

模糊聚类辨识算法

邓 辉 孙增圻 孙富春

(清华大学计算机系智能技术及系统国家重点实验室·北京, 100084)

摘要: 采用模糊输入聚类算法来辨识系统的模型, 通过两个模糊聚类准则函数求得聚类中心和模糊规则数, 然后求出各个子空间动态模型, 再用隶属函数光滑地把他们连接成一个全局动态模糊模型. 这种模型可以转化成状态空间模型, 从而进行控制器的设计和稳定性分析.

关键词: 模糊聚类; T-S 模型; 辨识

文献标识码: A

The Fuzzy Cluster Identification Algorithm

DENG Hui, SUN Zengqi and SUN Fuchun

(State Key Lab of Intelligent Technology and Systems, Department of Computer Science and Technology, Tsinghua University · Beijing, 100084 P.R.China)

Abstract: This paper identifies the model with fuzzy input type cluster algorithm. Firstly, get the fuzzy cluster center and fuzzy rule numbers with solving two fuzzy cluster objective function, then get every fuzzy local dynamic model, at last, use the membership functions to smoothly connect the local models together to form a global system model. This kind of fuzzy model can be transformed into the state space so that many stability analysis methods and controller design can be realized.

Key words: fuzzy cluster; T-S model; identification

1 引言(Introduction)

现在,许多实际控制系统都具有复杂、非线性动态特性.当得不到它们的模型或仅部分可得到时,模糊控制便可发挥它的作用.传统的模糊控制首先是用模糊集合表示定性知识,然后用模糊推理规则和近似推理来构造出一个模糊决策.这种方法的优点是由一组模糊规则表示的定性知识可以直接用来构造控制器.缺点是很难在推理机制中直接包含一个目标函数准则,也就是没有办法用一个全局稳定的目标函数准则来保证当把所有模糊控制规则合并到一起时,能得到一个全局稳定的闭环系统.这个问题的本质原因也就是缺乏合适的分解模型.

近来,有两种新的基于模糊模型的模糊控制方法提出来.第一种方法是使用模糊函数近似(Wang and Mendel^[1],1992;Zeng and Singh^[2],1994),其主要思想是构造一个模糊系统来近似一个期望的控制到给定目标.这种模糊系统可以看成是从输入空间到输出空间的映射.实际上,这个思想非常相似于一个静态 BP 神经网络,可以称之为一个静态模糊模型.这种模糊模型方法的缺点是系统的一些重要动态特

性不能表示出来,很难基于这种模糊模型设计控制器.第二种模糊模型(Takagi and Sugeno^[3],1985; Sugeno and Yasukawa^[4],1993;Tanaka and Sugeno^[5],1992)的主要思想是在每个局部子空间内来构造一个线性方程表示其规则,然后用隶属函数来把各个局部子空间线性方程合并起来,形成一个全局近似函数.这种模糊模型的优点是可以对模糊控制系统进行稳定性分析^[5].缺点是模型不容易用于控制设计,由于每个模型的输出仅是局部子空间输出,所以局部子空间模型不是真正的动态模型.合并这些输出可以构造出一个全局输入输出模型,可以用来稳定性分析,但不能使用状态空间分析方法来进行反馈控制器设计.

我们的工作构造一组局部子空间动态模型来表示每个局部子空间区域的动态特性,然后用隶属函数来连接这些局部子空间模型成一个全局动态模型.结果模型非常近似一个动态回归神经网络,因此,可以称之为动态模糊模型.这种方法是适合于非线性系统的控制器的设计.本文主要工作是建模,建模方法可以分为两步,一是隶属函数的辨识,这包括

模糊规则数的确定和隶属函数参数的确定,二是局部子空间线性模型的辨识,包括局部子空间规则映射的结构和参数确定,以及全局规则的插值.

