

基于 HBF 神经网络的自适应观测器

闻 新¹, 张兴旺¹, 张 威²

(1. 沈阳航空航天大学航空航天工程学部, 辽宁沈阳 110136; 2. 北京石油化工学院信息工程系, 北京 102617)

摘 要: 传统的 RBF(Radial Basis Function)神经元基函数通常把高斯类型与单一宽度作为每个神经元的激活函数, 这些特性限制了网络神经元的性能, 特别是在处理复杂的非线性建模问题上. 为了克服这个限制, 本文应用了具有类似 RBF 网络, 但激活函数不同-超基函数 HBF(Hyper Basis Function)的网络. 结合 RBF 网络, 分析了 HBF 网络的结构、基函数形式及基函数对网络的影响, 利用决策树算法计算了网络中心. 在此基础上, 提出了一种基于 HBF 神经网络的自适应观测器设计方法, 并通过引入 Lyapunov 函数, 证明了这种观测器设计方法的稳定性; 最后通过仿真验证了这种 HBF 神经网络观测器能很好地观测系统的状态值.

关键词: 状态估计; HBF (Hyper Basis Function) 神经网络; 非线性系统; 决策树

中图分类号: TP 183 **文献标识码:** A **文章编号:** 0372-2112 (2015)07-1315-05

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2015.07.010

Adaptive Observer Based on HBF Neural Networks

WEN Xin¹, ZHANG Xing-wang¹, ZHANG Wei²

(1. Faculty of Aerospace Engineering, Shenyang Aerospace University, Shenyang, Liaoning 110136, China;

2. College of Information Engineering, Beijing Institute of Petrochemical Technology, Beijing 102617, China)

Abstract: Conventional RBF neuron network is usually based on Gaussian activation function with single width for each activation function. This feature restricts neuron performance for modeling the complex nonlinear problems. To accommodate limitation of a single scale, this paper applies neural network with similar but yet different activation function—HBF(Hyper Basis Function). The state for nonlinear systems is estimated by using HBF neural networks. Combined with RBF (Radial Basis Function) networks, the structure of networks, the form of its basis functions and its influence on HBF(Hyper Basis Function) are analyzed. Decision tree algorithm is used to determine the network center. Then a design method of adaptive observer based on HBF neural networks is proposed. The Lyapunov function is introduced to prove the stability and the conditions of the bounded error of the observer. And this HBF neural network is turned out to observe system state very well by simulation.

Key words: state estimation; HBF(hyper basis function) neural networks; nonlinear system; decision tree

1 引言

在复杂的系统工程中, 状态观测器已经被广泛应用于过程监控、故障检测及故障诊断等领域, 如文献[1]利用已知的余度对动力学的影响形成残差, 构建了水下航行器故障诊断观测器, 文献[2]基于 RBF 神经网络构造飞行器控制系统观测器, 文献[3]将一种未知输入观测器应用于卫星姿态控制系统的状态估计. 但是, 通常在设计或构造观测器时, 必须遵守一些限定条件, 如要求动态系统满足可观性等等, 因此复杂非线性动态系统的观测器设计面临许多复杂的计算. 神经网络具有自学习和自适应能力等特征, 受到控制领域学者的高度关注,

在近几年里, 基于神经网络的观测器设计方面的研究成果层出不穷, 如文献[4]提出了一种基于 RBF 神经网络鲁棒滑模状态观测器, 文献[5]采用小波神经元构造系统的状态观测器等等. 由于神经网络在训练过程中存在学习速度慢等问题, 所以利用神经网络设计非线性系统观测器的研究还有待探索和完善^[5~9].

HBF 神经网络是泛化的 RBF 网络, 与 RBF 网络比较, 具有较强的泛化能力, 因此它能够以较高的精度实现函数的逼近. 本文利用 HBF 神经网络理论, 提出了一种基于 HBF 神经网络的自适应状态观测器设计方法, 并应用于非线性系统进行仿真验证.

