



洋流干扰下低速欠驱动AUV的三维路径规划

陈世同, 鲁子瑜

引用本文:

陈世同, 鲁子瑜. 洋流干扰下低速欠驱动AUV的三维路径规划[J]. 智能系统学报, 2025, 20(2): 425-434.

CHEN Shitong, LU Ziyu. 3D path planning for low-speed underdriven AUV under ocean current disturbance[J]. *CAAI Transactions on Intelligent Systems*, 2025, 20(2): 425-434.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202311004>

您可能感兴趣的其他文章

基于改进粒子滤波的AUV海底地形辅助定位方法

Seabed terrain-aided positioning method based on improved particle filtering for AUVs

智能系统学报. 2020, 15(3): 553-559 <https://dx.doi.org/10.11992/tis.201903027>

弹性网络核极限学习机的多标记学习算法

Multi-label learning algorithm of an elastic net kernel extreme learning machine

智能系统学报. 2019, 14(4): 831-842 <https://dx.doi.org/10.11992/tis.201806005>

采用相关滤波的水下海参目标跟踪

Underwater sea cucumber target tracking algorithm based on correlation filtering

智能系统学报. 2019, 14(3): 525-532 <https://dx.doi.org/10.11992/tis.201711037>

自主水面货船研究现状与展望

Research status and prospects of autonomous surface cargo ships

智能系统学报. 2019, 14(1): 57-70 <https://dx.doi.org/10.11992/tis.201806028>

事件驱动的强化学习多智能体编队控制

Event-triggered reinforcement learning formation control for multi-agent

智能系统学报. 2019, 14(1): 93-98 <https://dx.doi.org/10.11992/tis.201807010>

欠驱动船舶神经网络自适应路径跟踪控制

Neural network adaptive path tracking control for underactuated ships

智能系统学报. 2018, 13(2): 254-260 <https://dx.doi.org/10.11992/tis.201611011>

DOI: 10.11992/tis.202311004

网络出版地址: <https://link.cnki.net/urlid/23.1538.tp.20241223.1245.004>

洋流干扰下低速欠驱动 AUV 的三维路径规划

陈世同, 鲁子瑜

(哈尔滨工程大学 智能科学与工程学院, 黑龙江 哈尔滨 150001)

摘要: 海洋洋流对低速欠驱动水下运载器 (automatic underwater vehicle, AUV) 的航行影响巨大, 会增加航行时间、增大能源消耗以及改变航行轨迹, 故在洋流干扰的情况下规划出一条最优航行路线有着重要的意义。本文主要分析了洋流对 AUV 的影响机理, 由于传统的强化 Q 网络 (deep Q-network, DQN) 路径规划算法在复杂环境下存在过估计的问题, 提出了基于优先经验回放方法的改进 DQN 路径规划算法, 同时对动作设计和奖励函数进行优化。在基于 S57 海图数据建立的三维海洋环境下, 利用地球与空间研究机构 (earth & space research, ESR) 提供的洋流数据进行路径规划仿真。实验结果表明, 在充分考虑洋流干扰的情况下, 改进后的 DQN 算法能够规划出较优的全局路径规划, 提供一条时间最短且能耗最低的航行路线, 为 AUV 水下航行提供参考。

关键词: 自主水下运载器; 强化学习; 洋流干扰; 路径规划; 三维海洋环境; 强化 Q 网络; S57 海图; 奖励函数
中图分类号: TP242.6 **文献标志码:** A **文章编号:** 1673-4785(2025)02-0425-10

中文引用格式: 陈世同, 鲁子瑜. 洋流干扰下低速欠驱动 AUV 的三维路径规划 [J]. 智能系统学报, 2025, 20(2): 425-434.

英文引用格式: CHEN Shitong, LU Ziyu. 3D path planning for low-speed underdriven AUV under ocean current disturbance[J]. CAAI transactions on intelligent systems, 2025, 20(2): 425-434.

3D path planning for low-speed underdriven AUV under ocean current disturbance

CHEN Shitong, LU Ziyu

(College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin 150001, China)

Abstract: Ocean currents, which have a substantial impact on the navigation of low-speed, underdriven AUVs, can increase navigation time, raise energy consumption, and change the navigation trajectory. Therefore, planning an optimal navigation route that accounts for the disturbance of ocean currents is of considerable importance. This study mainly analyzes the mechanism by which ocean currents influence AUVs and proposes an improved DQN path planning algorithm based on the prioritized experience replay method. This modification addresses the problem of overestimation, which is a common issue when using a traditional DQN path planning algorithm. Additionally, the action design and reward functions are optimized. Path planning simulations are conducted in a 3D ocean environment, which is established based on S57 chart data and ocean current data provided by Earth & Space Research. Experimental results show that the improved DQN algorithm generates a more effective global path planning, offering a navigation route that minimizes time and energy consumption. This work provides a valuable reference for underwater AUV navigation, fully considering the impact of ocean current disturbances.

Keywords: automatic underwater vehicle; reinforcement learning; ocean current disturbance; path planning; 3D marine environment; deep Q-network; S57 charts; reward function

低速欠驱动水下运载器 (automatic underwater vehicle, AUV) 是一类为执行水下任务的自航载

体。其包括自主导航与控制、自主决策、自主作业等关键系统, 可以自主执行各种特殊任务^[1]。AUV 具有体积小、机动能力强、智能化程度高、作业风险低、低功耗、长续航等特点, 广泛地运用

收稿日期: 2023-11-02. 网络出版日期: 2024-12-24.

通信作者: 陈世同. E-mail: chenshitong@hrbeu.edu.cn.

©《智能系统学报》编辑部版权所有

于海洋环境探测、海洋安全保证等关键领域。路径规划作为自主式水下航行器的核心技术之一,是AUV执行任务时安全性、可靠性的重要支撑^[2]。

然而,海洋环境的复杂性和多变性对AUV的航行和路径规划提出了巨大的挑战。洋流作为海洋环境中主要的干扰源之一,会对AUV的航行产生显著的影响。洋流干扰引起AUV的航行问题包括路径偏移、能耗增加以及任务执行效率下降等,这些问题都限制了AUV在实际海洋环境中的应用^[3]。刘甲等^[4]提出了低速AUV在航渡过程中通过增加航路控制点的方法减小洋流的影响。Liu等^[5]提出了一种基于分布估计算法(estimation of distribution algorithm, EDA)的固定高度直方图方法来处理动态环境下AUV的路径规划问题。

由于目前AUV的路径规划常用自身控制应对洋流干扰的方法需要额外消耗更多的能源从而导致航行时间增大^[6],故本文在上述文献思想的基础之上,不使用当前常用的AUV自身控制应对洋流的方法,而是使用优先经验回放方法改进强化Q网络(deep Q-network, DQN)算法来解决应对洋流干扰情况下的全局路径规划的问题,从而规划出一条时间最短、能耗最低的AUV航行路线。

