

基于粗糙集理论的模糊神经网络 及其在化纤生产过程中的应用

陈双叶, 易继锴

(北京工业大学 电子信息与控制工程学院, 北京 100022)

[摘要] 提出一种基于粗糙集理论的模糊神经网络系统, 首先运用粗糙集理论来发现大量样本数据中的概略化的规则, 然后根据这些规则来设计神经网络的结构模型, 并利用神经网络技术对模型进行训练。化纤工业中抽丝冷却侧吹风过程的模拟仿真实验, 证明了该方法的有效性和可行性。

[关键词] 粗糙集; 模糊逻辑; 神经网络; 规则获取

1 引言

在复杂应用领域中, 如何从大量的观察和实验数据中, 特别是对于含有大量噪声的不准确、不完整且无先验知识的数据中, 获取知识、表达知识、推理决策一直是智能信息处理研究的重要任务。粗糙集理论、模糊逻辑和人工神经网络以它们独特的方法已成为这一领域研究的重要工具。由于他们的研究方法不同, 因而具有各自不同的特点^[1]。粗糙集是模拟人的抽象逻辑思维, 其要点是将知识与分类联系在一起, 直接从给定问题的描述集合出发, 通过不可分辨关系和不可分辨类确定给定问题的近似域, 从而找出问题的内在规律。模糊逻辑是建立在隶属函数和模糊规则基础上的, 以 if-then 规则形式来表达知识, 依赖于领域专家的经验知识。神经网络利用非线性映射的思想和并行处理的方法, 通过学习实现全局逼近。神经网络可以通过输入输出的数据样本对网络进行训练, 实现有导师学习, 但不能确定数据样本中哪些知识是冗余的, 哪些知识是有用的。同样模糊逻辑也存在规则获取的难题。粗糙集理论可以描绘知识表达中不同属性的重要性, 通过简约去掉冗余知识, 进行知识表达

空间简约, 直接从训练数据中提取精简规则。可见, 他们各有所长, 因此探索他们之间的有机结合, 无疑对智能信息处理研究具有十分重要的意义。

2 基于粗糙集理论的模糊神经网络技术

2.1 粗糙集理论的特点

粗糙集 (rough sets) 理论^[2~4]最早是由波兰学者, Z. Pawlak 于 1982 年提出的。近几年来, 粗糙集理论已成为人工智能领域中的一个新学术热点, 在机器学习、知识获取、知识发现和决策分析等领域得到了广泛的研究与应用^[5~8]。其主要特点是将知识和分类联系在一起, 在保持决策系统分类能力不变的前提下, 通过知识简约, 导出概念的分类规则。粗糙集理论的出发点在于, 根据目前已有的对给定问题的知识将问题的论域进行划分, 然后对划分后的每一个组成部分确定其对某一概念的支持程度, 即肯定支持此概念, 肯定不支持此概念和可能支持此概念。在粗糙集理论中以上三种情况分别用三个近似集合来表示: 正域 (positive region), 负域 (negative region) 和边界 (bound-

[收稿日期] 2001-06-01; 修回日期 2001-09-12

[基金项目] “八六三”国家高技术计划资助项目 (863-306-ZT01-07-1)

[作者简介] 陈双叶 (1970-), 男, 湖南邵阳市人, 北京工业大学博士研究生

ary)。

应用 Z. Pawlak 提出的决策系统 (decision system, DS) 对问题进行描述。设 $S = (U, A, \{V_a\}, a)$ 为知识表达系统。其中 U 为对象的非空有限集合, 称为论域。 A 为属性有限集合, V_a 为属性 $a \in A$ 的值域。 $a: U \rightarrow V_a$ 为一单映射。在论域 U 中任取一元素, 取属性集合 A 中的属性 a , 存在唯一值 V_a 与其对应。如果 A 由条件属性集合 C 和结论属性集合 D 组成, 且满足 $C \cup D = A, C \cap D = \phi$, 则称 S 为决策系统。为使表示简单和描述问题方便, 假设结论属性集合 D 只包含一个元素, 那么决策系统可以简化表示为 $(U, C \cup \{d\})$ 。在决策系统 $S = (U, C \cup \{d\})$ 中, $B \subseteq C$ 是条件属性集合 C 的一个子集, 则 B 在 U 上的不可分辨关系可以定义为:

$$\text{ind}(B, \{d\}) = \{(x_1, x_2) \in U \times V \mid d(x_1) = d(x_2), \text{ or } \forall a \in B, a(x_1) = a(x_2)\}。 (1)$$

