https://doi.org/10.1051/jnwpu/20213930477

基于 RBF 网络 Q 学习的 AUV 路径跟踪控制方法

李泽宇, 刘卫东, 李乐, 张文博, 郭利伟

(西北工业大学 航海学院, 陕西 西安 710072)

摘 要:水下回收过程中,AUV 航行速度受到多种因素影响而产生变化,艉部操纵舵效随之改变,直接影响了 AUV 回收路径跟踪控制性能。根据 AUV 航行状态,采用强化学习方法对 AUV 控制器进行自主学习优化,能够改善 AUV 航向及深度响应的性能指标,提高路径跟踪控制性能。建立 AUV 路径跟踪导引律,设计航向及俯仰运动滑模控制器,保证系统对外扰动的鲁棒性;采用 Q 学习方法,根据 AUV 航速、跟踪误差及其变化率,对滑模控制参数进行离线训练优化,搭建 RBF 网络加快训练过程,避免"维数灾"现象;将训练得到的 RBF-Q 学习网络应用于在线控制,与传统滑模控制器进行跟踪控制对比。仿真结果验证了算法的有效性。

关 键 词:自主水下航行器;路径跟踪;强化学习;RBF 神经网络

中图分类号:TP273 **文献标志码**:A **文章编号**:1000-2758(2021)03-0477-07

水下回收技术极大地提高了 AUV 续航能力,路径精确跟踪是实现水下回收的关键[1]。欠驱动 AUV 通过速度和姿态角耦合控制调整自身航行位置,从而实现路径跟踪控制。在回收的不同阶段,AUV 的实际航行速度会产生较大变化,为路径精确跟踪带来困难[2-3]。

许多传统控制算法都曾被应用到 AUV 路径跟踪控制问题当中,Min 等[4]设计了 AUV 路径点跟踪 PD 控制器,鲁棒性较差;张磊[5]采用遗传算法对模糊跟踪控制器进行了优化,算法性能对专家知识依赖较大;王宏健等[6]采用二阶滤波避免了反步法解析求导的复杂过程,能够实现 AUV 路径的精确跟踪,但系统鲁棒性较弱;王金强等[7-8]针对一种新型飞翼式 AUV 的位置跟踪问题,采用自适应神经网络对滑模控制器进行优化补偿,但未考虑 AUV 航速变化对路径跟踪控制的影响,且一组控制参数难以满足不同航行工况的控制需求。近年来随着相关技术进步,模型预测和强化学习等基于数据的算法同样被应用到 AUV 的跟踪控制当中。Shen 等[9]采用滚动优化方法对 AUV 进行了路径优化与跟踪控制一体化设计;Sun 等[10]采用深度强化学习方法设计了 AUV 三维路径跟踪控制器;Shi 等[11]采用多逆伪 Q 学习算法设计了轨迹跟踪控制算法;姚绪梁等[12]采用模型预测控制与滑模控制方法对 AUV 直线路径进行了仿真跟踪。这些方法对 AUV 模型准确性依赖较低,对外干扰也具有良好的鲁棒性,但算法结构复杂,经大量运算才能得到最优控制律,难以直接应用到工程实践。若将传统算法与学习网络相结合,对多种航行工况进行离线训练,将训练后的学习参数应用于在线控制,可降低 AUV 实时优化的运算量,加快实时运动控制收敛速度,提高 AUV 在不同航速及外扰动下的路径跟踪性能。

1 跟踪导引律及姿态控制器设计

欠驱动 AUV 具有左右对称性,AUV 与水下回收装置处在同一深度上,有利于 AUV 回收对接。AUV 水平面路径跟踪控制原理如图 1 所示,已知一条 AUV 全局回收路径曲线 S[3],离散化路径点序列为 $P(p_1,p_2,\cdots,p_n)$,其中 $p_i(\xi_i,\zeta_0,\eta_i)$,$i=1,2,\cdots,n$,路径点深度 ζ_0 保持不变。AUV 路径跟踪控制转

