

阮光维^{1,2} 李铁山^{1,3} 于仁海¹ 刘琪¹

基于执行依赖启发式动态规划的船舶减摇鳍在线最优控制

摘要

针对船舶线性横摇系统,设计了一种基于执行依赖启发式动态规划(ADHDP)方法的在线学习最优减摇鳍控制器.在设计过程中直接使用输入输出数据获取系统状态值.利用评价网络来逼近针对船舶减摇鳍控制系统设计的性能指标函数,并通过执行网络获得最优控制律,这两个网络都是多层前馈神经网络,即反向传播(BP)神经网络.在训练过程中,这两个神经网络不仅可以使用时测量数据,也可以减少船舶横摇模型的内部误差和不确定性干扰的影响,从而提高系统的鲁棒性.最后,仿真结果表明所提出的ADHDP控制器对于降低船舶横摇有很好的控制效果.

关键词

执行依赖启发式动态规划(ADHDP);自适应动态规划;船舶减摇鳍;最优控制

中图分类号 U664.36;TP273

文献标志码 A

收稿日期 2020-11-01

资助项目 国家自然科学基金(51939001,61976033);大连市重点学科重大课题科技创新基金(2018J11CY022);辽宁省兴辽英才计划高水平创新创业团队(XLYC1908018,XLYC1807046);辽宁省自然科学基金(20180550082,2019-ZD-0151);中央高校基本科研业务费项目(3132019345)

作者简介

阮光维,男,博士生,研究方向为智能船舶自动减摇鳍控制. nguyenquangduy@vimaru.edu.vn

李铁山(通信作者),男,博士,教授,主要从事智能船舶控制理论与技术、非线性系统智能控制理论与应用研究. tieshanli@126.com

1 大连海事大学 航海学院,大连,116026

2 越南海事大学 航海学院,越南 海防,180000

3 电子科技大学 自动化工程学院,成都,611731

0 引言

当船舶在海上航行时,由波浪、风和海流等外部干扰的影响而产生的横摇运动^[1]将极大地减弱船舶的航行效率,船员的日常活动和船舶安全也受到严重影响.因此,在船舶运动控制领域,如何尽可能地减小船舶的横摇动态具有重要意义.在最近的几十年中,随着船舶运动控制技术的不断发展,减摇鳍以其形式小巧、维修简单、减摇效率高优点被广泛应用于船舶减摇,并且获得了良好的控制性能.然而,该装置的效率取决于控制器设计,控制器的好坏将对减摇鳍系统有很大的影响.因此,减摇鳍控制系统对减摇鳍能否稳定工作有着极其重要的作用.目前,已经有许多控制方法被用来设计船舶减摇鳍控制器.例如:文献[2]提出了一种基于后推技术和闭环增益整形算法的非线性控制器;文献[3]提出了一种PID方法,用于船舶减摇鳍控制器设计,同时基于蒙特卡罗优化技术实现了船舶转向;文献[4]基于非线性扰动观测器,将自适应后推方法应用于非线性船舶横摇系统,并在降低船舶横摇方面获得了较好控制性能;文献[5]提出了一种简化的模糊控制器,用于船舶横摇系统,旨在提高船舶减摇鳍控制器的实时执行效率.

近年来,最优控制已在许多控制领域中得到广泛应用,并成为电机控制、机器人控制等自动控制领域的重要方法之一.在船舶减摇控制领域中,最优控制也有许多相关成果^[6-8].Adaptive Dynamic Programming (ADP)方法是由Miller等^[9]提出并广泛应用于控制领域的,其通常在结构上需要三个神经网络,分别是模型网络、评价网络和执行网络.ADP方法的基本思想是通过使用函数近似的结构(模糊模型、神经网络等)来逼近最优控制策略和最优性能指标函数.Miller等^[9]介绍了两种ADP结构,分别是双启发式动态规划(DHP)和启发式动态规划(HDP).后来,随着ADP方法的进一步发展,又有了其他结构,例如执行依赖启发式动态规划(ADHDP)、全局双重启发式动态规划(GDHP)、执行依赖双发式动态规划(ADDHP)、神经网络动态规划结构(NDP)^[10]、近似动态规划结构^[11]、自适应评价设计结构^[12]等.

受Si等^[13]和Xiao等^[14]的启发,本文提出了一种基于ADHDP(Action Dependent Heuristic Dynamic Programming)方法的新型船舶减

摇鳍在线学习最优控制方法. 首先对船舶横摇运动模型建立单自由度模型, 利用最优控制理论, 设计了基于 ADHDP 方法的船舶减摇鳍控制器. 在设计过程中不采用系统模型来获取未来的系统状态值, 而直接采用输入输出数据, 并使用两个 BP 神经网络来对最优控制律和性能指标函数进行逼近. 评价网络将用于逼近性能指标函数, 而执行网络将用于获取最优控制信号. 在训练过程中, 这两个神经网络不仅可以使使用实时测量数据, 也可以减少内部模型误差和不确定性干扰的影响, 从而提高控制精度和系统的鲁棒性.