2 HBF 神经网络

HBF(Hyper Basis Function)网络结构与 RBF 网络结构类似,但 HBF 网络是一种泛化的 RBF 网络,其网络的输出函数:

$$y_i = f(\mathbf{x}_i) = \sum_{j=1}^J w_j h_j(\mathbf{x}_i, \mathbf{c}_j, \sigma_j) + b_i \quad (1a)$$

式中, $\mathbf{x}_i \in R^n$ 为网络的第 i 个输入 (n_x 表示最大输入维数), w_j 为第 j 个基函数的连接权值, $h_j(\ast)$ 代表第 j 个神经元的基函数, \mathbf{c}_j 为第 j 个基函数的中心, $\sigma_j \in R^1$ 代表 \mathbf{x}_i 与 \mathbf{c}_j 之间的相似度, b_i 为常数.

令 $h_0 = 1, b_i = w_{i0}$, 则式(1a)可简化为

$$y_i = f(\mathbf{x}_i) = \sum_{j=0}^J w_{ij} h_j(\mathbf{x}_i, \mathbf{c}_j, \sigma_j) \quad (1b)$$

由式(1b)可见, HBF 网络的径向函数带有加权系数,使得网络具有插值决策功能.

2.1 网络基函数

在计算输入神经元与中心神经元间距离时, HBF 网络采用 Mahalanobis-like 距离^[10~12], 这主要体现在基函数的计算中.

$$h_j(\mathbf{x}_i, \mathbf{c}_j, \Sigma_j) = e^{-0.5(\mathbf{x}_i - \mathbf{c}_j)^T \Sigma_j^{-1} (\mathbf{x}_i - \mathbf{c}_j)} \quad (2)$$

式中, Σ_j 为正定方阵, 在数据进行局部缩放和定向时, 其用来表示 \mathbf{x}_i 与 \mathbf{c}_j 间的相似度, 一般有四种形式.

形式 1 网络中所有神经元都成球状, 且有相同的尺寸 $\sigma \in R^1$:

$$\Sigma_j = (1/\sigma^2) \mathbf{I}, j = 1, 2, \dots, J, \mathbf{I} \text{ 为同维单位矩阵}$$

形式 2 网络中所有神经元都成球状, 但具有不同的尺寸 $\sigma_j \in R^1$:

$$\Sigma_j = (1/\sigma_j^2) \mathbf{I}, j = 1, 2, \dots, J$$

形式 3 网络中每个神经元成椭球状, 具有变化的尺寸, 但与初始输入坐标取向一致:

$\Sigma_j = \text{diag}(1/\sigma_{1j}^2, 1/\sigma_{2j}^2, \dots, 1/\sigma_{n_x j}^2), j = 1, 2, \dots, J, n_x$ 为输入向量的最大维数

形式 4 权矩阵 Σ_j 为满阵, 每个神经元成椭球状, 且具有变化的尺寸

$$\Sigma_j = \begin{pmatrix} 1/\sigma_{11}^2 & \cdots & 1/\sigma_{1n_x}^2 \\ \vdots & \ddots & \vdots \\ 1/\sigma_{n_x 1}^2 & \cdots & 1/\sigma_{n_x n_x}^2 \end{pmatrix}$$

形式 1 所需确定的参数最少, 但会导致欠拟合; 形式 2 在 RBF 网络中使用较多, 但其不能对数据进行局部缩放; 形式 4 具有较好的灵活性和较多的优化参数, 但模型过高的自由度会导致严重的过拟合; 形式 3 保留了形式 4 的优点, 并且克服了形式 4 的缺点, 所以本文采用形式 3 作为基函数.

2.2 算法

HBF 网络学习算法有很多, 如无监督竞争学习法、LVQ 学习法、K-均值聚类法以及决策树法等^[6,7], 本文采用决策树方法来确定网络中心. 决策树(或分级树)分类的实质是在学习过程中选择能产生最大信息增益的属性特征对特征空间进行分割, 这些特征属性及其相应的值构成了各类别的决策边界, 这些边界通过特征值将该空间划分成互斥的决策区域, 每个区域赋予一个标识, 然后再通过合取和析取功能把各类别的决策区域进行整合^[13]. 对于二元决策树, 可以选用 Quinlan's C4.5 算法进行计算^[6].

