

基于多传感器的神经网络模式识别方法*

罗中良¹, 麦宜佳¹, 余剑峰²

(1. 佛山科学技术学院自动化系, 广东 佛山 528000;

2. Department of Statistics, Yale University, New Haven, CT 06510, USA)

摘要: 提出基于样本信息量不减少和减小误差的原则下, 提取样本数据的综合指标的处理方法, 从而克服传统识别方法中样本数据过大, 而且存在误差和干扰, 严重影响识别速度和效果的困难. 改善了神经网络模式识别的效果.

关键词: 方差分析; 神经网络; 模式识别

中图分类号: TP18 **文献标识码:** A **文章编号:** 0529-6579 (2001) 05-0021-04

模式识别中, 由专业知识及有关经验同时结合对象的特点选择样本的特征指标(自变量), 并采用多传感器测量. 这时由于目标的复杂性, 特征指标往往过多, 从而使神经网络进行模式识别过于复杂, 同时在测量学习样本的特征指标时存在误差, 所以在用神经网络进行模式识别时, 当样本特征指标太多时, 应该在不减少信息量的同时对数据进行处理, 以提高神经网络识别的性能.

1 样本数据处理方法

用多传感器获取反映样本不同属性的原始测量数据, 以此和目标库中的标准目标参数进行比较或运算来确定目标的类型, 这种比较或运算往往是非线性的, 用神经网络识别是上等之选. 设有 n 个样品, 每个样品用多传感器测出其 m 个特征指标, 此时共有 $n \times m$ 个数据, 当 m 过大时, 指标之间往往互有影响, 而且都存在误差. 从 m 个指标中寻找出很少几个能反映样本特征的综合性指标, 同时又能尽可能多地反映原来资料的信息(也就是方差尽可能的大), 显然这是处理数据的好方法.

设每个标准目标的特征指标通过多传感器测量, 并标准化处理为一向量 $p_i = (x_{i1}, x_{i2}, \dots, x_{im})$, 将 n 个标准目标(样本)的数据组成一个矩阵, 本文称之为样本矩阵, 记为 $P = (p_1, p_2, \dots, p_n)^T_{n \times m}$. 令 p_i 是取自 P 的随机样本, 并且 $E(P) = \mu$, $\text{Var}(P) = V$. 求 P 的线性函数 $a^T P$ 使得 $a^T P$

方差(即包含的信息)尽可能的大, 由于 $\text{Var}(ca^T P) = c^2 a^T \text{Var}(P) a$, 所以对 a 不加限制时, 问题变成没有意义了. 于是限制 $a^T a = 1$. 令 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0 = \lambda_{r+1} = \dots = \lambda_m$ 是 V 地全部特征根, γ_i 是对应的特征向量, 且 $\|\gamma_i\| = 1$, 于是有下面地定理.

定理

$$\begin{aligned} \max_{a^T a = 1} a^T V a &= \lambda_1 = \gamma_1^T V \gamma_1 \\ \max_{a^T a = 1} \{a^T V a\} &= \gamma_j^T V \gamma_j, j = 2, 3, \dots, r \\ a^T \gamma_i &= 0, i = 1, 2, \dots, j-1 \end{aligned}$$

也即是独立的综合指标包含的最多信息依次为 $\lambda_1, \lambda_2, \dots, \lambda_r$.

证明 因是半正定, 可分解为 $V = Q^T \text{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0) Q$. 其中, Q 为正交方阵, 且

$$Q = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix}$$

$$\begin{aligned} \text{故} \quad \max_{a^T a = 1} a^T V a &= \max_{a \neq 0} \frac{a^T V a}{a^T a} = \\ &= \max_{a \neq 0} \frac{a^T Q^T \text{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0) Q a}{a^T Q^T Q a} = \\ &= \max_{a \neq 0} \frac{a^T Q^T \text{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0) a}{a^T a} = \lambda_1 \end{aligned}$$

同理可证其它结论.

又因为 $\gamma_i^T V \gamma_j = 0, i \neq j$, 所以这些含有最多信息的综合指标 $\gamma_1^T P, \gamma_2^T P, \dots, \gamma_r^T P$ 是不相关的. 证毕.

* 收稿日期: 2001-02-16; 作者简介: 罗中良 (1968-), 男, 讲师.

