

连续小波过程神经网络及其仿真研究

李 洋, 钟诗胜, 张 艳

(哈尔滨工业大学 机电工程学院, 黑龙江 哈尔滨 150001)

摘 要:在小波分析和过程神经网络理论的基础上,提出了连续小波过程神经网络模型,其隐层为过程神经元,隐层激活函数采用小波函数.该网络结合了小波变换良好的时—频局域化性质及过程神经网络可以处理连续输入信号的特点,因而学习能力强,精度高.给出了小波过程神经网络学习算法,并以航空发动机滑油系统状态监测为例,分别利用传统BP网络和小波过程神经网络进行预测.结果表明,小波过程神经网络收敛速度快,精度高,优于BP网络的预测能力,同时也为航空发动机滑油系统状态监测问题提供了一种有效的方法.

关键词:过程神经元;连续小波过程神经网络;学习算法;发动机滑油监测

中图分类号:TP183 **文献标识码:**A **文章编号:**1673-4785(2007)06-0077-05

Research and simulation of a continuous wavelet process neural network

LI Yang, ZHONG Shi-sheng, ZHANG Yan

(Department of Mechanical and Electrical Engineering, Harbin Institute of Technology, Harbin 150001, China)

Abstract: A continuous wavelet process neural networks (CWPNN) model is proposed based on wavelet analysis and process neural network theory in which the hidden layer is composed of process neurons and the hidden layer activation function is a wavelets function. This network not only has excellent learning capability but also high accuracy since it combines the time-frequency local property of wavelet analysis with the ability of process neural networks to deal with continuous input signals. The corresponding learning algorithm is given. The network was compared with a traditional BP neural network while monitoring the state of an aeroengine lubricating oil system. The results demonstrated the faster convergence rate and higher accuracy of the network. Prediction capability was superior to a BP neural network, providing a more effective method for state monitoring of aeroengine lubricating oil systems.

Key words: process neuron; continuous wavelet process neural networks; learning algorithm; monitoring of aeroengine lubricating oil system

在实际问题中,许多系统的输入是依赖于时间变化的连续函数,其输出值既依赖于输入函数的空间聚合,又与时间的累计效应密切相关.针对这一问题,文献[1]中提出了过程神经网络模型.过程神经元与传统神经元不同之处在于过程神经元的输入和权值可以是时变的,其聚合运算中既有对空间的多输入聚合,也有对时间过程的积累.过程神经网络是人工神经网络在时间域上的扩展,放宽了传统神经网络模型对输入的同步瞬时限制,因此,对

于求解与过程有关的众多实际问题有着广泛的应用价值^[2].但是在神经网络的应用过程中,网络参数的优化和结构的构造等缺乏有效的理论性指导方法,而这些都将直接影响网络的实际应用效果.而小波分析理论为神经网络的构造和参数的选取提供了一个行之有效的方法.本文将过程神经网络和小波分析理论相结合,在文献[1-4]的基础上,提出了连续小波过程神经网络模型.其输入和权值是时变的,其聚合运算既有对空间的多输入聚合,也有对时间过程的积累.连续小波过程神经网络隐层神经元的激励函数采用小波基函数,隐层阈值由小波的平移与伸缩参数替换.连续小波过程神经网络结合了小波

收稿日期:2007-01-18.

基金项目:国家自然科学基金资助项目(60572174).

变换良好的时频局域化性质及过程神经网络可以处理时变信号的能力,从而避免了传统神经网络在解决大容量非线性时变系统的信号处理问题时的不适应性.由于连续小波过程神经网络引入了平移和伸缩因子,因此具有更强的函数逼近能力.

1 连续小波过程神经网络

1.1 过程神经元

过程神经元是由过程输入信号加权,时间、空间二维聚合和阈值激励输出 4 部分运算组成.与传统神经元不同之处在于过程神经元的输入和连接权可以是时变函数,过程神经元增加了一个对于时间的聚合算子,其聚合运算既有对输入信号的空间加权聚集,亦有对时间过程效应的累积.过程神经元模型的结构如图 1 所示.

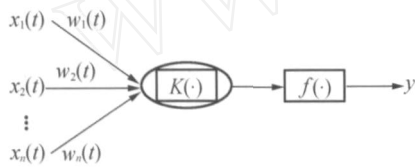


图 1 过程神经元结构图

Fig. 1 Structure of process neuron

过程神经元输入与输出之间的映射关系为

$$y = f((W(t) \oplus X(t)) \odot K(\cdot)). \quad (1)$$

式中: $X(t) = (x_1(t), x_2(t), \dots, x_n(t))$ 为过程神经元的输入矢量, $W(t) = (w_1(t), w_2(t), \dots, w_n(t))$ 为相应的连接权矢量, \oplus 表示某种空间聚合运算, 可取

为 $A(t) = W(t) \oplus X(t) = \sum_{i=1}^n x_i(t) w_i(t)$, \odot 表示某种时间聚合运算, 可取为 $A(t) \odot K(\cdot) = \int_0^T A(t) k(\cdot) dt$, 为过程神经元的输出阈值.

