基于确定学习的机器人任务空间自适应神经网络控制

吴玉香 [1] 王 聪 [1]

摘 要 针对产生回归轨迹的连续非线性动态系统, 确定学习可实现未知闭环系统动态的局部准确逼近. 基于确定学习理论, 本文使用径向基函数 (Radial basis function, RBF) 神经网络为机器人任务空间跟踪控制设计了一种新的自适应神经网络控制算法, 不仅实现了闭环系统所有信号的最终一致有界, 而且在稳定的控制过程中, 沿着回归跟踪轨迹实现了部分神经网络权值收敛到最优值以及未知闭环系统动态的局部准确逼近. 学过的知识以时不变且空间分布的方式表达、以常值神经网络权值的方式存储, 可以用来改进系统的控制性能, 也可以应用到后续相同或相似的控制任务中, 节约时间和能量. 最后, 用仿真说明了所设计控制算法的正确性和有效性.

关键词 确定学习, 径向基函数神经网络, 自适应神经网络控制, 机器人

引用格式 吴玉香, 王聪. 基于确定学习的机器人任务空间自适应神经网络控制. 自动化学报, 2013, **39**(6): 806−815

DOI 10.3724/SP.J.1004.2013.00806

Deterministic Learning Based Adaptive Network Control of Robot in Task Space

WU Yu-Xiang[1] WANG Cong[1]

Abstract Deterministic learning can achieve locally-accurate approximation of the unknown closed-loop system dynamics while attempting to control a class of nonlinear systems producing recurrent trajectories. Based on deterministic learning, an adaptive neural control algorithm is proposed for unknown robots in task space using radial basis function (RBF) networks. The designed adaptive neural controller can not only guarantee all signals in the closed-loop system uniformly ultimately bounded, but also achieve convergence of partial network weights to their optimal values. It can also learn the unknown closed-loop system dynamics in a stable control process along recurrent tracking orbits. The learned knowledge stored as constant network weights can be reused in a same or similar control task to improve the control performance and to save time and energy. Simulation results demonstrate the effectiveness of the proposed approach.

Key words Deterministic learning, Radial basis function (RBF) networks, adaptive neural network control, robot

Citation Wu Yu-Xiang, Wang Cong. Deterministic learning based adaptive network control of robot in task space. *Acta Automatica Sinica*, 2013, **39**(6): 806−815

自适应神经网络已经在机器人控制系统中得到了广泛的研究, 但大部分自适应神经网络控制器的设计都是针对关节空间进行的[1−4]. 文献 [1] 引进滤波器跟踪误差, 根据机器人动态设计自适应神经网络控制器, 保证了闭环系统所有信号的最终一致有界. 文献 [2] 使用线性观测器, 为未知机器人系统设计了自适应神经网络控制器, 保证了系统跟踪误差、观测器误差的最终一致有界以及神经网络权值的有界. 文献 [3] 针对柔性关节机器人操作手的轨迹跟踪问题, 提出了一种基于奇异摄动理论的机器人神经网络控制方法, 在一般框架下证明了系统跟踪误差最终一致有界. 文献 [4] 为水下机器人设计了自适应神经网络控制器, 证明了控制系统的跟踪误差一致稳定有界. 而机器人在实际工作中期望的位置或轨迹往往是定义在任务空间的. 关节空间控制器的主要限制在于要通过解逆运动学方程才能获得关节空间的期望位置或轨迹. 直接在任务空间中施以控制行为, 将会得到优于在关节空间内控制的性能. 基于神经网络的一致逼近能力, 几种任务空间的神经网络控制器被提出[5−8]. 文献 [5] 设计的自适应神经网络控制器既不需要计算逆动态模型也不需要计算雅可比矩阵的逆, 实现了渐近跟踪. 文献 [6] 针对具有不确定运动学的机器人, 设计的自适应神经网络控制器实现了跟踪误差的有界, 且界可以任意小. 文献 [7] 基于类 Lyapunov 方法, 设计的自适应神经网络控制器在有限确定时间内实现了稳定跟踪. 文献 [8] 针对自由漂浮空间机器人, 在任务空间内设计了一个基于径向基函数 (Radial basis function, RBF) 神经网络的鲁棒智能控制器, 实现了轨迹跟踪误差的全局渐近稳定.

以上讨论的这些神经网络控制方法, 无论是关节空间的还是任务空间的, 都是基于神经网络的通

收稿日期 2011-12-16 录用日期 2012-10-25
Manuscript received December 16, 2011; accepted October 25, 2012
国家自然科学基金 (60934001, 61225014, 61075082) 资助
Supported by National Natural Science Foundation of China (60934001, 61225014, 61075082)
本文责任编委 孙长银
Recommended by Associate Editor SUN Chang-Yin
1. 华南理工大学自动化科学与工程学院 广州 510640
1. College of Automation Science and Engineering, South China University of Technology, Guangzhou 510640

用逼近性而展开的, 但对于神经网络能否学习到系统未知动态并没有进行深入探讨. 持续激励条件与神经网络估计参数 (即神经网络的权值) 的收敛性息息相关. 而神经网络隐含层输出的函数递归矩阵的持续激励条件一般很难满足, 已有的许多研究放松了对持续激励条件的要求, 神经网络权值的收敛性也就不能得到保证. 在这种情况下, 神经网络的通用逼近性并没有得到保证, 以至对于经历过多次的同一控制任务, 在每次执行时, 神经网络仍需要重新进行冗余而繁琐的训练, 浪费时间和能量, 此时, 神经网络的学习能力是相当有限的. 相比于已有的神经网络理论成果, 新兴的确定学习理论对于神经网络的学习能力有着更深入的研究.

确定学习[9−10] 运用自适应控制和动力学系统的概念与方法, 研究未知动态环境下的知识获取、表达、存储和利用等问题. 针对产生回归轨迹的连续非线性动态系统, 借助该理论建立的自适应神经网络控制器在稳定的控制过程中能够对未知的闭环系统进行动态学习. 目前, 确定学习的研究和应用还处于新兴阶段, 虽然取得了一些研究成果[9−15], 但还有许多问题亟待研究与解决. 文献 [10] 研究了单位控制增益的简单非线性系统的确定学习. 运用反步设计方法, 文献 [11] 研究了更一般形式非线性系统的确定学习. 文献 [12] 研究了一类严格反馈非线性系统的确定学习. 文献 [13] 利用确定学习理论实现了微小振动故障的快速检测. 文献 [14] 研究了持续激励与确定学习速度和精度之间的关系. 文献 [15] 研究了噪声环境下的确定学习问题. 文献 [16] 基于确定学习, 研究了离散动态系统的建模与控制问题. 确定学习的基本要素包括: 1) 使用 RBF 神经网络; 2) 对于回归状态轨迹满足部分持续激励条件; 3) 在回归轨迹的邻域内实现对未知非线性闭环系统动态的局部准确逼近 (局部准确建模); 4) 所学的知识以时不变且空间分布的方式表达、以常值神经网络权值的方式存储, 并可在动态环境下实现闭环神经网络控制或者动态模式的快速识别. 回归轨迹代表了一大类从非线性系统产生的轨迹, 不仅包括周期轨迹, 还包括拟周期轨迹、概周期轨迹甚至部分混沌轨迹[17].

机器人特别是智能机器人已成为学术界和工业界共同研究的热点问题, 智能机器人的研究仍是以模拟、复制人类智能为最高目标. 如何使机器人具有类似人类的学习、记忆以及再利用等能力, 是目前乃至今后一段时间内的热门课题. 确定学习能够实现机器人未知闭环系统动态的 “学习”、“表达与存储” 以及 “再利用”, 为具有类似人类智能的机器人提供了一种简单而有效的解决方法. 确定学习在机器人领域的应用仅取得了一些初步的研究成果[18−20]. 文献 [18] 利用反步法实现了电驱动机械臂的跟踪控制与确定学习; 文献 [19] 利用确定学习实现了人的步态识别; 文献 [20] 利用反步法实现了机器人的鲁棒自适应神经网络控制与确定学习.

本文基于确定学习理论提出了一种新的机器人任务空间控制与学习策略, 不仅实现了闭环系统所有信号的最终一致有界, 而且在稳定的控制过程中, 实现了部分神经网络权值收敛到最优值以及未知闭环系统动态的局部准确逼近. 最后, 用两连杆平面机器人证明了所提控制方法的正确性和有效性.

1 预备知识

1.1 局部 RBF 神经网络

RBF 神经网络可以用下面的形式来描述[21]:

$$f_{nn}(\boldsymbol{X}) = \sum_{i=1}^{N} w_i s_i(\boldsymbol{X}) = \boldsymbol{W}^{\mathrm{T}} \boldsymbol{S}(\boldsymbol{X}) \tag{1}$$

其中, $\boldsymbol{X} \in \Omega_X \subset \mathbf{R}^q$ 为输入向量, $\boldsymbol{W} = [w_1, \cdots, w_N]^{\mathrm{T}} \in \mathbf{R}^N$ 为神经网络的权值向量, $N > 1$ 为神经网络的节点数, $\boldsymbol{S}(\boldsymbol{X}) = [s_1(\|\boldsymbol{X} - \boldsymbol{\xi}_1\|), \cdots, s_N(\|\boldsymbol{X} - \boldsymbol{\xi}_N\|)]^{\mathrm{T}}$, 高斯函数 $s_i(\cdot) = \exp[-\|\boldsymbol{X} - \boldsymbol{\xi}_i\|^2 / \eta_i^2]$ 为最普遍使用的径向基函数, 其中, $\boldsymbol{\xi}_i$ 为中心点, η_i 为中心宽度, 当 $\|\boldsymbol{X}\| \to \infty$ 时, 有 $s_i(\|\boldsymbol{X} - \boldsymbol{\xi}_i\|) \to 0$.

