文章编号:1000 - 4653(2026)04 - 0039 - 08

# 基于改进近端策略优化的航标自航方法研究

薛德来[1*], 朱景鹏[1], 吴川博[2]
(1. 交通运输部北海航海保障中心 烟台航标处, 山东 烟台 264000;
2. 河北科技大学 机械工程学院, 河北 石家庄 050018)

**摘 要**:在近海港口航道繁忙水域,海上实体航标极易发生碰撞,导致标体损坏,无法正常运行。为解决现有方案针对航标自航研究领域空白,不能实现自主航行避碰,未能有效结合人工智能技术使其智能化航行工作等问题,提出一种改进近端策略优化并结合速度障碍的航标自航方法。近端策略优化算法采用离线学习方式提高网络更新效率,设计航标自航动作空间,能更快实现模型收敛;但在模型收敛时容易陷入局部最优,为避免模型陷入局部最优,结合速度障碍算法表示状态空间,用速度向量设计奖励函数,使其在计算距离奖励函数基础上,增加基于位置和速度信息的奖励函数,提高收敛速度,并满足航标在近海港口航道中自适应地选择最佳速度;为提高算法鲁棒性,增加熵奖励函数对近端策略优化算法改进,以激励更多探索。试验表明,增加熵奖励函数与未增加相比,在平均航行步长、平均航行速度等指标上分别提升了 9.3% 、10.2% ,且模型测试成功率保持在 99% 以上。该方法实现了航标无人操控、自主航行及避碰的安全作业模式,降低航标在繁忙水域被碰撞损坏的风险,减少人工巡检和维护成本。对加快构建现代化航海保障体系、保障海上航行安全具有重要战略意义。

**关键词**:航行安全; 航标自主航行; 改进近端策略优化; 熵奖励; 航标避碰; 航标保护
**中图分类号**:U644.4; U675.96 **文献标志码**:A **DOI**:10.3969/j.issn.1000 - 4653.2026.04.005

# Autonomous navigation method for aids to navigation based on improved proximal policy optimization

*Xue Delai*[1*], *Zhu Jingpeng*[1], *Wu Chuanbo*[2]
(1. Yantai AtoN Division, Northern Navigation Service Center, Maritime Safety Administration, People's Republic of China, Yantai 264000, China; 2. School of Mechanical Engineering, Hebei University of Science and Technology, Shijiazhuang 050018, China)

**Abstract**: In the busy waters of nearshore port channels, physical Aids to Navigation (AtoN) are highly susceptible to collisions, which may cause structural damage and prevent them from operating normally. To address the research gap in autonomous navigation of AtoN, including the inability of existing methods to achieve autonomous navigation and collision avoidance and their insufficient integration with artificial intelligence technologies for intelligent navigation, this paper proposes an autonomous navigation method for AtoN based on improved Proximal Policy Optimization (PPO) combined with the Velocity Obstacle (VO) algorithm. The PPO algorithm was improved by introducing an offline learning strategy to enhance network update efficiency. By designing the autonomous navigation action space for AtoN, the model can achieve faster convergence. However, the model is prone to falling into local optima during convergence. To avoid this problem, the VO algorithm is incorporated to represent the state space, and a reward function is designed using velocity vectors. On the basis of calculating the distance-based reward function, an additional reward function based on position and velocity information is introduced to improve convergence speed and enable AtoN to adaptively select the optimal velocity in

**收稿日期**:2025-12-10
**基金项目**:河北省自然科学基金青年项目(E2026208041);河北省高等学校科学研究青年拔尖人才项目(BJ2026336)资助
**第一作者(*通信作者)**:薛德来(1999—),男,助理工程师,研究方向为航标技术管理,(E-mail)2035822746@qq.com
**引用格式**:薛德来,朱景鹏,吴川博. 基于改进近端策略优化的航标自航方法研究[J]. 中国航海,2026,49(4):39 - 46.
Xue D L, Zhu J P, Wu C B. Autonomous navigation method for aids to navigation based on improved proximal policy optimization[J]. Navigation of China,2026,49(4):39 - 46. (in Chinese)

nearshore port channels. To further enhance algorithm robustness, an entropy reward term is added to the PPO algorithm to encourage greater exploration. Experimental results show that, compared with the method without the entropy reward term, the proposed method improves the average navigation step length and average navigation velocity by 9.3% and 10.2%, respectively, while maintaining a model test success rate of over 99%. The proposed method enables a safe operational mode featuring unmanned control, autonomous navigation, and collision avoidance for AtoN. It can reduce the collision damage risk of AtoN in busy waters and lower the costs of manual inspection and maintenance. The study is of significance for accelerating the construction of a modern maritime navigation support system and ensuring maritime navigation safety.