本文的其余部分安排如下: 第 1 节提供了船舶横摇系统的线性模型和随机波浪扰动模型; 第 2 节给出了用于船舶横摇系统的 ADHDP 控制器设计; 第 3 节给出仿真结果; 第 4 节是总结.

1 准备知识和问题表述

1.1 船舶横摇线性模型

在海上航行时, 船舶的横摇运动主要是由波浪引起的. 当船舶的横摇角度小时, 可根据 Conolly 理论建立船舶线性横摇模型. 该模型被广泛用于船舶减摇鳍控制^[15], 通常描述如下:

$$(M_x + \Delta M_x)\ddot{\varphi} + 2N_u\dot{\varphi} + D_p h_m \varphi = -D_p h_m \alpha_1 - K_c, \quad (1)$$

其中, M_x 和 ΔM_x 分别是船舶的惯性力矩和附加的惯性力矩, φ 是船舶横摇角, $\dot{\varphi}$ 是船舶横摇角速度, h_m 是船的初始偏心高度, D_p 表示船舶的排水量, $2N_u$ 表示船侧倾运动的阻尼系数, α_1 代表有效波浪的倾角, K_c 表示作用在船上的力矩, 通常由减摇鳍产生, 可以用下面的等式表示:

$$K_c = 2Ll_f \cos(\alpha), \quad (2)$$

其中 L 是船的升力系数, 由减摇鳍产生, l_f 是减摇鳍的横摇力臂, α 是鳍角度, 这是减摇鳍的垂直轴与中心线之间的角度.

将式(1)重新整理得:

$$\ddot{\varphi} = \frac{1}{(M_x + \Delta M_x)} (-2N_u\dot{\varphi} - D_p h_m \varphi - D_p h_m \alpha_1 - K_c). \quad (3)$$

定义状态变量 $\mathbf{x} = [x_1, x_2]^T = [\varphi, \dot{\varphi}]^T$, 则式(3)的状态空间形式表示为

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = A_1 x_1 + A_2 x_2 + B_1 (D_p h_m \alpha_1 + K_c), \end{cases} \quad (4)$$

其中 A_1, A_2, B_1 是系数, 分别表示如下:

$$A_1 = -\frac{D_p h_m}{(M_x + \Delta M_x)}, \quad A_2 = -\frac{2N_u}{(M_x + \Delta M_x)},$$

$$B_1 = -\frac{1}{(M_x + \Delta M_x)}. \quad (5)$$

由式(4)得如下状态空间方程:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{C}\mathbf{w}(t), \quad (6)$$

其中 $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ A_1 & A_2 \end{bmatrix} \in \mathbf{R}^{2 \times 2}$ 是变量矩阵, $\mathbf{B} = \mathbf{C} =$

$\begin{bmatrix} 0 \\ B_1 \end{bmatrix} \in \mathbf{R}^{2 \times 1}$ 是输入矩阵. 此外, \mathbf{w} 表示外部干扰, 并且 \mathbf{u} 表示系统的输入信号, 即由减摇鳍产生的控制力矩.

1.2 波浪干扰模型

本文采用了基于 Pierson-Moscowitz 谱的波动模型^[16], 其描述如下:

$$S_t(\omega_a) = \frac{0.0081 \times g^2}{\omega_a^5} \exp\left(-\frac{3.11}{H_{1/3}^2 \omega_a^4}\right), \quad (7)$$

其中, $H_{1/3}$ 是波高, ω_a 是波频率, g 表示重力加速度.

根据随机过程理论, 可将波的倾斜角模型描述如下:

$$a(t) = \sum_{l=1}^n (\sqrt{2S_p(\omega_{al}) \Delta\omega_a} \cos(\omega_{al}t + \mu_l)), \quad (8)$$

其中 μ_l 表示波倾斜角初始相位, 其值在 $0 \sim 2\pi$ 之间随机变化. 为方便计算, 本文选择 $\mu_l = 0$, 则能谱与波高谱之间的关系式为

$$S_p(\omega_a) = G_1^2 G_2^2 \frac{\omega_a^4}{g^2} S_t(\omega_a), \quad (9)$$

其中 G_1 和 G_2 是系数, 根据船体的形状而变化. 考虑到船舶航速和航向对船舶所遭遇的海浪频率的影响, 可以将船舶的所谓相遇频率描述如下:

$$\omega_e = \omega_a - \frac{\omega_a^2}{g} V \cos(\beta), \quad (10)$$

其中 V 是船速, β 是海浪的方向. 根据能量等价原理, 遇到频率能谱密度函数与自然频率能谱密度函数有以下关系:

$$S_p(\omega_e) = S_p(\omega_a) / \left(1 - \frac{2\omega_a}{g} V \cos(\beta)\right), \quad (11)$$

因此实际仿真中的波倾角模型如下:

$$a_e(t) = \sum_{l=1}^n (\sqrt{2S_p(\omega_{el}) \Delta\omega_e} \cos(\omega_{el}t + \mu_l) \sin(\beta)). \quad (12)$$