不可分辨关系实际上是一个等价关系。通过一个不可分辨关系, 可以得到决策系统的一个划分。假设 $\text{ind}(B)$ 把 U 划分为 k 个等价类 (X_1, X_2, \dots, X_k) , 则记 $U/\text{ind}(B) = \{X_1, X_2, \dots, X_k\}$, $\{X_i\}$ 表示 U 上的一个等价类的描述。对于 $X \subseteq U, B \subseteq A$ 定义 $\underline{B}(X)$ 为 X 的 B 下近似, $\overline{B}(X)$ 为 X 的 B 上近似。其中:

$$\underline{B}(X) = \cup \{r \in U/\text{ind}(B), r \subseteq X\},$$

$$\overline{B}(X) = \cup \{r \in U/\text{ind}(B), r \cap X \neq \phi\}。$$

定义 $\text{pos}_B(X) = \underline{B}(X)$ 为 X 在 B 下的正域, $\text{neg}_B(X) = U - \overline{B}(X)$ 为 X 在 B 下的负域, $\text{bn}_B(X) = \overline{B}(X) - \underline{B}(X)$ 为 X 在 B 下的边界。正域、负域和边界的概念可以直观地解释如下: 正域 $\text{pos}_B(X)$ 是在论域 U 中根据不可分辨关系 $\text{ind}(B)$ 可确定划归入 X 的不可分辨类元素的集合。负域 $\text{neg}_B(X)$ 是肯定不属于 X 的不可分辨类的元素集合。边界 $\text{bn}_B(X)$ 则表示可能属于 X , 也可能不属于 X 的元素集合。对于分类来说, 并非所有的条件属性都是必要的。简约可以理解为在不丢失信息的前提下, 可以最简单地表示决策系统的结论属性对条件属性集合的依赖和关联。条件属性集合 C 的简约是 C 的一个非空子集 C' , 存在且只存在 $\text{ind}(C', \{d\}) = \text{ind}(C, \{d\})$ 。 C 的所有简约的集合记作 $\text{red}(C)$, C 的所有简约的交集称为核, 记作 Cor

$(C) = \cap \text{red}(C)$ 。如果条件属性 C 中的某属性 a 相对于不可分辨类 $U/\text{ind}(B) = (X_1, X_2, \dots, X_k)$ 中 x_i 是可去除的, 则表明 a 的存在与否不影响 X_i 的结论值。相对简约则可定义为: 对于不可分辨类 $U/\text{ind}(B) = (X_1, X_2, \dots, X_k)$, 存在 $C' \subseteq C$, 且对于 $\forall a \in C', a$ 为相对于 X_i 不可去除, 则 C' 为 C 相对于不可分辨类 X_i 的简约。相对于 X_i 的所有简约集合记作 $\text{red}(C, X_i)$ 。通过一组相对简约, 可以得到决策系统 $S = (U, C \cup \{d\})$ 中最简单的规则集。每个相对简约就是一条或一组规则的条件。