**Key words**: navigation safety; autonomous navigation of AtoN; improved proximal policy optimization; entropy reward; AtoN collision avoidance; AtoN protection

海洋大国背景下,海事领域发展迅速,在加快建设海洋强国中占举足轻重的地位。党的十八大到党的二十大工作报告中,皆对海洋强国战略做出重要指示,对于海洋资源的开发和利用受到广泛重视。提到海事,必不可少的就是船舶,作为海洋开发利用、海上贸易的中坚力量,其安全航行占据突出地位。保障船舶安全航行,除了船舶驾驶操作等内部因素,还需外部辅助设备,航标作为航海保障的重要组成部分在其中发挥重要角色。

2024年8月22日,国际航标协会(International Association of Lighthouse Authorities, IALA)于法国宣布《国际航标组织公约》即日生效,国际航标组织正式成立。目前已有包括中国在内的35个国家加入,成为新组织成员国[1]。航标是海上交通安全保障体系的重要组成部分。是保证船舶安全航行,保护航道健康稳定的重要手段。截至目前,我国已建成全球规模最大的航海保障体系,实现了沿海港口航道"一图覆盖",沿海水上安全通信"一网贯通"。

当前对于航标技术提升以及航标保护的研究,研究人员做出大量努力。为保证航标服务质量,节约成本,以信息化带动航标管理现代化,在"数字海事"背景下推动数字化航标发展[2]。受潮涨潮落影响导致航标发生飘移,提出潮流场作用下航标飘移计算方法[3]。为提升航标漂移监测效果,提出一种基于高分辨率遥感影像的船舶航标漂移监测方法[4]。针对航标受风浪、气候或撞击等因素的影响,提出一种基于机器视觉的轻量化航标倾斜检测方法[5]。

大量研究表明数字系统方法可有效提升航标监测效率,对于航标故障问题能够及时响应,但是提高监测效率的同时应该同步提高航标实体安全技术以及故障维修效率。本研究针对航标实体安全技术领域研究空白以及提升故障维修效率,提出一种改进近端策略优化结合速度障碍的航标自航方法,使其可以自主航行满足作业需求,并在发生灯器损坏等故障可实现自主返航维修,节省人力物力,节约时间成本。

## 1 问题现状

自航式航标研究中,由于灯船[6]船体与移动船艇型体一致,在现有无人艇技术基础上[7-8],容易实现自航式智能灯船技术。智能灯船控制系统研究中,一是当前灯船结构单一,仅实现发光引航,少数搭载雾号、雷达应答器、无线电指向标等,为非自航船,只有极少数为自航船。二是灯船结构较大,发生碰撞会造成较大损失。2025年10月铜陵段大通水道#248红浮被撞漂失,影响货船进港造成数十万经济损失;2025年11月南京段大胜关桥区红浮被撞,船体损坏,灯器丢失,合计损失达6万余元。迫切需要开发一种避碰技术,使得灯船搭载推进器成为智能灯船。为此需要研究一种自主航行避碰[9]的智能灯船控制方法,实现灯船无人作业、自主避碰、自主返航的安全作业模式。

现有技术的缺点表现在,一是灯船大多为非自航船,未搭载推进器等动力装置,船体移动需要航标船以及人工辅助;二是目前灯船几乎没有结合人工智能相关技术,自主航行避碰多应用于无人船,灯船自主航行可参考技术较少;三是当前自主航行技术在船舶领域的应用大多停留在仿真实验阶段,由于工程应用周期长、涉及知识面广和实现困难等因素,一直未取得重大突破,所以未能在航标领域进行研究;四是当前灯船检修皆需海事工作人员出海并登上灯船检修,在风浪大的恶劣海域环境,严重威胁海事工作人员的人身安全;五是在遭遇船体靠近时无法自主避碰,造成严重损失。因此研究一种自主航行避碰的智能航标方法尤为重要。

## 2 关键技术

为解决现有航标体在繁忙水域易发生碰撞,无法实现自主航行避碰从而导致后续抢修工作耗费大量人力物力财力等问题,本文提出一种航标自航方法研究,该方法以近端策略优化算法(Proximal Policy Optimization, PPO)为基础[10],改进其奖励函数

再辅以扩展策略,速度障碍算法(Velocity Obstacle, VO)对近端策略优化算法中动作选择进行向量表示,以速度向量限制动作选择范围,减少动作误差不稳定问题,同时也能有效解决网络过高估计问题。速度障碍算法[11]相比应用于多无人艇领域的相互速度障碍算法以及最优相互避碰算法的优点:一是计算量少,以更小的算力实现最优决策;二是更适合智能灯船等单体应用对象,响应速度更快,航行效率更高;三是能够对近端策略优化算法网络起到正向激励作用,削弱高估计问题和误差不稳定问题。

### 2.1 改进近端策略优化算法

近端策略优化是当前最先进的深度强化学习算法之一,在自主航行领域应用广泛,并且已经在很多项目中得到应用。是一种无模型基于策略优化梯度的算法。该算法旨在学习一种策略,可以根据训练期间的经验最大化获得累积奖励。

