



控制理论与应用  
*Control Theory & Applications*  
ISSN 1000-8152, CN 44-1240/TP

## 《控制理论与应用》网络首发论文

题目：基于递归神经网络的再入飞行器最优姿态控制  
作者：吉月辉，周海亮，车适行，高强  
网络首发日期：2020-10-16  
引用格式：吉月辉，周海亮，车适行，高强. 基于递归神经网络的再入飞行器最优姿态控制[J/OL]. 控制理论与应用.  
<https://kns.cnki.net/kcms/detail/44.1240.tp.20201015.1443.036.html>



**网络首发：**在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

**出版确认：**纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

## 基于递归神经网络的再入飞行器最优姿态控制

吉月辉<sup>1†</sup>, 周海亮<sup>2</sup>, 车适行<sup>1</sup>, 高强<sup>1</sup>

(1. 天津理工大学 电气电子工程学院;天津市复杂系统控制理论及应用重点实验室, 天津 300384;

2. 天津市计量监督检测科学研究院, 天津 300192)

**摘要:** 针对再入飞行器的姿态跟踪问题, 基于递归神经网络提出最优跟踪控制, 采用反步法和递归神经网络, 设计自适应前馈控制, 将再入飞行器的最优姿态跟踪问题转化为等价的姿态角误差/角速率误差最优调节问题. 采用自适应动态规划技术, 解决最优调节问题. 引入神经网络估计最优控制中的代价函数, 推导最优反馈控制律, 同时保证Hamilton-Jacobi-Isaacs (HJI)方程估计误差最小化. 采用Lyapunov 理论, 保证闭环系统中所有信号, 包括姿态角跟踪误差是一致最终有界的. 在MATLAB/SIMULINK中仿真验证了所提出控制策略的有效性.

**关键词:** 再入飞行器;最优控制;自适应动态规划;递归神经网络;姿态跟踪

**引用格式:** 吉月辉, 周海亮, 车适行等. 基于递归神经网络的再入飞行器最优姿态控制. 控制理论与应用, 2020, 37(x): xxx – xxx

## Recurrent neural network-based optimal attitude control of reentry vehicle

JI Yuehui<sup>1†</sup>, ZHOU Hai-liang<sup>2</sup>, CHE Shi-xing<sup>1</sup>, GAO Qiang<sup>1</sup>

(1. School of Electrical and Electronic Engineering, and Tianjin Key Laboratory for Control Theory & Applications in Complicated Systems, Tianjin University of Technology, Tianjin City 300384, China;

2. Tianjin Institute of Metrological Supervision and Testing(TIMST), Tianjin City 300192, China)

**Abstract:** An optimal control is proposed based on recurrent neural networks (RNNs) for the attitude tracking problem of reentry vehicle. Firstly, backstepping and RNNs are introduced to accomplish the adaptive feedforward control. The optimal attitude tracking problem of the reentry vehicle is transformed into the equivalent optimal regulation problem for attitude angle error/angular rate error. Then, adaptive dynamic programming is adopted to fulfill the optimal regulation problem. The neural network is utilized to estimate the cost function in the optimal control, subsequently the optimal feedback control law is constructed, and the estimation error in HJI equation is minimized. The stability analysis based on Lyapunov theory can ensure that all the signals in the closed-loop system, especially attitude angle error, are uniformly ultimately bounded. The effectiveness of the proposed control strategy is verified by numerical simulation in MATLAB/SIMULINK environment.

**Key words:** reentry vehicle; optimal control; adaptive dynamic programming; recurrent neural networks; attitude tracking control

**Citation:** JI Yuehui, ZHOU Hailiang, CHE Shixing, et al. Recurrent neural network-based optimal attitude control of reentry vehicle. *Control Theory & Applications*, 2020, 37(x): xxx – xxx

### 1 引言

在航天任务中, 姿态控制是再入飞行器控制系统的重要环节之一, 实现精确的姿态控制可以提供足够的气动力, 保证再入飞行器能够稳定地跟踪制导参考指令. 再入飞行器飞行速度快, 在飞行过程中易受到

风扰、大气密度和地球引力变化的影响. 由于气动参数受高度和速度变化的影响, 再入飞行器的气动参数是时变的. 因此, 再入飞行器的旋转运动方程具有强非线性、耦合性和不确定性等特点<sup>[1]</sup>.

已有多种控制方法用于再入飞行器的姿态控制器

<sup>†</sup> 通信作者. E-mail: jiyuehuitju@126.com.

天津市教委科研项目(2017KJ249) 资助.

Supported by the Research Project of Tianjin Municipal Education Commission(2017KJ249)

设计: 反馈线性化<sup>[2]</sup>、滑模控制<sup>[3-4]</sup>、自抗扰控制<sup>[5-6]</sup>、反步法等. 反步法通过对再入飞行器进行递推系统化和结构化设计, 获得良好的全局或局部稳定性. 针对存在输入约束、模型不确定性和外界干扰的可重复使用运载火箭, 基于自适应滤波反步法, 提出再入段姿态跟踪策略<sup>[7]</sup>. 针对升力体再入飞行器发生执行机构故障情况, 基于反步法设计姿态容错控制律, 无需飞行器转动惯量信息, 同时有效抑制外界干扰<sup>[8]</sup>. 但上述研究均未考虑系统的最优控制性能. 近年来, 最优控制设计的研究成果斐然, 对再入飞行器的姿态跟踪而言, 最优控制可优化姿态跟踪精度、控制力矩等性能指标, 已成为现代控制的研究热点之一. 传统非线性最优系统的设计需要推导Hamilton-Jacobi-Isaacs (HJI) 方程的解, 但解析推导HJI方程较困难. 为此, 采用自适应动态规划技术, 逼近HJI方程的最优解<sup>[9-11]</sup>.

针对再入飞行器的不确定问题, 神经网络(Neural Networks, NNs)、模糊逻辑系统(Fuzzy Logic Systems, FLs)因其对非线性函数的近似特性, 用于逼近难以精确建模的复杂非线性系统. 综合反步法和神经网络, 实现气动特性不确定性、外部干扰和执行器饱和情况下高超声速再入飞行器的姿态跟踪控制问题. 借助径向基函数神经网络(Radical Basis Function Neural Networks, RBFNN)的通用逼近能力, 设计自适应扰动观测器在线估计外部扰动<sup>[12]</sup>. 针对具有气动-伺服弹性和扰动的可重复使用运载火箭, 采用模糊逻辑系统设计扰动观测器, 结合自适应滑模技术, 设计姿态跟踪控制器<sup>[13]</sup>. 与传统的函数逼近器相比, 递归神经网络(Recurrent Neural Networks, RNNs)具有优越性能, 包括动态性能和存储信息的能力. 在递归神经元中引入内部反馈回路, 无需外部延迟反馈就可捕获动态信息. 通过自身操作可处理时变输入或时变输出. 因此, RNNs 是性能良好的动态映射, 可采用较少神经元以高精度逼近非线性函数.

为此, 针对再入飞行器, 实现基于RNNs的自适应最优姿态控制, 由前馈控制和反馈控制两部分组成. 本文的创新点在于: 1) 基于RNN和反步法, 设计前馈控制器, 将再入飞行器的姿态跟踪问题转化为姿态角跟踪误差/角速率跟踪误差系统的最优反馈控制问题. 其中, RNNs 估计再入飞行器中的未知扰动和非线性函数. 2) 基于自适应动态规划, 设计最优反馈控制器, 保证姿态角可有界地跟踪制导参考信号, 闭环系统所有信号是有界的, 同时保证代价函数最小化.

本文章节安排如下: 第2部分阐述了高超声速再入飞行器模型与递归网络的预备知识. 第3部分设计前馈控制器和最优反馈控制器. 第4部分讨论闭环系统的稳定性. 第5部分中通过仿真研究, 验证了所提出方法的有效性. 第6部分是论文的结论.

