

改进的 Elman 模型与递归反传控制神经网络*

时小虎¹, 梁艳春^{1,2+}, 徐旭³

¹(吉林大学 计算机科学与技术学院 国家教育部符号计算与知识工程重点实验室,吉林 长春 130012)

²(中国科学院 计算技术研究所 智能信息处理重点实验室,北京 100080)

³(吉林大学 数学学院,吉林 长春 130012)

An Improved Elman Model and Recurrent Back-Propagation Control Neural Networks

SHI Xiao-Hu¹, LIANG Yan-Chun^{1,2+}, XU Xu³

¹(Key Laboratory of Symbol Computation and Knowledge Engineering of the Ministry of Education, College of Computer Science and Technology, Jilin University, Changchun 130012, China)

²(Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, The Chinese Academy of Sciences, Beijing 100080, China)

³(College of Mathematics Science, Jilin University, Changchun 130012, China)

+Corresponding author: E-mail: ycliang@public.cc.jl.cn

<http://www.jlu.edu.cn>

Received 2002-08-20; Accepted 2002-09-17

Shi XH, Liang YC, Xu X. An improved Elman model and recurrent back-propagation control neural networks. *Journal of Software*, 2003,14(6):1110~1119.

<http://www.jos.org.cn/1000-9825/14/1110.htm>

Abstract: Two improved Elman neural networks, output-input feedback Elman network and output-hidden feedback Elman network are presented based on the Elman neural network. By using the output-input feedback Elman network as a passageway of the error back propagation, a recurrent back propagation control neural network model is developed. The stability of the improved Elman neural networks is proved in the sense of Lyapunov stability theory. The optimal adaptive learning rates are obtained, which can guarantee the stable convergence of the improved Elman networks. The ultrasonic motor is simulated by using the Elman and improved Elman networks respectively. Besides simulating the speed of the ultrasonic motor successfully, some useful results are also obtained. According to the results, the different network models based on the sampling situation in the fieldwork can be chosen. Numerical results show that the recurrent back propagation control neural network controller has good effectiveness for various kinds of reference speeds of the ultrasonic motor.

Key words: neural network; feedback; Lyapunov stability

摘要: 在 Elman 网络的基础上提出了两种改进网络:输出-输入反馈 Elman 网络和输出-隐层反馈 Elman 网络

* Supported by the National Natural Science Foundation of China under Grant No.19872027 (国家自然科学基金)

第一作者简介: 时小虎(1974—),男,河北玉田人,博士生,助教,主要研究领域为人工神经网络,模糊系统,计算机应用。

模型,并以前者作为误差反传的通道,建立了递归反向传播控制神经网络模型.在 Lyapunov 稳定性意义下分别给出了改进网络的稳定性证明,得到了保证网络稳定收敛的最佳自适应学习速率.分别用 Elman 网络及其改进网络对超声马达进行了模拟.利用改进的 Elman 网络模型,除了可以较好地模拟马达速度以外,还得到了一些有意义的结果,据此可以根据现场数据采样的情况,选用不同的网络模型.模拟实验结果表明,递归反向传播控制神经网络对多种形式的超声马达参考速度都有很好的控制效果.

关键词: 神经网络;反馈;李雅普诺夫稳定性

中图法分类号: TP18 文献标识码: A

人工神经网络迄今已经经历了半个世纪的研究,出现了数十种主要的网络结构和各种各样的网络学习算法.自 20 世纪 80 年代中期以来,在极短的时间内,神经网络就以其独特的非传统表达方式和固有的学习能力,在系统建模、辨识与控制中获得了广泛的应用,引起了控制界的普遍重视,并取得了一系列重要结果^[1,2].超声马达 (ultrasonic motor,简称 USM)是 20 世纪 80 年代初发展起来的一种新型驱动器,具有许多优良特性,如小型、轻量、低噪声、低成本、低转速、大力矩输出、定位精确、响应速度快、不受磁场影响等.近年来,研究人员在马达的特征、工作原理、实验、应用等多方面做了大量的工作^[3-6],使其得到了迅速的发展.超声马达的工作状态要受到许多因素的影响,其速度曲线有很强的非线性,因此,如何建立实用的模型以及如何正确调整马达的输入参数,实现快速、精确的速度、位置控制,具有十分重要的意义.于是人们很自然地想到用神经网络方法对超声马达进行模拟与控制.近年来,已经有许多人员从事这方面的工作^[7-9].他们的工作主要是对不同类型的超声马达采用不同的神经网络模型进行辨识和控制.

本文在部分递归的 Elman^[10-15]网络的基础上提出了具有输出-输入反馈机制和输出-隐层反馈机制的改进 Elman——OIF Elman(output-input feedback Elman)和 OHF Elman(output-hidden feedback Elman)网络模型,在这两个改进的网络模型中分别增加了输出-输入层的反馈和输出-隐层的反馈.针对超声马达的控制问题,本文建立了递归反向传播控制(recurrent back propagation control,简称 RBPC)神经网络模型.在该模型中,用 OIF Elman 网络提供误差的反向传播通道,对某些马达的逆动态病态问题提供了可行的解决方案.对于两种改进 Elman 网络和 RBPC 网络模型,分别证明了其在 Lyapunov 能量意义下的稳定性,给出了保证控制系统快速、稳定收敛的可变自适应学习率.

文中分别利用 Elman 网络及其两种改进网络对超声马达进行了模拟,得到了一些有意义的结果.利用本文提出的 RBPC 网络对超声马达进行了控制,对多种形式的参考速度的模拟结果表明, RBPC 控制方法具有很好的控制效果.与已有的方法相比,本文提出的方法可以使速度波动的幅值大为减小.