算法结构见图1,主体网络模型为 Actor 网络模型和 Critic 网络模型,两个网络模型皆为双网络。Actor 网络模型双网络表示为 $\pi_\theta$,$\pi_{\theta_{\text{old}}}$。首先在 $\pi_\theta$ 网络策略下选择动作,并将动作、状态、奖励等信息储存在经验池中,循环此步骤,从经验池中随机抽取样本更新 $\pi_{\theta_{\text{old}}}$。在训练循环时,经验池储存样本输入 Critic 网络模型,由样本数据评价 Actor 策略的准确性,计算优势函数 $\hat{A}A_t$ 以及状态值。循环训练梯度更新。该算法由一个参与者(actor)$\pi_\theta(\cdot\mid s_t)$ 和一个评估者(critic)$V(s_t)$ 组成。前者在时间 $t$ 处输出下一个动作的概率分布,后者估计该状态的预期累积奖励(标量)。由于 actor 和 critic 都将状态作为输入,因此可以在提取高级特征的两个网络之间共享骨干架构。旨在使策略选择具有更高"优势"的行动,即具有比评估者预测高的累积奖励。同时,抑制更新策略次数,避免出现过优化问题。针对过优化问题,将结合速度障碍算法进行改进。

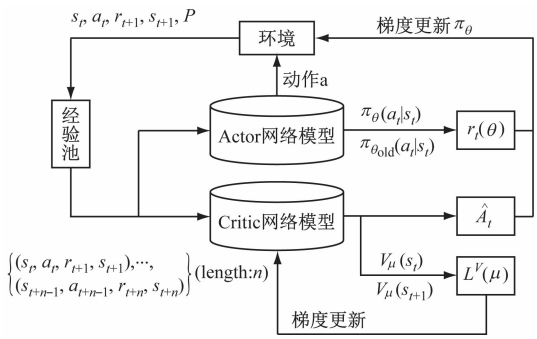



图1 PPO 算法结构框架

Fig. 1 Structural framework of the PPO algorithm

PPO 定义式方程式为

$$L_{t,\text{CLIP+VF}}(\theta) = \hat{E}_t[L_{t,\text{CLIP}}(\theta) - c_1 K_{\text{L}} L_{t,\text{VF}}(\theta)] \tag{1}$$

式(1)中,$c_1$ 是超参,$c_1 \in (0,1)$,$K_{\text{L}}$ 为散度系数,$L_{t,\text{CLIP}}(\theta)$ 为损失函数,$L_{t,\text{VF}}(\theta)$ 为监督学习函数,分别衡量策略评估(critic)和探索(exploration)准确性的重要性。

损失函数激发行为概率最大化,从而导致行为正面优势,损失函数式为

$$L_{\text{CLIP}}(\theta) = \hat{E}_t\left[\min\left(r_t(\theta)\hat{A}_t, \text{clip}\begin{pmatrix} r_t(\theta), \\ 1-\varepsilon, \\ 1+\varepsilon \end{pmatrix}\hat{A}_t\right)\right] \tag{2}$$

式(2)中,$\hat{E}_t$ 表示求期望估计,$\varepsilon$ 是超参,$\varepsilon \in (0,1)$。

$$r_t(\theta) = \frac{\pi_\theta(a_t \mid s_t)}{\pi_{\theta_{\text{old}}}(a_t \mid s_t)} \tag{3}$$

优势函数式为

$$\hat{A}_t = -V(s_t) + r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \cdots + \gamma^{T-t+1} r_{T-1} + \gamma^{T-t} V(s_T) \tag{4}$$

式(4)中,$\hat{A}_t$ 是估计优势,$-V(s_t)$ 是估计的初始状态值,$\gamma^{T-t}V(s_T)$ 是估计的最终状态值,$\gamma$ 为超参,$\gamma \in (0,1)$。

为对优势有一个良好的估计,需要一个可以预测给定状态值的评估者。该模型为具有简单均方误差(Mean Squared Error, MSE)损失的监督学习:

$$L_{t,\text{VF}} = \text{MSE}\begin{pmatrix} r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \\ \cdots + \gamma^{T-t+1} r_{T-1} + \\ \gamma^{T-t} V(s_T), V(s_t) \end{pmatrix} = \| \hat{A}_t \|_2 \tag{5}$$

最后,对于策略高熵,提出算法改进,增加熵奖励函数,以激励更多探索。标准熵式为

$$S[\pi_\theta](s_t) = -\int \pi_\theta(a_t \mid s_t) \ln(\pi_\theta(a_t \mid s_t))\,\text{d}a_t \tag{6}$$

得到改进近端策略优化(enhanced proximal policy optimization, EPPO)目标式方程式为

$$L_{t,\text{CLIP+VF}}(\theta) = \hat{E}_t[L_{t,\text{CLIP}}(\theta) - c_1 K_{\text{L}} L_{t,\text{VF}}(\theta) + c_2 S[\pi_\theta](s_t)] \tag{7}$$

式(7)中,$c_2$ 为超参,$c_2 \in (0,1)$。

### 2.2 速度障碍算法

VO 算法是基于速度避碰的简洁式算法之一,采用几何方式算法解释更加直观。

考虑由 $A$ 和 $B$ 组成的系统。$A$ 和 $B$ 的速度分别是 $V_A$ 和 $V_B$,半径分别是 $R_A$ 和 $R_B$,此时时间为 $t_0$。其中 $A$ 代表灯船,$B$ 代表障碍物。首先将 $B$ 映射到 $A$ 的构型空间。可将 $A$ 视为质点,同时将 $B$ 的半径

