

基于深度强化学习的船舶非线性横摇控制

钟齐鸣, 秦雷洪, 张松涛*, 南晓峰

(哈尔滨工程大学 智能科学与工程学院, 哈尔滨 150000)

摘要: [目的]为解决船舶非线性横摇运动中存在不确定参数摄动和海浪外界干扰等问题, [方法]提出基于双延迟深度确定性策略梯度(TD3)的控制方法, 建立船舶非线性横摇模型, 综合考虑海况、航速等因素设计状态空间和奖励函数以增强对动态环境的理解。同时设置基于深度确定性策略梯度(DDPG)的控制方法作为对照, 比较2种算法在不同海况下的减摇效果。[结果]结果表明: 2种算法均能有效减小横摇。在同工况下, TD3算法的减摇效果都略优于DDPG算法, 且TD3算法的输出控制信号表现出更为平滑和稳定的特性, 基于TD3算法设计的控制器对船舶的非线性横摇减摇具有较好的控制效果和很好的自适应能力。[结论]研究成果可为船舶非线性横摇控制提供一定参考。

关键词: 深度强化学习; 横摇减摇; 非线性控制

中图分类号: U674.951 **文献标志码:** A **【DOI】** 10.13788/j.cnki.cbge.2025.07.13

Ship Nonlinear Roll Reduction Control Based on Deep Reinforcement Learning

ZHONG Qiming, QIN Leihong, ZHANG Songtao*, NAN Xiaofeng

(College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin 150000, China)

Abstract: [Purpose] In order to solve the problems of uncertain parameter perturbations and external disturbances caused by waves in the nonlinear roll motion of ships, [Method] a control method based on double delay depth deterministic strategy gradient (TD3) is proposed to establish a nonlinear roll model of ships. The state space and reward function are designed to comprehensively consider factors such as sea conditions and speed to enhance the understanding of dynamic environments. Simultaneously, a control method based on depth deterministic policy gradient (DDPG) is set as a control to compare the anti roll effects of the two algorithms under different sea conditions. [Result] The results indicate that both algorithms can effectively reduce roll. Under the same operating conditions, the anti roll effect of TD3 algorithm is slightly better than that of DDPG algorithm, and the output control signal of TD3 algorithm shows smoother and more stable characteristics. The controller designed based on TD3 algorithm has good control effect and good adaptability for nonlinear roll anti roll of ships. [Conclusion] The research results can provide some references for nonlinear roll control of ships.

Key words: deep reinforcement learning; roll reduction; nonlinear control

0 引言

船舶在海上航行的过程中, 受到波浪和海流等多方面影响, 从而产生各类摇荡运动。其中, 船舶

横摇是威胁船舶安全的主要因素。剧烈的横摇运动不仅影响船舶的舒适性, 还可能导致货物损失和人员伤亡; 严重的非线性横摇运动在波浪激励或其他

收稿日期: 2024-08-19; 修回日期: 2024-11-26

基金项目: 预研项目(船舶多操纵面智能控制技术); 研究生课程思政案例省级建设项目(《船舶减摇原理与装置》); 哈尔滨工程大学本科教改项目(JG2023B0411)。

作者简介: 钟齐鸣(1998—), 男, 硕士研究生。研究方向: 船舶智能控制。

通信作者: 张松涛(1981—), 男, 副教授、研究生导师。研究方向: 船舶运动与航行姿态控制。

激励扰动下甚至会导致船舶的倾覆^[1]。因此, 寻找有效的方法来减少横摇运动一直是船舶运动控制领域研究的重点。各种类型的减摇装置(如舳龙骨、减摇水舱和减摇鳍等)已被用于减小横摇运动。对于中高速航行的船舶, 减摇鳍是减摇效果最好的减摇装置, 其根据船舶的横摇角和横摇角速度来控制机械角度以减小横摇运动^[2]。减摇鳍的减摇效果主要取决于鳍型线和控制策略, 前者决定鳍的水动力特性, 后者通过控制鳍的运动产生升力来减少船舶横摇^[3]。因此, 亟须选择合适的减摇鳍控制器以减小船舶在波浪中的横摇运动。

传统的船舶横摇控制方法主要依赖于比例-积分-微分(Proportional-Integral-Derivative, PID)控制器或者基于线性控制理论的控制方法。然而, 在复杂多变的海洋环境中, 这些传统方法往往难以取得令人满意的控制效果。因此, 近年来越来越多的研究开始关注于应用其他技术来解决船舶横摇控制的问题。FANG等^[4]提出基于神经网络理论的自适应PID控制器, 用于调整最佳减摇鳍角度以减少船舶在波浪中的横摇运动, 比较了传统PID控制器和神经网络PID控制器在规则波和短波浪中的控制性能, 并进一步研究了2种控制器在不同波浪方向下的横摇减少效果和稳定性。研究表明, 神经网络PID控制器在长时间连续操作中可能因累积误差或参数漂移导致性能下降。NEJIM^[5]提出一种自适应的PD控制方法以应对环境的干扰, 通过比较线性与非线性模型的仿真结果, 展示了所提出控制方案的有效性。然而, 自适应控制器非常复杂, 需要多次调整和大量分析才能实现对目标的控制。LIU等^[6]提出一种整合路径跟踪和横摇稳定的模型预测控制(Model Predictive Control, MPC)方法, 通过神经动力学优化提高计算效率并保证闭环稳定性; 为应对MPC实施中的计算强度和闭环稳定性问题, 还提出使用投影神经网络投影神经网络(Projection Neural Network, PNN)优化MPC方法制定的目标函数, 从而提高计算效率。然而, 并未深入分析模型误差对控制性能的影响。LI等^[7]将神经网络和自适应技术相结合, 提出了一种基于自适应神经网络的船舶横摇控制设计方法, 实现了船舶横摇的稳定控制。考虑到实际船舶横摇运动控制中不仅存在模型参数的不确定性等内部不确定性, 还存在外部未知波浪环境的影响, 因此在控制设计中必须考虑未知外部扰动的不利影响。然而, 自适应神经网络无法有效地重建未知的外部环境扰动。SONG等^[8]提出一种基于遗传算法(Genetic Algorithm, GA)和径向基函数人工神经网络(Radial Basis Function Artificial Neural Network, RBF-ANN)的控制策略, 用于解决传统减摇鳍在低速甚至零速时减摇性能迅速下降的问题。

