

基于数据-物理双驱动的强化学习无人船航迹跟踪控制

白伟伟^{1,2} 雷家虎¹ 林玉芳¹ 李荣辉³

(1. 大连海事大学航海学院, 大连 116026; 2. 广东海洋大学机械工程学院, 湛江 524091;
3. 广东海洋大学船舶与海运学院, 湛江 524091)

摘要: 无人船(Unmanned surface vessels, USVs)作为海洋智慧牧场的核心装备之一,在实际作业过程中需直面海洋环境随机扰动、自身物理约束及动力学参数未知等多重困难。为此,提出一种考虑输入约束与外界环境干扰的数据-物理双驱动强化学习最优控制方法。首先,针对无人船未知动力学系统建模难、传统数据驱动方法可解释性不足的难题,提出一种数据-物理双驱动(Data-physcis dual-driven, DPD)的建模方法,构建融合船舶动力学物理定律与循环神经网络(Recurrent neural networks, RNN)数据学习能力的数据-物理双驱动模型,实现未知复杂动力学特性的精准重构与模型物理可解释性的增强。其次,采用反步法前馈控制(Feedforward control, FC)与最优控制(Optimal control, OC)方法相结合的思路设计控制框架,提出一种基于强化学习的优化控制方法,提升无人船与复杂海洋环境的持续交互能力,达到跟踪精度与控制资源损耗的动态平衡,同时在控制器设计中考虑了船体承载力对控制输入的物理约束。最后,通过 Lyapunov 稳定性理论证明闭环系统内所有信号是一致最终有界的。通过二级海况下的对比仿真试验结果表明,所提方法的跟踪精度较物理机理驱动(Physics-principle-driven, PPD)方法整体提升 42.6%,控制输入成本较 PPD 方法整体降低 26.9%,较传统 FC 整体降低 9.6%,充分验证了方法的有效性和优越性。

关键词: 无人船; 航迹跟踪; 输入约束; 强化学习; 数据-物理双驱动

中图分类号: TP183; TP273 **文献标识码:** A **文章编号:** 1000-1298(2026)19-0054-09

OSID:



Reinforcement Learning-based Trajectory Tracking Control for Unmanned Surface Vessels via Data – Physics Dual-driven Approach

Bai Weiwei^{1,2} Lei Jiahu¹ Lin Yufang¹ Li Ronghui³

(1. Navigation College, Dalian Maritime University, Dalian 116026, China

2. School of Mechanical Engineering, Guangdong Ocean University, Zhanjiang 524091, China

3. Naval Architecture and Shipping College, Guangdong Ocean University, Zhanjiang 524091, China)

Abstract: A data – physics dual-driven (DPD) reinforcement learning (RL)-based optimal control method considering input constraints and external environmental disturbances was proposed for unmanned surface vessels (USVs) applied in intelligent marine ranching. Firstly, to address the challenges of modeling USVs’ unknown dynamics and insufficient interpretability of traditional data-driven methods, a DPD model was constructed by integrating the physical laws of ship dynamics with the data learning capability of recurrent neural networks (RNN), enabling accurate reconstruction of unknown complex dynamic characteristics and enhancing the model’s physical interpretability. Secondly, a control framework was designed via the combination of backstepping and optimal control methods. An RL-based optimal control strategy was proposed to improve the USVs’ continuous interaction capability with complex marine environments, achieving a dynamic balance between tracking accuracy and control resource consumption. Meanwhile, the physical constraints imposed by the hull’s load-bearing capacity on control inputs were incorporated into the controller design. Thirdly, the uniform ultimate boundedness (UUB) of all signals in the closed-loop system was proved by using Lyapunov stability theory. Finally,

收稿日期: 2026-03-16 修回日期: 2026-03-24

基金项目: 国家自然科学基金项目(52271360、52571405)、兴辽英才计划项目(XLYC2403032)、辽宁省自然科学基金项目(2025-MS-093)和辽宁省科技计划联合计划项目(重点研发计划项目)(2025JH2/101800233)

作者简介: 白伟伟(1989—),男,副教授,博士,主要从事优化控制、强化学习及其在船舶控制的应用研究, E-mail: baiweiwei_dl@163.com

comparative simulation results under Level 2 sea conditions demonstrated that the proposed method improved tracking accuracy by 42.6% compared with the physics-principle-driven (PPD) method, and reduced control input cost by 26.9% relative to the PPD method and 9.6% compared with traditional backstepping control, which fully verified the effectiveness and superiority of the proposed method.

Key words: unmanned surface vessels; trajectory tracking; input constraints; reinforcement learning; data - physics dual-driven

0 引言

无人船作为现代海洋智慧牧场的核心装备之一,承担水质监测、精准投喂与苗种投放等关键任务,依靠自主导航、智能决策等技术优势,不仅能够摆脱渔业作业强度大、劳动力短缺的双重困境,而且有效实现了作业效率与运营成本的双重优化^[1-2]。然而,无人船在海洋智慧牧场的高效作业与跟踪精度、控制效率及环境交互密切相关,三者是衡量无人船智能作业能力的核心维度,直接决定现代渔业转型升级过程中牧场的作业效率与运营成本^[3-4]。

针对无人船航迹跟踪问题,已有相关的研究成果^[5-7],然而无人船在智慧海洋牧场作业期间,既要直面风浪流随机扰动、水动力耦合效应等外部困难,又需应对控制输入受船体承载能力限制形成的严格物理约束,这导致传统方法难以实现跟踪精度与控制代价的动态平衡^[8]。针对上述问题,郭子杰等^[9]提出一种兼顾系统控制性能和节能效应的最优控制方法,通过构建多维度指标的代价函数,并动态调整策略,以实现控制代价的最小化。在此基础之上,Wang 等^[10]进一步引入强化学习技术,借助智能体与环境的持续交互迭代优化控制策略,能够适应海洋环境中参数时变、扰动未知的复杂场景,为求解最优反馈控制输入提供了一种全新思路。由此可见,基于强化学习的最优控制方法能进一步实现跟踪精度、控制效率以及环境交互能力的协同提升。

然而,现有最优控制方法仍存在一定的局限,主要表现为依赖物理机理驱动或数据驱动单一建模思路^[11]。其中,基于物理机理模型设计无人船航迹跟踪控制器,虽然具备理论可解释性,但需通过复杂的水池试验获取水动力参数,成本高且难以适配复杂海况下的未建模动态^[12-13]。而数据驱动方法借助神经网络等模型重构系统特性,无需依赖物理参数,但存在可解释性弱、泛化能力有限等不足^[14-15]。为破解上述难题,融合物理机理与数据驱动的动力学系统建模思路成为研究热点,数据-物理双驱动方法既能够利用船舶动力学方程嵌入先验知识,降低数据依赖与模型复杂度,又可通过数据驱动模块补偿未建模动态,显著提升模型适配性^[16-17]。

鉴于此,为解决海洋智慧牧场中具有未知动态

的无人船航迹跟踪问题,本文提出一种兼顾输入约束和外界环境干扰的数据-物理双驱动强化学习最优控制方法。通过稳定性分析以及二级海况下的无人船航迹跟踪仿真试验,验证所提方法在跟踪精度与控制代价方面的综合优势,以期为海洋智慧牧场无人船的高效智能作业提供技术支撑。

1 无人船数学模型

无人船实际水面航行时,受风浪流扰动、水动力耦合效应等影响,运动特性复杂。如图 1 所示,通过建立惯性坐标系 $Oxyz$ 与附体坐标系 $O_bx_by_bz_b$ 描述其空间位置和姿态,其运动状态通常表现为纵荡、横荡、垂荡(直线运动)及纵摇、横摇、艏摇(旋转运动)的六自由度耦合组合。本文主要考虑输入约束下全驱动无人船在 Oxy 平面上的航迹跟踪控制问题,因此可将其运动状态简化为仅包含纵荡、横荡与艏摇的三自由度运动。根据几何学方法构建三自由度运动学模型,具体表达式为

$$\begin{cases} \dot{x} = u \cos \psi - v \sin \psi \\ \dot{y} = u \sin \psi + v \cos \psi \\ \dot{\psi} = r \end{cases} \quad (1)$$

式中 $\dot{x}$ —— x 方向速度, m/s
 $\dot{y}$ —— y 方向速度, m/s
 ψ ——艏摇角, rad
 u ——纵荡速度, m/s
 v ——横荡速度, m/s
 $\dot{\psi}$ 、 r ——艏摇角速度, rad/s

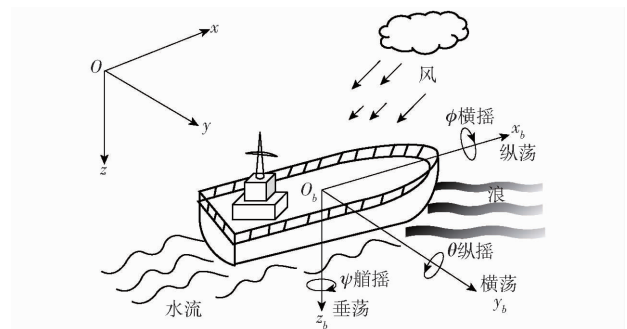


图 1 无人船坐标系

Fig. 1 Coordinate system of USV

根据船舶水动力学定律与动量守恒原理,其三自由度动力学模型表示为

$$\begin{cases} m_u \dot{u} = m_v v r - d_u u + \tau_u + \tau_{wu} \\ m_v \dot{v} = -m_u u r - d_v v + \tau_v + \tau_{wv} \\ m_r \dot{r} = -(m_v - m_u) v u - d_r r + \tau_r + \tau_{wr} \end{cases} \quad (2)$$