使用上述公式进行模拟, 当 $(1 - 2\omega_a V \cos(\beta)) / g = 0$ 时, $S_p(\omega_e)$ 中将会产生断点, 因此具有以下关系式:

$$S_p(\omega_a) \Delta\omega_a = S_p(\omega_e) \Delta\omega_e. \quad (13)$$

将式(13)代入式(12), 可得:

$$a_e(t) = \sum_{l=1}^n (\sqrt{2S_p(\omega_{al})} \Delta\omega_a \cos(\omega_{al}t + \mu_l) \sin(\beta)). \quad (14)$$

图 1 所示为随机海浪,其中 $H_{1/3} = 5.8$, $n = 28$, 波浪方向 $\beta = 45^\circ$, 船速 $V = 7.8$ m/s

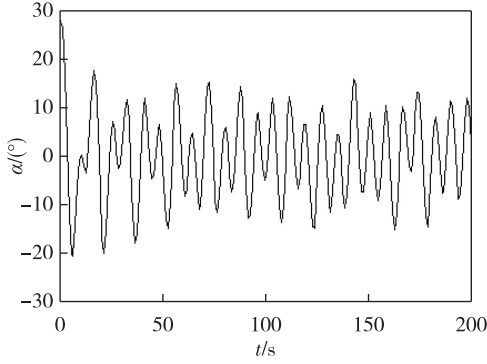


图 1 随机海浪
Fig. 1 Random waves

2 船舶减摇鳍系统的 ADHDP 控制器设计

线性模型(6)可以写成如下的离散时间形式:

$$\mathbf{x}(t+1) = \mathbf{A}_d \mathbf{x}(t) + \mathbf{B}_d \mathbf{u}(t) + \mathbf{C}_d \mathbf{w}(t), \quad (15)$$

其中 $\mathbf{x}(t)$ 是系统的状态变量, \mathbf{A}_d 是系统状态矩阵, $\mathbf{B}_d, \mathbf{C}_d$ 是系统的输入矩阵, $\mathbf{w}(t)$ 表示外部干扰, $\mathbf{u}(t)$ 用于表示系统的输入信号. 本文的任务是设计最优控制器 $\mathbf{u}(t)$ 以最大程度地减少如下的性能指标函数:

$$J^*(\mathbf{x}(t), \mathbf{u}(t)) = \min_{\mathbf{u}(t)} \left(\sum_{i=t}^{\infty} \gamma^{i-t} U(\mathbf{x}(i), \mathbf{u}(i)) \right), \quad (16)$$

其中 γ 是性能指标的折扣因子, $0 < \gamma \leq 1$. 为简单起见, 本文中 $\gamma = 1$. 函数 $U(\mathbf{x}(i), \mathbf{u}(i)) = \mathbf{x}^T(i) \mathbf{Q} \mathbf{x}(i) + \mathbf{u}^T(i) \mathbf{R} \mathbf{u}(i)$ 是效用函数, 其中 \mathbf{Q} 是一个半正定矩阵, 而 \mathbf{R} 是一个正定矩阵. 根据 Bellman 原理, 最优性能指标函数 $J^*(\mathbf{x}(t), \mathbf{u}(t))$ 可以写成如下形式:

$$J^*(\mathbf{x}(t), \mathbf{u}(t)) = \min_{\mathbf{u}(t)} (U(\mathbf{x}(t), \mathbf{u}(t)) + J(\mathbf{x}(t+1), \mathbf{u}(t+1))), \quad (17)$$

该方程也被称为离散时间 Hamilton-Jacobi-Bellman (DTHJB) 方程. 通过最小化性能指标函数, 可以得到最优控制策略:

$$\mathbf{u}^*(t) = \arg \min_{\mathbf{u}(t)} (U(\mathbf{x}(t), \mathbf{u}(t)) + J(\mathbf{x}(t+1), \mathbf{u}(t+1))). \quad (18)$$

ADHDP 在结构上需要使用两个神经网络: 一个是评价网络用于估计性能指标函数 $J(\mathbf{x}(t), \mathbf{u}(t))$, 另一个是执行网络用于逼近控制信号 $\mathbf{u}(t)$. 图 2 为

船舶减摇鳍系统最优 ADHDP 控制器示意图.

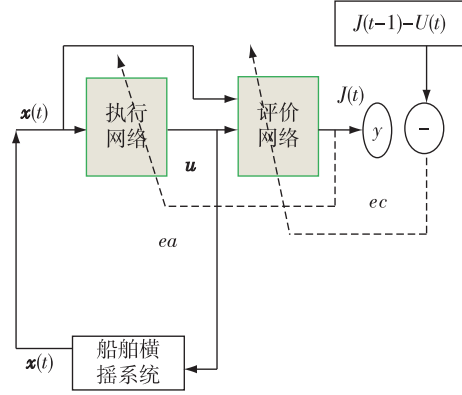


图 2 船舶减摇鳍系统的 ADHDP 结构
Fig. 2 ADHDP structure for ship fin stabilizer system

图 2 中, 实线表示信号流, 虚线表示评价网和执行网的参数更新路径. 系统的状态变量是执行网的输入, 它将输出控制信号 $\mathbf{u}(t)$ 之后, 该控制信号与系统的状态变量相结合, 成为评价网的输入, 它将在 (17) 中输出目标函数的近似值. $\mathbf{u}(t)$ 也是减摇鳍系统的控制输入. 评价网和执行网是基于 BP 神经网络的非线性多层反馈网络. 下一部分将给出评价网和执行网的详细结构.