依据文献 [21] 的结果, 对于一个径向基函数网络 (1), 在有足够多的节点, 有适当构建的节点中心及中心宽度的情况下, 它能够无限逼近任意的连续函数 $f(\boldsymbol{X}) : \Omega_X \to \mathbf{R}$, 紧集 $\Omega_X \subset \mathbf{R}^q$, 即

$$f(\boldsymbol{X}) = \boldsymbol{W}^{*\mathrm{T}} \boldsymbol{S}(\boldsymbol{X}) + \varepsilon(\boldsymbol{X}), \quad \forall \boldsymbol{X} \in \Omega_X \tag{2}$$

其中, $\boldsymbol{W}^*$ 为理想的神经网络权值, $\varepsilon(\boldsymbol{X})$ 为逼近误差, 满足 $|\varepsilon(\boldsymbol{X})| < \varepsilon^*$ ($\varepsilon^* > 0$ 是很小正常数).

根据 RBF 神经网络的局部特性, 对于在紧集 Ω_X 中任意有界轨迹 $\boldsymbol{X}_\xi(t)$, $f(\boldsymbol{X})$ 能够被有限数目的、位于轨迹 $\boldsymbol{X}_\xi(t)$ 局部区域内的神经元逼近, 即

$$f(\boldsymbol{X}) = \boldsymbol{W}_\xi^{*\mathrm{T}} \boldsymbol{S}_\xi(\boldsymbol{X}) + \varepsilon_\xi(\boldsymbol{X}) \tag{3}$$

其中, $\boldsymbol{S}_\xi(\boldsymbol{X}) = [s_{j1}(\boldsymbol{X}), \cdots, s_{j\xi}(\boldsymbol{X})]^{\mathrm{T}} \in \mathbf{R}^{N_\xi}$, $N_\xi < N$, $|s_{ji}| > \mu$ ($ji = j1, \cdots, j\xi$), $\mu > 0$ 是很小的正常数, $\boldsymbol{W}_\xi^* = [w_{j1}^*, \cdots, w_{j\xi}^*]^{\mathrm{T}}$, $\varepsilon_\xi(\boldsymbol{X})$ 是逼近误差, $||\varepsilon_\xi| - |\varepsilon||$ 的值很小.

基于已有的 RBF 神经网络 PE (Persistent excitation) 属性的结论[22−23], 文献 [9−10] 已经证明对于中心构建在规则区域内的局部 RBF 神经网络 $\boldsymbol{W}^{\mathrm{T}} \boldsymbol{S}(\boldsymbol{X})$, 几乎任何回归轨迹 $\boldsymbol{X}(t)$ 都能使其递归子矢量 $\boldsymbol{S}_\xi(\boldsymbol{X})$ 满足 PE 条件.

1.2 一类线性时变系统的稳定性分析

本部分主要介绍一类线性时变 (Linear time-varying, LTV) 系统的指数稳定性, 该类线性时变系统的指数稳定性与非线性系统的确定学习息息相关.

考虑如下通用的 LTV 系统:

$$\begin{bmatrix} \dot{\boldsymbol{e}} \\ \dot{\boldsymbol{\eta}} \end{bmatrix} = \begin{bmatrix} A(t) & B^{\mathrm{T}}(t) \\ -C(t) & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{e} \\ \boldsymbol{\eta} \end{bmatrix} \tag{4}$$

其中, $\boldsymbol{e} \in \mathbf{R}^n$ 是状态跟踪误差, $\boldsymbol{\eta} \in \mathbf{R}^p$ 是参数估计误差, $A(t) \in \mathbf{R}^{n\times n}$, $B(t) \in \mathbf{R}^{p\times n}$, $C(t) \in \mathbf{R}^{p\times n}$. 对于单位控制增益的简单非线性系统的确定学习, 在式 (4) 中, $A(t)$ 是时不变的[9−10], 这类线性时变系统的稳定性和收敛性在一些文章中已经讨论过了[9−10, 24]. 对于具有未知仿射项的非线性系统, 实现确定学习的困难在于闭环控制系统中出现了未知的仿射项, 使得式 (4) 中的 $A(t)$ 是时变的.

为实现含未知仿射项的非线性系统的确定学习, 文献 [11] 将与之相关的 LTV 系统转化为

$$\begin{bmatrix} \dot{\boldsymbol{e}}_1 \\ \dot{\boldsymbol{e}}_2 \\ \dot{\boldsymbol{\eta}} \end{bmatrix} = \left[\begin{array}{cc|c} & & 0 \\ & A(t) & S^{\mathrm{T}}(t) \\ \hline 0 & -\Gamma S(t)G(t) & 0 \end{array}\right] \begin{bmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \\ \boldsymbol{\eta} \end{bmatrix} \tag{5}$$

其中, $\boldsymbol{e}_1 \in \mathbf{R}^{(n-q)}, \boldsymbol{e}_2 \in \mathbf{R}^q$, $\boldsymbol{\eta} \in \mathbf{R}^p$, $A(\cdot) : [0,\infty) \to \mathbf{R}^{n\times n}$, $S(\cdot) : [0,\infty) \to \mathbf{R}^{p\times q}$, $G(\cdot) : [0,\infty) \to \mathbf{R}^{q\times q}$, $\Gamma = \Gamma^{\mathrm{T}} > 0$. 定义 $\boldsymbol{e} = [\boldsymbol{e}_1^{\mathrm{T}}, \boldsymbol{e}_2^{\mathrm{T}}]^{\mathrm{T}}, B(t) = [0 \ \ S(t)] \in \mathbf{R}^{p\times n}$, $H(t) = \mathrm{block-diag}\{I \ \ G(t)\} \in \mathbf{R}^{n\times n}$, 这里 diag 表示块对角阵形式, $C(t) = \Gamma B(t) H(t)$.

假设 1[25]. 存在一个 $\phi_{\mathrm{M}} > 0$ 以便对于所有 $t \geq 0$, 有 $\max\left\{\|B(t)\|, \left\|\frac{\mathrm{d}B(t)}{\mathrm{d}t}\right\|\right\} \leq \phi_{\mathrm{M}}$.

假设 2[25]. 存在对称矩阵 $P(t)$ 和 $Q(t)$ 以便 $-Q(t) = \dot{P}(t) + P(t)A(t) + A^{\mathrm{T}}(t)P(t)$. 而且, $\exists p_m, q_m, p_M$ 和 $q_M > 0$ 使得 $p_m I \leq P(t) \leq p_M I$ 和 $q_m I \leq Q(t) \leq q_M I$.

引理 1[11]. 考虑系统 (5), 在紧集 Ω 内假设 1 和假设 2 满足, 如果 $S(t)$ 满足 PE 条件, 那么系统 (5) 在紧集 Ω 内是一致指数稳定的.

2 机器人任务空间的自适应神经网络控制与学习

2.1 问题描述

非冗余的 n 关节机器人, 任务空间的动态方程可表示为[26]

$$M_x(q)\ddot{\boldsymbol{x}} + V_x(q,\dot{q})\dot{\boldsymbol{x}} + \boldsymbol{G}_x(q) + \boldsymbol{F}_x(\dot{q}) = \boldsymbol{\tau}_x \tag{6}$$

其中, $M_x(q) = J^{-\mathrm{T}}(q)M(q)J^{-1}(q)$, $\boldsymbol{G}_x(q) = J^{-\mathrm{T}}(q)\boldsymbol{G}(q)$, $V_x(q,\dot{q}) = J^{-\mathrm{T}}(q)V_m(q,\dot{q})J^{-1}(q) - J^{-\mathrm{T}}(q)M(q)J^{-1}(q)\dot{J}(q)J^{-1}(q)$, $\boldsymbol{F}_x(\dot{q}) = J^{-\mathrm{T}}(q)\boldsymbol{F}(\dot{q})$, $\boldsymbol{\tau}_x = J^{-\mathrm{T}}(q)\boldsymbol{\tau}$. $J(q) \in \mathbf{R}^{n\times n}$ 为雅可比矩阵, $\boldsymbol{x} \in \mathbf{R}^n$ 为机器人任务空间的位置矢量.

动态方程 (6) 具有如下的结构属性[26].

性质 1. 惯性矩阵 $M_x(q)$ 为正定对称矩阵, 且有界, 即满足 $m_1 I \leq M_x(q) \leq m_2 I, \forall q \in \mathbf{R}^n$. 其中, I 为适当维数的单位矩阵, $m_2 \geq m_1 > 0$.

性质 2. 矩阵 $[\dot{M}_x(q) - 2V_x(q,\dot{q})]$ 是斜对称的.