## 2 问题描述

### 2.1 高超声速再入飞行器数学模型

机体坐标系下, 再入飞行器旋转运动方程为

$$\begin{cases} \dot{\theta} = R\omega + d_\theta, \\ \dot{\omega} = -I_0^{-1}\Omega I_0\omega + I_0^{-1}M + d_\omega, \end{cases} \quad (1)$$

其中:  $\theta = [\alpha, \beta, \sigma]^T \in \mathbb{R}^3$  是姿态角向量, 分别是攻角、侧滑角和倾斜角;  $\omega = [p, q, r]^T \in \mathbb{R}^3$  是角速率向量, 分别是滚转角速率、俯仰角速率、偏航角速率;  $M = [M_x, M_y, M_z]^T \in \mathbb{R}^3$  是控制力矩输入向量, 分别是滚转力矩、俯仰力矩和偏航力矩;  $R = \begin{bmatrix} -\cos\alpha\tan\beta & 1 & -\sin\alpha\tan\beta \\ \sin\alpha & 0 & -\cos\alpha \\ -\cos\alpha\cos\beta & -\sin\beta & -\sin\alpha\cos\beta \end{bmatrix} \in \mathbb{R}^{3 \times 3}$ ;

$I_0 = \begin{bmatrix} I_{xx} & 0 & -I_{xz} \\ 0 & I_{yy} & 0 \\ I_{xz} & 0 & I_{zz} \end{bmatrix} \in \mathbb{R}^{3 \times 3}$  是标称的转动惯量矩

阵;  $\Omega = \begin{bmatrix} 0 & -r & q \\ r & 0 & -p \\ -q & p & 0 \end{bmatrix} \in \mathbb{R}^{3 \times 3}$ ;  $d_\theta \in \mathbb{R}^3$  是未知扰动向量;  $d_\omega \in \mathbb{R}^3$  是总力矩扰动矢量, 包括气动力矩不确定性和力矩扰动.

再入飞行器旋转运动方程可转化为严反馈形式:

$$\begin{cases} \dot{\theta} = f_\theta(\theta) + g_\theta(\theta)\omega + d_\theta, \\ \dot{\omega} = f_\omega(\theta, \omega) + g_\omega(\theta, \omega)M + d_\omega, \end{cases} \quad (2)$$

其中:  $f_\theta = 0, g_\theta = R, f_\omega = -I_0^{-1}\Omega I_0\omega, g_\omega = -I_0^{-1}$ .

本文的控制目标为: 针对再入飞行器的旋转动态(1), 基于递归神经网络设计最优控制器, 使得姿态角以最优方式跟踪有界的参考信号, 保证闭环系统中所有信号是一致最终有界的(Uniformly Ultimate Bounded, UUB), 同时最小化代价函数.

### 2.2 递归神经网络

如图1所示, 递归神经网络由输入层、具有反馈单元的隐含层和输出层组成. 与传统的径向基函数神经网络相比, 递归神经网络在隐含层中引入内部反馈收集内部状态信息, 提高对未知非线性光滑函数的逼近能力.

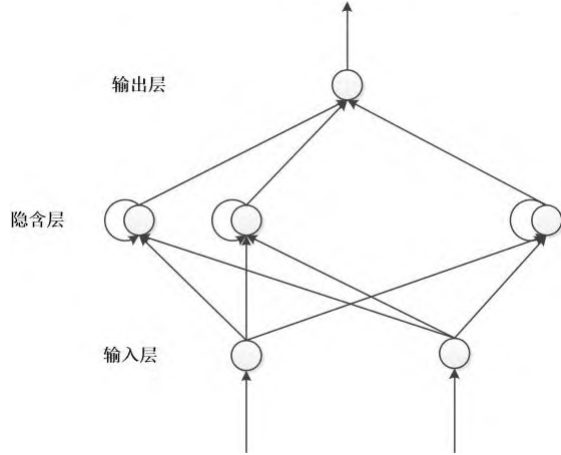


图 1 递归神经网络结构图

Fig. 1 The flow chart of RNNs

## 1) 输入层.

对输入层的每个节点 $i$ , 输入和输出为:

$$\begin{aligned} \text{net}_i(N) &= x_i(N), \\ O_i(N) &= \chi_i(\text{net}_i(N)), \end{aligned} \quad (3)$$

其中:  $x_i$ 是第 $i$ 个输入节点,  $N$ 是迭代次数,  $\chi_i$ 是激活函数.

## 2) 隐含层.

对隐含层的每个节点 $j$ , 输入和输出为:

$$\text{net}_j(N) = w_j O_j(N-1) + \sum_i w_{ji} O_i(N), \quad (4)$$

$$O_j(N) = \chi_j(\text{net}_j(N)),$$

其中:  $w_j$ 是递归权重,  $w_{ji}$ 是输入层和隐含层的权重,  $\chi_j$ 是激活函数.

## 3) 输出层.

对输出层的每个节点 $k$ , 输入和输出是:

$$\begin{aligned} \text{net}_k(N) &= \sum_j w_{kj} O_j(N), \\ O_k(N) &= \chi_k(\text{net}_k(N)), \end{aligned} \quad (5)$$

其中:  $w_{kj}$ 是隐含层和输出层的权重,  $\chi_k$ 是激活函数.

递归神经网络的输出可表示为:

$$\Gamma_k(w_{ok}, w_R, w_H) = w_{ok}^T \Omega_H(w_R, w_H) \quad (6)$$

其中:  $w_{ok} = [w_{k1} \cdots w_{kj}]^T$ ,  $\Omega_H = [O_1 \cdots O_j]^T$   
 $w_R = [w_1 \cdots w_j]^T$ ,  $w_{Hj} = [w_{j1} \cdots w_{ji}]^T$

**假设 1** [14] 对递归神经网络, 存在最优权重参数  $w_{ok}^*$ ,  $w_R^*$ ,  $w_H^*$  以估计未知函数  $d(\cdot)$ , 即  $d = w_o^{*T} \Omega_H^*(w_R^*, w_H^*) + \varepsilon$ , 其中  $\varepsilon$  是估计误差.

假设1说明递归神经网络具有较强的学习能力和自适应能力, 可充分逼近复杂的非线性函数, 学习和适应不确定系统的动态特性. 未知函数可表示为:

$$\hat{d} = \hat{w}_o^T \hat{\Omega}_H(\hat{w}_R, \hat{w}_H) \quad (7)$$

逼近误差为:

$$\begin{aligned} d - \hat{d} &= w_o^{*T} \Omega_H^* - \hat{w}_o^T \hat{\Omega}_H + \varepsilon \\ &= w_o^{*T} (\hat{\Omega}_H + \tilde{\Omega}_H) - \hat{w}_o^T \hat{\Omega}_H + \varepsilon \\ &= \tilde{w}_o^T \hat{\Omega}_H + \hat{w}_o^T \tilde{\Omega}_H + \tilde{w}_o^T \tilde{\Omega}_H + \varepsilon \end{aligned} \quad (8)$$

其中:  $\tilde{w}_o = w_o^* - \hat{w}_o$ ,  $\tilde{\Omega}_H = \Omega_H^* - \hat{\Omega}_H$ , 逼近总误差为  $\varepsilon_0 = \tilde{w}_o^T \tilde{\Omega}_H + \varepsilon$ .

采用泰勒展开, 将非线性函数线性化:

$$\begin{aligned} \tilde{\Omega}_H &= \frac{\delta \tilde{\Omega}_H}{\delta w_R} \Big|_{w_R=\hat{w}_R} (w_R^* - \hat{w}_R) \\ &+ \frac{\delta \tilde{\Omega}_H}{\delta w_H} \Big|_{w_H=\hat{w}_H} (w_H^* - \hat{w}_H) + \varepsilon_D \\ &= D\tilde{\Omega}_{HR} \cdot \tilde{w}_R + D\tilde{\Omega}_{HH} \cdot \tilde{w}_H + \varepsilon_D \end{aligned} \quad (9)$$

其中:  $D\tilde{\Omega}_{HR}$ ,  $D\tilde{\Omega}_{HH}$  是相应的偏导矩阵,  $\tilde{w}_R = w_R^* - \hat{w}_R$ ,  $\tilde{w}_H = w_H^* - \hat{w}_H$ ,  $\varepsilon_D$  是高阶项.

未知函数逼近误差(8) 可表示为:

$$\begin{aligned} d - \hat{d} &= \tilde{w}_o^T \hat{\Omega}_H + \hat{w}_o^T (D\tilde{\Omega}_{HR} \cdot \tilde{w}_R + D\tilde{\Omega}_{HH} \cdot \tilde{w}_H \\ &+ \varepsilon_D) + \varepsilon_0 \\ &= \tilde{w}_o^T \hat{\Omega}_H + \hat{w}_o^T (D\tilde{\Omega}_{HR} \cdot \tilde{w}_R + D\tilde{\Omega}_{HH} \cdot \tilde{w}_H) \\ &+ \Delta \end{aligned} \quad (10)$$

假设总匹配误差  $\Delta = \hat{w}_o^T \varepsilon_D + \varepsilon_0$  有界, 即  $|\Delta| \leq \bar{\Delta}$ , 其中  $\bar{\Delta}$  为正常数.