1 Elman 网络及其改进网络

1.1 Elman 网络

Elman 网络结构如图 1 所示,它除了输入层、隐层、输出层之外,还有一个特殊的联系单元.联系单元是用来记忆隐层单元以前时刻的输出值的,可认为是一个时延算子.因此,这里前馈连接部分可以进行连接权修正,而递归部分则是固定的,即不能进行学习修正.设网络的输入层为 r 个节点,隐层和联系单元为 n 个节点,输出层为 m 个节点,则网络输入 u 为 r 维向量,隐层输出 x 及联系单元输出 x_c 为 n 维向量,网络输出 y 为 m 维向量,连接权 W^{11} 为 $n \times n$ 维矩阵, W^{12} 为 $n \times r$ 维矩阵, W^{13} 为 $m \times n$ 维矩阵.网络的数学模型为

$$x(k) = f(W^{11}x_c(k) + W^{12}u(k-1)), \quad (1)$$

$$x_c(k) = \alpha x_c(k-1) + x(k-1), \quad (2)$$

$$y(k) = W^{13}x(k), \quad (3)$$

其中 $f(x)$ 多取为 sigmoid 函数,即 $f(x) = \frac{1}{1 + e^{-x}}$, $0 \leq \alpha < 1$ 为自连接反馈增益因子.当 α 固定为 0 时,此网络为标准的 Elman 网络;当 α 不为 0 时,为修改的 Elman 网络.设第 k 步系统的实际输出为 $y_d(k)$,定义误差函数为

$E(k) = \frac{1}{2} (y_d(k) - y(k))^T (y_d(k) - y(k))$, 将 E 对连接权 W^{I3}, W^{I2}, W^{I1} 分别求偏导, 由梯度下降法可得 Elman 网络的学习算法:

$$\Delta w_{ij}^{I3} = \eta_3 \delta_i^0 x_j(k), \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n, \tag{4}$$

$$\Delta w_{jq}^{I2} = \eta_2 \delta_j^h u_q(k-1), \quad j = 1, 2, \dots, n; q = 1, 2, \dots, r, \tag{5}$$

$$\Delta w_{jl}^{I1} = \eta_1 \sum_{i=1}^m (\delta_i^0 w_{ij}^{I3}) \frac{\partial x_j(k)}{\partial w_{jl}^{I1}}, \quad j = 1, 2, \dots, n; l = 1, 2, \dots, n, \tag{6}$$

其中 η_1, η_2, η_3 分别是 W^{I1}, W^{I2}, W^{I3} 的学习步长.

$$\delta_i^0 = (y_{d,i}(k) - y_i(k)), \tag{7}$$

$$\delta_j^h = \sum_{i=1}^m (\delta_i^0 w_{ij}^{I3}) f'_j(\cdot), \tag{8}$$

$$\frac{\partial x_j(k)}{\partial w_{jl}^{I1}} = f'_j(\cdot) x_l(k-1) + \alpha \frac{\partial x_j(k-1)}{\partial w_{jl}^{I1}}. \tag{9}$$

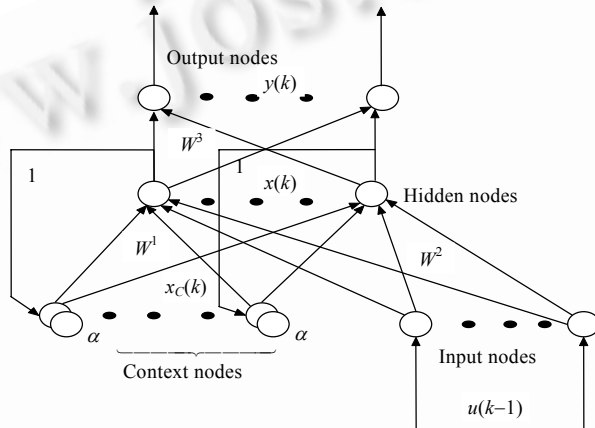


Fig.1 Architecture of the Elman network

图 1 Elman 网络结构示意图

1.2 改进的Elman网络

在 Elman 网络模型中,只计入了隐层节点的反馈,而没有考虑输出层节点的反馈.因为各层神经元的反馈信息都会影响网络的信号处理能力,所以我们考虑增加输出层节点的反馈,本文给出了两个改进 Elman 网络模型,都计入了输出节点的反馈,分别称为具有输出-输入反馈机制的改进 Elman 网络和具有输出-隐层反馈机制的改进 Elman 网络,OIF Elman 网络和 OHF Elman 网络结构示意图分别如图 2 和图 3 所示.其中 OIF Elman 网络增加了输出节点的反馈,称之为联系单元 2,并将它放在第 1 层,与输入单元和联系单元一起作为隐层节点的输入, $0 \leq \gamma < 1$ 为它的自连接反馈增益因子,它的权值记为 W^{I4} . 而 OHF Elman 网络只是将联系单元 2 放在第 2 层上,与隐层节点一起作为输出层的输入,它的权值记为 W^{I5} . 下面我们分别给出它们的数学模型和学习公式.