令 $\boldsymbol{z}_1 = \boldsymbol{x}, \boldsymbol{z}_2 = \dot{\boldsymbol{x}}$, 则式 (6) 可表示为

$$\begin{aligned} \dot{\boldsymbol{z}}_1 &= \boldsymbol{z}_2 \\ \dot{\boldsymbol{z}}_2 &= \boldsymbol{F}_1 + G_1\boldsymbol{\tau}_x \end{aligned} \tag{7}$$

其中, $\boldsymbol{F}_1 = -M_x^{-1}(V_x\dot{\boldsymbol{x}} + \boldsymbol{G}_x + \boldsymbol{F}_x)$, $G_1 = M_x^{-1}$, 并假定是完全未知的.

性质 3. 根据性质 1, $G_1(\cdot)$ 是正定对称有界的, 即存在正常数 $g_2 \geq g_1 > 0$, 使得 $g_2 I \geq G_1(\cdot) \geq g_1 I$.

性质 4. 根据性质 1, $\dot{G}_1^{-1} = \dot{M}_x$ 也是有界的, 即存在正常数 $g_3 > 0$, 使得 $||\dot{G}_1^{-1}(\cdot)|| \leq g_3$.

控制目标: 对于给定的机器人任务空间期望有界光滑回归参考轨迹, 使用 RBF 神经网络设计自适应神经网络控制器, 实现机器人任务空间的控制与学习. 文中假定参考轨迹的一阶和二阶导数也是有界光滑的回归轨迹.

2.2 自适应神经网络控制器设计

对于系统 (7) 及有界光滑的回归参考轨迹 $\varphi_d(\boldsymbol{x}_d(0))$, 即 $\boldsymbol{X}_d(t) = \boldsymbol{Z}_d(t) = [\boldsymbol{x}_d^{\mathrm{T}}(t), \dot{\boldsymbol{x}}_d^{\mathrm{T}}(t)]^{\mathrm{T}}$,使用高斯 RBF 神经网络的控制器设计过程如下.

定义滤波器跟踪误差

$$\boldsymbol{r} = \dot{\boldsymbol{e}} + \beta\boldsymbol{e} \tag{8}$$

其中, $\boldsymbol{e}(t) = \boldsymbol{x}(t) - \boldsymbol{x}_d(t)$, $\dot{\boldsymbol{e}}(t) = \dot{\boldsymbol{x}}(t) - \dot{\boldsymbol{x}}_d(t)$, β 为正定对角设计矩阵, 对 $\boldsymbol{r}(t)$ 求导得:

$$\dot{\boldsymbol{r}} = \dot{\boldsymbol{z}}_2 - \ddot{\boldsymbol{x}}_d + \beta\dot{\boldsymbol{e}} = \boldsymbol{F}_1 + G_1\boldsymbol{\tau}_x - \ddot{\boldsymbol{x}}_d + \beta\dot{\boldsymbol{e}} \tag{9}$$

选择控制输入

$$\boldsymbol{\tau}_x = -K_v\boldsymbol{r} - \hat{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z}) \tag{10}$$

其中, $K_v = \mathrm{diag}\{k_{vi}\}$ $(k_{vi} > 0)$ 为控制增益矩阵, $S(\boldsymbol{Z}) = \mathrm{diag}\{\boldsymbol{S}_1(\boldsymbol{Z}), \cdots, \boldsymbol{S}_n(\boldsymbol{Z})\}$, $\hat{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z}) = [\hat{\boldsymbol{W}}_1^{\mathrm{T}}\boldsymbol{S}_1(\boldsymbol{Z}), \cdots, \hat{\boldsymbol{W}}_n^{\mathrm{T}}\boldsymbol{S}_n(\boldsymbol{Z})]^{\mathrm{T}}$ (后面的 $\boldsymbol{W}^{*\mathrm{T}}S(\boldsymbol{Z})$, $\tilde{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$, $\tilde{\boldsymbol{W}}_\xi^{\mathrm{T}}S_\xi(\boldsymbol{Z})$, $\hat{\boldsymbol{W}}_\xi^{\mathrm{T}}S_\xi(\boldsymbol{Z})$, $\bar{\boldsymbol{W}}_\xi^{\mathrm{T}}S_\xi(\boldsymbol{Z})$, $\bar{\boldsymbol{W}}_{\bar{\xi}}^{\mathrm{T}}S_{\bar{\xi}}(\boldsymbol{Z})$, $\bar{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$ 的定义类同) 被用来逼近不确定函数:

$$\boldsymbol{F}_{Gx}(\boldsymbol{Z}) = G_1^{-1}(\boldsymbol{F}_1 - \ddot{\boldsymbol{x}}_d + \beta\dot{\boldsymbol{e}}) = \boldsymbol{W}^{*\mathrm{T}}S(\boldsymbol{Z}) + \boldsymbol{\varepsilon}(\boldsymbol{Z}) \tag{11}$$

其中, $\boldsymbol{F}_{Gx}(\boldsymbol{Z}) = [F_{Gx1}(\boldsymbol{Z}), \cdots, F_{Gxn}(\boldsymbol{Z})]^{\mathrm{T}}$, 神经网络的输入为 $\boldsymbol{Z} = [\boldsymbol{z}_1^{\mathrm{T}}, \boldsymbol{z}_2^{\mathrm{T}}]^{\mathrm{T}} \in \Omega \subset \mathbf{R}^{2n}$, $\boldsymbol{W}^*$ 为理想的神经网络权值, $\boldsymbol{\varepsilon}(\boldsymbol{Z})$ 为神经网络逼近误差, 在紧集 Ω 内有 $\|\boldsymbol{\varepsilon}(\boldsymbol{Z})\| < \varepsilon^*$ $(\varepsilon^* > 0)$.

选择权值更新律为

$$\dot{\hat{\boldsymbol{W}}} = \Gamma S(\boldsymbol{Z})\boldsymbol{r} - \sigma\Gamma\hat{\boldsymbol{W}} \tag{12}$$

其中, $\Gamma = \Gamma^{\mathrm{T}} > 0$ 为对角设计矩阵, $\sigma > 0$ 是一个很小正数.

由被控系统 (7), 滤波器跟踪误差 (8), 控制器 (10) 以及神经网络权值更新律 (12) 构成的闭环系统可表示为

$$\begin{cases} \dot{\boldsymbol{r}} = G_1[-K_v\boldsymbol{r} - \tilde{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z}) + \boldsymbol{\varepsilon}] \\ \dot{\hat{\boldsymbol{W}}} = \dot{\tilde{\boldsymbol{W}}} = \Gamma S(\boldsymbol{Z})\boldsymbol{r} - \sigma\Gamma\hat{\boldsymbol{W}} \end{cases} \tag{13}$$

其中, $\tilde{\boldsymbol{W}} = \hat{\boldsymbol{W}} - \boldsymbol{W}^*$.

定理 1 (UUB 和跟踪收敛). 考虑闭环系统 (13), 对给定的从初始条件 $\boldsymbol{x}_d(0) \in \Omega_d$ (Ω_d 为紧集) 开始的任意回归参考轨迹 $\varphi_d(\boldsymbol{x}_d(0))$, 有初始条件 $\boldsymbol{x}(0) \in \Omega_0$ (Ω_0 为紧集) 和 $\hat{\boldsymbol{W}}(0) = 0$, 则有:

1) 闭环系统所有信号保持最终一致有界;

2) 存在一个有限时间 T_1, 使得在控制器参数选择适当的情况下, 对所有的 $t \geq T_1$, 状态跟踪误差 $[\boldsymbol{e}^{\mathrm{T}}, \dot{\boldsymbol{e}}^{\mathrm{T}}]^{\mathrm{T}}$ 收敛到零的一个小邻域内.

证明. 1) 选取李亚普诺夫函数

$$V = \boldsymbol{r}^{\mathrm{T}}G_1^{-1}\frac{\boldsymbol{r}}{2} + \tilde{\boldsymbol{W}}^{\mathrm{T}}\Gamma^{-1}\frac{\tilde{\boldsymbol{W}}}{2} \tag{14}$$

对式 (14) 求导得:

$$\dot{V} = -\boldsymbol{r}^{\mathrm{T}}K_v\boldsymbol{r} + \boldsymbol{r}^{\mathrm{T}}\boldsymbol{\varepsilon} - \sigma\tilde{\boldsymbol{W}}^{\mathrm{T}}\hat{\boldsymbol{W}} + \boldsymbol{r}^{\mathrm{T}}\dot{G}_1^{-1}\frac{\boldsymbol{r}}{2} \tag{15}$$

令 $K_v = K_{v1} + K_{v2} + g_3 I/2$, 其中 $K_{v1}, K_{v2} > 0$, 并使用不等式 $k\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha} + \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\beta}/4k \geq \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta}$, 可得:

$$\boldsymbol{r}^{\mathrm{T}}\boldsymbol{\varepsilon} \leq \lambda_{\min}(K_{v2})\left\|\boldsymbol{r}\right\|^2 + \frac{\varepsilon^{*2}}{4\lambda_{\min}(K_{v2})} \tag{16}$$

又因为

$$-\sigma\tilde{\boldsymbol{W}}^{\mathrm{T}}\hat{\boldsymbol{W}} \leq \frac{-\sigma\left\|\tilde{\boldsymbol{W}}\right\|^2}{2} + \frac{\sigma\left\|\boldsymbol{W}^*\right\|^2}{2} \tag{17}$$

故式 (15) 可表示为

$$\begin{aligned} \dot{V} \leq -\lambda_{\min}(K_{v1})\left\|\boldsymbol{r}\right\|^2 - \frac{\sigma||\tilde{\boldsymbol{W}}||^2}{2} + \\ \frac{\sigma||\boldsymbol{W}^*||^2}{2} + \frac{\varepsilon^{*2}}{4\lambda_{\min}(K_{v2})} \end{aligned} \tag{18}$$

从上式可知, 要使 $\dot{V}$ 为负定的, 必须满足 $\|\boldsymbol{r}\| > \frac{\varepsilon^*}{2\sqrt{\lambda_{\min}(K_{v1})\lambda_{\min}(K_{v2})}} + \sqrt{\frac{\sigma}{2\lambda_{\min}(K_{v1})}}\left\|\boldsymbol{W}^*\right\|$ 或者 $||\tilde{\boldsymbol{W}}|| > \frac{\varepsilon^*}{\sqrt{2\sigma\lambda_{\min}(K_{v2})}} + ||\boldsymbol{W}^*||$.