1 洋流对低速欠驱动AUV航行的影响

1.1 低速欠驱动AUV的特点

低速欠驱动AUV是使用最广的一类水下机器人,其特点是速度相对较慢且独立控制输入的数量小于系统位形空间的自由度^[7],使其更适合执行一些需要精确控制、细致观测且航行周期较长的任务^[8]。

本文低速欠驱动AUV的前推力来自于AUV尾部的螺旋桨,横摆力矩和俯仰力矩分别来源于AUV尾部的垂直方向舵和水平方向舵,在主流的欠驱动AUV中,这种配置最为常见。

1.2 洋流运动对低速欠驱动AUV的影响机理

洋流是指海洋中的水流运动,并在全球范围内形成了复杂的循环系统。太阳辐射是洋流形成的主要动力来源,地球的自转也会影响着洋流的形成,此外海底的地形也会深刻影响着洋流的流向^[9]。

一般情况下,水平方向的洋流在全球范围内速度可达每小时数十千米。而在垂直方向上的洋流是由海水的密度差异、温度差异或盐度差异所驱动,通常以每小时几厘米到每小时几分米的速度进行^[10]。故为了简化洋流的运动,垂直于水平面的方向的流速 ω 忽略不计^[11]。假设洋流的纬向速度(zonal velocity) u 和经向速度(meridional velocity) v 沿深度均匀分布,且忽略黏性力、离心力和非线性惯性力。将洋流对AUV的作用力分解为沿着AUV纵轴(x 轴)方向、横轴(z 轴)和垂直于AUV运动平面(y 轴)方向,如图1所示。洋流在垂直于水平面的方向的流速 ω 忽略不计,即忽略沿AUV的 y 轴方向的洋流作用力,故在分析洋流对AUV航行轨迹的影响主要从AUV的 x 轴方向和 z 轴方向进行考虑^[12]。

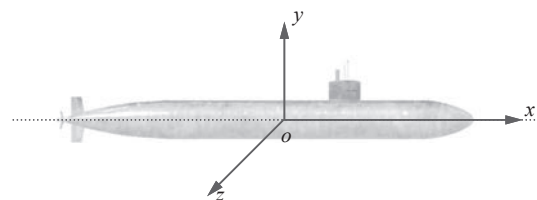


图1 低速欠驱动AUV的坐标分解

Fig. 1 Coordinate decomposition of low-speed underdriven AUV

图2从AUV上方俯视AUV,并进行坐标分解。由于洋流作用在低速欠驱动AUV上的流体动力,可以等效为AUV的 x 轴和 z 轴方向分别叠加洋流的加速度。其中 z 轴方向上的分解加速度则会影响AUV的航行轨迹,使得AUV无法沿着既定航线直线航行。

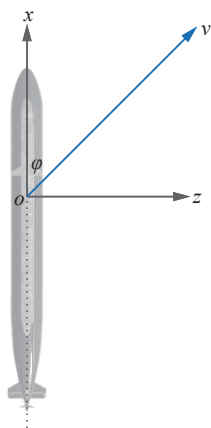


图2 在洋流影响下的AUV坐标分解

Fig. 2 Decomposition of AUV coordinates under the influence of ocean currents

如图2所示,设低速欠驱动AUV的航速为 v_0 ,航向为沿着 x 轴正方向直线航行, v_l 为洋流的速度, v_l 在图2中标注的方向为假定的洋流方向, v_l 与 x 轴的夹角为 φ ,AUV的航行轨迹的航迹方向为 η ,则存在以下关系:

沿航路 x 轴方向的速度分量为

$$v_x = v_0 + v_l \cdot \cos \varphi$$

沿 z 轴的速度分量为

$$v_z = v_l \cdot \sin \varphi$$

AUV 航迹的实际方向为

$$\eta = \arctan \frac{v_z}{v_x}$$

AUV 的实际航速为

$$v = \sqrt{v_x^2 + v_z^2}$$

通过对上式的分析,当洋流的流速 $v_l > 0$ 时,就会对 AUV 的航行产生影响。如图 3 所示, v 表示速度为 2 节 (knot, kn) 的洋流,方向为蓝色箭头方向。假设低速欠驱动 AUV 的航速为 5 kn,并沿着 v_0 所示方向航行。由于洋流的影响,低速欠驱动 AUV 实际上会沿着图 3 中 v_l 的方向航行,此时洋流对 AUV 的航向影响角 φ 为 21.80° 。郭兴海等^[13]指出 AUV 在进行路径规划时若不考虑洋流影响会导致计算结果与实际存在较大偏差。故低速欠驱动 AUV 在进行路径规划时需要考虑洋流的影响。

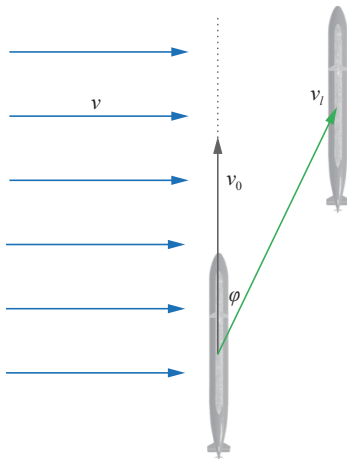


图 3 洋流对 AUV 的航行影响

Fig. 3 Impact of ocean currents on AUV navigation

所以,洋流对低速欠驱动 AUV 的影响主要分为 3 个主要方面:一是洋流会引起 AUV 的路径偏移,AUV 在航行的过程中会受到洋流的推动力影响而偏离预定的航线,使其无法沿着既定的规划航线进行航行;二是洋流的干扰会影响 AUV 的航行姿态和稳定性,尤其是对欠驱动的 AUV 影响更为明显,使其难以保持稳定的航行状态;三是洋流的干扰也会显著增加 AUV 的能耗,洋流对 AUV 的航行产生阻力,需要额外的能量来克服这种阻力^[14]。

2 洋流干扰下 AUV 路径规划算法

为了尽可能减少 AUV 在执行任务时受到的洋流干扰。本文首先建立对应任务海区的海底地

形模型;接着在地形模型上增加洋流模型并构建海洋动态三维模型;最后采用优先经验回放策略改进传统 DQN 算法后,对 AUV 的路径进行规划。

2.1 三维海洋环境及洋流模型构建

在 AUV 执行任务之前,需要先确定 AUV 执行任务的海域。首先获取任务海域的 S-57 电子海图^[15],图 4 为 S-57 数据交换结构。S-57 数据由一个或多个逻辑记录组成,每个逻辑记录由头标区、目次区和字段区构成。首条记录为数据描述记录 (data descriptive record, DDR),其后的各个记录为数据记录 (data record, DR)。记录的种类主要有以下 5 种:数据集描述、目录、数据字典、特征和空间信息。