增加 $R_A$，从而得到 $A'$ 和 $B'$。表示各自的位置和圆心的速度。

碰撞锥是碰撞时的相对速度集合[12]，见式(8)和式(9)。

$$C_{A,B} = \{V_{A,B} \mid \boldsymbol{\lambda}_{A,B} \cap B' \neq \varnothing\} \tag{8}$$

$$V_{A,B} = V_A - V_B \tag{9}$$

式(8)和式(9)中，$V_{A,B}$ 表示 $A'$ 相对于 $B'$ 的相对速度，$\lambda_{A,B}$ 是 $A'$ 相对于 $B'$ 的向量，见图2。

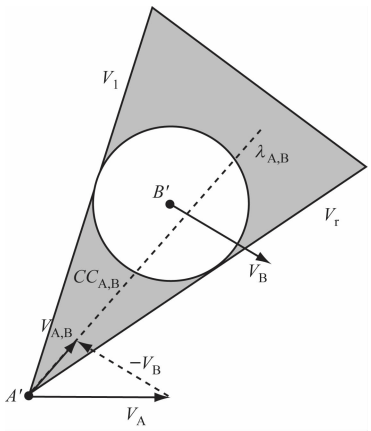



图2 $A$，$B$ 相对速度 $V_{A,B}$ 及 $VO$ 锥 $C_{A,B}$

Fig. 2 $A$, $B$ relative velocity $V_{A,B}$ and $VO$ cone $C_{A,B}$

最终，将 $V_O$ 定义为

$$V_O = C_{A,B} \oplus V_B \tag{10}$$

式(10)中，⊕代表闵可夫斯基向量加。

$VO$ 将 $A'$ 的绝对速度表示为安全区域和非安全碰撞区域，满足

$$A(t) \cap B(t) = \varnothing, \text{ if } V_{A(t)} \notin V_O(t) \tag{11}$$

### 2.3 航标自航算法原理

VO 算法速度向量表示 EPPO 算法运作机制可以解决避碰路径规划难题[13]。改进后的 EPPO-VO 算法结构，见图3。

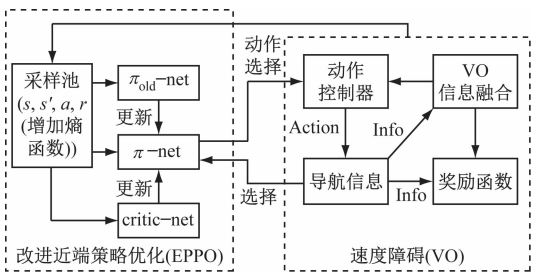



图3 EPPO-VO 算法结构框图

Fig. 3 Structural block diagram of EPPO-VO algorithm

本研究一是可以实现灯船智能化航行，满足作业任务需求；二是减少航标船投放工作，节省大量人力物力，最重要的是保障海事工作人员的人身安全；三是采用电驱践行绿色海洋、绿色航行目标；四是在遭遇碰撞局面时，实现自主避碰。基于本研究制定工程应用，理论与实际相结合形成一种新的航标智能化方案。

## 3 试验验证

### 3.1 研究场景设置

仿真场景设置灯船的最大速度为 1.5 m/s。在之前研究工作中，通过实船试验测定团队无人艇速度在 1.2 m/s—1.5 m/s 时既可以满足算法验证需求又能够保证船体发生碰撞时不会受到严重损伤[14]。可视化界面中多边形表示静态障碍物，可视为近海养殖浮漂网箱等，黑色球形表示动态障碍物，可视为近海渔船等，因近海货船渔船在港口航道内保持直线低速行驶，故动态障碍物运动轨迹设置为匀速直线，且航行轨迹与航标灯船航行轨迹交叉用以验证避碰效果，速度设置为 0.6 m/s。

### 3.2 研究平台设计

试验对象本研究以航标灯船为例，灯船是航道的重要组成部分，其外观形态更适合船艇模型。工程应用研究中，首先对灯船进行加装改进，保持原有功能不变的情况下，搭载信号传输与传感器系统，控制系统，动力推进系统和能源供应系统等；其次基于提出的智能灯船自主避碰方法，设计一种算法验证的工程应用方案，依据灯船传感器收集环境信息和灯船自身当前信息，通过通信协议对灯船控制系统发送运动指令。决策模型(EPPO-VO 训练得到的决策模型)依据获取的环境信息计算下一时刻位置，计算推力和力矩，动力推进系统根据运动指令推力分配。能源供应采用太阳能发电板与蓄电池相结合的方式保证动力需求；再次，智能灯船航行至目标位置，依靠相关定位技术，控制艉推配合艏侧推工作，直至智能灯船航行至准确位置并进行姿态调整；最后，智能灯船实现正常作业，如遇到灯船发生故障问题，向岸基工作人员发出检修信号，以便工作人员根据检修信号准备检修工作，同时码头重新派出智能灯船接替故障灯船位置。故障灯船检修信号发出并触发动力装置，自主航行至码头接受检修工作。