(1) 计算网络中心和宽度

用“决策树叶”表示每个决策区域 \mathfrak{R}_j

$$\mathfrak{R}_j = [\min(x_{1j}), \max(x_{1j})] \cdots [\min(x_{n_x j}), \max(x_{n_x j})] \quad (3)$$

计算网络中心 $\mathbf{c}_j = (c_{1j}, \dots, c_{n_x j})$

$$c_{ij} = (\min(x_{ij}) + \max(x_{ij}))/2, \quad i = 1, \dots, n_x \quad (4)$$

计算网络的内核宽度

$$\sigma_{ij} = (\max(x_{ij}) - \min(x_{ij}))/2, \quad i = 1, \dots, n_x \quad (5)$$

通过中心 \mathbf{c}_j 和矩阵 Σ_j 所描述的内核位置和形状, 可以用 EM 算法计算求得.

(2) 计算权值

假设网络隐藏层共有 K 个神经元, $\mathbf{x}^\mu, \mathbf{y}^\mu, \mu = 1, \dots, M$ 分别是训练样本集的特征向量与目标向量. 网络的误差函数

$$E(\mathbf{W}) = \|\mathbf{H}\mathbf{W} - \mathbf{Y}\|^2 \quad (6)$$

其中, \mathbf{W} 为输出层的权值矩阵; $\mathbf{H} = (H_{ij}) = (h_j(\mathbf{x}^\mu, \mathbf{c}_j, \sigma_j))$, H_{ij} 为第 μ 个输入向量 \mathbf{x}^μ 对应的第 j 个基函数的输出; $\mathbf{Y} = (Y_{ij})$, Y_{ij} 第 μ 个目标向量 \mathbf{y}^μ 的第 j 个分量.

求解输出权值向量:

$$\mathbf{W} = \mathbf{H}^+ \mathbf{Y} \quad (7)$$

其中 \mathbf{H}^+ 为矩阵 \mathbf{H} 的伪逆(或广义逆), 可通过奇异值分解(SVD)得到.

3 HBF 网络的自适应观测器

3.1 自适应观测器设计

考虑非线性系统^[14]:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x} + \mathbf{g}(\mathbf{x}, \mathbf{u}) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) \end{cases} \quad (8)$$

式中, $\mathbf{g}(\mathbf{x}, \mathbf{u})$ 为非线性函数向量, $\mathbf{C} \in R^{m \times n}$ 为定常矩阵, $\mathbf{A} \in R^{n \times n}$ 为 Hurwitz 矩阵, 以保证 (\mathbf{A}, \mathbf{C}) 可观测.

针对式(8)非线性系统, 构造图 1 所示 HBF 神经网络观测器.

图 1 所示的状态观测器可描述为:

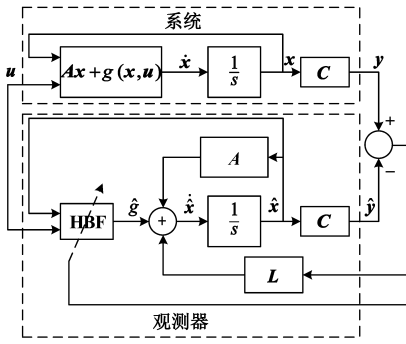


图1 非线性系统神经网络观测器模型

$$\begin{cases} \dot{\hat{x}}(t) = A\hat{x} + \hat{g}(\hat{x}, u) + L(y - \hat{y}) \\ \hat{y}(t) = C\hat{x}(t) \end{cases} \quad (9)$$

式中, L 为观测器增益, 它使得 $(A - LC)$ 为渐进稳定的 Hurwitz 矩阵。

定义状态误差 $e(t)$ 和残差 $e_y(t)$:

$$\begin{aligned} e(t) &= x(t) - \hat{x}(t) \\ e_y(t) &= y - \hat{y} = Ce(t) \end{aligned} \quad (10)$$