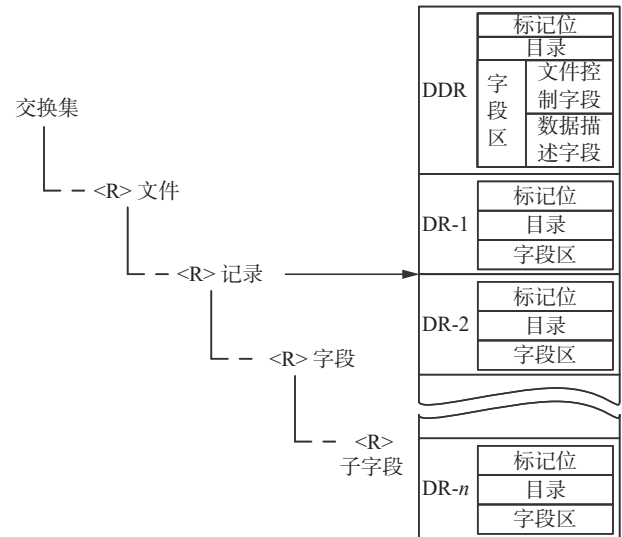
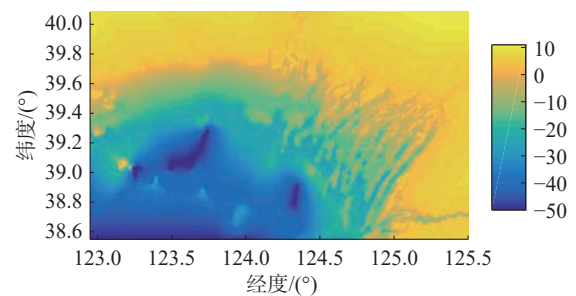


图 4 S-57 数据交换结构

Fig. 4 Structure of S-57 transfer file

在 S-57 电子海图的数据交换结构中,水深数据存储于 DEPART 字段中。在 DEPART 字段中包含 DRVAL1 子字段,其定义为深度范围最小值。

本文通过提取 DEPART 字段中含有最小水深数据的 DRVAL1 子字段。依据 DRVAL1 建模生成的任务海域的二维和三维海底地形如图 5 所示。



(a) 二维海底地形图

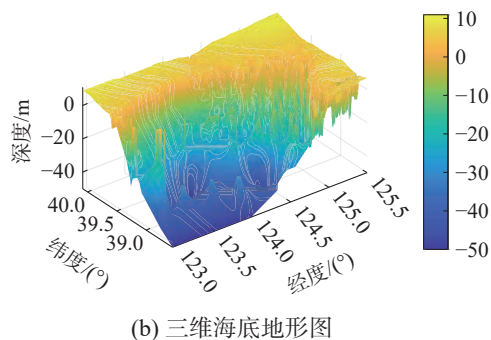


图5 某海域海底地形模型

Fig. 5 Topographic model of the seabed in a sea area

接着,在三维地形模型中构建动态洋流模型^[16],并考虑 AUV 执行任务的时间和总结当前海域的洋流规律。之后选择特定时刻的洋流数据绘制成区域洋流图^[17]。图 6 为图 5 海域的洋流情况(只考虑水平方向的洋流),图中蓝色箭头的方向表示洋流的方向,蓝色箭头的长短表示洋流的强弱,即箭头越长表示该区域的洋流强度越大,没有蓝色箭头的区域则为陆地部分。

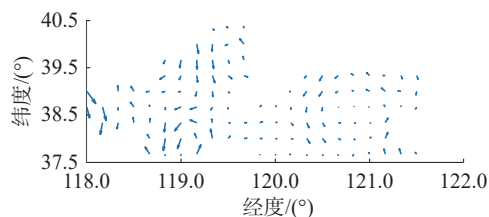


图6 指定海域的洋流情况

Fig. 6 Currents in the designated sea area

最后在图 6 的洋流图中提取图 5 海域的洋流数据,绘制成包含洋流的三维海洋环境图^[18],如图 7 所示。图中红色箭头的方向表示当前海域的洋流运动方向,箭头的长短表示洋流的相对大小,箭头上方的数值为当前洋流流速大小,单位为节。

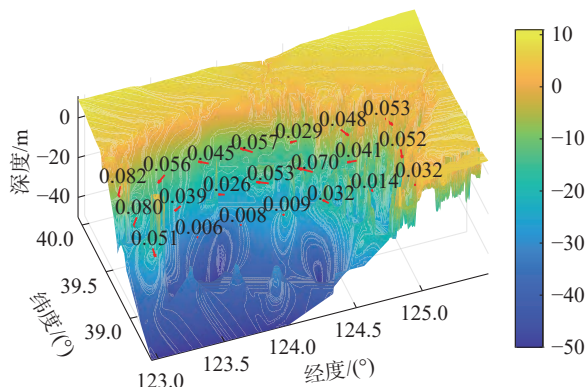


图7 含洋流的三维海洋环境

Fig. 7 3D marine environment with ocean currents

根据图 7 中洋流的运动及对应的海底地形情况,可以分析出该片海域的洋流总体上呈自东向

西运动,且海岸处的洋流大小平均为 0.05 kn 左右,而在东经 38.55°~东经 39.00°和北纬 123.5°~北纬 124.5°的海域洋流流速基本低于 0.01 kn。所以在该海域进行 AUV 路径规划的时候可以针对此海域洋流具有海岸处洋流流速比远离海岸处洋流流速更快的特点进行优化,从而达到减少 AUV 的航行时间和降低 AUV 能耗的目的^[19]。

2.2 在洋流干扰下的三维路径规划方法

DQN 算法是深度强化学习 (deep reinforcement learning, DRL) 的一种方法^[20],其基本原理结合了深度学习和强化学习,其核心是使用深度神经网络作为策略函数^[21-22]。Q 学习 (Q-learning) 算法构造了一个可用于深度学习的目标函数^[23]。DQN 算法可以使用价值函数来近似 Q 值^[24]。如图 8 所示, DQN 算法中包含 2 种结构相同但参数不同的网络,分别用于更新 Q 估计值和 Q 现实值^[25-26]。传统强化学习算法仅采用表格作为策略函数,在面对复杂环境时建立表格会变得十分困难^[27]。而 DQN 算法由于采用的具有强大拟合能力的深度神经网络,可以替代传统的由表格作为的策略函数,使得 DQN 算法在复杂场景中具有传统强化学习算法无可比拟的优势^[28-29]。所以在面对动态海洋环境进行路径规划决策时,使用 DQN 算法可以很好地解决复杂环境的问题^[30]。