其中

$$\begin{aligned} m_u &= m - X_{\dot{u}} & m_v &= m - Y_{\dot{v}} \\ m_r &= I_z - N_{\dot{r}} \\ d_u &= -X_u - X_{u|u|} |u| - X_{uuu} u^2 \\ d_v &= -Y_v - Y_{v|v|} |v| - Y_{vvv} v^2 \\ d_r &= -N_r - N_{r|v|} |v| - N_{r|r|} |r| \end{aligned}$$

式中

- m ——无人船质量, kg
- I_z ——转动惯量, $\text{kg} \cdot \text{m}^2$
- $\dot{u}$ ——纵荡加速度, m/s^2
- $\dot{v}$ ——横荡加速度, m/s^2
- $\dot{r}$ ——艏摇角加速度, rad/s^2
- $d_{\text{}}/$ ——水动力阻尼参数
- $\tau_{w\text{}}/$ ——无人船纵、横、艏向干扰力, N
- $\tau_{\text{}}/$ ——无人船纵、横、艏向控制输入, N
- $X_{\text{}}/$ ——附体坐标 x_a 轴方向的水动力导数
- $Y_{\text{}}/$ ——附体坐标 y_a 轴方向的水动力导数
- $N_{\text{}}/$ ——附体坐标 z_a 轴方向的水动力导数

下标 $\text{}/$ 表示 u, v, r , 下标 $\text{\#}$ 表示求导自变量, 包含线性项与非线性耦合项。

各自由度控制输入 $\tau_{\text{}}/$ 需考虑实际输入约束, 其定义为

$$\tau_{\text{}}/ = \begin{cases} \bar{\tau}_{\text{}}/ & (a_{\text{}}/ \geq \bar{\tau}_{\text{}}/) \\ a_{\text{}}/ & (\underline{\tau}_{\text{}}/ < a_{\text{}}/ < \bar{\tau}_{\text{}}/) \\ \underline{\tau}_{\text{}}/ & (a_{\text{}}/ \leq \underline{\tau}_{\text{}}/) \end{cases} \quad (3)$$

- 式中
- $a_{\text{}}/$ ——无人船纵、横、艏向理想控制输入, N
 - $\bar{\tau}_{\text{}}/$ ——无人船各自由度控制输入最大值, N
 - $\underline{\tau}_{\text{}}/$ ——无人船各自由度控制输入最小值, N

为便于后续控制器设计与稳定性分析, 定义 $\boldsymbol{a} = [a_u, a_v, a_r]^T$, 并将全驱动无人船的数学模型简化为

$$\begin{cases} \dot{\boldsymbol{\eta}} = \boldsymbol{J}(\boldsymbol{\psi}) \boldsymbol{\nu} \\ \boldsymbol{M} \dot{\boldsymbol{\nu}} + \boldsymbol{C}(\boldsymbol{\nu}) \boldsymbol{\nu} + \boldsymbol{D} \boldsymbol{\nu} = \boldsymbol{\tau} + \boldsymbol{\tau}_w \end{cases} \quad (4)$$

其中

$$\begin{aligned} \boldsymbol{\eta} &= [x, y, \psi]^T & \boldsymbol{\nu} &= [u, v, r]^T \\ \boldsymbol{\tau} &= [\tau_u, \tau_v, \tau_r]^T & \boldsymbol{\tau}_w &= [\tau_{wu}, \tau_{wv}, \tau_{wr}]^T \\ \boldsymbol{J}(\boldsymbol{\psi}) &= \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ \boldsymbol{C}(\boldsymbol{\nu}) &= \begin{bmatrix} 0 & 0 & -m_v v \\ 0 & 0 & m_u u \\ m_v v & -m_u u & 0 \end{bmatrix} \\ \boldsymbol{M} &= \begin{bmatrix} m_u & 0 & 0 \\ 0 & m_v & 0 \\ 0 & 0 & m_r \end{bmatrix} & \boldsymbol{D} &= \begin{bmatrix} d_u & 0 & 0 \\ 0 & d_v & 0 \\ 0 & 0 & d_r \end{bmatrix} \end{aligned}$$

本文的控制目标是: 设计一种考虑输入约束的数据-物理双驱动强化学习无人船最优航迹跟踪控制器, 需同时满足:

- (1) 确保闭环系统所有信号一致最终有界。
- (2) 在系统输出 $\boldsymbol{\eta} = [x, y, \psi]^T$ 跟踪参考信号 $\boldsymbol{\eta}_d = [x_d, y_d, \psi_d]^T$ 的前提下, 最小化代价函数 C 以实现最优控制。

为实现控制目标, 本文引入以下假设:

- (1) 参考信号 $\boldsymbol{\eta}_d$ 以及其关于时间 t 的一阶导数 $\dot{\boldsymbol{\eta}}_d$ 、二阶导数 $\ddot{\boldsymbol{\eta}}_d$ 均有界。
- (2) 外部干扰 $\boldsymbol{\tau}_w$ 有界, 即存在正常数 $\bar{\boldsymbol{\tau}}_w$, 使得 $\|\boldsymbol{\tau}_w\| \leq \bar{\boldsymbol{\tau}}_w$ 。

2 无人船的数据-物理双驱动模型

由式(2)可知, 物理机理建模需通过大量高成本的水池试验获取精准参数, 实施难度较高; 现有数据驱动无需这些参数即可设计控制器, 但存在可解释性弱、泛化能力有限等缺陷。为此, 本文提出一种基于 RNN 的数据-物理双驱动模型, 融合船舶动力学物理定律与 RNN 数据学习能力, 实现未知动力学精准重构的同时, 弥补数据驱动技术的固有局限性。

2.1 基于 RNN 的数据-物理双驱动建模

此时, 不考虑输入约束的无人船动力学方程的标称形式可表示为

$$\dot{\boldsymbol{\nu}} = -\boldsymbol{M}^{-1}(\boldsymbol{C}(\boldsymbol{\nu}) + \boldsymbol{D}) \boldsymbol{\nu} + \boldsymbol{M}^{-1} \boldsymbol{a} \quad (5)$$

结合文献[15]的方法, 可将上述物理方程重构为如下 RNN 模型

$$\dot{\boldsymbol{\nu}} = \boldsymbol{\kappa}_1^{*T} \boldsymbol{\nu} + \boldsymbol{\kappa}_2^{*T} \boldsymbol{\varphi}(\boldsymbol{\nu}) + \boldsymbol{\kappa}_3^{*T} + \boldsymbol{\kappa}_4^{*T} \boldsymbol{a} + \boldsymbol{\varepsilon} \quad (6)$$

其中 $\boldsymbol{\varphi}(\boldsymbol{\nu}) = [\varphi_1(\boldsymbol{\nu}), \varphi_2(\boldsymbol{\nu}), \varphi_3(\boldsymbol{\nu})]^T$

式中 $\varphi_i(\boldsymbol{\nu})$ ——RNN 基函数, $i = 1, 2, 3$

$\boldsymbol{\kappa}_j^*$ ——RNN 理想权值, $j = 1, 2, 3, 4$

$\boldsymbol{\varepsilon}$ ——重构误差

本文选取 $\varphi_i(\boldsymbol{\nu}) = \tanh(\boldsymbol{\nu})$, 且其满足

$$|\varphi_i(x_1) - \varphi_i(x_2)| \leq b |x_1 - x_2| \quad (7)$$

式中 b ——正常数

x_1, x_2 ——基函数定义域内的自变量

定义 RNN 权值误差 $\tilde{\boldsymbol{\kappa}}_j$ 为

$$\tilde{\boldsymbol{\kappa}}_j = \boldsymbol{\kappa}_j^* - \hat{\boldsymbol{\kappa}}_j \quad (8)$$

式中 $\hat{\boldsymbol{\kappa}}_j$ ——RNN 权值估计值

定义无人船建模误差 $\tilde{\boldsymbol{\nu}}$ 为

$$\tilde{\boldsymbol{\nu}} = \boldsymbol{\nu} - \hat{\boldsymbol{\nu}} = [\tilde{u}, \tilde{v}, \tilde{r}]^T \quad (9)$$

式中 $\hat{\boldsymbol{\nu}}$ ——无人船 RNN 模型估计值

$\tilde{u}, \tilde{v}, \tilde{r}$ ——纵荡、横荡、艏摇速度方向的建模误差

基于式(8)、(9), 对式(6)对应的 RNN 模型进

行估计,可得

$$\hat{\mathbf{v}} = \hat{\mathbf{\kappa}}_1^T \hat{\mathbf{v}} + \hat{\mathbf{\kappa}}_2^T \boldsymbol{\varphi}(\hat{\mathbf{v}}) + \hat{\mathbf{\kappa}}_3^T + \hat{\mathbf{\kappa}}_4^T \mathbf{a} - \boldsymbol{\xi} \quad (10)$$

其中

$$\boldsymbol{\xi} = P\tilde{\mathbf{v}} + \frac{\delta\tilde{\mathbf{v}}}{\tilde{\mathbf{v}}^T \tilde{\mathbf{v}} + 1}$$