OIF Elman 网络的数学模型为

$$x(k) = f(W^{I1} x_c(k) + W^{I2} u(k-1) + W^{I4} y_c(k)), \tag{10}$$

$$x_c(k) = \alpha x_c(k-1) + x(k-1), \tag{11}$$

$$y_c(k) = \gamma y_c(k-1) + y(k-1), \tag{12}$$

$$y(k) = W^{I3} x(k). \tag{13}$$

OHF Elman 网络的数学模型为

$$x(k) = f(W^{I1} x_c(k) + W^{I2} u(k-1)), \tag{14}$$

$$x_c(k) = \alpha x_c(k-1) + x(k-1), \tag{15}$$

$$y_c(k) = \gamma y_c(k-1) + y(k-1), \tag{16}$$

$$y(k) = W^{13}x(k) + W^{15}y_c(k). \tag{17}$$

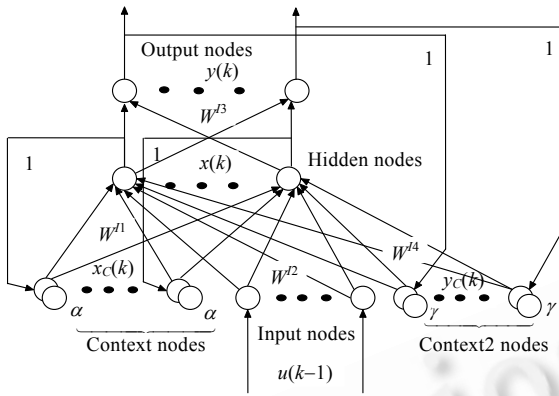


Fig.2 Architecture of the OIF Elman network
图 2 OIF Elman 网络结构示意图

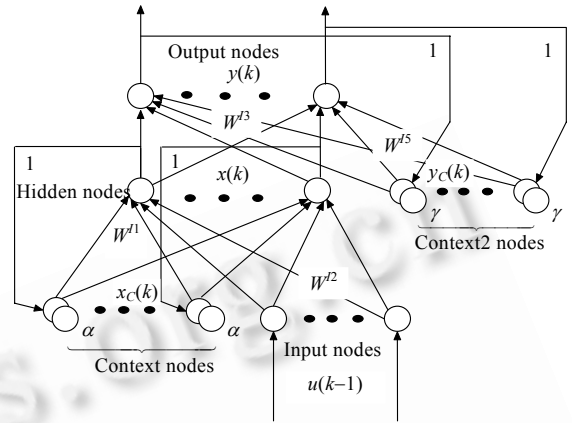


Fig.3 Architecture of the OHF Elman network
图 3 OHF Elman 网络结构示意图

OIF Elman 网络的学习算法:

OIF Elman 网络只是在 Elman 网络的隐层节点中增加了联系单元 2 作为输入,所以对权值 W^{11}, W^{12}, W^{13} 的修正公式与 Elman 网络的相同,分别由上一节中的式(6)、式(5)和式(4)给出,对于权值 W^{14} 的修正公式其推导和 Elman 网络类似,直接给出如下:

$$\Delta w_{js}^{14} = \eta_4 \sum_{i=1}^m (\delta_i^0 w_{ij}^{13}) \frac{\partial x_j(k)}{\partial w_{js}^{14}}, \quad j = 1, 2, \dots, n; s = 1, 2, \dots, m, \tag{18}$$

$$\frac{\partial x_j(k)}{\partial w_{js}^{14}} = f'_j(\cdot) y_s(k-1) + \gamma \frac{\partial x_j(k-1)}{\partial w_{js}^{14}}, \quad j = 1, 2, \dots, n; s = 1, 2, \dots, m, \tag{19}$$

其中 η_4 是 W^{14} 的学习步长, δ_i^0 由式(7)给定.

同样,对于 OHF Elman 网络, W^{11}, W^{12}, W^{13} 的权值修正公式也与 Elman 网络相同,而 W^{15} 的权值修正公式为

$$\Delta w_{ih}^{15} = \eta_5 \delta_i^0 y_{c,h}(k), \quad i = 1, 2, \dots, m; h = 1, 2, \dots, m, \tag{20}$$

其中 η_5 是 W^{15} 的学习步长, δ_i^0 由式(7)给出.

1.3 Elman及其改进网络的稳定性

学习率的选择对神经网络的运行性能至关重要.太小的学习率会使网络的运行太慢,太大的学习率则会使网络的运行不稳定^[16].为了高效地训练神经网络,我们从离散型的 Lyapunov 稳定性方法出发,给出可以保证神经网络稳定收敛的自适应学习率.我们首先以 OIF Elman 网络为例,讨论它的收敛性.

定理 1. 设 OIF Elman 的权值修正由上面的学习算法确定,则对于该网络的收敛性,有如下结论:

(a) 若 $0 < \eta_3(k) < \frac{2}{n}$, 则权值 W^{13} 的迭代学习过程(4)稳定收敛;

(b) 若 $0 < \eta_2(k) < 8 \left[nr \left| \max_k u_k(k) \right| \left| \max_{ij} (W_{ij}^{13}(k)) \right| \right]^{-1}$, 则权值 W^{12} 的迭代学习过程(5)稳定收敛;

(c) 若 $0 < \eta_1(k) < 32(1-\alpha)^2 \left[n^2 \left| \max_{ij} (W_{ij}^{13}(k)) \right|^2 \right]^{-1}$, 则权值 W^{11} 的迭代学习过程(6)稳定收敛;

(d) 若 $0 < \eta_4(k) < 32 \left[mn \cdot \left| \max_{ij} (W_{ij}^{I3}(k)) \right|^2 \left| \max_s y_{c,s}(k) \right|^2 \right]^{-1}$, 则权值 W^{I4} 的迭代学习过程(18,19)稳定收敛.

证明:由于篇幅所限,我们只证明(a)和(c),而(b)和(d)的证明类似.首先证明(a):定义 Lyapunov 能量函数为 $E(k) = \frac{1}{2} \sum_{i=1}^m e_i^2(k)$, 其中 $e_i(k) = y_{d,i}(k) - y_i(k)$, 于是 $\Delta E(k) = E(k+1) - E(k) = \frac{1}{2} \sum_{i=1}^m [e_i^2(k+1) - e_i^2(k)]$, 学习过程的误差可以表示为

$$e_i(k+1) = e_i(k) + \sum_{j=1}^n \frac{\partial e_i(k)}{\partial W_{ij}^{I3}} \Delta W_{ij}^{I3} = e_i(k) - \sum_{j=1}^n \frac{\partial y_i(k)}{\partial W_{ij}^{I3}} \Delta W_{ij}^{I3}. \tag{21}$$

