

基于 BP 神经网络的皮棉杂质在线检测方法

丁天怀, 李 勇, 苗君哲, 郑颖航
(清华大学)

摘 要: 为改进因皮棉含杂影响因素多, 现有杂质含量数学统计模型不精确, 测量精度不高的缺点, 用人工神经网络方法建立了基于 BP 神经网络的皮棉杂质含量数学模型。用自行设计的皮棉杂质测量系统提取皮棉杂质图像特征参数并进行处理。针对 BP 神经网络收敛慢的特点, 在实际算法中引入了动量项, 从而提高了网络收敛速度。试验结果表明, BP 神经网络模型拟合结果的相对剩余标准差为 1.76%, 拟合精度明显优于数理统计模型的拟合精度。

关键词: 皮棉杂质; 在线检测; BP 神经网络; 数据处理

中图分类号: TP183

文献标识码: A

文章编号: 1002-6819(2003)02-0137-03

1 引 言

皮棉含杂量的在线检测是棉花加工质量控制系统的核心, 传统的“机械称重法”^[1]因测量周期长, 检测过程烦琐, 显然不能满足在线检测的要求。皮棉杂质种类、形状、大小与分布存在随机性, 是皮棉含杂量在线检测的难点。20 世纪 70 年代美国科学家 R. L. Barker 和 D. W. Lyons 首次提出用图像处理方法对皮棉杂质进行分级, 并指出皮棉含杂量与杂质个数相关^[1,2]; 之后 R. A. Taylor 对杂质图像的研究表明皮棉含杂量还受杂质面积的影响^[3~5]。在以上研究的基础上 W. S. Anthony, B. D. Farah 等研制了基于图像处理方法的皮棉含杂量在线测量系统^[6~9]。虽然文献[3], [6], [7]和[8]给出了采用二元多项式拟合的含杂量经验公式, 但由于受回归方程精确度的限制, 测量精度不高。文献[9]没有考虑杂质分布的随机性, 给出的经验公式存在一定的局限性。本文引入了基于人工神经网络模型的数据处理方法, 解决了数理统计模型中回归方程不精确的问题, 与数理统计方法比较测量精度有明显提高。

2 皮棉杂质含量检测的 BP 神经网络模型

BP 神经网络(简称 BP 网络)是人工神经网络中最常用的一种, 已经广泛应用于模式识别、函数逼近、信号处理和系统控制等领域^[10]。BP 网络可以看成是一个从输入到输出的高度非线性映射, BP 算法把一组样本的 I/O 问题变成一个非线性优化的问题。如果输入结点数为 n , 输出结点数为 m , 则网络是从 R^n 到 R^m 的映射, 即:

$$F: R^n \rightarrow R^m \quad Y = F(X) \quad (1)$$

对于样本集合输入 X 和输出 Y , 可以认为存在某个映射满足:

$$y_i = G(x_i) \quad i = 1, 2, \dots, k \quad (2)$$

因此可利用 BP 网络的映射关系来处理皮棉杂质实验数据, 使 F 是 G 的最佳逼近。根据这一思路建立的皮棉杂质含量 BP 网络模型如图 1 所示。

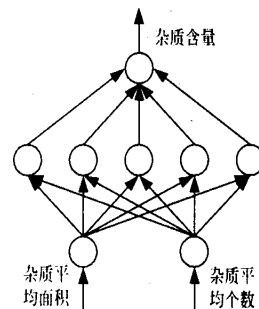


图 1 杂质含量 BP 神经网络模型

Fig. 1 Cotton trash content model based on BP neural network

BP 网络由 3 层组成, 其中输入层的两个节点分别对应杂质平均面积和杂质平均个数, 输出层对应皮棉杂质含量, 中间层包含了 5 个隐层节点。

网络中第 k 层第 j 个神经元的输入输出关系为:

$$I_j^k = \sum_{i=1}^N w_{ij}^k x_i^{k-1} - \theta_j^k \quad (3)$$

$$y_j^k = f(I_j^k) \quad (4)$$

式中 y_j^k ——第 k 层第 j 个神经元的输出; y_{ij}^k ——从第 $(k-1)$ 层第 i 个神经元到第 k 层第 j 个神经元之间的连接权值; θ_j^k ——阈值; $f(I)$ ——激活函数, 选用了 sigmoid 函数:

$$f(I) = 1/[1 + \exp(-I)] \quad (5)$$

针对 BP 算法学习速度慢, 容易陷入局部最小的缺点, 实际算法中引入了动量项, 减少了过调量, 从而改善了收敛性。改进的 BP 算法计算公式为:

$$\Delta w_{ij}(n) = a \Delta w_{ij}(n-1) + \eta \delta_j(n) y_i(n) \quad (6)$$

上式中第 2 项为常规 BP 算法的修正量, 第 1 项即为动量项, a 为学习率, η 为学习步长。改进的 BP 算法框图如图 2。

收稿日期: 2002-08-17

作者简介: 丁天怀(1945-), 男, 教授, 博士生导师, 系主任, 北京市海淀区清华园 清华大学精密仪器系, 100084。Email: dingth@p.m.tsinghua.edu.cn

