

# 基于规则产生准则与修剪策略的 D-FNN 算法研究\*

左 军, 周 灵, 李晓东

(佛山科学技术学院电子与信息工程学院, 广东 佛山 528000)

**摘 要:** 提出了一种 D-FNN 结构及其学习算法, 该 D-FNN 的结构基于径向基神经网络。模糊规则的产生由输出误差或可容纳边界的有效半径决定。修剪技术的应用, 使得网络结构能够保持紧凑, 学习速度快, 确保系统的泛化能力。对所提算法作了详细探讨, 并与相关算法作比较, 从而发现了 D-FNN 的独特思想。编写了 D-FNN 的仿真程序, 对具体案例进行了仿真。结果表明, D-FNN 具有紧凑的结构和优秀的性能。

**关键词:** 动态模糊神经网络; 径向基函数; 模糊规则; 修剪策略

**中图分类号:** TP301.6   **文献标志码:** A   **文章编号:** 0529-6579(2015)05-0043-06

## The Rule Criteria and Pruning Strategy Based on D-FNN Algorithm Research

ZUO Jun, ZHOU Ling, LI Xiaodong

(School of Electronic and Information Engineering, Foshan University, Foshan 528000, China)

**Abstract:** A new structure for D-FNN and its learning algorithm are put forward. The structure of this D-FNN is based on RBF neural network. In the new algorithm and structure, generation of fuzzy rule is determined by the output error and the effective radius of the accommodate boundary. At the same time, the application of pruning technology makes a simple network structure, fast learning speed and generalization ability for system. The new algorithm is discussed in detail and compared with correlated algorithms. By these technology methods, the unique advantage of D-FNN is found. At last, simulation program for D-FNN are wrote and the concrete cases are run in the program. Simulation results show that the new D-FNN has a compact structure and excellent performance.

**Key words:** dynamic fuzzy neural network (D-FNN); radial basis function (RBF); fuzzy rule; pruning strategy

模糊逻辑理论提供了一种数学工具来获取人们认知过程, 模糊推理则形式上是利用规则进行推理, 其逻辑取值可在 0 和 1 之间连续取值, 然而不存在正式的框架来选择模糊系统的各种参数。神经网络具有学习能力、自适应能力、并行处理能力以及泛化能力等, 具有构建与人们认知有关的各种行为的潜能, 但神经网络没有明确的物理意义。把模糊逻辑和神经网络这两种方法结合起来, 取各自优

点从而构成的动态模糊神经网络, 在消费电子、工业控制、系统辨识、模式识别、图像处理等许多领域得到了广泛的应用<sup>[1-4]</sup>。

## 1 RBF 神经网络与修剪技术理论分析

### 1.1 模糊神经网络的结构

D-FNN 的结构是基于扩展的 RBF 神经网络。RBF 神经网络具有单隐层的三层前馈网络, RBF

\* 收稿日期: 2014-12-18

基金项目: 广东省自然科学基金资助项目 (S2011020002719)

作者简介: 左军 (1958 年生), 男; 研究方向: 智能计算、信息安全、数据库应用; E-mail: 13392209106@126.com

结构如图 1 所示。Poggio 和 Girosi 已经证明, RBF 网络是连续函数的最佳逼近, 它能以任意精度逼近任一连续函数<sup>[2]</sup>。

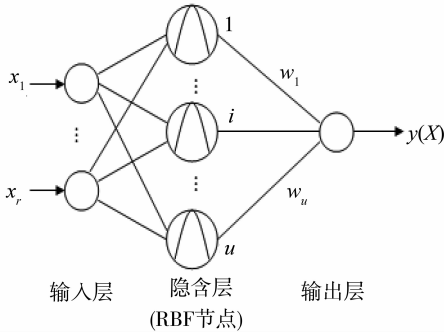


图 1 径向基神经网络结构

Fig. 1 RBF neural network structure

一个具有  $r$  个输入和一个输出的 RBF 神经网络, 该网络可以看成如下形式的映射  $f: R^r \rightarrow R^s$  :

$$y(X) = w_0 + \sum_{i=1}^u w_i R_i(X - C_i) \quad (1)$$

其中,  $X \in R^r$  是输入向量,  $R_i(\cdot)$  是基函数,  $\|\cdot\|$  表示输入空间上的欧氏范数,  $w_i (0 \leq i \leq u)$  是权值,  $C_i \in R^r (1 \leq i \leq u)$  是 RBF 的中心,  $u$  是 RBF 的单元数。