## 3 自适应RNN 最优控制

在本节中, 针对再入飞行器的姿态角子系统、角速率子系统分别设计自适应RNN最优控制  $u_i = u_{bi} + u_{ai}$ ,  $i = \theta, \omega$ ,  $u_{bi}$  是基于反步法和RNN的前馈跟踪控制, 将姿态跟踪问题转化为跟踪误差的镇定问题;  $u_{ai}$  是基于自适应动态规划(Adaptive Dynamic Programming, ADP)的最优反馈控制, 保证闭环系统的稳定性和代价函数的最小化. 控制系统框图如图2所示.

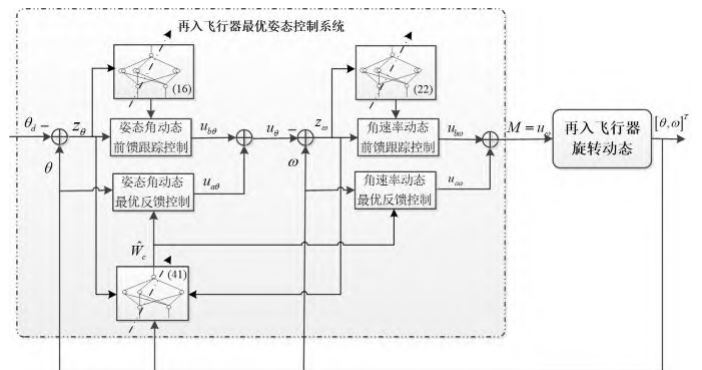


图 2 控制系统框图

Fig. 2 Block diagram of control system



### 3.1 前馈跟踪控制

为设计前馈跟踪控制, 引入如下坐标变换:

$$\begin{aligned} z_\theta &= \theta - \theta_d, \\ z_\omega &= \omega - u_\theta, \end{aligned} \quad (11)$$

其中:  $z_\theta, z_\omega$  是姿态角跟踪误差、角速率跟踪误差,  $u_\theta, u_\omega$  分别是姿态角、角速率子系统的自适应RNN最优控制律:

$$u_i = u_{bi} + u_{ai}. \quad (12)$$

第1步 姿态角跟踪误差 $z_\theta$ , 其动态如下:

$$\begin{aligned} \dot{z}_\theta &= \dot{\theta} - \dot{\theta}_d = g_\theta(\theta)\omega + d_\theta - \dot{\theta}_d \\ &= g_\theta(\theta)(z_\omega + u_\theta) + d_\theta - \dot{\theta}_d \end{aligned} \quad (13)$$

选择Lyapunov候选函数:

$$\begin{aligned} V_\theta &= \frac{1}{2} z_\theta^T z_\theta + \frac{1}{2} \text{tr}(\tilde{w}_{o\theta}^T \lambda_{o\theta}^{-1} \tilde{w}_{o\theta}) \\ &\quad + \frac{1}{2} \text{tr}(\tilde{w}_{R\theta}^T \lambda_{R\theta}^{-1} \tilde{w}_{R\theta}) + \frac{1}{2} \text{tr}(\tilde{w}_{H\theta}^T \lambda_{H\theta}^{-1} \tilde{w}_{H\theta}) \end{aligned} \quad (14)$$

其中:  $\lambda_{o\theta} = \lambda_{o\theta}^T > 0, \lambda_{R\theta} = \lambda_{R\theta}^T > 0, \lambda_{H\theta} = \lambda_{H\theta}^T > 0$  是待设计的控制增益矩阵.

Lyapunov候选函数的一阶导数为:

$$\begin{aligned} \dot{V}_\theta &= z_\theta^T \dot{z}_\theta + \text{tr}(\tilde{w}_{o\theta}^T \lambda_{o\theta}^{-1} \dot{\tilde{w}}_{o\theta}) \\ &\quad + \text{tr}(\tilde{w}_{R\theta}^T \lambda_{R\theta}^{-1} \dot{\tilde{w}}_{R\theta}) + \text{tr}(\tilde{w}_{H\theta}^T \lambda_{H\theta}^{-1} \dot{\tilde{w}}_{H\theta}) \\ &= z_\theta^T (g_\theta(\theta)(z_\omega + u_\theta) + d_\theta - \dot{\theta}_d) \\ &\quad + \text{tr}(\tilde{w}_{o\theta}^T \lambda_{o\theta}^{-1} \dot{\tilde{w}}_{o\theta}) \\ &\quad + \text{tr}(\tilde{w}_{R\theta}^T \lambda_{R\theta}^{-1} \dot{\tilde{w}}_{R\theta}) + \text{tr}(\tilde{w}_{H\theta}^T \lambda_{H\theta}^{-1} \dot{\tilde{w}}_{H\theta}) \end{aligned} \quad (15)$$

采用RNN估计未知非线性函数 $D_\theta = d_\theta$ :

$$\hat{D}_\theta = \hat{w}_{o\theta}^T \hat{\Omega}_{H\theta}(z_\theta, \hat{w}_{R\theta}, \hat{w}_{H\theta}) \quad (16)$$

设计虚拟控制和自适应权重更新律如下:

$$\begin{aligned} u_{b\theta} &= g_\theta^{-1}(-k_\theta z_\theta + \dot{\theta}_d \\ &\quad - \hat{w}_{o\theta}^T \hat{\Omega}_{H\theta}(z_\theta, \hat{w}_{R\theta}, \hat{w}_{H\theta})), \\ \dot{\hat{w}}_{o\theta} &= \lambda_{o\theta} z_\theta \hat{\Omega}_{H\theta}, \\ \dot{\hat{w}}_{R\theta} &= \lambda_{R\theta} z_\theta D \tilde{\Omega}_{HR\theta} \cdot \hat{w}_{o\theta}, \\ \dot{\hat{w}}_{H\theta} &= \lambda_{H\theta} z_\theta D \tilde{\Omega}_{HH\theta} \cdot \hat{w}_{o\theta}, \end{aligned} \quad (17)$$

其中 $k_\theta$  是控制增益.

因此, Lyapunov函数的一阶导数动态为:

$$\begin{aligned} \dot{V}_\theta &= z_\theta^T (g_\theta(\theta)(z_\omega + u_{b\theta} + u_{a\theta}) + D_\theta - \dot{\theta}_d) \\ &\quad + \text{tr}(\tilde{w}_{o\theta}^T \lambda_{o\theta}^{-1} \dot{\tilde{w}}_{o\theta}) + \text{tr}(\tilde{w}_{R\theta}^T \lambda_{R\theta}^{-1} \dot{\tilde{w}}_{R\theta}) \\ &\quad + \text{tr}(\tilde{w}_{H\theta}^T \lambda_{H\theta}^{-1} \dot{\tilde{w}}_{H\theta}) \\ &= z_\theta^T (-k_\theta z_\theta + g_\theta(\theta)(z_\omega + u_{a\theta}) + \tilde{w}_{o\theta}^T \hat{\Omega}_{H\theta} \\ &\quad + \hat{w}_{o\theta}^T (D \tilde{\Omega}_{HR\theta} \cdot \tilde{w}_{R\theta} + D \tilde{\Omega}_{HH\theta} \cdot \tilde{w}_{H\theta}) + \Delta_\theta) \\ &\quad + \text{tr}(\tilde{w}_{o\theta}^T \lambda_{o\theta}^{-1} \dot{\tilde{w}}_{o\theta}) + \text{tr}(\tilde{w}_{R\theta}^T \lambda_{R\theta}^{-1} \dot{\tilde{w}}_{R\theta}) \\ &\quad + \text{tr}(\tilde{w}_{H\theta}^T \lambda_{H\theta}^{-1} \dot{\tilde{w}}_{H\theta}) \\ &= z_\theta^T (-k_\theta z_\theta + g_\theta(\theta)(z_\omega + u_{a\theta}) + \Delta_\theta) \end{aligned} \quad (18)$$

其中 $\Delta_\theta$  为姿态角子系统的有界匹配误差, 满足:  $\Delta_\theta = D_\theta - \hat{D}_\theta - \tilde{w}_{o\theta}^T \hat{\Omega}_{H\theta} - \hat{w}_{o\theta}^T (D \tilde{\Omega}_{HR\theta} \cdot \tilde{w}_{R\theta} + D \tilde{\Omega}_{HH\theta} \cdot \tilde{w}_{H\theta})$

第2步  $z_\omega$  的动态为:

$$\begin{aligned} \dot{z}_\omega &= \dot{\omega} - \dot{u}_\theta \\ &= f_\omega(\theta, \omega) + g_\omega(\theta, \omega)(u_{b\omega} + u_{a\omega}) + d_\omega - \dot{u}_\theta \\ &= f_\omega(\theta_d, u_\theta) + h_\omega(z_\theta, z_\omega) + g_\omega(\theta, \omega)(u_{b\omega} + u_{a\omega}) \\ &\quad + d_\omega - \dot{u}_\theta \end{aligned} \quad (19)$$

其中:  $f_\omega(\theta_d, u_\theta) = -I_0^{-1} \Omega(u_\theta) I_0 u_\theta, h_\omega(z_\theta, z_\omega) = f_\omega(\theta, \omega) - f_\omega(\theta_d, u_\theta)$ .