通过船模试验和仿真验证了该方法的有效性, 在复杂海况下展现了良好的鲁棒性和减摇效果。尽管上述研究在船舶横摇控制中取得了显著成果, 但仍存在少量问题。虽然神经网络和自适应技术有所进展, 但在处理外部环境扰动和系统参数变化时仍然存在局限性。此外, 许多方法过于依赖精确模型, 难以在复杂海况下保持良好的控制性能。为了克服这些局限, 本文提出了一种基于双延迟深度确定性策略梯度算法(Twin Delayed Deep Deterministic Policy Gradient, TD3)的深度强化学习方法, 该方法能够在高度非线性和不确定的环境下自适应调整控制策略。通过自主学习, 实时优化减摇鳍角度, 适应变化的海况, 克服了传统方法中的调参困难。此外, TD3无须依赖精确数学模型或大量人工干预, 具有更强的灵活性和适应性。

针对船舶非线性横摇运动中, 传统控制算法因不确定参数摄动和海浪外界干扰等因素导致控制效果不佳的问题, 结合TD3算法原理和减摇鳍控制特性, 提出了一种基于双延迟深度确定性策略梯度的控制方法。通过设置网络结构和奖励函数来训练网络以实现期望输出。为证明基于TD3算法设计的控制器的有效性, 将该控制器的性能与基于DDPG算法设计的控制器进行比较。最后给出时域仿真结果, 验证TD3算法在船舶非线性横摇减摇的有效性。

1 波浪上横摇运动数学模型

1.1 海浪模型

在船舶航行过程中, 影响船舶的姿态和稳定航行的主要因素是由海浪引起的干扰力矩。采用谱分析的方法来估计波浪产生的随机干扰力矩。本文选择皮尔逊-莫斯科维茨(Pierson-Moskowitz, P-M)谱作为海浪能量谱函数^[9], 其表达式为

$$S_{\zeta}(\omega) = \frac{8.11 \times 10^{-3} g^2}{\omega^5} \exp\left(-\frac{3.11}{h_{1/3}^2 \omega^4}\right) \quad (1)$$

式中: $S_{\zeta}(\omega)$ 为波能谱密度; g 为重力加速度; ω 为波浪角频率; $h_{1/3}$ 为海浪波高有义值。

海浪的谐波幅值 ζ 是均匀微小量, 表达式为

$$\zeta = \sqrt{2S(\omega)\Delta\omega} \quad (2)$$

式中: $\Delta\omega$ 为频率间隔。

当船以一定航速和遭遇浪向航行时, 船体实际感受到的波浪角频率已不同于波浪的自然角频率, 而是遭遇角频率。遭遇角频率与海浪的自然角频率之间的关系式为

$$\omega_e = \omega - \frac{\omega^2}{g} v \cos \beta \quad (3)$$

式中: ω_e 为遭遇角频率; v 为船舶航速; β 为船舶相对于海浪的遭遇角。

船舶横摇程度的强弱与作用于船舶的海浪波倾

角的大小有关。因此,需要将海浪谱密度转换为波倾角谱^[10]。波倾角谱密度的表达式为

$$S_{\alpha}(\omega) = \frac{\omega^4}{g^2} S_{\zeta}(\omega) \quad (4)$$

$$\alpha(t) = \sin \beta \sum_{j=1}^N \sqrt{2 \int_{\omega_{j-1}}^{\omega_j} S_{\alpha}(\omega) d\omega} \cos(\omega_j t + \varepsilon_j) \quad (5)$$

式(4)和式(5)中: $S_{\alpha}(\omega)$ 为波倾角谱密度; $\alpha(t)$ 为波倾角。

有义波高为2 m, 航速 v 为16.7 kn, 遭遇角为60°时的随机海浪波倾角仿真曲线见图1。

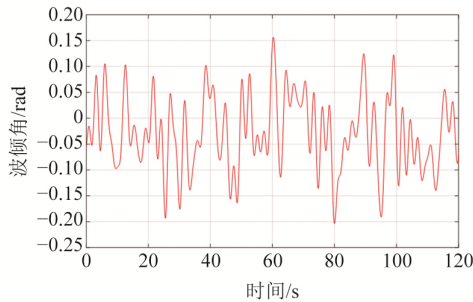


图1 海浪波倾角仿真曲线

Fig. 1 Simulation Curve of Wave Slope Angle

1.2 船舶横摇运动数学建模

以船舶质心为原点 G , x 、 y 和 z 分别为船舶的横摇轴、纵摇轴和艏摇轴。船舶横摇模型见图2。图2中: ϕ 为横摇角; α 为波倾角。假设减摇鳍位于船中部靠近质心的位置, 并且与其他运动的耦合较小, 此时可以认为与其他运动的耦合效应相对较小, 主要只考虑与横摇自由度的耦合^[11]。