不考虑偏置量, 采用高斯函数作为激活函数时, 则式 (1) 可以写为如下的表达式:

$$y(X) = \sum_{i=1}^u w_i R_i \left( -\frac{\|X - C_i\|^2}{\sigma^2} \right) \quad (2)$$

如果高斯函数的输出采用归一化, 则径向基网络产生如下归一化输出响应:

$$y(X) = \frac{\sum_{i=1}^u w_i R_i \left( -\frac{\|X - C_i\|^2}{\sigma^2} \right)}{\sum_{i=1}^u R_i \left( -\frac{\|X - C_i\|^2}{\sigma^2} \right)} \quad (3)$$

高斯函数不仅适合于全局映射而且适合于细化局部的特征, 而且能够达到快速的学习。因此以高斯函数作为激活函数的神经网络称为径向基 (RBF) 神经网络。由于 RBF 神经网络具有最好的逼近特性, 所以本文的研究是基于动态径向基神经网络<sup>[5]</sup>。

本文提出了 D-FNN 结构及其学习算法, 该模糊神经网络的结构基于扩展的径向基神经网络。其学习算法能够使结构的辨识和参数的调整同时进行, 并且 D-FNN 能使系统结构小 (使用修剪技术)、速度快, 可用于实时建模与控制。D-FNN 的结构如图 2 所示, 共 5 层。在图 2 中,  $x_1, x_2, \dots, x_r$  是输入的语言变量,  $y$  是系统的输出,  $MF_{ij}$  是第  $i$  个

输入变量的第  $j$  个隶属函数,  $R_j$  表示第  $j$  条模糊规则,  $N_j$  是第  $j$  个归一化节点,  $w_j$  是第  $j$  个规则的结果参数或者连接权,  $u$  指系统总的规则数。

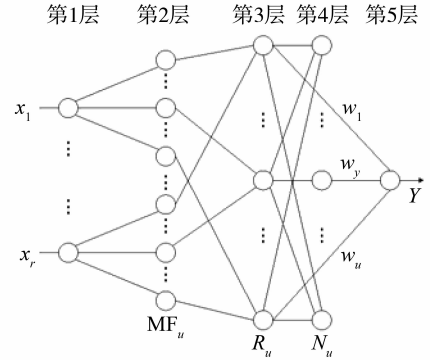


图 2 动态模糊神经网络结构

Fig. 2 D-FNN structure

## 1.2 规则产生准则理论分析

如果 D-FNN 规则数太少, 系统的性能将表现得很差; 规则数太多, 增加系统的复杂性, 导致系统的泛化能力很差。

误差判断可以表达为: 对于第  $i$  个观测数据  $(X_i, t_i)$ , 其中,  $X_i$  是输入向量,  $t_i$  是期望的输出, 计算 D-FNN 现有结构的全部输出为  $y_i$ 。

定义

$$\|e_i\| = \|t_i - y_i\| \quad (4)$$

如果

$$\|e_i\| > k_e \quad (5)$$

需要增加一条新的模糊规则。这里的  $k_e$  值是根据 D-FNN 期望的精度预先选定的。

可容纳边界判定可表达为: 对于第  $i$  个观测数据  $(X_i, t_i)$  计算输入值  $X_i$  和现有 RBF 单元的中心  $C_j$  之间的距离  $d_i(j)$ , 即

$$d_i(j) = \|X_i - C_j\|, j = 1, 2, \dots, u \quad (6)$$

其中,  $u$  是现有的模糊规则或者 RBF 单元的数量。

找出

$$d_{\min} = \operatorname{argmin}(d_i(j)) \quad (7)$$

如果

$$d_{\min} > k_d \quad (8)$$

需要增加一条新的模糊规则, 否则, 观测数据  $X_i$  可以由现有的最近的 RBF 单元表示<sup>[6]</sup>。这里,  $k_d$  是可容纳边界的有效半径。