对于定位技术，本研究提出两种方案：一是灯船加装锚定装置，依赖北斗卫星定位确定作业位置后，启动锚定装置，将沉石抛入海底；二是针对算法进行改进：静止状态下，受波浪流的影响会发生偏离，设置触发动力装置的阈值 $\delta$，$\delta \in [5,10)$ m；若智能灯船偏离位置距离目标位置坐标 $L < 5$ m，则不影响作业，无需触发动力装置调整位置；若智能灯船偏离位置距离目标位置坐标 $L \in [5,10)$ m，则触发动力装置航行至作业位置；若偏离位置 $L > 10$ m 范围，则触

发报警装置和锚定装置,出现这种情景说明推进系统发生故障,无法自主航行至目标位置,需及时进行救助维护,见图 4。方案一:在作业过程中更加可靠,灯船位移概率低,同时也存在一定的不足,在发生故障需自主返航维修时,锚定装置的动力可能不足以将埋在海底的沉石绞至灯船,无法实现自主返航阶段。方案二:灯船位移概率高,但推进器阶段性调整智能灯船的位置姿态,相对于长时间在海水中静止,减少故障发生率,且能够提供自主返航维修需求。

仿真试验中,弱化外部因素对试验结果的影响,将灯船表示为有避碰半径的圆形区域,运动模型采用双桨模型,使其验证本研究所提航标自航方法。

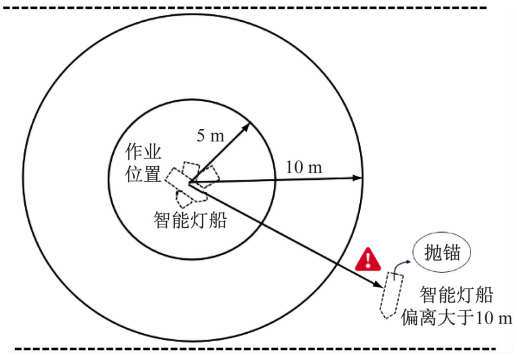



图 4 阈值区域划分

Fig. 4 Threshold regional division

### 3.3 奖励函数设计

算法奖励函数($R_{vo}$)[15]。

$$R_{vo}=\begin{cases}P_2+P_3\cdot R_{dd}+P_4\cdot R_t & ,v_t\in V_0 \text{ and } t_{\min}>0.1\\ P_2+P_1\cdot P_4\cdot R_t & ,v_t\in V_0 \text{ and } t_{\min}<0.1\\ P_5+P_6\cdot R_{dd} & ,v_t\notin V_0\end{cases} \tag{12}$$

$$R_{dd}=-\frac{dd}{dd_{\max}} \tag{13}$$

$$R_t=-\frac{0.2}{t_{\min}+0.2} \tag{14}$$

式(12)—式(14)中,$P_*$是常量,设置值如表 1 所示。$dd=\|v_t-v_{t,\mathrm{des}}\|$,$dd_{\max}=2$,$R_{dd}\in(0,1)$,$R_t\in(0,1)$。$t_{\min}$表示发生碰撞最短时间。

**表 1 奖励函数参数值**

**Tab. 1 Parameter values of the reward function**

| 参数 | 数值 |
|---|---|
| $P_1$ | 3.6 |
| $P_2$ | 0.3 |
| $P_3$ | 0.6 |
| $P_4$ | 1.8 |
| $P_5$ | 0.3 |
| $P_6$ | 1.8 |

### 3.4 研究方法验证

试验场景示意图表示完整的航标流程作业,见图 5。第一阶段为灯船自主航行至作业位置,第二阶段表示灯器出现故障时自主返航。在航标实际工作中,灯船返航的同时岸基派出功能正常的灯船接替工作。在灯船检测到灯器发生故障时,即使向岸基工作人员反馈,同时向码头指定维修地点返航。另外,在接收到报警的同时,岸基向等待作业的灯船发出任务指令,灯船自主航行至故障灯船位置接替其工作。验证所提航标自航方法在仿真环境中实现自主航行避碰任务的可行性与有效性,使用两个指标评估避碰算法性能反映算法模型效率[16],通过结果可知 EPPO-VO 模型效率优于 PPO-VO,如表 2 所示,平均航行步长缩短 9.3%,平均航行速度提升 10.2%。