2.1 评价网

在 ADHDP 方法中, 本文使用了评价网来逼近性能指标函数 $J(t)$. 这里设计一个 3-6-1 结构的评价网 (图 3), 其中包含 3 个输入神经元、6 个隐含层神经元和 1 个输出神经元. 输入信号 $\mathbf{x}_1, \mathbf{x}_2$ 是系统的状态 (船舶横摇角和船舶横摇角速度), 控制信号 $\mathbf{u}(t)$ 从执行网的输出信号中获得. 评价网的激活函数采用 sigmoid 函数, 评价网结构如下所示:

$$\tan \text{sig}(x) = \frac{1 - e^{-x}}{1 + e^{-x}}. \quad (19)$$

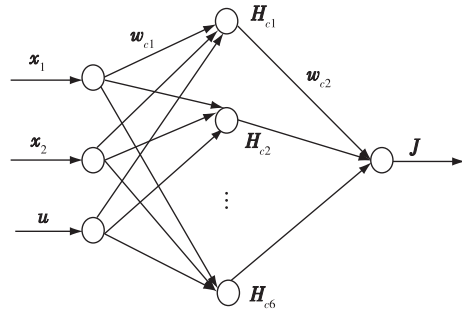


图 3 评价网络的结构
Fig. 3 Structure of critic network

评价网的输出层可以表示为

$$\begin{cases} \mathbf{n}_c(t) = \mathbf{w}_{c1} \mathbf{I}_c(t), \\ \mathbf{h}_c(t) = \tan \operatorname{sig}(\mathbf{w}_{c1} \mathbf{I}_c(t)) = \frac{1 - e^{-n_c(t)}}{1 + e^{-n_c(t)}}, \\ \mathbf{J}(t) = \mathbf{w}_{c2} \mathbf{h}_c(t), \end{cases} \quad (20)$$

其中 $\mathbf{I}_c(t) \in \mathbf{R}^{3 \times 1}$ 是输入向量, $\mathbf{J}(t)$ 是输出向量, $\mathbf{h}_c(t) \in \mathbf{R}^{1 \times 6}$ 是隐含层的输出向量, $\mathbf{w}_{c1} \in \mathbf{R}^{6 \times 3}$ 是输入到隐含层的权重向量, $\mathbf{w}_{c2} \in \mathbf{R}^{1 \times 6}$ 是隐含层到输出层的权重向量.

在评价网中,前向过程计算目标函数的近似值,而后向过程则更新评价网的权重矩阵.定义评价网误差函数为

$$\mathbf{e}_c(t) = \gamma \mathbf{J}(t) - \mathbf{J}(t-1) + \mathbf{U}(t), \quad (21)$$

评价网中的目标函数为

$$\mathbf{E}_c(t) = \frac{1}{2} \mathbf{e}_c^2(t). \quad (22)$$

评价网络的权重更新算法基于梯度下降方法,它是通过应用链式规则获得的.输入层到隐含层的权值更新表示为

$$\mathbf{w}_{c1}(t+1) = \mathbf{w}_{c1}(t) + \Delta \mathbf{w}_{c1}(t), \quad (23)$$

其中 $\Delta \mathbf{w}_{c1}(t) \in \mathbf{R}^{6 \times 3}$,

$$\begin{aligned} \Delta \mathbf{w}_{c1}^{ij}(t) &= -l_c \left(\frac{\partial \mathbf{E}_c(t)}{\partial \mathbf{w}_{c1}^{ij}(t)} \right) = \\ &= -l_c \left(\frac{\partial \mathbf{E}_c(t)}{\partial \mathbf{J}(t)} \frac{\partial \mathbf{J}(t)}{\partial \mathbf{h}_{ci}(t)} \frac{\partial \mathbf{h}_{ci}(t)}{\partial \mathbf{n}_{ci}(t)} \frac{\partial \mathbf{n}_{ci}(t)}{\partial \mathbf{w}_{c1}^{ij}(t)} \right) = \\ &= -l_c \gamma \mathbf{e}_c(t) \mathbf{w}_{c2}^i(t) \left[\frac{1}{2} (1 - \mathbf{h}_{ci}^2(t)) \right] \mathbf{x}_j(t), \end{aligned} \quad (24)$$

其中 i 是矩阵行号, j 是矩阵列号, $l_c > 0$ 是评价网的学习率.