所以

$$\Delta E(k) = \frac{1}{2} \sum_{i=1}^m e_i^2(k) \left[\left(1 - \eta_3(k) \left[\frac{\partial y_i(k)}{\partial W_i^{I3}} \right]^T \left[\frac{\partial y_i(k)}{\partial W_i^{I3}} \right] \right)^2 - 1 \right] = \frac{1}{2} \sum_{i=1}^m e_i^2(k) \left[\left(1 - \eta_3(k) \left\| \frac{\partial y_i(k)}{\partial W_i^{I3}} \right\|^2 \right)^2 - 1 \right], \tag{22}$$

其中 W_i^{I3} 表示一维向量, $\|\cdot\|$ 表示欧式范数.注意到 Elman 网络的隐层单元采用 sigmoid 函数,所以 $\left| \frac{\partial y_i(k)}{\partial W_{ij}^{I3}} \right| = |x_j(k)| < 1, i=1,2,\dots,m; j=1,2,\dots,n$, 按照欧氏范数的定义,有 $\left\| \frac{\partial y_i(k)}{\partial W_i^{I3}} \right\| < \sqrt{n}$, 从而当 $0 < \eta_3(k) < \frac{2}{n}$ 时,由式(22),有 $\Delta E(k) < 0$. 又 $E(k) > 0$, 根据离散型的 Lyapunov 稳定性方法,当 $t \rightarrow \infty$ 时,训练误差将收敛到 0,(a)证毕.

下面证明(c),定义 $Q_{jl}(k) = \frac{\partial x_j(k)}{\partial W_{jl}^{II}}, Q_{jl}(0) = 0, j=1,2,\dots,n; l=1,2,\dots,n$, 于是由式(9)有

$$Q_{jl}(k) \approx f'_j(k)x_l(k) + \alpha Q_{jl}(k-1), j=1,2,\dots,n; l=1,2,\dots,n,$$

将 $Q_{jl}(k-1), Q_{jl}(k-2), \dots$, 按照上式展开可得到

$$Q_{jl}(k) \approx \sum_{p=1}^k \alpha^{p-1} f'_j(k-p+1)x_l(k-p+1), j=1,2,\dots,n; l=1,2,\dots,n. \tag{23}$$

注意到 $0 < f'_j(\cdot) \leq 1/4, 0 \leq \alpha < 1, 0 < x_l(k-p+1) < 1$, 所以

$$\left| Q_{jl}(k) \right| \leq \sum_{p=1}^k \alpha^{p-1} \|f'_j(k-p+1)\| |x_l(k-p+1)| < \frac{1}{4} \sum_{p=1}^k \alpha^{p-1} < \frac{1}{4(1-\alpha)}, j=1,2,\dots,n; l=1,2,\dots,n. \tag{24}$$

类似式(22),有

$$\Delta E(k) = \frac{1}{2} \sum_{i=1}^m e_i^2(k) \left[\left(1 - \eta_1(k) \left[\frac{\partial y_i(k)}{\partial W^{II}} \right]^T \left[\frac{\partial y_i(k)}{\partial W^{II}} \right] \right)^2 - 1 \right] = \frac{1}{2} \sum_{i=1}^m e_i^2(k) \left[\left(1 - \eta_1(k) \left\| \frac{\partial y_i(k)}{\partial W^{II}} \right\|^2 \right)^2 - 1 \right]. \tag{25}$$

此处把 W^{II} 看成一个 $n \times n$ 维向量.注意到

$$\left| \frac{\partial y_i(k)}{\partial W_{jl}^{II}} \right| = \left| \frac{\partial y_i(k)}{\partial x_j(k)} \frac{\partial x_j(k)}{\partial W_{jl}^{II}} \right| = \left| W_{ij}^{I3}(k) Q_{jl}(k) \right| \leq \frac{1}{4(1-\alpha)} \left| \max_{ij} (W_{ij}^{I3}(k)) \right|, i=1,2,\dots,n; j=1,2,\dots,n, \tag{26}$$

所以 $\left\| \frac{\partial y_i(k)}{\partial W^{II}} \right\| \leq \frac{n}{4(1-\alpha)} \left| \max_{ij} (W_{ij}^{I3}(k)) \right|, i=1,2,\dots,n; j=1,2,\dots,n$, 从而当 $0 < \eta_1(k) < \frac{32(1-\alpha)^2}{n^2 \left| \max_{ij} (W_{ij}^{I3}(k)) \right|^2}$ 时,由式(25),有

$\Delta E(k) < 0$. 根据离散型的 Lyapunov 稳定性方法,当 $t \rightarrow \infty$ 时训练误差将收敛到 0,(c)证毕. □

由上述定理可知,当 $\eta_3(k)$ 满足 $0 < \eta_3(k) < \frac{2}{n}$ 时,可以保证网络训练收敛,但由式(22)可知,当

$\left(1 - \eta_3(k) \left\| \frac{\partial y_i(k)}{\partial W_i^{I3}} \right\|^2 \right)^2 = 0$ 时, $\Delta E(k)$ 最小,即网络以最快速度收敛.若近似地认为 $\left\| \frac{\partial y_i(k)}{\partial W_i^{I3}} \right\| = \sqrt{n}$, 则取 $\eta_3(k) =$

$\hat{\eta}_3(k) = \frac{1}{n}$, 可以保证网络以较快速度收敛.同理,分别取 $\hat{\eta}_2(k) = 4 \left[nr \left| \max_k u_k(k) \right| \left| \max_{ij} (W_{ij}^{I3}(k)) \right| \right]^{-1}$, $\hat{\eta}_1(k) =$

$16(1-\alpha)^2 \left[n^2 \left| \max_i(W_i^{13}(k)) \right|^2 \right]^{-1}$, $\hat{\eta}_4(k) = 16 \left[mn \left| \max_{ij}(W_{ij}^3(k)) \right|^2 \left| y_{C,s}(k) \right|^2 \right]^{-1}$,可以保证网络以较快速度收敛.

