



(12) 发明专利

(10) 授权公告号 CN 115657689 B

(45) 授权公告日 2023.04.07

(21) 申请号 202211670756.8

G06N 3/08 (2023.01)

(22) 申请日 2022.12.26

G06N 3/04 (2023.01)

(65) 同一申请的已公布的文献号

审查员 许玉姗

申请公布号 CN 115657689 A

(43) 申请公布日 2023.01.31

(73) 专利权人 安徽大学

地址 230601 安徽省合肥市经开区九龙路
111号

(72) 发明人 曹翔 孙长银 吴巧云

(74) 专利代理机构 南京钟山专利代理有限公司

32252

专利代理师 张力

(51) Int. Cl.

G05D 1/02 (2020.01)

G06Q 10/04 (2023.01)

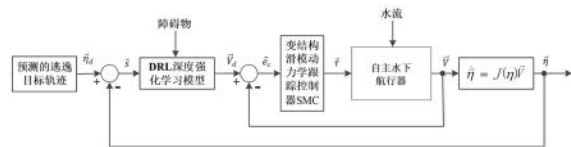
权利要求书3页 说明书9页 附图3页

(54) 发明名称

一种基于轨迹预测的自主水下航行器目标跟踪控制方法

(57) 摘要

本发明公开了一种基于轨迹预测的自主水下航行器目标跟踪控制方法,基于TPENN神经网络进行目标轨迹预测,得到预测的逃逸目标轨迹;根据预测的逃逸目标轨迹,基于深度强化进行动态目标跟踪控制,实现安全避障、水流补偿和轨迹跟踪。本发明将轨迹预测问题转化为时间序列的预测问题,将TPENN神经网络模型引入目标轨迹预测之中,利用TPENN神经网络挖掘轨迹数据的内部相关性,解决轨迹预测难题;在动态逃逸目标轨迹预测的基础上,针对自主水下航行器非线性模型不确定特性和水流影响,借助强化学习方法的非线性、强学习,不依赖模型的特性,同时引入深度学习算法克服强化学习的维度灾难问题,从而实现航行器自主动态目标跟踪控制。



1. 一种基于轨迹预测的自主水下航行器目标跟踪控制方法,其特征在于,包括步骤1、基于TPENN神经网络进行目标轨迹预测,得到预测的逃逸目标轨迹;所述TPENN神经网络包括输入层、隐含层、输出层和承接层;

其中,承接层的激活函数为恒等函数,用于存储记忆隐含层单元上一时刻的输出,当作一步延时算子;

假设在时刻 t ,TPENN神经网络的输入是 $U(t)$,隐含层的输出为 $X(t)$,整个网络的输出为 $Y(t)$, $\omega^1, \omega^2, \omega^3$ 分别表示输入层与隐含层、承接层与隐含层、隐含层与输出层之间的连接权值, b 和 c 分别代表隐含层和输出层的偏置,则神经网络的计算过程表示为

$$\begin{cases} X(t) = f_2(\omega^1 U(t) + b), & t = 0 \\ X(t) = f_2(\omega^1 U(t) + \omega^2 X(t-1) + b), & t > 0 \end{cases} \quad (6)$$

$$Y(t) = g(\omega^3 X(t) + c) \quad (7)$$

神经网络采用时间反向传播算法进行学习和训练,其目标函数为

$$E(t) = \frac{1}{2} (Y_d(t) - Y(t))^T (Y_d(t) - Y(t)) \quad (8)$$

$Y_d(t)$ 表示预期的输出;

由于神经网络的输入是具有时间序列的历史数据,因此引入时间回归因子 $f_{TP}(t)$ 得到神经网络改进后的目标函数 $E_{TP}(t)$:

$$f_{TP}(t) = \frac{1}{1 + e^{(\alpha - 2\alpha t \cdot \lambda^{-1})}} \quad (9)$$

$$E_{TP}(t) = \sum_{k=1}^t f_{TP}(\lambda) E(\lambda) \quad (10)$$

其中, α 表示是打折率, λ 表示样本数量, t 表示迭代次数;

步骤2、根据预测的逃逸目标轨迹,基于深度强化进行动态目标跟踪控制,实现安全避障、水流补偿和轨迹跟踪;

步骤2所述控制过程具体为:

从预测的逃逸目标轨迹中获取期望的自主水下航行器位姿矢量 $\eta_d = [x_d \ y_d \ z_d \ \phi_d \ \theta_d \ \psi_d]^T$,与自主水下航行器当前状态矢量 $\eta = [x \ y \ z \ \phi \ \theta \ \psi]^T$ 比较,产生航向以及位置轨迹误差,并与自主水下航行器的状态信息整合生成状态误差向量,输入DRL深度强化学习模型中,根据强化学习算法输出最优的运动速度控制信号矢量

$\vec{V}_d = [u_d \ v_d \ w_d \ p_d \ q_d \ r_d]^T$,并与自主水下航行器输出速度信号矢量 \vec{V} 比较,产生自主水下航行器跟踪速度误差 \vec{e}_v ,输入变结构滑模动力学跟踪控制器SMC,由此产生自主水下航行器的动力学推进控制信号 $\vec{\tau}$,驱动自主水下航行器的速度矢量矩阵 \vec{V} 跟踪自主水下航行器跟踪的运动学速度控制信号 \vec{V}_d ,通过坐标转换 $\dot{\eta} = J(\eta)\vec{V}$ 得到AUV实际位姿矢量 $\eta = [x \ y \ z \ \phi \ \theta \ \psi]^T$,从而实现自主水下航行器对动态目标的高效精准跟踪;

其中, x_d, y_d, z_d 分别为自主水下航行器在惯性坐标系中期望的位置; ϕ_d, θ_d, ψ_d 分别为自

主水下航行器对惯性坐标系中期望的横倾角、纵倾角和艏向角；

x y z 分别为自主水下航行器在惯性坐标系中当前的位置； ϕ θ ψ 分别为自主水下航行器对惯性坐标系中当前的横倾角、纵倾角和艏向角；

u_d v_d w_d 分别为自主水下航行器线速度矢量在载体坐标系中的3个分量； p_d q_d r_d 分别为自主水下航行器角速度矢量在载体坐标系中的3个分量；