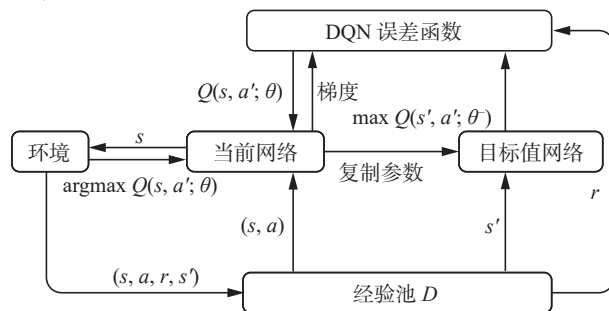


图8 DQN 算法

Fig. 8 DQN algorithm

DQN 算法使用均方差更新网络权重来定义目标函数,公式为

$$L(\omega) = E[r + \gamma \max_{a'} Q(s', a', \omega') - Q(s, a, \omega)]^2$$

本节使用优先经验回放方法在传统 DQN 算法 (算法 1) 的基础上进行改进,并分别设计智能体 (低速欠驱动 AUV) 的动作空间和奖励函数,使得智能体可以根据执行海域的地形及洋流特点规划出航行距离和航行时间都较优的航行轨迹。

算法 1 传统 DQN 算法

输入 智能体的当前状态 s 。

输出 训练好的智能体模型。

1) 将重放记忆 D 的容量大小初始化为 N

- 2) 用随机权重 θ 初始化动作函数 Q
- 3) 用权重 $\theta^- = \theta$ 初始化目标价值函数 Q^-
- 4) for episode=1, 2, ..., M do:
初始化序列 $s_1 = \{x_1\}$
预处理序列 $\varphi_1 = \varphi(s_1)$
- 5) for $t = 1, 2, \dots, T$ do:
6) 以概率 ε 选择一个随机动作
否则选择 $a_t = \arg \max Q(\varphi(s_t), a, \theta)$
- 7) 在环境中执行 a_t 并观察奖励函数 r_t 和 x_{t+1}
- 8) 设置 $s_{t+1} = s_t, a_t, x_{t+1}$ 和预处理 $\varphi_{t+1} = \varphi(s_{t+1})$
- 9) 在 D 中存储样本 $(\varphi_t, a_t, r_t, \varphi_{t+1})$
- 10) 在 D 中抽取样本 $(\varphi_j, a_j, r_j, \varphi_{j+1})$

$$y_j = r_j \text{ if } s \text{ is end}$$

$$y_j = r_j + \gamma \max_{a'} Q'(\varphi_{j+1}, a'; \theta^-)$$

$$y_j = r_j + \gamma Q(s'_t, \max_{a'} Q'(s'_t, a'; \theta); \theta^-)$$
- 11) 设置网络参数 θ 执行梯度下降步骤

$$(y_j - Q(\varphi_j, a_j; \theta))^2$$
- 12) 每 C 步重置 $Q' = Q$
- 13) end for
- 14) 更新智能体的当前状态 s_t
- 15) end for

上述过程会不断进行迭代, 直到达到最大迭代次数 M 后终止训练程序。

2.2.1 改进 DQN 算法

传统 DQN 算法中, 在计算时间差分目标 (temporal difference target, TD target) 时用到了最大化, 即式 (1), 造成了计算的结果大于真实值。所以在网络迭代的过程中, 这种最大化造成的过估计会随着迭代次数的增加而不断累积, 从而导致传统 DQN 算法在处理复杂环境下的路径规划时容易陷入过估计, 最终导致无法给出合适的航行轨迹^[31-32]。

为了解决上述过估计的问题并提高智能体的学习效率, 本文使用优先经验回放 (prioritized experience reply, PER) 策略改进传统 DQN 算法。PER 的优先级是基于经验的时间差误差 (temporal difference error, TD-error), 用以表示智能体在执行某个动作后所得的奖励与预期奖励之间的差异。TD-error 越大表明这个经验对智能体的学习起到正向作用。在经验回放时, 优先级高的经验会被智能体优先执行, 以便更快地学习到正向经验, 从而改进策略, 缩短迭代次数。

只有经验被重放后, PER 才会更新当前的 TD-error。当需要解决一些复杂的问题, 尤其是本文的三维海洋环境下的路径规划时, 为了确保 TD-error 为零时的经验也能够被智能体采样学习

到, 采用转移 i 的概率, 公式为

$$p(i) = \frac{p_i^\alpha}{\sum_k p_k^\alpha}$$

式中: p_i 表示第 i 个经验样本的优先级, α 则决定使用优先级的大小。 p_i 公式为

$$p_i = |\delta_i| + \varepsilon$$

式中: δ_i 表示 TD-error, ε 用以防止经验样本 TD-error 的值为 0。

为了消除采样带来的偏差, 引入重要性采样。其目的是为了确定 $f(x)$ 在 $p(x)$ 分布下的期望。为了使样本符合 $p(x)$ 分布, 还需要额外引入 $q(x)$ 分布, 最终的表达式为

$$E[f(x)] = \int_x f(x)p(x)dx = \int_x \omega(x)f(x)q(x)dx = \int_x g(x)q(x)dx \quad (2)$$

式中 $\omega(x)$ 表示重要度采样权重。将 $p(x) = 1/N$ 和 $q(x) = p(j)$ 代入式 (2) 后得到重要度采样权重函数, 公式为

$$\omega_i = \left(\frac{1}{N \cdot P(j)} \right)^\beta$$

式中 β 用来消除学习后的偏差。最后, 归一化重要度采样权重公式为

$$\omega_i = \frac{(N \cdot P(j))^{-\beta}}{\max(x_i) \cdot \omega_i}$$

在三维海洋环境下, 通过使用优先经验回放方法改进传统 DQN 算法训练智能体 (低速欠驱动 AUV) 的算法流程如图 9 所示。

2.2.2 动作设计

动作选择是指智能体根据策略选取下一步动作的过程, 是强化学习核心要素之一。智能体通过采取一定的策略选取下一步的动作。选择不同的动作来获取不同的奖励值, 在不同的环境状态选择回报越高的动作具有更高的价值, 高回报使得智能体在下次训练中更倾向于选择该动作^[33]。在本文中, 智能体能够采用 27 个方向的动作, 如图 10 所示。动作数值分别为 0、1、2、...、25、26, 对应的表示方向如图 10 中色块中的数值所示, 例如动作数值 16 表示对应的方向为沿 y 轴的正方向运动。需要额外指出的是, 动作数值 13 表示的是静止状态, 但在实际的路径规划过程中, 不存在静止状态, 所以动作数值 13 在本实验中是不存在的数值, 为了避免出现动作 13, 会在奖励函数中对出现动作 13 时的情况进行惩罚。所以在本文中, 智能体能够采用的动作方向为 26 个, 即动作数值 0 ~ 12、14 ~ 26。