由此得到 $\boldsymbol{r}$ 和 $\tilde{\boldsymbol{W}}$ 是最终一致有界的, 且界为

$$\begin{aligned} \left\|\boldsymbol{r}\right\| \leq \frac{\varepsilon^*}{2\sqrt{\lambda_{\min}(K_{v1})\lambda_{\min}(K_{v2})}} + \\ \sqrt{\frac{\sigma}{2\lambda_{\min}(K_{v1})}}\left\|\boldsymbol{W}^*\right\| \end{aligned} \tag{19}$$

$$\left\|\tilde{\boldsymbol{W}}\right\| \leq \frac{\varepsilon^*}{\sqrt{2\sigma\lambda_{\min}(K_{v2})}} + \left\|\boldsymbol{W}^*\right\| := \tilde{W}^* \tag{20}$$

适当选取设计矩阵 β 可保证 $\boldsymbol{r} = \dot{\boldsymbol{e}} + \beta\boldsymbol{e}$ 是稳定的. 因为 $\boldsymbol{r}$, $\boldsymbol{x}_d, \dot{\boldsymbol{x}}_d$ 是有界的, 所以 $\boldsymbol{z}_1, \boldsymbol{z}_2$ 是有界的, 且对于所有 $\boldsymbol{Z}$, $S(\boldsymbol{Z})$ 是有界的, 故知 $\boldsymbol{\tau}_x$ 也是有界的. 由此可得, 闭环系统所有信号是最终一致有界的.

2) 选取李亚普诺夫函数

$$V_r = \boldsymbol{r}^{\mathrm{T}}G_1^{-1}\frac{\boldsymbol{r}}{2} \tag{21}$$

对 V_r 求导得:

$$\begin{aligned} \dot{V}_r = \frac{\boldsymbol{r}^{\mathrm{T}}\dot{\boldsymbol{r}}}{2} + \frac{\boldsymbol{r}^{\mathrm{T}}\dot{G}_1^{-1}\boldsymbol{r}}{2} \leq -\boldsymbol{r}^{\mathrm{T}}K_v\boldsymbol{r} + \left\|\boldsymbol{r}\right\|\left\|\boldsymbol{\varepsilon}\right\| + \\ \left\|\boldsymbol{r}\right\|\left|\left|\tilde{\boldsymbol{W}}\right|\right|\left\|S(\boldsymbol{Z})\right\| + \frac{\boldsymbol{r}^{\mathrm{T}}\dot{G}_1^{-1}\boldsymbol{r}}{2} \end{aligned} \tag{22}$$

令 $K_v = \bar{K}_{v1} + 2\bar{K}_{v2} + g_3 I/2$, 其中 $\bar{K}_{v1}, \bar{K}_{v2} > 0$, 利用三角不等式, 有:

$$\begin{aligned} \left\|\boldsymbol{r}\right\|\left\|\boldsymbol{\varepsilon}\right\| \leq \lambda_{\min}(\bar{K}_{v2})\left\|\boldsymbol{r}\right\|^2 + \frac{\varepsilon^{*^2}}{4\lambda_{\min}(\bar{K}_{v2})} \\ \left\|\boldsymbol{r}\right\|\left\|\tilde{\boldsymbol{W}}\right\|\left\|S(\boldsymbol{Z})\right\| \leq \lambda_{\min}(\bar{K}_{v2})\left\|\boldsymbol{r}\right\|^2 + \frac{\tilde{W}^{*^2}s^{*^2}}{4\lambda_{\min}(\bar{K}_{v2})} \end{aligned} \tag{23}$$

其中, s^* 是 $\|S(\boldsymbol{Z})\|$ 的上界[23], $\tilde{W}^*$ 是 $||\tilde{\boldsymbol{W}}||$ 的上界 (式 (20) 定义的), 由此, 式 (22) 可表示为

$$\dot{V}_r \leq -\lambda_{\min}(\bar{K}_{v1})\boldsymbol{r}^{\mathrm{T}}\boldsymbol{r} + \frac{\tilde{W}^{*2}s^{*2} + \varepsilon^{*2}}{4\lambda_{\min}(\bar{K}_{v2})} \tag{24}$$

令 $\delta = (\tilde{W}^{*2}s^{*2} + \varepsilon^{*2})/4\lambda_{\min}(\bar{K}_{v2})$, 显然只要选择足够大的 $\bar{K}_{v2}$, δ 就能足够小, 于是有:

$$\dot{V}_r \leq -2\lambda_{\min}(\bar{K}_{v1})g_1 V_r + \delta \tag{25}$$

令 $c = \lambda_{\min}(\bar{K}_{v1})g_1$, $p = \delta/2c > 0$, 则式 (25) 满足:

$$0 \leq V_r(t) < p + (V_r(0) - p)\exp(-2ct) \tag{26}$$

即

$$\boldsymbol{r}^{\mathrm{T}}\boldsymbol{r} < 2g_2 p + 2g_2 V_r(0)\exp(-2ct) \tag{27}$$

由上面的不等式可知, 给定 $\gamma > \sqrt{2g_2p} = \sqrt{\delta/\lambda_{\min}(\bar{K}_{v1})}$, 存在一个由 δ 和 $\lambda_{\min}(\bar{K}_{v1})$ 决定的有限时间 T_1, 使得对所有的 $t \geq T_1$, $\boldsymbol{r}$ 满足:

$$\|\boldsymbol{r}\| < \gamma \tag{28}$$

这里, γ 为一个小残差集的大小, 可由 K_v 决定. 因此, 当选取足够大的 K_v 时, 对所有的 $t \geq T_1$, 滤波器跟踪误差 $\boldsymbol{r}$ 能足够小, 即状态跟踪误差 $[\boldsymbol{e}^{\mathrm{T}}, \dot{\boldsymbol{e}}^{\mathrm{T}}]^{\mathrm{T}}$ 能够收敛到零的一个小邻域内[27], 即 $\boldsymbol{Z}(t)|_{t\geq \mathrm{T}_1}$ 收敛到 $\boldsymbol{Z}_d(t)|_{t\geq \mathrm{T}_1}$. □

注 1. 式 (11) 中最优权值 $\boldsymbol{W}^*$ 被定义为对所有的 $\boldsymbol{X} \in \Omega_x \subset \mathbf{R}^q$ 使逼近误差 $|\varepsilon|$ 最小的值, 即 $\boldsymbol{W}^* \triangleq \arg\min\limits_{\boldsymbol{W}\in\mathbf{R}^N} \left\{ \sup\limits_{\boldsymbol{X}\in\Omega_x} \left| f(\boldsymbol{X}) - \boldsymbol{W}^{\mathrm{T}}\boldsymbol{S}(\boldsymbol{X}) \right| \right\}$. 对于不同的目标函数, $\boldsymbol{W}^*$ 是不同的. 本文提出的控制方法能学到的 $\boldsymbol{W}^*$ 对应的目标函数是 $\boldsymbol{F}_{Gx}(\boldsymbol{Z})$.

注 2. 定理 1 表明经过时间 T_1 后, 系统轨迹 $\boldsymbol{Z}(t) = [\boldsymbol{z}_1^{\mathrm{T}}(t), \boldsymbol{z}_2^{\mathrm{T}}(t)]^{\mathrm{T}}$ 变成与 $\boldsymbol{Z}_d(t) = \boldsymbol{X}_d(t) = [\boldsymbol{x}_d^{\mathrm{T}}(t), \dot{\boldsymbol{x}}_d^{\mathrm{T}}(t)]^{\mathrm{T}}$ 一样的回归轨迹. 因为 $\boldsymbol{Z} = [\boldsymbol{z}_1^{\mathrm{T}}, \boldsymbol{z}_2^{\mathrm{T}}]^{\mathrm{T}}$ 被选为 RBF 神经网络的输入, 根据文献 [10] 中的定理 2.7, $S(\boldsymbol{Z})$ 将满足部分 PE 条件, 即沿着跟踪轨迹 $\boldsymbol{Z}(t)|_{t\geq \mathrm{T}_1}$, $S_\xi(\boldsymbol{Z})$ 满足 PE 条件.