所述步骤2还包括，将时变立体水流模型加入到载体坐标系的速度矢量中进行控制补偿；

所述时变立体水流模型中，自主水下航行器动力学方程的动力项为：水动力+推进机构推力+水流干扰力。

2. 根据权利要求1所述的一种基于轨迹预测的自主水下航行器目标跟踪控制方法，其特征在于，所述步骤1将逃逸目标当前时刻 t 时间序列轨迹坐标输入已学习的TPENN神经网络，得到逃逸目标未来时刻的预测轨迹坐标。

3. 根据权利要求1所述的一种基于轨迹预测的自主水下航行器目标跟踪控制方法，其特征在于，所述TPENN神经网络的学习方式为：

输入训练样本数据，根据BPTT算法，得到 $E_{TP}(t)$ 的偏导数，其权重计算如式(11)所示：

$$\begin{cases} \Delta\omega_{iq}^1(t) = \sum_{\lambda=1}^t \eta_1 f_{TP}(\lambda) \delta_j^h u_q(\lambda-1) \\ \Delta\omega_{ii}^2(t) = \sum_{\lambda=1}^t \eta_2 f_{TP}(\lambda) \sum_{i=1}^m (\delta_i^0 \omega_{ij}^3) \frac{\partial x_j(\lambda)}{\partial \omega_{ii}^2} \\ \Delta\omega_{ij}^3(t) = \sum_{\lambda=1}^t \eta_3 f_{TP}(\lambda) \delta_i^0 x_j(\lambda) \end{cases} \quad (11)$$

η_1 、 η_2 、 η_3 是 ω^1 、 ω^2 、 ω^3 学习步；

$i=1, 2, \dots, m$ ； $j=1, 2, \dots, n$ ； $q=1, 2, \dots, r$ ； $l=1, 2, \dots, n$ ；

δ_j^h 、 δ_i^0 分别表示第 j 层输出层的误差、第 i 层输入层的误差；

通过神经网络的训练，得到TPENN神经网络的各项参数。

4. 根据权利要求1所述的一种基于轨迹预测的自主水下航行器目标跟踪控制方法，其特征在于，所述步骤1通过水下声呐与光学传感器得到逃逸目标所有历史时刻位姿数据，作为TPENN神经网络学习训练样本。

5. 根据权利要求1所述的一种基于轨迹预测的自主水下航行器目标跟踪控制方法，其特征在于，所述DRL深度强化学习模型包括DRL避障模块、导航模块和动作调度器；

避障模块，用于DRL通过双Q-network学习模式实现避开障碍物；

导航模块，用于通过双Q-network学习的方式完成自主水下航行器导航策略的在线训练，通过在线学习解决目标跟踪的最终策略；

动作调度器，用于对导航模块在线部分进行有效的探索和加速训练，并生成自主水下航行器运动学控制指令。

6. 根据权利要求5所述的一种基于轨迹预测的自主水下航行器目标跟踪控制方法，其特征在于，所述避障模块设置碰撞负奖励，将避障目标表达为强化学习的一般目标，以寻求一种使总奖励最大化的策略，经过训练，由双Q-network得到学习Q函数，最终收敛于最优Q函数附近；

所述导航模块使用双Q-network中的一个网络分别将获得的局部障碍物信息、预测轨迹信息转化为奖励,并将奖励结果立即反馈以改进导航策略;同时,使用另一个完全连接的Q-network在线学习的行为价值函数;

所述动作调度器利用从避障模块、导航模块学习到的策略,在重置缓冲区中积累训练经验,随着时间的推移,调度器逐渐从静态学习策略转移到在线学习策略,在线探索和优化一直持续到产生稳定的自主水下航行器跟踪的运动学速度控制信号。

一种基于轨迹预测的自主水下航行器目标跟踪控制方法

技术领域

[0001] 本发明属于水下航行器控制技术领域,具体涉及一种基于轨迹预测的自主水下航行器目标跟踪控制方法。

背景技术

[0002] 深海水下追踪与博弈是未来水下无人战争的重要形式。自主水下航行器自带能源动力、航行范围广而深的特性,使其成为水下无人战争的重要载体,在海洋工程与海防军事领域有着十分广泛的应用前景,发挥着不可替代的作用。特别是基于水下动态目标跟踪控制技术,在国家海防安全领域得到高度关注,是水下逃逸目标追踪控制、敌对目标拦截围捕及各种水下博弈对抗研究的关键与核心技术。

[0003] 目标跟踪是指自主水下航行器在控制律的激励下持续跟踪轨迹连续变化的目标,强调在指定时间到达指定位置,是一种严格考虑时间约束的动态跟踪。由于自主水下航行器自身非线性和模型不确定性的存在,早期普遍使用的PID控制、反步控制等依赖自主水下航行器模型的控制策略,很难实现精确的水下目标跟踪。近年来,针对自主水下航行器非线性和模型不确定性特性,新型变结构滑模控制、自适应控制、智能控制等得到了广泛关注。

[0004] 然而,在水下目标跟踪控制研究中,通常将自主水下航行器的性能设置为与目标性能相同,甚至更优,采用跟随的方式能够实现对目标的跟踪。当自主水下航行器的性能比逃逸目标更差时,再采用跟随的方式很难实现目标跟踪。自主水下航行器系统本身的非线性、强耦合、模型不确定特性,水下海流及外界扰动因素影响,使得准确有效的自主水下航行器目标跟踪控制律的设计非常困难。现有控制策略中,要么将自主水下航行器简化为线性系统,要么训练样本难获取,无法满足在线跟踪目标的需求。因此,自主水下航行器在复杂海洋环境下的动态目标跟踪控制仍有很大的提升空间。

发明内容

[0005] 本发明所要解决的技术问题是针对上述现有技术的不足,提供一种基于轨迹预测的自主水下航行器目标跟踪控制方法。

[0006] 为实现上述技术目的,本发明采取的技术方案为:

[0007] 一种基于轨迹预测的自主水下航行器目标跟踪控制方法,包括

[0008] 步骤1、基于TPENN神经网络进行目标轨迹预测,得到预测的逃逸目标轨迹;

[0009] 步骤2、根据预测的逃逸目标轨迹,基于深度强化进行动态目标跟踪控制,实现安全避障、水流补偿和轨迹跟踪。

[0010] 为优化上述技术方案,采取的具体措施还包括:

[0011] 上述的步骤1将逃逸目标当前时刻 t 时间序列轨迹坐标输入已学习的TPENN神经网络,得到逃逸目标未来时刻的预测轨迹坐标。

[0012] 上述的TPENN神经网络包括输入层、隐含层、输出层和承接层;

[0013] 其中,承接层的激活函数为恒等函数,用于存储记忆隐含层单元上一时刻的输出,

当作一步延时算子；

[0014] 假设在时刻 t , TPENN神经网络的输入是 $U(t)$, 隐含层的输出为 $X(t)$, 整个网络的输出为 $Y(t)$, $\omega^1, \omega^2, \omega^3$ 分别表示输入层与隐含层、承接层与隐含层、隐含层与输出层之间的连接权值, b 和 c 分别代表隐含层和输出层的偏置, 则神经网络的计算过程表示为

$$[0015] \begin{cases} X(t) = f_2(\omega^1 U(t) + b), & t = 0 \\ X(t) = f_2(\omega^1 U(t) + \omega^2 X(t-1) + b), & t > 0 \end{cases} \quad (6)$$

$$[0016] Y(t) = g(\omega^3 X(t) + c) \quad (7)$$

[0017] 神经网络采用时间反向传播算法进行学习和训练, 其目标函数为

$$[0018] E(t) = \frac{1}{2} (Y_d(t) - Y(t))^T (Y_d(t) - Y(t)) \quad (8)$$

[0019] $Y_d(t)$ 表示预期的输出；

[0020] 由于神经网络的输入是具有时间序列的历史数据, 因此引入时间回归因子 $f_{TP}(t)$ 得到神经网络改进后的目标函数 $E_{TP}(t)$ ：

$$[0021] f_{TP}(t) = \frac{1}{1 + e^{(\alpha - 2\alpha t - \lambda^{-1})}} \quad (9)$$

$$[0022] E_{TP}(t) = \sum_{k=1}^t f_{TP}(\lambda) E(\lambda) \quad (10)$$

[0023] 其中, α 表示是打折率, λ 表示样本数量, t 表示迭代次数。

[0024] 上述的TPENN神经网络的学习方式为：

[0025] 输入训练样本数据, 根据BPTT算法, 得到 $E_{TP}(t)$ 的偏导数, 其权重计算如式(11)所示：

$$[0026] \begin{cases} \Delta \omega_q^1(t) = \sum_{\lambda=1}^t \eta_1 f_{TP}(\lambda) \delta_j^h u_q(\lambda-1) \\ \Delta \omega_l^2(t) = \sum_{\lambda=1}^t \eta_2 f_{TP}(\lambda) \sum_{i=1}^m (\delta_i^0 \omega_y^3) \frac{\partial x_j(\lambda)}{\partial \omega_l^2} \\ \Delta \omega_y^3(t) = \sum_{\lambda=1}^t \eta_3 f_{TP}(\lambda) \delta_i^0 x_j(\lambda) \end{cases} \quad (11)$$

[0027] η_1, η_2, η_3 是 $\omega^1, \omega^2, \omega^3$ 学习步长；

[0028] $i=1, 2, \dots, m; j=1, 2, \dots, n; q=1, 2, \dots, r; l=1, 2, \dots, n$ ；

[0029] δ_j^h, δ_i^0 分别表示第 j 层输出层的误差、第 i 层输入层的误差；

[0030] 通过神经网络的训练, 得到TPENN神经网络的各项参数。

[0031] 上述的步骤1通过水下声呐与光学传感器得到逃逸目标所有历史时刻位姿数据, 作为TPENN神经网络学习训练样本。

[0032] 上述的步骤2所述控制过程具体为：

[0033] 从预测的逃逸目标轨迹中获取期望的自主水下航行器位姿矢量 $\eta_d = [x_d \ y_d \ z_d \ \phi_d \ \theta_d \ \psi_d]^T$, 与自主水下航行器当前状态矢量 $\eta = [x \ y \ z \ \phi \ \theta \ \psi]^T$ 比较, 产生航向以及位置轨迹误差, 并与自主水下航行器的状态信息整合生成状态误差向量, 输入DRL深度强化学习模

型中,根据强化学习算法输出最优的运动速度控制信号矢量

$\bar{V}_d = [u_d \ v_d \ w_d \ p_d \ q_d \ r_d]^T$,并与自主水下航行器输出速度信号矢量 \bar{V} 比较,产生自主水下航行器跟踪速度误差 \bar{e}_c ,输入变结构滑模动力学跟踪控制器SMC,由此产生自主水下航行器的动力学推进控制信号 $\bar{\tau}$,驱动自主水下航行器的速度矢量矩阵 \bar{V} 跟踪 \bar{V}_d ,通过坐标转换 $\dot{\eta} = J(\eta)\bar{V}$ 得到AUV实际位姿矢量 $\eta = [x \ y \ z \ \phi \ \theta \ \psi]^T$,从而实现自主水下航行器对动态目标的高效精准跟踪;

[0034] 其中, $x_d \ y_d \ z_d$ 分别为自主水下航行器在惯性坐标系中期望的位置; $\phi_d \ \theta_d \ \psi_d$ 分别为自主水下航行器对惯性坐标系中期望的横倾角、纵倾角和艏向角;

[0035] $x \ y \ z$ 分别为自主水下航行器在惯性坐标系中当前的位置; $\phi \ \theta \ \psi$ 分别为自主水下航行器对惯性坐标系中当前的横倾角、纵倾角和艏向角;

[0036] $u_d \ v_d \ w_d$ 分别为自主水下航行器线速度矢量在载体坐标系中的3个分量; $p_d \ q_d \ r_d$ 分别为自主水下航行器角速度矢量在载体坐标系中的3个分量。

[0037] 上述的DRL深度强化学习模型包括DRL避障模块、导航模块和动作调度器;

[0038] 避障模块,用于DRL通过双Q-network学习模式实现避开障碍物;

[0039] 导航模块,用于通过双Q-network学习的方式完成自主水下航行器导航策略的在线训练,通过在线学习解决目标跟踪的最终策略;