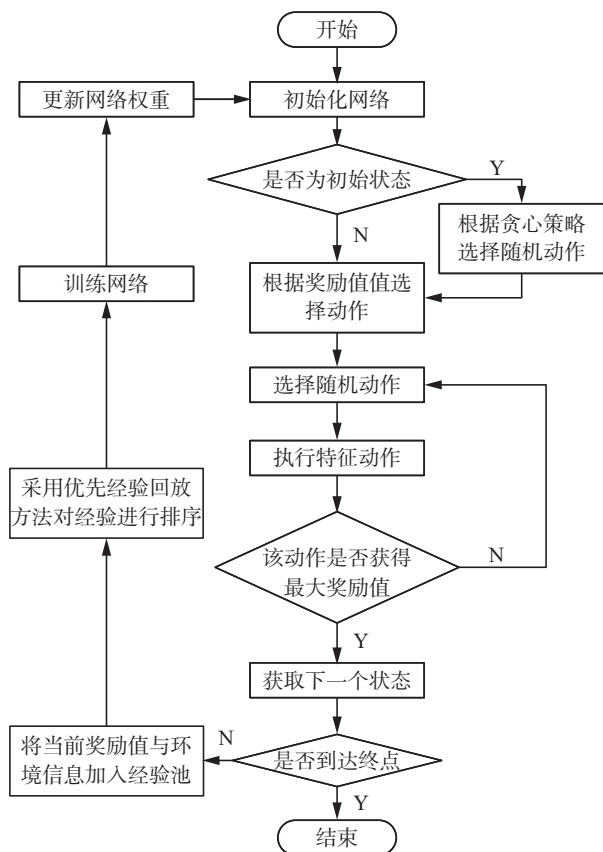


图 9 改进 DQN 算法模型流程

Fig. 9 Improved DQN algorithm model flow

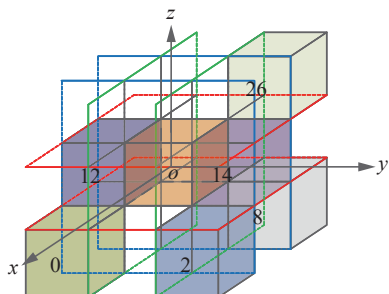


图 10 动作空间设计

Fig. 10 Motion space design

此外, 由于本文使用的是低速欠驱动 AUV, 而图 10 给出的智能体的动作方向设定是把 AUV 当作质点来处理的。这种方法忽略了低速欠驱动 AUV 本身的方向, 可能会出现绘制的轨迹在实际中 AUV 无法执行的情况。例如当前 AUV 航行的方向以图 10 中的 13 往 14 的方向且此时的位置坐标在 13, 如果下一步需要执行的动作数值为 0, 则是不可能发生的, 如果执行的动作数值为 2 则是正常的。所以为了避免出现与实际不符的动作数值, 会在下文中以奖励函数惩罚的方式进行避免。

2.2.3 奖励函数

奖励函数是影响智能体路径规划质量的关键因素, 也是强化学习的核心要素之一。通过设置

合适的奖励函数可以极大地提高智能体搜寻路径的正确率。奖励函数 R 在状态 S 下的 t 时刻到 $t+1$ 时刻的状态所能获得的期望奖励是 $R_t = E[R_{t+1}|S_t = s]$, 累积奖励函数值 G_t , 也就是总回报值。

$$G_t = R_{t+1} + \gamma R_{t+2} + \dots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}$$

式中 γ 是折扣系数, 表示未来时刻回报奖励的衰减效果, $\gamma \rightarrow 1$ 则衰减效果越小, 表示越有远见, 从而避免陷入局部最优。

本文中奖励函数涉及多个方面, 包括 AUV 路径规划的目标和约束、低速欠驱动 AUV 航行特点的约束以及任务海域中洋流速度与方向的约束等。

1) AUV 路径规划的目标和约束。AUV 路径规划的首要目标是到达任务预设的终点, 其次需要考虑的是执行的动作是否朝向目标点, 最后需要考虑 AUV 航行规划出的轨迹所需要的时间长短。

首先, 如果 AUV 在限定能量要求下到达终点, 则给予大量正向奖励值。如果 AUV 未能在能量耗尽前到达终点或撞到障碍物, 则给予大量惩罚。

其次, 为了让 AUV 在任意状态下都倾向于选择朝向目标点的动作, 则需要对 AUV 选择朝向目标点运动的动作给予正奖励, 反之则给予相应的惩罚。则 AUV 朝向目标点的奖励函数为

$$R_{k2} = \frac{r_0}{r} \cdot [\sqrt{x^2 + y^2 + z^2} - \sqrt{x'^2 + y'^2 + z'^2}]$$

式中: R_{k2} 为 AUV 执行动作的奖励函数, r_0 为初始状态下 AUV 起点到达目标点的直线距离, r 为当前状态距离终点的直线距离, $\sqrt{x^2 + y^2 + z^2}$ 表示的是前一个状态 AUV 距终点的距离, $\sqrt{x'^2 + y'^2 + z'^2}$ 表示的是执行当前动作之后的距终点的距离。 r_0/r 用于反映 AUV 距目标点的远近程度, AUV 距目标越近, 则 AUV 获得的奖励值越高。 $\sqrt{x^2 + y^2 + z^2} - \sqrt{x'^2 + y'^2 + z'^2}$ 反映的是如果 AUV 当前状态与前一个状态相比是远离目标, 则结果为负, 即给予惩罚; 当 AUV 当前状态与上个状态相比是接近目标, 则结果为正, 即获得正向奖励。

最后, 需要计算 AUV 执行每步动作所需要的时间。由于 AUV 执行的动作方向有 26 个, 相应的只有 3 个不同的单位长度, 即 1、 $\sqrt{2}$ 和 $\sqrt{3}$ 。再根据 AUV 当前速度与洋流的合速度计算当前 AUV 的实际速度, 即可获得每步 AUV 所需的航行时间, 并给予相应的奖励或惩罚。

2) 低速欠驱动 AUV 航行约束。虽然本文在

使用改进 DQN 算法对 AUV 进行路径规划时是把 AUV 当作质点处理的,但是实际情况下是不可以的。在对低速欠驱动 AUV 进行路径规划时,需要尽量避免 AUV 频繁地进行短距离大角度的改变航向。例如,如图 11 所示,假设当前 AUV 的航行方向为 v ,则在下一步的动作数值选择中,若选择 2、5、8、11、14、17、20、23、26 这 9 个动作数值时给予奖励,选择其他动作则给予惩罚。从而避免 AUV 执行不合理的动作。

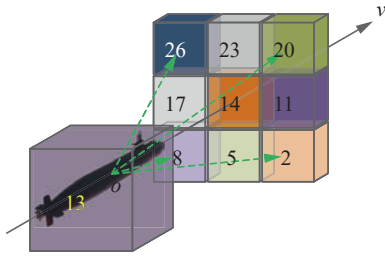


图 11 低速欠驱动 AUV 航行约束

Fig. 11 Schematic of low-speed underdriven AUV navigation constraints

3) 洋流速度与方向约束。任务海区的洋流速度与方向,会对低速欠驱动 AUV 的航行轨迹产生极大的影响。当 AUV 的航行方向与洋流方向的夹角为锐角时,会提升 AUV 的航行速度并减少其能量消耗,则给予奖励;当 AUV 处于逆流航行时,则根据实际情况给予适当的惩罚。并且,对于洋流会推动 AUV 产生侧向推动从而导致 AUV 偏离航线的情况,会对 AUV 执行每步操作时的状态进行修正,即 AUV 每步运行后的实时坐标是在经过动作数值选择和洋流干扰叠加后的坐标。