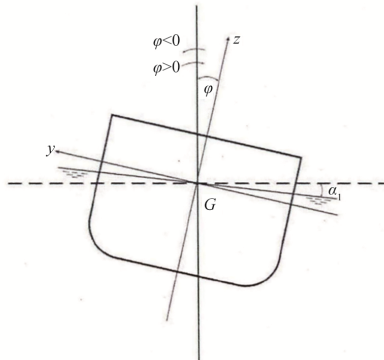


图2 船舶横摇模型

Fig. 2 Ship Rolling Model

在船舶横摇角度较小时, 船舶线性横摇模型能比较好地表达船舶的横摇运动; 但是对于大角度横摇, 非线性因素比较突出, 线性横摇模型不能很好地反映船舶横摇特性^[12]。引起非线性的原因很多, 大致可以分为2类: 1) 环境条件的非线性; 2) 船体运动的非线性。本文仅考虑由于横摇角度过大产生的复原力矩非线性和阻尼力矩非线性。

非线性横摇运动方程式的重要特点是不适用叠加原理。因此, 不能像线性解那样, 把波浪中非线性横摇看成是自由衰减振荡运动和强迫振荡运动的叠加, 它们有更复杂的运动形式^[13]。

复原力矩非线性是由静稳性曲线的形状确定的。在小角度情况下, 复原力矩与倾角成正比, 随着倾角增加, 复原力矩的大小必须根据静稳性曲线的形状来确定。船舶横摇过大而倾覆的原因之一可能是静稳性曲线的非线性。

横摇阻尼力矩本质上是是非线性的, 只是在横摇角度(角速度)较小时, 线性成分占主要部分, 才可以作为线性问题处理。

船舶的非线性横摇模型表达式为

$$-Dh\alpha_1 = (J_{\phi\phi} + \Delta J_{\phi\phi})\ddot{\phi} + B_1\dot{\phi} + B_2\dot{\phi}|\dot{\phi}| + C_1\phi + C_3\phi^3 + C_5\phi^5 \quad (6)$$

式中: $J_{\phi\phi}$ 为船舶自身的转动惯量; $\Delta J_{\phi\phi}$ 为附加转动惯量; $\ddot{\phi}$ 为船舶的横摇角加速度; B_1 和 B_2 为非线性阻尼力矩系数; $\dot{\phi}$ 为船舶的横摇角速度; C_1 、 C_3 和 C_5 为非线性横摇复原力矩的系数; ϕ 为船舶横摇角; D 为船舶排水量; h 为船舶的横稳心高; α_1 为船舶在遭遇频率的有效波倾角。

1.3 鳍上升力模型

减摇鳍是最常见的减摇设备, 减摇效果最高可达90%。其通过感知船舶横摇运动并产生相应的反作用力, 来抑制或减小船舶的横摇摆动, 从而提高船舶的稳定性和航行性能^[14]。减摇鳍模型参数见表1。

表1 减摇鳍模型参数

Tab. 1 Parameters of Fin Stabilizer Model

参数	数值
鳍面积/m ²	9
展长/m	4.450
平均弦长/m	2.023
展弦比	2.2
翼型	鱼尾鳍

根据升力公式, 可以获得减摇鳍上的升力为

$$L = \frac{1}{2} \rho A_F V^2 C_L \quad (7)$$

式中: ρ 为海水密度; A_F 为鳍减摇鳍的参考面积; V 为船体相对于来流的速度; C_L 为升力系数。

2 控制器算法及控制器设计

减摇鳍控制的核心是通过调整鳍角来产生对抗横摇的复原力矩, 这一控制过程本质上是一个连续动作问题。双延迟深度确定性策略梯度通过确定性策略函数直接输出动作, 而不依赖离散化或随机采样的方式, 这种连续的动作输出方式非常适合减摇鳍的高精度控制需求。此外, 船舶横摇运动是一个

高度非线性、多变量耦合的动力学系统, 减摇鳍控制需要综合考虑横摇角、横摇角速度、波浪方向、船速等多个因素进行精确控制。TD3具有自适应学习能力, 通过学习 $Q(s, a)$ 价值函数不断优化策略, 能够在复杂环境中优化鳍角的动作策略, 提高减摇效果。

2.1 双延迟深度确定性策略梯度

双延迟深度确定性策略梯度是强化学习中基于策略的梯度下降方法, 是基于深度确定性策略梯度 (Deep Deterministic Policy Gradient, DDPG) 的改进版本。与DDPG算法一样, TD3算法也是基于Actor-Critic架构、面向连续动作空间的深度强化学习算法, 常用于解决连续动作空间中的强化学习问题^[15]。针对DDPG算法容易造成策略估计过于乐观、目标函数方差较大等问题, TD3算法通过以下2种技术手段来缓解这些问题。

1) 双重网络

由于DDPG算法的复杂性, 稍微不合适的超参数设置或其他调整可能会导致模型无法有效学习, 甚至出现不稳定或发散的情况。其中最常见的问题是模型学习到的 Q 值会显得比实际情况更高。如果高估是均匀的, 对于智能体最终的决策不会产生影响; 如果是非均匀的, 对于智能体最终的决策会产生显著影响。然而实际上网络的高估通常是非均匀的。TD3网络结构见图3。图3中: S 和 S' 分别为当前状态和下一个状态; A 和 A' 分别为当前动作和下一个动作。