收稿日期:2020-09-07 **基金项目**:国家自然科学基金(61903304)与国家重点研发计划项目(2016YFC0301700)资助

作者简介:李泽宇(1992—),西北工业大学博士研究生,主要从事水下航行器回收控制方法研究。

通信作者:李乐(1986—),西北工业大学助理教授,主要从事水下机器人协同控制研究。e-mail: leli@nwpu.edu.cn

化以下 2 个子问题组合控制：即调整目标路径航向 ψ_{cmd} 的导引控制，以及对 ψ_{cmd} 的航向角跟踪控制。

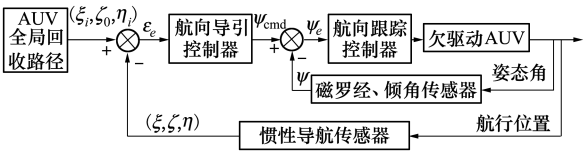


图 1 欠驱动 AUV 水平面路径跟踪控制原理图

为实现水平面路径跟踪，AUV 需要保持在指令深度 ζ_0 处航行，控制原理如图 2 所示。

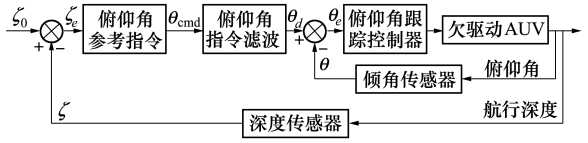


图 2 AUV 深度控制原理图

1.1 AUV 路径跟踪导引律

水平面路径跟踪示意如图 3 所示，$p_i(\xi_i,\zeta_0,\eta_i)$ 和 $p_{i+1}(\xi_{i+1},\zeta_0,\eta_{i+1})$ 为相邻的 2 个路径点，两点连线即为期望路径，ψ_0，ψ 分别表示路径切线方向和 AUV 当前航向，ε_e 表示 AUV 当前位置 (ξ,ζ,η) 距路径段 p_ip_{i+1} 的位置偏移，计算如(1)式所示。采用如(2)式的 AUV 目标路径航向导引控制律，其中 c_0，c_1，c_2，ε_0，ε_{max}，ψ_{max} 均为大于零的常数，$\dot{\psi}$ 为 AUV 航

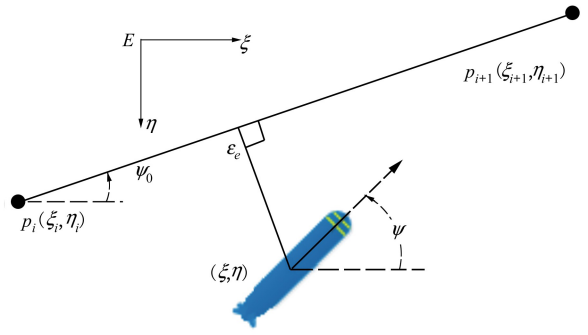


图 3 AUV 路径跟踪示意图

向角速度，当 AUV 在路径段 p_ip_{i+1} 右侧，ε_e 取正，反之 ε_e 取负。

$$\varepsilon_e = \frac{|\xi(\eta_{i+1}-\eta_i)+\eta(\xi_{i+1}-\xi_i)+(\eta_i\xi_{i+1}-\eta_{i+1}\xi_i)|}{\sqrt{(\xi_{i+1}-\xi_i)^2+(\eta_{i+1}-\eta_i)^2}} \tag{1}$$

$$\psi_{cmd} = \begin{cases} \psi_0 + \dfrac{2}{1+e^{-c_1\varepsilon_e-c_2\varepsilon_e}} - 1 + \dot{\psi} & \varepsilon_0 > |\varepsilon_e| \\ \psi_0 + c_0\varepsilon_e + \dot{\psi} & \varepsilon_{max} > |\varepsilon_e| \geqslant \varepsilon_0 \\ \psi_0 + \psi_{max} & |\varepsilon_e| \geqslant \varepsilon_{max} \end{cases} \tag{2}$$

1.2 AUV 航向滑模控制律

AUV 航向导引律(2)给出了 AUV 跟踪期望路径所需要的指令航向角度，而对指令航向角的精确稳定跟踪是实现 AUV 全局路径跟踪的关键。AUV 水平面偏航运动方程如(3)式所示：

$$\begin{aligned} &\dot{\psi} = \omega_y \\ &a_1^{\psi}\dot{\omega}_y = a_2^{\psi}\omega_y + a_3^{\psi}\delta_r + \hat{d}_{\psi} \\ &a_1^{\psi} = J_{yy} + \lambda_{55},\ a_2^{\psi} = -\left(mz_cv_z + mx_cv_x - \frac{1}{2}\rho SL^2vm_y^{\omega_y}\right) \\ &a_3^{\psi} = \frac{1}{2}\rho v^2SLm_y^{\delta_r} \end{aligned} \tag{3}$$

式中：ω_y，$\dot{\omega}_y$ 为航向角速度及加速度；δ_r 为艉部垂直舵角；$\hat{d}_{\psi}$ 为系统未建模特性及外噪声干扰；v，v_x，v_z 为 AUV 速度及分量，其余变量为总体参数[13]。

定义滑模面函数如(4)式，其中 $c_{\psi}>0$，$\psi_e=\psi_d-\psi$，ψ_d 表示航向控制指令，与(2)式中 ψ_{cmd} 数值大小相同。取航向跟踪控制律如(5)式所示，则 AUV 航向跟踪误差 ψ_e 在有限时间内渐进收敛。

$$s_{\psi} = c_{\psi}\psi_e + \dot{\psi}_e,\ c_{\psi} > 0 \tag{4}$$

$$\delta_r = \frac{1}{a_3^{\psi}}\left[-(a_2^{\psi}+a_1^{\psi}c_{\psi})\omega_y + a_1^{\psi}(k_{\psi}s_{\psi}+\tau_{\psi}\mathrm{sat}(s_{\psi}))\right] \tag{5}$$