## 1.3 修剪技术理论分析

修剪技术, 即在学习进行时, 检测到不活跃的模糊规则并加以剔除, 使网络不会持续增长, 可获

得更为紧凑的 D-FNN 结构。在本文中我们将采用 ERR（误差下降率）方法作为一种为修剪技术来选择重要的模糊规则。这种方法本质上是一种正交变换的方法，其要点是用 QR 分解把矩阵  $H$  变换成一组正交基向量。这个变换使我们可以计算出每一个基向量对期望输出能量值的贡献大小<sup>[7]</sup>。

假定理想的输出为  $T = (t_1, t_2, \dots, t_n) \in \mathbf{R}^n$ ， $Y$  为系统的输出。我们的目标是：给定  $\Psi \in \mathbf{R}^{(r+1) \times n}$  和  $T \in \mathbf{R}^n$  之间有如下关系<sup>[8]</sup>：

$$Y = W \times \Psi \quad (9)$$

$$\bar{E} = \|T - Y\| \quad (10)$$

误差能量用  $\bar{E}^T \bar{E}$  表示。给定  $n$  对输入 - 输出  $\{X(i), t(i), i = 1, 2, \dots, n\}$ ，考虑式 (9) 和式 (10) 作为线性回归模型的一种特殊情况：

$$D = H\theta + E \quad (11)$$

其中， $D = T^T \in \mathbf{R}^n$  为期望输出， $H = \Psi^T = (h_1 \dots h_v) \in \mathbf{R}^{n \times v}$ ， $v = n \times (r + 1)$  称为回归向量， $\theta = W^T \in \mathbf{R}^n$  包含实参数并且假定  $E \in \mathbf{R}^n$  是与回归量  $h_i$  不相关的一个误差向量<sup>[9]</sup>。

对于任意矩阵  $H$ ，如果它的行数大于列数，通过 QR 分解：

$$H = QA \quad (12)$$

可把  $H$  变换成一组正交基向量  $Q = (q_1, q_2, \dots, q_v) \in \mathbf{R}^{n \times v}$ ，其维数与  $H$  相同，各列向量  $q_i$  构成正交集，并且  $A \in \mathbf{R}^{n \times v}$  为一个上三角矩阵。通过这一变换，有可能从每一基向量计算每一个个体对期望输出能量的贡献<sup>[3]</sup>。

$$err_i = \frac{(q_i^T D)^2}{q_i^T q_i D^T D} \quad i = 1, 2, \dots, v \quad (13)$$

把新得到的  $err_i (i = 1, 2, \dots, (r + 1)u)$  重新排列为矩阵  $\Delta = (\delta_1, \delta_2, \dots, \delta_u) \in \mathbf{R}^{(r+1) \times u}$ ， $\Delta$  的第  $i$  列  $\delta_i$  就是与第  $i$  个规则相关的  $(r + 1)$  个误差下降率。进一步定义

$$\eta_i = \sqrt{\frac{\delta_i^T \delta_i}{r + 1}} \quad (14)$$

那么， $\eta_i$  反映第  $i$  个规则的重要性， $\eta_i$  值越大，表示第  $i$  个规则越重要<sup>[5]</sup>。如果

$$\eta_i < k_{err} \quad (15)$$

其中， $k_{err}$  表示预设的阈值，则第  $i$  个规则可剔除<sup>[10]</sup>。

## 2 实验结果与分析

本文将通过 Hermite 函数逼近以及 Mackey-Glass 时间序列预测实例来验证 D-FNN 算法的有效

性，并通过 D-FNN 算法与其他算法的比较，来检验 D-FNN 在学习效率和性能方面的特点。

一般情况下函数逼近的精度由平方  $L^2$  - 范数来度量，即可以由下式表示

$$\|f\|^2 = \int_{X \in K} |f(X)|^2 dX \quad (16)$$

当一个动态模糊神经网络用于函数逼近问题时，一般形式的表达如下

$$f \approx \sum_{i=1}^u w_i R_i(X, w_i) \quad (17)$$

其中， $w_i$  是权值， $R_i(X, w_i)$  是基函数， $u$  是基函数的个数。一旦  $R_i(X, w_i)$  被选定，权值的选择就可以用线性代数中的标准方法求解。