2.3 从自适应神经网络控制中学习

利用高斯 RBF 神经网络的局部属性, 在有限时间 T_1 后, 沿着跟踪轨迹 $\boldsymbol{Z}(t)|_{t\geq \mathrm{T}_1}$, 闭环系统 (13) 可表示为

$$\begin{cases} \dot{\boldsymbol{r}} = G_1[-K_v\boldsymbol{r} - \tilde{\boldsymbol{W}}_\xi^{\mathrm{T}} S_\xi(\boldsymbol{Z}) + \boldsymbol{\varepsilon}_\xi] \\ \dot{\hat{\boldsymbol{W}}}_\xi = \dot{\tilde{\boldsymbol{W}}}_\xi = \Gamma_\xi S_\xi(\boldsymbol{Z})\boldsymbol{r} - \sigma\Gamma_\xi\hat{\boldsymbol{W}}_\xi \end{cases} \tag{29}$$

$$\dot{\hat{\boldsymbol{W}}}_{\bar{\xi}} = \dot{\tilde{\boldsymbol{W}}}_{\bar{\xi}} = \Gamma_{\bar{\xi}} S_{\bar{\xi}}(\boldsymbol{Z})\boldsymbol{r} - \sigma\Gamma_{\bar{\xi}}\hat{\boldsymbol{W}}_{\bar{\xi}} \tag{30}$$

其中, $S_\xi(\boldsymbol{Z})$ 为 $S(\boldsymbol{Z})$ 的子矢量, $\hat{\boldsymbol{W}}_\xi$ 为相应的权值子矢量; 下标 ξ 代表靠近轨迹 $\boldsymbol{Z}(t)|_{t\geq \mathrm{T}_1}$ 的区域, 下标 $\bar{\xi}$ 代表远离轨迹 $\boldsymbol{Z}(t)|_{t\geq T_1}$ 的区域: $\boldsymbol{\varepsilon}_\xi$ 是局部逼近误差, 且 $\|\boldsymbol{\varepsilon}_\xi\|$ 的值很小.

定理 2 (学习和知识的表达). 考虑闭环系统 (29), 对给定的从初始条件 $\boldsymbol{x}_d(0) \in \Omega_d$ 开始的任意回归参考轨迹 $\varphi_d(\boldsymbol{x}_d(0))$, 有初始条件 $\boldsymbol{x}(0) \in \Omega_0$ 和 $\hat{\boldsymbol{W}}(0) = 0$, 选取适当的控制参数, 则有:

沿着跟踪轨迹 $\boldsymbol{Z}(t)|_{t\geq \mathrm{T}_1}$, 部分神经网络权值估计 $\hat{\boldsymbol{W}}_\xi$ 收敛到最优值 $\boldsymbol{W}_\xi^*$ 附近的一个小邻域内, 闭环系统的未知动态 $\boldsymbol{F}_{Gx}(\boldsymbol{Z})$ 可被 $\hat{\boldsymbol{W}}^{\mathrm{T}} S(\boldsymbol{Z})$ 和 $\bar{\boldsymbol{W}}^{\mathrm{T}} S(\boldsymbol{Z})$ 局部准确逼近, 其中 $\bar{\boldsymbol{W}}$ 定义为

$$\bar{\boldsymbol{W}} = \underset{t\in[t_a,t_b]}{\mathrm{mean}}(\hat{\boldsymbol{W}}(t)) \tag{31}$$

这里, $[t_a, t_b]$ $(t_b > t_a > T_2 > T_1)$ 代表 $\hat{\boldsymbol{W}}$ 收敛后 (瞬态过程后) 的一个时间段.

证明. 令 $\bar{\boldsymbol{z}} = G_1^{-1}\boldsymbol{r} = M_x\boldsymbol{r}$ 以及 $\boldsymbol{\eta} = \tilde{\boldsymbol{W}}_\xi$, 子系统 (29) 可转化为

$$\begin{cases} \dot{\bar{\boldsymbol{z}}} = -[K_v G_1 - G_1^{-1}\dot{G}_1]\bar{\boldsymbol{z}} - \boldsymbol{\eta}^{\mathrm{T}} S_\xi(\boldsymbol{Z}) + \boldsymbol{\varepsilon}_\xi(\boldsymbol{Z}) \\ \dot{\boldsymbol{\eta}} = \Gamma_\xi S_\xi(\boldsymbol{Z}) G_1\bar{\boldsymbol{z}} - \sigma\Gamma_\xi\hat{\boldsymbol{W}}_\xi \end{cases} \tag{32}$$

将式 (32) 写成矩阵形式

$$\begin{bmatrix} \dot{\bar{\boldsymbol{z}}} \\ \dot{\boldsymbol{\eta}} \end{bmatrix} = \begin{bmatrix} A(t) & B(t) \\ -C(t) & 0 \end{bmatrix} \begin{bmatrix} \bar{\boldsymbol{z}} \\ \boldsymbol{\eta} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_\xi \\ -\sigma\Gamma_\xi\hat{\boldsymbol{W}}_\xi \end{bmatrix} \tag{33}$$

因为 $\|\boldsymbol{\varepsilon}_\xi\|$ 和 $\|\sigma\Gamma_\xi\hat{\boldsymbol{W}}_\xi\|$ 很小, 系统 (33) 能够看成系统 (5) 的扰动系统[28]. 其中

$$A(t) = [-K_v G_1 + \dot{G}_1^{-1} G_1] \tag{34}$$

$$B(t) = [-S_\xi^{\mathrm{T}}(\boldsymbol{Z})] \tag{35}$$

$$C(t) = [-\Gamma_\xi S_\xi(\boldsymbol{Z}) G_1] \tag{36}$$

选取 $P(t) = [G_1]$, 则有:

$$\dot{P} + PA + A^{\mathrm{T}}P = [-2G_1(K_v - \dot{G}_1^{-1})G_1 + \dot{G}_1] \tag{37}$$

假设 1 很容易满足, 据性质 3 和 $\dot{G}_1(\cdot), \dot{G}_1^{-1}(\cdot)$ 的有界性, 可选择 K_v 使得 $2G_1(K_v - \dot{G}_1^{-1})G_1 - \dot{G}_1$ 为严格正定的, 故 $\dot{P} + A^{\mathrm{T}}P + PA$ 的负定被保证, 因此假设 2 也满足了.

经过时间 T_1 后, 神经网络的输入 $\boldsymbol{Z}(t)$ 跟踪上参考回归轨迹 $\boldsymbol{Z}_d(t)$, 根据文献 [10] 中的定理 2.7, $S(\boldsymbol{Z})$ 将满足部分 PE 条件, 即沿着跟踪轨迹 $\boldsymbol{Z}(t)|_{t\geq T_1}$, $S_\xi(\boldsymbol{Z})$ 满足 PE 条件. 根据引理 1 可知, 系统 (33) 的正常系统一致指数稳定. 对于扰动系统 (33), 使用文献 [28] 中的引理 4.6, 参数误差 $\boldsymbol{\eta} = \tilde{\boldsymbol{W}}_\xi$ 在有限时间 T_2 $(T_2 > T_1)$ 内指数收敛到零附近的一个小邻域内. 邻域的大小由神经网络的逼近能力以及状态跟踪误差决定.

$\hat{\boldsymbol{W}}_\xi$ 收敛到 $\boldsymbol{W}_\xi^*$ 附近的小邻域内意指沿着跟踪轨迹 $\boldsymbol{Z}(t)|_{t\geq \mathrm{T}_2}$, 未知闭环系统动态 $\boldsymbol{F}_{Gx}(\boldsymbol{Z})$ 能够用递归子矢量 $S_\xi(\boldsymbol{Z})$ 表示, 且误差很小, 即

$$\boldsymbol{F}_{Gx}(\boldsymbol{Z}) = \hat{\boldsymbol{W}}_\xi^{\mathrm{T}} S_\xi(\boldsymbol{Z}) + \boldsymbol{\varepsilon}_\xi \tag{38}$$

其中, $|\|\boldsymbol{\varepsilon}_\xi\| - \varepsilon^*|$ 很小. 根据 $\bar{\boldsymbol{W}} = \underset{t\in[t_a,t_b]}{\mathrm{mean}}(\hat{\boldsymbol{W}}(t))$ 选取 $\bar{\boldsymbol{W}}$, 由 RBF 神经网络的局部属性, 式 (38) 可表示为

$$\boldsymbol{F}_{Gx}(\boldsymbol{Z}) = \bar{\boldsymbol{W}}_\xi^{\mathrm{T}} S_\xi(\boldsymbol{Z}) + \bar{\boldsymbol{\varepsilon}}_\xi \tag{39}$$

其中, $\bar{\boldsymbol{W}}_\xi = [\bar{w}_{j1}, \cdots, \bar{w}_{j\xi}]^{\mathrm{T}}$ 为 $\bar{\boldsymbol{W}}$ 的子矢量, $\bar{\boldsymbol{\varepsilon}}_\xi$ 是闭环系统动态由 $\bar{\boldsymbol{W}}_\xi^{\mathrm{T}} S_\xi(\boldsymbol{Z})$ 逼近时的误差, 在瞬态过程结束以后, $|\|\bar{\boldsymbol{\varepsilon}}_\xi\| - \|\boldsymbol{\varepsilon}_\xi\||$ 是很小的.