2 动态模糊模型(Dynamic fuzzy model)

我们用下面的模糊模型来表示一个多输入多输出离散时间复杂系统,这个系统包含了近似推理规则和局部子空间分解模型:

$$R^l: \text{If } z_1 \text{ is } F_1^l \text{ and } \dots z_n \text{ is } F_n^l,$$

$$\text{Then } y(t+1) = A(q^{-1})y(t) + B(q^{-1})u(t) + C_l. \quad (2.1)$$

$$\begin{cases} A(q^{-1}) = A_1^l + A_2^l q^{-1} + \dots + A_n^l q^{-n+1}, \\ A_k^l = [a_{kij}^l]_{p \times p}, \\ B(q^{-1}) = B_1^l + B_2^l q^{-1} + \dots + B_n^l q^{-n+1}, \\ B_k^l = [b_{kij}^l]_{p \times g}, \\ l = 1, 2, \dots, m, \bar{n} = p \times n + g \times n. \end{cases} \quad (2.2)$$

这里 R^l 表示第 l 个近似推理规则 (A_l, B_l, C_l) 是模型系统第 l 个局部子空间模型, m 是近似推理规则数, $y(t) \in \mathbb{R}^p$ 是系统的输出变量, $u(t) \in \mathbb{R}^g$ 是系统的输入变量, q^{-1} 为延迟因子,即 $q^{-1}y(t) = y(t-1)$,且

$$\begin{aligned} \mathcal{Z}(t) &= (y(t)^T, y(t-1)^T, \dots, y(t-n+1)^T, \\ u(t)^T, u(t-1)^T, \dots, u(t-n+1)^T)^T. \end{aligned} \quad (2.3)$$

系统的最后输出由所有局部子空间模型加权后求得,因为式(2.1)的模型仅表示在某一局部子区域系统的特性,故可以称之为一个模糊动态局部子空间线性模型.

令 $\mu(\mathcal{Z}(t)) \cong \mu(z)$ 是推得的模糊集合 F^l 的隶属函数,这里 $F^l = \prod_{i=1}^{\bar{n}} F_i^l$. 再用中央平均去模糊器,乘积推理和单值模糊化,动态模糊模型(2.1)可以转化成下面的分解全局模型

$$\begin{aligned} y(t+1) &= A(q^{-1}, \mu(z))y(t) + \\ &B(q^{-1}, \mu(z))u(t) + C(\mu(z)), \end{aligned} \quad (2.4)$$

$$\text{其中 } A(q^{-1}, \mu(z)) = \sum_{l=1}^m \mu_l A(q^{-1}),$$

$$B(q^{-1}, \mu(z)) = \sum_{l=1}^m \mu_l B(q^{-1}),$$

$$C(q^{-1}, \mu(z)) = \sum_{l=1}^m \mu_l C(q^{-1}),$$

$$\mu(z) = (\mu_1(z), \mu_2(z), \dots, \mu_m(z)).$$

这种模型主要有以下三个优点:

1) 模糊模型包括两方面的知识,一是由模糊 IF-THEN 规则表示的定性知识,另一个是有局部子空间动态模型表示的定量知识.模糊模型(2.4)把这两种知识综合到一个数学框架结构中来表示复杂系统.

2) 这种方法可以认为是非线性控制中分段线性化方法的一个扩展,它把输入空间分成各个模糊子空间,用隶属函数来把它们光滑连成一个全局非线性系统的模糊模型.而不象分段线性化那样在操作点附近划分成清晰的线性子空间,因此不可能光滑地把它连接起来合并成一个全局系统模型.由于模型(2.4)中每个局部子空间的模型都是一个线性模型,因此,可以将线性系统理论用来进行非线性系统的分析和设计.

3) 可以证明^[1],模糊模型(2.4)是一个通用函数近似器,在一个紧集上,任何连续函数都可以被这个模糊模型精确地近似到任意精度.