系统稳定性证明如下：取李雅普诺夫函数为 $V=\frac{1}{2}s_{\psi}^2$，其一阶微分为

$$\begin{aligned} \dot{V} = s_{\psi}\dot{s}_{\psi} &= s_{\psi}(c_{\psi}\dot{\psi}_e + \ddot{\psi}_e) = s_{\psi}(-c_{\psi}\omega_y - \dot{\omega}_y) = \\ &s_{\psi}\left(-c_{\psi}\omega_y - \frac{a_2^{\psi}}{a_1^{\psi}}\omega_y - \frac{a_3^{\psi}}{a_1^{\psi}}\delta_r\right) = \\ &-a_1^{\psi}(k_{\psi}s_{\psi}^2 + \tau_{\psi}s_{\psi}\mathrm{sat}(s_{\psi})) \end{aligned} \tag{6}$$

式中：$\mathrm{sat}(\cdot)$ 为饱和函数。根据图 1，ψ_{cmd} 是路径跟踪大回路导引控制器的输出，而 ψ_d 是航向控制小回路的指令输入，二者数值相同，但控制含义不同。在设计航向控制器时，ψ_d 常取为常值，通过设置(5)式中的控制器参数，调整控制性能，然后实现对时变航向指令的跟踪。(6)式中，AUV 的转动惯量参数 $a_1^{\psi}>0$，当 $s_{\psi}\neq 0$ 时，若取参数 k_{ψ}，τ_{ψ} 均为大于零的正常数，则

$$\dot{V} = -a_1^{\psi}(k_{\psi}s_{\psi}^2 + \tau_{\psi}s_{\psi}\mathrm{sat}(s_{\psi})) < 0$$

恒成立,因此,根据 Barbalat 引理,对于运动方程(3)式、滑模切换面(4)式以及控制律(5)式组成的闭环系统,系统具有渐进稳定性,误差在有限时间内收敛。

1.3　AUV 俯仰跟踪控制律

AUV 俯仰运动方程如(7)式所示,ω_z,$\dot{\omega}_z$ 分别为、俯仰角速度及其加速度,δ_e 为艉部水平舵角,α 为 AUV 攻角,G 为重力,v_y 为 AUV 垂向速度,$\hat{d}_{\theta}$ 为系统未建模特性和外部干扰,其余总体参数详情参见文献[13]。

$$\begin{aligned}
&\dot{\theta} = \omega_z\\
&a_1^{\theta}\dot{\omega}_z = a_2^{\theta}\omega_z + a_3^{\theta}\delta_e + a_4^{\theta} + a_5^{\theta} + \hat{d}_{\theta}\\
&\dot{\zeta} = v_x\sin\theta + v_y\cos\theta = v\sin(\theta - \alpha)\\
&\alpha = -\arctan\left(\frac{v_y}{v_x}\right), v = \sqrt{v_x^2 + v_y^2}
\end{aligned} \tag{7}$$

式中

$$a_1^{\theta} = J_{zz} + \lambda_{66}, a_3^{\theta} = \frac{1}{2}\rho v^2 SLm_z^{\delta_e}$$

$$a_2^{\theta} = -\left(mx_c v_x + my_c v_y - \frac{1}{2}\rho SL^2 v m_z^{\omega_z}\right)$$

$$a_4^{\theta} = G(y_c\sin\theta - x_c\cos\theta), a_5^{\theta} = \frac{1}{2}\rho v^2 SLm_z^{\alpha}\alpha$$

AUV 回收路径指令深度为 ζ_0,深度误差 $\zeta_e = \zeta_0 - \zeta$,选取理想深度动态偏差特性为

$$\dot{\zeta}_e + \lambda\zeta_e = 0 \qquad \lambda > 0 \tag{8}$$

则可得到:$\dot{\zeta} = \dot{\zeta}_0 + \lambda(\zeta_0 - \zeta)$,与(8)式中运动学方程联立求解可得指令俯仰角 θ_{cmd} 如(9)式所示,其中要求 $|(\dot{\zeta}_0 + \lambda(\zeta_0 - \zeta))/v| \leqslant 1$,需要对深度偏差做饱和限幅,如(10)式所示,采用二阶环节对 θ_{cmd} 进行滤波,避免俯仰角指令出现阶跃振荡:$\omega_0^2\theta_d + 2\omega_0\dot{\theta}_d + \ddot{\theta}_d = \omega_0^2\theta_{\mathrm{cmd}}$,其中 ω_0 为角频率,$\theta_d(0) = 0$,$\dot{\theta}_d(0) = \dot{\theta}(0)$。

$$\theta_{\mathrm{cmd}} = \arcsin\frac{\dot{\zeta}_0 + \lambda(\zeta_0 - \zeta)}{v} + \alpha \tag{9}$$

$$\theta_{\mathrm{cmd}} = \arcsin\frac{\dot{\zeta}_0 + \lambda k_0\mathrm{sat}[(\zeta_0 - \zeta)/k_0]}{v} + \alpha \tag{10}$$