由高斯 RBF 神经网络的局部属性, 沿着跟踪轨迹 $\boldsymbol{Z}(t)|_{t\geq \mathrm{T}_2}$, $S_{\bar{\xi}}$ 和 $\bar{\boldsymbol{W}}_{\bar{\xi}}^{\mathrm{T}} S_{\bar{\xi}}(\boldsymbol{Z})$ 都非常小, 这意

味着沿着跟踪轨迹 $\boldsymbol{Z}(t)|_{t\geq T_2}$, 整个 RBF 神经网络 $\bar{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$ 能够逼近未知闭环系统动态 $\boldsymbol{F}_{Gx}(\boldsymbol{Z})$

$$\begin{aligned}\boldsymbol{F}_{Gx}(\boldsymbol{Z}) &= \bar{\boldsymbol{W}}_{\xi}^{\mathrm{T}}S_{\xi}(\boldsymbol{Z}) + \bar{\boldsymbol{W}}_{\bar{\xi}}^{\mathrm{T}}S_{\bar{\xi}}(\boldsymbol{Z}) + \boldsymbol{\varepsilon}_1 = \\ &\quad \bar{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z}) + \boldsymbol{\varepsilon}_2\end{aligned} \tag{40}$$

其中, $|\|\boldsymbol{\varepsilon}_2\| - \|\bar{\boldsymbol{\varepsilon}}_{\xi}\||$ 是很小的, 因此可知沿着跟踪轨迹 $\boldsymbol{Z}(t)|_{t\geq T_2}$, $\bar{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$ 能够逼近未知的非线性函数 (未知闭环系统动态) $\boldsymbol{F}_{Gx}(\boldsymbol{Z})$. □

注 3.　定理 2 表明在对回归参考轨迹的跟踪控制过程中能够实现未知闭环系统动态的确定学习 (即部分神经网络权值收敛到最优值), 学过的知识以时不变且空间分布的方式表达、以常值神经网络权值的方式存储. 但在传统的神经网络控制中, 由于不满足 PE 条件, 神经网络权值的收敛性不能得到保证, 而使用时变神经网络权值表达和存储学过的知识一般是很困难的. 虽然在有些神经网络文献中, 处理方法是提出一种算法进行训练, 再提出停止训练的条件, 然后, 认为此时神经网络已实现近似, 但这种方法目前还缺乏相应的理论证明. 而通过确定学习, 过去经验或知识的表达与存储变成了一件很容易的事情.

注 4.　本文提出的基于确定学习的控制方法包括两个阶段, 第一阶段是控制与学习, 首先是设计自适应神经网络控制器, 使系统状态轨迹收敛到期望的回归轨迹, 实现系统的稳定; 然后是学习, 由于部分持续激励条件的满足, 实现了部分神经网络权值收敛到最优值及未知闭环系统动态的局部准确逼近, 即确定学习. 第二阶段是学习后的控制, 系统控制性能的提高主要体现在这个阶段上, 通过控制与学习后, 神经网络学习到的知识最终存储在常值神经网络中, 在运用所学到的知识进行后续控制时, 系统就避免了自适应环节, 不仅减小了系统的计算复杂性, 还提高了响应速度.

3　仿真研究

以两连杆平面机器人为例, 任务空间的动态方程为[26]

$$M_x(q)\ddot{\boldsymbol{x}} + V_x(q,\dot{q})\dot{\boldsymbol{x}} + \boldsymbol{G}_x(q) + \boldsymbol{F}_x(\dot{q}) = \boldsymbol{\tau}_x \tag{41}$$

其中, $M_x(q) = J^{-\mathrm{T}}(q)M(q)J^{-1}(q)$, $\boldsymbol{G}_x(q) = J^{-\mathrm{T}}(q)\boldsymbol{G}(q)$, $V_x(q,\dot{q}) = J^{-\mathrm{T}}(q)V_m(q,\dot{q})J^{-1}(q) - J^{-\mathrm{T}}(q)M(q)J^{-1}(q)\dot{J}(q)J^{-1}(q)$, $\boldsymbol{F}_x(\dot{q}) = J^{-\mathrm{T}}(q)\boldsymbol{F}(\dot{q})$, $\boldsymbol{\tau}_x = J^{-\mathrm{T}}(q)\boldsymbol{\tau}$.

$$M(q) = \begin{bmatrix} l_2^2m_2 + l_1^2(m_1+m_2) + 2l_1l_2m_2\cos(q_2) & l_2^2m_2 + l_1l_2m_2\cos(q_2) \\ l_2^2m_2 + l_1l_2m_2\cos(q_2) & l_2^2m_2 \end{bmatrix}$$

$$V_m(q,\dot{q}) = \begin{bmatrix} -l_1l_2m_2\sin(q_2)\dot{q}_2 & -l_1l_2m_2\sin(q_2)(\dot{q}_1+\dot{q}_2) \\ l_1l_2m_2\sin(q_2)\dot{q}_1 & 0 \end{bmatrix}$$

$$G(q) = \begin{bmatrix} l_1(m_1+m_2)g\cos(q_1) + l_2m_2g\cos(q_1+q_2) \\ l_2m_2g\cos(q_1+q_2) \end{bmatrix}$$

$$J(q) = \begin{bmatrix} -l_1\sin(q_1) - l_2\sin(q_1+q_2) & -l_2\sin(q_1+q_2) \\ l_1\cos(q_1) + l_2\cos(q_1+q_2) & l_2\cos(q_1+q_2) \end{bmatrix}$$

其中, q_1 和 q_2 为关节 1 和关节 2 的角度, m_1 和 m_2 为连杆 1、2 的质量; l_1 和 l_2 为连杆 1、2 的长度; g 为重力加速度. 在仿真中摩擦力表示为[29]

$$\boldsymbol{F}(q) = \begin{bmatrix} 12\dot{q}_1 + 0.5\mathrm{sgn}(\dot{q}_1) & 12\dot{q}_2 + 0.5\mathrm{sgn}(\dot{q}_2) \end{bmatrix}^{\mathrm{T}}$$

参考轨迹为 $\boldsymbol{x}_d(t) = [x_{d1}(t), x_{d2}(t)]^{\mathrm{T}} = [0.8\sin t, 0.8\cos t]^{\mathrm{T}}$. 本文提出的控制器设计过程归纳如下. 首先, 定义状态跟踪误差为 $\boldsymbol{e} = \boldsymbol{x} - \boldsymbol{x}_d$ 和 $\dot{\boldsymbol{e}} = \dot{\boldsymbol{x}} - \dot{\boldsymbol{x}}_d$, 则有:

$$\boldsymbol{r} = \dot{\boldsymbol{e}} + \lambda\boldsymbol{e} \tag{42}$$

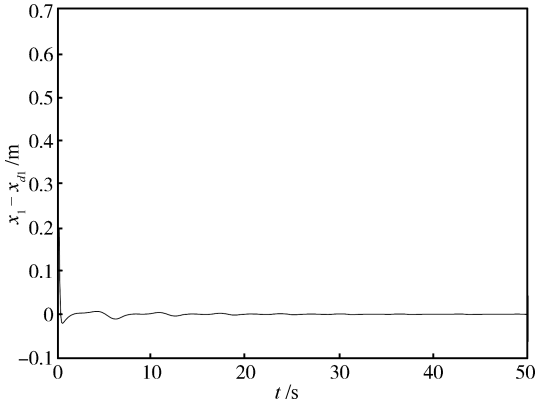
选取控制输入

$$\boldsymbol{\tau}_x = -k_v\boldsymbol{r} - \hat{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z}) \tag{43}$$

选取权值更新律

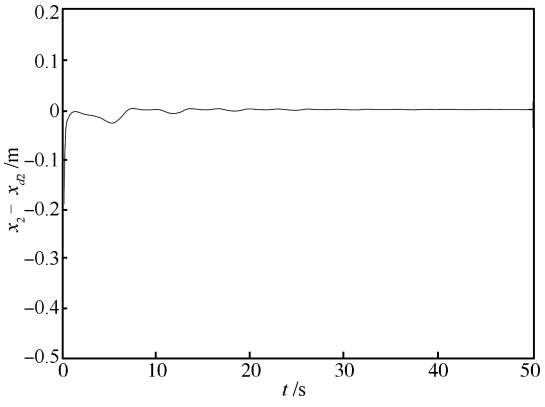
$$\dot{\hat{\boldsymbol{W}}} = \Gamma S(\boldsymbol{Z})\boldsymbol{r} - \sigma\Gamma\hat{\boldsymbol{W}} \tag{44}$$

两连杆平面机器人的系统参数为[23] $l_1 = l_2 = 1.0\,\mathrm{m}$, $m_1 = 0.8\,\mathrm{kg}$, $m_2 = 2.3\,\mathrm{kg}$, $g = 9.8\,\mathrm{m/s^2}$. 控制器参数选择: 神经网络节点数 $N = 5\times5\times5\times5 = 625$, 均匀分布在 $[-1.5\ \ 1.5]\times[-1.5\ \ 1.5]\times[-1.5\ \ 1.5]\times[-1.5\ \ 1.5]$ 范围内, 宽度 $\eta = 0.7$, $\Gamma = \mathrm{diag}\{10\}$, $K_v = \mathrm{diag}\{100\}$, $\beta = \mathrm{diag}\{10\}$, $\sigma = 0.0001$, 初始权值 $\hat{\boldsymbol{W}}(0) = 0$. 初始状态 $\boldsymbol{X}(0) = [0.6083, 0.3196; 0, 0]$. 图 1 为跟踪误差曲线, 图 2 为部分神经网络权值收敛, 图 3 为函数逼近曲线, 图 4 为控制输入.