式中 $\boldsymbol{\xi}$ ——重构误差补偿反馈项

P ——设计的线性反馈项增益系数

δ ——非线性反馈项的增益参数估计值

重构误差补偿反馈项主要由线性项和非线性项协同实现误差补偿,两者结合既保证了误差补偿的快速性,又提升了补偿的稳定性与精度,有效修正 RNN 模型与实际系统间的偏差。

定义非线性反馈项的增益参数误差 $\tilde{\delta}$ 为

$$\tilde{\delta} = \delta^* - \hat{\delta} \quad (11)$$

式中 δ^* ——理想非线性反馈项的增益参数

RNN 权值和非线性反馈项的增益参数估计值更新率设计为

$$\begin{cases} \dot{\hat{\mathbf{\kappa}}}_j = -\dot{\tilde{\mathbf{\kappa}}}_j = \Gamma_j \Delta_j \tilde{\mathbf{v}}^T \\ \dot{\hat{\delta}} = -\dot{\tilde{\delta}} = -Q \frac{\tilde{\mathbf{v}}^T \tilde{\mathbf{v}}}{\tilde{\mathbf{v}}^T \tilde{\mathbf{v}} + 1} \end{cases} \quad (12)$$

其中

$$\Delta_j = \begin{cases} \hat{\mathbf{v}} & (j=1) \\ \boldsymbol{\varphi}(\hat{\mathbf{v}}) & (j=2) \\ 1 & (j=3) \\ \mathbf{a} & (j=4) \end{cases}$$

式中 Γ_j ——预设的正定矩阵

Q ——预设的正常数

对式(9)进行求导,可得

$$\begin{aligned} \dot{\tilde{\mathbf{v}}} &= \mathbf{\kappa}_1^{*T} \tilde{\mathbf{v}} + \hat{\mathbf{\kappa}}_1^T \hat{\mathbf{v}} + \mathbf{\kappa}_2^{*T} \tilde{\boldsymbol{\varphi}}(\tilde{\mathbf{v}}) + \boldsymbol{\varepsilon} + \hat{\mathbf{\kappa}}_2^T \boldsymbol{\varphi}(\hat{\mathbf{v}}) + \\ &\quad \hat{\mathbf{\kappa}}_3^T + \hat{\mathbf{\kappa}}_4^T \mathbf{a} + P\tilde{\mathbf{v}} + \frac{\delta^* \tilde{\mathbf{v}}}{\tilde{\mathbf{v}}^T \tilde{\mathbf{v}} + 1} - \frac{\delta\tilde{\mathbf{v}}}{\tilde{\mathbf{v}}^T \tilde{\mathbf{v}} + 1} \end{aligned} \quad (13)$$

其中

$$\tilde{\boldsymbol{\varphi}}(\tilde{\mathbf{v}}) = \boldsymbol{\varphi}(\mathbf{v}) - \boldsymbol{\varphi}(\hat{\mathbf{v}})$$

2.2 稳定性证明

选取 Lyapunov 函数证明无人船的数据-物理双驱动模型的稳定性,即

$$V_1 = \frac{1}{2} \tilde{\mathbf{v}}^T \tilde{\mathbf{v}} + \frac{1}{2} \text{tr} \left(\sum_j^4 \Gamma_j^{-1} \tilde{\mathbf{\kappa}}_j^T \tilde{\mathbf{\kappa}}_j \right) + \frac{1}{2Q} \tilde{\delta}^2 \quad (14)$$

其中, tr 为矩阵迹的求解符号。

对 V_1 求导,并将式(12)、(13)代入,可得

$$\begin{aligned} \dot{V}_1 &= \tilde{\mathbf{v}}^T \mathbf{\kappa}_1^{*T} \tilde{\mathbf{v}} + \tilde{\mathbf{v}}^T \mathbf{\kappa}_2^{*T} \tilde{\boldsymbol{\varphi}}(\tilde{\mathbf{v}}) + \tilde{\mathbf{v}}^T \boldsymbol{\varepsilon} + \\ &\quad \tilde{\mathbf{v}}^T P \tilde{\mathbf{v}} + \frac{\delta^* \tilde{\mathbf{v}}^T \tilde{\mathbf{v}}}{\tilde{\mathbf{v}}^T \tilde{\mathbf{v}} + 1} \end{aligned} \quad (15)$$

根据文献[15],假设重构误差向量 $\boldsymbol{\varepsilon}$ 满足 $\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} \leq \delta^* \tilde{\mathbf{v}}^T \tilde{\mathbf{v}}$ 。结合式(7)和 Young's 不等式,式(15)右边第 2、3 项可变换为

$$\begin{cases} \tilde{\mathbf{v}}^T \mathbf{\kappa}_2^{*T} \tilde{\boldsymbol{\varphi}}(\tilde{\mathbf{v}}) \leq \frac{1}{2} \tilde{\mathbf{v}}^T \mathbf{\kappa}_2^{*T} \mathbf{\kappa}_2^* \tilde{\mathbf{v}} + \frac{1}{2} b^2 \tilde{\mathbf{v}}^T \tilde{\mathbf{v}} \\ \tilde{\mathbf{v}}^T \boldsymbol{\varepsilon} \leq \frac{1}{2} \tilde{\mathbf{v}}^T \tilde{\mathbf{v}} + \frac{1}{2} \delta^* \tilde{\mathbf{v}}^T \tilde{\mathbf{v}} \end{cases} \quad (16)$$

式(15)转换为

$$\begin{aligned} \dot{V}_1 &\leq \frac{1}{2} \tilde{\mathbf{v}}^T [2\mathbf{\kappa}_1^{*T} + \mathbf{\kappa}_2^{*T} \mathbf{\kappa}_2^* + \\ &\quad (b^2 + 1 + 3\delta^* + 2P) \mathbf{I}_{3 \times 3}] \tilde{\mathbf{v}} \end{aligned} \quad (17)$$

式中 $\mathbf{I}_{3 \times 3}$ —— 3×3 的单位矩阵

通过合理选取增益参数 P ,可使 $\dot{V}_1 < 0$ 。最终,只考虑外界环境干扰的无人船数据-物理双驱动模型可表示为

$$\begin{cases} \dot{\boldsymbol{\eta}} = \mathbf{J}(\boldsymbol{\psi}) \hat{\mathbf{v}} \\ \dot{\hat{\mathbf{v}}} = \hat{\mathbf{\kappa}}_1^T \hat{\mathbf{v}} + \hat{\mathbf{\kappa}}_2^T \boldsymbol{\varphi}(\hat{\mathbf{v}}) + \hat{\mathbf{\kappa}}_3^T - \boldsymbol{\xi} + \hat{\mathbf{\kappa}}_4^T (\mathbf{a} + \boldsymbol{\tau}_w) \end{cases} \quad (18)$$

3 航迹跟踪控制输入设计

基于上述数据-物理双驱动模型,首先采用 Lyapunov 反步法设计前馈控制输入 $\mathbf{a}_f$ 。然后,采用强化学习方法设计出最优反馈控制输入 $\mathbf{a}_o^*$ 。最终,总控制输入为 $\mathbf{a} = \mathbf{a}_f + \mathbf{a}_o^*$ 。

3.1 前馈控制输入设计

(1) 定义跟踪误差 $\boldsymbol{\eta}_e$ 为

$$\boldsymbol{\eta}_e = \boldsymbol{\eta} - \boldsymbol{\eta}_d \quad (19)$$

对式(19)求导,结合式(18)可得

$$\dot{\boldsymbol{\eta}}_e = \mathbf{J}(\boldsymbol{\psi}) \hat{\mathbf{v}} - \dot{\boldsymbol{\eta}}_d \quad (20)$$

设计前馈虚拟控制输入 $\mathbf{v}_d$ 为

$$\mathbf{v}_d = \mathbf{J}^{-1}(\boldsymbol{\psi}) (-c_1 \boldsymbol{\eta}_e + \dot{\boldsymbol{\eta}}_d) \quad (21)$$

式中 c_1 ——设计参数

(2) 定义速度误差 $\mathbf{v}_e$ 为

$$\mathbf{v}_e = \hat{\mathbf{v}} - \mathbf{v}_d \quad (22)$$

结合式(18)、(20),对式(22)求导得

$$\begin{aligned} \dot{\mathbf{v}}_e &= \mathbf{G}(\mathbf{v}_e) + \mathbf{g}(\mathbf{v}_d) + \hat{\mathbf{\kappa}}_4^T (\mathbf{a}_f + \mathbf{a}_o^*) + \\ &\quad \hat{\mathbf{\kappa}}_4^T \boldsymbol{\tau}_w - \dot{\mathbf{v}}_d \end{aligned} \quad (23)$$

其中

$$\begin{aligned} \mathbf{G}(\mathbf{v}_e) &= \mathbf{g}(\hat{\mathbf{v}}) - \mathbf{g}(\mathbf{v}_d) \\ \mathbf{g}(\hat{\mathbf{v}}) &= \hat{\mathbf{\kappa}}_1^T \hat{\mathbf{v}} + \hat{\mathbf{\kappa}}_2^T \boldsymbol{\varphi}(\hat{\mathbf{v}}) + \hat{\mathbf{\kappa}}_3^T - \boldsymbol{\xi} \\ \mathbf{g}(\mathbf{v}_d) &= \hat{\mathbf{\kappa}}_1^T \mathbf{v}_d + \hat{\mathbf{\kappa}}_2^T \boldsymbol{\varphi}(\mathbf{v}_d) + \hat{\mathbf{\kappa}}_3^T - \boldsymbol{\xi} \end{aligned}$$