1.2 小波分析

函数 $(x) \in L^2(R)$ 若满足如下条件:

$$C = \frac{\int_R |\hat{x}(\omega)|^2 d\omega}{\int_R |\hat{x}(\omega)|^2 d\omega} < \infty, \quad (2)$$

式中: $\hat{x}(\omega)$ 为 (t) 的 Fourier 变换, 则函数 (x) 成为一个基本小波或母小波, 其通过伸缩平移可产生一组小波函数基:

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right). \quad (3)$$

式中: a, b 分别为伸缩、平移尺度因子. 对于任意函数 $f(x) \in L^2(R)$, 其连续小波变换可表示为

$$W_f(a, b) = (f, \psi_{a,b}) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} \psi_{a,b}(x) f(x) dx. \quad (4)$$

若小波满足允许条件, 则连续小波变换存在着逆变换:

$$f(t) = \frac{1}{C} \int_0^{\infty} \frac{da}{a^2} \int_{-\infty}^{+\infty} W_f(a, b) \psi_{a,b}(t) db = \frac{1}{C} \int_0^{\infty} \frac{da}{a^2} \int_{-\infty}^{+\infty} W_f(a, b) \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) db. \quad (5)$$

1.3 连续小波过程神经网络模型

由若干个过程神经元按一定的拓扑结构组成并且隐层激活函数采用小波函数的网络称为连续小波过程神经网络. 本文所提出的网络为 3 层连续小波过程神经网络, 设网络的输入层有 n 个单元, 隐层由 m 个小波过程神经元组成, 输出层只有 1 个单元, 网络拓扑结构为 $n - m - 1$, 如图 2 所示.

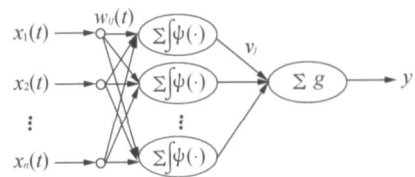


图 2 连续小波过程神经网络

Fig. 2 Continuous wavelet process neural network

连续小波过程神经网络输入与输出之间的关系为

$$y = g\left(\sum_{j=1}^m v_j \int_0^T \frac{\sum_{i=1}^n x_i(t) w_{ij}(t) dt - b_j}{a_j}\right). \quad (6)$$

式中: $X(t) = \{x_1(t), x_2(t), \dots, x_n(t)\}$ 为连续小波过程神经网络的输入函数矢量, $w_{ij}(t)$ 为输入层第 i 个单元与隐层第 j 个过程神经元的连接权函数, v_j 为隐层第 j 个过程神经元与输出层单元的连接权值, 为小波函数, g 为输出层单元的激励函数, 文中取为线性函数, y 为网络输出.

2 连续小波过程神经网络学习算法

2.1 函数正交基展开

由式(6)可以看出, 过程神经网络的时间累积运算比较复杂, 为此引入一组合适的正交基函数将输入函数和连接权函数同时展开, 以达到简化时间累积运算的目的.

设 $b_1(t), b_2(t), \dots, b_k(t), \dots$ 为 $C[0, T]$ 空间上一组标准正交基函数, $X(t) = \{x_1(t), x_2(t), \dots,$

$x_n(t)$ 为输入空间的任意函数,则 $x_i(t)$ 可表示为^[5]

$$x_i(t) = \sum_{k=1}^L c_i^k b_k(t), c_i^k \in R.$$

同理, $w_{ij}(t)$ 也可以由同一组标准正交基函数展开为

$$w_{ij}(t) = \sum_{k=1}^L w_{ij}^k b_k(t), w_{ij}^k \in R.$$

由数学分析理论知,对 $A > 0, \exists L_i$, 满足

$$\sup_{0 \leq t \leq T} \left| x_i(t) - \sum_{l=1}^L c_i^l b_l(t) \right| < A.$$

令 $L = \max\{L_1, L_2, \dots, L_n\}$, 则:

$$\begin{aligned} x_i(t) &= \sum_{k=1}^L c_i^k b_k(t), c_i^k \in R, \\ w_{ij}(t) &= \sum_{l=1}^L w_{ij}^l b_l(t), w_{ij}^l \in R. \end{aligned}$$