[0040] 动作调度器,用于对导航模块在线部分进行有效的探索和加速训练,并生成自主水下航行器运动学控制指令。

[0041] 上述的避障模块设置碰撞负奖励,将避障目标表达为强化学习的一般目标,即寻求一种使总奖励最大化的策略,经过训练,由双Q-network得到学习Q函数,最终收敛于最优Q函数附近;

[0042] 所述导航模块使用双Q-network中的一个网络分别将获得的局部障碍物信息、预测轨迹信息转化为奖励,并将奖励结果立即反馈以改进导航策略;同时,使用另一个完全连接的Q-network在线学习的行为价值函数。

[0043] 所述动作调度器利用从避障模块、导航模块学习到的策略,在重置缓冲区中积累训练经验,随着时间的推移,调度器逐渐从静态学习策略转移到在线学习策略,在线探索和优化一直持续到产生稳定的自主水下航行器跟踪的运动学速度控制信号。

[0044] 上述的步骤2进行动态目标跟踪控制时,将时变立体水流模型加入到载体坐标系的速度矢量中进行控制补偿;

[0045] 所述时变立体水流模型中,自主水下航行器动力学方程的动力项为:水动力+推进机构推力+水流干扰力。

[0046] 本发明具有以下有益效果:

[0047] 在本发明中将轨迹预测问题转化为时间序列的预测问题,将TPENN(Time Profit Elman Neural Network)神经网络模型引入目标轨迹预测之中,利用TPENN神经网络挖掘轨迹数据的内部相关性,解决轨迹预测难题;

[0048] 在动态逃逸目标轨迹预测的基础上,针对自主水下航行器非线性模型不确定特性和水流影响,借助强化学习方法的非线性、强学习,不依赖模型的特性,同时引入深度学习算法克服强化学习的维度灾难问题,从而实现航行器自主动态目标跟踪控制。

附图说明

- [0049] 图1为本发明的逃逸目标位置计算图；
 [0050] 图2为本发明的TPENN学习训练与运动轨迹预测生成图；
 [0051] 图3为本发明的深度强化学习目标跟踪控制图；
 [0052] 图4为本发明的DRL深度强化学习控制模型图；
 [0053] 图5为本发明的TPENN轨迹预测流程图。

具体实施方式

[0054] 为了使本发明的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，对本发明进行进一步详细说明。应当理解，此处所描述的具体实施例仅用以解释本发明，并不用于限定本发明。

[0055] 本发明中的步骤虽然用标号进行了排列，但并不用于限定步骤的先后次序，除非明确说明了步骤的次序或者某步骤的执行需要其他步骤作为基础，否则步骤的相对次序是可以调整的。可以理解，本文中所使用的术语“和/或”涉及且涵盖相关联的所列项目中的一者或一者以上的任何和所有可能的组合。

[0056] 如图1-5所示，一种基于轨迹预测的自主水下航行器目标跟踪控制方法，包括：

[0057] 步骤1、基于TPENN神经网络进行目标轨迹预测，得到预测的逃逸目标轨迹；

[0058] 由于自主水下航行器和逃逸目标的航行速度不同，特别是逃逸目标航行速度快时，如果只是对动态目标被动跟随，很难跟踪上目标，这时需要准确预测逃逸目标未来的逃逸路径，实现主动跟踪。

[0059] 在水下动态目标跟踪控制中，逃逸目标的运动状态具有一定的连续性，其运动轨迹不会发生跳变，这就说明其在某一时刻的位置与过去时刻的位置具有一定的相关性。

[0060] 在目标之前的轨迹上按照一定的时间间隔选取若干个位置点，根据轨迹连续性的性质，利用前若干时刻的位置，对下一时刻的位置进行预测。

[0061] 在本发明中将轨迹预测问题转化为时间序列的预测问题，将TPENN (Time Profit Elman Neural Network) 神经网络模型引入目标轨迹预测之中，利用TPENN神经网络挖掘轨迹数据的内部相关性，解决轨迹预测难题。具体方法如下：

[0062] ①逃逸目标坐标获取

[0063] 如图1所示，通过水下声呐与光学传感器，自主水下航行器与目标物及各类障碍物每一时刻的相对距离是确定的，根据坐标变换可以得到逃逸目标在每个时刻的位置坐标，从而得到逃逸目标所有历史时刻位姿数据，进而得到轨迹预测神经网络学习训练样本。

[0064] 设 (x_e, y_e) 为逃逸目标在惯性栅格坐标系中的坐标， (x_r, y_r) 为自主水下航行器在惯性栅格坐标系中的坐标，相互转换方式为式 (1)：

$$[0065] \quad \begin{cases} x_e = x_r + r \cdot \cos(\theta_l + \theta_r) \\ y_e = y_r + r \cdot \sin(\theta_l + \theta_r) \end{cases}$$

[0066] (1)

[0067] 其中， r 为动态目标与自主水下航行器之间的距离；

[0068] θ_l 是逃逸目标在自主水下航行器载体坐标系中的矢量角， θ_r 为自主水下航行器水平方向角。

[0069] ②轨迹预测生成

[0070] ENN(Elman Neural Network)神经网络模型是一种具有多层动态递归结构的循环神经网络,对历史数据具有高度敏感性和动态记忆能力,特别适合处理时间序列问题,为了提高模型的学习收敛速度,本发明将时间收益因子引入到ENN神经网络的目标函数中,提出一种改进的ENN神经模型TPENN,实现逃逸目标运动轨迹在线预测。动态逃逸目标运动轨迹预测生成过程见图2。它由在线学习训练与运动轨迹在线预测生成两大部分组成。具体描述如下:

[0071] 本发明利用截取的时间连续的逃逸目标航行数据,每个数据包包含该时刻逃逸目标的三维坐标及其航向角数据。以对x坐标进行预测为例,利用历史数据包中t时刻前n个连续时刻的轨迹数据,预测t时刻x坐标。样本数据构建完成之后,进而利用样本数据训练TPENN神经网络。

[0072] 然后应用当前时刻的逃逸目标轨迹数据,由学习收敛后的TPENN神经网络直接输出未来时刻系列轨迹坐标与姿态角数值,预测生成动态目标的未来逃逸轨迹。

[0073] 自主水下航行器目标跟踪是先预测目标的运动轨迹,然后自主水下航行器根据预测的轨迹开始的跟踪上目标。

[0074] 通过目标估计预测可以提高跟踪效率。目标预测的流程,如图5所示,本发明实施例部分提供一种基于TPENN的目标轨迹预测方法,包括:

[0075] 1:构建样本数据

[0076] 在惯性系中,目标的位置用 x_e, y_e 表示。本发明认为的变化 x_e, y_e 坐标与航向角直接相关,而俯仰角的变化又影响坐标的变化。因此,在预测目标的轨迹时,坐标 x_e, y_e 是独立预测的。坐标 x_e, y_e 利用自主水下航行器航向角数据进行预测。以x坐标的预测为例,利用前n个时刻的轨迹数据来预测 t_{n+1} 的x坐标,函数映射关系为

$$[0077] \quad x_d = f_2(L_1, L_2, \dots, L_n) \quad (3)$$

[0078] $L_i = (x_i, a_i)^T, i \in [1, n+1]$ 是目标在 t_i 时刻的x坐标和航向角, x_d 是自主水下航行器在时间 t_{n+1} 的横坐标值。

[0079] 选取 L_1-L_n 为第一组输入数据,输出数据为 x_{d+1} 。以 L_2-L_{n+1} 为第二组输入数据,输出数据为 x_{d+2} 。以此类推,依次类推,形成式(4)和(5)所示的训练样本输入数据和训练样本输出数据。测试数据的构建方法与训练样本数据一致。

$$[0080] \quad \begin{bmatrix} L_1 & L_2 & \dots & L_m \\ L_2 & L_3 & \dots & L_{m+1} \\ \cdot & \cdot & \cdot & \cdot \\ L_n & L_{n+1} & \dots & L_{n+m} \end{bmatrix} \quad (4)$$

$$[0081] \quad [x_{n+1} \quad x_{n+2} \quad \dots \quad x_{n+m+1}] \quad (5)$$

[0082] 在完成样本数据构建后,为防止各个数据的取值范围过大对预测误差的不利影响,对所有数据进行归一化处理。通过训练TPENN,模型可以学习目标在一定情况下的运动规律,进而预测目标的运动轨迹。

[0083] 2:确定TPENN神经网络结构

[0084] TPENN神经网络通过内部反馈结构,可以更好地处理时间序列问题。

[0085] 网络一般由输入层、隐含层、输出层和承接层组成。

[0086] 其中承接层的激活函数一般为恒等函数,用于存储记忆隐含层单元上一时刻的输出,可以当作一步延时算子。

[0087] 假设在时刻 t ,TPENN神经网络的输入是 $U(t)$,隐含层的输出为 $X(t)$,整个网络的输出为 $Y(t)$, $\omega^1, \omega^2, \omega^3$ 分别表示输入层与隐含层、承接层与隐含层、隐含层与输出层之间的连接权值。

[0088] b 和 c 分别代表隐含层和输出层的偏置,因此神经网络的计算过程可表示为

$$[0089] \begin{cases} X(t) = f_2(\omega^1 U(t) + b), & t = 0 \\ X(t) = f_2(\omega^1 U(t) + \omega^2 X(t-1) + b), & t > 0 \end{cases} \quad (6)$$

$$[0090] Y(t) = g(\omega^3 X(t) + c) \quad (7)$$

[0091] 神经网络采用时间反向传播算法进行学习和训练,其目标函数为

$$[0092] E(t) = \frac{1}{2} (Y_d(t) - Y(t))^T (Y_d(t) - Y(t)) \quad (8)$$

[0093] $Y_d(t)$ 表示预期的输出。

[0094] 当网络的输入是时间序列数据时,数据距离当前预测时刻越近,对预测值的影响越大。在最小二乘数中引入时间反馈因子 $f_{TP}(t)$ 。

[0095] 时间收益因子可以增强较近时间点的样本数据对预测结果的影响,而降低较远时间点的数据对预测结果的影响。

[0096] 由于神经网络的输入是具有时间序列的历史数据,因此本发明考虑引入时间回归因子 $f_{TP}(t)$ 转化为神经网络的目标函数。

[0097] 提出了一种基于时间回报因子的TPENN模型。

[0098] 下式为时间收益因素和改进后的目标函数 $E_{TP}(t)$:

$$[0099] f_{TP}(t) = \frac{1}{1 + e^{(\alpha - 2\alpha t \cdot \lambda^{-1})}} \quad (9)$$

$$[0100] E_{TP}(t) = \sum_{k=1}^t f_{TP}(\lambda) E(\lambda) \quad (10)$$

[0101] α 表示是打折率, λ 表示样本数量, t 表示迭代次数。

[0102] 3: 训练TPENN神经网络

[0103] 在完成样本数据构建后,确定TPENN神经网络结构,利用样本数据对神经网络进行训练,输入训练样本,根据BPTT(Back Propagation Through Time)算法,得到 $E_{TP}(t)$ 的偏导数,其权重计算如式(11)所示:

$$[0104] \quad \begin{cases} \Delta\omega_{iq}^1(t) = \sum_{\lambda=1}^t \eta_1 f_{TP}(\lambda) \delta_j^h u_q(\lambda-1) \\ \Delta\omega_{ij}^2(t) = \sum_{\lambda=1}^t \eta_2 f_{TP}(\lambda) \sum_{i=1}^m (\delta_i^0 \omega_{ij}^3) \frac{\partial x_j(\lambda)}{\partial \omega_{ij}^2} \\ \Delta\omega_{ij}^3(t) = \sum_{\lambda=1}^t \eta_3 f_{TP}(\lambda) \delta_i^0 x_j(\lambda) \end{cases} \quad (11)$$

[0105] η_1, η_2, η_3 是 $\omega^1, \omega^2, \omega^3$ 学习步长; $i=1, 2, \dots, m; j=1, 2, \dots, n; q=1, 2, \dots, r; l=1, 2, \dots, n$.

