



(12) 发明专利

(10) 授权公告号 CN 109696830 B

(45) 授权公告日 2021.12.03

(21) 申请号 201910098125.5	CN 103645736 A, 2014.03.19
(22) 申请日 2019.01.31	CN 104763694 A, 2015.07.08
(65) 同一申请的已公布的文献号 申请公布号 CN 109696830 A	CN 108919640 A, 2018.11.30
(43) 申请公布日 2019.04.30	WO 2018057828 A2, 2018.03.29
(73) 专利权人 天津大学 地址 300072 天津市南开区卫津路92号	US 2013325774 A1, 2013.12.05
(72) 发明人 鲜斌 张浩楠 张旭	CN 108803321 A, 2018.11.13
(74) 专利代理机构 天津市北洋有限责任专利代 理事务所 12201 代理人 刘国威	CN 107765553 A, 2018.03.06
(51) Int. Cl. G05B 13/04 (2006.01)	CN 108803349 A, 2018.11.13
(56) 对比文件	CN 106227043 A, 2016.12.14
CN 107357166 A, 2017.11.17	WO 2014047142 A1, 2014.03.27
CN 104932514 A, 2015.09.23	DE 202017106132 U1, 2017.12.21
CN 104932267 A, 2015.09.23	US 2013185437 A1, 2013.07.18
CN 109143855 A, 2019.01.04	WO 2017136613 A1, 2017.08.10
CN 106595671 A, 2017.04.26	US 9082079 B1, 2015.07.14
CN 107885086 A, 2018.04.06	鲜斌等. 基于神经网络前馈的无人直升机非 线性鲁棒控制设计.《传感器与微系统》.2018,第 37卷(第4期),104-107. (续)
CN 108008627 A, 2018.05.08	审查员 张琦
CN 109164821 A, 2019.01.08	

权利要求书2页 说明书8页 附图4页

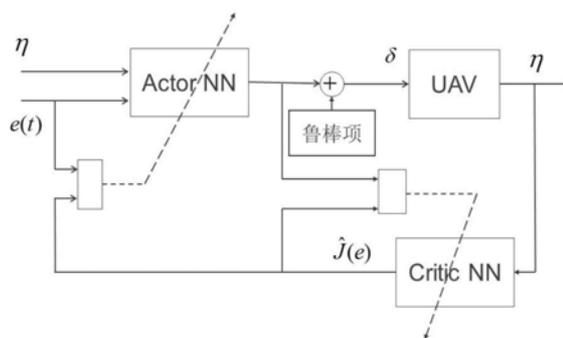
(54) 发明名称

小型无人直升机的强化学习自适应控制方法

(57) 摘要

本发明涉及小型无人直升机的非线性控制，为提出一种基于强化学习的自适应控制方法，实现在小型直升机具有系统参数不确定性和外界扰动的情况下，仍能保持飞行姿态稳定。为此，本发明采用的技术方案是，小型无人直升机的强化学习自适应控制方法，以强化学习自适应控制算法为基础，结合评价网执行网体系结构，用于小型无人直升机的姿态系统控制中，包括以下步骤：步骤1) 确定小型无人直升机的坐标系定义；步骤2) 确定小型无人直升机姿态动力学模型；步

骤3) 定义姿态角跟踪误差并整理动力学误差模型；步骤4) 控制律设计。本发明主要应用于小型无人直升机的非线性控制场合。



CN 109696830 B

[接上页]

**(56) 对比文件**

赵丹丹. 基于模型参考滑模控制的四旋翼无人飞行器设计.《中国优秀硕士学位论文全文数据库 工程科技 II 辑》.2017,C031-827.

孙景亮等. 基于自适应动态规划的导弹制导律研究综述.《自动化学报》.2017,第43卷(第7期),1101-1113.

屈秋霞等. 针对时变轨迹的非线性仿射系统的鲁棒近似最优跟踪控制.《控制理论与应用》.2016,第33卷(第1期),77-84.

Fenza, Giuseppe等. Building Adaptive

Tutoring Model using Artificial Neural Networks and Reinforcement Learning.《17th IEEE International Conference on Advanced Learning Technologies》.2017,

Imanberdiyev, N.等. Autonomous navigation of UAV by using real-time model-based reinforcement learning.《2016 14th International Conference on Control, Automation, Robotics and Vision》.2016,

鲜斌等. 小型无人直升机姿态非线性鲁棒控制设计.《控制理论与应用》.2014,第31卷(第4期),409-416.

1. 一种小型无人直升机的强化学习自适应控制方法,其特征是,以强化学习自适应控制算法为基础,结合评价网执行网体系结构,用于小型无直升人的姿态系统控制中,包括以下步骤:

步骤1) 确定小型无人直升机的坐标系定义;

小型无人直升机坐标系定义主要涉及两个坐标系,惯性坐标系  $\{I\} = \{0_I, x_I, y_I, z_I\}$  和机体坐标系  $\{B\} = \{0_B, x_B, y_B, z_B\}$ , 其中  $0_i$  ( $i = I, B$ ) 表示坐标系原点,  $x_i, y_i, z_i$  ( $i = I, B$ ) 分别对应坐标系三个主轴方向的单位矢量,各坐标系的定义均遵循右手定则,同时定义直升机姿态角在坐标系  $\{I\}$  下表示为  $\eta = [\phi, \theta, \psi]^T$ ,  $\phi, \theta, \psi$  分别对应滚转角、俯仰角和偏航角,目标轨迹姿态角在坐标系  $\{I\}$  下表示为  $\eta_d = [\phi_d, \theta_d, \psi_d]^T$ ,  $\phi_d, \theta_d, \psi_d$  分别对应目标旋转角、目标俯仰角和目标偏航角;

步骤2) 确定小型无人直升机姿态动力学模型;

通过分析小型无人直升机作用原理,用拉格朗日方程来描述其姿态动力学模型为:

$$M(\eta)\ddot{\eta} + C(\eta, \dot{\eta})\dot{\eta} + G(\eta) + d = S^{-T}(AD\delta(t) + B) \quad (1)$$

其中  $M(\eta)$  代表惯性矩阵,  $C(\eta, \dot{\eta})$  代表科氏力矩阵,  $G(\eta)$  为重力力矩向量,  $d$  代表未知扰动向量,  $S$  代表角速度变换矩阵,  $A, B$  代表旋翼动力学相关矩阵,  $D$  代表旋翼挥舞角动力学相关矩阵,  $\delta(t) = [\delta_{lat}(t) \ \delta_{lon}(t) \ \delta_{ped}(t)]^T$  代表控制输入,  $\delta_{lat}(t)$  代表横向周期变距,  $\delta_{lon}(t)$  代表纵向周期变距,  $\delta_{ped}(t)$  代表尾桨总距,角速度变换矩阵  $S$  表示为:

$$S = \begin{bmatrix} 1 & \sin(\phi)\tan(\theta) & \cos(\phi)\tan(\theta) \\ 0 & \cos(\phi) & -\sin(\phi) \\ 0 & \sin(\phi)/\cos(\theta) & \cos(\phi)/\cos(\theta) \end{bmatrix} \quad (2)$$