根据正交函数的性质,有

$$\int_0^T b_k(t) b_l(t) dt = \begin{cases} 1, & k = l, \\ 0, & k \neq l. \end{cases}$$

因此,式(6)可以转化为

$$y = g \left(\sum_{j=1}^m v_j \left(\sum_{i=1}^n \sum_{k=1}^L \frac{c_i^k w_{ij}^k - b_j}{a_j} \right) \right). \quad (7)$$

2.2 学习过程

给定 P 个学习样本函数: $(x_{p1}(t), x_{p2}(t), \dots, x_{pn}(t), d_p), p = 1, 2, \dots, P$; 其中, $x_{pi}(t)$ 的第 1 个下标表示学习样本序号,第 2 个下标表示输入函数向量分量序号; d_p 为对应于输入 $(x_{p1}(t), x_{p2}(t), \dots, x_{pn}(t))$ 网络的期望输出. 则连续小波过程神经网络的误差函数可表示为

$$\begin{aligned} E &= \frac{1}{2} \sum_{p=1}^P (y_p - d_p)^2 = \\ &= \frac{1}{2} \sum_{p=1}^P \left[g \left(\sum_{j=1}^m v_j \left(\sum_{i=1}^n \sum_{k=1}^L \frac{c_{pi}^k w_{ij}^k - b_j}{a_j} \right) \right) - d_p \right]^2. \end{aligned} \quad (8)$$

根据梯度下降算法,连续小波过程神经网络的待训练参数调整规则为

$$\begin{cases} w_{ij}^k(s+1) = w_{ij}^k(s) + \eta \frac{\partial E}{\partial w_{ij}^k}, \\ v_j(s+1) = v_j(s) + \eta \frac{\partial E}{\partial v_j}, \\ a_j(s+1) = a_j(s) + \eta \frac{\partial E}{\partial a_j}, \\ b_j(s+1) = b_j(s) + \eta \frac{\partial E}{\partial b_j}. \end{cases} \quad (9)$$

式中: η 为学习速率, s 为迭代次数. 为了简便起见, 记 $u = \left(\sum_{i=1}^n \sum_{k=1}^L c_{pi}^k w_{ij}^k - b_j \right) / a_j$, 则 w_{ij}^k, v_j, a_j, b_j 可分别表示为

$$\begin{cases} w_{ij}^k = - \frac{\partial E}{\partial w_{ij}^k} = - (y - d) \left(\sum_{j=1}^m v_j \left(u \right) \right) \sum_{i=1}^n \sum_{k=1}^L \frac{c_{pi}^k}{a_j}, \\ v_j = - \frac{\partial E}{\partial v_j} = - (y - d) \left(u \right), \\ a_j = - \frac{\partial E}{\partial a_j} = - (y - d) \left(\sum_{j=1}^m v_j \left(u \right) \left(- \frac{u}{a_j} \right) \right), \\ b_j = - \frac{\partial E}{\partial b_j} = - (y - d) \left(\sum_{j=1}^m v_j \left(u \right) \left(- \frac{1}{a_j} \right) \right). \end{cases} \quad (10)$$

文中神经网络的隐层激励函数采用 Morlet 小波,它是余弦调制的高斯波. Morlet 小波在时、频域局部特性都较好,表示为

$$\psi(x) = \cos \left(1.75 \frac{x-b}{a} \right) \exp \left(- \left(\frac{x-b}{a} \right)^2 / 2 \right). \quad (11)$$

2.3 学习算法描述

连续小波过程神经网络学习算法可完整描述为

- 1) 选取一组合适的正交基函数 $b_k(t) (k = 1, 2, \dots, L)$, 将连续小波过程神经网络的输入函数矢量 $X(t)$ 和输入层与隐层之间的连接权函数 $w_{ij}(t) (i = 1, 2, \dots, n; j = 1, 2, \dots, m)$ 展开;
- 2) 初始化连续小波过程神经网络待训练参数 w_{ij}^k, v_j, a_j 和 b_j ;
- 3) 给定网络学习误差精度 ϵ , 学习迭代次数 $s = 0$, 最大学习迭代次数 M , 学习速率 η ;
- 4) 根据式(8) 计算连续小波过程神经网络的误差函数 E , 如果 $E < \epsilon$ 或 $s > M$ 则转 6);
- 5) 根据式(9) 的规则修正连续小波过程神经网络待训练参数 w_{ij}^k, v_j, a_j 和 b_j , 否则, $s = s + 1$, 转 4);
- 6) 输出学习结果, 结束.