[0106] δ_j^h, δ_i^0 分别表示第 j 层输出层的误差、第 i 层输入层的误差;

[0107] 通过神经网络的训练, 得到 TPENN 神经网络的各项参数。

[0108] 4: 轨迹预测

[0109] 在某一时刻, 目标的运动状态具有绝对连续性。它的运动轨迹不跳跃, 这说明它在某一时刻的位置与它在最后时刻的位置有特定的相关性。在目标轨迹上按一定的时间间隔选择 n 个位置点, t_{n+1} 时刻的目标位置是未知的。根据轨迹连续性的性质, 前 n 个的位置可以用来预测时刻 t_{n+1} 的位置。

[0110] 本发明将轨迹预测问题转化为时间序列预测问题。利用 TPENN 挖掘轨迹数据的内容相关性, 解决轨迹预测问题。

[0111] 步骤 2、根据预测的逃逸目标轨迹, 基于深度强化进行动态目标跟踪控制, 实现安全避障、水流补偿和轨迹跟踪。

[0112] 在动态逃逸目标轨迹预测的基础上, 针对自主水下航行器非线性模型不确定特性和水流影响, 借助强化学习方法的非线性、强学习, 不依赖模型的特性, 同时引入深度学习算法克服强化学习的维度灾难问题, 从而实现航行器自主动态目标跟踪控制。

[0113] 目标跟踪控制系统分为三个子任务: 安全避障、水流补偿和轨迹跟踪。

[0114] 安全避障任务是使自主水下航行器跟踪过程中远离障碍物;

[0115] 水流补偿就是跟踪过程中补偿水流的影响;

[0116] 而轨迹跟踪任务是使自主水下航行器能够按照预测生成的轨迹航行, 且不受自主水下航行器模型不确定与非线性耦合影响。

[0117] 图 3 为深度强化学习自主水下航行器动态目标跟踪控制的整体方案。

[0118] 图 4 为深度强化学习模型的基本结构。

[0119] ① 深度强化学习轨迹跟踪控制原理

[0120] 深度强化学习控制器是由一个外环运动学控制器与内环动力学跟踪控制器构成, 控制原理如图 3 所示。

[0121] 外环 DRL (Deep Reinforcement Learning) 运动学跟踪控制器负责产生自主水下航行器的运动速度控制信号;

[0122] 而内环变结构滑模控制器 SMC (Sliding Mode Control) 负责产生自主水下航行器的动力学推进控制信号 (推力与力矩)。

[0123] 整个控制过程为:

[0124] 从预测的逃逸目标轨迹中获取期望的自主水下航行器位姿矢量 $\eta_d = [x_d \ y_d \ z_d \ \phi_d \ \theta_d \ \psi_d]^T$, 与自主水下航行器当前状态矢量 $\eta = [x \ y \ z \ \phi \ \theta \ \psi]^T$ 比较, 产生航向以及位置轨迹误

差,并与自主水下航行器的状态信息整合生成状态误差向量,输入DRL深度强化学习模型中,根据强化学习算法输出最优的运动学速度控制信号矢量

$\bar{V}_d = [u_d \ v_d \ w_d \ p_d \ q_d \ r_d]^T$,并与自主水下航行器输出速度信号矢量 \bar{V} 比较,产生自主水下航行器跟踪速度误差 \bar{e}_v ,输入内环变结构滑模控制器SMC,由此产生自主水下航行器的动力学推进控制信号(推力与力矩) $\bar{\tau}$,驱动自主水下航行器的速度矢量矩阵 \bar{V} 跟踪 \bar{V}_d ,通过坐标转换 $\dot{\eta} = J(\eta)\bar{V}$ 得到AUV实际位姿矢量 $\eta = [x \ y \ z \ \phi \ \theta \ \psi]^T$,从而构成一个双闭环串级跟踪控制系统,实现自主水下航行器对动态目标的高效精准跟踪;

[0125] 其中, $x_d \ y_d \ z_d$ 分别为自主水下航行器在惯性坐标系中期望的位置; $\phi_d \ \theta_d \ \psi_d$ 分别为自主水下航行器对惯性坐标系中期望的横倾角、纵倾角和艏向角;

[0126] $x \ y \ z$ 分别为自主水下航行器在惯性坐标系中当前的位置; $\phi \ \theta \ \psi$ 分别为自主水下航行器对惯性坐标系中当前的横倾角、纵倾角和艏向角;

[0127] $u_d \ v_d \ w_d$ 分别为自主水下航行器线速度矢量在载体坐标系中的3个分量; $p_d \ q_d \ r_d$ 分别为自主水下航行器角速度矢量在载体坐标系中的3个分量。

[0128] ②深度强化学习模型设计

[0129] 图4为深度强化学习模型的基本结构。

[0130] DRL控制模型包含三个模块:避障模块、导航模块和动作调度器。

[0131] 避障模块是DRL通过双Q-network (two-stream Q-network) 学习模式实现避开障碍物。设置碰撞负奖励,将避障目标表达为强化学习的一般目标,即寻求一种使总奖励最大化的策略。经过充分的训练,由双Q-network近似得到学习Q函数,最终收敛于最优Q函数附近。配备预先训练过的避障模块的自主水下航行器将自动避开移动的障碍物。

[0132] 导航模块提供了通过在线学习解决目标跟踪的最终策略,该策略也遵循DRL算法。轨迹导航模块使用双Q-network中的一个网络分别将获得的局部障碍物信息、预测轨迹信息转化为奖励(跟踪预测轨迹设置为积极的奖励,碰撞设置为消极的奖励,并在每一步设置了小的时间惩罚),并将奖励结果立即反馈以改进导航策略;同时,使用另一个完全连接的Q-network近似在线学习的行为价值函数。通过双Q-network学习的方式完成自主水下航行器导航策略的在线训练。

[0133] 动作调度器:

[0134] 旨在对轨迹导航模块在线部分进行有效的探索和加速训练,并生成自主水下航行器运动学控制指令。动作调度器利用从避障模块、轨迹导航模块学习到的策略,在重置缓冲区中积累足够的高质量训练经验。随着时间的推移,调度器逐渐从静态学习策略转移到在线学习策略。在线探索和优化一直持续到产生稳定的自主水下航行器跟踪的运动学速度控制信号 \bar{V}_d 。③水流补偿模型设计

