

# 基于 MC 方法和 BP 网络的印刷质量 控制图模式识别研究

初红艳, 李 鹏, 蔡力钢, 李风光

(北京工业大学 北京市先进制造技术重点实验室, 北京 100124)

摘 要: 建立了印刷质量控制图模式的数学模型, 基于 Monte Carlo (MC) 方法模拟印刷质量数据, 生成仿真样本, 并使用标准变换和线性编码方法进行预处理, 既不丢失样本数据的内在特征, 又大大降低了数据复杂度. 通过实验确定了结构为 24-18-16-4 的 4 层反向传播 (back propagation, BP) 网络模型, 并采用比例共轭梯度训练算法, 提高了网络的稳定性和收敛速度. 在对控制图模式识别时, 采用不同训练样本容量的实验方案, 模式识别正确率达 95.87%. 结果表明, 该方法可以提高印刷企业的质量控制水平和自动化程度.

关键词: 印刷质量控制图; 模式识别; Monte Carlo 方法; 反向传播神经网络

中图分类号: TS 805.3

文献标志码: A

文章编号: 0254-0037(2011)06-0816-06

要实现印刷品的高质量、高效率, 印刷的工艺过程就必须处于稳定受控的状态. 统计过程控制 (statistical process control, SPC) 是一种面向过程的质量管理方法, 能实时监控印刷过程中的异常波动, 使质量控制做到预防为主<sup>[1-2]</sup>. 随着人工智能技术的发展, 印刷企业自动化、信息化程度不断提高, 通过建立反向传播 (back propagation, BP) 神经网络对 SPC 控制图进行识别, 不需要精确描述变量间的非线性映射关系, 就能实现控制图的识别由人工观察到智能的判断, 这将大大提高印刷企业质量管理水平和自动化程度<sup>[3-4]</sup>.

## 1 模式的分类及数学模型

### 1.1 印刷质量控制图模式的分类

本研究将印刷质量控制图分为单一墨区质量控制图和整幅画面质量控制图. 单一墨区质量控制图是对样张每个墨区的质量特性值进行监控, 而整幅画面质量控制图则是从全局角度, 对样张所有墨区的质量特性值取均值加以监控. 目前, 根据异常模式的不同形态, 一般将控制图的异常划分为阶跃模式、趋势模式、周期模式等, 阶跃异常又可分为向上阶跃和向下阶跃, 趋势异常又可分为上升趋势和下降趋势<sup>[5-7]</sup>. 正常模式与发生异常时基本模式如图 1 所示. 图中 CL 为中心线 (control limit), UCL 为上控制线 (upper control limit), LCL 为下控制线 (lower control limit). 这种分类方式, 基本涵盖了印刷过程

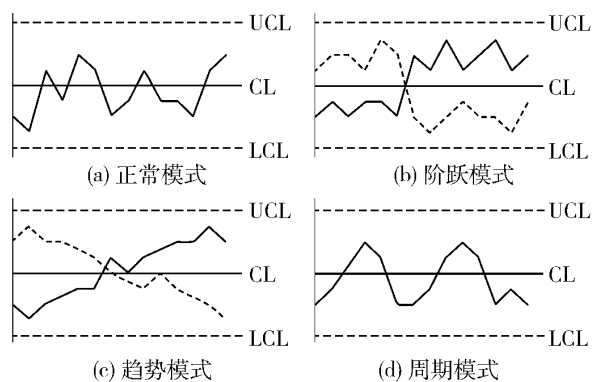


图 1 印刷质量控制图基本模式分类

Fig. 1 Basic patterns of control chart printing quality

收稿日期: 2009-07-07.

基金项目: 国家“十一五”科技支撑计划课题 (2006BAF03B01); 北京市教委科研计划项目 (KM200910005006).

作者简介: 初红艳 (1972—), 女, 黑龙江哈尔滨人, 博士, 副教授.

的大部分异常情况,比如环境温湿度变化、墨量变化、印刷机的性能失稳、水墨失衡、工人误操作等。

## 1.2 数学模型

要深入研究印刷模式的类型,就要建立各个模式相应的数学模型。数学模型一方面要符合印刷过程实际,具有明显的物理意义,另一方面要便于识别研究。

$t$ 时刻的印刷质量特性值可描述<sup>[8-9]</sup>为

$$x(t) = \mu + r(t) + d(t) \quad (1)$$

式中, $x(t)$ 为 $t$ 时刻测量得到的印刷质量特性值; $\mu$ 为受控情况下印刷质量特性值的均值; $r(t)$ 为 $t$ 时刻印刷质量特性值的正态分布随机偏差,为偶然因素引起的随机干扰, $r(t) \sim N(0, \sigma^2)$ ;  $d(t)$ 为 $t$ 时刻异常因素导致的印刷质量特性值的偏差,直接反映了各个模式的区别。