3 航空发动机滑油系统仿真

为降低维修成本,世界各大航空公司均采取了“视情维修”的维修方式,即根据发动机状态来确定发动机是否需要返厂修理及进行何种维修,这样可以在保障安全的前提下,大大提高发动机的装机使用时间,降低运营成本. 航空发动机的滑油系统是保证发动机正常工作的一个重要系统. 滑油系统的故障在发动机各类故障中占较高的比例,并且滑油具有循环使用的特点,携带着发动机大量内部信息. 通过对滑油系统进行监测和技术诊断,可提前发现发动机的内部故障,减少不必要的零部件更换,同时能够进行故障的早期预报,避免恶性事故的发生,保证

飞行安全,是实施“视情维修”的重要手段之一^[6-7]. 本文依据现有的飞机发动机滑油金属含量历史数据分别用连续小波过程神经网络和传统BP网络进行了仿真预测,并将预测结果进行了比较和讨论.

本文采用的数据来自通用航空公司的某型号发动机从一次换油到另一次换油的完整工作阶段的滑油光谱分析数据. 因为在滑油监测中以Fe的金属含量为主要监测对象,因此,本文以铁含量的时间序列分析为例. 按照光谱分析的时间排序得到金属含量数据序列155个. Fe含量时间序列如图3所示.

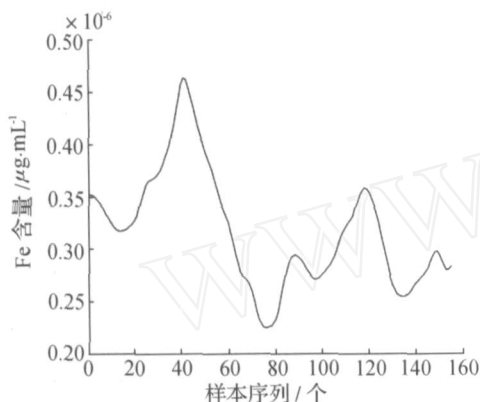


图3 Fe含量时间序列

Fig. 3 Time series of Fe content

由于连续小波过程神经网络的输入是一个连续函数,为此将连续的6个时间序列数据进行拟合,构成一个时序函数作为连续小波过程神经网络的输入函数,以第7个数据作为输出. 连续小波过程神经网络的拓扑结构取为1-11-1. 其平移与伸缩参数由文献[8]确定. 为了简化连续小波过程神经网络的计算,将网络的输入函数及连接权函数由勒让德基函数展开,基函数个数为6个. 共得到149组样本,前100组样本作为小波过程神经网络的训练样本,用后49组非训练样本进行测试. 网络误差精度设为0.005,学习速率为0.1,最大迭代次数10000次. 网络经1932次学习迭代后收敛. 为进行对比分析,在同样条件下,采用3层BP网络进行预测,网络经3929次学习迭代后收敛. 2种网络模型的学习误差曲线如图4所示. 为测试学习完成连续小波过程神经网络的泛化能力,用后49组非训练样本进行测试. 2种模型的预测结果见图5. 由图5可以看出, BP网络对于像发动机滑油系统这样复杂的时变系统,其泛化能力低于连续小波过程神经网络,而连续小波过程神经网络能够较好地航空发动机滑油系统进行状态监测. 在误差控制方面, BP网络的预测

结果有很大的误差,连续小波过程神经网络的误差却控制在较小的范围内. BP网络预测结果的平均相对误差为3.41%,连续小波过程神经网络预测结果的平均相对误差仅为0.97%.