3 皮棉杂质测量系统

为对杂质面积和杂质个数进行测量,建立了如图 3 所示的皮棉杂质测量系统,包括硬件和软件两部分。

系统中图像采集卡的图像最大采集分辨率为 $700 \times 560 \times 8 \text{ bit}$;黑白 CCD 摄像机的水平清晰度为 600 线,总像素为 473820;光照箱为 $300 \times 200 \times 200 \text{ mm}^3$,箱内装有 4 只白炽灯。光从与窗口法线方向呈 45° 角的方向对压紧在采样窗口的棉样进行照射。CCD 前加装滤色镜,选通 $600 \sim 700 \text{ nm}$ 的光线,以达到最佳的图像效果。

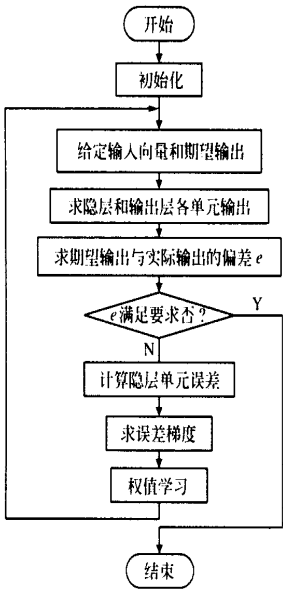
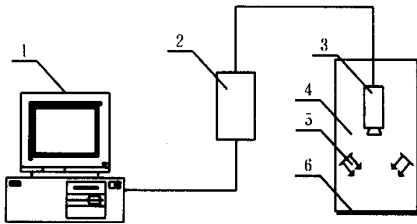


图 2 改进 BP 算法框图
Fig 2 Improved BP algorithm flow chart



1. 计算机 2. 图像采集卡 3. CCD 摄像机
4. 光照箱 5. 光源 6. 采样窗口
图 3 皮棉杂质测量系统
Fig 3 Cotton trash measurement system

系统图像处理软件首先对捕捉到的皮棉杂质原始图像与同等光照条件下获取的标准白板光照图进行差运算以消除光照不均匀对图像的影响;其次利用对数拉伸法对图像进行灰度拉伸来增大杂质和皮棉背景的灰度差;再次对拉伸后的图像采用中值滤波方法进行平滑滤波衰减噪声;最后利用可变区域比例阈值法对皮棉杂质图像进行分割并提取杂质面积和杂质个数等特征参数。

4 试验数据处理与分析

用上述测量系统对 50 个标准棉样进行多次采样平

均,提取杂质面积和杂质个数,部分试验数据见表 1。其中: X, Y 为试验测得的杂质平均面积和平均个数; Z 为标准棉样的含杂率。

表 1 皮棉杂质检测数据

Table 1 Measured data of cotton trash content							
数据点	1	2	3	4	5	6	7 ...
$Z/\%$	0.7	0.74	0.83	0.91	1.01	1.16	1.38 ...
$X/\text{像素}$	66.3	73.2	78.1	85.7	91.4	124.1	143.9 ...
$Y/\text{个}$	3.9	4.1	4.4	4.9	5.6	8.4	9 ...

文献[8]给出了含杂量数理统计基本模型:

$$Z = a + bX + cY + dX^2 + eY^2 + fXY \quad (7)$$

对上述试验数据用最小二乘法回归的经验公式如下:

$$Z = 0.2491 - 8.1683 \times 10^{-3}X + 0.1506Y + 3.1264 \times 10^{-5}X^2 - 6.0159 \times 10^{-4}Y^2 + 6.1565 \times 10^{-4}XY \quad (8)$$

在图 1 所示的杂质含量 BP 网络模型中,选用改进 BP 算法,取 $\epsilon = 0.005, \alpha = 0.95, \eta = 0.3$,对 50 组试验数据进行学习。用学习好的 BP 网络和回归经验公式拟合结果如表 2 所示。表中, Z_{nc} 和 δ_1 分别为神经网络拟合值和拟合值相对误差; Z_{rc} 和 δ_2 分别为数理统计拟合值和拟合值相对误差。

表 2 两种处理方法的拟合结果

Table 2 Fitted values of double processing methods					
数据点	$Z/\%$	$Z_{nc}/\%$	$Z_{rc}/\%$	$\delta_1/\%$	$\delta_2/\%$
1	0.70	0.6893	0.7512	1.5319	7.319
2	0.74	0.7471	0.7793	0.9526	5.305
3	0.83	0.8039	0.8153	3.1368	1.774
4	0.91	0.8956	0.8735	1.5865	4.011
5	1.01	0.9706	0.9448	3.9008	6.457
6	1.16	1.1902	1.2097	2.6056	4.2838
7	1.38	1.3679	1.2879	0.8783	6.6723
...