对于 Elman 网络,相当于 $W^{14} = 0$,而 W^{11} , W^{12} 和 W^{13} 都与 OIF Elman 网络具有相同的结论.对于 OHF Elman 网络的 W^{11} , W^{12} 和 W^{13} ,仍然可以得到与 OIF Elman 网络相同的结论.对于 W^{15} ,与 OIF Elman 网络类似,相应于式(25),可以得到下面的式(27).

$$\Delta E(k) = \frac{1}{2} \sum_{i=1}^m e_i^2(k) \left[\left(1 - \eta_5(k) \left[\frac{\partial y_i(k)}{\partial W^{15}} \right]^T \left[\frac{\partial y_i(k)}{\partial W^{15}} \right] \right)^2 - 1 \right] = \frac{1}{2} \sum_{i=1}^m e_i^2(k) \left[\left(1 - \eta_5(k) \left\| \frac{\partial y_i(k)}{\partial W^{15}} \right\|^2 \right)^2 - 1 \right]. \quad (27)$$

而在 OHF Elman 网络中,

$$\left| \frac{\partial y_{i'}(k)}{\partial W_{ih}^{15}} \right| = \left| \delta_{i'} y_{C,h}(k) \right| = \delta_{i'} \left| y_{C,h}(k) \right| \leq \delta_{i'} \left| \max_h(y_{C,h}(k)) \right|, \quad i' = 1, 2, \dots, m; i = 1, 2, \dots, m; h = 1, 2, \dots, m,$$

其中 $\delta_{i'} = \begin{cases} 1, & i = i' \\ 0, & i \neq i' \end{cases}$,而 $\left\| \frac{\partial y(k)}{\partial W^{15}} \right\| \leq \sqrt{m} \left| \max_h(y_{C,h}(k)) \right|$,由式(27),当 $0 < \eta_5(k) < 2 \left[m \left| \max_h(y_{C,h}(k)) \right|^2 \right]^{-1}$ 时,有 $\Delta E(k) < 0$,注意到 $E(k) > 0$,根据离散型的 Lyapunov 稳定性方法,当 $t \rightarrow \infty$ 时训练误差将收敛到 0.经过与 OIF Elman 网络类似的讨论可知,当 $\eta_5(k)$ 取 $\hat{\eta}_5(k) = \left[m \left| \max_h(y_{C,h}(k)) \right|^2 \right]^{-1}$ 时,网络以较快速度收敛.

2 基于 RBPC 神经网络模型的超声马达控制

2.1 RBPC神经网络的结构及学习算法

在本节中,我们针对超声马达建立了一个神经网络控制器,即递归反向传播控制 RBPC 神经网络模型.图 4 表示了对超声马达的整个控制结构.NNI 为神经网络辨识器,采用的是上一节提出的 OIF Elman 神经网络,它的作用是提供误差反传通道.NNI 的输入是 NNC 的输出 $u(k)$,NNI 的输出是 $\hat{y}(k)$,在控制过程中,按上一节的学习算法,经在线学习获得 $\hat{y}(k)$,训练误差为 $\hat{e}(k) = y(k) - \hat{y}(k)$.NNC 为神经网络控制器,它只有一个输入,即 $e_c(k-1)$,为上一次的系统误差.有 3 个隐单元,采用 sigmoid 函数;一个输出单元,为 USM 的控制参数 $u(k)$.在学习过程中,当考虑误差对权的导数时,要考虑输入即上一次的系统误差对权的导数,在此意义下,我们称其为递归反传(recurrent back propagation).NNC 隐层节点的输出为 $v_i(k) = f(e(k-1)w_i^{c1}(k))$,输出层节点的输出为 $u(k) = \sum_{i=1}^3 v_i(k)w_i^{c2}(k)$.其中 w^{c1} , w^{c2} 分别为第 1 层和第 2 层的权值; v, u 分别为隐层和输出层的输出. u 作为 USM 的输入, $y(k)$ 为 USM 的输出.设 USM 的期望输出为 $y_d(k)$,令其与 USM 的实际输出之差为 $e_c(k)$,即 $e_c(k) = y_d(k) - y(k)$.定义系统误差函数为

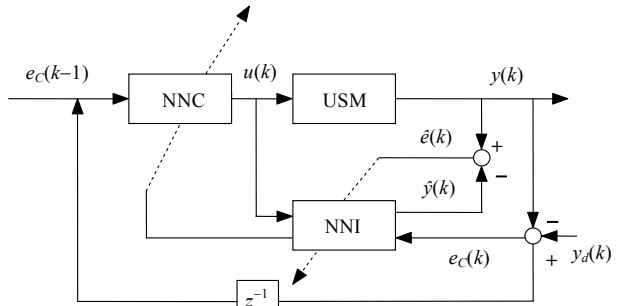


Fig.4 Speed control system
图 4 速度控制系统

在控制的过程中,要求始终保持 $\Delta E < 0$.注意到, $\Delta E = \frac{\partial E}{\partial w^{c1}} \Delta w^{c1} + \frac{\partial E}{\partial w^{c2}} \Delta w^{c2}$,考虑 E 对 w^{c1} , w^{c2} 的偏导数,为简明起见,以 w^c 代表 w^{c1} 和 w^{c2} ,即 $\frac{\partial E}{\partial w^c}$ 代表 $\frac{\partial E}{\partial w^{c1}}$ 和 $\frac{\partial E}{\partial w^{c2}}$.因为输入 $e_c(k-1)$ 也是 w^c 的函数,所以这是一个递归过程.具体分析如下:

$$\frac{\partial E}{\partial w^c} = -e(k) \frac{\partial y(k)}{\partial w^c} = -e_c(k) \frac{\partial y(k)}{\partial u} \frac{\partial u(k)}{\partial w^c} \approx -e_c(k) \frac{\partial \hat{y}(k)}{\partial u} \frac{\partial u(k)}{\partial w^c}, \quad (28)$$

$$\hat{y}(k) = \sum_{i=1}^n x_i w_i^{I3} = \sum_{i=1}^n f \left(\sum_{j=1}^n x_{C,j}(k) w_{ij}^{I1} + u(k) w_i^{I2} + y_C(k) w_i^{I4} \right) w_i^{I3}, \quad (29)$$

$$\frac{\partial \hat{y}(k)}{\partial u} = \sum_{i=1}^n (x_i (1 - x_i) w_i^{I3} w_i^{I2} + \frac{\partial y_C(k)}{\partial u} w_i^{I4}), \quad (30)$$

