

文章编号: 1001-0920(2001)01-0076-04

粗神经网络及其在数据融合中的应用

张兆礼, 孙圣和

(哈尔滨工业大学 自动化测试与控制系, 黑龙江 哈尔滨 150001)

摘要: 给出一种新的神经网络——粗神经网络结构, 并给出了基于粗神经网络的多传感器数据融合模型, 阐述了用于数据融合的粗神经网络的结构和训练方法。分析和仿真结果表明, 新模型不仅能解决传统模型所能解决的问题, 而且能解决传感器输出为二值或一个范围的多传感器数据融合问题。

关键词: 粗神经网络; 粗集理论; 多传感器数据融合

中图分类号: TP 18 **文献标识码:** A

Rough Neural Network and Its Application in Multisensor Data Fusion

ZHANG Zhao-li, SUN Sheng-he

(Department of Automatic Test and Control, Harbin Institute of Technology, Harbin 150001, China)

Abstract: A new data fusion model based on the rough neural network is proposed. The architecture and training method are presented. Analyses and experiments illustrate that this new model can solve not only the problem which traditional models can solve, but also the data fusion problem when the sensor's outputs are two values or a range of values.

Key words: rough neural network; rough set theory; multisensor data fusion

1 引言

数据融合技术能够从多源信号中提取信息, 减小信号的不确定度并给出信源更加完善的描述。用于实现多传感器数据融合的技术已得到了深入的研究, 例如假设检验理论、D-S 证据推理理论、模糊集理论以及专家系统等理论和方法, 都在数据融合中得到了应用^[1-4]。多传感器数据融合的概念源于人类利用多种感觉器官存活的能力, 而神经网络的提出则是模仿人类的推理和模仿能力, 因此多传感器数据融合和神经网络能够很好地联系在一起。

近几年来, 用神经网络理论和技术实现融合过程引起不少学者的研究兴趣^[5-7]。但是传统的神经

网络, 例如前馈神经网络、反馈神经网络以及自组织神经网络中, 其每个神经元的输出在训练以及应用中都是一个精确值, 即使在模糊神经网络中, 神经元也是用一个精确值来代替一个固定的隶属度。

在多传感器数据融合中, 有些传感器的输出有时并非一个值。这包括两层含义: 一是传感器同时输出两个值或者一个范围; 二是在观测时间间隔内, 传感器的输出是变化的。针对这种情况, 本文提出一种新的神经网络结构, 称为粗神经网络, 并给出了基于粗神经网络的多传感器数据融合模型。该模型不仅能解决传统神经网络所能解决的问题, 且可应用于传感器的输出不只是一个数值的情况。

收稿日期: 1999-12-10; 修回日期: 2000-04-07

作者简介: 张兆礼(1972—), 男, 山东莱芜人, 博士生, 从事神经网络、传感器信息处理的研究; 孙圣和(1937—), 男, 山东荣城人, 教授, 博士生导师, 从事电子测量与仪器、信号处理与系统辨识的研究。

2 粗神经网络

2.1 粗神经网络结构的提出

粗集理论是用于数据分析的一个数学分支^[8]。它可以对不完整、不精确和不确定数据进行分析 and 推理,发现数据间的关系,提取有用特征,是智能信息处理有力的数学工具^[9,10]。

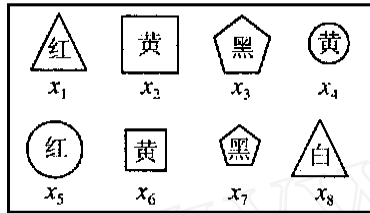


图 1 论域 U

给定论域 $U = \{x_1, x_2, \dots, x_8\}$ 如图 1 所示,则存在等价关系 R 如下

- $E_1 = \{x_1, x_5\}$ 红
- $E_2 = \{x_2, x_4, x_6\}$ 黄
- $E_3 = \{x_3, x_7\}$ 黑
- $E_4 = \{x_8\}$ 白

显然, R 等价关系实际上是论域 U 上按颜色分类的结果。考虑集合 $x = \{x_1, x_8\}$ 为论域 U 上的三角形集合,则集合 x 是 R 等价关系不可分的,即 x 为 R 上的粗集。根据粗集理论,可用 R 等价关系将集合 x 表示如下: $R_-(x) = E_4, R^-(x) = E_4 \cup E_1$ 。其中 $R_-(x), R^-(x)$ 为集合 x 的 R 上、下近似。根据该思想,可将粗集理论与人工神经网络技术结合起来。人工神经网络利用粗集理论中粗集的概念,把一个神经元分成上下两层分别代表粗糙集合的上下近似,组成一个粗神经元,继而组成整个粗神经网络。