综上,通过充分考虑低速欠驱动 AUV 的航行特点、路径规划的目标以及洋流速度与方向的约束等方面设计出的奖励函数,可以极大地提高改进 DQN 算法训练效果的上限。

3 实验结果和分析

3.1 实验平台描述

本文以低速欠驱动 AUV 为研究对象,并设定 AUV 航行时的能耗恒定。选取东经 $122.95^\circ \sim 125.5^\circ$ 和北纬 $38.55^\circ \sim 40.0833^\circ$ 的海域作为实验海域,该海域的地形和洋流情况分别如图 5、6 所示。其中,海底地形数据来自于 S-57 电子海图,洋流数据来自于地球与空间研究 (earth & space research, ESR) 机构,洋流数据的采集日期为 2023 年 6 月 17 日。图 7 为实验海域的地形与洋流叠加图。

本文使用优先经验回放方法改进的 DQN 算法作为 AUV 的路径规划算法, AUV 的参数设置如表 1 所示。

表 1 AUV 参数设置
Table 1 AUV parameter setting

| 参数名 | 参数值 |
|--------|---------|
| 学习效率 | 0.000 4 |
| 折扣系数 | 0.99 |
| 贪心率 | 0.01 |
| 状态空间维度 | 142 |
| 动作空间维度 | 27 |
| 总训练回合数 | 13 600 |

为了训练 AUV 在任务海域进行路径规划,分别让 AUV 使用改进 DQN 算法先后在无洋流的环境和有洋流的环境进行路径规划训练。训练方法是在该海域中随机生成智能体的起点与目标点,起点与目标点坐标在一定的区间内随机选择,如表 2 所示。

表 2 训练参数设置
Table 2 Training parameter setting

| 参数名 | 参数值 |
|---------|-------------------------------------|
| 起点经度范围 | 东经 $124.00^\circ \sim 124.50^\circ$ |
| 起点纬度范围 | 北纬 $39.25^\circ \sim 39.70^\circ$ |
| 起点深度范围 | $-20 \sim 0$ m |
| 目标点经度范围 | 东经 $123.00^\circ \sim 123.67^\circ$ |
| 目标点纬度范围 | 北纬 $38.55^\circ \sim 38.90^\circ$ |
| 目标点深度范围 | $-48 \sim -25$ m |

3.2 实验结果

本文先后对无洋流环境和有洋流环境进行了路径规划的训练。得到了 2 种环境下的损失函数曲线,如图 12 所示。

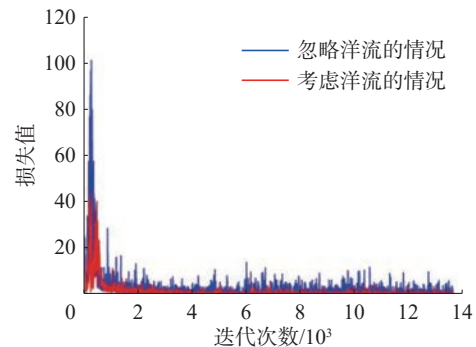


图 12 有无洋流干扰下的损失函数曲线

Fig. 12 Loss function curves with and without current interference

随机选定起点坐标 (东经 124.225° , 北纬 39.458° , 0.0 m) 和目标点坐标 (东经 123.314° , 北纬 38.695° , 43.5 m), 并分别在无洋流影响下的环境和

有洋流影响下的环境的进行路径规划测试,测试结果如图13、14所示。

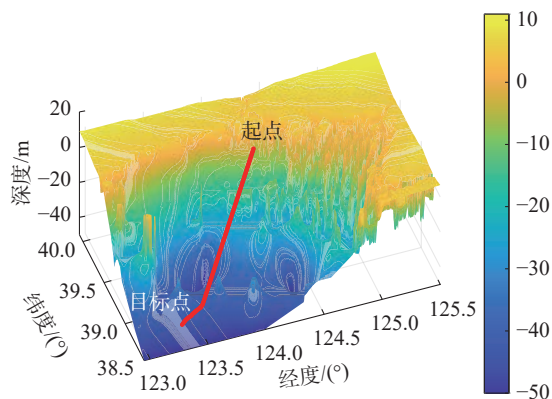


图13 无洋流干扰下的路径规划

Fig. 13 Path planning without current disturbance

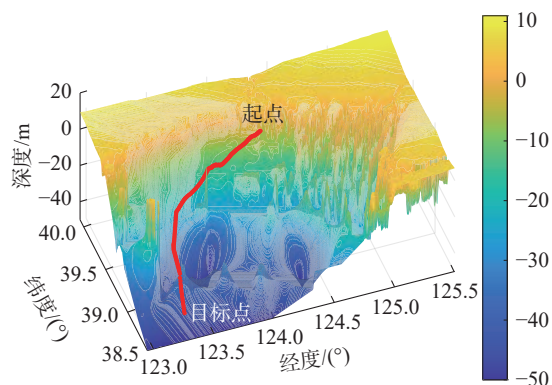


图14 有洋流干扰下的路径规划

Fig. 14 Path planning with current disturbance

2种情况下AUV随时间距终点的直线距离以及航行时对应的实时航速如图15所示。

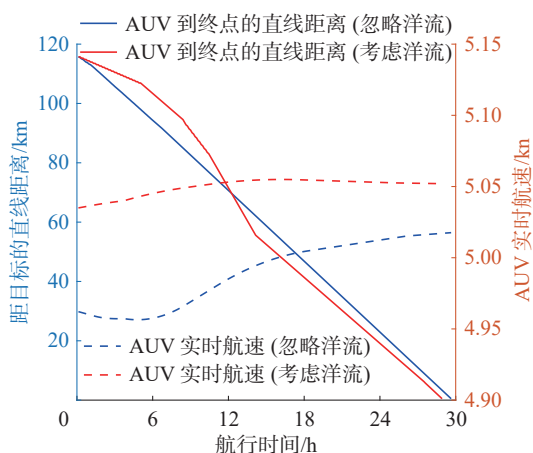


图15 AUV随时间到终点的直线距离和实时速度

Fig. 15 Straight-line distance from AUV over time to the end point and real-time velocity

3.3 实验结果分析

先后对无洋流干扰下的海洋环境与有洋流干扰下的海洋环境进行路径规划实验,并分别迭代13 600次之后,两者的损失数值都相对较小且趋

于稳定。如图13、14所示,相比于不考虑洋流干扰的AUV航线,在考虑洋流干扰的情况下AUV会更多地选择顺洋流航行的航线。如图15所示,虽然考虑洋流干扰的AUV实际航行的距离更远,但是其航速在整个航行期间都大于不考虑洋流干扰的AUV,并最终比不考虑洋流干扰的AUV节约5.29%的时间到达终点。由于本文假定航行时AUV的能耗恒定,即考虑洋流干扰下的路径规划比不考虑洋流干扰下的路径规划,AUV可以节约5.29%的能量消耗。故使用优先经验回放方法改进的DQN算法可以给AUV提供一条考虑洋流干扰情况下的全局路径规划,并且航行时间更短、能耗更低。