步骤3) 定义姿态角跟踪误差并整理动力学误差模型;

定义系统姿态跟踪误差  $e_1$  及其滤波误差  $e_2$  为:

$$\begin{aligned} e_1 &= \eta - \eta_d \\ e_2 &= \dot{e}_1 + \lambda e_1 \end{aligned} \quad (3)$$

其中  $\lambda = [\lambda_1, \lambda_2, \lambda_3]^T$  为正常数阵,引入辅助矩阵  $\Omega = S^{-T}AD$ , 输入转矩量  $\tau^T = \Omega \delta(t)$ , 定义性能指标函数为:

$$J(e_1) = \int_t^{\infty} r(e_1(s), \tau^T(s)) ds \quad (4)$$

其中,  $r(e_1, \tau^T) = e_1^T Q e_1 + \tau^T R \tau$ ,  $Q, R$  为正定矩阵, 需要保证系统的性能指标函数最优, 定义哈密顿函数为如下形式:

$$H(e_1, \tau^T) = r(e_1, \tau^T) + \frac{\partial J(e_1)}{\partial e_1} \dot{e}_1 \quad (5)$$

引入评价网  $\hat{J}(e_1)$  来近似性能指标函数  $J(e_1)$ :

$$\hat{J}(e_1) = \hat{W}_1 \varphi_1(e_1) \quad (6)$$

其中,  $\hat{W}_1$  为权重向量,  $\varphi_1(e_1)$  为径向基函数, 设计权重更新律为:

$$\dot{\hat{W}}_1 = -a_1 \sigma_1 (\sigma_1^T \hat{W}_1 + e_1^T Q e_1 + \tau^T R \tau) \quad (7)$$

其中,  $a_1$  为评价网络的自适应增益,  $\sigma = \nabla \varphi_1 \dot{e}_1$ ,  $\sigma_1 = \frac{\sigma}{\sigma^T \sigma + 1}$  为辅助变量;

对  $e_2$  求一阶时间导数, 并将式 (1) 代入整理, 得到滤波误差的开环动态方程为:

$$M(\eta) \dot{e}_2 = -\frac{1}{2} \dot{M}(\eta) e_2 - e_1 + \tau - d + N(x) \quad (8)$$

其中, 辅助函数  $N(x)$ ,  $x = [\eta^T, \dot{\eta}^T, e_1^T, \dot{e}_1^T]^T$  定义为:

$$N(x) = \frac{1}{2} \dot{M}(\eta) e_2 + M(\eta) \lambda \dot{e}_1 - M(\eta) \ddot{\eta}_d - C(\eta, \dot{\eta}) \dot{\eta} - G(\eta) + e_1 + S^{-T} B \quad (9)$$

步骤4) 控制律设计;

引入执行网络  $\hat{N}(x)$  来逼近系统辅助函数  $N(x)$ , 表示为:

$$\hat{N}(x) = \hat{W}_2^T \varphi_2(x) \quad (10)$$

其中,  $\hat{W}_2$  为权重向量,  $\varphi_2(x)$  为径向基函数, 设计权重更新律为:

$$\dot{\hat{W}}_2 = -a_2 \varphi_2(x) (\hat{W}_2^T \varphi_2(x) + k_l \nabla \varphi_1^T \hat{W}_1) \quad (11)$$

其中,  $a_2$  为执行网络的自适应增益;

根据以上分析, 设计设计控制输入  $\delta(t)$  为:

$$\delta(t) = \Omega^{-1} \left( -k_r e_2 - \Gamma \operatorname{sgn}(e_2) - \hat{W}_2^T \varphi_2(x) - \frac{k_v e_2}{e_1^T e_1 + \beta} \right) \quad (12)$$

其中  $k_r$ ,  $\Gamma$ ,  $k_v$ ,  $\beta$  均为正常数阵,  $\operatorname{sgn}(\cdot)$  为标准符号函数;

以上述控制律进行小型无人直升机的姿态控制。

## 小型无人直升机的强化学习自适应控制方法

### 技术领域

[0001] 本发明涉及一种小型无人直升机的非线性控制方法,特别是涉及一种基于强化学习的小型无人直升机自适应控制方法。具体讲,涉及小型无人直升机的强化学习自适应控制方法。

### 背景技术

[0002] 无人机(Unmanned Aerial Vehicle,简称UAV),是装配了必要的数据处理单元、传感器、自动控制器以及通信设备的飞行器系统,能够在远距离遥控下或者自主控制下执行预先设定的飞行任务。无人直升机能够做到:1)可垂直起降、可定点悬停、可在空中原地转弯;2)能够近地机动飞行;3)起飞和降落无需专用的跑道;4)若直升机在空中发生故障,可以利用自身旋翼特性自转下滑,最终安全降落。其在民用领域和军事上均发挥着重要的价值,受到各国的普遍重视(图书:国防工业出版社;著者:贺天鹏,张俊,曾国奇,等;出版年月:2016;文章题目:无人直升机系统设计;页码:1-21)。然而,由于无人直升机是欠驱动且具有显著动态耦合的高度非线性系统,在控制器设计和实施过程中需要对这些动态耦合加以考虑和解决,因此其控制器的设计非常具有挑战性。

[0003] 目前,针对小型无人直升机系统的控制方法主要包括线性控制、非线性控制和智能控制等多种设计方法。线性控制方法是目前常用的无人机控制方法,刘昊等人基于简化的无人直升机线性模型设计了LQR(linear quadratic regulator,线性二次调节器)控制器。该方法通过数值仿真验证了控制器的有效性,但这种线性控制方法是基于在无人直升机工作点附近的近似线性化进行分析设计的,只能在平衡点附近使系统实现稳定,因而不可避免地限制了控制器的应用范围(期刊:IEEE Transactions on Industrial Electronics;著者:Liu H,Lu G,Zhong Y;出版年月:2013;文章题目:Robust LQR Attitude Control of a 3-DOF Laboratory Helicopter for Aggressive Maneuvers;页码:4627-4636)。