式中 $\dot{\mathbf{v}}_d$ ——虚拟控制输入的导数

结合式(22),设计前馈控制输入为

$$\mathbf{a}_f = \frac{1}{\hat{\mathbf{\kappa}}_4^T} (-c_2 \mathbf{v}_e - \mathbf{J}(\boldsymbol{\psi}) \boldsymbol{\eta}_e - \mathbf{g}(\mathbf{v}_d) + \dot{\mathbf{v}}_d) \quad (24)$$

式中 c_2 ——设计参数

构造 Lyapunov 函数

$$V_2 = \frac{1}{2} \boldsymbol{\eta}_e^T + \frac{1}{2} \mathbf{v}_e^T \mathbf{v}_e \quad (25)$$

对 V_2 求导,代入式(19)、(22)、(23)得

$$\dot{V}_2 = -c_1 \boldsymbol{\eta}_e^2 - c_2 \boldsymbol{v}_e^2 + \boldsymbol{v}_e^T (\boldsymbol{G}(\boldsymbol{v}_e) + \hat{\boldsymbol{\kappa}}_4^T \boldsymbol{a}_o^*) + \boldsymbol{v}_e^T \hat{\boldsymbol{\kappa}}_4^T \boldsymbol{\tau}_w \quad (26)$$

根据第 1.1 节的假设 2, 结合 Young's 不等式, 式(26)右侧第 4 项可放缩为

$$\boldsymbol{v}_e^T \hat{\boldsymbol{\kappa}}_4^T \boldsymbol{\tau}_w \leq \frac{1}{2} \boldsymbol{v}_e^2 + \frac{1}{2} \|\hat{\boldsymbol{\kappa}}_4^T\|^2 \|\boldsymbol{\tau}_w\|^2 \quad (27)$$

将式(27)代入式(26)得

$$\dot{V}_2 \leq -c_1 \boldsymbol{\eta}_e^2 - \left(c_2 - \frac{1}{2}\right) \boldsymbol{v}_e^2 + \boldsymbol{v}_e^T (\boldsymbol{G}(\boldsymbol{v}_e) + \hat{\boldsymbol{\kappa}}_4^T \boldsymbol{a}_o^*) + \frac{1}{2} \|\hat{\boldsymbol{\kappa}}_4^T\|^2 \|\boldsymbol{\tau}_w\|^2 \quad (28)$$

由式(28)可知, 仅依靠前馈控制输入 $\boldsymbol{a}_f$ 无法确保无人船系统的稳定性。因此, 需进一步设计最优反馈控制输入 $\boldsymbol{a}_o^*$, 使总控制输入 $\boldsymbol{a}$ 满足系统稳定性要求。

3.2 最优反馈控制输入设计

设计最优反馈控制输入 $\boldsymbol{a}_o^*$, 需同时实现两个目标: 使如下误差系统稳定和最小化代价函数

$$\dot{\boldsymbol{v}}_e = \boldsymbol{G}(\boldsymbol{v}_e) + \hat{\boldsymbol{\kappa}}_4^T \boldsymbol{a}_o^* \quad (29)$$

为此, 定义系统式(29)的代价函数为

$$C(\boldsymbol{v}_e) = \int_t^\infty U(\boldsymbol{v}_e(s), \boldsymbol{a}_o(s)) ds \quad (30)$$

其中 $U(\boldsymbol{v}_e, \boldsymbol{a}_o) = \boldsymbol{v}_e^T \boldsymbol{\Theta} \boldsymbol{v}_e + \boldsymbol{a}_o^T \boldsymbol{\Lambda} \boldsymbol{a}_o$

式中 $U(\boldsymbol{v}_e, \boldsymbol{a}_o)$ ——基本效用函数

$\boldsymbol{\Theta}, \boldsymbol{\Lambda}$ ——预设的设计参数

$\boldsymbol{a}_o$ ——容许的控制输入

s ——积分变量

根据 Bellman 最优性原理, 最优代价函数可表示为

$$C^*(\boldsymbol{v}_e) = \min_{\boldsymbol{a}_o} C(\boldsymbol{v}_e) \quad (31)$$

若 $C^*(\boldsymbol{v}_e)$ 是连续可微, 可得到 Lyapunov 方程 $U(\boldsymbol{v}_e, \boldsymbol{a}_o) + \nabla_{\boldsymbol{v}_e}^T C^*(\boldsymbol{v}_e) (\boldsymbol{G}(\boldsymbol{v}_e) + \hat{\boldsymbol{\kappa}}_4^T \boldsymbol{a}_o) = 0$ 。其中, $\nabla_{\boldsymbol{v}_e} C^*(\boldsymbol{v}_e)$ 是 $C^*(\boldsymbol{v}_e)$ 关于 $\boldsymbol{v}_e$ 的偏导数。

基于 $\nabla_{\boldsymbol{v}_e} C^*(\boldsymbol{v}_e)$, 定义 Hamilton 函数为

$$H(\boldsymbol{v}_e, \boldsymbol{a}_o, \nabla_{\boldsymbol{v}_e} C^*(\boldsymbol{v}_e)) = U(\boldsymbol{v}_e, \boldsymbol{a}_o) + \nabla_{\boldsymbol{v}_e}^T C^*(\boldsymbol{v}_e) (\boldsymbol{G}(\boldsymbol{v}_e) + \hat{\boldsymbol{\kappa}}_4^T \boldsymbol{a}_o) \quad (32)$$

通过求解 Hamilton 函数关于容许的控制输入 $\boldsymbol{a}_o$ 的偏导数为零 $\left(\frac{\partial H}{\partial \boldsymbol{a}_o} = 0\right)$, 其中 H 为 $H(\boldsymbol{v}_e, \boldsymbol{a}_o, \nabla_{\boldsymbol{v}_e} C^*(\boldsymbol{v}_e))$ 简写, 解得最优反馈控制输入

$$\boldsymbol{a}_o^* = -\frac{1}{2} \boldsymbol{\Lambda}^{-1} \hat{\boldsymbol{\kappa}}_4^T \nabla_{\boldsymbol{v}_e} C^*(\boldsymbol{v}_e) \quad (33)$$

将式(33)代入式(32), 可得 HJB 方程

$$H(\boldsymbol{v}_e, \boldsymbol{a}_o^*, \nabla_{\boldsymbol{v}_e} C^*(\boldsymbol{v}_e)) = U(\boldsymbol{v}_e, \boldsymbol{a}_o^*) + \nabla_{\boldsymbol{v}_e}^T C^*(\boldsymbol{v}_e) (\boldsymbol{G}(\boldsymbol{v}_e) + \hat{\boldsymbol{\kappa}}_4^T \boldsymbol{a}_o^*) \quad (34)$$

由于 HJB 方程固有的非线性特性, 其解析解难以直接求解。因此, 通过评价网络来逼近代价函数, 从而实现近似最优反馈控制输入的设计, 其表达式为

$$C^*(\boldsymbol{v}_e) = \boldsymbol{\kappa}_c^{*T} \boldsymbol{\varphi}_c(\boldsymbol{v}_e) + \varepsilon_c(\boldsymbol{v}_e) \quad (35)$$

其中

$$\begin{aligned} \boldsymbol{\varphi}_c(\boldsymbol{v}_e) &= [\varphi_1(\boldsymbol{v}_e), \varphi_2(\boldsymbol{v}_e), \varphi_3(\boldsymbol{v}_e), \varphi_4(\boldsymbol{v}_e), \varphi_5(\boldsymbol{v}_e)]^T \\ \varphi_k(\boldsymbol{v}_e) &= \exp\left(-\frac{\|\boldsymbol{v}_e - \boldsymbol{\mu}_k\|^2}{2\omega^2}\right) \end{aligned}$$

式中 $\boldsymbol{\kappa}_c^*$ ——评价网络理想权值

$\boldsymbol{\mu}_k$ ——基函数的中心值, $k=1, 2, 3, 4, 5$

$\varphi_k(\boldsymbol{v}_e)$ ——评价网络基函数

ω ——基函数的宽度

$\varepsilon_c(\boldsymbol{v}_e)$ ——逼近误差

则最优代价函数的梯度为

$$\nabla_{\boldsymbol{v}_e} C^*(\boldsymbol{v}_e) = \nabla_{\boldsymbol{v}_e}^T \boldsymbol{\varphi}_c(\boldsymbol{v}_e) \boldsymbol{\kappa}_c^* + \nabla_{\boldsymbol{v}_e} \varepsilon_c(\boldsymbol{v}_e) \quad (36)$$

将式(36)分别代入(33)、(34)得

$$\begin{cases} \boldsymbol{a}_o^* = -\frac{1}{2} \boldsymbol{\Lambda}^{-1} \hat{\boldsymbol{\kappa}}_4^T (\nabla_{\boldsymbol{v}_e}^T \boldsymbol{\varphi}_c(\boldsymbol{v}_e) \boldsymbol{\kappa}_c^* + \nabla_{\boldsymbol{v}_e} \varepsilon_c(\boldsymbol{v}_e)) \\ H = \boldsymbol{v}_e^T \boldsymbol{\Theta} \boldsymbol{v}_e - \frac{1}{4} \boldsymbol{\kappa}_c^{*T} \nabla_{\boldsymbol{v}_e} \boldsymbol{\varphi}_c \boldsymbol{E} \nabla_{\boldsymbol{v}_e}^T \boldsymbol{\varphi}_c \boldsymbol{\kappa}_c^* + \hat{\boldsymbol{\kappa}}_4^T \nabla_{\boldsymbol{v}_e} \boldsymbol{\varphi}_c \boldsymbol{G}(\boldsymbol{v}_e) + \varepsilon_{HJB} = 0 \end{cases} \quad (37)$$