隐含层到输出层的权重向量更新表达式为

$$\mathbf{w}_{c2}(t+1) = \mathbf{w}_{c2}(t) + \Delta \mathbf{w}_{c2}(t), \quad (25)$$

其中 $\Delta \mathbf{w}_{c2}(t) \in \mathbf{R}^{1 \times 6}$,

$$\begin{aligned} \Delta \mathbf{w}_{c2}^i(t) &= -l_c \left(\frac{\partial \mathbf{E}_c(t)}{\partial \mathbf{w}_{c2}^i(t)} \right) = \\ &= -l_c \left(\frac{\partial \mathbf{E}_c(t)}{\partial \mathbf{J}(t)} \frac{\partial \mathbf{J}(t)}{\partial \mathbf{w}_{c2}^i(t)} \right) = -l_c \gamma \mathbf{e}_c(t) \mathbf{h}_{ci}(t), \end{aligned} \quad (26)$$

其中 i 是权重矩阵的行数.

注1 在式(24)和(26)中, $\partial \mathbf{E}_c(t) / \partial \mathbf{J}(t)$ 由式(21)和(22)计算,而 $\partial \mathbf{J}(t) / \partial \mathbf{w}_c(t)$ 与状态变量和控制输入有关.

2.2 执行网

本文采用执行网获取控制信号 $\mathbf{u}(t)$,设计了一

个2-6-1结构的执行网络(图4),其中包含2个输入神经元、6个隐含层神经元和1个输出神经元.包括 $\mathbf{x}_1, \mathbf{x}_2$ 的输入信号是系统的状态变量,输出信号是系统的控制信号,同时使用 sigmoid 函数作为激活函数.

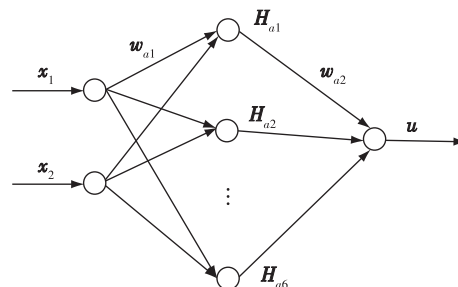


图4 执行网络的结构

Fig. 4 Structure of action network

执行网络的输出信号可以确定为

$$\begin{cases} \mathbf{n}_a = \mathbf{w}_{a1} \mathbf{I}_a(t), \\ \mathbf{h}_a(t) = \tan \operatorname{sig}(\mathbf{w}_{a1} \mathbf{I}_a(t)) = \frac{1 - e^{-n_a}}{1 + e^{-n_a}}, \\ \mathbf{v}(t) = \mathbf{w}_{a2} \mathbf{h}_a(t), \\ \mathbf{u}(t) = \tan \operatorname{sig}(\mathbf{v}(t)), \end{cases} \quad (27)$$

其中 $\mathbf{I}_a(t) \in \mathbf{R}^{2 \times 1}$ 是输入向量, $\mathbf{u}(t)$ 是输出向量, $\mathbf{h}_a(t)$ 是隐含层的输出向量, $\mathbf{w}_{a1} \in \mathbf{R}^{6 \times 2}$ 是输入层到隐含层的权重向量, $\mathbf{w}_{a2} \in \mathbf{R}^{1 \times 6}$ 是隐含层到输出层的权重向量.

定义执行网误差函数为

$$\mathbf{e}_a(t) = \mathbf{J}(t). \quad (28)$$

最小化下面的目标函数,以更新执行网络中的权重:

$$\mathbf{E}_a(t) = \frac{1}{2} \mathbf{e}_a^2(t). \quad (29)$$

执行网络的权重更新算法是基于梯度下降方法,通过应用链式规则给出的.如下所示输入层到隐含层的权值更新表示为

$$\mathbf{w}_{a1}(t+1) = \mathbf{w}_{a1}(t) + \Delta \mathbf{w}_{a1}(t), \quad (30)$$

其中 $\Delta \mathbf{w}_{a1}(t) \in \mathbf{R}^{6 \times 3}$,

$$\begin{aligned} \Delta \mathbf{w}_{a1}^{ij}(t) &= -l_a \left(\frac{\partial \mathbf{E}_a(t)}{\partial \mathbf{w}_{a1}^{ij}(t)} \right) = \\ &= -l_a \left(\frac{\partial \mathbf{E}_a(t)}{\partial \mathbf{J}(t)} \frac{\partial \mathbf{J}(t)}{\partial \mathbf{u}(t)} \frac{\partial \mathbf{u}(t)}{\partial \mathbf{v}(t)} \frac{\partial \mathbf{v}(t)}{\partial \mathbf{h}_a(t)} \frac{\partial \mathbf{h}_a(t)}{\partial \mathbf{n}_a(t)} \frac{\partial \mathbf{n}_a(t)}{\partial \mathbf{w}_{a1}^{ij}(t)} \right) = \\ &= -l_a \gamma \mathbf{e}_a(t) \left[\frac{1}{2} (1 - \mathbf{u}^2(t)) \right] \mathbf{w}_{a2}^i \left[\frac{1}{2} (1 - \mathbf{h}_{ai}^2(t)) \right] \cdot \\ &= \mathbf{x}_j(t) \sum_{i=1}^{H_l} \left[\mathbf{w}_{c2}^i \frac{1}{2} (1 - \mathbf{h}_{ci}^2(t)) \mathbf{w}_{c1}^{i,n} \right], \end{aligned} \quad (31)$$

其中 i 是矩阵行数, j 是矩阵列数, $l_a > 0$ 是执行网的学习率, H_l 是执行网中隐藏节点的数量, n 是评价网的输入信号的数量.