使用修剪技术与规则产生准则的 D-FNN 算法仿真实验。被逼近的函数为如下的 Hermite 多项式

$$f(x) = 1.1(1 - x + 2x^2) \exp\left(-\frac{x^2}{2}\right) \quad (18)$$

为了训练学习这个被逼近的函数，使用随机样本函数在区间  $[-4, 4]$  内来产生 200 个输入输出数据作为训练集。仿真结果如图 3 所示。表 1 列出了用均方根误差表达的结构和性能方面的对比。D-FNN 只需要 6 条规则，就能达到相同的训练误差，比 OLS、M-RAN 和 RANEKF 算法所需要的规则数都少<sup>[11-13]</sup>。由图 3 (a)、图 3 (b) 可以看出，由于使用了修剪技术，使得 D-FNN 的结构很小，但其泛化能力很强。仿真结果表明：与 OLS、M-RAN 和 RANEKF 算法比较，D-FNN 无论在结构上还是性能上都具有明显的优势。

Mackey 等人在 1977 年提出的 Mackey-Glass 混沌时间序列模型，是用来描述白血病时，血液中白血细胞的数量变化情况。许多研究者对这种典型的时间序列预测问题进行了研究。其方程用如下的表达式所示

$$x(t + 1) = (1 - a)x(t) + \frac{bx(t - \tau)}{1 + x^{10}(t - \tau)} \quad (19)$$

表 1 不同算法的均方根误差对比  
Table 1 RMSE compare different algorithms

	D-FNN	OLS	M-RAN	RANEKF
模糊规则数	6	7	8	13
可调参数数量	24	22	25	40
训练次数	200	200	200	200
均方根误差	0.005 3	0.009	0.043 7	0.026 1

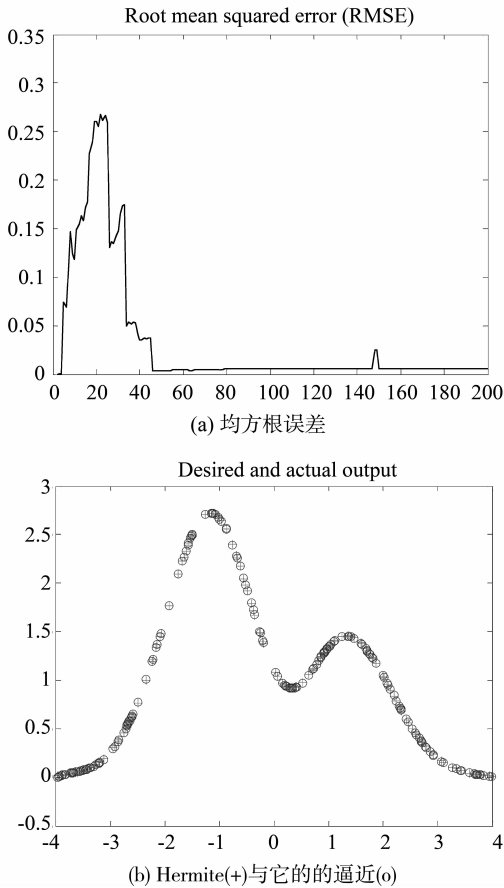


图 3 Hermite 函数逼近

Fig. 3 Hermite function approximation

由于新陈代谢生理机制方面的原因, 血液中白细胞的数量变化情况在方程 (19) 中表现出了某种混沌特性与周期性。当  $\tau > 17$  时, 方程 (19) 表现出混沌现象, 而当  $\tau < 17$  时, 方程 (19) 表现出周斯性。 $\tau$  值越大混沌现象越严重。为了能够在相同的基础上进行比较, 参数选择为:  $a = 0.1$ 、 $b = 0.2$  和  $\tau = 17$ 。预测模型表示如下:

$$\begin{aligned} x(t+p) &= f[x(t), x(t-\Delta t), \\ & x(t-2\Delta t), x(t-3\Delta t)] \end{aligned} \quad (20)$$

为了研究训练和测试情况, 用方程 (19) 在  $t=0$  到  $t=6000$  之间产生 6000 个数据。方程 (19) 初始条件选为: 当  $t < 0$  时,  $x(t) = 0, x(0) = 1.2$ 。6000 个数据将作为输入输出数据来训练预测模型式 (20), 使预测模型逼近式 (19), 并测试其性能。