3.4 在虚拟仿真环境下的应用

为了检验改进DQN算法在复杂仿真环境中对低速欠驱动AUV进行路径规划的实用性,通过选定东经117.57°~东经126.91°和北纬35.67°~北纬41.00°的海域作为仿真实验海域,如图16为构建的可视化仿真实验海域。对应的二维彩色海底地形与洋流情况如图17所示。

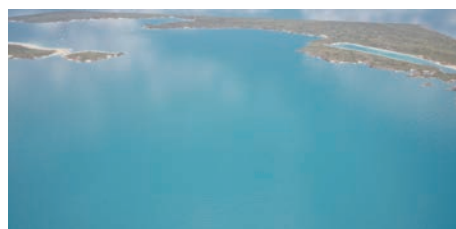


图16 仿真实验海域

Fig. 16 Simulated experimental sea area

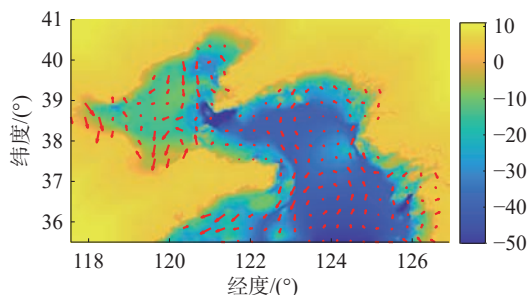


图17 仿真实验海域洋流模型

Fig. 17 Simulation of experimental ocean current models

如图17的洋流运动及对应的海底地形情况,可以分析出该片海域的海况比较复杂,使用此海域可以验证改进DQN算法在面对复杂海况时是否可以给AUV绘制出一条时间更短、能耗更低的航线。

通过设置起点坐标(东经120.487°,北纬39.825°,0.0 m)和终点坐标(东经124.566°,北纬38.864°,-42.8 m),使用改进DQN算法在仿真实验

海域进行路径规划,如图 18 所示。由图 18 中的给出的 AUV 路径规划轨迹可以看出,改进 DQN 算法可以在实验海域完成路径规划。并且结合当前海域的洋流情况,可以看出改进 DQN 给出的全局路径规划轨迹考虑了当前海域复杂的海洋运动特点,从而使得 AUV 航行时间更短、能耗更低。

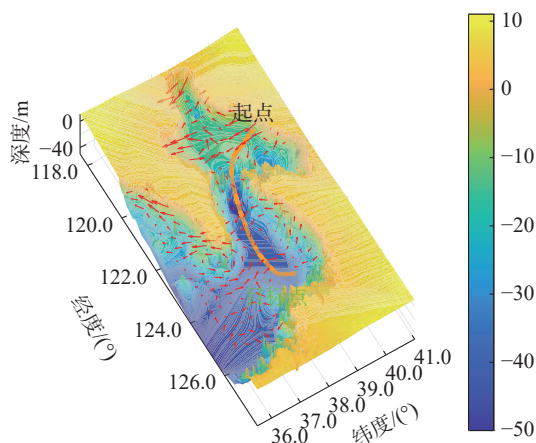


图 18 仿真实验海域的路径规划轨迹

Fig. 18 Simulation of path-planning trajectories in the experimental sea area

4 结束语

本文主要研究的是在洋流干扰的情况下如何让低速欠驱动 AUV 的路径规划更为合理,使得低速欠驱动 AUV 可以降低航行能耗并减少航行时间。

首先通过介绍洋流的运动并简化,并分析洋流运动对低速欠驱动 AUV 的影响机理,在此基础上建立指定任务海域三维地形模型及洋流运动模型,之后使用优先经验回放方法对传统 DQN 算法进行改进,从而避免传统 DQN 算法存在过估计的问题,并优化相应的奖励函数。最终 AUV 可以在有洋流干扰的环境下完成路径规划的任务。通过构建虚拟仿真海洋环境,并建立低速欠驱动 AUV 模型,将迭代后的路径规划算法应用在仿真 AUV 上,同样可以完成路径规划任务。实验结果表明,使用优先经验回放方法改进的 DQN 算法可以给 AUV 提供一条考虑洋流干扰情况下的全局路径规划,且航行时间更短、能耗更低。

在未来的工作中会进一步增加海洋动态环境的复杂程度,例如增加运动的物体等障碍物和海洋垂直方向运动的内波,提高低速欠驱动 AUV 在复杂多变的海域中的路径规划和避障能力。

参考文献:

[1] LI Xiaohong, YU Shuanghe. Three-dimensional path

planning for AUVs in ocean currents environment based on an improved compression factor particle swarm optimization algorithm[J]. *Ocean engineering*, 2023, 280: 114610.

[2] 李娟, 张韵, 陈涛. 改进 RRT 算法在未知三维环境下 AUV 目标搜索中的应用[J]. *智能系统学报*, 2022, 17(2): 368–375.

LI Juan, ZHANG Yun, CHEN Tao. Application of the improved RRT algorithm to AUV target search in an unknown 3D environment[J]. *CAAI transactions on intelligent systems*, 2022, 17(2): 368–375.

[3] BI Anyuan, ZHAO Fengye, ZHANG Xiantao, et al. Combined depth control strategy for low-speed and long-range autonomous underwater vehicles[J]. *Journal of marine science and engineering*, 2020, 8(3): 181.

[4] 刘甲, 周伟江, 马伟建. 低速 AUV 航渡过程中减少洋流影响的方法[J]. *舰船科学技术*, 2020, 42(3): 88–92, 97.

LIU Jia, ZHOU Weijiang, MA Weijian. Adaptive decision methods to reduce effect on ocean current to low-speed AUV sailing[J]. *Ship science and technology*, 2020, 42(3): 88–92, 97.

[5] LIU Rundong, CHEN Zonggan, WANG Zijia, et al. Intelligent path planning for AUVs in dynamic environments: an EDA-based learning fixed height histogram approach [J]. *IEEE access*, 2019, 7: 185433–185446.

[6] 应泽光, 何琪. 基于改进 A* 算法的无人艇复杂水域路径规划[J]. *机电技术*, 2022, 45(5): 33–35.

YING Zeguang, HE Qi. Complex water path planning for unmanned boats based on improved A* algorithm[J]. *Mechanical & electrical technology*, 2022, 45(5): 33–35.

[7] WANG Yanlong, LIANG Xu, LI Baoan, et al. Research and implementation of global path planning for unmanned surface vehicle based on electronic chart[C]//International Conference on Mechatronics and Intelligent Robotics. Kunming: Springer, 2017: 534–539.