其中

$$\frac{\partial y_C(k)}{\partial u} = \gamma \frac{\partial y_C(k-1)}{\partial u} + \frac{\partial \hat{y}(k-1)}{\partial u}, \quad (31)$$

令

$$\frac{\partial y_C(0)}{\partial u} = 0; \frac{\partial \hat{y}(0)}{\partial u} = 0, \quad (32)$$

则 $\frac{\partial y_C(k)}{\partial u}$ 可由式(30)~式(32)递推得到.

注意到 $\frac{\partial e_C(k-1)}{\partial w_i^{C2}} \approx -\frac{\partial \hat{y}(k-1)}{\partial u} \frac{\partial u(k-1)}{\partial w_i^{C2}}$, 于是有

$$\frac{\partial u(k)}{\partial w_i^{C2}} \approx v_i(k) - \sum_{j=1}^3 v_j(k) (1 - v_j(k)) w_j^{C1} \frac{\partial \hat{y}(k-1)}{\partial u} \frac{\partial u(k-1)}{\partial w_i^{C2}} w_j^{C2}. \quad (33)$$

同理可得

$$\frac{\partial u(k)}{\partial w_i^{C1}} = e_C(k) v_i(k) (1 - v_i(k)) w_i^{C2} - \sum_{j=1}^3 v_j(k) (1 - v_j(k)) w_j^{C2} \frac{\partial \hat{y}(k-1)}{\partial u} \frac{\partial u(k-1)}{\partial w_i^{C1}} w_j^{C1}. \quad (34)$$

欲使 $\Delta E < 0$, 只须令 $\Delta w^C = -\eta^C \frac{\partial E}{\partial w^C}$, 这里, η^C 为学习步长. 总结起来, 可以给出 RBPC 的学习算法为

$$w^C(k+1) = w^C(k) + \Delta w^C(k), \quad \Delta w^{C1}(k) = -\eta^{C1} \frac{\partial E}{\partial w^{C1}}, \quad \Delta w^{C2}(k) = -\eta^{C2} \frac{\partial E}{\partial w^{C2}}, \quad (35)$$

其中 η^{C1} , η^{C2} 分别为 w^{C1} , w^{C2} 的学习步长, $\frac{\partial E}{\partial w^{C2}}$ 由式(28)、式(30)~式(33)给出, $\frac{\partial E}{\partial w^{C1}}$ 由式(28)、式(30)~式(32)、式(34)给出.

2.2 RBPC网络的稳定性

仿照上一节的稳定性证明, 下面讨论 RBPC 网络的稳定性. 我们首先讨论 w^{C2} 学习算法的稳定性. 由

$$e_C(k+1) = e_C(k) + \sum_{i=1}^3 \frac{\partial e_C(k)}{\partial w_i^{C2}} \Delta w_i^{C2} = e_C(k) - \sum_{i=1}^3 \frac{\partial y(k)}{\partial w_i^{C2}} (\eta^{C2} e_C(k) \frac{\partial y_C(k)}{\partial w_i^{C2}}) = e_C(k) \left[1 - \eta^{C2} \left(\frac{\partial \hat{y}(k)}{\partial u} \right)^2 \left\| \frac{\partial u(k)}{\partial w^{C2}} \right\|^2 \right],$$
 有

$$\Delta E(k) = \frac{1}{2} e_C^2(k+1) - \frac{1}{2} e_C^2(k) = \frac{1}{2} e_C^2(k) \left[\left(1 - \eta^{C2} \left(\frac{\partial \hat{y}(k)}{\partial u} \right)^2 \left\| \frac{\partial u(k)}{\partial w^{C2}} \right\|^2 \right)^2 - 1 \right], \quad (36)$$

易见, 当 $0 < \eta^{C2} < 2 \left[\left(\frac{\partial \hat{y}(k)}{\partial u} \right)^2 \left\| \frac{\partial u(k)}{\partial w^{C2}} \right\|^2 \right]^{-1}$ 时, $\Delta E(k) < 0$, 根据离散型的 Lyapunov 稳定性方法, 当 $t \rightarrow \infty$ 时训练误差

将收敛到 0. 同理, 当 $0 < \eta^{C1} < 2 \left[\left(\frac{\partial \hat{y}(k)}{\partial u} \right)^2 \left\| \frac{\partial u(k)}{\partial w^{C1}} \right\|^2 \right]^{-1}$ 时, 能够保证学习算法中 w^{C1} 的收敛性. 其中 $\frac{\partial \hat{y}(k)}{\partial u}$, $\frac{\partial u(k)}{\partial w^{C2}}$,

$\frac{\partial u(k)}{\partial w^{C1}}$ 分别由式(30)、式(33)、式(34)给出. 由式(36)易见, 当 $\eta^{C2} = \hat{\eta}^{C2} = \left[\left(\frac{\partial \hat{y}(k)}{\partial u} \right)^2 \left\| \frac{\partial u(k)}{\partial w^{C2}} \right\|^2 \right]^{-1}$ 时, $\Delta E(k)$ 最小, 即能

保证以最快速度收敛. 同理, 当 $\eta^{C1} = \hat{\eta}^{C1} = \left[\left(\frac{\partial \hat{y}(k)}{\partial u} \right)^2 \left\| \frac{\partial u(k)}{\partial w^{C1}} \right\|^2 \right]^{-1}$ 时, $\Delta E(k)$ 最小, 能够保证以最快速度收敛.