设  $p = 6$ 、 $\Delta t = 6$ , 在  $111 \leq t \leq 610$  之间随机选择 500 对数据样本, 在  $111 \leq t \leq 1110$  之间随机选择 1000 对数据样本来作为训练数据。500 和 1000 个随机样本用来检验被辨识的数据模型。从  $t = 111$  到  $t = 610$  的训练结果和前 6 步预测结果由

图 4 所示。选择参数不同, 产生的模糊神经网络结构也不同, 其训练均方根误差和测试均方根误差列于表 2 中。

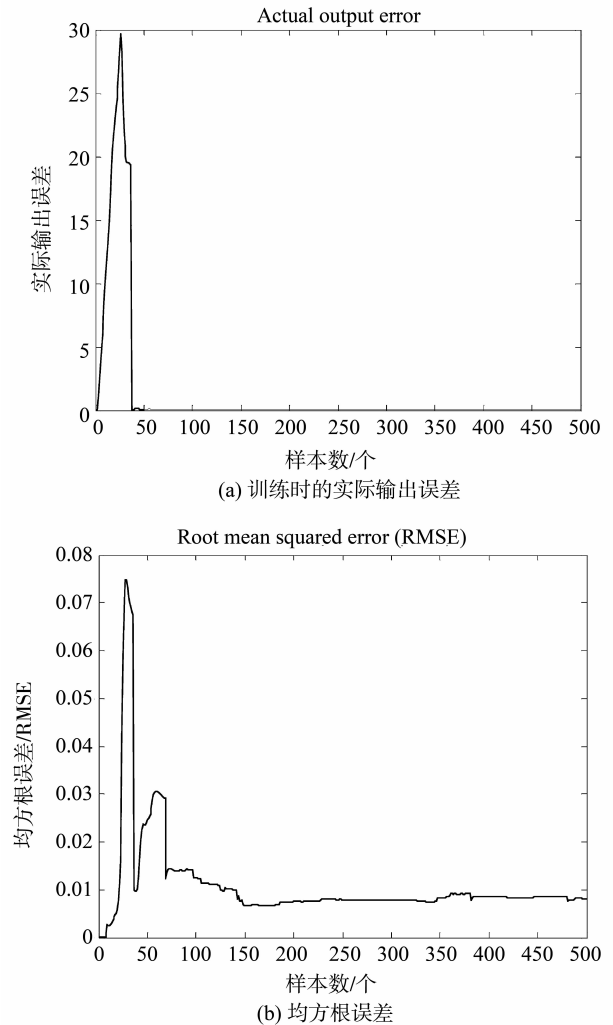


图 4 Mackey-Glass 时间序列预测

Fig. 4 Mackey-Glass time series prediction

表 2 两种不同结构训练和测试仿真结果  
Table 2 Two different structures for training and testing simulation results

结构	训练的 RMSE	测试的 RMSE	规则数
1	0.013 1	0.013 2	6
2	0.008 3	0.008 4	10

图 5 说明了 D-FNN 泛化性能的测试, 其中  $p = 6$ , D-FNN 模糊规则数为 10。测试时, 采用前 500 数据样本用于训练模型, 后 500 数据样本用于预测模型。由表 2 可以看出, 因为模型测试误差几乎等同模型训练误差, 这说明系统无论是采样 6 条模糊规则还是 10 条模糊规则所构建, 模糊神经网络

络都具有很强的泛化能力，图 5 也说明这一点。表 3 列出了 D-FNN 与 RBF-AFS、OLS 和 ANFIS 的比较结果<sup>[14-15]</sup>。