隐含层到输出层的权重更新表达式为

$$\mathbf{w}_{a2}(t+1) = \mathbf{w}_{a2}(t) + \Delta \mathbf{w}_{a2}(t), \quad (32)$$

其中 $\Delta \mathbf{w}_{a2}(t) \in \mathbf{R}^{1 \times 6}$,

$$\begin{aligned} \Delta \mathbf{w}_{a2}^i(t) = & -l_a \left(\frac{\partial E_a}{\partial \mathbf{w}_{a2}^i} \right) = -l_a \left(\frac{\partial E_a}{\partial \mathbf{J}} \frac{\partial \mathbf{J}}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial \mathbf{v}} \frac{\partial \mathbf{v}}{\partial \mathbf{w}_{a1}^i} \right) = \\ & -l_a \gamma e_a(t) \left[\frac{1}{2} (1 - \mathbf{u}^2(t)) \right] \mathbf{h}_{ai}(t) \cdot \\ & \sum_{i=1}^{H_l} \left[\mathbf{w}_{c2}^i \frac{1}{2} (1 - \mathbf{h}_{ci}^2(t)) \mathbf{w}_{c1}^{i,n} \right]. \quad (33) \end{aligned}$$

注 2 在式(31)和(33)中, $\partial E_a(t)/\partial \mathbf{J}(t)$ 由式(28)和(29)计算, 而 $\partial \mathbf{J}(t)/\partial \mathbf{u}(t)$ 通过链式规则获得, 该规则与执行网中的权重矩阵相关^[14].

3 仿真研究

本节将对容器进行仿真, 船舶参数如表 1 所示.

表 1 船舶参数

Table 1 Vessel parameters

参数	数值
船两柱间长/m	79.38
船宽/m	9.6
吃水/m	3.1
排水量/t	1100
减摇鳍面积/m ²	3.8
减摇鳍力臂/m	5.6
升力梯度系数	0.06
船舶进水角/(°)	48
船舶初稳性高度/m	1.012
设计航速/knot	18

船舶横摇动态的状态空间方程如下:

$$\begin{aligned} \mathbf{x}(t+1) = & \begin{bmatrix} 0 & 1 \\ -0.56 & -0.18 \end{bmatrix} \mathbf{x}(t) + \\ & \begin{bmatrix} 0 \\ 0.05 \end{bmatrix} \mathbf{u}(t) + \begin{bmatrix} 0 \\ 0.05 \end{bmatrix} \mathbf{w}(t), \quad (34) \end{aligned}$$

效用函数可以设计如下:

$$U(\mathbf{x}(i), \mathbf{u}(i)) = \mathbf{x}^T(i) \mathbf{Q} \mathbf{x}(i) + \mathbf{u}^T(i) \mathbf{R} \mathbf{u}(i), \quad (35)$$

其中 \mathbf{Q} 是一个半正定矩阵, 可以选择 $\mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, \mathbf{R}

是一个正定矩阵, 可以选择 $\mathbf{R} = [1]$.

本文设计了具有两个神经网络的 ADHDP 控制器, 每个神经网络的参数如下: 评价网为 3-6-1 结构,

具有 3 个输入神经元、6 个隐含层神经元和 1 个输出神经元, 执行网为 2-6-1 结构, 有 2 个输入神经元、6 个隐含层神经元和 1 个输出神经元; 评价网和执行网的学习率 $l_c = l_a = 0.005$; 初始权重矩阵 $\mathbf{w}_{c1} \in \mathbf{R}^{6 \times 3}$ 和 $\mathbf{w}_{a1} \in \mathbf{R}^{6 \times 2}$ 分别为评价网和执行网中输入层到隐含层的权值矩阵, 初始权重矩阵 $\mathbf{w}_{c2} \in \mathbf{R}^{1 \times 6}$ 和 $\mathbf{w}_{a2} \in \mathbf{R}^{1 \times 6}$ 分别为评价网和执行网中的隐含层到输出层的权值矩阵, 权值在 $[-1, 1]$ 中随机选择. 初始权重矩阵选择如下:

$$\mathbf{w}_{a01} = \begin{bmatrix} -0.8317 & -0.8883 \\ -0.7334 & -0.0205 \\ -0.6522 & 0.0317 \\ -0.2195 & -0.1437 \\ 0.6640 & 0.2987 \\ 0.6081 & 0.2380 \end{bmatrix}$$

$$\mathbf{w}_{a02} = \begin{bmatrix} -0.4529 \\ -0.1441 \\ -0.9654 \\ 0.9614 \\ -0.6534 \\ -0.7778 \end{bmatrix},$$

$$\mathbf{w}_{c01} = \begin{bmatrix} -0.1326 & -0.9634 & 0.6574 \\ -0.7001 & 0.3316 & 0.6430 \\ 0.9005 & -0.9623 & 0.4568 \\ 0.4361 & -0.8488 & -0.7025 \\ 0.0163 & 0.0550 & 0.3161 \\ -0.1380 & -0.8668 & 0.0526 \end{bmatrix},$$

$$\mathbf{w}_{c02} = \begin{bmatrix} 0.1481 \\ 0.0467 \\ 0.5794 \\ -0.0733 \\ -0.1311 \\ 0.8189 \end{bmatrix}.$$