3 数值模拟结果与讨论

根据本文的方法,我们对纵向振动超声马达^[17,18]进行了辨识和速度控制的数值模拟.马达的具体参数为:驱动频率为 27.8(KHZ),驱动电压为 300(V),额定输出力矩为 2.5(kg.cm),旋转速度为 3.8(m/s).NNI 中的增益因子 α 和 γ 都取 0.4.

3.1 Elman及其改进网络对超声马达的模拟

我们分别用 Elman 网络及其两种改进的 Elman 网络,即 OIF Elman 网络和 OHF Elman 网络对超声马达模型进行了模拟,并且进行了比较,模拟的结构图如图 5 所示.

图 6 是这几种模型在每秒取样 1 000 次情况下的模拟曲线,图 7 是这几种模型在每秒取样 300 次情况下的模拟曲线.为考察它们的鲁棒性,在 0.4 秒时对马达施加了一个随机电压扰动.考察模拟计算结果,可以看到这样的现象:当采样次数较高时,Elman 模型优于 Elman 的改进网络模型;而当采样次数较低时,Elman 的改进网络模型优于 Elman 模型.

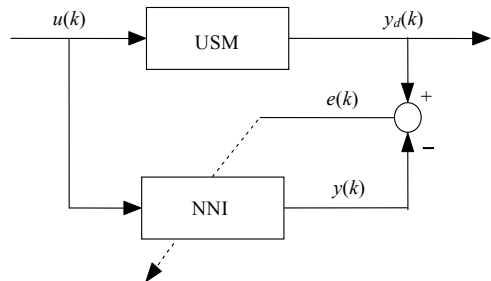


Fig.5 Block diagram of the identification model

图 5 辨识模型的结构图

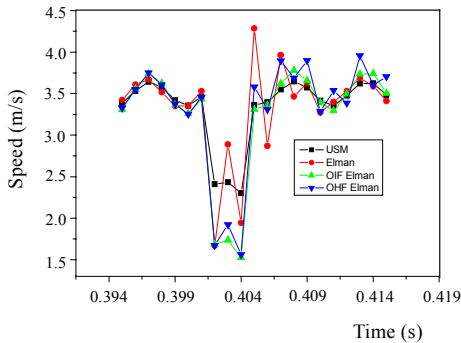


Fig.6 Distributed simulation speed curves for samples of 1000 times/s

图 6 取样 1000 次/秒时扰动的模拟曲线

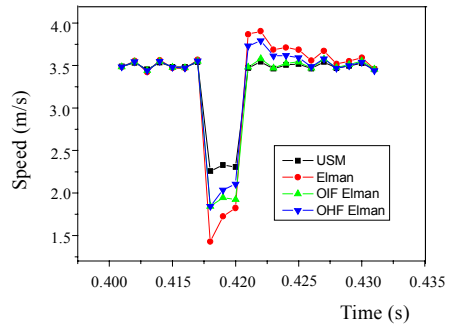


Fig.7 Distributed simulation speed curves for samples of 300 times/s

图 7 取样 300 次/秒时扰动的模拟曲线

例如,在采样次数为每秒取样 1000 次时,Elman 模型比它的两种改进网络模型效果要好,它们的误差幅度分别为 Elman:0.123,OIF Elman:0.154,OHF Elman:0.149;而当每秒采样 300 次时,正好相反,Elman 网络的两种改进的结果比原模型有了显著的改善,它们的误差幅度分别为 Elman:0.252,OIF Elman:0.098,OHF Elman:0.141.其中误差幅度定义为稳定之后的最大误差减去最小误差.通过较多的模拟实验观察到,当采样次数高于 1 000 次时,Elman 模型的效果都要好于它的改进网络模型,当采样次数低于 300 次时,Elman 的改进网络模型优于 Elman 模型.它们在采样次数为 600 次时效果大致相同.这个观察结果使我们可以根据现场实际采样的情况来决定选用什么样的网络模型.

3.2 RBPC网络对超声马达的控制

图 8 表示取控制值为 3.6m/s,外加力矩为 1N.cm 时,采用两种不同的控制方案得到的马达的速度控制曲线,其中虚线表示采用文献[6]的方法得到的速度控制曲线,实线表示采用本文方法得到的速度控制曲线.数值模拟结果表明,采用两种方案虽然都可以获得相对稳定的速度控制曲线,但总是带有小的波动.由图 8 可见,采用本文方法所获得的速度曲线,其波动的幅度明显减小.我们定义速度的波动为 $\zeta = (V_{max} - V_{min}) / V_{aver} \times 100\%$,其中

V_{\max} , V_{\min} , V_{aver} 分别表示速度的最大值、最小值和平均值.由图 8 可见,采用文献[6]方法所得到的速度波动较大,其波动平均值达到 5.7%.而采用本文方法得到的波动平均值仅为 1.9%,与文献[6]相比,精度可以提高 2 倍左右.为了进一步验证本文方法对不同参考速度的控制效果,我们考察了参考速度值随时间变化的速度控制曲线.图 9 和图 10 分别显示了对于参考速度按阶梯状变化和按正弦规律变化两种情形的速度控制曲线.从图中可见,本文给出的速度控制方案具有很好的控制精度.