考虑如下的隶属函数:

定义 2.1 一组函数

$$\begin{aligned} \mu_l(z, \bar{z}_l, \sigma_l) &= (\mu_l(z, \bar{z}_1, \sigma_1), \mu_l(z, \bar{z}_2, \sigma_2), \\ &\dots, \mu_l(z, \bar{z}_m, \sigma_m)) \end{aligned}$$

只要满足下面的条件,就可以称之为钟形隶属函数,

$$1) 0 \leq \mu_l(z, \bar{z}_l, \sigma_l) \leq 1, \mu_l(\bar{z}_l, \bar{z}_l, \sigma_l) = 1;$$

$$2) \sum_{l=1}^m \mu_l(z, \bar{z}_l, \sigma_l) = 1;$$

$$3) \text{给定一组 } \varepsilon_l \text{ 和 } \eta_l, \text{ 且 } \varepsilon_l > 0 \text{ 若存在一组 } \sigma_l \text{ 满足}$$

$$\|z - \bar{z}_l\| > \eta_l \Rightarrow \mu_l(z, \bar{z}_l, \sigma_l) < \varepsilon_l.$$

这里 \bar{z}_l 是钟形隶属函数的中心, σ_l 是钟形隶属函数的衰减因子.条件 1) 2) 保证了当 z 离中心 \bar{z}_l 很近时,第 l 个子空间线性模型起着主要作用,条件 3) 则保证了当 z 离中心 \bar{z}_l 很远时,由于钟形隶属函数的衰减因子的选择,则第 l 个子空间线性模型的作用非常小.

本文采用如下的隶属函数

$$\mu_l(z, \bar{z}_l, \sigma_l) = \left(\sum_{j=1}^m \frac{\|z - \bar{z}_j\| \sigma_j}{\|z - \bar{z}_j\| \sigma_j} \right)^{-1}. \quad (2.5)$$

由于它是对输入变量的模糊聚类,因此把它称之为输入类钟形隶属函数(ITBSMF),这种隶属函数可以不采用任何递推自学习算法而直接由模糊聚类算法求得.

3 隶属函数的辨识(Identification of membership function)

主要思想是采用模糊聚类的方法得到规则数和隶属函数的参数,在下面的讨论中,主要采用钟形隶属函数(BSMF).

由式(2.1)可以看出,一个 IF-THEN 规则包括两

部分,前件部分和后件部分,前件部分对应式(2.4)中的隶属函数,后件部分对应局部子空间的线性模型.因此,钟形隶属函数(BSMF)的辨识也就是前件部分的辨识,子空间的划分个数对应着规则数,子空间的中心对应着钟形隶属函数(BSMF)的中心,子空间重叠的程度对应着 BSMF 衰减因子.

用模糊聚类算法,我们可以一次得到规则数, BSMF 的中心和衰减因子,首先,我们定义如下的目标函数

$$J(\mu, \bar{z}, \alpha) = w_1 \left(\sum_{t=1}^N \sum_{l=1}^m \mu_l(t)^p \|x(t) - \bar{z}_l\|^2 + w_2 \sum_{t=1}^N \sum_{l=1}^m \mu_l(t)^p e_l(t)^2 \right), \quad (3.1)$$

对于所有的 t , 有 $\sum_{l=1}^m \mu_l(t) = 1$,

$$e_l(t) = y(t)^T - \varphi(t-1)^T \alpha^l, \quad (3.2)$$

其中

$$\begin{aligned} \varphi(t-1)^T &= ((1, 1, \dots, 1), y(t-1)^T, \\ &\dots, y(t-n)^T, u(t-1)^T, \\ &\dots, u(t-n)^T), \end{aligned}$$

$$\alpha^l = [\alpha_1^l \quad \alpha_2^l \quad \dots \quad \alpha_p^l], \alpha_J^l = [\alpha_{J1}^l \quad \dots \quad \alpha_{J\bar{n}}^l],$$

$$\bar{n} = p + p \times n + g \times n.$$

这里 m 是规则数, $\bar{z} = [\bar{z}_1 \quad \bar{z}_2 \quad \dots \quad \bar{z}_m]$ 是各个聚类空间的中心, $\|x(t) - \bar{z}_l\|$ 是各个输入点到中心的距离, $y(t) = \varphi(t-1)^T \alpha^l$, $l = 1, 2, \dots, m$ 是 m 个线性子空间模型估计方程. ω 用来调节隶属函数的形状, w_1 和 w_2 是加权因子.