对于模式识别时所选取的实际标准样本 $\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n)^T_{n \times m}$, 不妨设 $\bar{\mathbf{P}} = \frac{1}{n} \mathbf{P}^T \mathbf{I} = 0$, 则协方差阵的极大似然估计为 $\mathbf{S} = \frac{1}{n} \mathbf{P}^T \mathbf{P}$, 将 $\mathbf{S}$ 代替协方差阵 $\mathbf{V}$ 就可得到样本的综合信息指标. 一般地, 样本特征指标采用传感器测量时都存在误差(可能由随机噪声引起), 所以如果选取全部综合指标, 一则不能简化网络, 二则对测量误差不能克服. 选择综合指标的标准为^[1]: $\frac{\lambda_1 + \lambda_2 + \dots + \lambda_l}{\text{tr}(\mathbf{V})} \geq 85\%$. 这样就减小了测量误差或干扰的影响, 从而可以保证模式识别效果的提高.

2 神经网络模式识别

1985 年美国 MIT 的 PDP 小组提出 BP 模型是近年来应用最广泛的网络之一. 其标准的学习算法是一种简单的最速下降静态寻优算法^[2], 即:

$$\Delta W(n) = \eta \frac{\partial E}{\partial w}(n-1) + \alpha \Delta W(n-1)$$

其中, η 是学习速率, α 是惯性因子.

当 BP 网络用于模式识别时, 将传感器检测待识别目标的特征指标并标准化为输入向量 $\mathbf{p}_i = (x_{i1}, x_{i2}, \dots, x_{im})$. 用上面介绍的样本数据处理方法, 选出代表最多信息的几个综合指标作为输入, 并抽出了 n 个样本, 对每种模式定出一个目标输出作为属于该模式的样本的目标输出. 先随机地选出 n_1 个样本作为学习教材. 经训练后可建立

模式识别的智能识别系统. 为了检验所建立识别系统可将剩下的 $n - n_1$ 个样本输入识别系统看是否能作出正确的模式识别, 若不能, 则要重新调节学习速率 η , 惯性因子 α 和隐蔽节点数.

3 神经网络模式识别实例

模式识别时, 通过测量反映不同对象或目标的多种特征参数, 再用各种方法来推断待识别的对象类型. 本文以神经网络对茶叶品质进行等级识别, 来说明上述数据处理方法的有效性. 茶叶品质由多种特征决定, 它们可采用多种不同的传感器或生化分析来测量. 下面是商业部颁布的 4 种炒青绿茶标准样 (1~6 级) 和 6 个市售样, 凭着对茶叶的专业知识和经验, 人们选出了它的 10 个特征指标, 其原始数据^[3]如表 1. 将表 1 中 30 组样本数据的前 24 个样本作为神经网络的学习教材, 后 6 个样本作为预测样本, 并将学习教材标准化.

首先, 采用一般的 BP 神经网络对茶叶品质等级进行识别, 将茶叶的 10 个特征指标都作为神经网络的输入, 将 6 个样品等级的目标输出依次定为:

0.9, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1 (等级 1)
 0.1, 0.9, 0.1, 0.1, 0.1, 0.1, 0.1 (等级 2)
 0.1, 0.1, 0.9, 0.1, 0.1, 0.1, 0.1 (等级 3)
 0.1, 0.1, 0.1, 0.9, 0.1, 0.1, 0.1 (等级 4)
 0.1, 0.1, 0.1, 0.1, 0.9, 0.1, 0.1 (等级 5)
 0.1, 0.1, 0.1, 0.1, 0.1, 0.9, 0.1 (等级 6)

表 1 商业部颁炒青绿茶标准样和市售样的感官级别及理化指标

Tab. 1 The original data of the physical and chemistry index of green-tea issued by ministry of commerce

指标	$w(\text{氨基酸})$ %	$w(\text{咖啡碱})$ %	$w(\text{水浸出物})$ %	$w(\text{粗纤维})$ %	叶底 L	叶底 a	叶底 b	茶汤 L	茶汤 a	茶汤 b	
婺炒青	1	2.49	4.13	39.61	11.36	29.1	-2.6	13.1	89.8	-1.8	14.3
	2	2.16	4.11	39.43	12.11	27.5	-1.9	12.3	87.4	-1.6	18.1
	3	1.87	3.87	39.40	13.18	26.8	-1.8	11.9	87.9	-1.7	18.3
	4	1.74	3.65	39.12	14.92	26.5	-1.6	11.5	87.8	-1.8	19.3
	5	1.39	3.35	39.15	14.92	26.3	-1.3	11.3	88.4	-1.5	21.8
	6	1.33	3.02	39.08	16.3	26.7	-1.1	11.5	84.4	-1.4	22
屯炒青	1	2.45	4	39.53	11.2	30.5	-1.7	12.8	91.2	-1.6	18
	2	2.18	4.13	39.37	11.88	29.2	-1.3	12.6	88	-1.5	18.7
	3	1.88	3.73	39.06	13.02	28.4	-1.4	12.3	87.8	-2	18.8
	4	1.86	3.34	38.8	13.78	26.8	-1.2	11.8	87.5	-2	18.9
	5	1.62	3.15	37.9	14.66	27.2	-1.4	11.5	87.4	-2	20
	6	1.51	2.96	36.57	16.43	27.2	-1.4	12.1	86.7	-1.9	19.7