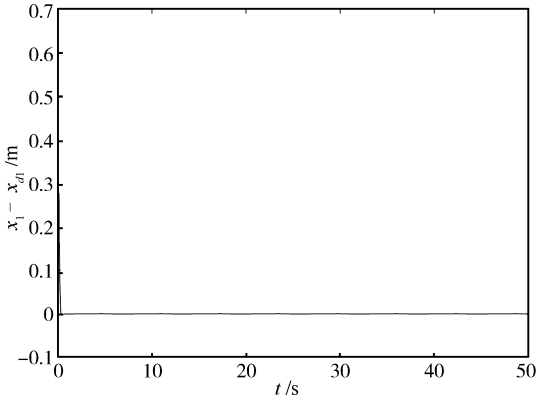
(a) 使用 $\hat{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$ 时 x_1 的跟踪误差

(a) Tracking error of x_1 using $\hat{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$



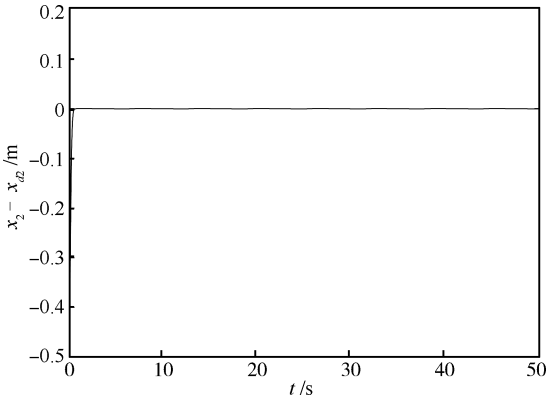
(b) 使用 $\hat{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$ 时 x_2 的跟踪误差

(b) Tracking error of x_2 using $\hat{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$



(c) 使用 $\bar{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$ 时 x_1 的跟踪误差

(c) Tracking error of x_1 using using $\bar{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$

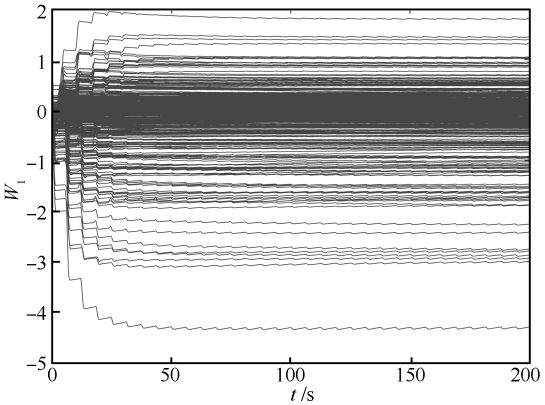


(d) 使用 $\bar{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$ 时 x_2 的跟踪误差

(d) Tracking error of x_2 using $\bar{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$

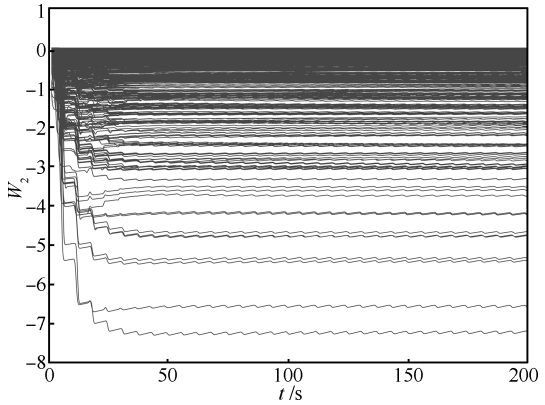
图 1 状态跟踪误差曲线

Fig. 1 State tracking error curves



(a) 部分权值 $\hat{W}_{1\xi}$ 收敛

(a) Partial parameter convergence $\hat{W}_{1\xi}$

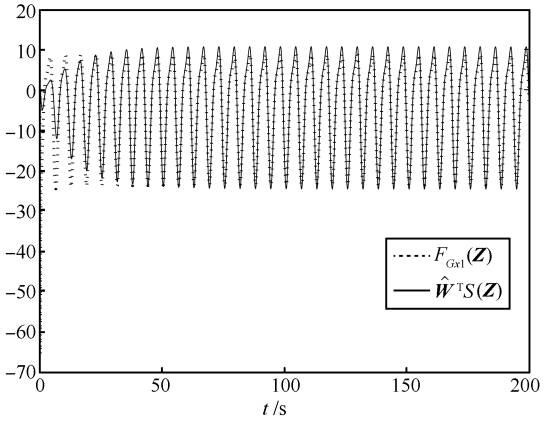


(b) 部分权值 $\hat{W}_{2\xi}$ 收敛

(b) Partial parameter convergence $\hat{W}_{2\xi}$

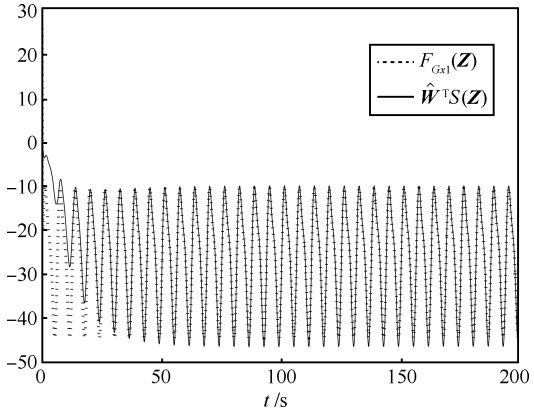
图 2 部分权值 $\hat{\boldsymbol{W}}_{\xi}$ 收敛

Fig. 2 Partial parameter convergence $\hat{\boldsymbol{W}}_{\xi}$



(a) 函数 $F_{Gx1}(\boldsymbol{Z})$ 的逼近

(a) Function approximation of $F_{Gx1}(\boldsymbol{Z})$

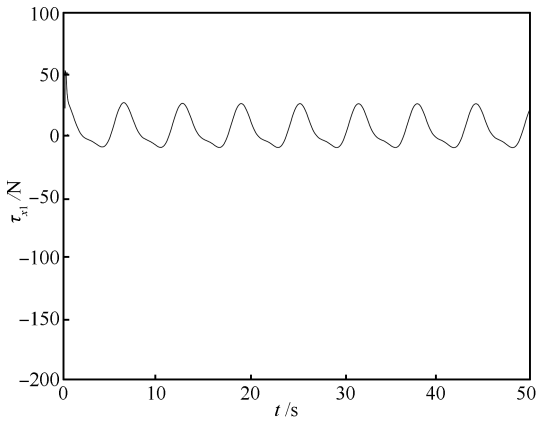


(b) 函数 $F_{Gx2}(\boldsymbol{Z})$ 的逼近

(b) Function approximation of $F_{Gx2}(\boldsymbol{Z})$

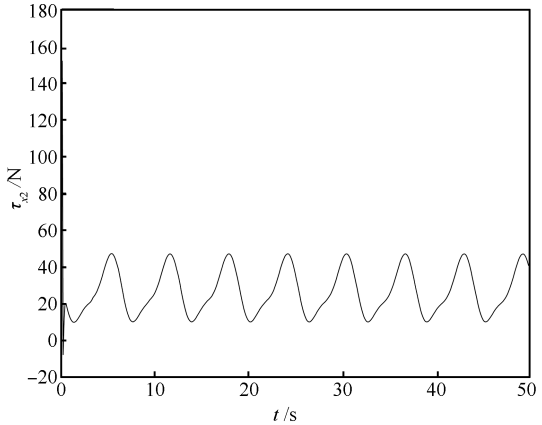
图 3　函数逼近

Fig. 3　Function approximation



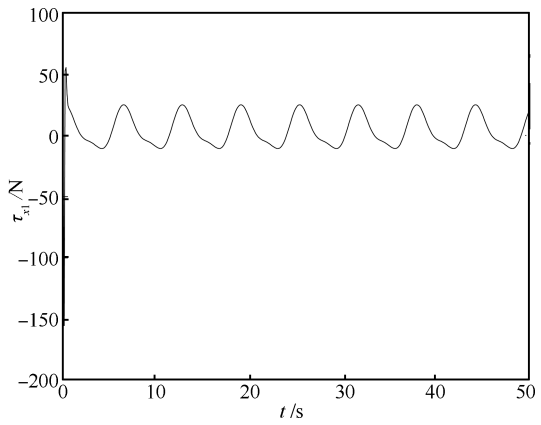
(a) 使用 $\hat{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$ 时控制输入 τ_{x1}

(a) Control input τ_{x1} using $\hat{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$



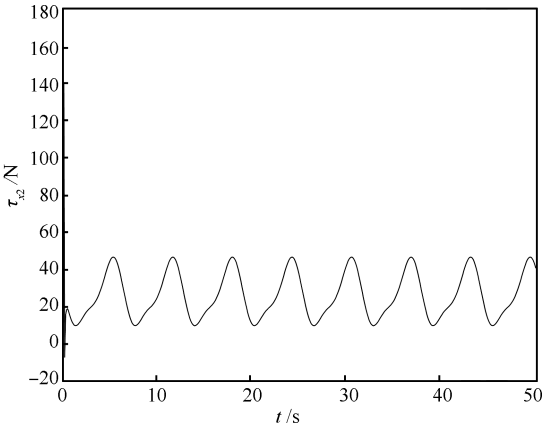
(b) 使用 $\hat{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$ 时控制输入 τ_{x2}

(b) Control input τ_{x2} using $\hat{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$



(c) 使用 $\bar{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$ 时控制输入 τ_{x1}

(c) Control input τ_{x1} using $\bar{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$



(d) 使用 $\bar{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$ 时控制输入 τ_{x2}

(d) Control input τ_{x2} using $\bar{\boldsymbol{W}}^{\mathrm{T}}S(\boldsymbol{Z})$

图 4　控制输入 τ_x

Fig. 4　Control input τ_x

从仿真结果可以清楚地看出, 提出的控制算法不仅成功地驱动仿真系统的状态收敛到参考状态附近很小的邻域内, 而且在跟踪控制过程中实现了确定学习. 跟踪误差曲线图 1 说明, 使用带权值在线更新的神经控制器使系统跟踪上参考轨迹的时间大于使用学习后带常数权值的神经控制器使系统跟踪上参考轨迹的时间. 由此可知使用学过的知识可以快速地使系统轨迹跟踪上参考轨迹, 改进系统的控制性能, 节省时间和能量.