2.2 基于粗糙集理论的规则知识获取

规则知识获取可以看作是一个知识表达、提取有用属性、简化属性表达和推理规则的一种过程。从前面的论证中, 可知一个决策系统 S 经过不可分辨关系 $\text{ind}(B)$ 得到一个划分 $U/\text{ind}(B) = (X_1, X_2, \dots, X_k)$ 。对于其中的任何一个不可分辨类 $X_i, X_i \rightarrow \{d\}$ 在形式上表示一条或一组规则。因此规则知识的获取, 可以通过计算相对简约或简约的方法来获取。通过相对简约的计算可以获取比较精简的规则, 但计算量比较大。因此计算简约成为获取知识规则的有效途径。对于决策系统 $S = (U, A, \{V_a\}, a)$, 其中 $A = C \cup D$ 。则对于任一 $X_i, X_i \in U, a \in C$, 若 $a(X_i) \neq a(X_j)$, 则以所有属性 a 值的集合 (a_{ij}) 为元素的矩阵, 称为属性值的可辨识矩阵, 记作 $M(U, B) = (k_{ij})_{n \times n}$, 其中, n 为论域 U 的个数, k_{ij} 为:

$$k_{ij} = \{a \in B, \forall (x_i, x_j) \in U \times V, a(x_i) \neq a(x_j)\}。 (2)$$

在辨识矩阵的基础上, 可以定义分辨函数 $f(s)$ 。定义 $f(s) = \bigwedge_{1 \leq i, j \leq n} c_{ij}$, 其中 c_{ij} 为可辨识矩阵的元素。 $\bigvee c_{ij}$ 表示 c_{ij} 中所有属性的析取运算, \bigwedge 表示合取运算。假设 $f(s)$ 经运算后, 化为析取范式 $f'(s)$, 即 $f'(s) = \bigwedge_{1 \leq i \leq k} \tau_i$ 。其中 $\tau_i \in 2^C, \bigwedge \tau_i$ 为 $f(s)$ 的一个合取子式, k 为 $f(s)$ 中合取子式的个数, 那么 $\text{red}(A) = \{\tau_i: 1 \leq i \leq k\}$ 。下面介绍一简单决策系统来说明上述概念和规则的获取。由信息表 1, 根据结论值 e 来划分, 可得决策矩阵表 2 和表 3。

在表 2 中, 以属性值为 1 的决策矩阵为例, 分辨函数可按式计算:

表1 信息表

Table 1 information table

U	A				
	a	b	c	d	e
1	3	0	0	5	0
2	4	0	1	6	1
3	4	1	0	5	0
4	3	2	0	6	1
5	4	2	0	5	1
6	4	0	0	5	0
7	4	2	1	6	1

表2 属性值为1的决策矩阵

Table 2 the decision matrix of attribute value 1

	1	3	6
2	(a,4) (c,1) (d,6)	(b,0) (c,1) (d,6)	(c,1) (d,6)
4	(b,2) (d,6)	(a,3) (b,2) (d,6)	(a,3) (b,2) (d,6)
5	(a,4) (b,2)	(b,2)	(b,2)
7	(a,4) (b,2) (c,1) (d,6)	(b,2) (c,1) (d,6)	(b,2) (c,1) (d,6)

表3 属性值为0的决策矩阵

Table 3 the decision matrix of attribute value 0

	2	4	5	7
1	(a, 3) (c, 0) (d, 5)	(b, 0) (d, 5)	(a, 3) (b, 0)	(a, 3) (b, 0) (c, 0) (d, 5)
3	(b, 1) (c, 0) (d, 5)	(a, 4) (b, 1) (d, 5)	(b, 1)	(b, 1) (c, 0) (d, 5)
6	(c, 0) (d, 5)	(a, 4) (b, 0) (d, 5)	(b, 0)	(b, 0) (c, 0) (d, 5)

$$[(a, 4) \vee (c, 1) \vee (d, 6)] \wedge [(b, 0) \vee (c, 1) \vee (d, 6)] \wedge [(c, 1) \vee (d, 6)] \wedge [(b, 2) \vee (d, 6)] \wedge [(a, 3) \vee (b, 2) \vee (d, 6)] \wedge [(a, 3) \vee (b, 2) \vee (d, 6)] \wedge [(a, 4) \vee (b, 2)] \wedge (b, 2) \wedge (b, 2) \wedge [(a, 4) \vee (b, 2) \vee (c, 1) \vee (d, 6)] \wedge [(b, 2) \vee (c, 1) \vee (d, 6)] \wedge [(b, 2) \vee (c, 1) \vee (d, 6)] = [(c, 1) \vee (d, 6)] \wedge (b, 2) \wedge [(b, 2) \vee (c, 1) \vee (d, 6)] = [(c, 1) \vee (d, 6)] \wedge (b, 2).$$