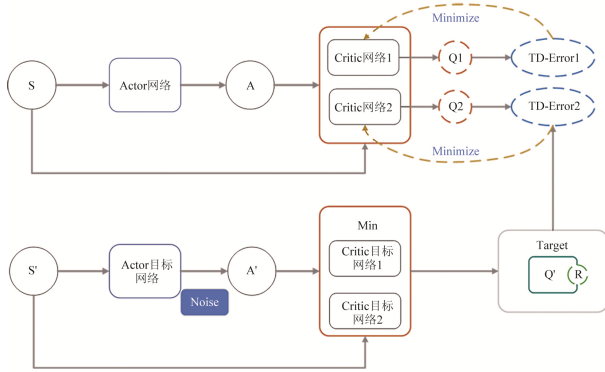


图3 TD3 网络结构

Fig. 3 Architecture of TD3 Network

TD3采用了双 Q 网络 (Double Q-Networks) 的策略, 通过使用2个独立的 Q 网络来估计动作值。在计算目标 Q 值时, 不是简单地取最大值, 而是选择2个 Q 网络中估计值较小的一个, 从而避免了过度乐观的估计^[16]。目标 Q 值可以表示为

$$y = r + \gamma \min_{i=1,2} Q_{\theta_i'}[s', \pi_{\phi_i}(s')] \quad (8)$$

式中: y 为目标 Q 值, 用于更新 Q 网络的参数; r 为智能体在当前状态 S 执行当前动作 A 后获得的即时反馈奖励; γ 为折扣因子, 用于权衡即时奖励和未来奖励的相对重要性, γ 越大, 算法越重视未来的奖励; $Q_{\theta_i'}[s', \pi_{\phi_i}(s')]$ 为在下一状态 s' 下执行动作 $\pi_{\phi_i}(s')$ 的预期累积奖励。

2) 目标策略平滑化

确定性策略往往会过度拟合, 使得价值估计中的峰值过于集中。在更新Critic网络时, 确定性策略的学习目标容易受到函数逼近误差的影响, 导致目标估计方差较大, 进而使得估计值不准确。为了减小这种由误差引起的方差, 可以引入正则化策略即目标策略平滑化^[17]。尽管函数逼近在一定程度上隐含地实现了这一点, 但TD3通过修改训练过程来显式地强制这种相似动作之间的关系。具体的实现方法是利用目标动作周围的区域来计算目标值, 从而有助于平滑价值估计。目标策略平滑化的目标 Q 值可以表示为

$$y = r + E \left\{ Q_{\theta_i'}[s', \pi_{\phi_i}(s') + \omega] \right\} \quad (9)$$

式中: ω 为引入的随机噪声, 添加在目标 Q 值的期望计算中, 用于目标动作的探索。

实际操作中, 通过向目标动作中加入随机噪声 ω 并在小批量中求平均值来平滑策略的期望, 减少由于函数逼近误差引入的偏差。该随机噪声 ω 服从正态分布, 其值限制在 $[-c, c]$ 的范围内。因此, 式(9)可修改为

$$\begin{cases} y = r + \gamma Q_{\theta_i'}[s', \pi_{\phi_i}(s') + \omega] \\ \omega \sim \text{clip}[N(0, \sigma), -c, c] \end{cases} \quad (10)$$

式中: clip 为裁剪函数; N 为正态分布; σ 为噪声强度的标准差。

3) 延迟策略更新

在DDPG中, 策略网络和 Q 网络同步更新, 容易导致策略网络过度优化当前 Q 网络输出的 Q 值。这会引发 Q 值过估计的问题。因为Actor网络是通过最大化累积期望回报来更新的, 需要利用Critic网络来进行评估。如果Critic网络非常不稳定, 那么Actor网络自然也会出现震荡。因此通过延迟策略更新, TD3减少了这种同步更新的频率, 使得 Q 网络有更多时间进行准确的价值评估, 从而降低了 Q 值过估计的风险。

2.2 算法伪代码

双延迟深度确定性策略梯度的伪代码见表2。

表2 双延迟深度确定性策略梯度的伪代码

Tab. 2 Pseudocode for Double-Delay Depth Deterministic Policy Gradients

序号	TD3 算法说明
1	初始化 Critic 网络 $Q_{\theta_1}, Q_{\theta_2}$ 和 Actor 网络 π_{ϕ} 的随机参数 θ_1, θ_2, ϕ
2	初始化目标网络参数
3	初始化经验池
4	循环
5	通过探索策略选择动作, 同时获得奖励 r 和新状态 s'
6	智能体执行动作与环境交互, 将信息以元组的形式存储到经验池中
7	从经验池缓存中采样一个数量为 N 的小批量元组进行网络训练
8	通过 Actor 目标网络获取状态 s' 下的动作
9	通过 Critic 目标网络获得 $s', \bar{a}$ 下 2 个目标 Q 值函数, 并计算目标值
10	更新 Critic 训练网络参数
11	循环
12	通过策略梯度更新 Actor 训练网络参数
13	软更新目标网络参数
14	结束循环

2.3 控制器设计

控制系统流程图见图4。与传统控制器相比, 使用TD3 agent替代PID控制器实现对减摇鳍鳍角的输出。将横摇角度、横摇角速度、横摇角加速度作为状态输入, 减摇鳍的鳍角作为输出动作。Actor网络根据当前状态输出动作值, 在训练过程中, 会不断