注 5. 当系统有小的测量噪声时, 也能保证确定学习的实现, 但噪声对系统的稳态性能和学习都有一定的影响. 详细的分析见文献 [15].

注 6. 文献 [5−8] 提出的自适应神经网络控制器也实现了机器人任务空间的跟踪控制, 但它们没有实现对未知闭环系统动态的真正学习. 本文设计的自适应神经网络控制器, 理论研究和仿真实验都验证了它不仅实现了机器人任务空间的跟踪控制, 而且在稳定的控制过程中实现了部分神经网络权值收敛到最优值及未知闭环系统动态的局部准确逼近. 这是其他自适应神经网络控制器不具有的特点.

4 结论

本文针对系统参数未知的机器人, 基于确定学习理论在任务空间提出了一种新的自适应神经网络控制与学习算法. 提出的控制算法在对回归轨迹进行稳定的跟踪控制中, 通过自主学习, 实现了部分神经网络权值收敛到最优值及机器人未知闭环系统动态的局部准确逼近, 即确定学习. 学过的知识以时不变且空间分布的方式表达、以常值神经网络权值的方式存储, 可以用来改进系统的控制性能, 也可以应用于后续相同或相似的控制任务中, 节省时间和能量. 也就是说, 本文所提出的控制与学习策略能够实现知识的 “学习”、“表达与存储” 以及 “再利用”, 具有像人类一样的学习能力.

References

1 Lewis F L, Liu K, Yesildirek A. Neural net robot controller with guaranteed tracking performance. *IEEE Transactions on Neural Networks*, 1995, **6**(3): 703−715

2 Sun F C, Sun Z Q, Woo P Y. Neural network-based adaptive controller design of robotic manipulators with an observer. *IEEE Transactions on Neural Networks*, 2001, **12**(1): 54−67

3 Peng Ji-Gen, Ni Yuan-Hua, Qiao Hong. Neural network control of flexible-joint robot manipulators. *Acta Automatica Sinica*, 2007, **33**(2): 175−180
(彭济根, 倪元华, 乔红. 柔性关节机操手的神经网络控制. 自动化学报, 2007, **33**(2): 175−180)

4 Yu Jian-Cheng, Li Qiang, Zhang Ai-Qun, Wang Xiao-Hui. Neural network adaptive control for underwater vehicles. *Control Theory & Applications*, 2008, **25**(1): 9−13
(俞建成, 李强, 张艾群, 王晓辉. 水下机器人的神经网络自适应控制. 控制理论与应用, 2008, **25**(1): 9−13)

5 Ge S S, Hang C C, Woon L C. Adaptive neural network control of robot manipulators in task space. *IEEE Transactions on Industrial Electronics*, 1997, **44**(6): 746−752

6 Cheng L, Hou Z G, Tan M. Adaptive neural network tracking control of manipulators using quaternion feedback. In: Proceedings of the 2008 IEEE International Conference on Robotics and Automation. Pasadena, CA: IEEE, 2008. 3371−3376

7 Liu Ying-Zhuo, Wang Yue-Chao, Xi Ning. Application of Lyapunov-like methodology for a humanoid robot tracking in task space. *Control Theory & Applications*, 2004, **21**(3): 351−356
(刘英卓, 王越超, 席宁. 类 Lyapunov 理论在类人形机器人任务空间内跟踪的应用. 控制理论与应用, 2004, **21**(3): 351−356)

8 Feng Bao-Min, Ma Guang-Cheng, Wen Qi-Yong, Wang Chang-Hong. Design of robust intelligent controller for space robot in task space. *Journal of Astronautics*, 2007, **28**(4): 914−919
(丰保民, 马广程, 温奇咏, 王常虹. 任务空间内空间机器人鲁棒智能控制器设计. 宇航学报, 2007, **28**(4): 914−919)

9 Wang C, Hill D J. Learning from neural control. *IEEE Transactions on Neural Networks*, 2006, **17**(1): 130−146

10 Wang C, Hill D J. *Deterministic Learning Theory for Identification, Recognition and Control*. Boca Raton: CRC Press, 2009

11 Liu T F, Wang C, Hill D J. Learning from neural control of nonlinear systems in normal form. *Systems & Control Letters*, 2009, **58**(9): 633−638

12 Wang C, Wang M, Liu T F, Hill D J. Learning from ISS-modular adaptive NN control of nonlinear strict-feedback systems. *IEEE Transactions on Neural Networks and Learning Systems*, 2012, **23**(10): 1539−1550

13 Wang C, Chen T R. Rapid detection of small oscillation faults via deterministic learning. *IEEE Transactions on Neural Networks*, 2011, **22**(8): 1284−1296

14 Yuan C Z, Wang C. Persistency of excitation and performance of deterministic learning. *Systems & Control Letters*, 2011, **60**(12): 952−959

15 Yuan C Z, Wang C. Performance of deterministic learning in noisy environments. *Neurocomputing*, 2012, **78**(1): 72−82

16 Wang Cong, Chen Tian-Rui, Liu Teng-Fei. Deterministic learning and data-based modeling and control. *Acta Automatica Sinica*, 2009, **35**(6): 693−706
(王聪, 陈填锐, 刘腾飞. 确定学习与基于数据的建模及控制. 自动化学报, 2009, **35**(6): 693−706)

17 Shilnikov L P. *Methods of Qualitative Theory in Nonlinear Dynamics, Part II*. Singapore: World Scientific, 2001

18 Wu Y X, He Q Z, Wang C. Adaptive neural learning control of rigid-link electrically-driven robot manipulators. In: Proceedings of the 30th Chinese Control Conference. Yantai, China: IEEE, 2011. 6616−6622

19 Zeng W, Wang C. Human gait recognition via deterministic learning. *Neural Networks*, 2012, **35**: 92−102

20 Wang C, Gu W J, Xue Z G. Deterministic learning from robust adaptive NN control of robot manipulators. In: Proceedings of the 8th World Congress on Intelligent Control and Automation. Jinan, China: IEEE, 2010. 526−531

21 Park J, Sandberg I W. Universal approximation using radial-basis-function networks. *Neural Computation*, 1991, **3**(2): 246−257

22 Gorinevsky D. On the persistency of excitation in radial basis function network identification of nonlinear systems. *IEEE Transactions on Neural Networks*, 1995, **6**(5): 1237−1244

23 Kurdila A J, Narcowich F J, Ward J D. Persistency of excitation in identification using radial basis function approximants. *SIAM Journal on Control and Optimization*, 1995, **33**(2): 625−641

24 Farrell J A. Stability and approximator convergence in nonparametric nonlinear adaptive control. *IEEE Transactions on Neural Networks*, 1998, **9**(5): 1008−1020

25 Loría A, Panteley E. Uniform exponential stability of linear time-varying systems: revisited. *Systems & Control Letters*, 2002, **47**(1): 13−24

26 Lewis F L, Abdallah C T, Dawson D M. *Control of Robot Manipulators*. New York: Macmillan, 1993

27 Slotine J J, Li W P. *Applied Nonlinear Control*. Englewood Cliffs: Prentice-Hall, 1991

28 Khalil H K. *Nonlinear Systems (Third edition)*. Upper Saddle River: Prentice-Hall, 2002

29 Huang S J, Lee J S. A stable self-organizing fuzzy controller for robotic motion control. *IEEE Transactions on Industrial Electronics*, 2000, **47**(2): 421−428

吴玉香 华南理工大学自动化学院副教授. 主要研究方向为非线性系统的自适应神经网络控制, 确定学习, 机器人控制及故障诊断. 本文通信作者.
E-mail: xyuwu@scut.edu.cn
(**WU Yu-Xiang** Associate professor at the School of Automation, South China University of Technology. Her research interest covers adaptive neural network control, deterministic learning, robot control, and fault diagnosis. Corresponding author of this paper.)

王 聪 华南理工大学自动化学院教授. 主要研究方向为非线性系统自适应神经网络控制与辨识, 确定学习理论, 动态模式识别, 基于模式的控制, 振动故障诊断及在航空航天, 生物医学工程等领域的应用. E-mail: wangcong@scut.edu.cn
(**WANG Cong** Professor at the School of Automation, South China University of Technology. His research interest covers adaptive neural network control/identification, deterministic learning theory, dynamical pattern recognition, pattern-based intelligent control, oscillation fault diagnosis, and the application in aerospace and biomedical engineering.)