其中 $\varepsilon_{HJB} = \nabla_{\boldsymbol{v}_e}^T \varepsilon_c (\boldsymbol{G}(\boldsymbol{v}_e) + \hat{\boldsymbol{\kappa}}_4^T \boldsymbol{a}_o^* + \frac{1}{4} \boldsymbol{E} \nabla_{\boldsymbol{v}_e} \varepsilon_c)$

$$\boldsymbol{E} = \hat{\boldsymbol{\kappa}}_4^T \boldsymbol{\Lambda}^{-1} \hat{\boldsymbol{\kappa}}_4$$

定义评价网络权值误差 $\tilde{\boldsymbol{\kappa}}_c$ 为

$$\tilde{\boldsymbol{\kappa}}_c = \boldsymbol{\kappa}_c^* - \hat{\boldsymbol{\kappa}}_c \quad (38)$$

式中 $\hat{\boldsymbol{\kappa}}_c$ ——评价网络权值估计值

利用评价网络对代价函数的梯度进行估计, 则有

$$\nabla_{\boldsymbol{v}_e} \hat{C}^*(\boldsymbol{v}_e) = \nabla_{\boldsymbol{v}_e}^T \boldsymbol{\varphi}_c(\boldsymbol{v}_e) \hat{\boldsymbol{\kappa}}_c \quad (39)$$

则最优反馈控制输入和 HJB 函数的估计值 $\hat{\boldsymbol{a}}_o^*$ 、 $\hat{H}$ 分别为

$$\begin{cases} \hat{\boldsymbol{a}}_o^* = -\frac{1}{2} \boldsymbol{\Lambda}^{-1} \hat{\boldsymbol{\kappa}}_4^T \nabla_{\boldsymbol{v}_e}^T \boldsymbol{\varphi}_c(\boldsymbol{v}_e) \hat{\boldsymbol{\kappa}}_c \\ \hat{H} = \boldsymbol{v}_e^T \boldsymbol{\Theta} \boldsymbol{v}_e - \frac{1}{4} \hat{\boldsymbol{\kappa}}_c^T \nabla_{\boldsymbol{v}_e} \boldsymbol{\varphi}_c \boldsymbol{E} \nabla_{\boldsymbol{v}_e}^T \boldsymbol{\varphi}_c \hat{\boldsymbol{\kappa}}_c + \hat{\boldsymbol{\kappa}}_4^T \nabla_{\boldsymbol{v}_e} \boldsymbol{\varphi}_c \boldsymbol{G}(\boldsymbol{v}_e) \end{cases} \quad (40)$$

为使 $\hat{H}$ 最小, 利用梯度下降法设计 $\tilde{\boldsymbol{\kappa}}_c$ 得

$$\tilde{\boldsymbol{\kappa}}_c = \beta \hat{H} \left[\nabla_{\boldsymbol{v}_e} \boldsymbol{\varphi}_c \left(\frac{\boldsymbol{E} \nabla_{\boldsymbol{v}_e}^T \boldsymbol{\varphi}_c \hat{\boldsymbol{\kappa}}_c}{2} - \boldsymbol{G}(\boldsymbol{v}_e) \right) \right] \quad (41)$$

式中 β ——评价网络学习率

将式(40)代入式(29), 可得近似误差系统

$$\dot{\boldsymbol{v}}_e = \boldsymbol{G}(\boldsymbol{v}_e) + \hat{\boldsymbol{\kappa}}_4^T \hat{\boldsymbol{a}}_o^* \quad (42)$$

根据式(37)可知

$$\boldsymbol{\nu}_e^T \boldsymbol{\Theta} \boldsymbol{\nu}_e = \boldsymbol{\kappa}_c^{*T} \nabla_{\boldsymbol{\nu}_e} \boldsymbol{\varphi}_c \times \left(\frac{\boldsymbol{E} \nabla_{\boldsymbol{\nu}_e}^T \boldsymbol{\varphi}_c \boldsymbol{\kappa}_c^*}{4} - \boldsymbol{G}(\boldsymbol{\nu}_e) \right) - \boldsymbol{\varepsilon}_{HJB}$$

将其代入式(40),则评价网络误差的导数为

$$\begin{aligned} \dot{\tilde{\boldsymbol{\kappa}}}_c = -\dot{\hat{\boldsymbol{\kappa}}}_c = -\beta & \left[\nabla_{\boldsymbol{\nu}_e} \boldsymbol{\varphi}_c \left(\dot{\boldsymbol{\nu}}_e + \frac{\boldsymbol{E} \nabla_{\boldsymbol{\nu}_e} \boldsymbol{\varepsilon}_c}{2} + \right. \right. \\ & \left. \left. \frac{\boldsymbol{E} \nabla_{\boldsymbol{\nu}_e}^T \boldsymbol{\varphi}_c \tilde{\boldsymbol{\kappa}}_c}{2} \right) \right] \left[\tilde{\boldsymbol{\kappa}}_c^T \nabla_{\boldsymbol{\nu}_e} \boldsymbol{\varphi}_c \left(\dot{\boldsymbol{\nu}}_e + \frac{\boldsymbol{E} \nabla_{\boldsymbol{\nu}_e} \boldsymbol{\varepsilon}_c}{2} \right) + \right. \\ & \left. \frac{1}{4} \tilde{\boldsymbol{\kappa}}_c^T \nabla_{\boldsymbol{\nu}_e} \boldsymbol{\varphi}_c \boldsymbol{E} \nabla_{\boldsymbol{\nu}_e}^T \boldsymbol{\varphi}_c \tilde{\boldsymbol{\kappa}}_c + \boldsymbol{\varepsilon}_{HJB} \right] \end{aligned} \quad (43)$$

3.3 稳定性证明

在分析前馈控制输入稳定性的基础上,构造 Lyapunov 函数以证明整个闭环系统稳定,即

$$V = V_2 + \frac{1}{2\beta} \tilde{\boldsymbol{\kappa}}_c^T \tilde{\boldsymbol{\kappa}}_c \quad (44)$$

对 V 求导,并将式(28)、(29)、(43)代入,可得

$$\begin{aligned} \dot{V} \leq & -c_1 \boldsymbol{\eta}_e^2 - \left(c_2 - \frac{1}{2} \right) \boldsymbol{\nu}_e^2 + \boldsymbol{\nu}_e^T \dot{\boldsymbol{\nu}}_e + \\ & \frac{1}{2} \|\tilde{\boldsymbol{\kappa}}_4^T\|^2 \|\boldsymbol{\tau}_w\|^2 - \tilde{\boldsymbol{\kappa}}_c^T \nabla_{\boldsymbol{\nu}_e} \boldsymbol{\varphi}_c \times \\ & \left(\dot{\boldsymbol{\nu}}_e + \frac{\boldsymbol{E} \nabla_{\boldsymbol{\nu}_e} \boldsymbol{\varepsilon}_c}{2} + \frac{\boldsymbol{E} \nabla_{\boldsymbol{\nu}_e}^T \boldsymbol{\varphi}_c \tilde{\boldsymbol{\kappa}}_c}{2} \right) \times \\ & \left[\tilde{\boldsymbol{\kappa}}_c^T \nabla_{\boldsymbol{\nu}_e} \boldsymbol{\varphi}_c \left(\dot{\boldsymbol{\nu}}_e + \frac{\boldsymbol{E} \nabla_{\boldsymbol{\nu}_e} \boldsymbol{\varepsilon}_c}{2} + \frac{\boldsymbol{E} \nabla_{\boldsymbol{\nu}_e}^T \boldsymbol{\varphi}_c \tilde{\boldsymbol{\kappa}}_c}{4} \right) + \boldsymbol{\varepsilon}_{HJB} \right] \end{aligned} \quad (45)$$

根据文献[9],假设: $\|\nabla_{\boldsymbol{\nu}_e} \boldsymbol{\varphi}_c\| \leq \Phi_{\varphi_c}$ 、 $\|\nabla_{\boldsymbol{\nu}_e} \boldsymbol{\varepsilon}_c\| \leq \Phi_{\varepsilon_c}$ 、 $\|\boldsymbol{\varepsilon}_{HJB}\| \leq \Phi_{\varepsilon_{HJB}}$ 、 $\|\tilde{\boldsymbol{\kappa}}_4^T\| \leq \Phi_{\tilde{\kappa}_4}$ 、 $\|\dot{\boldsymbol{\nu}}_e\| \leq b\sqrt{\|\boldsymbol{\nu}_e\|}$ 和 $\|\boldsymbol{E}\| \leq \Phi_{\tilde{\kappa}_4}^2 \|\boldsymbol{\Lambda}^{-1}\| \leq \Phi_E$, 其中 Φ_{φ_c} 、 Φ_{ε_c} 、 $\Phi_{\varepsilon_{HJB}}$ 、 $\Phi_{\tilde{\kappa}_4}$ 、 b 和 Φ_E 都是正参数。应用 Young's 不等式,把式(45)各项展开,以其中一项为例说明放缩过程

$$\begin{aligned} & \left(\tilde{\boldsymbol{\kappa}}_c^T \nabla_{\boldsymbol{\nu}_e} \boldsymbol{\varphi}_c \dot{\boldsymbol{\nu}}_e \right) \left(\tilde{\boldsymbol{\kappa}}_c^T \nabla_{\boldsymbol{\nu}_e} \boldsymbol{\varphi}_c \frac{\boldsymbol{E} \nabla_{\boldsymbol{\nu}_e} \boldsymbol{\varepsilon}_c}{2} \right) \leq \\ & \frac{\|\tilde{\boldsymbol{\kappa}}_c\|^4 \Phi_{\varphi_c}^4}{4} + \frac{\|\tilde{\boldsymbol{\kappa}}_c\|^2 \Phi_{\varphi_c}^2 \Phi_E^2 \Phi_{\varepsilon_c}^2}{8} + \\ & \frac{b^4 \|\boldsymbol{\nu}_e\|^2}{8} + \frac{b^2 \|\boldsymbol{\nu}_e\| \Phi_E^2 \Phi_{\varepsilon_c}^2}{8} \end{aligned} \quad (46)$$