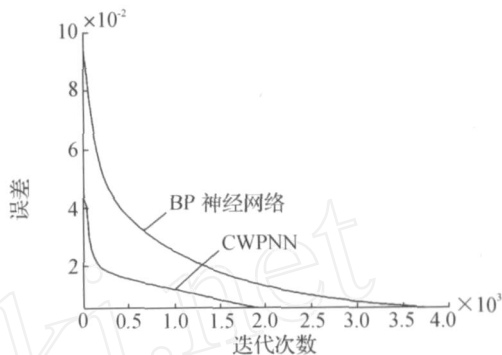


图4 学习误差曲线

Fig. 4 Learning error curve

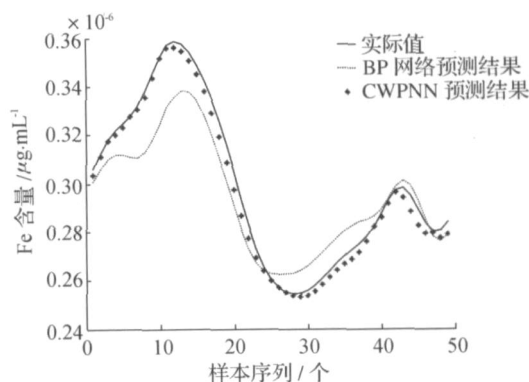


图5 Fe含量预测结果

Fig. 5 Result of Fe content forecasting

事实上,连续小波过程神经网络可借助于小波分析理论来指导网络的初始化和参数选取,使网络具有较快的收敛速度和较高的泛化精度. 小波函数具有良好的时-频局部特性,并且对于不同的逼近函数,小波函数有更广泛的选择性,即可以根据逼近函数的特性来选择不同的小波函数. 特别地,对于突变函数,小波函数具有逐步精细的描述特性,使得函数的逼近效果更好. 从而大大提高网络模型的建模能力,不仅收敛速度快,而且精度有所提高.

4 结论

1) 本文提出了连续小波过程神经网络模型,它是基于小波分析和过程神经网络建立的一类新型前馈网络,由于它结合了小波变换良好的时频局部化性质和过程神经网络处理时变信号的能力,因而具有较强的逼近和容错能力.

2)通过对发动机滑油系统的仿真预测,证明了连续小波过程神经网络能够较好地解决复杂的时变问题,同时也验证了连续小波过程神经网络具有较快的收敛速度和较好的预测精度。

参考文献:

- [1]何新贵,梁久祯. 过程神经网络的若干理论问题[J]. 中国工程科学, 2000, 2(12): 40 - 44.
HE Xingui, LIANG Jiuzhen. Some theoretical issues on procedure neural networks [J]. Engineering Science, 2000, 2(12): 40 - 44.
- [2]许少华,何新贵. 基于函数正交基展开的过程神经网络学习算法[J]. 计算机学报, 2004, 27(5): 645 - 650.
XU Shaohua, HE Xingui. Learning algorithms of process neural networks based on orthogonal function basis expansion [J]. Chinese Journal of Computers, 2004, 27(5): 645 - 650.
- [3]何新贵,梁久祯,许少华. 过程神经网络的训练及其应用[J]. 中国工程科学, 2001, 3(4): 31 - 35.
HE Xingui, LIANG Jiuzhen, XU Shaohua. Learning and applications of procedure neural networks[J]. Engineering Science, 2001, 3(4): 31 - 35.
- [4]ZHANG Q H, BENVENISTE A. Wavelet networks[J]. IEEE Transactions on Neural Networks, 1992, 3(6): 889 - 898.
- [5]柳重堪. 正交函数及其应用[M]. 北京:国防工业出版社, 1982.
- [6]王 广,李 军. 基于神经网络的航空发动机滑油监测分析[J]. 润滑与密封, 2003(5): 123 - 125.
WANG Guang, LI Jun. Aeroengine lube monitoring analysis based on neural networks[J]. Lubrication Engineering, 2003, 5: 123 - 125.
- [7]赵学智,邹春华,陈统坚,等. 小波神经网络的参数初始化研究[J]. 华南理工大学学报, 2003, 31(2): 77 - 79.
ZHAO Xuezhi, ZOU Chunhua, CHEN Tongjian, et al. A research on the initialization of parameters of wavelet neural networks[J]. Journal of South China University of Technology, 2003, 31(2): 77 - 79.

作者简介:



李 洋,男,1978 年生,博士研究生,主要研究方向为人工智能及其应用、故障诊断等,发表学术论文 6 篇,2 篇被 SCI 检索,3 篇被 EI 检索,1 篇被 ISTP 检索。

E-mail:ansysman@126.com.



钟诗胜,男,1964 年生,哈尔滨工业大学威海分校副校长,教授,博士生导师,博士,主要研究方向为制造业信息化、人工智能理论与应用、数控设备等,主持并完成国家 863 研究项目 2 项,参加国家 863 研究项目 3 项,主持和参加国家自然科学基金项目各 1 项,主持过多项省(部)级研究

项目,获黑龙江省科技进步奖 2 项,黑龙江省教学成果奖 1 项,国家版权局计算机软件著作权登记 2 套,获专利 1 项,获黑龙江省 CIMS 应用示范工程先进工作者称号,出版专著 1 部,发表学术论文 100 余篇,30 余篇被 SCI、EI 检索。

E-mail:zhongss@hit.edu.cn.



张 艳,女,1977 年生,博士研究生,主要研究方向为人工智能、智能设计及决策支持系统,发表学术论文 1 篇,并被 EI 检索。

E-mail:zhangyanjenny@163.com.