式(3.1)的第一部分要求在模型输入空间内,各个输入点到各自聚类中心之间的距离尽可能地小,第二部分则要求模糊系统的输出和各个子空间输出之间的误差也是最小的.在这种情况下,基于这种目标函数的模糊聚类算法可以认为是一种动态特性聚类算法.

给定加权因子 w_1, w_2 和形状因子 ω , 则最小化目标函数 $J(\mu, \bar{z}, \alpha)$ 的必要条件是^[6]:

$$\bar{z}_l = \frac{\sum_{t=1}^N \mu_l(t)^p x(t)}{\sum_{t=1}^N \mu_l(t)^p}, \quad l = 1, 2, \dots, m, \quad (3.3a)$$

$$\alpha_l = [\Phi^T D_l \Phi]^{-1} \Phi^T D_l Y, \quad l = 1, 2, \dots, m, \quad (3.3b)$$

$$\Phi = [\varphi(1) \quad \varphi(2) \quad \dots \quad \varphi(N)]^T,$$

$$Y = [y(1) \quad y(2) \quad \dots \quad y(N)],$$

$$D_l = \text{diag}[\mu_l(t)]_{N \times N},$$

$$\begin{cases} \mu_l(t) = \left\{ \sum_{j=1}^m \frac{(w_1 \|x(t) - \bar{z}_l\|^2 + w_2 e_l(t)^2)}{(w_1 \|x(t) - \bar{z}_j\|^2 + w_2 e_j(t)^2)} \right\}^{-1}, \\ \sigma = 1/(\bar{\omega} - 1), \quad l = 1, 2, \dots, m. \end{cases} \quad (3.3c)$$

因此模糊输入空间聚类算法可以分为如下几步:

第一步:选定加权因子 w_1, w_2 和 ω , 迭代结束

阈值 $\epsilon > 0$ 和初始隶属函数 $\mu_l^{(0)}, \sum_{l=1}^m \mu_l^{(0)} = 1$.

第二步:如果 $I_l = 0$, 则根据式(3.3)更新 $\mu^{(k)} \rightarrow \mu^{(k+1)}$, 否则

$$\mu_i^{(k+1)}(t) = 0, \quad \forall i \in \tilde{I}_l \text{ and } \sum_{i \in I_l} \mu_i^{(k+1)}(t) = 1,$$

这里

$$I_l \cong \{i \mid 1 \leq i \leq m, w_1 \|x(t) - \bar{z}_l\|^2 + w_2 e_l(t)^2 = 0\},$$

$$\tilde{I}_l \cong \{1, 2, \dots, m\} - I_l.$$

第三步:如果 $\|\mu^{(k+1)} - \mu^{(k)}\| \leq \epsilon$, 则停止循环, 否则转第二步.

w_1, w_2 和 ω 是可调参数. 如果 w_1 比 w_2 大, 则意味着平均聚类准确度的权值要大一些, 反之, 则更注重方程聚类的准确度. ω 是 BSMF 的形状参数, 决定着衰减因子. 在这个算法中, 已事先假定好规则数已知, 用模糊输入空间聚类算法, 我们可以求得 BSMF. 我们用下面这个目标函数来求规则数,

$$h_z(m) = \sum_{t=1}^N \sum_{l=1}^m \mu_l(t)^p (\|x(t) - \bar{z}_l\|^2 - \|\bar{z}_l - \bar{z}\|^2),$$

这里 $\bar{z} = \frac{1}{N} \sum_{t=1}^N x(t)$.

m 逐渐增大, 直到 $h_z(m)$ 迭代最小值时, 规则数也即确定了. 最优的聚类算法是使得在每个聚类中心内部的偏差最小和聚类之间的偏差最大. 用目标函数 $h_z(m)$ 和模糊输入空间聚类算法, 可以得到如下的钟形隶属函数(BSMF)的辨识算法:

第一步:令 $m = 1$, 选一个迭代结束阈值 $\epsilon > 0$.