其余各项采用相同的方式处理,可得

$$\begin{aligned} \dot{V} \leq & -c_1 \|\boldsymbol{\eta}_e\|^2 - c_3 \|\boldsymbol{\nu}_e\|^2 + c_4 \|\boldsymbol{\nu}_e\| - \\ & c_5 \|\tilde{\boldsymbol{\kappa}}_c\|^4 - c_6 \|\tilde{\boldsymbol{\kappa}}_c\|^2 + c_7 \end{aligned} \quad (47)$$

其中

$$\begin{aligned} c_3 &= \frac{17b^4}{16} + c_2 - 1 \\ c_4 &= b^2 \left(\frac{1}{2} - \frac{\Phi_{\varepsilon_c}^2 \Phi_E^2}{4} - \frac{\Phi_{\varepsilon_{HJB}}}{2} \right) \\ c_5 &= \Phi_{\varphi_c}^4 \left(\frac{5}{2} + \frac{\Phi_E^2}{2} + \frac{3\Phi_E^4}{32} \right) \end{aligned}$$

$$\begin{aligned} c_6 &= \Phi_{\varphi_c}^2 \left(\frac{11\Phi_E^2 \Phi_{\varepsilon_c}^2}{16} + \frac{3\Phi_E^4 \Phi_{\varepsilon_c}^2}{32} + \frac{3\Phi_{\varepsilon_{HJB}}}{2} + \frac{\Phi_E^2 \Phi_{\varepsilon_{HJB}}}{4} \right) \\ c_7 &= \frac{\Phi_{\tilde{\kappa}_4}^2 \boldsymbol{\tau}_w^2}{2} - \frac{\Phi_E^4 \Phi_{\varepsilon_c}^4}{16} - \frac{\Phi_E^2 \Phi_{\varepsilon_c}^2 \Phi_{\varepsilon_{HJB}}}{4} \end{aligned}$$

最终,满足以下条件

$$\|\boldsymbol{\nu}_e\| > \frac{c_4 + \sqrt{c_4^2 + 4c_3 c_7}}{2c_3}$$

则 $\dot{V} < 0$ 。

4 仿真

为验证数据-物理双驱动最优控制方法(DPD-OC)的有效性和优越性,本文以文献[18]中的 CyberShip II(CS2)无人船为对象,开展3类方法的对比仿真试验。实验中,除所提方法外,选取基于物理机理的最优控制方法(PPD-OC)、数据-物理双驱动反步法控制方法(DPD-FC)作为对比方法,CS2无人船的参数如表1所示。

表1 CS2无人船参数

Tab.1 USVS Parameters of CS2

参数	数值
质量 m/kg	23.8
转动惯量 $I_z/(\text{kg}\cdot\text{m}^2)$	1.76
纵荡速度水动力导数 X_u	-0.722 5
横荡速度水动力导数 Y_v	-0.861 2
船摇角速度水动力导数 N_r	-1.9
非线性水动力导数 $X_{u u }$	-1.327 4
非线性水动力导数 $Y_{v v }$	-36.282 3
非线性水动力导数 $N_{r r }$	-0.75
运动耦合水动力导数 $Y_{v r }$	-8.05
运动耦合水动力导数 $N_{r v }$	0.08
纵荡线加速度水动力导数 $X_{\dot{u}}$	-2
横荡线加速度水动力导数 $Y_{\dot{v}}$	-10
船摇角加速度水动力导数 $N_{\dot{r}}$	-1
纵荡速度立方水动力导数 X_{uuu}	-5.866 4
控制输入最大值 $\bar{\tau}_u, \bar{\tau}_v/\text{N}$	50
控制输入最小值 $\underline{\tau}_u, \underline{\tau}_v/\text{N}$	-50
船摇控制输入最大值 $\bar{\tau}_r/(\text{N}\cdot\text{m})$	20
船摇控制输入最小值 $\underline{\tau}_r/(\text{N}\cdot\text{m})$	-20

参考信号 $\boldsymbol{\eta}_d$ 为

$$\begin{cases} x_d = 10\sin(0.1t) \\ y_d = 5\cos(0.05t) \\ \psi_d = \sin(0.1t) \end{cases}$$

为还原真实海洋作业场景,本次仿真设定为二级海况,同步引入含风速 2 m/s 、风向 50° 的海风扰动,以及风致不规则波浪干扰,对应的扰动波形如图2所示。二者的数学模型均参照文献[19]的物

理机理模型构建:风谱遵循 NORSOK 标准,波谱采用 JONSWAP 模型。

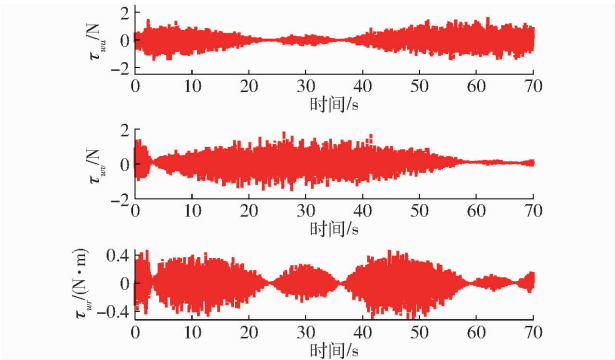


图 2 扰动波形
Fig. 2 Disturbance waveform

首先,以实船动力学系统为研究对象,构建数据-物理双驱动模型,参照文献[20-24],建模所用参数如表 2 所列。图 3 的建模误差曲线显示,误差经初始阶段快速波动后,迅速收敛并稳定于零点的有界紧集内,体现了模型的高精度拟合性能。图 4 呈现了 RNN 权重范数与非线性反馈项参数的变化过程,可见权重范数收敛性良好,反馈项参数亦稳定收敛至固定值。这充分证明了该数据-物理双驱动模型的稳定性。随后,针对无人船最优航迹跟踪控制输入,其设计所需的参数如表 3 所示。CS2 无人船的初始位置、初始姿态以及初始速度设定为 $(-2\text{ m}, 3\text{ m}, \pi/2\text{ rad}, 3\text{ m/s}, 0, 0)$ 。

表 2 数据-物理双驱动建模参数	
Tab. 2 Parameters of DPD-based modeling	
参数	数值
RNN 权值更新率矩阵 Γ_1	$\begin{bmatrix} 1 & 0.2 & 0.3 \\ 0.2 & 1 & 0.5 \\ 0.3 & 0.5 & 1 \end{bmatrix}$
RNN 权值更新率矩阵 Γ_2	$\begin{bmatrix} 1 & 0.2 & 0.1 \\ 0.2 & 1 & 0.3 \\ 0.1 & 0.3 & 1 \end{bmatrix}$
RNN 权值更新率矩阵 Γ_3	$2\mathbf{I}_{3 \times 3}$
RNN 权值更新率矩阵 Γ_4	$\begin{bmatrix} 0.2 & 0.1 & 0.1 \\ 0.1 & 0.2 & 0.1 \\ 0.1 & 0.1 & 0.2 \end{bmatrix}$
RNN 权值 $\hat{\kappa}_i (i=1,2)$ 初始值	$\mathbf{O}_{3 \times 3}$
RNN 权值 $\hat{\kappa}_3$ 初始值	$\mathbf{O}_{3 \times 1}$
RNN 权值 $\hat{\kappa}_4$ 初始值	$\begin{bmatrix} 25 & 0 & 0 \\ 0 & 24 & 3 \\ 0 & 3 & 4 \end{bmatrix}$
非线性反馈项更新率参数 Q	1
非线性反馈项参数 $\hat{\delta}$ 初始值	0.1
线性反馈项增益系数 P	1.2