由式(8)和式(9)可得:

$$\dot{e}(t) = \dot{x}(t) - \dot{\hat{x}}(t) = (A - LC)e(t) + g(x, u) - \hat{g}(\hat{x}, u) \quad (11)$$

3.2 稳定性分析

根据神经网络逼近性能, 在给定逼近误差 $\epsilon(x) > 0$ 情况下, 非线性函数 $g(x, u)$ 可表示为:

$$g(x, u) = W^T f(x, u) + \epsilon(x) \quad (\|W\|_F \leq W_M, \text{即保证 } W \text{ 有界}) \quad (12)$$

则由网络估计得:

$$\hat{g}(\hat{x}, u) = \hat{W}^T f(\hat{x}, u) \quad (13)$$

将式(12)和式(13)代入式(11)可得:

$$\begin{aligned} \dot{e}(t) &= A_c e(t) + e_W^T f(\hat{x}, u) \\ &+ W^T [f(x, u) - f(\hat{x}, u)] + \epsilon(x) \end{aligned} \quad (14)$$

其中, $e_W = W - \hat{W}$, $A_c = A - LC$ 。

根据误差反馈算法^[15]可得

$$\dot{\hat{W}} = -\eta \frac{\partial J}{\partial \hat{W}} - \rho \|e_y\| \hat{W} \quad (15)$$

$J = \frac{1}{2} e_y^T e_y$, η 为学习率, ρ 为衰减系数。

修正后的网络权值为

$$\dot{\hat{W}} = -\eta f(\hat{x}, u) e_y^T C A_c^{-1} - \rho \|e_y\| \hat{W} \quad (16)$$

对式(16)进行微分得

$$\dot{e}_W = \eta f(\hat{x}, u) e_y^T C A_c^{-1} + \rho \|e_y\| \hat{W} \quad (17)$$

引入正定的 Lyapunov 函数

$$L = \frac{1}{2} e^T P e + \frac{1}{2} \text{tr}(e_W^T e_W) \quad (18)$$

其中, P 为正定矩阵, 且对任意正定矩阵 Q 满足 $A_c^T P +$

$$P A_c = -Q.$$

对式(18)进行微分

$$\dot{L} = e^T P \dot{e} + \text{tr}(e_W^T \dot{e}_W) \quad (19)$$

将式(14)和式(17)代入式(19)可得:

$$\begin{aligned} \dot{L} &= e^T P [e_W^T f(\hat{x}, u) + W^T (f(x, u) - f(\hat{x}, u)) + \epsilon(x)] \\ &+ e^T P A_c e + \text{tr}(e_W^T \eta f(\hat{x}, u) e_y^T C A_c^{-1} \\ &+ e_W^T \rho \|e_y\| \hat{W}) \end{aligned} \quad (20)$$

令 $\varphi = W^T (f(x, u) - f(\hat{x}, u)) + \epsilon(x)$, 且 φ 有界, $\|\varphi\| \leq \Phi$, $\delta = \eta C^T C A_c^{-1}$, 式(20)可简化为:

$$\begin{aligned} \dot{L} &= -\frac{1}{2} e^T Q e + e^T P [e_W^T f(\hat{x}, u) + \varphi] \\ &+ \text{tr}(e_W^T f(\hat{x}, u) e^T \delta + e_W^T \rho \|Ce\| (W - e_W)) \end{aligned} \quad (21)$$

根据下列不等式

$$\begin{aligned} \text{tr}(e_W^T (W - e_W)) &\leq W_M \|e_W\| - \|e_W\|^2 \\ \text{tr}(e_W^T f(\hat{x}, u) e^T \delta) &\leq f_M \|e_W\| \|e\| \|\delta\| \end{aligned} \quad (22)$$

可得

$$\begin{aligned} \dot{L} &\leq -\frac{1}{2} \lambda_{\min}(Q) \|e\|^2 + \|e\| \|P\| (\|e_W\| f(\hat{x}, u) + \Phi) \\ &+ f_M \|e_W\| \|e\| \|\delta\| + \rho \|Ce\| (W_M \|e_W\| - \|e_W\|^2) \end{aligned} \quad (23)$$