[0004] 为克服线性控制方法的上述不足,许多学者采用非线性控制方法实现无人直升机的大范围控制。Raptis I A等人将无人机动态模型划分为两个子系统,内环为姿态控制系统,外环为位置控制系统,采用反步法设计了跟踪控制器,通过数值仿真验证了无人机的镇定、跟踪和高难度飞行实验(期刊:IEEE Transactions on Control Systems Technology;著者:Raptis I A,Valavanis K P,Moreno W A;出版年月:2011;文章题目:A Novel Nonlinear Backstepping Controller Design for Helicopters Using the Rotation Matrix;页码:465-473)。针对系统存在参数不确定性,蒙志君等人设计了一种基于扩张状态观测器的自抗扰反步控制器,通过引入扩张状态观测器来实时观测由模型不确定性和外界扰动组成的全部未知扰动,同时与反步法相结合,使无人直升机在存在外界扰动下均能对控制指令实现快速和准确的反应(期刊:控制理论与应用;著者:陈旭智,蒙志君,赵文龙,等;出版年月:2015;文章标题:适应扰动的无人直升机姿态跟踪控制;页码:1534-1539)。

[0005] 近年来智能控制方法也在无人直升机控制上得到了较多的应用,Nodland D等人

利用自适应神经网络对无人机模型不确定部分进行估计,只需要较少的模型信息,就能够实现无人机的轨迹跟踪控制(期刊:IEEE Transactions on Neural Networks and Learning Systems;著者:Nodland D,Zargarzadeh H;出版年月:2013;文章题目:Neural Network-based Optimal Adaptive Output Feedback Control of a Helicopter UAV;页码:1061-1073)。Abbeel P等人使用强化学习方法对无人机进行控制,利用强化学习算法对直升机飞行的实验数据进行学习训练,最终实现无人直升机的自主飞行和特技飞行(期刊:The International Journal of Robotics Research;著者:Abbeel P,Coates A,Ng A Y;出版年月:2010;文章题目:Autonomous Helicopter Aerobatics through Apprenticeship Learning;页码:1608-1639)。

## 发明内容

[0006] 为克服现有技术的不足,本发明旨在提出一种基于强化学习的自适应控制方法,实现在小型直升机具有系统参数不确定性和外界扰动的情况下,仍能保持飞行姿态稳定。为此,本发明采用的技术方案是,小型无人直升机的强化学习自适应控制方法,以强化学习自适应控制算法为基础,结合评价网执行网体系结构,用于小型无人直升机的姿态系统控制中,包括以下步骤:

[0007] 步骤1) 确定小型无人直升机的坐标系定义;

[0008] 小型无人直升机坐标系定义主要涉及两个坐标系,惯性坐标系 $\{I\} = \{O_I, x_I, y_I, z_I\}$ 和机体坐标系 $\{B\} = \{O_B, x_B, y_B, z_B\}$ ,其中 $O_i$  ( $i=I, B$ ) 表示坐标系原点,  $x_i, y_i, z_i$  ( $i=I, B$ ) 分别对应坐标系三个主轴方向的单位矢量,各坐标系的定义均遵循右手定则,同时定义直升机姿态角在坐标系 $\{I\}$ 下表示为 $\eta = [\phi, \theta, \psi]^T$ ,  $\phi, \theta, \psi$ 分别对应滚转角、俯仰角和偏航角,目标轨迹姿态角在坐标系 $\{I\}$ 下表示为 $\eta_d = [\phi_d, \theta_d, \psi_d]^T$ ,  $\phi_d, \theta_d, \psi_d$ 分别对应目标旋转角、目标俯仰角和目标偏航角;

[0009] 步骤2) 确定小型无人直升机姿态动力学模型;

[0010] 通过分析小型无人直升机作用原理,用拉格朗日方程来描述其姿态动力学模型为:

$$[0011] \quad M(\eta)\ddot{\eta} + C(\eta, \dot{\eta})\dot{\eta} + G(\eta) + d = S^{-T}(AD\delta(t) + B) \quad (1)$$

[0012] 其中 $M(\eta)$ 代表惯性矩阵,  $C(\eta, \dot{\eta})$ 代表科氏力矩阵,  $G(\eta)$ 为重力力矩向量,  $d$ 代表未知扰动向量,  $S$ 代表角速度变换矩阵,  $A, B$ 代表旋翼动力学相关矩阵,  $D$ 代表旋翼挥舞角动力学相关矩阵,  $\delta(t) = [\delta_{lat}(t) \delta_{lon}(t) \delta_{ped}(t)]^T$ 代表控制输入,  $\delta_{lat}(t)$ 代表横向周期变距,  $\delta_{lon}(t)$ 代表纵向周期变距,  $\delta_{ped}(t)$ 代表尾桨总距,角速度变换矩阵 $S$ 表示为:

$$[0013] \quad S = \begin{bmatrix} 1 & \sin(\phi)\tan(\theta) & \cos(\phi)\tan(\theta) \\ 0 & \cos(\phi) & -\sin(\phi) \\ 0 & \sin(\phi)/\cos(\theta) & \cos(\phi)/\cos(\theta) \end{bmatrix} \quad (2)$$

[0014] 步骤3) 定义姿态角跟踪误差并整理动力学误差模型;

[0015] 定义系统姿态跟踪误差 $e_1$ 及其滤波误差 $e_2$ 为:

$$[0016] \quad \begin{aligned} e_1 &= \eta - \eta_d \\ e_2 &= \dot{e}_1 + \lambda e_1 \end{aligned} \quad (3)$$

[0017] 其中 $\lambda = [\lambda_1, \lambda_2, \lambda_3]^T$ 为正常数阵,引入辅助矩阵 $\Omega = S^{-T}AD$ ,输入转矩量 $\tau^I = \Omega \delta$

(t), 定义性能指标函数为:

$$[0018] \quad J(e_1) = \int_t^{\infty} r(e_1(s), \tau^T(s)) ds \quad (4)$$

[0019] 其中,  $r(e_1, \tau^T) = e_1^T Q e_1 + \tau^T R \tau$ ,  $Q, R$  为正定矩阵, 需要保证系统的性能指标函数最优, 定义哈密尔顿函数为如下形式:

$$[0020] \quad H(e_1, \tau^T) = r(e_1, \tau^T) + \frac{\partial J(e_1)}{\partial e_1} \dot{e}_1 \quad (5)$$

[0021] 引入评价网  $\hat{J}(e_1)$  来近似性能指标函数  $J(e_1)$ :

$$[0022] \quad \hat{J}(e_1) = \hat{W}_1^T \varphi_1(e_1) \quad (6)$$

[0023] 其中,  $\hat{W}_1$  为权重向量,  $\varphi_1(e_1)$  为径向基函数, 设计权重更新律为:

$$[0024] \quad \dot{\hat{W}}_1 = -a_1 \sigma_1 (\sigma_1^T \hat{W}_1 + e_1^T Q e_1 + \tau^T R \tau) \quad (7)$$