为了验证所设计控制器的有效性, 本文设计了船舶减摇鳍的线性二次调节器控制器来进行比较. 线性二次调节器控制器 (Linear Quadratic Regulator, LQR) 表示如下:

$$\mathbf{u}_{lqr}^*(t) = -\mathbf{K} \mathbf{x}(t), \quad (36)$$

其中 $\mathbf{K} = [\mathbf{B}_d^T \mathbf{P} \mathbf{B}_d + \mathbf{R}]^{-1} \mathbf{B}_d^T \mathbf{P} \mathbf{A}_d$ 是状态反馈的增益, \mathbf{P} 是 Riccati 方程的半正定解, $\mathbf{P} = \mathbf{A}_d^T (\mathbf{P} - \mathbf{P} \mathbf{B}_d (\mathbf{B}_d^T \mathbf{P} \mathbf{B}_d + \mathbf{R})^{-1} \mathbf{B}_d^T \mathbf{P}) \mathbf{A}_d + \mathbf{Q}$.

图 5 为具有减摇鳍和不具有减摇鳍控制器的船舶横摇角及横摇速率的变化情况. 由图 5 可见, 设计

的减摇鳍控制器的有效性较好.图 6 为本文设计的控制器(ADHDP)和线性二次调节器控制器(LQR)的比较,表明本文设计的控制器获得比较好的控制效果.图 7 和图 8 分别为评价网络和执行网络的权重收敛过程,显然权重都收敛到常数,这表明所提出的控制器具有很好的性能.

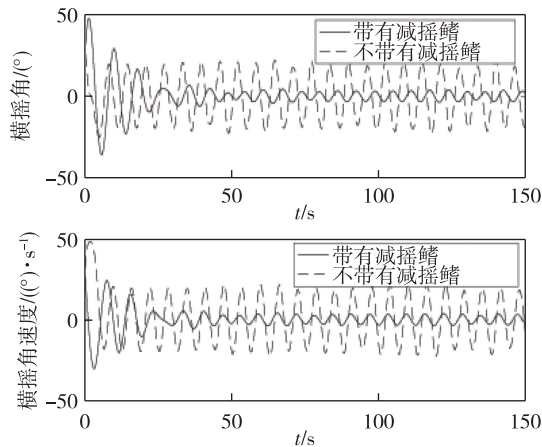


图 5 具有减摇鳍和不具有减摇鳍的横摇角和横摇速率的变化情况

Fig. 5 Roll angle and roll rate with and without fin stabilizer

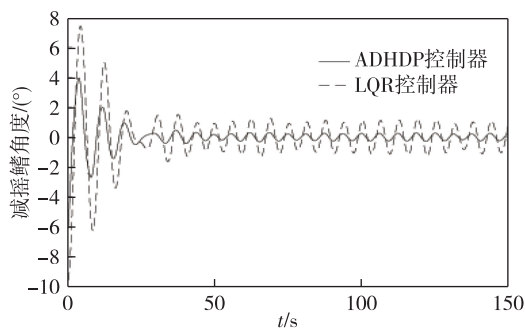


图 6 ADHDP 和 LQR 控制器对比

Fig. 6 Comparison between ADHDP and LQR controller

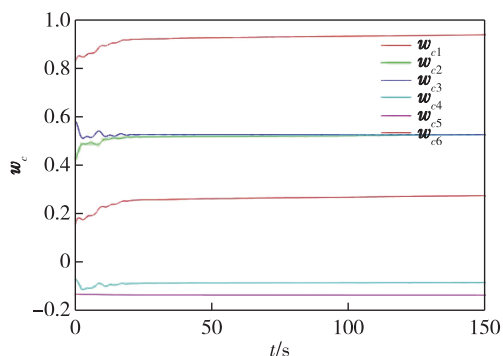


图 7 评价网络的权值矩阵

Fig. 7 Weight matrices of critic network

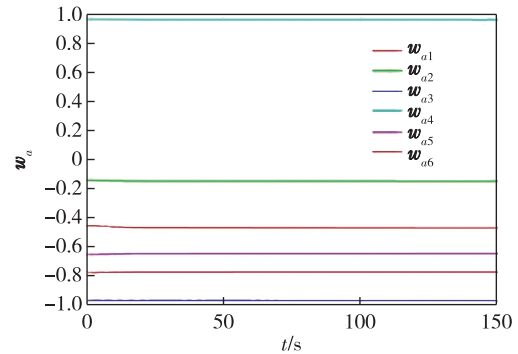


图 8 执行网络的权重矩阵

Fig. 8 Weight matrices of action network

4 结束语

本文提出了一种基于 ADHDP 方法的船舶减摇鳍系统的最优控制器.本文所提出的方法是利用评价网络来逼近针对船舶减摇鳍控制系统设计的性能指标函数,并通过执行网络来获得最优控制律. ADHDP 的优点是在线训练,ADHDP 结构中的两个神经网络既可以使用实时测量数据,同时也可以减少内部模型误差和不确定性干扰的影响,从而提高了鲁棒性.仿真结果验证了该方法良好的减摇效果.