粗神经网络与传统神经网络的最大不同主要表现为组成神经元和连接方式的不同,如图 2 所示。由此可见,每个粗神经元由上下两层组成。两个粗神经元的基本连接如图 2(a) 所示,它们有 15 种连接组合;常见组合如图 2(b) 所示。

粗神经元 j 的输入计算如下^[11]

$$\text{Input}_{w_j} = \sum_i^j (w_{UU_{ij}} \times \text{Output}_{u_i} + w_{LU_{ij}} \times \text{Output}_{l_i}) \quad (1)$$

$$\text{Input}_{l_j} = \sum_i^j (w_{LL_{ij}} \times \text{Output}_{l_i} + w_{UL_{ij}} \times \text{Output}_{u_i}) \quad (2)$$

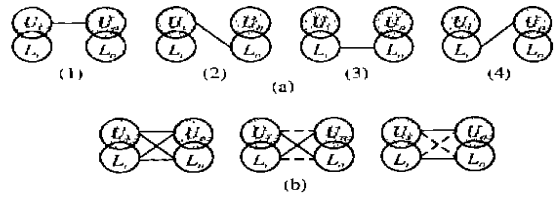


图 2 粗集神经元结构及其连接

(a) 基本连接 (b) 连接组合

其中 i 是和 j 相连的粗神经元。

粗神经元 j 的输出可计算如下

$$\text{Output}_{w_j} = \max \{f_U(\text{Input}_{w_j} - \theta_j), f_L(\text{Input}_{l_j} - \theta_j)\} \quad (3)$$

$$\text{Output}_{l_j} = \min \{f_U(\text{Input}_{w_j} - \theta_j), f_L(\text{Input}_{l_j} - \theta_j)\} \quad (4)$$

其中, θ_j 和 θ_j 为粗神经元 j 上下两层的阈值函数, f_U 和 f_L 为上下两层的传递函数,它们可以相同也可以不同。在本文中, f_U 和 f_L 均取为 Sigmoid 函数,有

$$f_U(x) = f_L(x) = \frac{1}{1 + e^{-x}} \quad (5)$$

当传递函数确定后,便可调节权值改变粗神经元上下两层的抑制和激励关系。在全互连的情况下,假定 $w_{UU} = w_{UL} = w_{LU} = w_{LL}$,则粗神经元的上下两层的输出相同。此时,一个粗神经元便相当于有两个相同输出的传统神经元。通过分别设定权值和连接情况,可以得到多种不同的神经元结构。

2.2 粗神经网络的学习算法

传统神经网络的学习方法一般是由 Widrow - Hoff 准则派生的,有

$$\Delta w_{ij} = \eta [\hat{o}_j(t) - o_j(t)] f'(w_j^T X) x_i \quad (6)$$

其中, η 为学习率, $\hat{o}_j(t)$ 是神经元 j 的理想输出, $o_j(t)$ 是神经元 j 的实际输出, x_i 是神经元 j 来自神经元 i 的输入, Δw_{ij} 是每次学习权值的修正值。对于粗神经网络,其神经元由上下两层构成,所以有其学习的特殊性。下面给出针对粗神经网络的 BP 学习算法。

定义粗神经网络的误差传递函数如下

$$E = \frac{1}{2} \sum_{p=1}^m [(o_{w_p} - \hat{o}_{w_p}) - (o_{l_p} - \hat{o}_{l_p})]^2 \quad (7)$$

则有

$$E = \frac{1}{2} \sum_{p=1}^m [(o_{w_p} - \alpha_{w_p}) - (\hat{o}_{w_p} - \hat{\alpha}_{w_p})]^2 \quad (8)$$

其中, o_{w_p} 为神经元 p 上层的实际输出, \hat{o}_{w_p} 为神经元 p 上层的理想输出, α_{w_p} 为神经元 p 下层的实际输出, $\hat{\alpha}_{w_p}$ 为神经元 p 下层的理想输出。定义

$$OU_p - OL_p = O_p, \quad \hat{OU}_p - \hat{OL}_p = \hat{O}_p \quad (9)$$

则有
$$E = \frac{1}{2} \sum_{p=1}^m (O_p - \hat{O}_p)^2 \quad (10)$$

至此,在做了初步处理后,粗神经网络的BP算法完全可以借用传统的神经网络BP学习算法。

3 基于粗神经网络的数据融合

粗神经网络不仅能处理传感器有一个输出值的情况,而且能处理有两个或一个范围的情况。因此,基于粗神经网络的数据融合能够代替基于传统神经网络的数据融合,并能解决其不能解决的问题。特别是在信源的特征能用粗糙集表示时,如在模式识别、语音信号处理、图象处理等领域。