记俯仰角误差 $\theta_e = \theta_d - \theta$,滑模面如(11)式所示,取俯仰跟踪控制律如(12)式所示,则俯仰角误差 θ_e 在有限时间内收敛。

$$s_{\theta} = c_{\theta}\theta_e + \dot{\theta}_e, c_{\theta} > 0 \tag{11}$$

$$\delta_e = \frac{(-a_2^{\theta} - a_1^{\theta}c_{\theta})\omega_z - a_4^{\theta} - a_5^{\theta} + a_1^{\theta}(k_{\theta}s_{\theta} + \tau_{\theta}\mathrm{sat}(s_{\theta}))}{a_3^{\theta}} \tag{12}$$

系统稳定性证明:取李雅普诺夫函数为 $V = \frac{1}{2}s_{\theta}^2$,一阶微分为

$$\begin{aligned}
\dot{V} = s_{\theta}\dot{s}_{\theta} = s_{\theta}(c_{\theta}\dot{\theta}_e + \ddot{\theta}_e) = s_{\theta}(-c_{\theta}\omega_z - \dot{\omega}_z) =\\
-a_1^{\theta}(k_{\theta}s_{\theta}^2 + \tau_{\theta}s_{\theta}\mathrm{sat}(s_{\theta})) \leqslant 0
\end{aligned} \tag{13}$$

与航向角控制同理,AUV 艉部舵面无法直接实现深度控制,只有通过图 2 中的双回路,在实现俯仰角指令 θ_d 的跟踪前提下,实现深度控制,因而 θ_d 也被认为是常值。根据 Barbalat 引理,对于运动方程式(7)、滑模切换面式(11)以及控制律式(12)组成的闭环系统,系统具有渐进稳定性,误差在有限时间内收敛。

2　RBF 网络 Q 学习参数优化

由航向及俯仰运动方程(3)式和(7)式可知,当航速保持恒定时,AUV 运动状态的响应特性主要与舵角 δ_r 和 δ_e 变化有关,舵角控制律(5)式和(12)式中,除 AUV 自身状态变量 ψ,w_y,θ,w_z 等外,滑模控制参数 $c_{(\cdot)}$,$k_{(\cdot)}$,$\tau_{(\cdot)}$ 的取值直接影响舵角变化过程,参数的选取决定了系统控制特性。

在 AUV 回收对接中,受转速指令、转向、浮潜运动以及时变扰流等影响,AUV 航速难以保持稳定状态,传统方法中同一组滑模控制参数不利于保证 AUV 全局路径跟踪控制性能。本节采用学习算法,根据 AUV 航行状态反馈,实时优化滑模控制参数,提高系统性能。

2.1　Q 学习算法参数定义

应用 Q 学习算法进行 AUV 控制优化,首要工作是根据滑模控制特点定义马尔科夫决策四元组 $\boldsymbol{E} = \{S, A, P, R\}$,即状态空间 S、动作空间 A、转移概率 P 和瞬时奖赏值 R。以航向控制优化为例说明如下:考虑到算法的通用性,选取 AUV 前向速度、航向角误差及其导数作为系统状态量,即 $\boldsymbol{s} = [v_x \quad \psi_e \quad \dot{\psi}_e]$,该状态空间是连续可观测的,每个状态变量均可由 AUV 携带的传感器实时采样测量;动作空间由滑模控制参数组成,即 $\boldsymbol{a} = [c_{\psi} \quad k_{\psi} \quad \tau_{\psi}]$,该空间为离散空间,每个参数的取

值个数分别为 c_n, k_n, τ_n。不考虑外扰动，当给定系统状态和动作时，AUV 下一时刻状态是确定的，所以转移概率 P 不需要定义。评判系统控制性能常常以超调量、稳态误差和响应时间为依据，这些指标与系统误差及其变化率密切相关，因此，瞬时奖赏值函数 R 选取与航向控制误差及其变化率有关的函数，如(15)式所示，其中常值系数 $r_{\psi_e}=1, r_{\dot{\psi}_e}=2$，学习经验累计函数 Q 值更新如(16)式所示，学习率 α 和折扣系数 γ 为正常数。

$$r = \exp(-r_{\psi_e}^2 \mid \psi_e \mid - r_{\dot{\psi}_e}^2 \mid \dot{\psi}_e \mid) \tag{14}$$

$$Q_{t+1}(\boldsymbol{s},\boldsymbol{a}) = Q_t(\boldsymbol{s},\boldsymbol{a}) + \alpha(r + \gamma \max_{a' \in A} Q(\boldsymbol{s}',\boldsymbol{a}') - Q_t(\boldsymbol{s},\boldsymbol{a})) \tag{15}$$