[0025] 其中,  $a_1$  为评价网络的自适应增益,  $\sigma = \nabla \varphi_1 \dot{e}_1$ ,  $\sigma_1 = \frac{\sigma}{\sigma^T \sigma + 1}$  为辅助变量。

[0026] 对  $e_2$  求一阶时间导数, 并将式 (1) 代入整理, 得到滤波误差的开环动态方程为:

$$[0027] \quad M(\eta) \dot{e}_2 = -\frac{1}{2} \dot{M}(\eta) e_2 - e_1 + \tau - d + N(x) \quad (8)$$

[0028] 其中, 辅助函数  $N(x)$ ,  $x = [\eta^T, \dot{\eta}^T, e_1^T, \dot{e}_1^T]^T$  定义为:

$$[0029] \quad N(x) = \frac{1}{2} \dot{M}(\eta) e_2 + M(\eta) \lambda \dot{e}_1 - M(\eta) \ddot{\eta}_d - C(\eta, \dot{\eta}) \dot{\eta} - G(\eta) + e_1 + S^{-T} B \quad (9)$$

[0030] 步骤4) 控制律设计;

[0031] 引入执行网络  $\hat{N}(x)$  来逼近系统不确定函数  $N(x)$ , 表示为:

$$[0032] \quad \hat{N}(x) = \hat{W}_2^T \varphi_2(x) \quad (10)$$

[0033] 其中,  $\hat{W}_2$  为权重向量,  $\varphi_2(x)$  为径向基函数, 设计权重更新律为:

$$[0034] \quad \dot{\hat{W}}_2 = -a_2 \varphi_2(x) (\hat{W}_2^T \varphi_2(x) + k_1 \nabla \varphi_1^T \hat{W}_1) \quad (11)$$

[0035] 其中,  $a_2$  为执行网络的自适应增益;

[0036] 根据以上分析, 设计设计控制输入  $\delta(t)$  为:

$$[0037] \quad \delta(t) = \Omega^{-1} (-k_r e_2 - \Gamma \operatorname{sgn}(e_2) - \hat{W}_2^T \varphi_2(x) - \frac{k_v e_2}{e_1^T e_1 + \beta}) \quad (12)$$

[0038] 其中  $k_r, \Gamma, k_v, \beta$  均为正常数阵,  $\operatorname{sgn}(\cdot)$  为标准符号函数;

[0039] 以上述控制律进行小型无人直升机的姿态控制。

[0040] 本发明的特点及有益效果是:

[0041] 1. 本发明采用基于强化学习的系统不确定性补偿方法, 减少了无人直升机未建模动态和模型不确定参数对系统控制性能的影响;

[0042] 2. 本发明易于实现, 可满足大部分飞行情况的需求。

[0043] 3.本发明使用了滑模鲁棒项,缩短了算法学习时间,同时提高了控制律对外界扰动的鲁棒性。

#### 附图说明:

[0044] 图1是本发明控制系统结构图;

[0045] 图2是本发明所采用的实验平台;

[0046] 图3是姿态跟踪实验时无人直升机的姿态角曲线图;

[0047] 图4是姿态跟踪实验时无人直升机的控制输入曲线图;

[0048] 图5是姿态跟踪实验时无人直升机的偏航角及跟踪误差曲线图;

[0049] 图6是姿态跟踪实验时无人直升机的评价网权值范数 $\|\hat{w}_1\|$ 曲线图;

[0050] 图7是姿态跟踪实验时无人直升机的执行网权值范数 $\|\hat{w}_2\|$ 曲线图。

#### 具体实施方式

[0051] 本发明采用的技术方案是,以强化学习自适应控制算法为基础,结合评价网执行网体系结构,用于小型无直升人的姿态系统控制中,包括以下步骤:

[0052] 步骤1) 确定小型无人直升机的坐标系定义;

[0053] 小型无人直升机坐标系定义主要涉及两个坐标系,惯性坐标系 $\{I\} = \{0_I, x_I, y_I, z_I\}$ 和机体坐标系 $\{B\} = \{0_B, x_B, y_B, z_B\}$ ,其中 $0_i$  ( $i = I, B$ ) 表示坐标系原点, $x_i, y_i, z_i$  ( $i = I, B$ ) 分别对应坐标系三个主轴方向的单位矢量,各坐标系的定义均遵循右手定则,同时定义直升机姿态角在坐标系 $\{I\}$ 下表示为 $\eta = [\phi, \theta, \psi]^T$ ,  $\phi, \theta, \psi$ 分别对应滚转角、俯仰角和偏航角,目标轨迹姿态角在坐标系 $\{I\}$ 下表示为 $\eta_d = [\phi_d, \theta_d, \psi_d]^T$ ,  $\phi_d, \theta_d, \psi_d$ 分别对应目标旋转角、目标俯仰角和目标偏航角;

[0054] 步骤2) 确定小型无人直升机姿态动力学模型;

[0055] 通过分析小型无人直升机作用原理,用拉格朗日方程来描述其姿态动力学模型为:

$$[0056] \quad M(\eta)\ddot{\eta} + C(\eta, \dot{\eta})\dot{\eta} + G(\eta) + d = S^{-T}(AD\delta(t) + B) \quad (1)$$

[0057] 其中 $M(\eta)$ 代表惯性矩阵, $C(\eta, \dot{\eta})$ 代表科氏力矩阵, $G(\eta)$ 为重力力矩向量, $d$ 代表未知扰动向量, $S$ 代表角速度变换矩阵, $A, B$ 代表旋翼动力学相关矩阵, $D$ 代表旋翼挥舞角动力学相关矩阵, $\delta(t) = [\delta_{lat}(t) \delta_{lon}(t) \delta_{ped}(t)]^T$ 代表控制输入, $\delta_{lat}(t)$ 代表横向周期变距, $\delta_{lon}(t)$ 代表纵向周期变距, $\delta_{ped}(t)$ 代表尾桨总距,角速度变换矩阵 $S$ 表示为:

$$[0058] \quad S = \begin{bmatrix} 1 & \sin(\phi)\tan(\theta) & \cos(\phi)\tan(\theta) \\ 0 & \cos(\phi) & -\sin(\phi) \\ 0 & \sin(\phi)/\cos(\theta) & \cos(\phi)/\cos(\theta) \end{bmatrix} \quad (2)$$