多传感器数据融合按结构可分为数据层融合、特征层融合和决策级融合。本文给出的是基于粗神经网络的数据层数据融合结构,如图3所示。粗神经网络为3层前馈神经网络,每个传感器的输出经过预处理后,与一个输入神经元相连。其中

$$\begin{cases} \text{Sensor}_i = (\text{Sensor}_{i_u}, \text{Sensor}_{i_l})^T \\ i = 1, 2, \dots, n \\ \text{Output}_i = (\text{Output}_{i_u}, \text{Output}_{i_l})^T \\ i = 1, 2, \dots, m \end{cases} \quad (11)$$

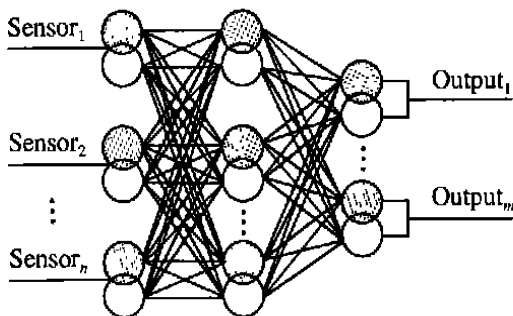


图3 基于粗神经网络的数据融合结构

4 仿真实验

本文的仿真试验是基于粗神经网络的数据融合在模式识别方面的应用。在仿真1中,粗神经网络的输入层有5个神经元,隐层有8个神经元,输出层有1个神经元;在仿真2中,输入层有2个神经元,隐层有5个神经元,输出层有2个神经元。

仿真1 线性高斯问题:有

$$\xi \sim N(0, 1) \quad H_1, \quad \eta \sim N(1, 1) \quad H_0$$

其中 $N(m, \sigma^2)$ 表示均值为 m , 方差为 σ^2 的高斯分布。试验目的是用基于粗神经网络的数据融合对给

出的数据点进行分类。对于这类问题,一个粗神经元相当于两个传统的神经元,粗神经元扩展成2倍的BP神经网络。试验结果以及与采用模糊神经网络、BP神经网络的比较如表1所示。

表1 试验1结果

	样本数	训练时间	检测概率
RNN	500	100	93.4%
BPNN	1000	85	93.0%
FNN	1000	150	85.7%

由试验结果可以看出,粗神经网络与BP神经网络相比,在取得接近的检测概率的情况下,样本数减少了一半,但训练时间稍长些。

仿真2 二维高斯分布数据分类问题:有

$$(\xi, \eta) \sim N(0, 0, 1, 1, 0) \quad H_1$$

$$(\xi, \eta) \sim N(1, 1, 1, 1, 0) \quad H_0$$

其中 $N(m_1, m_2, \sigma_1^2, \sigma_2^2, \rho)$ 是相关系数为 ρ , 均值分别为 m_1 和 m_2 , 方差分别为 σ_1^2 和 σ_2^2 的二元高斯分布。对于该问题,传感器分成两类:一是水平量传感器,输出的是水平方向的起点和终点;二是垂直量传感器,输出的是垂直方向的起点和终点。两类传感器的输出分别连入粗神经网络输入层两个神经元中。试验结果如表2所示。

表2 试验2结果

	样本数	训练时间	检测概率
RNN	500	100	90.6%
BPNN	×	×	×
FNN	1000	150	76.7%

由仿真结果可见,在二维高斯数据的识别中,粗神经网络取得的效果最好,而BP神经网络解决此类问题则很困难。

5 结 论

本文给出了粗神经网络结构的概念,并给出了基于此神经网络的数据融合模型,阐述了粗神经网络的结构和学习方法。仿真结果表明,该模型不仅能替代基于传统神经网络的数据融合模型,而且能解决传统模型不能解决的问题。当传感器的输出有两个值或是一个范围时,基于粗神经网络的数据融合是首选方法。本文的后续工作是对粗神经网络的结构做进一步研究,取得更适合于粗神经网络的算法。

(下转第82页)

设外部干扰 $\tau_d = \begin{bmatrix} q_1 \sin(t) \\ q_2 \cos(t) \end{bmatrix}$, 各参数的实际值为: m_1