(转下页)

(接上页)

指标		$w(\text{氨基酸})$ %	$w(\text{咖啡碱})$ %	$w(\text{水浸出物})$ %	$w(\text{粗纤维})$ %	叶底 L	叶底 a	叶底 b	茶汤 L	茶汤 a	茶汤 b
婺炒青	1	2.93	3.35	38.37	10.97	27.8	-1.3	11.5	90.4	-1.8	18.1
	2	2.54	3.31	37.51	12.01	27.3	-1.4	11.4	88.6	-2.1	22.5
	3	1.95	2.87	37.15	13.38	27.2	-1.3	11.3	88.3	-2.4	22.4
	4	1.86	2.57	35.94	14.61	26.3	-1.1	10.9	88.8	-2.5	22
	5	1.61	2.63	37.7	15.81	25.2	-1.2	10.3	87	-2.9	21.5
	6	1.47	2.2	36.73	16.95	25.9	-1.1	10.6	86	-2.7	20.6
遂炒青	1	2.74	3.76	39.69	11.56	29.5	-1.6	12.5	91.8	-1.8	17.1
	2	2.3	3.56	38.77	12.56	28.4	-1.4	12.3	90	-1.7	18.8
	3	1.92	3.51	38.6	14.06	27.6	-1.1	11.9	89	-1.6	20.2
	4	1.67	2.71	38.03	14.96	26.7	-1	11.7	88.7	-1.5	21.5
	5	1.61	2.63	37.7	15.81	26.3	-0.7	11.4	88.6	-1.5	21.9
	6	1.6	2.45	36.72	17.37	26.5	-0.6	11	88.7	-1.5	22.7
市售样	1	2.89	3.97	39.59	11.05	28.2	-1.6	12.7	90.6	-1.8	17.7
	2	2.19	3.85	39.71	12.48	28.5	-1.3	12.3	87.6	-1.8	21.4
	3	1.92	3.52	38.36	13.5	27.7	-1.1	12.4	86.6	-1.7	23.6
	4	1.78	2.89	38.04	14.02	26.4	-0.6	11.7	86.2	-1.7	25.2
	5	1.7	2.4	36.01	15.48	25.2	-0.8	11.2	86	-1.9	25.6
	6	1.42	2.43	34.32	17.14	24.6	-0.8	10.9	85.6	-1.6	26.7

本文选择 BP 网络的隐蔽层含 14 个节点，经过学习后建立了茶叶等级分类预测的智能识别系统。然后对表 1 中的后 6 个预测样本进行检测得到如表 2。

表 2 一般 BP 神经网络模式识别结果

Tab.2 The prediction results by original neural net model

神经网络的识别结果						识别样品等级	实际样品等级
0.774	0.153	0.127	0.102	0.110	0.085	1	1
0.280	0.366	0.296	0.079	0.201	0.054	2	2
0.065	0.487	0.778	0.149	0.011	0.359	3	3
0.117	0.081	0.095	0.523	0.417	0.079	4	4
0.077	0.172	0.071	0.404	0.510	0.138	5	5
0.129	0.094	0.054	0.096	0.069	0.943	6	6

根据对样本等级的目标输出，可知神经网络模式识别结果由表 2 一行中的最大值决定，最大值与次大值的差值大小可以衡量识别的好坏，差值越大，表明识别越明显。由表 2 可以看出，虽然 BP 网络能够正确地识别茶叶的等级，但是识别得比较勉强，而且收敛速度较慢，一旦测量误差大了一点，就很可能识别错误，如表 2 的第 2、4 个产品的等级识别就很模糊。