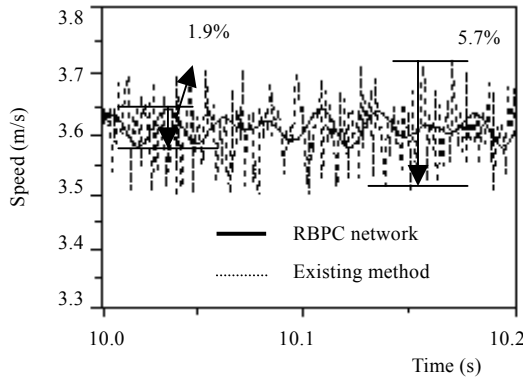


Fig.8 Motor speed control curves using different schemes
图 8 不同控制方案下的马达速度曲线

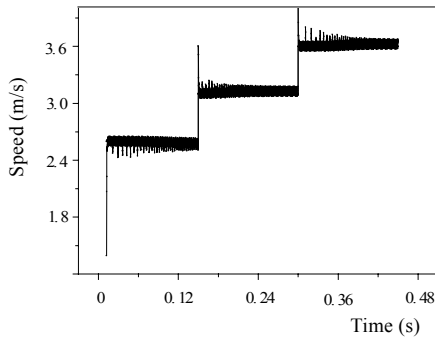


Fig.9 Speed curve for step-type control
图 9 控制值按阶梯变化时的速度曲线

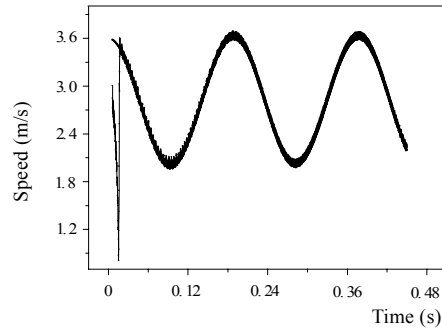


Fig.10 Speed curve for sinusoidal-type control
图 10 控制值按正弦变化时的速度曲线

4 结 论

本文首先在 Elman 网络的基础上提出了两种增加了输出层反馈的改进的 Elman 神经网络模型.通过采用这 3 种网络分别对超声马达进行了模拟,取得了较好的模拟效果,并且得到了下面的结论:当采样次数较高时,Elman 网络好于它的改进型;反之,当采样次数较低时,改进的 Elman 网络好于 Elman 网络.本文用 OIF Elman 网络作为误差反传的通道,建立了 RBPC 神经网络控制器,对超声马达进行了控制,对于常量、阶梯型和正弦型的参考速度,都可以得到很好的控制效果.与文献[6]提供的方法相比,本文的方法使速度波动的幅度值大为减小,提高了控制的精度.通过理论分析,分别对改进的 Elman 网络以及 RBPC 网络在 Lyapunov 稳定性意义下给出了控制网络自适应学习率的适用范围,从而保证了它们的稳定性.

References:

- [1] Sun ZQ, Zhang ZX, Deng ZD. Theory and Technology of Intelligent Control. Beijing: Tsinghua University Press; Guangxi University of Science and Technology Press, 1997 (in Chinese).
- [2] Zhang NY, Yan PF. Neural Networks and Fuzzy Control. Beijing: Tsinghua University Press, 1998 (in Chinese).

- [3] Senjyu T, Yokoda S, Uezato K. A study on high-efficiency drive of ultrasonic motors. *Electric Power Components and Systems*, 2001,29(3):179~189.
- [4] Uehino K. Piezoelectric motors: Overview. *Smart Materials and Structures*, 1998,7(3):273~285.
- [5] Hagood NW, McFarland AJ. Modeling of a piezoelectric rotary ultrasonic motor. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 1995,42(2):210~224.
- [6] Senjyu T, Miyazato H, Yokoda S, Uezato K. Speed control of ultrasonic motors using neural network. *IEEE Transactions on Power Electronics*, 1998,13(3):381~387.
- [7] Lin FJ, Wai RJ, Shyu KK, Liu TM. Recurrent fuzzy neural network control for piezoelectric ceramic linear ultrasonic motor drive. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 2001,48(4):900~913.
- [8] Senjyu T, Yokoda S, Uezato K. Speed control of ultrasonic motors using fuzzy neural network. *Journal of Intelligent Fuzzy System*, 2000,8(2):135~146.
- [9] Lin FJ, Wai RJ, Hong CM. Recurrent neural network control for LCC-resonant ultrasonic motor drive. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 2000,47(3):737~749.
- [10] Elman JL. Finding structure in time. *Cognitive Science*, 1990,14(2):179~211.
- [11] Pham DT, Liu X. Dynamic system modeling using partially recurrent neural networks. *Journal of Systems Engineering*, 1992,2(2):90~97.
- [12] Pham DT, Liu X. Training of Elman networks and dynamic system modeling. *International Journal of Systems Science*, 1996, 27(2):221~226.
- [13] Pham DT, Karaboga D. Training Elman and Jordan networks for system identification using genetic algorithms. *Artificial Intelligence in Engineering*, 1999,13(2):107~117.
- [14] Li X, Chen ZQ, Yuan ZZ, Chen GR. Generating chaos by an Elman network. *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications*, 2001,48(9):1126~1131.
- [15] Geqay R, Liu T. Nonlinear modelling and prediction with feedforward and recurrent networks. *Physica D*, 1997,108(1-2):119~134.
- [16] Zhou CG, Liang YC. *Computational Intelligence*. Changchun: Jilin University Press, 2001 (in Chinese).
- [17] Xu X, Liang YC, Shi XH. Analysis of frequency-temperature characteristics of ultrasonic motors. *Journal of Jilin University (Science Edition)*, 2002,40(2):109~113 (in Chinese with English abstract).
- [18] Sashida T. Trial construction and operation of an ultrasonic vibration driven motor. *Applied Physics*, 1982,51(6):713~720.

附中文参考文献:

- [1] 孙增圻,张再兴,邓志东.智能控制理论与技术.北京:清华大学出版社;广西科技大学出版社,1997.
- [2] 张乃尧,阎平凡.神经网络与模糊控制.北京:清华大学出版社,1998.
- [16] 周春光,梁艳春.计算智能.长春:吉林大学出版社,2001.
- [17] 徐旭,梁艳春,时小虎.超声马达的频率温度特性分析.吉林大学学报(理学版),2002,40(2):109~113.