$= 0.5 \text{ kg}$, $m_2 = 1.5 \text{ kg}$, $r_1 = 1 \text{ m}$, $r_2 = 0.8 \text{ m}$, 其中 m_1 和 m_2 的验前估计值为 $m_{10} = 0.4 \text{ kg}$, $m_{20} = 1.2 \text{ kg}$, 采样时间为 5 ms , 设期望轨迹为

$$q_{d1} = 0.5 \sin t - 0.5 \text{ (rad)}$$

$$q_{d2} = -0.5 \cos t + 0.5 \text{ (rad)}$$

初始关节位置和速度分别取为

$$q_1(0) = q_2(0) = 0.4$$

$$\dot{q}_1(0) = \dot{q}_2(0) = 0$$

仿真结果如图 1~4 所示。其中 $\lambda_1 = \lambda_2 = 10$, $K = \text{diag}(8, 12)$, $\eta = 0.4$, 网络初始权值均为 0.5, 各基函数宽度为 0.25, 中心在 $(-2, 2)$ 中随机选取。

由仿真结果可见, 本文设计的控制器能够有效补偿系统不确定性的影响, 保证机器人系统对期望轨迹的快速跟踪。

5 结 论

本文将变结构系统的理论和神经网络的非线性映射能力相结合, 提出一种基于神经网络的机械手自适应滑模控制器。该控制方案克服了常规滑模控制对系统不确定性界的要求, 同时可保证机械手位置矢量和速度矢量的跟踪误差渐近收敛于 0。

参考文献:

- [1] B H Nam, S J Lee, S W Lee. A neural network for the trajectory control of robotic manipulators with uncertainties[A]. Proc Amer Contr Conf[C]. Philadelphia, 1997. 3120-3123
- [2] R Carelli, E F Camacho, D Patino. A neural network feedforward adaptive controller for robots[J]. IEEE Trans on System, Man and Cybernetics, 1995, 25(9): 1281-1287.
- [3] Z Man, X H Yu, K Eshraghian *et al.* A robust adaptive sliding mode tracking control using a RBF neural network for robotic manipulators[A]. Proc of IEEE Int Conf Neural Network[C]. San Diego, 1995. 2403-2408
- [4] N Sadegh. Nonlinear identification and control via neural networks[A]. Control Systems with Dynamic Models, A SME Meeting[C]. Hawaii, 1991. 735-737.
- [5] R M Sanner, J J E Slotine. Gaussian networks for directive control[A]. IEEE Trans on Neural Networks [C]. 1992
- [6] Slotine J J E, Lee W. Applied nonlinear control[M]. Prentice Hall, 1991.
- [7] M Ertugrul, O Kaynak. Neural computation of the equivalent control in sliding mode for robot trajectory control applications[A]. Proc IEEE Int Conf Robot Auto[C]. Belgium, 1998. 2042-2047.
- [8] K K D Young. A variable structure model following control design for robotics application[J]. IEEE Trans on Auto Contr, 1988, 4(5): 556-561.

(上接第 78 页)

参考文献:

- [1] David L H, James L. An introduction to multisensor data fusion[J]. Proc of the IEEE, 1997, 85(1): 6-23
- [2] Waltz E. Data fusion for C³I: A tutorial—Command and control communications intelligence (C³I) handbook [M]. Palo Alto: EW Communications, 1986. 217-226
- [3] Waltz E, Llinas J. Multisensor data fusion[M]. Artech House, 1990
- [4] Crowley J L, Demazeu Y. Principles and techniques for sensor data fusion[J]. Signal Processing, 1993, 32: 5-27.
- [5] Cires J, Romo P A, Zufiria P J *et al.* Comparative analysis of some neural network architectures for data fusion[A]. Proc of IEEE Int Conf on Neural Networks [C]. 1995. 79-83
- [6] Wan W, Fraser D. Multisource data fusion with multiple self-organizing maps[J]. IEEE Trans on Geosci Remote Sensing, 1999, 37(3): 1344-1349.
- [7] Fukuda T. Multisensor integration system based on fuzzy inference and neural network [J]. Information Sciences, 1993, 71(1): 27-41.
- [8] Pawlak Z. Rough sets theoretical aspects of reasoning about data[M]. Poland Nowowiejska, 1990
- [9] 曾黄麟. 粗集理论及其应用[M]. 重庆: 重庆大学出版社, 1998
- [10] 王珏. Rough Set 约减与数据浓缩[J]. 高技术通讯, 1997, 7(1): 40-45
- [11] Lingras P. Comparison of neofuzzy and rough neural networks[J]. Information Sciences, 1998, 110: 207-215