定义如下Lyapunov 候选函数:

$$\begin{aligned} V &= V_\theta + \frac{1}{2} z_\omega^T z_\omega + \frac{1}{2} \text{tr}(\tilde{w}_{o\omega}^T \lambda_{o\omega}^{-1} \tilde{w}_{o\omega}) \\ &\quad + \frac{1}{2} \text{tr}(\tilde{w}_{R\omega}^T \lambda_{R\omega}^{-1} \tilde{w}_{R\omega}) + \frac{1}{2} \text{tr}(\tilde{w}_{H\omega}^T \lambda_{H\omega}^{-1} \tilde{w}_{H\omega}) \end{aligned} \quad (20)$$

其一阶导数为:

$$\begin{aligned} \dot{V} &= \dot{V}_\theta + z_\omega^T (f_\omega(\theta_d, u_\theta) + h_\omega(z_\theta, z_\omega) \\ &\quad + g_\omega(\theta, \omega)(u_{b\omega} + u_{a\omega}) + d_\omega - \dot{u}_\theta) \\ &\quad + \text{tr}(\tilde{w}_{o\omega}^T \lambda_{o\omega}^{-1} \dot{\tilde{w}}_{o\omega}) + \text{tr}(\tilde{w}_{R\omega}^T \lambda_{R\omega}^{-1} \dot{\tilde{w}}_{R\omega}) \\ &\quad + \text{tr}(\tilde{w}_{H\omega}^T \lambda_{H\omega}^{-1} \dot{\tilde{w}}_{H\omega}) \end{aligned} \quad (21)$$

采用RNN估计未知函数 $D_\omega = f_\omega(\theta_d, u_\theta) + d_\omega$ :

$$\hat{D}_\omega = \hat{w}_{o\omega}^T \hat{\Omega}_{H\omega}(z_\omega, \hat{w}_{R\omega}, \hat{w}_{H\omega}) \quad (22)$$

设计控制律和自适应权重更新律为:

$$\begin{aligned} u_{b\omega} &= g_\omega^{-1}(-k_\omega z_\omega + \dot{u}_\theta \\ &\quad - \hat{w}_{o\omega}^T \hat{\Omega}_{H\omega}(z_\omega, \hat{w}_{R\omega}, \hat{w}_{H\omega})), \\ \dot{\hat{w}}_{o\omega} &= \lambda_{o\omega} z_\omega \hat{\Omega}_{H\omega}, \\ \dot{\hat{w}}_{R\omega} &= \lambda_{R\omega} z_\omega D \tilde{\Omega}_{HR\omega} \cdot \hat{w}_{o\omega}, \\ \dot{\hat{w}}_{H\omega} &= \lambda_{H\omega} z_\omega D \tilde{\Omega}_{HH\omega} \cdot \hat{w}_{o\omega}, \end{aligned} \quad (23)$$

其中 $k_\omega$  是控制增益.

Lyapunov函数(20)的一阶导数为:

$$\begin{aligned}\dot{V} = & -k_\theta z_\theta^T z_\theta + z_\theta^T h_\theta(z_\theta) + z_\theta^T g_\theta(\theta) z_\omega \\ & + z_\theta^T g_\theta(\theta) u_{a\theta} + z_\theta^T \Delta_\theta \\ & + z_\omega^T (-k_\omega z_\omega - g_\theta^T z_\theta - \hat{w}_{o\omega}^T \hat{\Omega}_{H\omega}(z_\omega, \hat{w}_{R\omega}, \hat{w}_{H\omega}) \\ & + f_\omega(\theta_d, u_\theta) + h_\omega(z_\theta, z_\omega) + g_\omega(\theta, \omega) u_{a\omega} + d_\omega) \\ & + \text{tr}(\tilde{w}_{o\omega}^T \lambda_{o\omega}^{-1} \dot{\tilde{w}}_{o\omega}) + \text{tr}(\tilde{w}_{R\omega}^T \lambda_{R\omega}^{-1} \dot{\tilde{w}}_{R\omega}) \\ & + \text{tr}(\tilde{w}_{H\omega}^T \lambda_{H\omega}^{-1} \dot{\tilde{w}}_{H\omega}) \\ = & -k_\theta z_\theta^T z_\theta - k_\omega z_\omega^T z_\omega + z_\theta^T h_\theta(z_\theta) + z_\omega^T h_\omega(z_\theta, z_\omega) \\ & + z_\theta^T g_\theta(\theta) u_{a\theta} + z_\omega^T g_\omega(\theta, \omega) u_{a\omega} + z_\theta^T \Delta_\theta + z_\omega^T \Delta_\omega\end{aligned}\quad (24)$$

其中 $\Delta_\omega$ 为角速率子系统有界匹配误差,满足:  
 $\Delta_\omega = D_\omega - \hat{D}_\omega - \tilde{w}_{o\omega}^T \hat{\Omega}_{H\omega} - \hat{w}_{o\omega}^T (D \tilde{\Omega}_{HR\omega} \cdot \tilde{w}_{R\omega} + D \tilde{\Omega}_{HH\omega} \cdot \tilde{w}_{H\omega})$

因此,  $\dot{V}$ 动态满足:

$$\begin{aligned}\dot{V} = & -Z^T K Z + Z^T H(Z) + Z^T G U_a + Z^T \Delta \\ \leq & -Z^T (K - \frac{1}{2} I) Z + Z^T H(Z) + Z^T G U_a \\ & + \frac{1}{2} \|\Delta\|^2\end{aligned}\quad (25)$$

其中:  $Z = [z_\theta, z_\omega]^T$ ;  $H(Z) = [h_\theta(z_\theta), h_\omega(z_\theta, z_\omega)]^T$ ;  
 $U_a = [u_{a\theta}, u_{a\omega}]^T$ ;  $\Delta_\omega = [\Delta_\theta, \Delta_\omega]^T$ ;  $K = \text{diag}([k_\theta, k_\omega]^T)$ ;  $G = \text{diag}([g_\theta(\theta), g_\omega(\theta, \omega)]^T)$

设计最优反馈控制 $U_a$ , 稳定跟踪误差动态 $\dot{Z} = H(Z) + G U_a$ , 同最小化代价函数, 保证系统的最优性能。

### 3.2 最优反馈控制

考虑如下的跟踪误差动态:

$$\dot{Z} = H(Z) + G U_a \quad (26)$$

选择如下的代价函数:

$$J = \int_t^\infty (\Omega(Z(\tau)) + U_a^T(\tau) \Xi U_a(\tau)) d\tau \quad (27)$$

其中:  $\Omega(Z)$  是半正定函数,  $\Xi$  是正定矩阵。

定义相应的Hamilton 函数:

$$\begin{aligned}h(Z, U_a, J) = & (\nabla J(Z))^T (H(Z) + G U_a) \\ & + \Omega(Z) + U_a^T \Xi U_a\end{aligned}\quad (28)$$

其中 $\nabla J(Z)$  是 $J(Z)$ 相对于 $Z$ 的导数。

定义效用函数为:

$$r(Z, U_a) = \Omega(Z) + U_a^T \Xi U_a \quad (29)$$

选择如下的最优代价函数 $J^*$ :

$$J^* = \min_{U_a^*} \int_t^\infty (\Omega(Z(\tau)) + U_a^{*T}(\tau) \Xi U_a^*(\tau)) d\tau \quad (30)$$

若设计如下的状态反馈最优控制输入:

$$U_a^* = -\frac{1}{2} \Xi^{-1} G^T (\nabla J^*(Z)) \quad (31)$$

其中 $\nabla J^*(Z)$ 是 $J^*(Z)$ 相对于 $Z$ 的导数. 则闭环系统是稳定的, 代价函数(27)达到最优值 $J^*$ , 且Hamilton-Jacobi-Isaacs(HJI) 方程成立:

$$\begin{aligned}h(Z, U_a^*, J^*) \\ = & (\nabla J^*(Z))^T (H(Z) + G U_a^*) + \Omega(Z) + U_a^{*T} \Xi U_a^* \\ = & 0\end{aligned}\quad (32)$$