其中, $\lambda_{\min}(Q)$ 为矩阵 Q 的最小特征值。

进一步整理得

$$\begin{aligned} \dot{L} &\leq -\frac{1}{2} \lambda_{\min}(Q) \|e\|^2 \\ &+ \|e\| [\|P\| \Phi - \rho \|C\| \|e_W\|^2 \\ &+ \|e_W\| (\|P\| f_M + f_M \|\delta\| + \rho W_M \|C\|)] \end{aligned} \quad (24)$$

$$\text{令 } K_1 = \frac{\|\delta\|}{2},$$

$$K_2 = \frac{\|P\| f_M + f_M \|\delta\| + \rho W_M \|C\|}{2(\rho \|C\| - K_1^2)}$$

代入式(24)并整理得

$$\begin{aligned} \dot{L} &\leq -\frac{1}{2} \lambda_{\min}(Q) \|e\|^2 + [\|P\| \Phi + (\rho \|C\| - K_1^2) K_2^2 \\ &- (\rho \|C\| - K_1^2) (K_2 - \|e_W\|)^2 - K_1^2 \|e_W\|^2] \|e\| \end{aligned} \quad (25)$$

因此, 只要 $\|e\| > \frac{2}{\lambda_{\min}(Q)} [\|P\| \Phi + (\rho \|C\|$

$- K_1^2) K_2^2]$ 且 $\rho \geq \frac{K_1^2}{\|C\|}$, 就可保证 $\dot{L} \leq 0$, 即保证估计误差、权值误差和输出误差均有界。

4 仿真结果

考虑非线性系统, 其状态方程如式(8), 对应参数值为:

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, g(x, t) = \begin{pmatrix} -0.2 \sin t \\ \cos t - 3 \sin x_1 \end{pmatrix}, C = \begin{pmatrix} 1 & 1 \end{pmatrix}$$

由式(9), 设计 HBF 神经网络状态观测器, 设初始

参数为 $L = (-0.9 \ 3)^T$, $x = (0.2 \ 0)^T$, $\hat{x} = (0 \ 0.1)^T$. 图 2 和图 3 为仿真曲线, 通过仿真结果可知, HBF 神经网络观测器对非线性系统的状态变量具有较好的跟踪能力, 但由于状态变量初始化(按照经验选取初始值)的差异, 造成在开始阶段的估计误差相对较大, 在曲线拐点处的逼近误差较大, 是由于状态的变化率较大造成的.