为避免出现“维数灾”现象[14]，采用 RBF 神经网络建立系统状态 $\boldsymbol{s}$、动作 $\boldsymbol{a}$ 和 Q 值间的映射关系，其中输入层为状态量 $\boldsymbol{s}=[v_x \quad \psi_e \quad \dot{\psi}_e]$，输出层节点为 Q 值，每个 Q 值与动作空间中的待选滑模参数一一对应，输出层节点个数 $n = c_n + k_n + \tau_n$，隐层节点个数与输出层节点个数相同，网络损失函数设置为(16)式，误差代价函数为(17)式，网络连接权值修正函数采用反向传播的形式，详细过程参见文献[15]。

$$L(\boldsymbol{s},\boldsymbol{a}) = r + \gamma \max_{a \in A} Q(\boldsymbol{s}',\boldsymbol{a}) - Q(\boldsymbol{s},\boldsymbol{a}) \tag{16}$$

$$E = \frac{1}{2}\sum_{a \in A} L^2(\boldsymbol{s},\boldsymbol{a}) \tag{17}$$

2.2 姿态控制参数优化过程

AUV 当前状态 $\boldsymbol{s}$ 经过 RBF 网络正向传播得到所有控制参数对应的 $Q(\boldsymbol{s},\boldsymbol{a})$ 值，选取最大 Q 值对应的动作值 $\boldsymbol{a}$，将其代入到航向及俯仰控制器，AUV 运动状态发生改变，得到 AUV 新的状态 $\boldsymbol{s}'$ 并计算瞬时奖励值函数 r 的大小，根据(16)式和(17)式进行误差反向传播，更新 RBF 网络连接权值。由于预先并不能获取 $Q(\boldsymbol{s},\boldsymbol{a})$ 的期望值，需要通过(15)式进行 Q 值的迭代，所以将 AUV 新的状态 $\boldsymbol{s}'$ 输入至更新后的 RBF 网络，选取最大 Q 值 $\max_{a' \in A} Q(\boldsymbol{s}',\boldsymbol{a}')$，则(15)式中的主要参数均已得到，能够完成期望 Q 值的更新。随着迭代的进行，RBF 网络权值与 $Q(\boldsymbol{s},\boldsymbol{a})$ 值逐渐收敛稳定，学习过程结束。当输入新的航向角指令时，在上述 RBF 网络基础上继续进行学习更新，对姿态控制过程进行持续优化。

根据上述过程，RBF-Q 学习优化控制原理如图 4 所示，图中同时给出了俯仰角优化控制原理，定义的网络状态输入 $\boldsymbol{s}=[v_x \quad \theta_e \quad \dot{\theta}_e]$ 以及动作空间 $\boldsymbol{a}=[c_\theta \quad k_\theta \quad \tau_\theta]$。

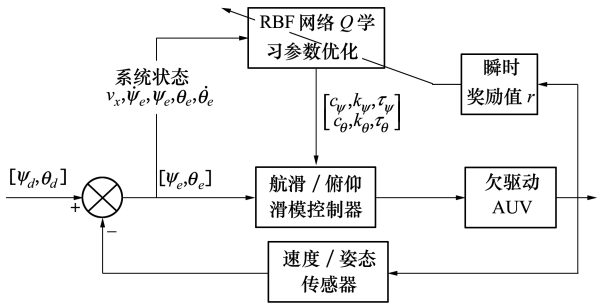


图 4 RBF-Q 学习参数优化控制原理图

3 仿真验证

建立数值仿真环境，依次进行 AUV 不同航速下的定深、定向控制仿真，以及全局路径跟踪控制仿真，将 RBF-Q 学习参数优化控制方法与传统滑模控制(SMC)方法进行对比，分析算法特点。

仿真中采用了某 AUV 空间六自由度非线性模型，舵角 δ_r 和 δ_e 限幅均为 ±15°，AUV 航速区间为 3.5 ~ 11 kn，外部扰动主要体现在 AUV 速度分量 v_x, v_y, v_z 上的噪声扰动。算法主要参数取值如下：学习率 $\alpha = 0.5$，折扣系数 $\gamma = 0.9$，航向及俯仰控制参数 $c_{(\cdot)}, k_{(\cdot)}$ 和 $\tau_{(\cdot)}$ 的取值个数为 10，综合考虑超调量、稳态误差及响应速度，AUV 航速为 8 kn 时，滑模控制律(5)式和(12)式中的最优参数取值如表 1 所示，其中 Δ 表示取值间隔。RBF-Q 学习算法需要对 AUV 不同航速、航向角度指令和俯仰角度指令进行离线学习训练，表 2 给出了 RBF-Q 学习网络训练的取值范围，为避免过拟合，AUV 航速增加了一定幅值的随机噪声。