采用ADP算法求解最优控制问题中HJI方程(32)的解, 引入评价网络估计近似代价函数。

**引理 1** 考虑跟踪误差系统(26)、代价函数(27)和最优反馈控制(31). 假设 $J_s(Z)$ 是一个连续可微的Lyapunov 候选函数, 满足 $J_s(Z) = (\nabla J_s(Z))^T (H(Z) + G U_a^*)$ ,  $\Lambda(Z)$ 是一个正定函数, 即 $\forall Z \neq 0, \Lambda(Z) > 0; Z = 0, \Lambda(Z) = 0$ , 且满足  $\lim_{Z \rightarrow \infty} \Lambda(Z) = \infty$  和 $(\nabla J(Z))^T \Lambda(Z) (\nabla J_s(Z)) = \Omega(Z) + U_a^{*T} \Xi U_a^*$

则下式成立:

$$\begin{aligned}(\nabla J_s(Z))^T (H(Z) + G U_a^*) \\ < -(\nabla J_s(Z))^T \Omega(Z) (\nabla J_s(Z))\end{aligned}\quad (33)$$

注1: 在已有文献中[15]、[16]多次引用引理1. 一般情况下, 假设闭环跟踪误差动态有界, 且其上界是系统跟踪误差的函数, 即 $\|H(Z) + G U_a^*\| \leq \tau \|\nabla J_s(Z)\|$ ,  $\tau > 0$ , 则 $\|(\nabla J_s(Z))^T (H(Z) + G U_a^*)\| \leq \tau \|\nabla J_s(Z)\|^2$ 成立. 考虑到 $(\nabla J_s(Z))^T (H(Z) + G U_a^*) < 0$ , 引理1成立。

采用理想的评价网络逼近最优代价函数:

$$J^*(Z) = W_c^{*T} \varphi_c(Z) + \varepsilon_c(Z) \quad (34)$$

其中:  $W_c^* \in \mathbb{R}^{N_c}$ 是评价网络的理想权重矩阵,  $\varphi_c(Z) : \mathbb{R}^n \rightarrow \mathbb{R}^{N_c}$ 是激活函数,  $N_c$ 是隐含层神经元数量,  $\varepsilon_c(Z)$ 为有限的逼近函数, 且 $\|\varepsilon_c\| \leq \bar{\varepsilon}_c$ ,  $\bar{\varepsilon}_c$ 为正常数。

最优控制器和Hamiltonian函数可设计为:

$$U_a^* = -\frac{1}{2} \Xi^{-1} G^T (\nabla \varphi_c^T(Z) W_c^* + \nabla \varepsilon_c(Z)) \quad (35)$$

$$\begin{aligned}h(Z, W_c^*) = & \Omega(Z) + (W_c^{*T} \varphi_c(Z)) H(Z) \\ & - \frac{1}{4} (\nabla \varphi_c^T(Z) W_c^*)^T G \Xi^{-1} G^T (\nabla \varphi_c^T(Z) W_c^*) + \tau_{HJB}\end{aligned}\quad (36)$$

其中 $\tau_{HJB}$  是HJI方程残差:

$$\begin{aligned}\tau_{HJB} = & (\nabla \varepsilon_c(Z))^T H(Z) \\ & - \frac{1}{2} (\nabla \varepsilon_c(Z))^T G \Xi^{-1} G^T \nabla \varphi_c^T(Z) W_c^*\end{aligned}$$

$$\begin{aligned}
& -\frac{1}{4}(\nabla \varepsilon_c(Z))^T G \Xi^{-1} G^T (\nabla \varepsilon_c(Z)) \\
& = (\nabla \varepsilon_c(Z))^T (H(Z) + G U_a^*) \\
& + \frac{1}{4}(\nabla \varepsilon_c(Z))^T G \Xi^{-1} G^T \nabla \varepsilon_c(Z)
\end{aligned}$$

定义实际评价网络为:

$$\hat{J}(Z) = \hat{W}_c^T \varphi_c(Z) \quad (37)$$

其中  $\hat{W}_c$  是神经网络权重的估计向量。

最优控制器可表示为:

$$\hat{U}_a(Z) = -\frac{1}{2} \Xi^{-1} G^T \nabla \varphi_c^T(Z) \hat{W}_c \quad (38)$$

则Hamilton函数的估计值为:

$$\begin{aligned}
\hat{h}(Z, \hat{W}_c) &= \Omega(Z) + \hat{W}_c^T \varphi_c(Z) H(Z) \\
& - \frac{1}{4}(\nabla \varphi_c^T(Z) \hat{W}_c)^T G \Xi^{-1} G^T \nabla \varphi_c^T(Z) \hat{W}_c \\
& = e_c
\end{aligned} \quad (39)$$

选择如下的目标函数:

$$E_c = \frac{1}{2} e_c^T e_c \quad (40)$$

设计评价网络中权重  $\hat{W}_c$  的更新律, 使得目标函数(40)最小:

$$\begin{aligned}
\dot{\hat{W}}_c &= -\frac{\eta \mu}{\rho^2} e_c + \frac{\eta}{2} \nabla \varphi_c(Z) G \Xi^{-1} G^T \nabla J_s(Z) \\
& + \eta \left( \frac{\mu}{4 \rho^2} (\nabla \varphi_c^T(Z) \hat{W}_c)^T G \Xi^{-1} G^T (\nabla \varphi_c^T(Z) \hat{W}_c) \right. \\
& \left. - (\lambda_2 \hat{W}_c - \lambda_1 \left( \frac{\mu}{\rho} \right)^T \hat{W}_c) \right)
\end{aligned} \quad (41)$$

其中:  $\mu = \nabla \varphi_c(Z) (H(Z) + G \hat{U}_a)$ ,  $\rho = 1 + \mu^T \mu$ ,  $\eta > 0$  是学习率,  $\lambda_1, \lambda_2$  是可调增益,  $\nabla J_s(Z)$  项在引理1中定义。

定义评价网络权重向量的估计误差:

$$\tilde{W}_c = W_c - \hat{W}_c \quad (42)$$

则权重的估计误差动态满足:

$$\begin{aligned}
\dot{\tilde{W}}_c &= -\eta \left( \frac{\mu}{\rho} \right) \left( \frac{\mu}{\rho} \right)^T \tilde{W}_c \\
& + \frac{\eta \mu}{\rho^2} \left( \frac{1}{4} (\nabla \varphi_c^T(Z) \tilde{W}_c)^T G \Xi^{-1} G^T (\nabla \varphi_c^T(Z) \tilde{W}_c) \right. \\
& - \tau_{HJB} - \frac{\eta}{2} \nabla \varphi_c(Z) G \Xi^{-1} G^T \nabla J_s(Z) \\
& - \eta \left( \frac{\mu}{4 \rho^2} (\nabla \varphi_c^T(Z) \hat{W}_c)^T G \Xi^{-1} G^T (\nabla \varphi_c^T(Z) \hat{W}_c) \right. \\
& \left. - (\lambda_2 \hat{W}_c - \lambda_1 \left( \frac{\mu}{\rho} \right)^T \hat{W}_c) \right)
\end{aligned} \quad (43)$$

#### 4 系统稳定性分析

定理1给出本文的主要结论和闭环系统的稳定性分析。

**假设2** 评价网络的理想权重  $W_c$ 、梯度项  $\nabla \varphi_c(Z)$  和  $\nabla \varepsilon_c(Z)$  有界, 即  $\|W_c\| \leq W_c^m$ ,  $\|\nabla \varphi_c(Z)\| \leq \varphi_c^m$ ,  $\|\nabla \varepsilon_c(Z)\| \leq \varepsilon_c^m$ . 因此, HJI方程的残差  $\tau_{HJB}$  有界, 即  $\|\tau_{HJB}\| \leq \tau_{HJB}^m$ .