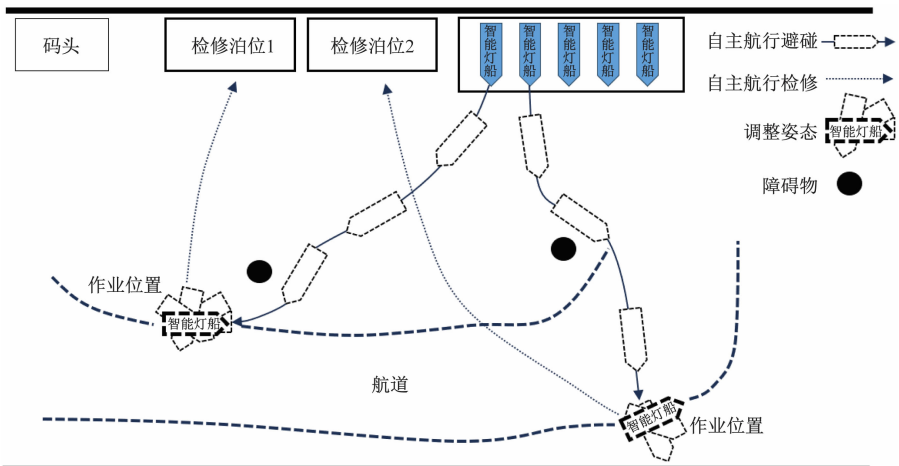



图 5 航标作业流程示意

Fig. 5 Schematic diagram of AtoN operation process

表 2 避碰算法航行数据比较结果

Tab. 2 Comparison results of navigation data for collision avoidance algorithms

| 算法 | 平均航行步长/step | 平均航行速度/(m/s) |
|---|---|---|
| PPO-VO | 129 | 0.916 |
| EPPO-VO | 117 | 1.009 |

验证灯船从岸基自主航行避碰至作业位置,执行作业任务。以两条灯船为例,下达任务指令后,灯船向着作业位置航行,航行过程中观测到障碍物时及时躲避,顺利到达目标位置。

### 3.4.1 试验场景一

灯船在无障碍的水域环境航行到达作业位置,见图 6。设置此场景目的是灯船布设或返航过程中,结合船舶自动识别系统(Automatic Identification-System,AIS)系统选取没有船只通航的时间段,可最大程度避免碰撞事故,也避免干扰船长决策。从轨迹可得路径规划接近于一条直线,接近最优轨迹,表示算法在路径规划实现最优解。

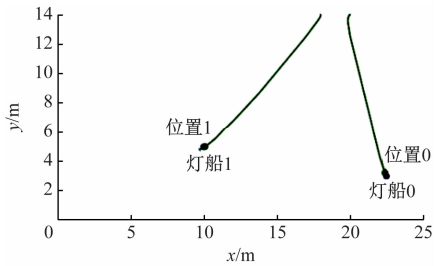



图 6 航标自主航行示意

Fig. 6 Schematic diagram of AtoN autonomous navigation

灯船在存在障碍物水域航行至作业位置,见图 7 和图 8。设置此场景目的是灯船布设或返航过程中,通航船只极少数场景下,灯浮标、浮漂、礁石等静态障碍物不可避免,需验证此场景下的自主航行避碰效果。

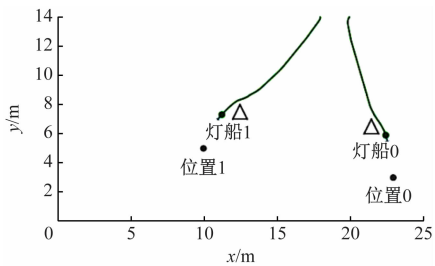



图 7 航标自主航行避碰示意一

Fig. 7 Schematic diagram 1 of AtoN autonomous navigation and collision avoidance

图 7 是灯船会遇障碍物时避障场景,在遇障碍物时能够及时调整姿态完成避碰,图 8 为完整避障过程到达指定作业位置,在完成避碰过程后,调整艏向继续向作业位置航行。

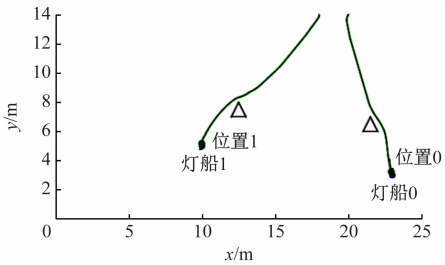



图 8 航标自主航行避碰示意二

Fig. 8 Schematic diagram 2 of AtoN autonomous navigation and collision avoidance

### 3.4.2 试验场景二

验证灯船在检测到灯器发生故障时自主返航等待修复,图 9 为灯船在无障碍的水域环境自主返航检修到达码头位置,从轨迹可得路径规划接近于一条直线,为最优轨迹。图 10 为灯船在存在障碍物水域自主返航检修到达码头位置,图 11 和图 12 是灯船同时会遇动态障碍物和静态障碍物时避障场景,设置此场景目的是在灯船布设或返航过程中,在繁忙港口不可避免地有通航船只通行,港口航道内船只保持低速直线运动,设置此场景可以验证航道内有船只时灯船自主航行避碰效果。在会遇船只时灯船航行优先级最低,默认对方不改变航行行为,由我方自主避碰。图 13 为完整避障过程到达码头检修位置。