表 1 AUV 航向及俯仰控制参数取值

传统滑模控制参数			Q 学习优化参数范围		
c_ψ	k_ψ	τ_ψ	$c_\psi, \Delta=0.3$	$k_\psi, \Delta=0.5$	$\tau_\psi, \Delta=0.02$
3	5	0.2	[1.8, 4.5]	[3, 7.5]	[0.12, 0.3]
c_θ	k_θ	τ_θ	$c_\theta, \Delta=0.3$	$k_\theta, \Delta=0.5$	$\tau_\theta, \Delta=0.02$
4	3.5	0.1	[2.8, 5.5]	[1.5, 6]	[0.02, 0.2]

表 2　RBF-Q 学习离线训练样本取值

指令参数	取值范围	取值间隔
航速 v_x/kn	[3.5,11]	1.5
航向角 ψ_d/(°)	[-90,90]	2.5
俯仰角 θ_d/(°)	[-40,40]	1.5

3.1　AUV 多航速定深、定向控制

在回收过程中,AUV 存在不同航速下的定深、定向航行状态,选取工况为:航向角 ψ 由 0°定向 30°,深度 ζ 由 175 m 定深 160 m,无外扰动条件下,选取 AUV 航速 v_x 为 8 kn 和 6 kn 进行对比,结果如图 5 和图 6 所示;在 6 kn 航速下考虑时变扰动,航向及深度指令不变,仿真结果如图 7 所示。

从图 5 可以看到,2 种算法下 AUV 定向、定深响应曲线近似重合,航向角 ψ、俯仰角 θ 及深度 ζ 响应均具有较小的超调量和稳态误差,表明在 8 kn 航速下,RBF-Q 学习算法对滑模控制器的改善效果很小。但当 AUV 航速变化为 6 kn 时,AUV 艉部舵效下降,传统滑模控制器的深度响应曲线出现超调量约 16.7%,如图 6 所示,RBF-Q 学习算法将深度响应的超调量降低至 6.6%,在航向角响应中,RBF-Q 学习算法比传统滑模算法响应时间缩短了 10 s,控制效果改善显著。随着外部扰动的增加,如图 7 所示,RBF-Q 学习控制器在稳态精度、响应速度和超调量方面,均比传统滑模算法具有更好的表现,二者在航向稳态误差相差了约 5.5°,深度误差相差了约 2 m,但这些控制效果的改善存在一定的代价,即 RBF-Q 学习算法根据 AUV 航速、跟踪误差及其变化率持续优化滑模控制器参数。在 8 kn 无扰动条件下,控制参数基本无变化,初始参数即在最优解附近。当航速减小为 6 kn,算法对控制器进行了短暂优化,之后随着误差收敛优化过程结束,达到稳定,而扰动的增加,使 RBF-Q 学习网络持续对参数进行优化,以降低扰动影响。