[0059] 步骤3) 定义姿态角跟踪误差并整理动力学误差模型;

[0060] 定义系统姿态跟踪误差 $e_1$ 及其滤波误差 $e_2$ 为:

$$[0061] \quad \begin{aligned} e_1 &= \eta - \eta_d \\ e_2 &= \dot{e}_1 + \lambda e_1 \end{aligned} \quad (3)$$

[0062] 其中  $\lambda = [\lambda_1, \lambda_2, \lambda_3]^T$  为正常数阵, 引入辅助矩阵  $\Omega = S^{-T}AD$ , 输入转矩量  $\tau^I = \Omega \delta(t)$ , 定义性能指标函数为:

$$[0063] \quad J(e_1) = \int_t^{\infty} r(e_1(s), \tau^I(s)) ds \quad (4)$$

[0064] 其中,  $r(e_1, \tau^I) = e_1^T Q e_1 + \tau^T R \tau$ ,  $Q, R$  为正定矩阵, 需要保证系统的性能指标函数最优, 定义哈密尔顿函数为如下形式:

$$[0065] \quad H(e_1, \tau^I) = r(e_1, \tau^I) + \frac{\partial J(e_1)}{\partial e_1} \dot{e}_1 \quad (5)$$

[0066] 引入评价网  $\hat{J}(e_1)$  来近似性能指标函数  $J(e_1)$ :

$$[0067] \quad \hat{J}(e_1) = \hat{W}_1^T \varphi_1(e_1) \quad (6)$$

[0068] 其中,  $\hat{W}_1$  为权重向量,  $\varphi_1(e_1)$  为径向基函数, 设计权重更新律为:

$$[0069] \quad \dot{\hat{W}}_1 = -a_1 \sigma_1 (\sigma_1^T \hat{W}_1 + e_1^T Q e_1 + \tau^T R \tau) \quad (7)$$

[0070] 其中,  $a_1$  为评价网络的自适应增益,  $\sigma = \nabla \varphi_1 \dot{e}_1$ ,  $\sigma_1 = \frac{\sigma}{\sigma^T \sigma + 1}$  为辅助变量。

[0071] 对  $e_2$  求一阶时间导数, 并将式 (1) 代入整理, 得到滤波误差的开环动态方程为:

$$[0072] \quad M(\eta) \dot{e}_2 = -\frac{1}{2} \dot{M}(\eta) e_2 - e_1 + \tau - d + N(x) \quad (8)$$

[0073] 其中, 辅助函数  $N(x)$ ,  $x = [\eta^T, \dot{\eta}^T, e_1^T, \dot{e}_1^T]^T$  定义为:

$$[0074] \quad N(x) = \frac{1}{2} \dot{M}(\eta) e_2 + M(\eta) \lambda \dot{e}_1 - M(\eta) \ddot{\eta}_d - C(\eta, \dot{\eta}) \dot{\eta} - G(\eta) + e_1 + S^{-T} B \quad (9)$$

[0075] 步骤4) 控制律设计;

[0076] 引入执行网络  $\hat{N}(x)$  来逼近系统不确定函数  $N(x)$ , 表示为:

$$[0077] \quad \hat{N}(x) = \hat{W}_2^T \varphi_2(x) \quad (10)$$

[0078] 其中,  $\hat{W}_2$  为权重向量,  $\varphi_2(x)$  为径向基函数, 设计权重更新律为:

$$[0079] \quad \dot{\hat{W}}_2 = -a_2 \varphi_2(x) (\hat{W}_2^T \varphi_2(x) + k_v \nabla \varphi_1^T \hat{W}_1) \quad (11)$$

[0080] 其中,  $a_2$  为执行网络的自适应增益。

[0081] 根据以上分析, 设计设计控制输入  $\delta(t)$  为:

$$[0082] \quad \delta(t) = \Omega^{-1} (-k_r e_2 - \Gamma \operatorname{sgn}(e_2) - \hat{W}_2^T \varphi_2(x) - \frac{k_v e_2}{e_1^T e_1 + \beta}) \quad (12)$$

[0083] 其中  $k_r, \Gamma, k_v, \beta$  均为正常数阵,  $\operatorname{sgn}(\cdot)$  为标准符号函数;

[0084] 以上述控制律进行小型无人直升机的姿态控制。

[0085] 下面结合实施例和附图对本发明做出详细说明。

[0086] 本发明旨在提供一种基于强化学习的自适应控制方法, 实现在小型直升机具有系统参数不确定性和外界扰动的情况下, 仍能保持飞行姿态稳定。

[0087] 本发明的小型无人直升机强化学习自适应控制采用的技术方案是,以强化学习自适应控制算法为基础,结合评价网执行网体系结构,用于小型无人直升机的姿态系统控制中,包括以下步骤:

[0088] 步骤1) 确定小型无人直升机的坐标系定义;

[0089] 小型无人直升机坐标系定义主要涉及两个坐标系,惯性坐标系 $\{I\} = \{0_I, x_I, y_I, z_I\}$ 和机体坐标系 $\{B\} = \{0_B, x_B, y_B, z_B\}$ ,其中 $0_i$  ( $i=I, B$ ) 表示坐标系原点,  $x_i, y_i, z_i$  ( $i=I, B$ ) 分别对应坐标系三个主轴方向的单位矢量,各坐标系的定义均遵循右手定则,同时定义直升机姿态角在坐标系 $\{I\}$ 下表示为 $\eta = [\phi, \theta, \psi]^T$ ,  $\phi, \theta, \psi$ 分别对应滚转角、俯仰角和偏航角,目标轨迹姿态角在坐标系 $\{I\}$ 下表示为 $\eta_d = [\phi_d, \theta_d, \psi_d]^T$ ,  $\phi_d, \theta_d, \psi_d$ 分别对应目标旋转角、目标俯仰角和目标偏航角;

[0090] 步骤2) 确定小型无人直升机姿态动力学模型;

[0091] 通过分析小型无人直升机作用原理,用拉格朗日方程来描述其姿态动力学模型为:

$$[0092] \quad M(\eta)\ddot{\eta} + C(\eta, \dot{\eta})\dot{\eta} + G(\eta) + d = S^{-T}(AD\delta(t) + B) \quad (1)$$