地学习和优化, 以输出更合适的动作。同时将Actor网络输出的动作与状态同时作为Critic网络的输入。TD3网络中有2个Critic网络, 这2个网络的主要功能是评估 Q 值, 也就是根据给定的状态和动作来估计未来的奖励。使用2个Critic网络可以减小估计的 Q 值的方差, 使结果更加稳定。

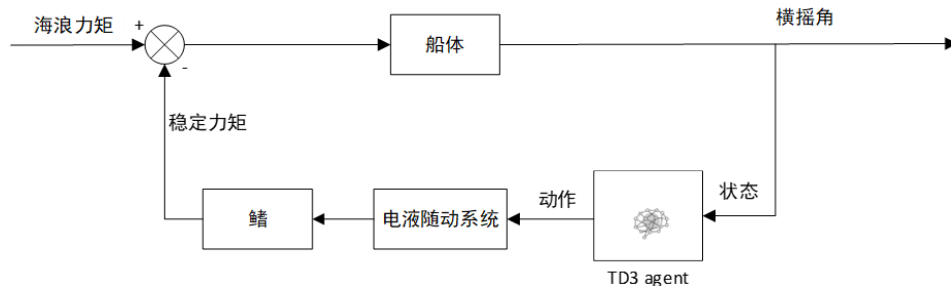


图4 控制系统流程图

Fig. 4 Control System Flowchart

2.4 奖励函数设计

在强化学习中, 奖励函数定义了智能体在环境中采取行动所获得的反馈, 引导智能体朝着期望的行为方向发展。因此, 奖励函数的设计直接影响着智能体的学习效率和最终性能, 选择合适的奖励函数是至关重要的。本文以横摇角度、横摇角速度、鳍角速度作为考虑因素。为了确保船舶的稳定性和舒适性, 希望船舶的横摇角度尽可能小, 所以对横摇角度的负值进行加权并采取分段函数处理。当横摇角度不小于 2° 时, 给予一个较大的惩罚权重; 当横摇角度小于 2° 时, 给予一个较小的惩罚权重。若

要避免剧烈的摇摆, 横摇角速度也应尽可能小, 对横摇角速度的平方的负值进行加权处理; 同时避免频繁和大幅度的控制动作, 以减少对系统的冲击; 对鳍角速度的平方的负值也进行加权。综上所述, 设计的具体奖励函数为

$$R = q_1(t)\phi(t)^2 + q_2\dot{\phi}(t)^2 + q_3\dot{\delta}(t)^2 \quad (11)$$

式中: q_2 为横摇角速度权重, 取0.9; q_3 为减摇鳍角速度权重, 取0.5; q_1 为横摇角权重。

横摇角权重相关设置为

$$q_1(t) = \begin{cases} q_{\text{large}} = 3 & |f(t)| \geq 2^\circ \\ q_{\text{small}} = 1.5 & |f(t)| < 2^\circ \end{cases} \quad (12)$$

3 仿真试验

为验证上述设计的TD3算法应用于船舶非线性横摇控制的有效性, 基于Matlab搭建仿真试验平台, 试验过程中涉及的运维船的船体主要参数见表3。

表3 船体主要参数

Tab.3 Main Parameters of the Hull

参数	数值
船长/m	91.00
垂线间长/m	86.32
型宽/m	20.00
型深/m	15.00
结构吃水/m	6.00
设计航速/kn	16.1
设计排水量/t	8 124

在仿真试验中, 设置DDPG的控制方法作为对照组, 比较2种算法的减摇效果。航行条件设定为航速16.7 kn, 遭遇角 60° ; 仿真时长为120 s, 控制周期为0.1 s, 打鳍角度限幅为 $\pm 20^\circ$, 打鳍角速度限幅为 $\pm 25^\circ$ 。分别在海况为3级(SSN3)、4级(SSN4)、5级(SSN5)的工况下进行仿真。海浪干扰力矩仿真曲线见图5。由图5可知, 随着浪级的升高, 海浪的周期也随之增大, 从而产生更大的干扰力和力矩, 使船体的运动更加剧烈。

Critic和Actor的网络结构分别见图6和图7。Critic网络中全连接层神经元个数设置为64, Actor网络中第1个全连接层神经元个数设置为64, 第2个