[0135] 水流是水相对稳定的非周期性流动,随季节、气候、海域、地形、深度变化而变化,是时间和空间的复杂函数,目前很难用精确的数学表达式描述其运动规律。但是,考虑水下航行器的航行能力有限,只能在特定的时间段和水域内作一定距离航行,当水流随时间和空间的变化是在较大范围内发生的,而在有限的特定水域和特定时间段内,水流的流速和流向都是比较稳定的。因此,可将水流的运动简化为变化较缓慢的流动。

$$[0136] \quad \psi(x, y) = 1 - \tanh\left(\frac{y - B(t) \cos(k(x - ct))}{(1 + k^2 B(t)^2 \sin^2(k(x - ct)))^{1/2}}\right) \quad (2)$$

[0137] 其中, x, y 为惯性坐标系中自主水下航行器平面坐标, $B(t)$ 为水流模型函数, k, c 为运动参数, 对水流函数分别在 x 方向与 y 方向求导, 可得到水流在 x 方向与 y 方向的分量。将水下航行器工作空间的三维水下环境按深度分层, 每层可看作一个二维平面坐标系统, 并对其进行栅格化处理, 每个栅格内的海流模型相同, 从而设计出三维水流模型。

[0138] 对于水流对自主水下航行器轨迹跟踪跟踪控制的影响, 可以通过将时变立体水流模型加入到载体坐标系的速度矢量中进行控制补偿。

[0139] 将水流的速度加入到自主水下航行器的运动控制中, 最后转变为一种附加的力和力矩的作用, 在水流存在的情况下, 水下航行器的受力除了其自身运动引起的自主水下航行器水动力(力矩)和推进器产生的推力(力矩)外, 还受到水流的干扰力(力矩), 因此可将自主水下航行器动力学方程的动力项变为: 水动力(力矩)+推进机构推力(力矩)+水流干扰力(力矩), 从而综合考虑水流的动力学影响, 使得动力学控制满足目标跟踪要求, 实现自主水下航行器高效可靠的跟踪控制。

[0140] 对于本领域技术人员而言, 显然本发明不限于上述示范性实施例的细节, 而且在不背离本发明的精神或基本特征的情况下, 能够以其他的具体形式实现本发明。因此, 无论从哪一点来看, 均应将实施例看作是示范性的, 而且是非限制性的, 本发明的范围由所附权利要求而不是上述说明限定, 因此旨在将落在权利要求的等同要件的含义和范围内的所有变化囊括在本发明内。不应将权利要求中的任何附图标记视为限制所涉及的权利要求。

[0141] 此外, 应当理解, 虽然本说明书按照实施方式加以描述, 但并非每个实施方式仅包含一个独立的技术方案, 说明书的这种叙述方式仅仅是为清楚起见, 本领域技术人员应当将说明书作为一个整体, 各实施例中的技术方案也可以经适当组合, 形成本领域技术人员可以理解的其他实施方式。