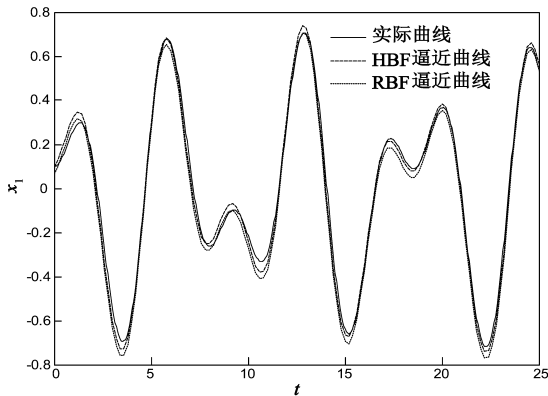


图2 x_1 与 \hat{x}_1 的状态估计曲线

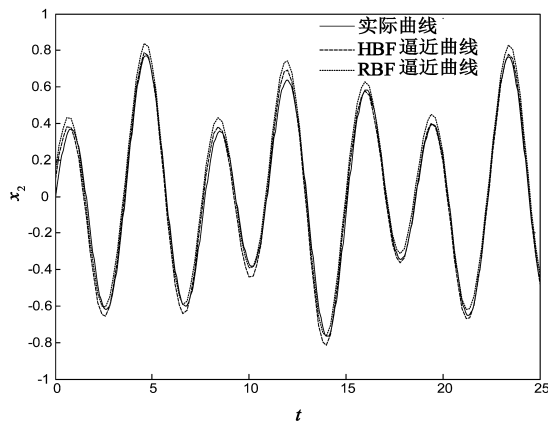


图3 x_2 与 \hat{x}_2 的状态估计曲线

图 4 为 HBF 网络观测器与传统的径向基网络观测器状态估计的输出误差曲线, 通过比较可以看出, 在初始阶段两种网络跟踪估计误差均相对较大, 但 HBF 网络跟踪估计性能提高较快. HBF 网络的平均估计误差为 0.0831, 而 RBF 网络的平均估计误差为 0.1024, 这归结为其超级基函数作用的结果.

5 结论

本文提出的自适应观测器设计方法, 因为利用了 HBF 网络, 一方面使得神经元数量较少, 能够降低传统观测器设计方法的复杂度; 另一方面, 因为网络对初值的选取具有较强的适应能力, 所以该方法对非线性动态神经系统具有较好的自适应逼近能力, 仿真结果证

明该方法设计的观测器具有设计简单和精度高等优点. 但由于网络本身的缺陷, 也面临很多问题需要进一步改进, 如需要优化的参数较多, 另外由于网络参数具有较大的选择性, 所以有时会导致数据的过拟合等问题, 这些都有待进一步完善.

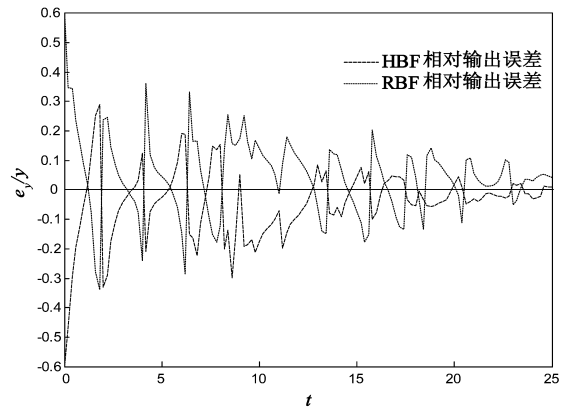


图4 输出 y 的相对误差

参考文献

- [1] 高立娥, 刘卫东, 路尧. X 字型舵水下航行器故障诊断观测器设计与仿真研究[J]. 弹箭与制导学报, 2008, 28(4): 222 - 224.
Gao Li-e, Liu Wei-dong, Lu Yao. Failure diagnose observer design and simulation for X-type rudder plane of underwater vehicle[J]. Journal of Projectiles, Rockets, Missiles and Guidance, 2008, 28(4): 222 - 224. (in Chinese)
- [2] 宋玉琴, 章卫国, 刘小雄. 基于 RBF 神经网络观测器飞控系统故障诊断[J]. 计算机仿真, 2010, 27(3): 85 - 88.
Song Yu-qin, Zhang Wei-guo, Liu Xiao-xiong. Fault diagnosis based on RBF neural network observer in flight control system [J]. Computer Simulation, 2010, 27(3): 85 - 88. (in Chinese)
- [3] 王振华, 张森, 沈毅. 卫星姿态控制系统执行器故障检测与分离[J]. 哈尔滨工业大学学报, 2013, 45(2): 72 - 76.
Wang Zhen-hua, Zhang Miao, Shen Yi. Actuator fault detection and isolation for the attitude control system of satellite [J]. Journal of Harbin Institute of Technology, 2013, 45(2): 72 - 76. (in Chinese)
- [4] 张袅娜, 张德江, 李兴广. 基于 RBF 神经网络的鲁棒滑模观测器设计[J]. 系统工程与电子技术, 2008, 30(12): 2455 - 2457.
Zhang Niao-na, Zhang De-jiang, Li Xing-guang. Design of robust sliding mode observer based on RBF neural network [J]. Systems Engineering and Electronics, 2008, 30(12): 2455 - 2457. (in Chinese)
- [5] 黄宜军, 熊炎, 刘宏兵. 基于小波网络观测器的飞控系统故障诊断[J]. 弹箭与制导学报, 2006, 26(3): 225 - 227.
Huang Yi-jun, Xiong Yan, Liu Hong-bing. Research on fault