上述参数的选取以保障系统稳定性、建模精度与控制性能的协同最优为核心,具体过程如下:首先

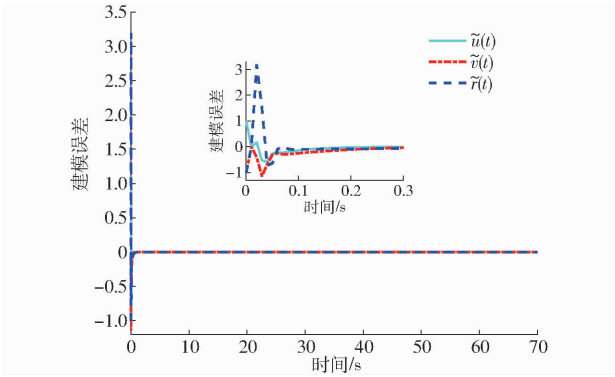


图 3 建模误差曲线
Fig. 3 Modeling error curves

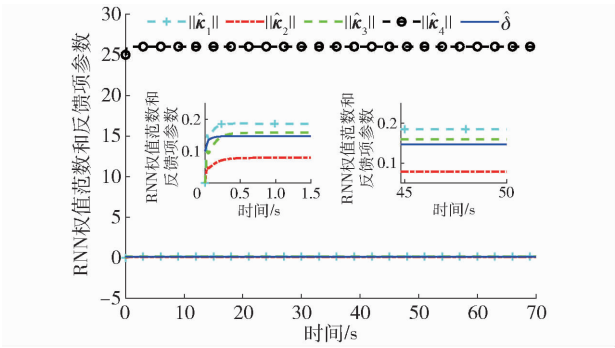


图 4 DPD 方法 RNN 权值范数与反馈项参数曲线
Fig. 4 RNN weight norm and parameter curves by DPD method

表 3 航迹跟踪控制参数	
Tab. 3 Parameters of trajectory tracking control	
参数	数值
前馈虚拟控制输入参数 c_1	1
前馈控制输入参数 c_2	3
代价函数预设矩阵 Θ	$\mathbf{I}_{3 \times 3}$
代价函数预设矩阵 Λ	$0.0004\mathbf{I}_{3 \times 3}$
基函数的中心值 μ_k	$-2, -1, 0, 1, 2$
基函数的宽度 ω	$\sqrt{2}$
评价网络学习率 β	0.1
评价网络权值 $\hat{\kappa}_e$ 初始值	$\mathbf{O}_{5 \times 1}$

借鉴文献[15,18]中针对 CS2 无人船的经典参数取值范围,结合本文数据-物理双驱动模型的融合结构及航迹跟踪控制需求,进行参数初始适配,确保初始取值满足第 2.2 节模型稳定性与第 3.3 节闭环系统稳定性的理论约束边界;在此基础上,通过多组对比试错仿真开展参数寻优,以跟踪误差最小化、控制代价最优为目标,同步验证不同参数组合下 Lyapunov 函数导数的充分条件是否成立,最终筛选出既满足一致最终有界性要求,又能兼顾建模精度与控制响应速度的最优参数组合。

最后,根据仿真结果,CS2 无人船在 DPD-OC、PPD-OC、DPD-FC 3 类控制方法下的航迹跟踪效果如图 5 所示,可见两类 DPD 方法的航迹曲线均与

参考信号高度吻合,而 PPD 方法的航迹存在一定偏离,直观印证了数据-物理双驱动方法的跟踪精度上的优势。图 6 的跟踪误差对比显示了初始阶段各方法误差均有短暂波动,但 DPD 方法的误差收敛速度快于 PPD 方法,并迅速稳定在零附近,其中 DPD-FC 的误差幅值最小,DPD-OC 的误差水平与之接近且整体波动更平缓。图 7 为考虑输入约束的控制输入对比,DPD-OC 的各自由度控制输入不仅数值显著低于另外两类方法,且波动幅度更小,充分体现了其控制资源消耗更低的特性。图 8 呈现了 OC 框架下的评价网络权值范数对比,DPD-OC 的权值范数远小于 PPD-OC,且快速收敛至稳定区间,这体现了 DPD-OC 的网络参数更新过程更趋平稳,既减少了参数波动带来的训练扰动,又切实规避了参数发散隐患,保障了控制方法的可靠性。

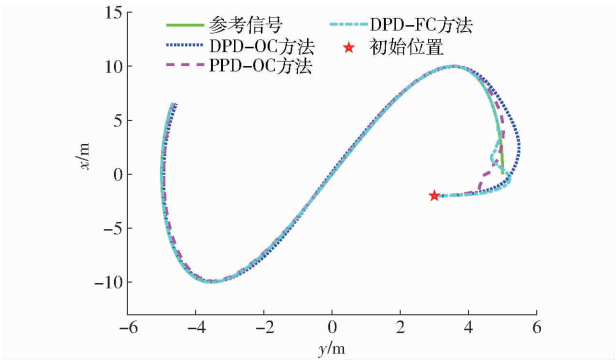


图 5 航迹跟踪对比

Fig. 5 Comparison of trajectory tracking

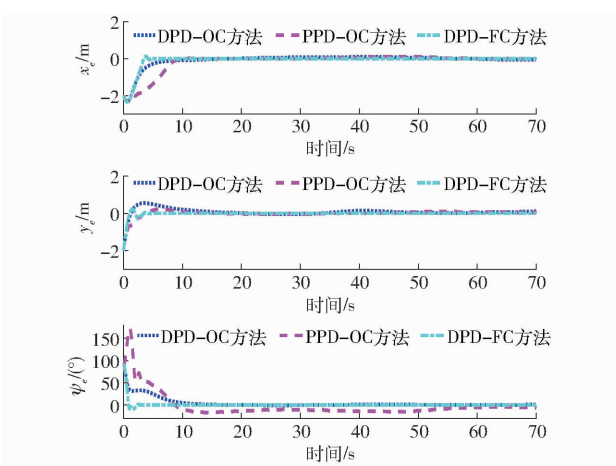


图 6 跟踪误差对比

Fig. 6 Comparison of tracking error

为量化 3 类方法的性能差异,采用文献[25]所提的平均绝对误差 (Mean absolute error, MAE) 衡量跟踪精度,平均绝对输入 (Mean absolute input, MAI) 评估控制代价,具体结果如表 4 所示。跟踪精度方面,DPD-FC 表现最优,DPD-OC 精度紧随其后且综合性能均衡,比 PPD-OC 精度整体提高 42.6%,收敛速度加快 2 倍以上;控制代价方面,DPD-OC

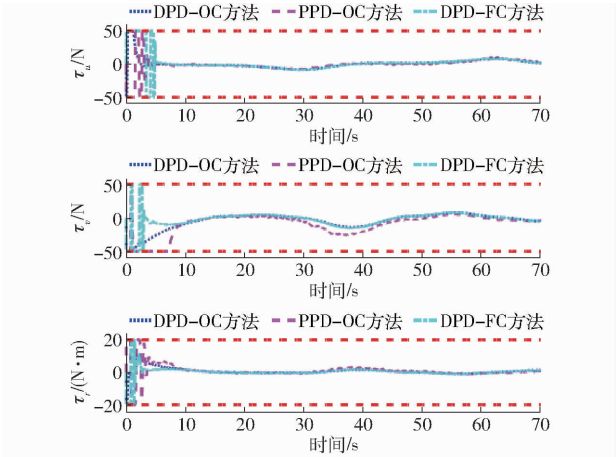


图 7 控制输入对比

Fig. 7 Comparison of control input

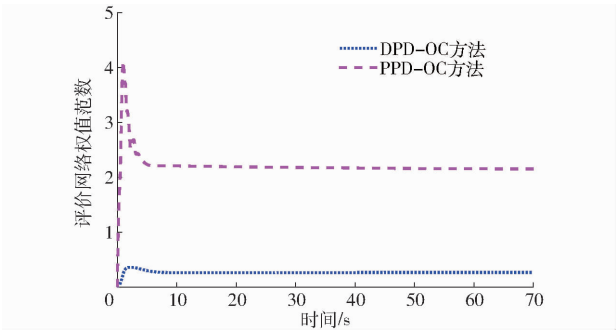


图 8 OC 方法的评价网络权值范数对比

Fig. 8 Comparison of critic NN weight norm by OC method

表 4 3 类控制方法下性能比较

Tab. 4 Performance comparison by three methods

指标	变量	DPD-OC	PPD-OC	DPD-FC
MAE	x_e/m	0.139 9	0.201 3	0.074 2
	y_e/m	0.113 7	0.074 0	0.018 3
	$\psi_e/(^{\circ})$	0.072 4	0.293 0	0.014 0
MAI	τ_u/N	4.376 7	5.155 7	6.718 9
	τ_r/N	7.827 2	11.176 1	6.829 1
	$\tau_r/(N \cdot m)$	1.122 5	1.909 0	1.202 5

的控制输入成本最低,比 DPD-FC、PPD-OC 的控制输入成本整体降低 9.6%、26.9%。

5 结论

(1)本文提出一种考虑输入约束与外界环境干扰的数据-物理双驱动强化学习最优控制方法,并以物理机理模型、传统反步法为对比策略开展性能验证。结果表明,该方法在保障优异跟踪精度的前提下,有效降低控制资源消耗,实现了跟踪性能与控制经济性的协同优化。

(2)该方法能够提升智慧海洋牧场无人船在复杂海况下的控制资源利用效率,达成跟踪性能与控制经济性的最优均衡,尤其适用于控制资源约束严格的实船作业场景。

(3)文中假设参考信号与外部干扰均满足有界条件,且无人船初始状态已知。但在实际应用中,若外界干扰超出预设边界或初始定位存在误差,可能会影响控制器的稳定性和跟踪精度。因此,后续需进一步研究在更宽松假设下的控制策略,以提升方法适配实际能力。

参 考 文 献

[1] 文莉莉, 邬满. 智慧海洋牧场关键技术研究及应用进展[J]. 广西科学院学报, 2024, 40(4):365-378.
Wen Lili, Wu Man. Research and application progress on key technologies of intelligent marine ranching[J]. Journal of Guangxi Academy of Sciences, 2024, 40(4):365-378. (in Chinese)

[2] 李俊, 李怀进, 汪振华, 等. 海洋牧场组合式无人监测船设计与试验[J]. 上海海洋大学学报, 2023, 32(5):1048-1058.
Li Jun, Li Huaijin, Wang Zhenhua, et al. Design and test of marine ranching combination-type unmanned monitoring vessel[J]. Journal of Shanghai Ocean University, 2023, 32(5):1048-1058. (in Chinese)

[3] Wang X, Liu Z T, Zhou P, et al. A saturation adaptive nonlinear integral sliding mode controller for ship permanent magnet propulsion motors[J]. Journal of Marine Science and Engineering, 2025, 13(5):976.