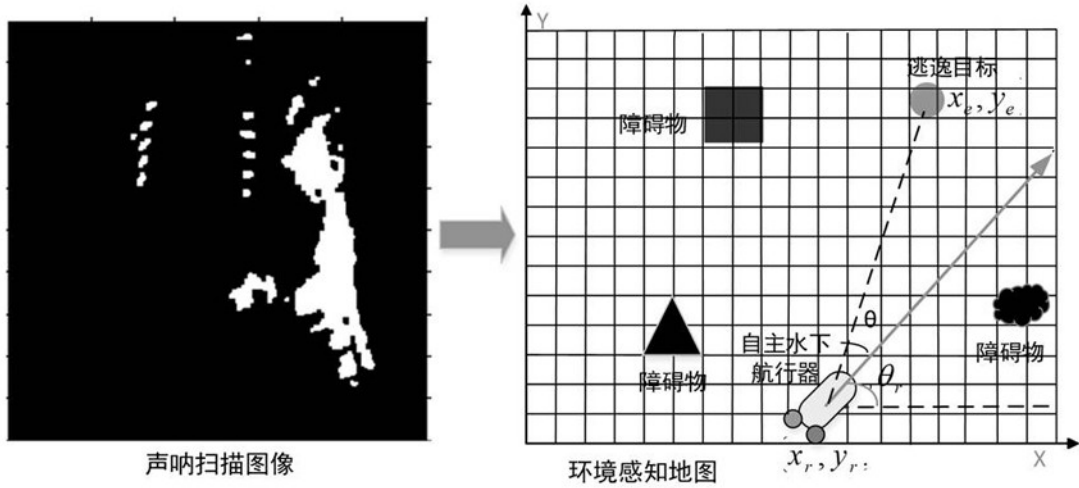


图1

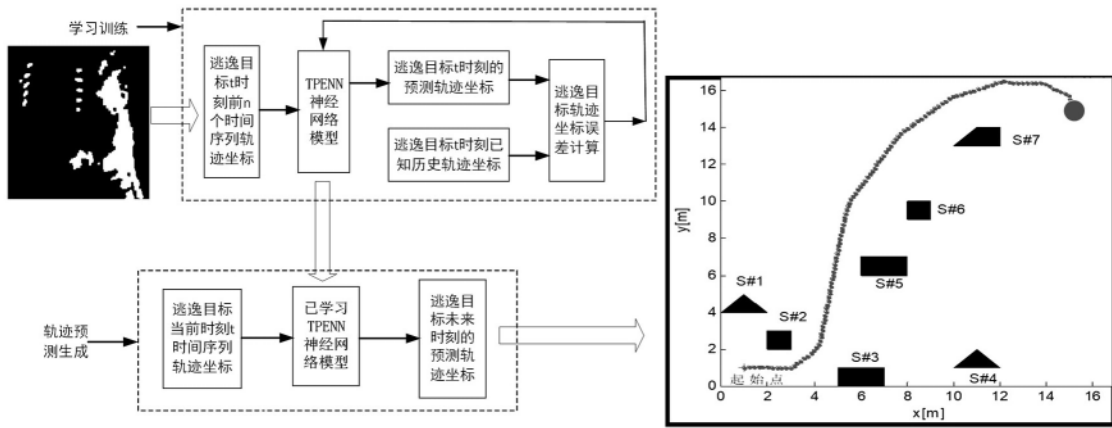


图2

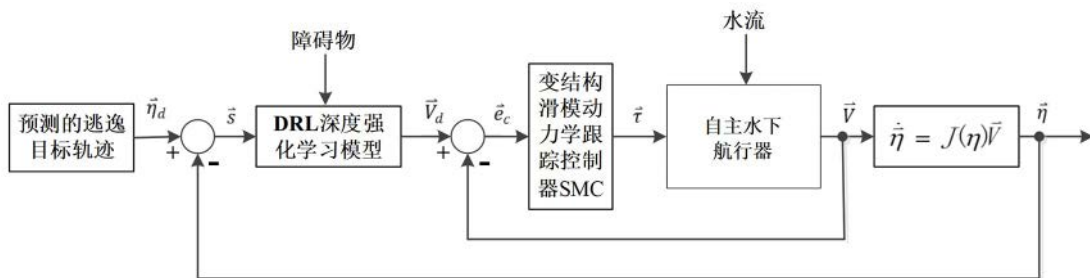


图3

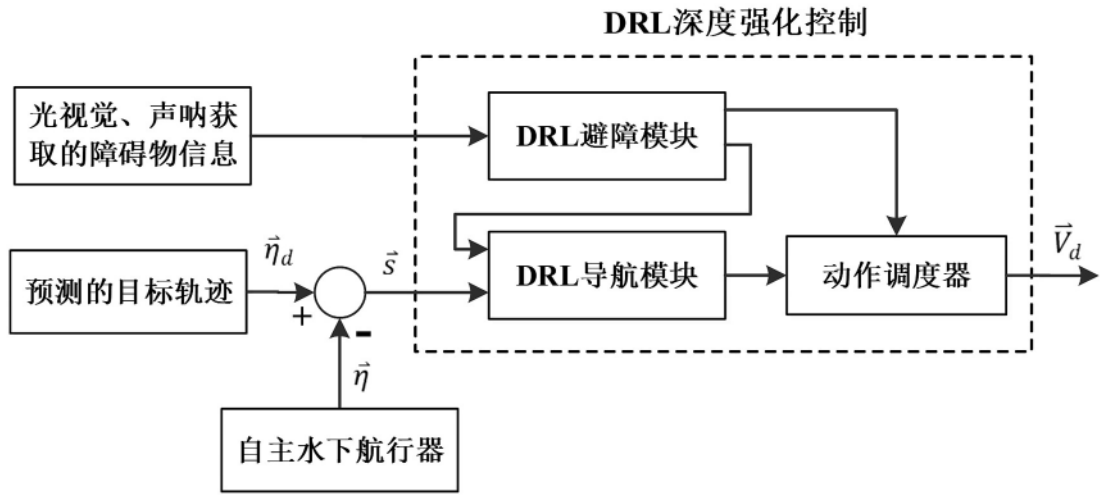


图4

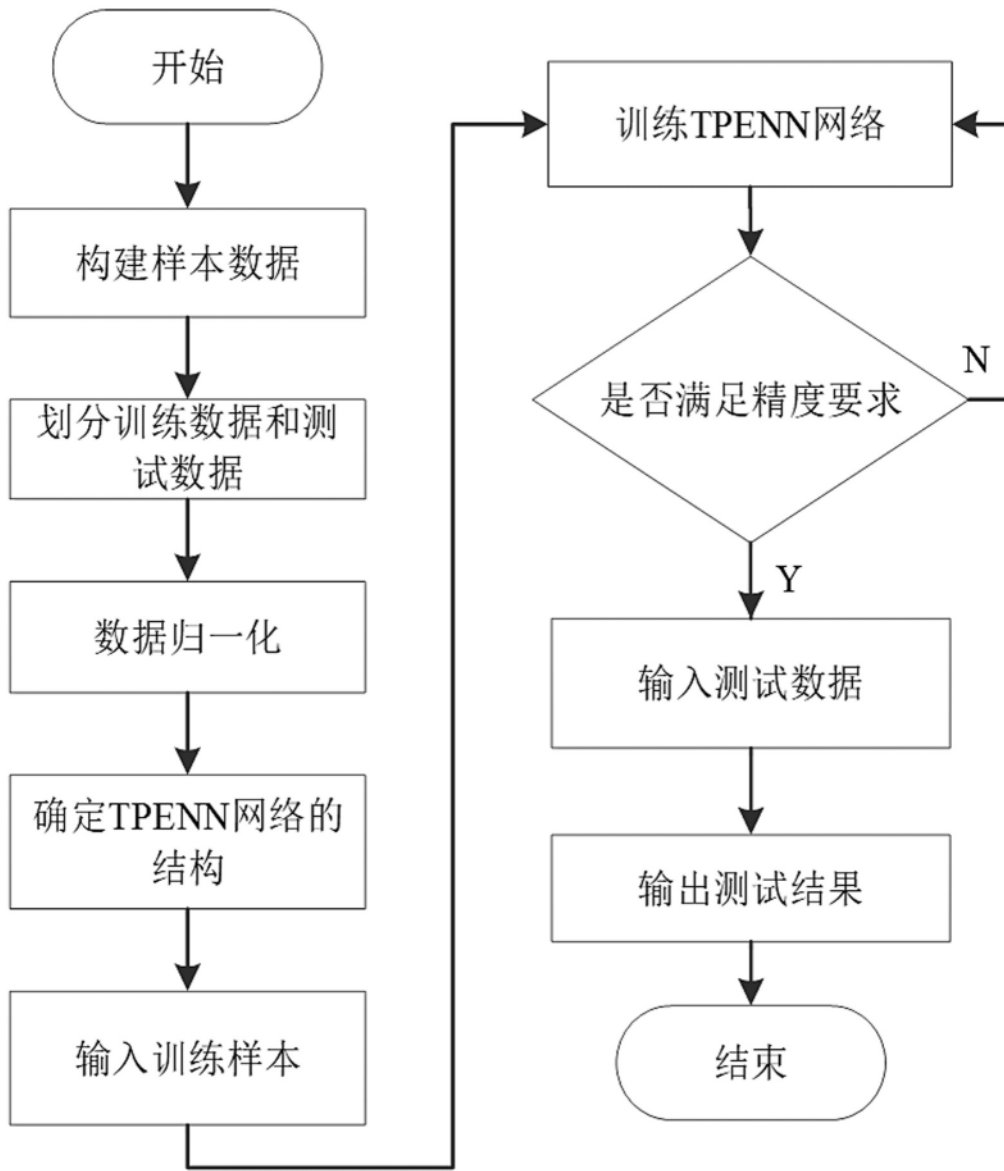


图5