第二步:用模糊输入空间聚类算法求得 BSMF.

第三步:如果 $\|h_z(m+1) - h_z(m)\| < \epsilon$ 则停止迭代, 否则 $m = m + 1$, 再转第二步.

4 局部子空间模型的辨识(Identification of the local model)

在所提的建模方法中, 我们把局部子空间模型参数的估计和结构的选择, 以及减小系统的输出和模糊模型的输出之间的误差一起考虑, 因此, 这是一个用全局辨识方法来得到模糊系统(2.4)的全局函数表示形式.

局部子空间模型的辨识对应着规则的插值. 为了能保持(2.4)式中局部子空间线性模型系统的局

部特性,因而在整个辨识过程中都坚持采用钟形隶属函数.如果把钟形隶属函数和局部子空间模型的参数一起调节来最小化全局误差,尽管我们能得到给定系统的一个较好全局逼近,但局部子空间模糊规则不再表示局部子空间的特性.结果模型是一个全局非线性函数模型而不是本文所讨论的模糊模型.

因为多输入多输出模糊模型(2.4)式可由 p 个多输入单输出系统表示.这里我们用 $y(t)$ 表示(2.4)式中一个子系统的输出.因此,其输出如下式所示:

$$y(t) = \sum_{i=1}^M v_i(t, \mu) \theta_i, \quad (4.1)$$

这里 $v_i(t, \mu)$ 是 μ 和式(2.4)一输出项的乘积.举个例子 $v_i(t, \mu) = \mu_f(x(t-1))y(t-1)$. θ_i 是式(2.4)局部子空间模型的参数, $\theta = (\theta_1, \theta_2, \dots, \theta_m)^T$, m 是空间的个数.

式(4.1)的辨识问题也就是从描述 $y(t)$ 的一组子式中选取一个最重要的子式.如果没有子式满足条件,则表明这些子式对模糊模型没有什么影响.

假定下面的模型由以上的结构算法求得

$$\hat{y}(t) = v^T(t, \mu) \hat{\theta}, \quad (4.2)$$

这里 $\hat{y}(t)$ 和 $\hat{\theta}$ 分别表示 $y(t)$ 和 θ 的估计值.

参数估计的目标也就是最小化下面的准则函数

$$J(\hat{\theta}) = \|Y - \Phi(\mu) \hat{\theta}\|^2, \quad (4.3)$$

这里

$$Y = (y(1), y(2), \dots, y(N))^T, \quad (4.4)$$

$$\Phi(\mu) = \begin{bmatrix} v_0(1, \mu) & v_1(1, \mu) & \dots & v_M(1, \mu) \\ v_0(2, \mu) & v_1(2, \mu) & \dots & v_M(2, \mu) \\ \vdots & \vdots & \ddots & \vdots \\ v_0(N, \mu) & v_1(N, \mu) & \dots & v_M(N, \mu) \end{bmatrix}. \quad (4.5)$$

(4.3)式的最小值为

$$\hat{\theta} = [\Phi(\mu)^T \Phi(\mu)]^{-1} \Phi(\mu)^T Y. \quad (4.6)$$

模型(4.2)是关于参数的线性模型.整个算法仅包括最小二乘算法.对于矩阵奇异的情况,可以采用SVD分解方法.辨识算法的收敛点可以看成是传统最小二乘算法的结果.

