, YANG G Q. Intelligent path planning in maritime environments based on deep Q-network [J]. *Electronic Measurement Technology*, 2024, 47(5): 77-84.
- [11] DONG Y S, ZOU X J. Mobile robot path planning based on improved DDPG reinforcement learning algorithm [C]. 2020 IEEE 11th International Conference on Software Engineering and Service Science(ICSESS). IEEE, 2020: 52-56.
- [12] LI L Y, WU D F, HUANG Y Q, et al. A path planning strategy unified with a COLREGS collision avoidance function based on deep reinforcement learning and artificial potential field[J]. *Applied Ocean Research*, 2021, 113: 102759.
- [13] 张永梅,赵家瑞,吴爱燕. 好奇心驱动的深度强化学习机器人路径规划算法[J]. *科学技术与工程*, 2022, 22(25): 11075-11083.
- ZHANG Y M, ZHAO J R, WU AI Y. Curiosity-driven deep reinforcement learning algorithm for robot path planning [J]. *Science Technology and Engineering*, 2022, 22(25): 11075-11083.
- [14] WANG Y, FANG Y L, LOU P, et al. Deep reinforcement learning based path planning for mobile robot in unknown environment [C]. *Journal of Physics: Conference Series*. IOP Publishing, 2020, 1576(1): 012009.
- [15] 邓修朋,崔建明,李敏,等. 深度强化学习在机器人路径规划中的应用[J]. *电子测量技术*, 2023, 46(6): 1-8.
- DENG X P, CUI J M, LI M, et al. Application of deep reinforcement learning in robot path planning[J]. *Electronic Measurement Technology*, 2023, 46(6): 1-8.
- [16] DUAN Z, WANG G, LIU Q, et al. USV path planning based on adaptive fuzzy reward [C]. *International Joint Conference on Rough Sets*. Cham: Springer Nature Switzerland, 2022: 117-131.
- [17] LI P, WANG Y, GAO Z. Path planning of mobile robot based on improved td3 algorithm[C]. 2022 IEEE International Conference on Mechatronics and Automation(ICMA). IEEE, 2022: 715-720.
- [18] 徐言民,张云雷,沈杰,等. 基于模糊集合理论的船舶碰撞危险度模型[J]. *舰船科学技术*, 2021, 43(4): 82-87.
- XU Y M, ZHANG Y L, SHEN J, et al. Collision risk model of ships based on fuzzy set theory[J]. *Ship Science and Technology*, 2021, 43(4): 82-87.
- [19] 刘淑霞,李立刚,金久才,等. 基于漂角估计的无人船局部动态避障方法[J]. *电光与控制*, 2023, 30(1): 103-108.
- LIU SH X, LI L G, JIN J C, et al. Local dynamic obstacle avoidance method for unmanned ships based on drift angle estimation [J]. *Opto-Electronic Engineering and Control*, 2023, 30(1): 103-108.
- [20] FUJIMOTO S, HOOF H, MEGER D. Addressing function approximation error in actor-critic methods[C]. *International Conference on Machine Learning*. PMLR, 2018: 1587-1596.

作者简介

韩志强, 硕士研究生, 主要研究方向为无人船路径规划与避碰。

E-mail: 15554422312@163.com

李立刚(通信作者), 副教授, 主要研究方向为无人船避障检测及控制、智能信息处理。

E-mail: upcllg@163.com