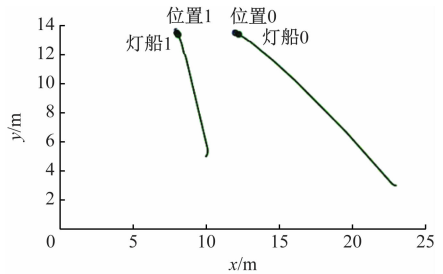



图 9 航标自主返航示意一

Fig. 9 Schematic diagram 1 of AtoN automatic return

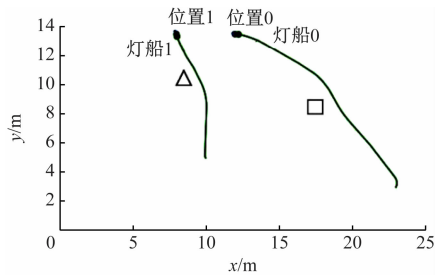



图 10 航标自主返航示意二

Fig. 10 Schematic diagram 2 of AtoN automatic return

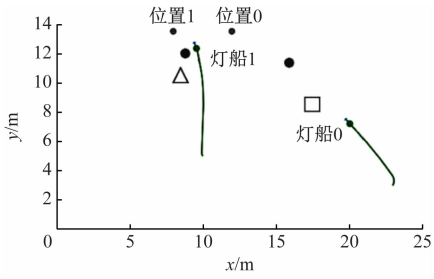



图 11 航标自主返航避碰示意一

Fig. 11 Schematic diagram 1 of AtoN automatic return and collision avoidance

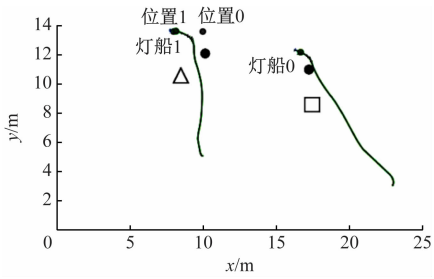



图 12 航标自主返航避碰示意二

Fig. 12 Schematic diagram 2 of AtoN automatic return and collision avoidance

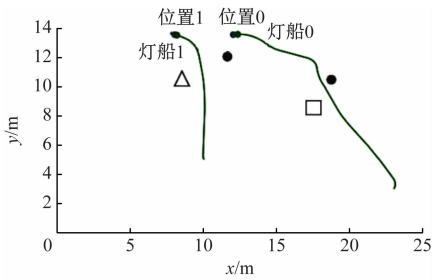



图 13 航标自主返航避碰示意三

Fig. 13 Schematic diagram 3 of AtoN automatic return and collision avoidance

### 3.4.3 未来试验设计

基于本研究以及团队研究工作,在未来工作中将实现仿真场景的实际验证,与相关领域专家团队合作,依赖与灯船船模相近的实际无人艇做工程试验验证,在此基础上不断探索灯船加装装备的改进以及通信功能的完善,逐步实现完整的航海保障技术方案。

## 4 总结与展望

本文针对航标自航方法开展研究,通过当前存在实际问题及航标保护的迫切需要,提出一种自航方法用以满足航标安全工作需求。针对航标痛点,经过理论研究与实地考察,通过大量试验形成有效解决方案,经过试验验证,所提算法训练所得避碰模型测试成功率达 99% 以上,对比未改进算法获得更短的航行步长和更快的航行速度,在工程应用表现为节省时间和能源,表明自航方法的实用性及可靠性。基于以上工作及先前研究工作,下一步将借助试验平台进行实际验证,形成工程技术方案,为海洋强国贡献一份力量。

## 参 考 文 献

[1] 曾繁屏,肖小春,杨弘矗,等. 国际航标组织转制和我国对策探讨[J]. 中国海事,2025,(1):22-24.
Zeng F P, Xiao X C, Yang H C, et al. Discussion on IALA′s change of status and China′s countermeasures [J]. China Maritime Safety, 2025, (1): 22-24. (in Chinese)

[2] 黄凤飞. "数字"航标运行保障中心建设研究[J]. 中国航海,2013,36(4):34-37.
Huang F F. On construction of digitized AtoN supporting center[J]. Navigation of China, 2013,36(4):34-37. (in Chinese)

[3] 周春辉,赵俊男,甘浪雄,等. 潮流场作用下的航标漂移计算方法研究[J]. 安全与环境学报,2021,21(1):217-223.
Zhou C H, Zhao J N, Gan L X, et al. On the early warning method of the navigation buoy drift under the tidal current field [J]. Journal of Safety and Environment, 2021,21(1):217-223. (in Chinese)

[4] 王晓晴. 基于高分辨率遥感影像的船舶航标漂移监测方法[J]. 舰船科学技术,2023,45(16):173-176.
Wang X Q. Ship beacon drift monitoring method based on high resolution remote sensing images [J]. Ship Science and Technology, 2023,45(16):173-176. (in Chinese)

[5] 赵建森,赵婧莹,韩冰,等. 基于机器视觉的航标倾斜检测方法[J]. 中国航海,2024,47(4):44-50.
Zhao J S, Zhao J Y, Han B, et al. Detection method of the buoy tilt based on machine vision[J]. Navigation of China, 2024,47(4):44-50. (in Chinese)

[6] 孙丽辉,张建伟,郑尚龙. 改善广州伶仃航道马友石水域通航条件的必要性及可行性探讨[J]. 航海技术,2025,(2):25-29.
Sun L H, Zhang J W, Zheng S L. Discussion on the necessity and feasibility of improving the navigational conditions of Mayoushi waters in Guangzhou Lingding navigation channel[J]. Marine Technology, 2025,(2):25-29. (in Chinese)