现在，用上面讨论的方法先对样本数据进行处理，再用神经网络识别对产品品质进行识别。把 24 个训练样本看作从一个 10 维的随机向量中抽出的随机样本，对其进行数据处理（此过程可在 S-Plus^[4]中实现），见表 3。

表 3 模式识别学习样本的方差

Tab.3 The variance of specimen in model identification

综合指标	1	2	3	4	5	6	7	8	9	10
方差	2.463	1.150	1.019	0.642	0.525	0.436	0.392	0.259	0.186	0.134
总比例/%	63.3	77.1	88.0	92.3	95.2	97.1	98.7	99.4	99.8	100.0

可以看出只需取前 3 个综合指标（因总比例大于 85%），且对应的 3 个特征向量为：

(-0.309, -0.367, -0.310, 0.357, -0.351, 0.301, -0.358, -0.248, -0.116, 0.340),
 (-0.459, 0.201, 0.378, 0.250, 0.000, 0.000, 0.207, -0.434, 0.555, 0.000),
 (0.134, -0.127, 0.000, 0.000, 0.201, 0.546, 0.000, 0.452, 0.564, 0.307)

由此可以求出训练样本的前 3 个综合指标,并以它们作为 BP 网络的输入,重新建立茶叶等级识别预测的智能识别系统,再对表 1 中的后 6 个预测样本进行等级识别,得到如表 4 数据.

表 4 对样本数据处理后的神经网络模式识别结果

Tab.4 The prediction results by modified neural net model

神经网络的识别结果						识别样品等级	实际样品等级
0.912	0.093	0.084	0.071	0.119	0.034	1	1
0.137	0.916	0.052	0.104	0.091	0.023	2	2
0.101	0.084	0.902	0.046	0.110	0.154	3	3
0.098	0.104	0.071	0.921	0.111	0.073	4	4
0.119	0.123	0.059	0.076	0.787	0.210	5	5
0.119	0.129	0.068	0.049	0.091	0.975	6	6

通过表 2 和表 4 的比较可以清楚地看出,经数据处理后的神经网络模式识别有着明显的改进.本文定义一个量 ρ 来衡量神经网络识别的好坏.由于目标输出已标准化,所以可以看出将 ρ 定义为最大数减去次大数比较合理.于是一般神经网络模式识别的 ρ 为: 0.621, 0.070, 0.291, 0.106, 0.106, 0.814, 而样本数据经过处理后的神经网络模式识别的 ρ 为: 0.793, 0.779, 0.748, 0.810, 0.577, 0.846 可以看出样本数据经过处理后的神经网络模式识别的 ρ 值显著地大于未处理时的 ρ 值,而且能容许大得多的测量误差,这样就可以更加清楚地看出本文的模式识别方法的效果(最明显的从第 2、4 个产品识别时的 ρ 值看).又由于神经网络的输入节点数明显地减少(例子中由 10 个减少为 3 个),故收敛速度大大提高.

4 结 论

用多传感器测量样本特征指标,并用人工神

经网络进行模式识别时,样本数据往往过大,影响识别速度和效果.本文提出基于样本信息量不减少和减小误差的原则下进行数据处理,减小了神经网络的输入节点,并通过对茶叶品质等级识别,说明此方法的有效实用.此方法可应用于所有样本型产品、目标的鉴定和识别中,亦可用于工业过程故障监测报警中.

参考文献:

- [1] 张尧庭,方开泰.多元统计分析引论[M].北京:科学出版社,1997. 28-30, 322-328.
- [2] 李士勇.模糊控制.神经控制和智能控制论[M].哈尔滨:哈尔滨工业大学出版社,1996.94-95.
- [3] 蔡煜东,官家文,甘骏人.茶叶品质评定的人工神经网络方法[J].茶叶科学,1993,13(1):45-50.
- [4] VENABLES W N, RIPLEY B D. Modern applied statistics with S - PLUS[M]. New York: Springer - Verlag Inc, 1997.328-346.

Method of Model Identification Using Neural Network Based on Multi - sensor

LUO Zhong-liang¹, MAI Yi-jia¹, YU Jian-feng²

(1. Department of Automation, Foshan University, Foshan Guangdong, 528000, China;

2. Department of Statistics, Yale University, New Haven, CT 06510, USA)

Abstract: Based on the criterion of lessening the errors with minimum loss of information, statistical methods are applied to capture some comprehensive variables from the sample. These methods greatly improve the identification results of the neural network.

Keywords: variance analysis; neural network; model identification