

文章编号: 1001-0920(2004)10-1129-04

基于遗传小波神经网络的冷轧轧制力预报研究

黄 敏, 王建辉, 顾树生

(东北大学 信息科学与工程学院, 辽宁 沈阳 110004)

摘 要: 依据小波函数的非线性逼近能力和神经网络的自学习特性, 提出一种适合高维输入的小波神经网络建模方法, 这种网络结构类似于多层感知器, 不同的是隐层神经元的激励函数为小波基函数。为使小波神经网络具有更高的学习精度和更快的收敛速度, 将遗传算法、小波神经网络和梯度下降算子结合起来, 提出一种遗传小波神经网络。将该网络应用于冷轧轧制力的预报, 仿真结果表明预报精度大为提高。

关键词: 小波神经网络; 混合遗传算法; 轧制力预报

中图分类号: TP183 **文献标识码:** A

Study on cold mill rolling force prediction based on wavelet neural network with genetic algorithm

HUAN GM in, WANG Jian-hui, GU Shu-sheng

(School of Information Science and Engineering, Northeastern University, Shenyang 110004, China
Correspondent: HUANG M in, E-mail: huangmzqb@163.com)

Abstract: Based on the function approximation ability of wavelet and the learning characteristic of neural network, a wavelet neural network (WNN) is introduced to handle the high dimension input problem. The structure of the WNN is similar to that of multi-layer perception, but the active function of hidden nodes is replaced by a wavelet base function. In order to obtain higher accuracy and faster speed, a wavelet neural network based on hybrid genetic algorithm (GAWNN) is put forward, which combines genetic algorithm with wavelet analysis and neural network and gradient descend operator. The application of GAWNN to cold mill rolling force prediction gives better results than typical model and force prediction precision is improved. Simulation results demonstrate the effectiveness of the methodology.

Key words: wavelet neural network; hybrid genetic algorithm; rolling force prediction

1 引 言

冷轧机轧制力的精确快速预报直接关系到冷轧产品的质量, 因此轧制力的预报精度对于实际轧钢生产具有重要意义。传统在线使用的冷轧机简单解析模型, 难以考虑轧制过程中摩擦、润滑、轧制速度对轧制压力的影响, 通常精度较差; 精度高的有限元算法和复杂的解析模型需要反复迭代, 运算时间较

长; 而传统的神经网络方法往往难以科学地确定其网络的结构, 且学习训练参数易陷入局部极小, 影响了预测的精度和可靠性。

本文建立了小波神经网络预报模型, 采用多值编码的混合遗传算法, 同时优化网络结构和网络参数, 以避免人为给定网络结构参数的弊端, 使网络达到最佳的逼近效果, 保证轧制力预报的可靠性。

收稿日期: 2003-10-13; 修回日期: 2003-12-25

基金项目: 国家自然科学基金资助项目(60274024)。

作者简介: 黄敏(1974—), 女, 辽宁铁岭人, 博士生, 从事小波分析、神经网络等研究; 顾树生(1939—), 男, 黑龙江绥化人, 教授, 博士生导师, 从事微粒子群、优化方法等研究。

2 基于小波神经网络的预报模型

2.1 基于多层感知器的小波神经网络结构

小波神经网络(WNN)^[1]是基于小波分析而构造的一类新型前馈网络,即用非线性小波基取代通常的神经元非线性激励函数(如 Sigmoid 函数)。通常意义下的高维小波神经网络映射学习容易产生“维数灾”问题,即随着网络输入维数的增加,网络所需训练样本呈指数增长,网络的收敛速度大大下降^[2]。本文根据多层感知器的思想,构建了一种适用于高维输入的小波神经网络。与文献[1,3]不同的是,所选用的小波函数为一元函数,从而避免了多元小波带来的维数灾问题

多输入多输出的小波神经网络结构如图 1 所示

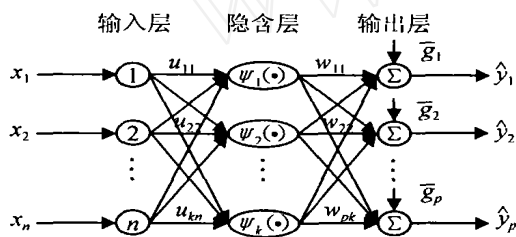


图 1 小波神经网络结构

网络输入为 $X = [x_1 \ x_2 \ \dots \ x_n]^T$, 网络输出为 $\hat{Y} = [\hat{y}_1 \ \hat{y}_2 \ \dots \ \hat{y}_p]^T$, $G = [\bar{g}_1 \ \bar{g}_2 \ \dots \ \bar{g}_p]^T$ 中各分量为各路输出的平均值。第 i 路的输出为

$$\hat{y}_i = \sum_{j=1}^k w_{ij} \Psi \left[\frac{\sum_{l=1}^n u_{jl} x_l - \tau_j}{a_j} \right] + \bar{g}_i \quad (1)$$