这样，类1最终的决策为：if $c = 1, d = 6$ then $e = 1$ else if $b = 2$ then $e = 1$

2.3 基于粗集理论的模糊神经网络^[9-13]

在 Mamdani 模型中，模糊系统的输出可由下式表示：

$$y = g(x) = \frac{\sum_{j=1}^m \left(\prod_{i=1}^n \mu_{A_i^j}(x_i^j) \right) \theta_j}{\sum_{j=1}^m \left(\prod_{i=1}^n \mu_{A_i^j}(x_i^j) \right)} = \sum_{j=1}^m P_j(x) \theta_j, \tag{3}$$

其中：
$$\mu_{A_i^j} = a_i^j \exp \left[-\frac{1}{2} \left(\frac{x_i - \bar{x}_i}{\theta_i} \right)^2 \right].$$

Wang 和 Mendel (1992) 曾给出模糊系统存在定理：给定任一连续函数 $f : U^n \rightarrow R$ 和任意 $\epsilon > 0$ ，存在一函数 $g(x)$ 使得 $\text{SUP}_{x \in U^n}, |g(x) - f(x)| < \epsilon$ 。同样由 BP 定理可以知道：给定任意 $\epsilon > 0$ 和任意 L_2 函数 $f : [0, 1]^n \rightarrow R^m$ ，存在一个 3 层 BP 网络，它可以在任意 ϵ 平方误差内逼近 f 。因此，可以用一等价的神经网络来表示模糊系统。模糊系统存在定理，虽然证明了 $g(x)$ 可以在任意 ϵ 范围内逼近 f ，但并没有给出如何构造这样一个模糊系统的方法。另外模糊规则的数目和对应各输入量的隶属函数的数目也没法直接求出。因此，将粗糙集理论、模糊系统和神经网络结合起来，充分利用各自的优点。根据粗集理论来确定规则和精简规则，利用这些规则来确定神经网络的结构模型。

根据以上论述，可以设计如图 1 所示的模糊神经网络结构模型。

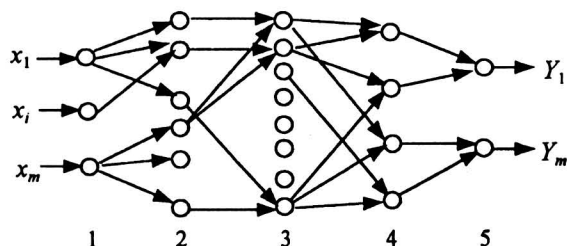


图1 模糊神经网络结构图

Fig.1 Fuzzy neural network structure

图1 模糊神经网络模型为 5 层结构，对应的是一个多输入多输出系统。

第 1 层为输入层，输入变量 $X = (x_1, x_2, \dots, x_m)^T$ 。

第 2 层表示规则的前件，实现输入变量 x_i 的模糊化，可得到 r_j 个模糊集合。 \bar{x}_{ij} 为输入变量 x_i 的第 j 个模糊集合的中心值，取作用函数为 Gauss 函数

$$\mu_i^j = \exp\left(-\left[\frac{x_i - \bar{x}_{ij}}{\theta_{ij}}\right]^2\right), \quad (4)$$