[4] 刘晓阳, 王文良, 郑欣宇, 等. 面向池塘养殖的水下海参原位计数无人船系统设计与试验[J]. 农业机械学报, 2025, 56(10):82-93.
Liu Xiaoyang, Wang Wenliang, Zheng Xinyu, et al. Design and testing of unmanned boat system for in-situ counting of sea cucumbers in pond aquaculture[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(10):82-93. (in Chinese)

[5] 沈跃, 刘铭晖, 沈亚运, 等. 基于MC-LADRC的水空两栖倾转多旋翼无人船水面起飞控制方法[J]. 农业机械学报, 2025,56(6):608-617.
Shen Yue, Liu Minghui, Shen Yayun, et al. Surface takeoff control of tilting multi rotor unmanned ship based on MC-LADRC[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(6):608-617. (in Chinese)

[6] Li T S, Zhao R, Chen C L, et al. Finite-time formation control of under-actuated ships using nonlinear sliding mode control[J]. IEEE Transactions on Cybernetics, 2018, 48(11):3243-3253.

[7] Tang Y J, Chen L Y, Mou J M, et al. Robust model predictive control for ship collision avoidance under multiple uncertainties[J]. IEEE Transactions on Transportation Electrification, 2024, 10(4): 10374-10387.

[8] Zhou L, Shen Z P, Nie Y, et al. Event-triggered adaptive dynamic programming for optimal tracking control of unmanned surface vessel with input constraints[J]. Transactions of the Institute of Measurement and Control, 2023, 45(15): 2835-2847.

[9] 郭子杰, 白伟伟, 周琪, 等. 基于性能指标约束的一类输入死区非线性系统最优控制[J]. 自动化学报, 2019, 45(11): 2128-2136.
Guo Zijie, Bai Weiwei, Zhou Qi, et al. Adaptive optimal control for a class of nonlinear systems with dead zone input and prescribed performance[J]. Acta Automatica Sinica, 2019, 45(11):2128-2136. (in Chinese)

[10] Wang Y H, Bai W W, Zhang W J. Cooperative optimization hunting control of multi-MSV based on reinforcement learning[J]. Applied Computational Intelligence and Soft Computing, 2025, 2025:1415549.

[11] Zhao B, Zhang S C, Liu D R. Self-triggered approximate optimal neuro-control for nonlinear systems through adaptive dynamic programming[J]. IEEE Transactions on Neural Networks and Learning Systems, 2025, 36(3): 4713-4723.

[12] Morawski L, Pomirski J. Ship track-keeping: experiments with a physical tanker model[J]. Control Engineering Practice, 1998, 6(6): 763-769.

[13] Tayfun C, Stephen P B. Nonlinear optimal tracking control with application to super-tankers for autopilot design[J]. Automatica, 2004, 40(11): 1845-1863.

[14] 王维坤, 刘加朋, 王保防, 等. 基于数据驱动的无人船轨迹跟踪自适应控制[J]. 中国舰船研究, 2025, 20(1):223-231.
Wang Weikun, Liu Jiapeng, Wang Baofang, et al. Data-driven unmanned surface vehicles trajectory tracking adaptive control[J]. Chinese Journal of Ship Research, 2025, 20(1):223-231. (in Chinese)

[15] Liu S L, Niu B, Zong G D, et al. Data-driven-based event-triggered optimal control of unknown nonlinear systems with input constraints[J]. Nonlinear Dynamics, 2022, 109(2):891-909.

[16] 闫洪迪, 崔承刚, 陈辉, 等. 基于融合物理模型与数据驱动模型的电池SOH估计方法[J]. 电源技术, 2025, 49(6): 1183-1191.
Yan Haodi, Cui Chenggang, Chen Hui, et al. Battery SOH estimation method based on fusion of physical model and data-driven model[J]. Chinese Journal of Power Sources, 2025, 49(6):1183-1191. (in Chinese)

[17] Qin B D, Wang Z, Fan W F, et al. A novel data-driven physical iterative modeling approach and its application in quantum instrumentation[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2024, 54(9):5667-5679.

Song Deping, Wang Cheng, Sun Dongxia, et al. Design and test of intelligent cotton planter based on Beidou navigation[J]. Journal of Intelligent Agricultural Mechanization, 2021, 2(1): 44 – 50, 63. (in Chinese)

[17] 陈建林, 薄克明, 王健平, 等. 水稻插秧机搭载北斗卫星导航系统驾驶作业适配性试验[J]. 农机科技推广, 2022(4): 44 – 45.

Chen Jianlin, Bo Keming, Wang Jianping, et al. Adaptability test of driving operation of rice transplanter equipped with Beidou satellite navigation system[J]. Agriculture Machinery Technology Extension, 2022(4): 44 – 45. (in Chinese)

[18] 武涛. 无人驾驶水稻直播机抗干扰精准跟踪及高效协同作业研究[D]. 上海: 上海交通大学, 2022.

Wu Tao. Research on anti-disturbance precise tracking and efficient cooperative operation of unmanned rice seeding machine [D]. Shanghai: Shanghai Jiao Tong University, 2022. (in Chinese)

[19] He X, Zhang D, Yang L, et al. Design and experiment of a GPS-based turn compensation system for improving the seeding uniformity of maize planter[J]. Computers and Electronics in Agriculture, 2021, 187: 106250.

[20] Lee K, Choi H, Kim J. Development of path generation and algorithm for autonomous combine harvester using dual GPS antenna[J]. Sensors, 2023, 23(10): 4944.

[21] Padhiary M, Saha D, Kumar R, et al. Enhancing precision agriculture: a comprehensive review of machine learning and AI vision applications in all-terrain vehicle for farm automation[J]. Smart Agricultural Technology, 2024, 8: 100483.

[22] 崔荣江, 荐世春, 杨继鲁, 等. 勺链式大蒜取种器的优化设计与试验[J]. 农机化研究, 2017, 39(2): 99 – 102, 107.

Cui Rongjiang, Jian Shichun, Yang Jilu, et al. Optimization design and test of take garlic spoon[J]. Journal of Agricultural Mechanization Research, 2017, 39(2): 99 – 102, 107. (in Chinese)

[23] 耿爱军, 栗晓宇, 侯加林, 等. 自动定向大蒜播种机的设计与试验[J]. 农业工程学报, 2018, 34(11): 17 – 25.

Geng Aijun, Li Xiaoyu, Hou Jialin, et al. Design and experiment of automatic directing garlic planter[J]. Transactions of the CSAE, 2018, 34(11): 17 – 25. (in Chinese)

[24] 查玲玲. 大蒜直立种植装置设计与研究[D]. 杭州: 浙江理工大学, 2023.

Zha Lingling. Design and research of a vertical planting device for garlic[D]. Hangzhou: Zhejiang Sci-Tech University, 2023. (in Chinese)

[25] GB/T 36197—2018 土壤质量土壤采样技术指南[S].

(上接第 62 页)

[18] Poger S, Thor I F, Petar V K. Adaptive maneuvering, with experiments, for a model ship in a marine control laboratory[J]. Automatica, 2005, 41(2): 289 – 298.

[19] Zhang G Q, Liu S, Zhang X K, et al. Event-triggered cooperative formation control for autonomous surface vehicles under the maritime search operation[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(11): 21392 – 21404.

[20] Wei Q L, Song R Z, Yan P F. Data-driven zero-sum neuro-optimal control for a class of continuous-time unknown nonlinear systems with disturbance using ADP[J]. IEEE Transactions on Neural Networks and Learning Systems, 2016, 27(2): 444 – 458.

[21] Lu C X, Meng D Y, Li H Y. Data-induced learning control for unknown nonlinear systems with full-tracking performances[J]. Science China Information Sciences, 2025, 68(8): 1 – 14.

[22] Yang X, He H B. Adaptive critic designs for event-triggered robust control of nonlinear systems with unknown dynamics[J]. IEEE Transactions on Cybernetics, 2018, 49(6): 2255 – 2267.

[23] Fu Y, Chai T Y. Online solution of two-player zero-sum games for continuous-time nonlinear systems with completely unknown dynamics[J]. IEEE Transactions on Neural Networks and Learning Systems, 2016, 27(12): 2577 – 2587.

[24] Durstewitz D, Koppe G, Thurm M I. Reconstructing computational system dynamics from neural data with recurrent neural networks[J]. Nature Reviews Neuroscience, 2023, 24(11): 693 – 710.

[25] Zhang G Q, Cai Y Z, Zhang W D. Robust neural control for dynamic positioning ships with the optimum-seeking guidance [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2017, 47(7): 1500 – 1509.