**定理1** 针对再入飞行器旋转运动模型(1), 自适应前馈控制输入(23), 最优反馈控制(38), 代价函数(27)和权重更新律(41)组成的闭环系统, 所有信号包括输出跟踪误差和评价网络的权重估计误差是UUB的, 并实现了预先定义的代价函数最小化。

证明: 定义Lyapunov候选函数:

$$V_{HJB} = \dot{V} + \tilde{W}_c^T \eta^{-1} \dot{\tilde{W}}_c + J_s(Z) \quad (44)$$

Lyapunov函数的一阶导数为:

$$\begin{aligned}
\dot{V}_{HJB} &= \dot{V} + \tilde{W}_c^T \eta^{-1} \dot{\tilde{W}}_c + (\nabla J_s(Z))^T \dot{Z} \\
& \leq -Z^T (K - \frac{1}{2} I) Z + Z^T H(Z) + Z^T G U_a \\
& + \frac{1}{2} \|\Delta\|^2 + \tilde{W}_c^T \eta^{-1} \dot{\tilde{W}}_c + (\nabla J_s(Z))^T \dot{Z}
\end{aligned} \quad (45)$$

其中评价网络的估计误差动态满足:

$$\begin{aligned}
& \tilde{W}_c^T \eta^{-1} \dot{\tilde{W}}_c \\
& = -\tilde{W}_c^T \left( \frac{\mu \mu^T}{\rho^2} \right) \tilde{W}_c - \tilde{W}_c^T \frac{\mu}{\rho^2} \tau_{HJB} \\
& + \tilde{W}_c^T \frac{\mu}{4 \rho^2} ((\nabla \varphi_c^T(Z) \tilde{W}_c)^T G \Xi^{-1} G^T (\nabla \varphi_c^T(Z) \tilde{W}_c)) \\
& - \frac{1}{2} \tilde{W}_c^T \nabla \varphi_c(Z) G \Xi^{-1} G^T \nabla J_s(Z) \\
& - \tilde{W}_c^T \left( \frac{\mu}{4 \rho^2} (\nabla \varphi_c^T(Z) \hat{W}_c)^T G \Xi^{-1} G^T (\nabla \varphi_c^T(Z) \hat{W}_c) \right. \\
& \left. - (\lambda_2 \hat{W}_c - \lambda_1 \left( \frac{\mu}{\rho} \right)^T \hat{W}_c) \right) \\
& = -\tilde{W}_c^T \left( \frac{\mu \mu^T}{\rho^2} \right) \tilde{W}_c - \tilde{W}_c^T \frac{\mu}{\rho^2} \tau_{HJB} \\
& + \tilde{W}_c^T \frac{\mu}{2 \rho^2} ((\nabla \varphi_c^T(Z) W_c)^T G \Xi^{-1} G^T (\nabla \varphi_c^T(Z) \tilde{W}_c)) \\
& - \tilde{W}_c^T \frac{\mu}{4 \rho^2} ((\nabla \varphi_c^T(Z) W_c)^T G \Xi^{-1} G^T (\nabla \varphi_c^T(Z) W_c)) \\
& - \frac{1}{2} \tilde{W}_c^T \nabla \varphi_c(Z) G \Xi^{-1} G^T \nabla J_s(Z) \\
& + \tilde{W}_c^T (\lambda_2 \hat{W}_c - \lambda_1 \left( \frac{\mu}{\rho} \right)^T \hat{W}_c)
\end{aligned} \quad (46)$$

注 意:  $\tilde{W}_c^T (\lambda_2 \hat{W}_c - \lambda_1 \left( \frac{\mu}{\rho} \right)^T \hat{W}_c) = \tilde{W}_c^T (\lambda_2 W_c - \lambda_2 \tilde{W}_c - \lambda_1 \left( \frac{\mu}{\rho} \right)^T W_c + \lambda_1 \left( \frac{\mu}{\rho} \right)^T \tilde{W}_c)$

定义  $\chi = [(\frac{\mu}{\rho})^T \tilde{W}_c, \tilde{W}_c]^T$ , 因此可推导:

$$\begin{aligned}
& \tilde{W}_c^T \eta^{-1} \dot{\tilde{W}}_c = -\chi^T T \chi + \chi^T E - \\
& \frac{1}{2} \tilde{W}_c^T \nabla \varphi_c(Z) G \Xi^{-1} G^T \nabla J_s(Z)
\end{aligned}$$

其

中:  $T =$

$$\begin{bmatrix} I & -\frac{1}{2}\lambda_1 \\ -\frac{1}{2}\lambda_1 & \lambda_2 - \frac{\mu}{2\rho^2}(\nabla\varphi_c^T(Z)W_c)^T G \Xi^{-1} G^T \nabla\varphi_c^T(Z) \end{bmatrix}, \quad \dot{E} = \begin{bmatrix} -\frac{\tau_{HJB}}{1 + \sigma^T \sigma} \\ (\lambda_2 - \lambda_1(\frac{\mu}{\rho}))^T \\ -\frac{\mu}{4\rho^2}(\nabla\varphi_c^T(Z)W_c)^T G \Xi^{-1} G^T \nabla\varphi_c^T(Z)W_c \end{bmatrix}$$

选择合适参数 $\lambda_1, \lambda_2$ 保证矩阵 $T$ 是正定的, 则:

$$\begin{aligned} \dot{V}_{HJB} \leq & -Z^T(K - \frac{1}{2}I)Z + Z^T H(Z) \\ & + Z^T G U_a + \frac{1}{2}\|\Delta\|^2 - \lambda_{\min}(T)\|\chi\|^2 \\ & + \|E\|\|\chi\| - \frac{1}{2}\widetilde{W}_c^T \nabla\varphi_c(Z) G \Xi^{-1} G^T \nabla J_s(Z) \\ & + (\nabla J_s(Z))^T \dot{Z} \end{aligned} \quad (47)$$

假设在最优控制器下, 下式成立 $\|H(Z) + G U_a^*\| \leq c\sqrt{\|Z\|}, c > 0$ <sup>[11, 15-16]</sup>. 由Young's 不等式, 则:

$$Z^T(H(Z) + G U_a^*) \leq \frac{\|Z\|^2}{2\zeta^2} + \frac{1}{2}\zeta^2 c^2 \|Z\|, \zeta > 0 \quad (48)$$

因此, 式(47)可表示为:

$$\begin{aligned} \dot{V}_{HJB} \leq & -(K_{\min} - \frac{1}{2} - \frac{1}{2\zeta^2})\|Z\|^2 + \frac{1}{2}\zeta^2 c^2 \|Z\| \\ & + \frac{1}{2}\|\Delta\|^2 - \lambda_{\min}(T)\|\chi\|^2 + \|E\|\|\chi\| \\ & - \frac{1}{2}\widetilde{W}_c^T \nabla\varphi_c(Z) G \Xi^{-1} G^T \nabla J_s(Z) \\ & + (\nabla J_s(Z))^T (H(Z) + G \hat{U}_a) \end{aligned} \quad (49)$$

其中控制器增益 $K_{\min} = \min\{K\}$ , 且满足 $K_{\min} > \frac{1}{2} + \frac{1}{2\zeta^2}$ .

考虑(35)和(38), 有:

$$U_a^* - \hat{U}_a = -\frac{1}{2}\Xi^{-1} G^T (\nabla\varphi_c^T(Z)\widetilde{W}_c + \nabla\epsilon_c(Z)) \quad (50)$$

式(49)为:

$$\begin{aligned} \dot{V}_{HJB} \leq & -(K_{\min} - \frac{1}{2} - \frac{1}{2\zeta^2})\|Z\|^2 + \frac{1}{2}\zeta^2 c^2 \|Z\| \\ & + \frac{1}{2}\|\Delta\|^2 - \lambda_{\min}(T)\|\chi\|^2 + \|E\|\|\chi\| \\ & + (\nabla J_s(Z))^T (H(Z) + G U_a^*) \\ & + \frac{1}{2}(\nabla J_s(Z))^T G \Xi^{-1} G^T \nabla\epsilon_c(Z) \\ \leq & -(K_{\min} - \frac{1}{2} - \frac{1}{2\zeta^2})\|Z\|^2 + \frac{1}{2}\zeta^2 c^2 \|Z\| \\ & + \frac{1}{2}\|\Delta\|^2 - \lambda_{\min}(T)\|\chi\|^2 + \|E\|\|\chi\| \\ & - \lambda_{\min}(\Lambda(Z))\|\nabla J_s(Z)\|^2 \\ & + \frac{1}{2}\|\nabla J_s(Z)\|\kappa\epsilon_c^m \end{aligned} \quad (51)$$

可以证明 $\|G \Xi^{-1} G^T\| \leq \kappa$ .

定义:

$$v = \frac{\zeta^4 c^4}{16(K_{\min} - \frac{1}{2} - \frac{1}{2\zeta^2})^2} + \frac{\|E\|^2}{4\lambda_{\min}(T)} + \frac{(\kappa\epsilon_c^m)^2}{16\lambda_{\min}(\Lambda(Z))} + \frac{1}{2}\|\Delta\|^2$$

$$\|Z\| \geq \sqrt{\frac{v}{(K_{\min} - \frac{1}{2} - \frac{1}{2\zeta^2})}} + \frac{\zeta^2 c^2}{4(K_{\min} - \frac{1}{2} - \frac{1}{2\zeta^2})} \quad (52)$$

或

$$\|\chi\| \geq \sqrt{\frac{v}{\lambda_{\min}(T)}} + \frac{\|E\|}{2\lambda_{\min}(T)} \quad (53)$$

或

$$\|\nabla J_s(Z)\| \geq \sqrt{\frac{v}{\lambda_{\min}(\Lambda(Z))}} + \frac{\kappa\epsilon_c^m}{4\lambda_{\min}(\Lambda(Z))} \quad (54)$$

则 $\dot{V}_{HJB} \leq 0$ 成立. 因此, 自适应最优控制器可保证闭环系统中所有信号, 包括输出跟踪误差是一致最终有界的.