其中: $\Psi(\bullet)$ 为小波基函数; u_{jl}, w_{ij}, a_j 和 τ_j 分别为权重因子、小波基伸缩因子和平移因子; k 表示小波基的个数; n 和 p 分别表示输入层和输出层的单元个数。图 1 的 WNN 优化问题可描述为: 求使误差函数

$$E = \frac{1}{2} \sum_{t=1}^m \sum_{i=1}^p (y_i^t - \hat{y}_i^t)^2 \quad (2)$$

满足精度要求的最小的 k 值和相应的网络参数。其中: y_i^t 和 \hat{y}_i^t 分别为第 t 组训练样本的期望输出和实际输出, m 为训练样本的组数

2.2 预报模型

本文以某钢厂的五机架冷轧为研究背景。冷轧机轧制力 WNN 模型预报, 通过现场实测得到轧制过程的设备及工艺参数与轧制压力一一对应的数据供模型训练。WNN 网络采用的是一种映射表示法, 根据对象的输入输出数据直接建模, 无需对象的先验知识, 网络只根据训练样本的输入输出数据自动

寻找其中的相互关系, 并具有自学习的特点。本文采用图 1 所示的 WNN 来建立预报模型。输入参量^[4]为

$$X = [h_0 \ h_1 \ \dots \ h_5 \ t_{f1} \ t_{f2} \ \dots \ t_{f5} \ k_{f0} \ k_{f1} \ \mu \ B]^T \quad (3)$$

其中: h_0 为来料厚度, $h_1 \sim h_5$ 为 1~5 轧机出口厚度, t_{f1} 为入口张力值, $t_{f2} \sim t_{f5}$ 为 1~5 轧机出口张力值, k_{f0} 为来料材料强度, k_{f1} 为材料硬化系数, μ 为摩擦系数, B 为板宽。合计 16 个输入参量

在不能准确提供 k_{f0} 和 k_{f1} 的情况下, 可用材料的主要化学成分代替, 如 C, Mn, Si, Ti, Ni, Cr 等。在带钢过焊缝时, 摩擦系数随着轧机速度的变化而变化, 计算公式为

$$\mu = a + b(1/v) + c(1/v)^2 \quad (4)$$

其中: a, b, c 值与所采用的润滑剂的品种有关; v 为轧制速度

输出层参量为 $\hat{Y} = [\hat{F}_1 \ \hat{F}_2 \ \hat{F}_3 \ \hat{F}_4 \ \hat{F}_5]^T$ 。5 个机架的轧制力。采用三层网络, 隐含层单元数由优化算法决定。轧制力预报模型如图 2 所示

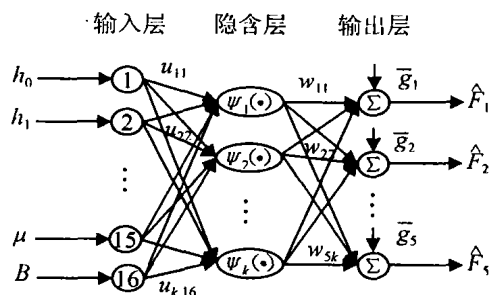


图 2 基于 WNN 的轧制力预报模型

3 基于多值编码的混合遗传算法小波神经网络优化

本文采用混合遗传算法(HGA), 即在遗传算法中引入梯度下降算子对图 2 的 WNN 进行优化, 构成了遗传小波神经网络(GAWNN)。采用一种多值编码方式, 同时优化 WNN 的网络拓扑结构和网络参数, 具有较快的收敛速度和较高的逼近精度

3.1 多值编码方式

首先对所优化的 WNN 网络参数和结构进行编码。本文采用多值编码方式, 即同时采用二进制和实值进行编码。每个染色体应包括的信息有 WNN 网络的结构及参数, 整个染色体分为 5 段: 第 1 段到第 4 段的基因位分别对应权重因子、平移因子和伸缩因子, 采用实值编码构成 WNN 的参数基因段; 该染色体的第 5 段为 WNN 的结构基因段, 代表 WNN 网络的结构, 采用二进制编码, 即用 0 或 1 作为标