其中 $i = 1, 2, \dots, m, j = 1, 2, \dots, r_i$ 。

第 3 层的每一个节点表示一条规则，这些规则是通过粗集理论的简约计算而得到的。和第 2 层节点之间的连接方式由经过简约计算得到的精简规则来确定。因此，第 2 层和第 3 层节点之间连接并不是全连接。该层节点的作用函数为：

$$\mu_j = \mu_{j_1}^{i_1} \mu_{j_2}^{i_2} \dots \mu_{j_c}^{i_c} \quad (5)$$

其中 $\mu_{j_1}^{i_1}, \mu_{j_2}^{i_2}, \dots, \mu_{j_c}^{i_c}$ 表示第 2 层节点的输出值。

$j_1, j_2, \dots, j_c \in (1, 2, \dots, m), i_1 \in (1, 2, \dots, r_{j_1}), i_2 \in (1, 2, \dots, r_{j_2}), \dots, i_c \in (1, 2, \dots, r_{j_c}),$

$j = 1, 2, \dots, p$ 为规则的条数。

第 4 层代表该规则的后件。由于第 3 层节点表示精简后的规则，因此第三层和第四层之间的连接也不为全连接，其连接结构由具体规则确定，第 4 层神经元的输入可表示为：

$$v_i^j = \sum_{k=1}^l w_{jk} a_{jk}, \quad (6)$$

$i = 1, 2, \dots, n$ ，为第 4 层的节点数， $j = 1, 2, \dots, q$ 为第 3 层节点数， j_1, j_2, \dots, j_l 为与节点相连的第 3 层节点号。 w_{jk} 为连接权值，表示该条规则的适用度。

第 4 层节点的输出为：

$$v_i^j = v_i^j / \sum_{j=1}^{q_i} v_i^j \quad (7)$$

第 5 层为输出层：

$$y_i = \sum_{k=1}^{q_i} w_{ik} v_i^k \quad (8)$$

定义误差函数为：

$$E_p = \frac{1}{2} \sum_{i=1}^n (y_{pi} - \hat{y}_{pi})^2 \quad (9)$$

根据梯度下降算法可以得到

$$\frac{\partial E_p}{\partial w_{ij}} = \frac{\partial E_p}{\partial y_{pi}} \frac{\partial y_{pi}}{\partial w_{ij}} = (y_{pi} - \hat{y}_{pi}) v_i^j \quad (10)$$

故 $w_{ij}(k+1) = w_{ij}(k) - \beta \left(\frac{\partial E_p}{\partial w_{ij}} \right) = w_{ij}(k) - \beta (y_{pi} - \hat{y}_{pi}) v_i^j$ ，其中 $\beta > 0$ 为学习速度。

这种基于粗集理论的模糊神经网络模型和学习算法，具有学习速度快，容错能力强，模型具有可

解释性等特点。

3 仿真实验研究

基于粗糙集的模糊神经网络是信息处理的重要智能技术，有着很好的应用前景。笔者以化纤工业抽丝冷却侧吹风非线性控制系统为例，对此方法进行仿真实验研究。侧吹风空气热湿过程的控制结构框图见图 2， T_0, H_0 分别为工艺要求的恒温和恒湿度的设定值。 T'_0 为根据 T_0, H_0 值经焓湿图 ($i \sim d$) 变换后的送风露点温度设定值。图中 T_1 为喷淋后的测量温度值， T 为送风温度。从图 2 可以看出，侧吹风系统是一个多输入、多输出的多变量控制系统，是一类具有参数耦合，时变特性，且难以建立精确数学模型的复杂的非线性系统。采用传统的控制方法很难达到好的控制效果。笔者采用基于粗糙集理论的模糊神经网络结构模型设计方法和学习算法，对 FNNC 进行了学习和训练。以 FNNC2 为例，仿真的具体过程如下：