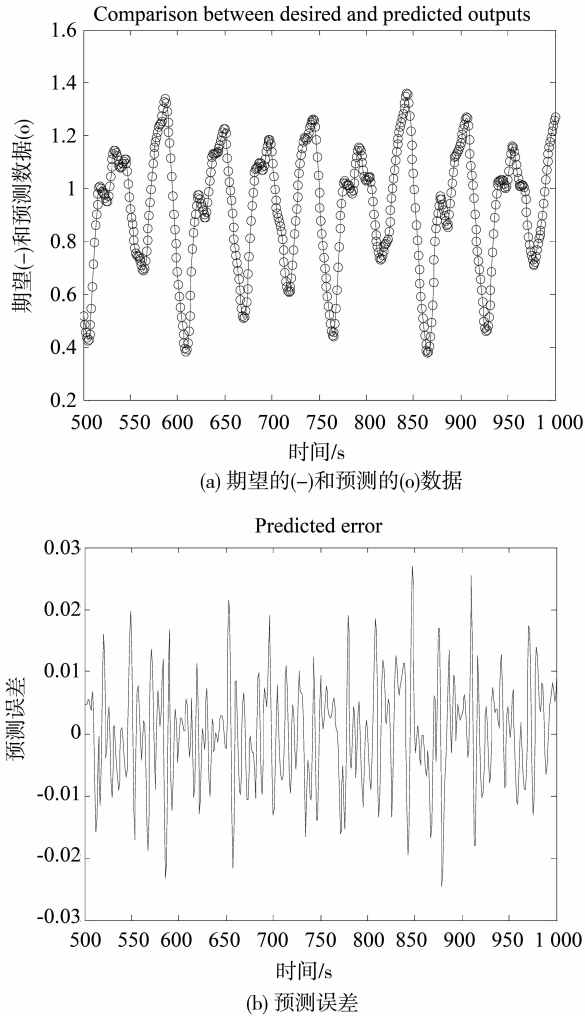


图 5 D-FNN 预测的测试结果

Fig. 5 D-FNN prediction of test results

表 3 不同的结构和性能比较 (训练条件:  $n=500$ 、 $p=6$ )

Table 3 Comparison of different structures and properties (Training conditions:  $n=500$ 、 $p=6$ )

方法	参数的数量	训练的均方根误差	测试的均方根误差	规则数
D-FNN	140	0.008 2	0.008 3	10
RBF-AFS	210	0.010 7	0.012 0	21
OLS	211	0.008 7	0.008 9	35
ANFIS	104	0.0016	0.0015	16

从比较结果来看，D-FNN 有 140 个可调参数，ANFIS 104 个可调参数，但 D-FNN 的性能也没有 ANFIS 性能好。原因是 ANFIS 一个学习过程是基

于整个学习样本，通过迭代学习方式进行的，可以达到整体的最优解。D-FNN 采用修剪技术在线学习方式，每个样本只学习一次，只能获得次优的结果。然而，D-FNN 也只能得到次优解的 OLS、RBF-AFS 相比，D-FNN 结构更小（修剪技术的使用，结构不会持续增长）而泛化能力更强。

### 3 结 论

本文把 D-FNN 算法与其他学习算法进行了比较，并深入探讨了这些算法之间的相互关系。由于修剪技术的使用使得 D-FNN 具有紧凑的结构和优秀的性能。ANFIS 及 RBF-AFS 一个学习过程是基于整个学习样本，而 D-FNN 采用占线学习方法，每个数据样本只学习一次，D-FNN 学习的快速性是以不能获得最优解为代价的。实验结果表明，对测试数据的误差与对训练数据的训练误差很接近，这说明 D-FNN 具有很强的泛化能力。直到今天，神经网络结构和泛化能力的研究仍然是神经网络的一个研究热点，引起研究人员的极大兴趣，有关这方面的工作还有待进一步的探究。动态模糊神经网络已经在消费电子、工业控制、系统辨识、模式识别、图像处理等许多领域得到了广泛的应用。

### 参考文献：

- [1] BORS G, GABBOUJ M. Minimal topology for a radial basis functions neural networks for pattern classification [J]. Digital Processing, 2012, 34:173 - 188.
- [2] 任爱红. 模糊随机过程函数列均方差一致 Henstock 积分的可积性[J]. 中山大学学报:自然科学版, 2010, 51(4): 41 - 44.
- [3] WU S Q, ER M J. Dynamic fuzzy neural networks: A novel approach to function approximation[J]. IEEE Trans Syst, Man, Cybern Part B, 2012, 30: 358 - 364.
- [4] ESPOSITO A, MARINARO M, ORICCHOI D, et al. Approximation of continuous and discontinuous mappings by a growing neural RBF-based algorithm [J]. Neural Networks, 2013,25:651 - 665.
- [5] MYOOD J, DARKEN C J. Fast Learning in network of locally-tuned processing units [J]. Neural Computation, 2013,1: 281 - 294.
- [6] 杨文光. 权值直接确定的三角型模糊前向神经网络 [J]. 中山大学学报:自然科学版,2013,52(2):33 - 37.
- [7] BORS A G, PITAS I. Median radial basis function neural network [J]. IEEE Trans Neural Networks, 2012, 23: 1351 - 1364.
- [8] HAYKIN S. Neural networks, a comprehensive founda-