## 5 仿真案例

在Matlab/Simulink环境中, 验证所提出的自适应RNN最优控制, 对再入飞行器大机动飞行时的有效性和控制性能. 再入飞行器参数选取X-33的物理参数, 即 $I_{xx} = 434270, I_{xz} = 17880, I_{zz} = 1131541$ . 为了验证控制器的鲁棒性, 仿真中考虑如下的外部扰动和力矩扰动:

$$d_\theta = \begin{bmatrix} \frac{1}{5}(0.1 + \sin t + \sin \frac{t}{5}) \\ \frac{1}{5}(0.1 + \sin t + \sin \frac{t}{5}) \\ \frac{1}{5}(0.1 + \sin t + \sin \frac{t}{5}) \end{bmatrix} \in \mathbb{R}^3,$$



$$d_\omega = \begin{bmatrix} \frac{1}{10}(0.1 + \sin t + \sin \frac{t}{5}) \\ \frac{1}{10}(0.1 + \sin t + \sin \frac{t}{5}) \\ \frac{1}{10}(0.1 + \sin t + \sin \frac{t}{5}) \end{bmatrix} \in \mathbb{R}^3,$$

再入飞行器的初始条件为:  $\theta(0) = [0 \ 0 \ 0]$ ,  $\omega(0) = [0 \ 0 \ 0]$ . 再入飞行时, 侧滑角制导参考信号保持在 0, 攻角和倾斜角的制导参考信号选为方波信号. 仿真参数选择如下: 前馈跟踪控制器中, 控制器增益选为  $k_\theta = 20$ ,  $k_\omega = 10$ . 在姿态角子系统、角速率子系统中分别引入 RNN 网络, 每个 RNN 网络由输入层、隐含层、输出层组成, 依次选用 1 个节点、3 个节点、1 个节点(每个节点为 3 维向量), 激活函数均选为  $\chi(s) = \frac{1}{1 + e^{-s}}$ . RNN 网络权重的初始值选取如下, 其中  $\hat{\omega}_{o\theta}(0) = [0.01, 0.01, 0]^T$ , 其他初始值选为 0 向量, 网络权重在给定初始值基础上更新. 其他参数选为  $\lambda_{o\theta} = \lambda_{R\theta} = \lambda_{H\theta} = \lambda_{o\omega} = \lambda_{R\omega} = \lambda_{H\omega} = 10$ . 在最优反馈控制器中, 半正定函数  $\Omega(Z)$  选为  $\Omega(Z) = \frac{1}{2}(z_\theta^2 + z_\omega^2)$ , 正定矩阵  $\Xi = I_6$ , 激活函数选为  $\varphi_c(Z) = [z_\theta, z_\omega]^T$ , 参数选为  $\eta = 10$ ,  $\lambda_1 = 1$ ,  $\lambda_2 = 10$ .