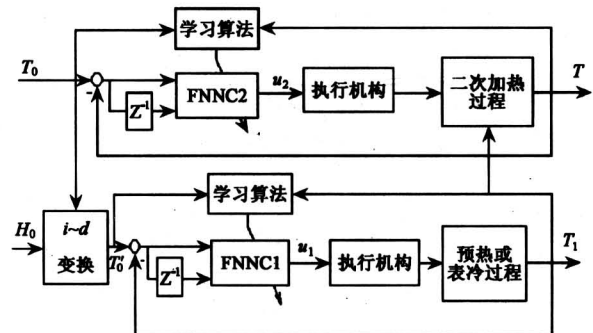
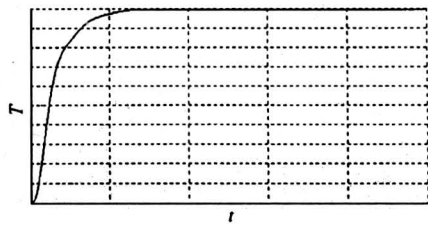


图 2 侧吹风模糊神经网络控制系统原理框图

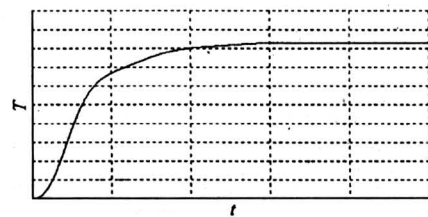
Fig.2 Fuzzy neural control system structure

设 $e(t), \dot{e}(t)$ 分别表示温度误差和温度误差变化率，那么 FNNC2 的输入变量为 $e(t), \dot{e}(t)$ ，论域为 $[0, 50]$ ；输出变量为 u_2 ，论域为 $[0, 50]$ 。FNNC2 采用图 1 所示的模糊神经网络结构模型。将变量 $e(t), \dot{e}(t)$ 及 u_2 均匀地分成 10 个等级，取各等级的中心值作为 c_{ij} ，这样可以构造第二层节点的作用函数。根据现场提供的数据样本，经筛选后共得到 500 组训练样本数据，在这些样本数据中，根据前面介绍的简约计算方法，可以很方便地得到 186 条规则。例如其中一条规则为： $e_1(t) = 4$ and $\dot{e}_1(t) = 1$ then $y_1 = 8 \mu = 0.45$ ， μ 为置信度。这条规则的含义可以解释为：当温度误差 $e(t)$ 和温度误差变化率 $\dot{e}(t)$ 分别在 $[15, 20]$ 和 $[0, 5]$

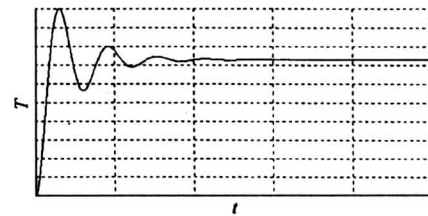
的区间取值时, 控制变量 u_2 的输出在 70~80 范围内。其中 $e_1(t)$, $\dot{e}_1(t)$ 和 y_1 为输入输出变量离散化后的变量, 规则的置信度表示在训练样本中该规则的可信程度。根据这些规则可以建立相应的第三层节点。 w_{ij} 的初值设置为 y 的 10 个等级的中心值, 用前面介绍的学习算法进行学习, 学习完后投入控制系统进行仿真实验。图 3 a 为 FNNC2 控制下的温度响应曲线, 图 3 b 为模糊控制得到的温度曲线, 图 3 c 为 PID 控制算法的温度响应曲线。



a. 采用基于粗糙集模糊神经网络控制策略的输出响应



b. 采用模糊逻辑控制策略的温度响应



c. 采用PID控制策略的输出响应

图3 仿真系统采用不同控制策略的温度输出响应

Fig. 3 Output response of simulated system with the different control strategies

4 结论

将粗糙集理论和模糊神经网络技术相结合, 运用粗糙集理论的不可分辨关系和不可分辨类的概念和简约计算方法, 从大量原始数据中发现精简的、概略化的规则, 根据所得的规则来建立神经网络模型和确定各隐层节点之间的连接关系。从而使网络从一开始就具有良好的拓扑结构, 因而网络的规模