全连接层神经元个数设置为128。2种算法中设置的相关参数信息见表4。

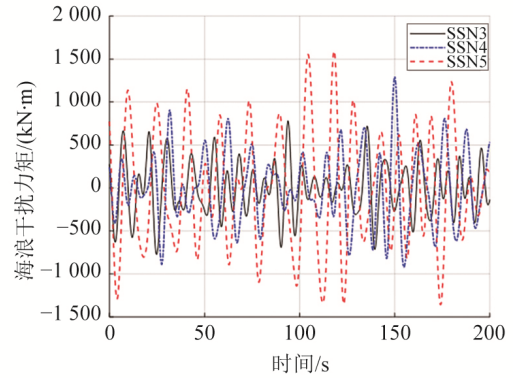


图5 海浪干扰力矩仿真曲线

Fig. 5 Wave Disturbance Moment Simulation Curve

表4 网络相关参数

Tab. 4 Network-Related Parameters

参数	数值
仿真时间/s	120
采样时间/s	0.1
学习率	0.001
折扣因子	0.99
平滑因子	0.001
批量大小	128

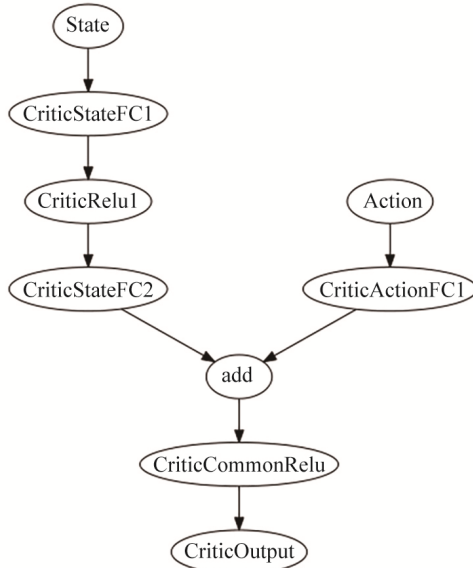


图6 Critic 网络结构

Fig. 6 Critic Network Architecture

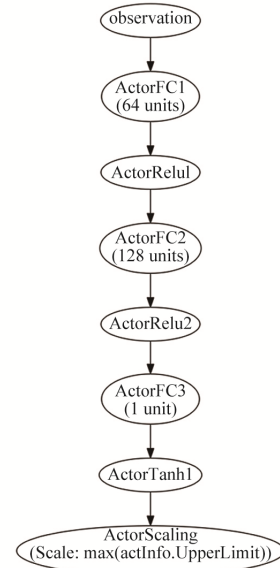


图7 Actor 网络结构

Fig. 7 Actor Network Architecture

考察减摇控制器的控制效果时,常用减摇率这一概念。减摇率 η 可以用公式近似计算得到,即

$$\eta = 100 \times \left(1 - \frac{\theta_1}{\theta_2} \right) \quad (13)$$

式中: θ_1 和 θ_2 分别为有、无减摇控制器时横摇角的均方根值。

DDPG算法、TD3算法和无控制情况下,船舶在各海况下横摇角的最大值、最小值、标准差和减摇率等4项数据的归纳统计见表5。由表5可知:随着海况增大,无控制情况下的横摇角度明显增大;同一海况下,2种控制算法均能实现有效减小横摇,且TD3减摇效果均大于DDPG。以SSN3工况为例,DDPG算法和TD3算法的标准差分别为0.836 5°和

0.759 9°,减摇率分别为82.20%和84.15%。在DDPG中,策略网络和价值网络同步更新,这可能导致策略网络过早收敛到次优解;尤其在训练初期,价值网络的估计尚不准确,策略网络的同步更新会受到错误价值估计的影响。TD3引入策略延迟更新机制,即策略网络的更新频率低于价值网络,提高了策略的优化质量,使得最终结果更接近全局最优解,因此,TD3减摇效果要优于DDPG。随着海况增大,2种算法控制效果均有减弱。以TD3为例,SSN3、SSN4和SSN5工况下减摇率分别为84.15%、79.50%和76.62%。这表明更恶劣的海况对控制器性能提出了更高的挑战,同时反映出TD3算法在复杂条件下仍然保持较高的减摇效果和良好的适应性。

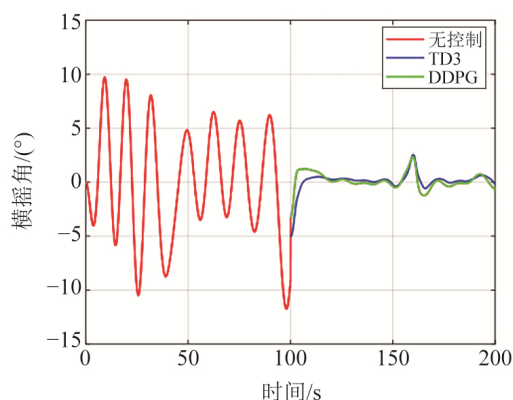
表5 不同工况横摇角度归纳统计

Tab. 5 Roll Angle Summary Statistics Under Different Conditions

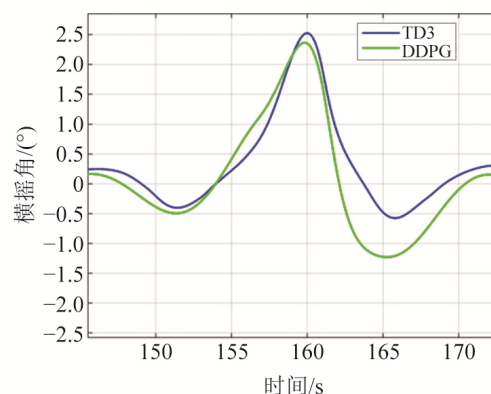
控制方式	海况	最大值/(°)	最小值/(°)	标准差/(°)	减摇率/%
无控制	SSN3	12.900	-11.730	4.735 1	—
	SSN4	16.410	-19.940	6.769 9	—
	SSN5	19.410	-25.490	8.492 0	—
DDPG	SSN3	2.368	-3.465	0.836 5	82.20
	SSN4	8.045	-8.291	1.659 6	75.57
	SSN5	11.830	-11.210	2.310 2	72.80
TD3	SSN3	2.527	-5.019	0.759 9	84.15
	SSN4	3.895	-9.694	1.387 9	79.50
	SSN5	4.896	-12.350	1.815 0	76.62

SSN3~SSN5工况下的横摇运动轨迹分别见图8~图10。由图8~图10可知,在SSN3工况下,2种算法都能很好实现减摇,SSN3工况仅在155 s~165 s出现较大波动,此时鳍角已打满,无法提供更多升力。在SSN4和SSN5工况下,船舶在100 s左右遭遇

了大角度横摇,此时DDPG的减摇效果明显低于TD3。原因可能是在复杂的高维连续动作任务中,DDPG的性能会因高估计偏差或策略不稳定而下降。从控制效果来看,TD3的控制效果相对更加平滑,而DDPG则在某些位置出现了较大的波动。