志, 表示该隐单元的有效性, 0 表示无效, 1 表示有效

构成染色体时, 可根据经验确定一个较大的 k 值 根据染色体中第 5 段各基因位的值, 便可确定该小波神经网络的结构, 其中 1 的个数即为隐单元的个数, 并有相应组数的有效参数

3.2 适应度函数

适应度函数的定义对混合算法具有全局性的影响, 是算法求得全局极值的关键 WNN 的一个重要特点是网络输出值与期望输出值之间的误差平方和越小, 表示网络性能越好 本文对 WNN 进行优化时, 在标准遗传算法中引入了梯度下降算子, 因此定义的适应度函数 $F(x)$ 为所求的误差平方和函数 $F(x) = E$, 其中 E 如式(2)所示

3.3 混合遗传算法

本文对 WNN 的参数和结构采用多值编码方式, 因此在进行某些遗传操作时, 也要分别采用两种操作

1) 交叉: 首先判断交叉的基因位置, 如果交叉的基因位于 WNN 的参数基因段, 则交叉操作可采用线性组合的方式, 将两个染色体对应交叉位的基因值组合生成新的基因; 如果交叉的基因位于 WNN 的结构基因段, 即二进制编码段, 则可按标准遗传算法中的单点交叉进行操作

2) 变异: 变异操作时需要判断该基因的编码方式, 如果该位是二进制码“0”或“1”, 即变异基因位于 WNN 的结构基因段, 则以一定的概率对其进行求反运算; 如果该位是实值编码, 则采用下式进行变异操作:

$$x_i^{r+1}(q) = x_i^r(q) + (\text{rand} - 0.5)F(x_{\text{best}})\alpha_r \quad (5)$$

其中: rand 为 0-1 区间的随机数, $F(x_{\text{best}})$ 为到第 r 代为止染色体群中具有最小适应度的染色体所对应的累计平方误差, α_r 为变异因子取值为 0.75, q 为变异的基因位

3) 梯度下降算子: 为了加快收敛速度, 在遗传算法中嵌入了梯度下降算子, 对交叉和变异后产生的新群体中的个体, 以一定的概率进行梯度下降优化 优化过程中应注意以下几点:

优化后染色体上的基因位必须保证在所限定的数值范围内, 如 WNN 的伸缩因子应为正值

梯度下降算子不改变 WNN 的结构, 只对网络参数进行优化

交叉和变异后的子代和父代共同构成子代

候选群体, 通过繁殖实现竞争, 保持了群体的多样性 梯度下降运算的子代较好地继承了其父代的特性, 因此只有子代进入候选群体^[5].

4) 选择: 采用适应值排序、最优策略保存与比例选择机制相结合的选择策略 这样既可抑制比例选择机制的过早收敛和停滞现象, 又能克服交叉和变异带来的随机漫游现象, 加快收敛速度

比例选择可采用如下算子:

$$P_{ri} = (F_{\max} - F_i) / \sum_{i=1}^{\text{Pop-size}} F_i \quad (6)$$

其中: F_i 为个体 i 的适应度, F_{\max} 为上代个体中最大的适应度, Pop-size 为群体规模, P_{ri} 为个体 i 被选中的概率 采用该复制算子可保证适应度较小的染色体在下一代中有较大的选择概率

本文混合算法中的遗传算法——交叉算子、变异算子和选择算子的作用是宏观搜索, 处理的是大范围的搜索问题, 同时优化 WNN 的参数和结构; 而梯度下降算子的作用是极值局部搜索, 处理的是小范围搜索问题和加速搜索问题, 它只优化 WNN 的参数而不改变其结构 两种方法的结合, 既能克服局部极小问题, 又能加快搜索速度

4 仿真结果及分析

4.1 样本数据处理

本文利用高精度多通道数据采集仪, 对现场不同钢种、不同规格的带钢进行大量测试分析; 在此基础上, 选取一个钢种的 300 组轧制数据作为学习样本, 并选取 50 组轧制数据作为测试样本, 对训练好的网络模型进行测试