5 仿真实例(Simulation)

考虑下面的两个 IF-THEN 规则

If $y(t) > 0$

Then $y(t+1) = 1.2y(t) - 1.3y(t-1) + 0.4u(t) - 0.22u(t-1) + \alpha(t+1)$,

If $y(t) < 0$

Then $y(t+1) = 0.34y(t) - 0.56y(t-1) + 1.4u(t) - 0.42u(t-1) + \alpha(t+1)$.

这里 $\alpha(t)$ 是一有界干扰,而且对于所有的 t , 有 $|\alpha(t)| \leq 0.05$. 我们使用随机信号作为输入信号,其中前 200 个数据用来建模,后 100 个数据用来测试.采用模糊输入空间聚类算法,求得结果如下:

$$y_m(t+1) = \sum_{l=1}^3 \mu_l f_l(x(t)),$$

$$x(t) = (y(t), y(t-1), u(t), u(t-1)),$$

这里,隶属函数和聚类中心为

$$\mu_l(z) = \left(\sum_{j=1}^3 \frac{\|z - \bar{z}_l\|^\sigma}{\|z - \bar{z}_j\|^\sigma} \right)^{-1},$$

$$l = 1, 2, 3, \sigma = 1/(\bar{\omega} - 1) = 1/(1.2 - 1),$$

$$\bar{z}_1 = (-6.6241), \bar{z}_2 = (2.6497), \bar{z}_3 = (12.6516).$$

局部子空间动态模型为

$$f_1(x(t)) = 0.3372y(t) - 0.5602y(t-1) + 1.4001u(t) - 0.4204u(t-1),$$

$$f_2(x(t)) = 1.1118y(t) - 1.0997y(t-1) + 0.6452u(t) - 0.1430u(t-1),$$

$$f_3(x(t)) = 1.2025y(t) - 1.3005y(t-1) + 0.3997u(t) - 0.2219u(t-1).$$

拟合曲线如图 1、图 2 所示.