(a) 横摇角全时域响应对比



(b) 横摇角局部放大响应对比

图8 SSN3 工况下横摇运动轨迹

Fig. 8 Roll Motion Trajectory Under SSN3 Condition

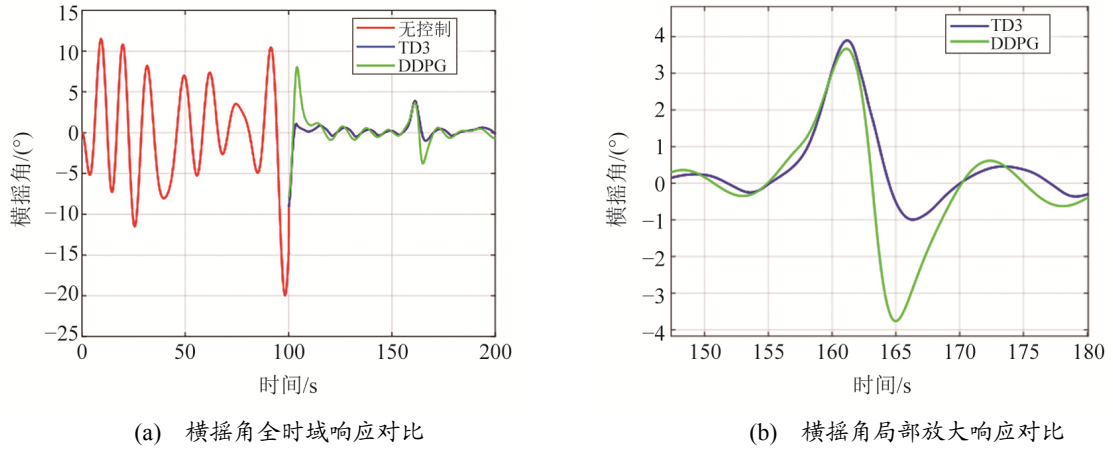


图9 SSN4 工况下横摇运动轨迹

Fig. 9 Roll Motion Trajectory Under SSN4 Condition

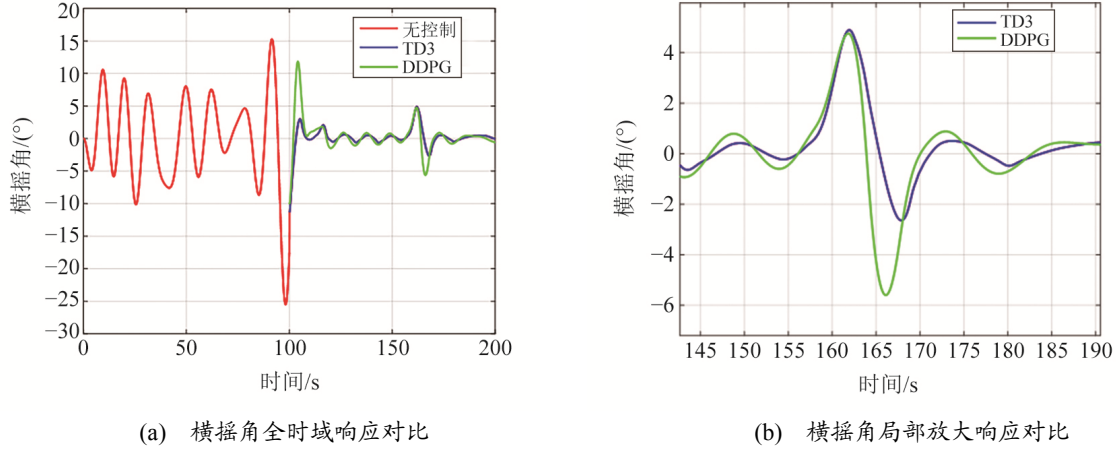


图10 SSN5 工况下横摇运动轨迹

Fig. 10 Roll Motion Trajectory Under SSN5 Condition

SSN3~SSN5工况下的鳍角变化分别见图11~图13。由图11~图13可知, 2种算法在同种工况下的控制效果均非常相似, 在幅值上略有差别; 但随着海况变大, DDPG算法在局部地方的鳍角变化波动较大。2种算法的主要区别集中在100 s左右, 在面对船舶大幅度横摇时, DDPG此时无法对减摇鳍进行准确控制, 从而导致减摇效果弱于TD3。从控制效果来看, TD3的控制效果相对更加平滑, 而DDPG则在某些位置出现了较大的波动, 鳍角变化更为激进, 且这种情况随着海况的增加越发明显。