- diagnosis for flight control systems based on wavelet network observer[J]. Journal of Projectiles, Rockets, Missiles and Guidance, 2006, 26(3): 225 - 227. (in Chinese)
- [6] Friedhelm Schwenker, Hans A. Kestler, Gu Ènther Palm. Three learning phases for radial-basis-function networks[J]. Neural Networks, 2001, 14(4 - 5): 439 - 458.
- [7] CHENG Caixia, SUN Fuchun, ZHOU Xinqun. One fire detection method using neural networks[J]. Tsinghua Science and Technology, 2011, 16(1): 31 - 35.
- [8] Dipak, M. Adhyaru. State observer design for nonlinear systems using neural network[J]. Applied Soft Computing, 2012, 12(8): 2530 - 2537.
- [9] 武宏伟, 戴琼海, 王普, 等. 基于神经网络的非线性系统的观测器设计[J]. 清华大学学报(自然科学版), 2000, 40(3): 44 - 47.
- Wu Hong-wei, Dai Qiong-hai, Wang Pu, et. al. Neural-network-based observers for nonlinear systems[J]. Journal of Tsinghua University(Science and Technology), 2000, 40(3): 44 - 47. (in Chinese)
- [10] M Valls, R Aler, O Fernandez. Using a Mahalanobis-like distance to train radial basis neural networks[J]. Computational Intelligence and Bioinspired Systems. 2005, 3512: 257 - 263.
- [11] Sergiy A Vorobyov, Andrzej Cichocki. Hyper radial basis function neural networks for interference cancellation with nonlinear processing of reference signal[J]. Digital Signal Processing, 2001, 11(3): 204 - 221.
- [12] Najdan Vukovic', Zoran Miljkovic'. A growing and pruning sequential learning algorithm of hyper basisfunction neural network for function approximation [J]. Neural Networks, 2013, 46(10): 210 - 226.
- [13] 李爱军, 罗四维, 黄华, 等. 基于决策树的神经网络[J]. 计算机研究与发展, 2005, 42(8): 1312 - 1317.
- Li Ai-jun, Luo Si-wei, Huang Hua, et. al. Decision tree based neural network design[J]. Journal of Computer Research and Development, 2005, 42(8): 1312 - 1317. (in Chinese)
- [14] 徐志超, 周玉国, 于凤满. 基于改进神经网络的非线性系统观测器设计[J]. 微型机与应用, 2011, 30(8): 76 - 78.
- Xu Zhi-chao, Zhou Yu-guo, Yu Feng-man. Design of observers for nonlinear systems based on improved neural-network[J]. Microcomputer & Its Applications, 2011, 30(8): 76 - 78. (in Chinese)
- [15] Ferdose Ahammad Shaik, Shubhi Purwar, Bhanu Pratap. Real-time implementation of Chebyshev neural network observer for twinrotor control system[J]. Expert Systems with Applications, 2011, 38(10): 13043 - 13049.

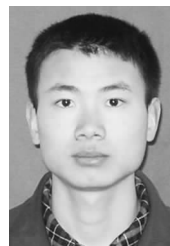
作者简介



闻 新 男, 1961 年 3 月生, 辽宁沈阳人,

博士生导师, 现任南京航空航天大学航天控制系

主任. 长期从事航天器总体设计工作及航天器智能故障诊断工作. 曾主持国家 921 工程“神舟飞船故障模拟与仿真实验室”建设, 组织完成若干国家“863”项目和总装支撑预研项目, 曾主持完成国家发改委“高精度卫星导航产品产业化”等项目. 曾担任中国航天科工集团公司研发中心主任, 中心副总师、主任和总指挥等职务.



张兴旺(通信作者) 男, 1989 年 7 月生, 河北保定人, 沈阳航空航天大学飞行器设计专业硕士研究生, 主要研究方向为智能检测与诊断及航天器故障诊断技术.

E-mail: aaron_zxw@163.com