- tion[M]. New York: Macmillan, 2012:465-477.
- [9] GIROSI F, POGGIO T. Networks and the best approximation property[J]. Biological Cybernetics, 2012, 63: 169-176.
- [10] ER M J, WU S Q. A fast learning algorithm for parsimonious fuzzy neural systems[J]. Fuzzy Sets and Systems, 2009, 126: 337-351.
- [11] CHEN S, COWAN C F N, GRANT P M. Orthogonal least squares algorithm for radial basis function network [J]. IEEE Trans Neural Networks, 2011, 22: 302-310.
- [12] LU Y, SUNDARARAJAN N, SARATCHANDRAN P. A sequential learning scheme for function approximation by Using minimal radial basis function networks[J]. Neural Computation, 2012, 19(2): 461-478.
- [13] KADIRKAMANATHAN V, NIRANJAN M. A function estimation approach to sequential learning with neural networks[J]. Neural Computation, 2011, 16(4): 954-975.
- [14] CHO K B, WANG B H. Radial basic function based adaptive fuzzy systems and their applications to system identification and prediction [J]. Fuzzy Sets and Systems, 2006, 26: 344-354.
- [15] JANG J S R. ANFIS: adaptive-network-based fuzzy Inference system [J]. IEEE Trans Syst, Man, Cybern, 2012, 23: 665-684.

(上接第 31 页)

于此思想,设计了遗忘因子按响应谱取指数平均的自调整方法及相应的迭代算法,在算法的收敛速度与收敛精度之间取得较好的平衡。仿真结果表明优化的频谱均衡算法较传统的均衡算法的均衡时间降低了近 50%,稳态误差减低了约 90%。

#### 参考文献:

- [1] 刘小勇,罗中良,罗明珠.基于自适应逆控制的冲击振动控制方法[J].中山大学学报:自然科学版,2008,47(5):49-53.
- [2] 许伯强,孙丽玲,李和明.基于高频率分辨力谱估计技术与优化算法的异步电动机转子故障检测新方法[J].中国电机工程学报,2013,33(3):140-147.
- [3] 陈家焱,陈章位,贺惠农,等.非高斯随机振动试验并行控制策略研究[J].机械工程学报,2012,48(4):193-198.
- [4] 冯炎,安宝坤.改进型加权噪声功率谱估计算法[J].重庆科技学院学报,2014,16(4):147-149.
- [5] 冯炎,尼玛扎西.基于频带间相关性的加权噪声功率谱估计[J].信息与电子工程,2010,8(4):431-435.
- [6] 张芄,蔡萍,武博.基于自适应逆控制的随机振动控制算法研究[J].振动与冲击,2013,32(22):163-167.
- [7] 刘力宏,李五一,杜晓坤,张书明.机载设备振动环境实测数据处理并行算法[J].科学技术与工程,2015,15(20):72-77.
- [8] 关广丰,王海涛,熊伟.随机振动功率谱再现自适应控制算法研究[J].振动与冲击,2011,30(3):1-4.
- [9] 赵淑红,朱光明.用小波变换谱均衡法提高地震资料的分辨率[J].西安科技大学学报,2007,27(2):255-259.
- [10] 罗中良,汪华斌,陈治明,等.面向随机振动功率谱估计的小波变换去噪算法理论分析[J].中山大学学报:自然科学版,2012,51(2):12-16.
- [11] 曲蕾,王京,杨斌虎.X滤波自适应逆控制算法及在液压位置控制系统中的应用[J].北京科技大学学报,2008(3):163-166.
- [12] 陈传仁,周熙襄.小波谱白化方法提高地震资料的分辨率[J].石油地球物理勘探,2000(6):54-59.
- [13] 孟新宇,宁辉,王道波,等.频域响应函数估计的非参数辨识法[J].振动与冲击,2013,32(14):174-179.