图 4 是两种处理方法的相对误差分布图。可见神经网络处理方法的相对误差分布图窄而尖;数理统计方法的相对误差分布图宽而平,这说明神经网络处理方法比数理统计处理方法误差小,精度高。

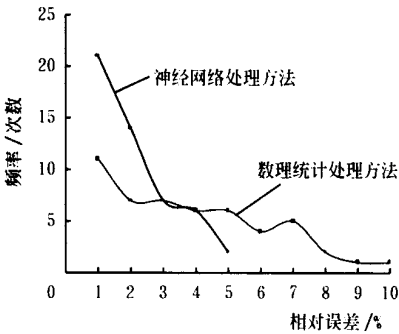


图 4 两种处理方法的相对误差分布图
Fig 4 Relative error distribution diagrams of double processing methods

两种处理方法的性能指标如表 3 所示。

表 3 两种处理方法的性能指标

Table 3 Performance indexes of double processing methods

处理方法	剩余标准差 S	相对剩余标准差 S_r
数理统计	0.135338	4.11%
神经网络	0.024892	1.76%

表 3 中, 剩余标准差 S 和相对剩余标准差 S_r 的计算式如下^[11]

$$S = \sqrt{(z_i - \hat{z}_i)^2 / (n - m - 1)}$$

(9)

$$S_r = S / \bar{z}$$

(10)

式中 n ——样本容量; m ——自变量个数。从表 3 可以看出神经网络处理方法的剩余标准差仅为数理统计方法的 18.4%, 相对剩余差则比数理统计方法小 2.35%。说明神经网络处理方法拟合结果准确度明显优于数理统计方法。

5 结 论

- 1) 在 BP 网络的学习过程中, 引入动量项可以减少过调量从而改善网络收敛性。
- 2) 利用 BP 网络对皮棉杂质试验数据进行处理, 精度明显高于数理统计模型的拟合精度, 从而为皮棉含杂量在线精确检测及棉花加工质量控制系统的实现奠定了基础。

[参 考 文 献]

[1] Barker R I, Lyons D W. Instrumented procedures for analysis of non-lint particles in cotton[J]. ASME Journal

of Engineering for Industry, 1976, 98(3): 845~ 848

[2] Lyons D W, Barker R I. Comparative analysis of trash contamination in cotton by optical scanning [J]. ASME Journal of Engineering for Industry, 1977, 99(1): 46~ 50

[3] Taylor R A. Using high-speed image analysis to estimate trash in cotton [J]. ASME Journal of Engineering for Industry, 1985, 107(2): 206~ 219

[4] Taylor R A, Godbey L C. Using digital images to measure and discriminate small particles in cotton [C]. Proceedings of SPIE—The International Society for Optical Engineering, 1991, 16~ 27.

[5] Taylor R A. Measuring leaf material in ginned cotton from surface images [C]. Proceedings of SPIE- The International Society for Optical Engineering, 1995, 218~ 225

[6] Anthony W S. Comparison of video color/trash meters for measurements on cotton [J]. Applied Engineering in Agriculture, 1989, 5(1): 18~ 23

[7] Anthony W S. Online assessment of foreign matter in cotton during ginning [J]. Applied Engineering in Agriculture, 1990, 6(3): 267~ 272

[8] 苗君哲. 用图像处理法在线检测棉花杂质的试验研究[D]. 北京: 清华大学精密仪器与机械学系, 2000

[9] Farah B D. Measurement of trash contents and grades in cotton using digital image analysis [C]. International Conference on Signal Processing Proceedings, 1996, 2, 20 Oct 14~ 18

[10] 王 雪, 丁天怀, 刘旭平. 基于遗传模糊神经网络算法的棉花轧花过程智能监控方法研究 [J]. 农业工程学报, 1998, 14(1): 204~ 208

[11] 钟继贵. 误差理论与数据处理 [M]. 北京: 水利电力出版社, 1993

On-line measurement of trash contents in cotton based on BP neural network

Ding Tianhuai, Li Yong, Miao Junzhe, Zheng Yinghang

(Department of Precision Instruments, Tsinghua University, Beijing 100084, China)

Abstract To improve the existing trash content, in cotton statistical models, which are inaccurate and imprecise in measurement as there are many factors influencing trash content in cotton, a model based on the BP neural network was constructed. A measurement system was designed to pick up image feature parameters of trash in cotton and to dispose these parameters. As for the slow convergence rate of BP algorithm, a momentum item was introduced into BP algorithm so that the convergence rate was increased. Experimental results show that the relative residual standard deviation of the fitted value of the BP neural network is 1.76% and the accuracy of its fitted value is much higher than that of other statistical models.

Key words: trash in cotton; on-line measurement; BP neural network; data processing