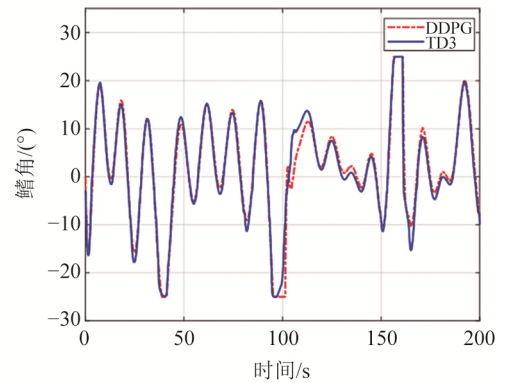


图11 SSN3 工况下鳍角变化

Fig. 11 Fin Angle Variation Under SSN3 Condition

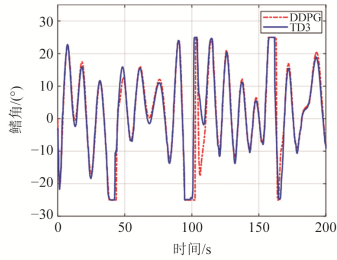


图 12 SSN4 工况下鳍角变化

Fig. 12 Fin Angle Variation Under SSN4 Condition

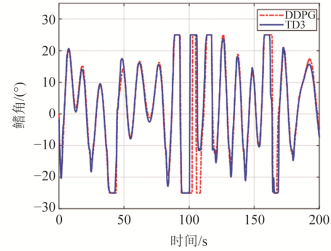


图 13 SSN5 工况下鳍角变化

Fig. 13 Fin Angle Variation Under SSN5 Condition

4 结论

本文提出基于双延迟深度确定性策略梯度 (TD3) 的控制方法, 建立船舶非线性横摇模型, 综合考虑海况、航速等因素设计状态空间和奖励函数以增强对动态环境的理解。结果表明, TD3 算法能够有效应对船舶在非线性横摇过程中遇到的各种扰动, 提供较为稳定的控制, 确保船舶在复杂海况下的横摇稳定性。基于 TD3 算法与 DDPG 算法设计的控制器均能有效的控制减摇幅度, 区别在于 TD3 算法的控制信号表现出更为平滑和稳定的特性, 避免了大幅波动, 确保了船舶横摇控制的连续性和可靠性。TD3 算法在不同工况下均表现出更优越的控制性能, 具有更好的适应性和鲁棒性, 能够稳定地保持船舶的横摇姿态。

参考文献:

- [1] PEREZ T. Ship Motion Control: Course Keeping and Roll Stabilisation Using Rudder and Fins[M]. Berlin: Springer Science & Business Media, 2006.
- [2] EYRES D J, BRUCE G J. Ship Construction[M]. Oxford: Butterworth-Heinemann, 2012.
- [3] SUN M X, LUAN T T, LIANG L H. RBF Neural Network Compensation-Based Adaptive Control for Lift-Feedback System of Ship Fin Stabilizers to Improve Anti-Rolling Effect[J]. Ocean Engineering, 2018, 163: 307-321.
- [4] FANG M C, ZHUO Y Z, LEE Z Y. The Application of the Self-Tuning Neural Network PID Controller on the Ship Roll Reduction in Random Waves[J]. Ocean Engineering, 2010, 37(7): 529-538.
- [5] NEJIM S. Design of Limited Authority Adaptive Ship

- Steering Autopilots[J]. International Journal of Adaptive Control and Signal Processing, 2000, 14(4): 381-391.
- [6] LIU C, WANG D Y, ZHANG Y X, et al. Model Predictive Control for Path Following and Roll Stabilization of Marine Vessels Based on Neurodynamic Optimization[J]. Ocean Engineering, 2020, 217: 107524.
- [7] LI R H, LI T S, BAI W W, et al. An Adaptive Neural Network Approach for Ship Roll Stabilization via Fin Control[J]. Neurocomputing, 2016, 173: 953-957.
- [8] SONG J G, LIANG L H, ZHANG S T, et al. Design and Experimental Investigation of a GA-Based Control Strategy for a Low-Speed Fin Stabilizer[J]. Ocean Engineering, 2020, 218: 108234.
- [9] 胡开业. 船舶在波浪中的大幅横摇运动及其运动稳定性研究[D]. 哈尔滨: 哈尔滨工程大学, 2011.
- HU K Y. Study on Large-Amplitude Roll Motion of Ships in Waves and its Motion Stability[D]. Harbin: Harbin Engineering University, 2011.
- [10] LUO J, ER G K, IU V P, et al. Non-Stationary Probabilistic Analysis of Non-Linear Ship Roll Motion Due to Modulated Periodic and Random Excitations[J]. Probabilistic Engineering Mechanics, 2024, 75: 103574.
- [11] 孙义洲, 许劲松, 叶睿. 水翼船的横摇控制[J]. 船舶工程, 2024, 46(1): 92-97.
- SUN Y Z, XU J S, YE R. Rolling Control for Hydrofoil Craft[J]. Ship Engineering, 2024, 46(1): 92-97.
- [12] MATSUI S, SUGIMOTO K, SHINOMOTO K. Simplified Estimation Formula for Frequency Response Function of Roll Motion of Ship in Waves[J]. Ocean Engineering, 2023, 276: 114187.
- [13] 李冰, 葛志刚. 船舶控制原理及其控制系统[M]. 哈尔滨: 哈尔滨工程大学出版社, 2021.
- LI B, QI Z G. Principles of Ship Control and its Control Systems[M]. Harbin: Harbin Engineering University Press, 2021.
- [14] 洪超, 陈莹霞. 船舶减摇技术现状及发展趋势[J]. 船舶工程, 2012, 34(增刊 2): 236-244.
- HONG C, CHENG Y X. Current Status and Development Trends of Ship Roll Reduction Technology[J]. Ship Engineering, 2012, 34(Suppl. 2): 236-244.
- [15] FUJIMOTO S, VAN H H, MEGER D. Addressing Function Approximation Error in Actor-Critic Methods[EB/OL]. [2018-06-21]. <https://arxiv.org/abs/1802.09477v3>.
- [16] HASSELT H V. Double Q-Learning[J]. Advances in Neural Information Processing Systems, 2010, 11: 2613-2621.
- [17] DUAN Y, CHEN X, HOUTHOOFT R, et al. Benchmarking Deep Reinforcement Learning for Continuous Control[C]// International Conference on Machine Learning. 2016.