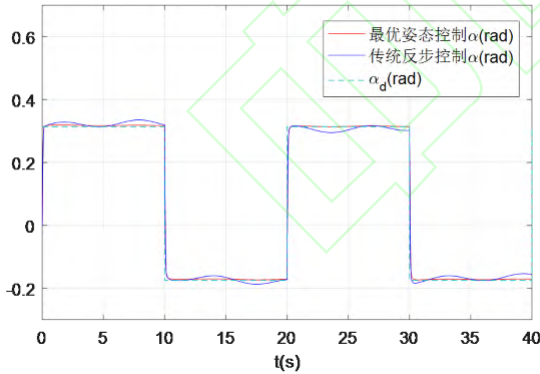


图 3 攻角跟踪轨迹

Fig. 3 The tracking trajectory of angle of attack

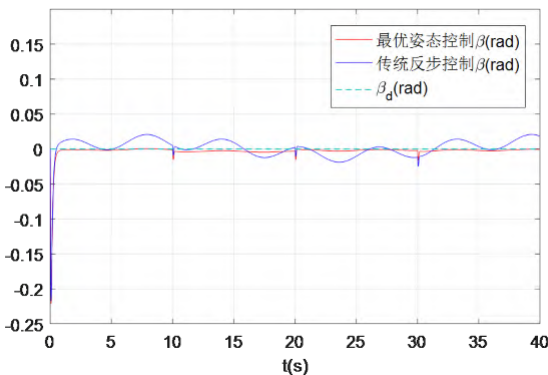


图 4 侧滑角跟踪轨迹

Fig. 4 The tracking trajectory of sideslip angle

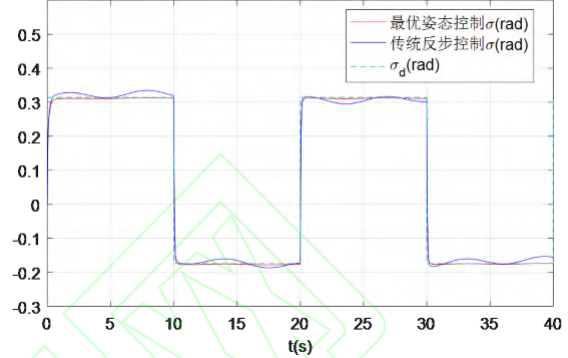


图 5 倾斜角跟踪轨迹

Fig. 5 The tracking trajectory of bank angle

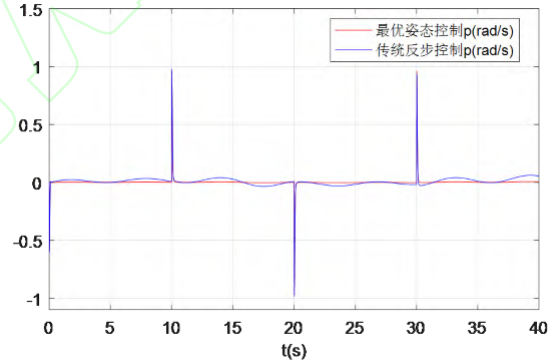


图 6 滚转角速率跟踪轨迹

Fig. 6 The tracking trajectory of roll rate

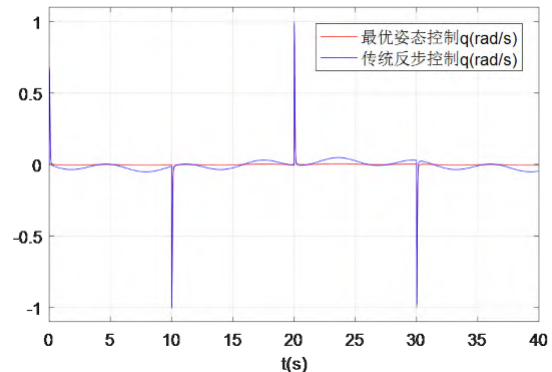


图 7 俯仰角速率跟踪轨迹

Fig. 7 The tracking trajectory of pitch rate

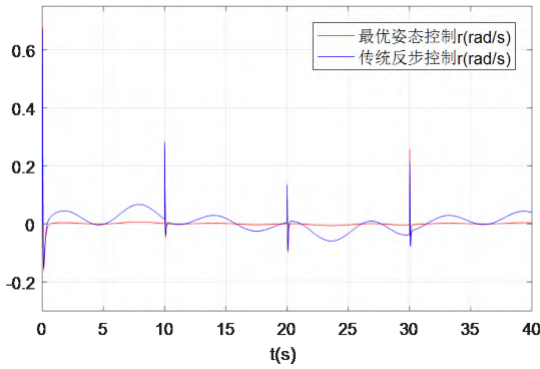


图 8 偏航角速率跟踪轨迹

Fig. 8 The tracking trajectory of yaw rate

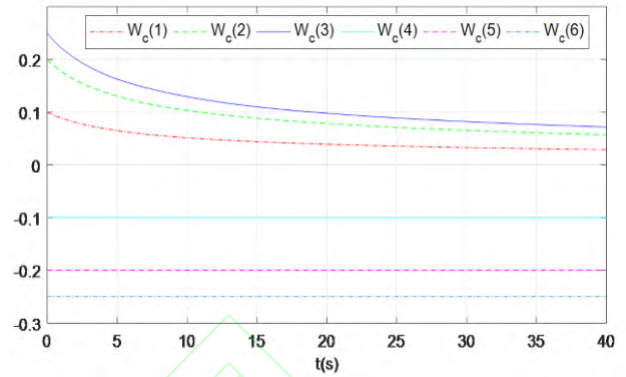


图 11 最优跟踪控制评价网络权重更新律

Fig. 11 Critic neural network weighting update law

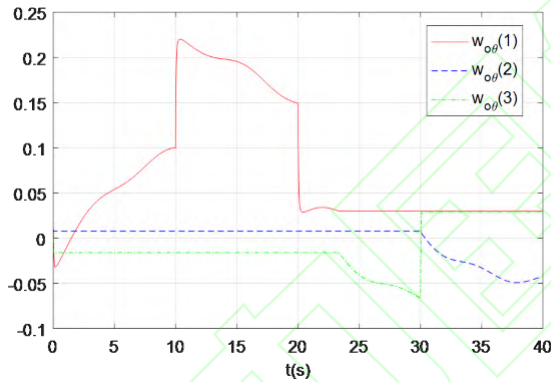


图 9 前馈跟踪控制RNNs权重更新律

Fig. 9 RNNs weighting update law

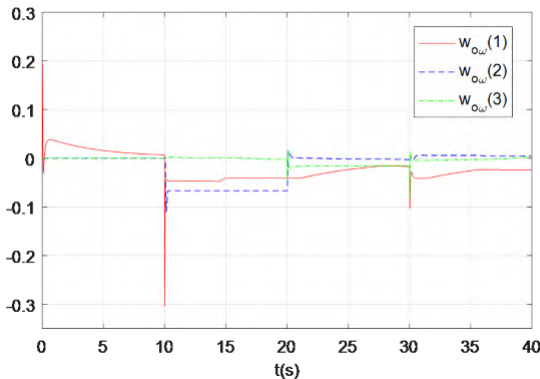


图 10 前馈跟踪控制RNNs权重更新律

Fig. 10 RNNs weighting update law

为验证所提出最优姿态控制的跟踪性能, 引入传统反步控制下再入飞行器的跟踪轨迹对比, 传统反步控制系统中选取和最优姿态控制反步部分相同的控制增益. 仿真结果如图3-11所示, 图3-图5是姿态角向量轨迹, 图6-图8是角速率向量跟踪轨迹, 图9和图10是前馈跟踪控制RNNs权重 $\hat{w}_{o\theta}$ 和 $\hat{w}_{o\omega}$ 更新律, 图11是最优跟踪控制评价网络权重 $\hat{W}_c$ 更新律. 根据姿态角响应过程可知, 存在干扰情况下再入飞行器可快速跟踪制导参考信号, 跟踪精度较高, 再入飞行器可完成飞行任务. 与传统反步控制相比, 姿态角跟踪误差较小, 角速率跟踪轨迹的暂态性能良好. 因此, 所提出的自适应RNN最优姿态控制可提供满意的跟踪性能和鲁棒性, 保证飞行器闭环系统的稳定性, 同时实现代价函数最小化.

## 6 结论

基于反步法和最优控制理论, 提出再入飞行器的RNNs自适应最优姿态控制方案, 该控制包括前馈跟踪控制器和最优调节器两部分. 稳定性分析证明, 所提出的控制方法既保证闭环系统中所有信号都是有界的, 实现代价函数最小化, 同时再入飞行器姿态角可跟踪制导参考信号.

## 参考文献:

- [1] WANG Y, YU J, MEI Y et al. Nonlinear dynamics of fixed-trim reentry vehicles with moving-mass roll control system. *Journal of systems engineering and electronics*, 2016, 27(6): 1249–1261.
- [2] SHI Li Nan, ZHANG Ran, LI Zhaoying et al. Inverse dynamic attitude control for re-entry vehicle based on output redefinition. *Journal of Beijing University of Aeronautics and Astronautics*, 2014, 40(9): 1291–1298.  
(史丽楠, 张冉, 李昭莹等. 基于输出重定义的再入飞行器动态逆姿态控制. 北京航空航天大学学报, 2014, 40(9): 1291–1298.)
- [3] WANG L, SHENG Y, and LIU X. High-order sliding mode attitude controller design for reentry flight. *Journal of systems engineering and electronics*, 2014, 25(5): 848–858.
- [4] TIAN B L, FAN W, SU R et al. Real-Time Trajectory and Attitude Coordination Control for Reusable Launch Vehicle in Reentry Phase. *IEEE transactions on industrial electronics*, 2015, 62(3): 1639–1650.

- [5] LI Zixing, LI Gaofeng. Moving centroid reentry vehicle modeling and active disturbance rejection roll control. *Acta aeronautica et astronautica sinica*, 2012, 33(11): 2121–2129.  
(李自行, 李高风. 移动质心再入飞行器建模及自抗扰滚动控制. 航空学报, 2012, 33(11): 2121–2129.)
- [6] ZHANG Huiping, YU Yue, WANG Honglun. Research on ADRC-Based Trajectory Linearization Control of Hypersonic Reentry Vehicle. *Transactions of Beijing Institute of Technology*, 2019, 39(8): 852–858.  
(张惠平, 余跃, 王宏伦. 基于自抗扰的高超再入飞行器轨迹线性化控制技术. 北京理工大学学报, 2019, 39(8): 852–858.)
- [7] ZONG Q, WANG F, TIAN B L. Nonlinear adaptive-filter backstepping flight control for reentry vehicle with input constraint and external disturbances. *Proceedings of the Institution of Mechanical Engineers*, 2014, 228(6): 889–907.
- [8] LIU Weixing, GENG Yunhai, WU Baolin. Inertia-free fault tolerant attitude tracking for a lifting reentry vehicle. *Journal of Astronautics*, 2018, 39(6): 683–689.  
(刘伟星, 耿云海, 吴宝林. 转动惯量未知的再入飞行器姿态容错控制. 宇航学报, 2018, 39(6): 683–689.)
- [9] SUN K K, LI Y M, TONG S C. Fuzzy Adaptive Output Feedback Optimal Control Design for Strict-Feedback Nonlinear Systems. *IEEE transactions on systems, man, and cybernetics: systems*, 2017, 47(1): 33–44.
- [10] HASSAN Z, TRAVIS D, SARANGAPANI J. Optimal Control of Nonlinear Continuous-Time Systems in Strict-Feedback Form. *IEEE transaction on neural networks and learning systems*, 2015, 26(10): 2535–2549.
- [11] TONG S C, SUN K K, SUI S. Observer-Based Adaptive Fuzzy Decentralized Optimal Control Design for Strict-Feedback Nonlinear Large-Scale Systems. *IEEE transaction on fuzzy systems*, 2018, 26(2): 569–584.
- [12] MA G F, CHEN C, LYU Y Y et al. Adaptive Backstep-ping-Based Neural Network Control for Hypersonic Reentry Vehicle with Input Constraints. *IEEE ACCESS*, 2018, 6: 1954–1966.
- [13] MAO Q, DOU L Q, YANG Z et al. Fuzzy Disturbance Observer-Based Adaptive Sliding Mode Control for Reusable Launch Vehicles with Aeroservoelastic Characteristic. *IEEE transactions on industrial informatics*, 2020, 16(2): 1214–1223.
- [14] FEI J T, LU C. Adaptive Sliding Mode Control of Dynamic Systems Using Double Loop Recurrent Neural Network Structure. *IEEE transactions of neural networks and learning system*, 2018, 29(4): 1275–1286.
- [15] SUN J L, LIU C S. Distributed Fuzzy Adaptive Back-stepping Optimal Control for Nonlinear Multi-missile Guidance Systems with Input Saturation. *IEEE transactions on fuzzy systems*, 2019, 27(3): 447–461.
- [16] SUN K, SUI S, TONG S. Fuzzy adaptive decentralized optimal control for strict feedback nonlinear large-scale systems. *IEEE transaction on cybernetics*, 2018, 48(4): 1326–1339.

#### 作者简介:

**吉月辉** 讲师, 目前研究方向为高超声速飞行器控制、故障诊断与容错控制, E-mail: jiyuehuitju@126.com;

**周海亮** 高级工程师, 目前研究方向为无人机控制, E-mail: gaoge2004@163.com;

**车适行** 硕士研究生, 目前研究方向为输出调节理论与应用, E-mail: cheshixing@126.com;

**高强** 教授, 硕士生导师, 目前研究方向为非线性控制、工业过程的故障诊断, E-mail: gaoqiang@tjut.edu.cn.