4.2 遗传小波神经网络模型优化结果

轧制压力预报模型采用一个输入层、一个隐层和一个输出层的 GAWNN 结构 输入层神经单元为 16 个, 输出层神经单元为 5 个, 分别对应于 5 个轧制力 隐层单元激活函数采用 Morlet 小波函数

$$\Psi(x) = \cos(1.75x) \exp(-x^2/2) \quad (7)$$

利用实测数据中的学习样本对模型进行学习训练, 模型优化后的隐单元数为 11, 所以最佳模型结构为 16-11-5

4.3 仿真结果与实测值解析模型计算值比较

为验证网络模型的泛化能力, 利用测试样本对 GAWNN 模型进行仿真 为作比较, 同时对实测值解析模型^[5]进行仿真, 结果如图 3 所示

从图 3 可以看出, 该网络模型具有很强的泛化能力, 其预报结果明显优于实测值解析模型预报结果 同时也可以看出, 第 1 机架的预报精度低于其他

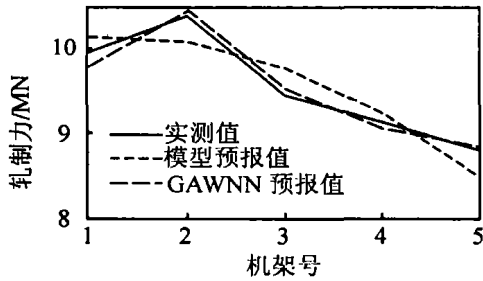


图3 实测值、模型预报值和 GAWNN 预报值结果比较
4 个机架, 主要原因是热轧钢卷(来料)带来的扰动(包括厚度波动和硬度波动)对第1 机架轧制力的影响较大, 所以预报精度较低

5 结 论

本文提出一种适合于高维输入的小波神经网络建模方法, 并应用于冷轧轧制压力预报。该模型具有很高的学习精度和良好的泛化能力, 其计算精度优于原轧制力模型。在遗传算法中引入梯度下降算子, 结合两种方法的长处, 同时对WNN 网络结构和网络参数进行优化。该算法既有较快的收敛性, 又能以

较大的概率得到全局极值

参考文献(References):

- [1] Zhang Q, Benvenise A. Wavelet network [J]. *IEEE Trans on Neural Network*, 1992, 3(6): 889-898
- [2] Delyon B, Judilsky A, Benvenise A. A accuracy analysis for wavelet approximations[J]. *IEEE Trans on Neural Network*, 1995, 6(2): 332-348
- [3] Zhang J, Walter G G, Miao Y B, et al. Wavelet neural networks for function learning[J]. *IEEE Trans on Signal Processing*, 1995, 43(6): 1485-1496
- [4] 孙登月, 朱光明, 杜凤山. 冷连轧机轧制力人工神经网络预报[J]. *冶金设备*, 2001, 6(3): 1-3
(Sun Dengyue, Zhu Guangming, Du Fengshan. Neural network predict for cold continuous rolling force [J]. *Metallurgical Equipment*, 2001, 6(3): 1-3)
- [5] 赵明旺. 基于遗传算法和最速下降法的函数优化混合数值算法[J]. *系统工程理论与实践*, 1997, 17(7): 59-64
(Zhao Mingwang. A hybrid numerical algorithm for function optimization based on genetic algorithm and steepest decent algorithm [J]. *Systems Engineering - Theory and Practice*, 1997, 17(7): 59-64)

(上接第 1128 页)

6 结 论

本文采用机器视觉来分析纸张的匀度特性, 给出了匀度分析的实验系统结构, 分析了标准匀度检测方法的缺陷, 提出一种改进的梯度指数纹理分析算法, 并对纸张匀度进行实验分析。实验结果表明, 该方法对于分析纸张匀度是可行性的, 并具有较高的鲁棒性

参考文献(References):

- [1] 阎东波, 刘焕彬. 纸页匀度的一种表征方法[J]. *中国造纸*, 1997, (6): 6-11.
(Dongpo Yan, Huanbin Liu. A expressing method of the paper evenness[J]. *China Pulp and Paper*, 1997, (6): 6-11.)
- [2] 夏丽峰, 马忻. 纸和纸板物理特性及其试验方法[M]. 北

京: 轻工业出版社, 1996 96-104

- [3] 贾云得. 机器视觉[M]. 北京: 科学出版社, 2000 53-56
- [4] Zhang Jianguo, Tan Tieniu. Brief review of invariant texture analysis methods [J]. *Pattern Recognition*, 2003, 36(3): 657-664
- [5] Kenneth R Castleman. *Digital Image Processing* [M]. Texas: Prentice-Hall Int Inc, 1996 71-75
- [6] Jean Philippe Bernie, Murray Douglas W J. Local grammage distribution and formation of paper by light transmission image analysis[J]. *Tappi J*, 1999, 79(1): 23-31
- [7] Chuns Lu, Pou Chung. Unsupervised texture segmentation via wavelet transform [J]. *Pattern Recognition*, 1997, 30(5): 729-742