参考文献

References

- [1] Perez T.Ship motion control:course keeping and roll stabilisation using rudder and fins [M]. London: Springer-Verlag,2005
- [2] 王新屏,张显库.基于 Backstepping 与闭环增益成形的减摇鳍控制 [J].大连海事大学学报,2008,34(3):89-92
WANG Xinping,ZHANG Xianku.Fin stabilizer control based on backstepping and closed-loop gain shaping algorithms [J].Journal of Dalian Maritime University,2008,34(3):89-92
- [3] Liang L H,Zhao P,Zhang S T.Roll reduction control during ship turns using fin stabilizers with PID controller based on Monte Carlo optimization [C] // IEEE International Conference on Mechatronics and Automation,2018:749-754
- [4] Bai W W,Li T S,Lu Z K.Adaptive backstepping-based nonlinear disturbance observer for fin stabilizer system [C] // International Joint Conference on Neural Networks,2014:1258-1264
- [5] Sun S A,Hu J Q,Yin J C,et al.Design of simplified fuzzy controller for ship fin stabilizer [C] // Proceedings of the 33rd Chinese Control Conference,2014:4534-4538
- [6] Yin Y,Xia L,Li H K.The stochastic optimal control of fin stabilizers under the ship IPMS networks [C] // IEEE International Conference on Mechatronics and Automation,

- 2010:969-972
- [7] Zhang P, Yang H S, Yang Y. Multi-dimensional Taylor network optimal control for ship roll stabilization [C] // IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference, 2017:966-970
- [8] Jin H Z, Liu Z Q, Qi Z G, et al. Design of fin stabilizers control system with optimal added resistance [C] // Proceedings of the 32nd Chinese Control Conference, 2013: 7525-7529
- [9] Miller W T, Sutton R S, Werbos P J. A menu of designs for reinforcement learning over time [M] // Neural Networks for Control. Cambridge, MA: The MIT Press, 1995
- [10] Mu C X, Wang D, He H B. Novel iterative neural dynamic programming for data-based approximate optimal control design [J]. Automatica, 2017, 81 : 240-252
- [11] Xie Q Q, Luo B, Tan F X. Optimal tracking control for ship course using approximate dynamic programming method [C] // Proceedings of the 32nd Chinese Control Conference, 2013: 2911-2916
- [12] Tan F X, Guan X P. Kernel-based adaptive critic designs for optimal control of nonlinear discrete-time system [C] // 2018 37th Chinese Control Conference, 2018: 2167-2172
- [13] Si J, Wang Y T. Online learning control by association and reinforcement [J]. IEEE Transactions on Neural Networks, 2001, 12(2) : 264-276
- [14] Xiao G Y, Zhang H G, Luo Y H. Online optimal control of unknown discrete-time nonlinear systems by using time-based adaptive dynamic programming [J]. Neurocomputing, 2015, 165: 163-170
- [15] Zhang S T, Zhao P, Liang L H. LQR-based ship roll reduction control using fin stabilizer [C] // IEEE International Conference on Mechatronics and Automation, 2018: 1031-1036
- [16] Li H, Guo C, Li X F. 3-D visual simulation of irregular ocean waves based on Matlab [J]. Acta Simulata Systematica Sinica, 2003, 7(15) : 1057-1059

Online optimal control for ship fin stabilizer system based on action dependent heuristic dynamic programming

NGUYEN Quangduy^{1,2} LI Tieshan^{1,3} YU Renhai¹ LIU Qi¹

1 Navigation College, Dalian Maritime University, Dalian 116026

2 Faculty of Navigation, Vietnam Maritime University, Haiphong 180000, Vietnam

3 School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 611731

Abstract When ships are sailing on the sea, roll motion will greatly reduce the safety of ships and cargo, as well as the health of the crew. Therefore, the ship roll stabilization device has become one of the indispensable equipment on the ship. As an active roll reduction device, fin stabilizer is widely used for roll reduction due to its good anti-rolling performance. In this paper, an online learning optimal controller based on action dependent heuristic dynamic programming (ADHDP) is proposed for the ship fin stabilizer system. The input and output data, instead of a system model, are used in the design to obtain the system state. Two back propagation neural networks, including a critic network and an action network, are used to approximate the performance function and obtain the control law, respectively. The two neural networks can use real-time measurement data, and reduce internal model error and the uncertainty disturbance, thus improve the robustness of the system. Finally, the effectiveness of the proposed ADHDP controller is validated by simulation results.

Key words action dependent heuristic dynamic programming (ADHDP); adaptive dynamic programming; ship fin stabilizer; optimal control