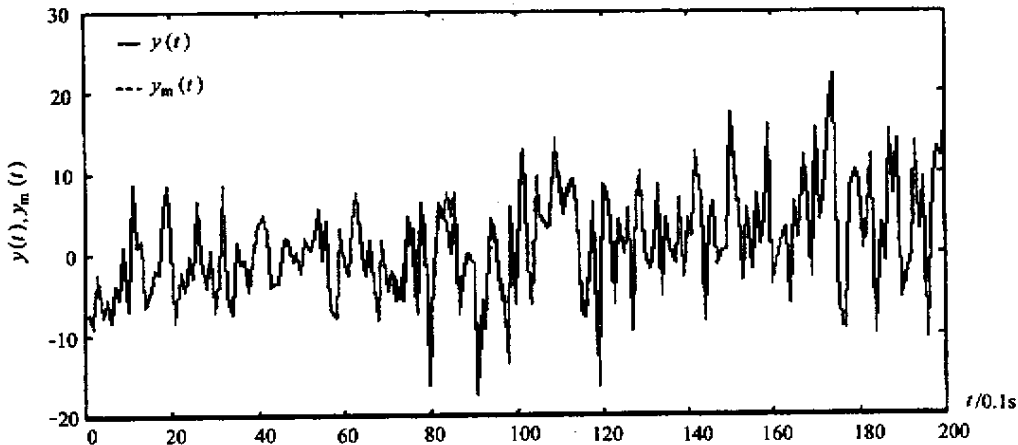


图 1 模糊聚类建模的拟合曲线
Fig. 1 Fitting curves of the modeling by fuzzy clustering

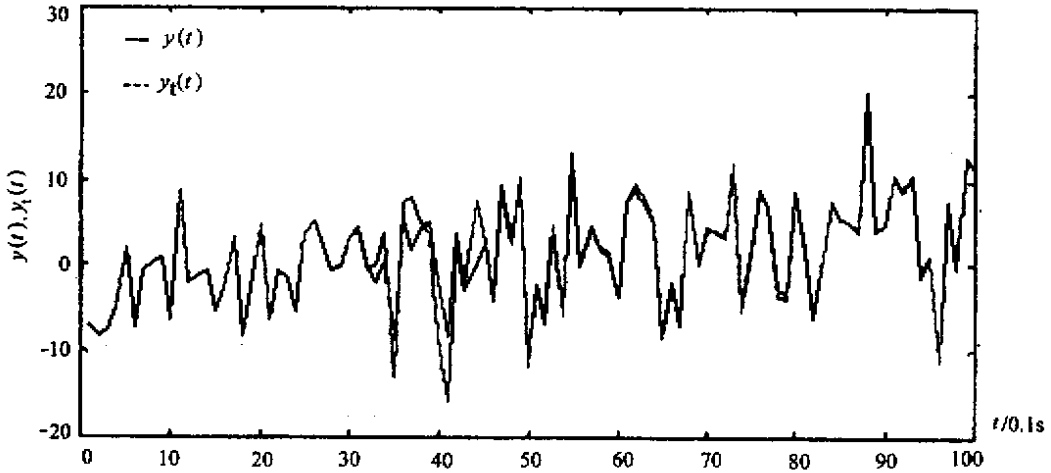


图 2 模糊聚类测试的拟合曲线

Fig. 2 Fitting curves for the fuzzy clustering test

从仿真结果可以表明,在 12.6516 这个大于 0 和 -6.6241 小于 0 的聚类中心附近辨识出的模型非常准确,在 2.6497 这个在 0 附近的聚类中心也即是两个聚类空间的边缘,迭代过程增加了一个局部子空间模型来弥补这个误差,仿真结果非常准确.

6 结论(Conclusion)

本文提出用模糊输入聚类算法来辨识系统的模型,用这种方法辨识出的系统模型,它将比较方便用于控制器的设计,我们将在以后的文章继续这个内容.

参考文献(References)

[1] Wang L X and Mendel J M. Fuzzy basis functions , universal approximation and orthogonal least squares learning[J]. IEEE Trans. Neural Networks , 1992 , 3(5) :807 - 814

[2] Zeng X J and Singh M G. Approximation theory of fuzzy systems [J]. IEEE Trans. on Fuzzy Systems , 1994 , 2(2) :162 - 176

[3] Takagi T and Sugeno M. Fuzzy identification of systems and its application to modeling and control[J]. IEEE Trans. Syst. , Man , and Cybern. , 1985 , 15(1) :116 - 132

[4] Sugeno M and Yasukawa T. Fuzzy logical based approach to qualitative modeling[J]. IEEE Trans. on Fuzzy Systems , 1993 , 1(1) :7 - 25

[5] Tanaka K and Sugeno M. Stability analysis and design of Fuzzy control systems[J]. Fuzzy Sets and Systems , 1992 , 45(2) :135 - 156

[6] Cao S G , Rees N W and Feng G. Analysis and design of uncertain fuzzy control systems , part I : fuzzy modelling and identification [A]. Proceedings of the 5th IEEE International Conference on Fuzzy Systems [C] , New Orleans , LA , 1996 :640 - 646

[7] Sun Z Q. Intelligent Control Theory and Application[M]. Beijing : Tsinghua University Press , 1997 (in Chinese)

[8] Cao S G and Rees N W. Identification of dynamic fuzzy models[J]. Fuzzy Sets and Systems , 1995 , 74(3) :307 - 320

本文作者简介

邓 辉 1972 年生,现在清华大学计算机科学与技术系计算机应用专业攻读博士学位,主要研究兴趣为模糊逻辑和神经网络控制,柔性空间机器人的建模与控制.

孙增圻 1943 年生,1966 年毕业于清华大学自动控制系,1981 年在瑞典获博士学位,现为清华大学计算机系教授,博士生导师,并兼任国家航天 863 遥科学及空间机器人专家组组长.发表论文 200 多篇,编著书 7 本.主要研究领域为神经网络控制,模糊控制,机器人智能控制与仿真等.

孙富春 见本刊 2001 年第 1 期第 141 页.