[0093] 其中 $M(\eta)$ 代表惯性矩阵,  $C(\eta, \dot{\eta})$ 代表科氏力矩阵,  $G(\eta)$ 为重力力矩向量,  $d$ 代表未知扰动向量,  $S$ 代表角速度变换矩阵,  $A, B$ 代表旋翼动力学相关矩阵,  $D$ 代表旋翼挥舞角动力学相关矩阵,  $\delta(t) = [\delta_{lat}(t) \delta_{lon}(t) \delta_{ped}(t)]^T$ 代表控制输入,  $\delta_{lat}(t)$ 代表横向周期变距,  $\delta_{lon}(t)$ 代表纵向周期变距,  $\delta_{ped}(t)$ 代表尾桨总距,角速度变换矩阵 $S$ 表示为:

$$[0094] \quad S = \begin{bmatrix} 1 & \sin(\phi)\tan(\theta) & \cos(\phi)\tan(\theta) \\ 0 & \cos(\phi) & -\sin(\phi) \\ 0 & \sin(\phi)/\cos(\theta) & \cos(\phi)/\cos(\theta) \end{bmatrix} \quad (2)$$

[0095] 步骤3) 定义姿态角跟踪误差并整理动力学误差模型;

[0096] 定义系统姿态跟踪误差 $e_1$ 及其滤波误差 $e_2$ 为:

$$[0097] \quad \begin{aligned} e_1 &= \eta - \eta_d \\ e_2 &= \dot{e}_1 + \lambda e_1 \end{aligned} \quad (3)$$

[0098] 其中 $\lambda = [\lambda_1, \lambda_2, \lambda_3]^T$ 为正常数阵,引入辅助矩阵 $\Omega = S^{-T}AD$ ,输入转矩量 $\tau^I = \Omega \delta(t)$ ,定义性能指标函数为:

$$[0099] \quad J(e_1) = \int_t^{\infty} r(e_1(s), \tau^I(s)) ds \quad (4)$$

[0100] 其中,  $r(e_1, \tau^I) = e_1^T Q e_1 + \tau^T R \tau$ ,  $Q, R$ 为正定矩阵,需要保证系统的性能指标函数最优,定义哈密顿函数为如下形式:

$$[0101] \quad H(e_1, \tau^I) = r(e_1, \tau^I) + \frac{\partial J(e_1)}{\partial e_1} \dot{e}_1 \quad (5)$$

[0102] 引入评价网 $\hat{J}(e_1)$ 来近似性能指标函数 $J(e_1)$ :

$$[0103] \quad \hat{J}(e_1) = \hat{W}_1 \varphi_1(e_1) \quad (6)$$

[0104] 其中,  $\hat{W}_1$ 为权重向量,  $\varphi_1(e_1)$ 为径向基函数,设计权重更新律为:

[0105]  $\dot{\hat{W}}_1 = -a_1 \sigma_1 (\sigma_1^T \hat{W}_1 + e_1^T Q e_1 + \tau^T R \tau)$  (7)

[0106] 其中,  $a_1$  为评价网络的自适应增益,  $\sigma = \nabla \varphi_1 \dot{e}_1$ ,  $\sigma_1 = \frac{\sigma}{\sigma^T \sigma + 1}$  为辅助变量。

[0107] 对  $e_2$  求一阶时间导数, 并将式 (1) 代入整理, 得到滤波误差的开环动态方程为:

[0108]  $M(\eta) \dot{e}_2 = -\frac{1}{2} \dot{M}(\eta) e_2 - e_1 + \tau - d + N(x)$  (8)

[0109] 其中, 辅助函数  $N(x)$ ,  $x = [\eta^T, \dot{\eta}^T, e_1^T, \dot{e}_1^T]^T$  定义为:

[0110] 
$$N(x) = \frac{1}{2} \dot{M}(\eta) e_2 + M(\eta) \lambda \dot{e}_1 - M(\eta) \ddot{\eta}_d - C(\eta, \dot{\eta}) \dot{\eta} - G(\eta) + e_1 + S^{-T} B$$
 (9)

[0111] 步骤4) 控制律设计;

[0112] 引入执行网络  $\hat{N}(x)$  来逼近系统不确定函数  $N(x)$ , 表示为:

[0113]  $\hat{N}(x) = \hat{W}_2^T \varphi_2(x)$  (10)

[0114] 其中,  $\hat{W}_2$  为权重向量,  $\varphi_2(x)$  为径向基函数, 设计权重更新律为:

[0115]  $\dot{\hat{W}}_2 = -a_2 \varphi_2(x) (W_2^T \varphi_2(x) + k_1 \nabla \varphi_1^T \hat{W}_1)$  (11)

[0116] 其中,  $a_2$  为执行网络的自适应增益。

[0117] 根据以上分析, 设计设计控制输入  $\delta(t)$  为:

[0118]  $\delta(t) = \Omega^{-1} (-k_r e_2 - \Gamma \text{sgn}(e_2) - \hat{W}_2 \varphi_2(x) - \frac{k_v e_2}{e_1^T e_1 + \beta})$  (12)

[0119] 其中  $k_r$ ,  $\Gamma$ ,  $k_v$ ,  $\beta$  均为正常数阵,  $\text{sgn}(\cdot)$  为标准符号函数;

[0120] 以上述控制律进行小型无人直升机的姿态控制。

[0121] 控制结构框图如图1所示。

[0122] 下面给出具体的实例:

[0123] 一、半实物仿真平台介绍

[0124] 为验证本文设计控制律的有效性, 使用本研究组设计开发的无人直升机飞行实验平台, 进行了实时镇定飞行实验。实验平台如图2所示, 该实验平台以基于MATLAB-RTW工具箱的xPC作为实时仿真环境, 采用自主设计的以ARM Cortex-M3为内核的惯性导航单元作为传感器, 该机载传感器最高更新频率为500Hz, 可提供三轴角速度及三轴姿态角信息, 其中滚转角和俯仰角测量精度为  $\pm 0.2^\circ$ , 偏航角的测量精度为  $\pm 0.5^\circ$ 。

[0125] 二、姿态跟踪飞行实验

[0126] 为了验证本发明中控制器的有效性及实用性, 利用本研究组自主设计开发的无人直升机姿态飞行实验平台, 进行了实时的偏航跟踪飞行实验。设置参考轨迹为:

[0127]  $\eta_d = \left[ 0 \quad 0 \quad -\frac{\pi}{6} \sin\left(\frac{\pi}{10} t\right) \right]^T$ 。

[0128] 控制律中各参数确定为:  $k_r = \text{diag} \{ [1.5, 1.5, 1.2]^T \}$ ,  $\lambda = \text{diag} \{ [0.57, 0.56, 0.54]^T \}$ ,  $\Gamma = \text{diag} \{ [0.02, 0.002, 0.003]^T \}$ ,  $k_v = \text{diag} \{ [2.4, 2.2, 0.5]^T \}$ ,  $\beta = \text{diag} \{ [2.4, 2.2, 0.5]^T \}$ 。