[7] 会鑫,王宁,王帅. 基于局部状态感知的无人艇深度强化学习路径规划[J]. 中国舰船研究,2026,21(3):317-327.
Hui X, Wang N, Wang S. Local perception-based path planning for unmanned surface vehicles using deep reinforcement learning [J]. Chinese Journal of Ship

Research, 2026,21(3):317-327. (in Chinese)

[8] 蔡庆,孙世平.无人艇自主决策技术特点及发展研究[J].舰船科学技术,2025,47(5):8-15.
Cai Q, Sun S P. Research on the characteristics and development of autonomous decision-making technology for unmanned surface vehicles[J]. Ship Science and Technology, 2025,47(5):8-15. (in Chinese)

[9] Li Y, Wu D, Wang H, et al. Dynamic collision avoidance for maritime autonomous surface ships based on deep Q-network with velocity obstacle method[J]. Ocean Engineering, 2025,320: 120335.

[10] Schulman J, Wolski F, Dhariwal P, et al. Proximal policy optimization algorithms[J]. 2017(7): 06347.

[11] Fiorini P, Shiller Z. Motion planning in dynamic environments using the relative velocity paradigm[C]// Robotics and Automation, 1993. Proceedings IEEE International Conference on Robotics and Automation. Atlanta, IEEE, 1993: 560-565.

[12] 熊勇,贺益雄,黄立文.基于速度障碍的多船自动避碰控制方法[J].中国航海,2015,38(3):46-51.
Xiong Y, He Y X, Huang L W. Multi-ship collision avoiding control considering velocity obstacle[J]. Navigation of China, 2015, 38(3): 46-51. (in Chinese)

[13] Feng S, Sun H, Yan X, et al. Dense reinforcement learning for safety validation of autonomous vehicles[J]. Nature,2023, 615: 620-627.

[14] Wu D, Yuan K, Huang Y, et al. Design and test of an improved active disturbance rejection control system for water sampling unmanned surface vehicle[J]. Ocean Engineering,2022, 245: 110367.

[15] 薛德来.近海多无人艇协同作业避碰策略研究[D].厦门:集美大学,2024.
Xue D L. Research on collision avoidance strategies for cooperative operations of multi-unmanned surface vehicles in offshore environments[D]. Xiamen: Jimei University, 2024. (in Chinese)

[16] Xue D, Wu D, Yamashita A S, et al. Proximal policy optimization with reciprocal velocity obstacle based collision avoidance path planning for multi-unmanned surface vehicles[J]. Ocean engineering, 2023, 273: 114005.

(上接第29页)

Li Q H, Ren J S, Hua Y. Parametric identification of ship maneuvering motion response model based on square root cubature kalman filtering[J]. Journal of System Simulation, 2024, 36(8): 1790-1799. (in Chinese)

[8] He H W, Zou Z J. Black-box modeling of ship maneuvering motion using system identification method based on BP neural network[C]//Proceedings of the ASME 2020 39th International Conference on Ocean, Offshore and Arctic Engineering. Florida: American Society of Mechanical Engineers (ASME), 2020: V06BT06A037.

[9] Luo W, Zhang Z. Modeling of ship maneuvering motion using neural networks[J]. Journal of Marine Science and Application, 2016, 15(4): 426-432.

[10] Jiang Y, Wang X G, Zou Z J, et al. Identification of coupled response models for ship steering and roll motion using support vector machines[J]. Applied Ocean Research, 2021, 110: 102607.

[11] Chen L, Yang P, Li S, et al. Grey-box identification modeling of ship maneuvering motion based on LS-SVM[J]. Ocean Engineering, 2022, 266: 112957.

[12] Jiang Y, Hou X R, Wang X G, et al. Identification modeling and prediction of ship maneuvering motion based on LSTM deep neural network[J]. Journal of Marine Science and Technology, 2022, 27: 125-137.

[13] Raissi M, Karniadakis G E. Hidden physics models: Machine learning of nonlinear partial differential equations[J]. Journal of Computational Physics, 2018, 357: 125-141.

[14] Raissi M, Perdikaris P, Karniadakis G E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations[J]. Journal of Computational Physics, 2019, 378: 686-707.

[15] Raissi M, Yazdani A, Karniadakis G E. Hidden fluid mechanics: learning velocity and pressure fields from flow visualizations[J]. Science, 2020, 367(6481): 1026-1030.

[16] Raissi M, Wang Z, Karniadakis G E, et al. Deep learning of vortex-induced vibrations[J]. Journal of Fluid Mechanics, 2019, 861: 119-137.

[17] Casado M H, Ferreiro R. Identification of the nonlinear ship model parameters based on the turning test trial and the backstepping procedure[J]. Ocean Engineering, 2005, 32: 1350-1369.

[18] Wang S, Wang L, Namkyun I, et al. Real-time Parameter Identification of Ship Maneuvering Response Model Based on Nonlinear Gaussian Filter[J]. Ocean Engineering, 2022, 247: 110471.