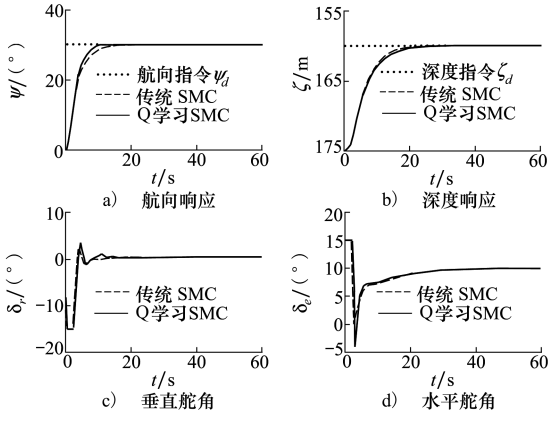


图 5　无扰动 8 kn 航速下 AUV 定向定深控制响应

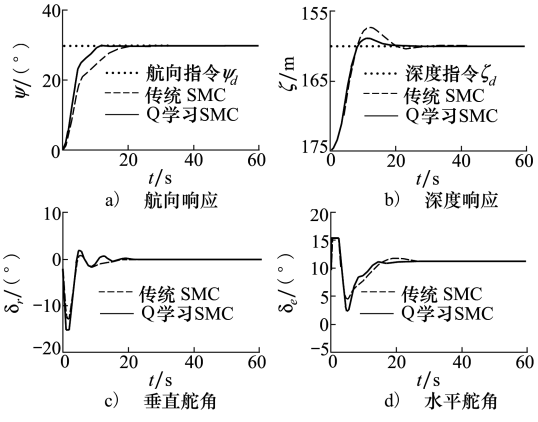


图 6　无扰动 6 kn 航速下 AUV 定向定深控制响应

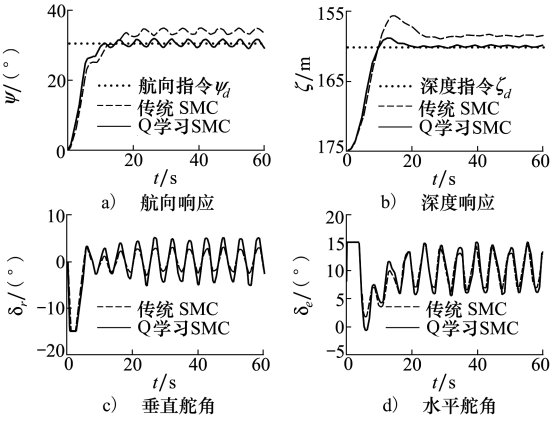


图 7　时变扰流,6 kn 航速下 AUV 定向定深控制响应

3.2　AUV 全局路径跟踪控制仿真

为分析 2 种算法在 AUV 路径跟踪控制中的性能,本节对先前研究成果[3]中的路径进行跟踪仿真,该路径为水平面内的避障回收路径,离散路径点个数为 100,航行深度设定为 $\zeta_0 = 120$ m,根据本文(1)式和(2)式进行跟踪导引,依次采用传统滑模控制方法和 RBF-Q 学习参数优化控制方法对路径进行跟踪仿真,结果如图 8 和图 9 所示。

图 8 给出了 AUV 在 E-$\xi\eta$ 平面内的航行轨迹,AUV 的起点与规划路径起点不重合,回收路径上存在多个障碍物,外部扰动为时变海流。在路径的起始阶段,传统滑模控制算法下的 AUV 轨迹更加接近期望轨迹,随着学习网络的持续优化,在路径中后

段,RBF-Q 学习算法控制下的 AUV 轨迹与期望轨迹基本重合,而传统滑模控制算法下的 AUV 轨迹始终与期望轨迹存在偏差。图 9 给出了 AUV 跟踪过程

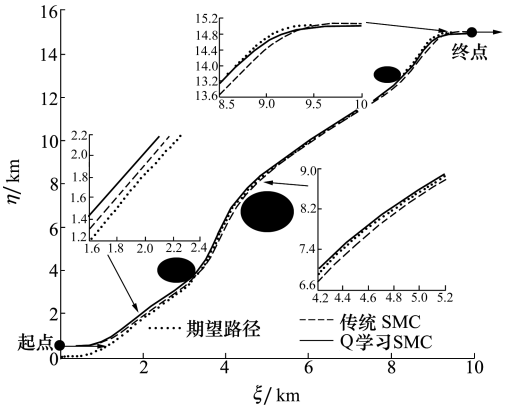


图 8 回收海域有障碍物时,AUV 路径跟踪控制曲线

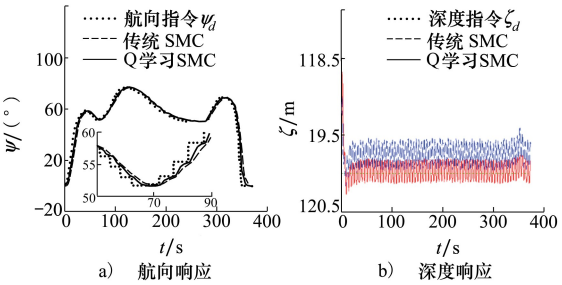


图 9 AUV 路径跟踪航向及深度控制响应

中航向角和深度响应曲线。为避开水中障碍并到达路径终点,AUV 进行了多次转向运动,RBF-Q 学习算法对时变航向指令具有准、更快的响应效果,而且在深度保持控制中,RBF-Q 学习算法的振荡幅值更小,稳态控制精度更高。

4 结 论

AUV 水下回收对接过程中,航行速度受推进器转速、姿态变化以及时变扰流等因素影响难以保持稳定,使艉部舵面的操纵性能发生改变,导致同一组控制参数无法满足复杂环境下的姿态控制和路径跟踪控制性能要求。针对上述问题,本文采用 RBF 网络 Q 学习算法对传统滑模控制参数进行优化,根据 AUV 当前航速、指令误差及其变化律选择相应的控制参数,经过大量离线训练获取学习网络权重参数,将其应用到在线控制当中,仿真结果表明 RBF-Q 学习算法能够有效改善 AUV 姿态控制和路径跟踪控制的性能,在相同条件下,RBF-Q 学算法比传统滑模算法具有更小的超调量和稳态误差,响应速度更快。在后续研究中,针对舵角控制输入中出现的非线性抖动问题作进一步优化,逐步进行 AUV 路径跟踪控制的实航试验研究。

参考文献:

[1] RIDAO P, CARRERAS M, RIBAS D, et al. Intervention AUVs: the next challenge[J]. Annual Reviews in Control, 2015, 40: 227-241

[2] SHI Y, SHEN C, FANG H, et al. Advanced control in marine mechatronic systems: a survey[J]. IEEE/ASME Trans on Mechatronics, 2017,22(3): 1121-1131

[3] LI Z, LIU W, GAO L, et al. Path planning method for AUV docking based on adaptive quantum-behaved particle swarm optimization[J]. IEEE Access, 2019,7:78665-78674