评价网和执行网参数确定为 $a_1 = \text{diag}\{[0.03, 0.04, 0.04]^T\}$ ,  $a_2 = \text{diag}\{[0.03, 0.04, 0.01]^T\}$ ,  $k_1 = \text{diag}\{[0.54, 0.56, 0.9]^T\}$ 。

[0129] 姿态跟踪控制实验情况如图3到图7所示,强化学习自适应控制取得了较好的控制效果。通过图3可知,在偏航跟踪实验中,滚转和俯仰角的控制误差保持在 $\pm 2^\circ$ 以内,偏航角控制误差保持在 $\pm 2.2^\circ$ 以内。图4是正则化后的控制输入变化曲线,图6和图7分别为评价网和执行网权重范数收敛曲线,均保持在合理的范围内,验证了本文所设计控制器的合理性。

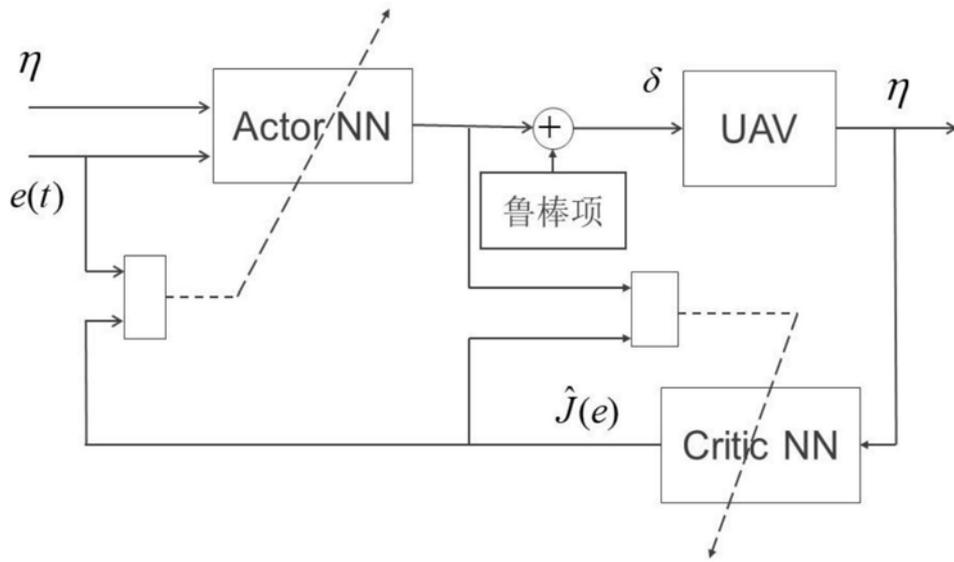


图1

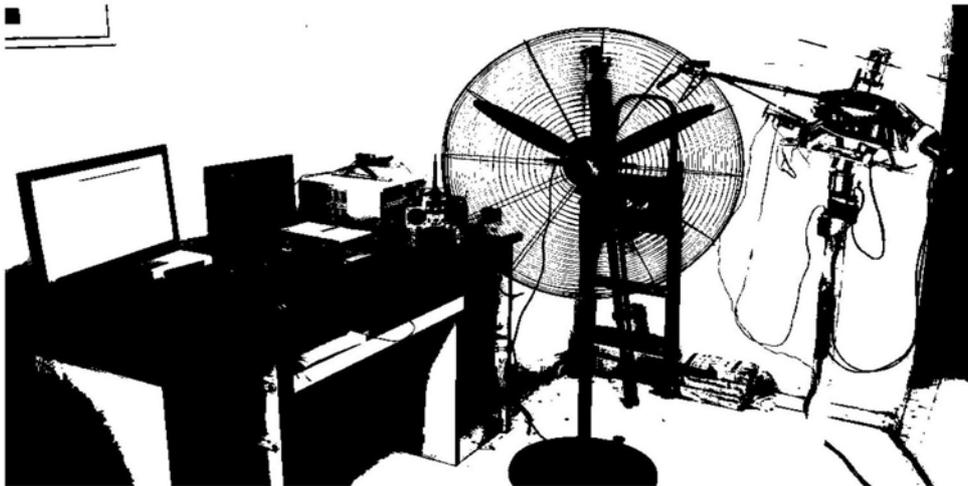


图2

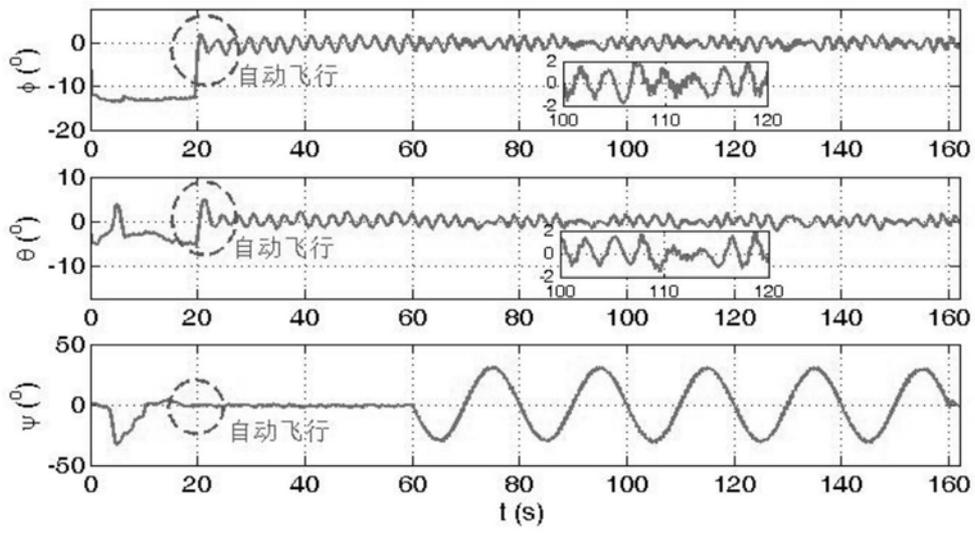


图3

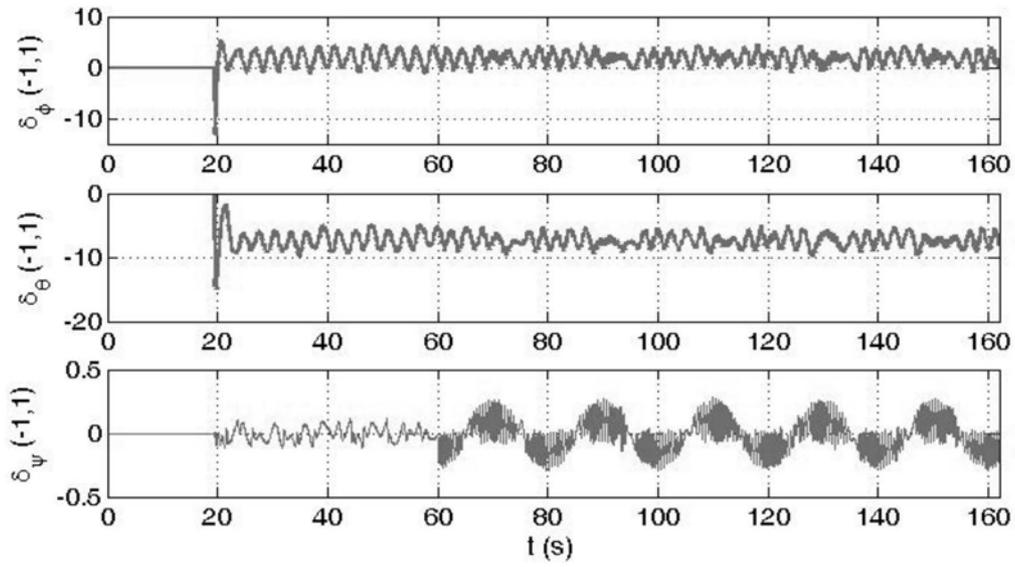


图4

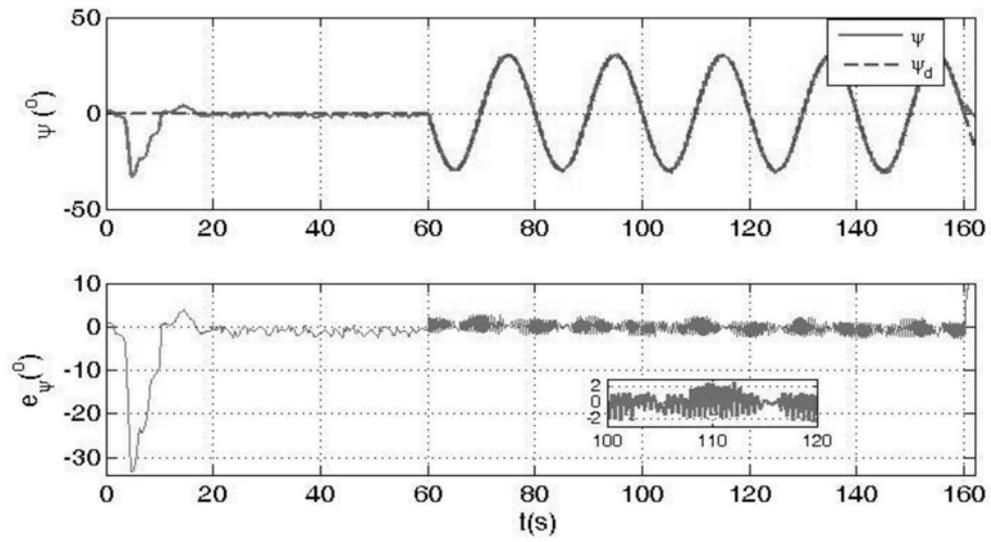


图5

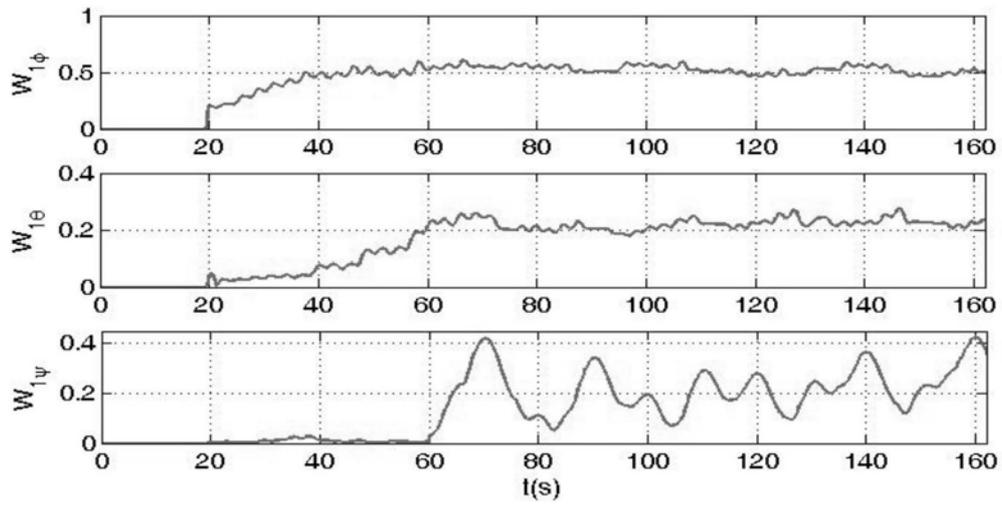


图6

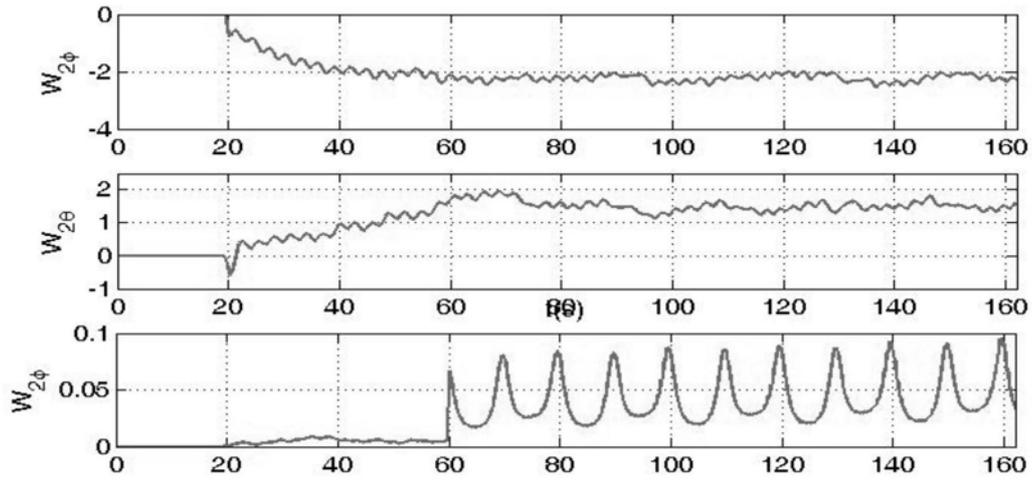


图7