[8] LI Ye, JIANG Yanqing, MA Shan, et al. Inverse speed analysis and low speed control of underwater vehicle[J]. *Journal of central south university*, 2014, 21(7): 2652–2659.

[9] DOHAN K. Ocean surface currents from satellite data[J]. *Journal of geophysical research (oceans)*, 2017, 122(4): 2647–2651.

[10] HU Siyuan, XIAO Shuai, YANG Jiachen, et al. AUV path planning considering ocean current disturbance based on cloud desktop technology[J]. *Sensors*, 2023, 23(17): 7510.

[11] 李慧, 赵琳, 毛英. 海况干扰下潜艇六自由度运动分析[J]. *哈尔滨工程大学学报*, 2017, 38(1): 94–100.

LI Hui, ZHAO Lin, MAO Ying. Analysis of six-degree-of-freedom motion in submarines under sea disturbance [J]. *Journal of Harbin engineering university*, 2017, 38(1): 94–100.

[12] XING Yuan, YOUNG R, NGUYEN G, et al. Optimal path planning for wireless power transfer robot using area division deep reinforcement learning[J]. *Wireless power*

- transfer, 2022, 9(1): 9921885.
- [13] 郭兴海, 计明军, 张卫丹, 等. 可变洋流环境中自主水下航行器动态路径规划的改进 QPSO 算法[J]. 系统工程理论与实践, 2021, 41(8): 2112–2124.
- GUO Xinghai, JI Mingjun, ZHANG Weidan, et al. Improved QPSO algorithm for dynamic path planning of autonomous underwater vehicles in variable ocean current environment[J]. Systems engineering-theory & practice, 2021, 41(8): 2112–2124.
- [14] KIANI F, SEYYEDABBASI A, ALIYEV R, et al. Adapted-RRT: novel hybrid method to solve three-dimensional path planning problem using sampling and metaheuristic-based algorithms[J]. Neural computing and applications, 2021, 33(22): 15569–15599.
- [15] 刘锋, 张严, 陈彦勇, 等. S-57 电子海图的快速读取及可视化存储[J]. 舰船科学技术, 2014, 36(7): 108–112.
- LIU Feng, ZHANG Yan, CHEN Yanyong, et al. Rapid reading and visual storage of the S-57 electronic chart[J]. Ship science and technology, 2014, 36(7): 108–112.
- [16] 扈震, 杨之江, 马振强. 基于 S-57 标准的电子海图三维可视化[J]. 地球科学, 2010, 35(3): 471–474.
- HU Zhen, YANG Zhijiang, MA Zhenqiang. Electronic navigation chart 3D visualization based on S-57[J]. Earth science, 2010, 35(3): 471–474.
- [17] HU Hao, ZHOU Yongjian, WANG Tonghao, et al. A multi-task algorithm for autonomous underwater vehicles 3D path planning[C]//2020 3rd International Conference on Unmanned Systems. Harbin: IEEE, 2020: 972–977.
- [18] KRIEG M, MOHSENI K. Dynamic modeling and control of biologically inspired vortex ring thrusters for underwater robot locomotion[J]. IEEE transactions on robotics, 2010, 26(3): 542–554.
- [19] BIJLSMA S J. Optimal ship routing with ocean current included[J]. Journal of navigation, 2010, 63(3): 565–568.
- [20] YANG Yang, LI Juntao, PENG Lingling. Multi-robot path planning based on a deep reinforcement learning DQN algorithm[J]. CAAI transactions on intelligence technology, 2020, 5(3): 177–183.
- [21] HAN J. An efficient approach to 3D path planning[J]. Information sciences, 2019, 478: 318–330.
- [22] LI Jianxin, CHEN Yiting, ZHAO Xiuniao, et al. An improved DQN path planning algorithm[J]. The journal of supercomputing, 2022, 78(1): 616–639.
- [23] QIAO Lei, ZHANG Weidong. Trajectory tracking control of AUVs via adaptive fast nonsingular integral terminal sliding mode control[J]. IEEE transactions on industrial informatics, 2020, 16(2): 1248–1258.
- [24] GU Yuwan, ZHU Zhitao, LYU Jidong, et al. DM-DQN: Dueling Munchausen deep Q network for robot path planning[J]. Complex & intelligent systems, 2023, 9(4): 4287–4300.
- [25] 赵玉新, 杜登辉, 成小会, 等. 基于强化学习的海洋移动观测网络观测路径规划方法[J]. 智能系统学报, 2022, 17(1): 192–200.
- ZHAO Yuxin, DU Denghui, CHENG Xiaohui, et al. Path planning for mobile ocean observation network based on reinforcement learning[J]. CAAI transactions on intelligent systems, 2022, 17(1): 192–200.
- [26] WU Keyu, WANG Han, ESFAHANI M A, et al. Achieving real-time path planning in unknown environments through deep neural networks[J]. IEEE transactions on intelligent transportation systems, 2022, 23(3): 2093–2102.
- [27] YANG Jian, XU Xin, YIN Dong, et al. A space mapping based 0–1 linear model for onboard conflict resolution of heterogeneous unmanned aerial vehicles[J]. IEEE transactions on vehicular technology, 2019, 68(8): 7455–7465.
- [28] PHUNG M D, HA Q P. Safety-enhanced UAV path planning with spherical vector-based particle swarm optimization[J]. Applied soft computing, 2021, 107: 107376.
- [29] ZHANG Jiaxin, LIU Meiqin, ZHANG Senlin, et al. Robust global route planning for an autonomous underwater vehicle in a stochastic environment[J]. Frontiers of information technology & electronic engineering, 2022, 23(11): 1658–1672.
- [30] WANG Jiankun, JIA Xiao, ZHANG Tianyi, et al. Deep neural network enhanced sampling-based path planning in 3D space[J]. IEEE transactions on automation science and engineering, 2022, 19(4): 3434–3443.
- [31] MELO A G, PINTO M F, MARCATO A L M, et al. Dynamic optimization and heuristics based online coverage path planning in 3D environment for UAVs[J]. Sensors, 2021, 21(4): 1108.
- [32] TAN Li, ZHANG Hongtao, SHI Jiaqi, et al. A robust multiple unmanned aerial vehicles 3D path planning strategy via improved particle swarm optimization[J]. Computers and electrical engineering, 2023, 111: 108947.
- [33] QI Yongqiang, LI Shuai, KE Yi. Three-dimensional path planning of constant thrust unmanned aerial vehicle based on artificial fluid method[J]. Discrete dynamics in nature and society, 2020: 4269193.

作者简介:



陈世同, 讲师, 主要研究方向为海洋观测探测技术、智能导航。曾获黑龙江省技术发明奖一等奖 1 次、三等奖 1 次, 获发明专利授权 5 项, 发表学术论文 10 余篇。E-mail: chenshitong@hrbeu.edu.cn。



鲁子瑜, 硕士研究生, 主要研究方向为强化学习算法、路径规划算法。E-mail: lycxlzy@hrbeu.edu.cn。