大大减小。这种基于粗糙集理论的模糊神经网络模型和学习算法, 具有学习速度快, 容错能力强, 模型具有可解释性等特点。

参考文献

- [1] 曾黄麟, 曾 谦. 基于粗糙集理论的神经网络[J]. 四川轻化工学院学报 1997, 13(1): 1~5
- [2] Pawlak Z. Rough sets [J]. International Journal of Information and Computer Science, 1982, 11(5): 341~356
- [3] Pawlak Z. Rough sets-theoretical aspects of reasoning about data[M]. Dordrecht: Kluwer Academic Publishers, 1991. 68~162
- [4] Slowiski R. Intelligent decision support: handbook of application of the rough sets-theory [M]. The Netherlands: Kluwer Academic Publishers, 1992. 1~235
- [5] Fibak J. Rough sets based decision algorithms for treatment of duodenal ulcer by HSV [J]. Bull Pol Acad Sci Biol, 1986, 34: 227~246
- [6] Rozek A M. Rough set dependency analysis among attributes in computer implementation of expert inference models [J]. Int J Man Machine Studies, 1989, (30): 457~473
- [7] Nguyen J. Classification based on optimal feature extraction and the theory of rough sets [D]. Thesis of S Master Degree, SDSU, 1995
- [8] Guan J W, Bell D A. Rough computational methods for informational systems [J]. Artificial Intelligence, 1998, 105(1): 77~103.
- [9] Yi Jikai. Fuzzy control technique based on genetic algorithms optimizing and its application [R]. International Conference on Intelligent Processing System, Oct 1997
- [10] Chen Shuangye. Self-learning fuzzy neural network and its application to fire auto-detecting in fire protection system [A]. Proceedings of the 3rd World Congress on Intelligent Control and Automation [C], July 2000
- [11] Yi Jikai. The FNN control and its application to chemical fiber industry [A]. Proceedings of the 3rd World Congress on Intelligent Control and Automation [C], July 2000
- [12] 易继错, 侯媛彬. 智能控制技术 [M]. 北京: 北京工业大学出版社, 1999
- [13] Yi Jikai, Chen Shuangye. The application of fuzzy neural networks to the temperature control system of oil-burning tunnel kiln [R]. International Conference on Intelligent Processing System, Oct 1997

(cont. on p.51)

Definition and Verification of Shape Meter Method

Zhang Jinzhi¹, Duan Chunhua¹, Zhu Jianqin²

(1. *Central Iron & Steel Research Institute, Beijing 100081, China;*

2. *Baoshan Steel & Iron Co. Ltd., Shanghai 201900, China*)

[**Abstract**] Shape survey and control mathematics model, analytical theory for shape stiffness and the shape and gauge coordination control law are called shape meter method. In this paper, two experimental methods had been adopted to verify the correctness and practicality of the shape meter method: one is to roll aluminum plate and calculate the shape stiffness of mill and rolled piece, then survey aluminum plate crown to verify shape stiffness equation; the other is to calculate survey data off-line of hot continuous roll, and verify the shape survey and control mathematics model by self-adaptation method.

[**Key words**] shape meter method; shape stiffness equation; survey and control mathematics model; follow-up calculation; self-adaptation

(cont. from p. 46)

A Fuzzy Neural Network Based on Rough Sets and Its Applications to Chemical Fiber Production

Chen Shuangye, Yi Jikai

(*Electronic Information & Control Engineering College, Beijing Polytechnic University, Beijing 100022, China*)

[**Abstract**] A fuzzy neural network based on rough sets is presented in this paper. First, a set of rough rules are found from the given training data by using rough sets theory, then the structure and model are designed according these rules, and then the model is trained by neural network technique. The experiments that simulate the control process of side-wind for chemical fiber are carried out. The results proved its efficiency and feasibility.

[**Key words**] rough sets; fuzzy logic; neural network; rules extracted

※

※

※

※

※

更 正

本刊 2001 年第 11 期中文目次页倒数第 2 行的“形象技术”应为“形象艺术”;同期第 61 页式(1)上半式应为 $\xi + 2\zeta\Omega\xi + \Omega^2\xi = Du$,