[4] MIN J K, WOON-KYUNG Baek, KYOUNGNAM Ha, et al. Way-point tracking for a hovering AUV by PID controller[C]// 15th International Conference on Control, Automation and Systems, BEXCO, Busan, Korea, 2015

[5] 张磊. 基于遗传算法优化的水下机器人路径跟踪模糊控制技术研究[D]. 杭州:浙江大学, 2017
ZHANG Lei. Research on fuzzy control of underwater vehicle path following based on genetic algorithm optimization[D]. Hangzhou: Zhejiang University, 2017 (in Chinese)

[6] 王宏健,陈子印,贾鹤鸣,等. 基于滤波反步法的欠驱动 AUV 三维路径跟踪控制[J]. 自动化学报, 2015,41(3): 631-645
WANG Hongjian, CHEN Ziyin, JIA Heming, et al. Three-dimensional path-following control of underactuated autonomous underwater vehicle with command filtered backstepping[J]. Acta Automatica Sinica, 2015,41(3):631-645 (in Chinese)

[7] 王金强,王聪,魏英杰,等. 未知海流干扰下自主水下航行器位置跟踪控制策略研究[J]. 兵工学报, 2019,40(3): 583-591
WANG Jingqiang, WANG Cong, WEI Yingjie, et al. Position tracking control of autonomous underwater vehicles in the disturbance of unknown ocean currents[J]. Acta Armamentarii, 2019,40(3): 583-591 (in Chinese)

[8] 王金强,王聪,魏英杰,等. 欠驱动 AUV 自适应神经网络反步滑模跟踪控制[J]. 华中科技大学学报, 2019, 47(12):

12-17
WANG Jingqiang, WANG Cong, WEI Yingjie, et al. Path following of an underactuated AUV based on adaptive neural network backstepping sliding mode control[J]. Journal of Huazhong University of Science and Technology, 2019,47(12):12-17 (in Chinese)
[9] SHEN C, SHI Y, BUCKHAM B. Integrated path planning and tracking control of an AUV: a unified receding horizon optimization approach[J].IEEE/ASME Trans on Mechatronics, 2017,22(3): 1163-1173
[10] SUN Y, ZHANG C, ZHANG G, et al. Three-dimensional path tracking control of autonomous underwater vehicle based on deep reinforcement learning[J]. Journal of Marine Science and Engineering, 2019, 7(12): 443
[11] SHI W, SONG S, WU C, et al. Multi pseudo q-learning-based deterministic policy gradient for tracking control of autonomous underwater vehicles[J]. IEEE Trans on Neural Networks and Learning Systems, 2019, 30(12): 3534-3546
[12] 姚绪梁,王晓伟. 基于 MPC 导引律的 AUV 路径跟踪和避障控制[J]. 北京航空航天大学学报, 2020,46(6): 1053-1062
YAO Xuliang, WANG Xiaowei. Path following and obstacle avoidance control of AUV based on MPC guidance law[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020,46(6): 1053-1062 (in Chinese)
[13] 严卫生. 鱼雷航行力学[M]. 西安:西北工业大学出版社, 2005
YAN Weisheng. Torpedo navigation mechanics[M]. Xi'an: Northwestern Polytechnical Press, 2005 (in Chinese)
[14] SUN Y, CHENG J, ZHANG G, et al. Mapless motion planning system for an autonomous underwater vehicle using policy gradient-based deep reinforcement learning[J]. Journal of Intelligent & Robotic Systems, 2019,96(3/4):591-601
[15] HAGAN M T, DEMUTH H B, BEALE M H. Neural network design[M]. Beijing: China Machine Press, 2002

Path following method for AUV based on Q-Learning and RBF neural network

LI Zeyu, LIU Weidong, LI Le, ZHANG Wenbo, GUO Liwei
(School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: In the underwater docking process, the oscillation on AUV velocity brings extra challenge on AUV path following. A Q-learning based Sliding Mode Control (SMC) method to increase the path following performances is proposed. Firstly, AUV guidance law is designed to reduce the path following error. Heading and depth sliding mode controllers are designed to track the guidance law. Then, according to AUV velocity, tracking error and the first derivative, the control parameters of SMC are optimized via Q-learning network. RBF neural network is built to accelerate the offline learning rate. Finally, numerical simulations are made to investigate the characteristics of the present method. Comparisons are made between the trained Q-learning based SMC and the traditional SMC. The results show the effectiveness of the present method.

Keywords: autonomous underwater vehicle; path following; reinforcement learning; neural network

引用格式:李泽宇, 刘卫东, 李乐, 等. 基于 RBF 网络 Q 学习的 AUV 路径跟踪控制方法[J]. 西北工业大学学报,2021,39(3): 477-483
LI Zeyu, LIU Weidong, LI Le, et al. Path following method for AUV based on Q-Learning and RBF neural network[J]. Journal of Northwestern Polytechnical University, 2021, 39(3): 477-483 (in Chinese)

© 2021 Journal of Northwestern Polytechnical University.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.