针对 $d(t)$ 建立各个模式的模型如下。

正常模式:  $d(t) = 0$ 。

阶跃异常模式:  $d(t) = \pm 1(t - t_0)\alpha$ 。式中, $+$ 代表向上阶跃, $-$ 代表向下阶跃; $t_0$ 为发生阶跃的起始点; $1(t - t_0)$ 为阶跃异常系数, $t \geq t_0$ 时, $1(t - t_0) = 1$ ;  $t < t_0$ 时, $1(t - t_0) = 0$ ;  $\alpha$ 为阶跃幅值。

趋势异常模式:  $d(t) = \pm v(t - t_0)\beta$ 。式中, $+$ 代表逐渐上升, $-$ 代表逐渐下降; $t_0$ 为发生趋势的起始点; $v$ 为趋势异常系数, $t \geq t_0$ 时, $v = 1$ ;  $t < t_0$ 时, $v = 0$ ;  $\beta$ 为趋势斜率。

周期异常模式:  $d(t) = A \sin(2\pi(t - t_0)/T)$ 。式中, $A$ 为波动的幅值; $t_0$ 为发生周期的起始点; $T$ 为波动的周期长度,即每个周期包含点子数量。

## 2 基于 Monte Carlo (MC) 方法和 BP 网络的模式识别

### 2.1 MC 方法生成样本数据

基于神经网络的模式识别,需要充足的训练样本数据以保证较高的识别率。当前一般采用MC方法<sup>[10-11]</sup>模拟实际生产过程,产生仿真数据用以训练识别。MC方法,又称统计试验方法,是指通过建立概率模型,利用随机抽样技术产生仿真数据的方法。在模拟数据时,随机数取值范围必须严格符合印刷生产的实际,而且需要较大的数据容量。

根据印刷质量控制图模式分类以及抽样检测的实际,对 $d(t)$ 中不同模式的参数进行取值,如表1所示。按照给定步长产生数据序列,利用MC方法,生成符合均匀分布的数据集,同时生成 $r(t) \sim N(0, \sigma^2)$ 的随机偏差数据集,其中 $\sigma$ 为原始样本数据的标准差,为不失一般性,本文取原始样本数据的 $\mu = 0$ ,  $\sigma = 1$ 。基于建好的数学模型,产生训练样本组 $x(t)$ 。

表1 MC方法生成随机数的取值范围

Table 1 Range of random data generated by Monte Carlo method

参数	取值范围	步长	意义
$t$	$1 \leq t \leq 24$	1	控制链长,即检测次序
$t_0$	$3 \leq t_0 \leq 9$	1	异常模式出现的初始位置
$\alpha$	$1.25 \leq \alpha \leq 3, -3 \leq \alpha \leq -1.25$	0.25	阶跃异常的阶跃幅值
$\beta$	$0.125 \leq \beta \leq 0.3, -0.3 \leq \beta \leq -0.125$	0.025	趋势异常的变化斜率
$A$	$1.25 \leq A \leq 3$	0.25	周期异常的周期幅值
$T$	$3 \leq T \leq 7$	1	周期异常的周期长度

### 2.2 数据的预处理

样本数据的描述形式对神经网络性能有着十分重要的影响。实验中,需要将MC方法模拟的数据进行预处理,以在保留数据内在特征的基础上,简化数据表示的复杂程度,从而提高网络的收敛性能和识别

率. 为增加网络模型的通用性 本文对原始样本数据进行 2 步处理 ,即标准化和线性编码.

首先通过公式  $X(t) = (x(t) - \mu) / \sigma$  进行标准化处理. 式中  $x(t)$  为原始样本数据;  $X(t)$  为标准化后的样本数据 ,此时  $X(t) \sim N(0, 1)$ .

在编码时 ,考虑  $X(t)$  的值在  $[-3, +3]$  的概率为 99.73% 将控制图区域划分为 14 个区域 ,每个区域对应一个编码值. 编码处理法则为

$$y(t) = \begin{cases} 3.5 - 0.25, & X(t) > 3 \\ -3 + 0.5k - 0.25, & -3 + 0.5(k-1) < X(t) \leq -3 + 0.5k, k=1, 2, \dots, 12 \\ -3 - 0.25, & X(t) \leq -3 \end{cases} \quad (2)$$

$y(t)$  为编码结果 ,即网络的输入.

### 2.3 建立 BP 网络模型

BP 网络是神经网络中最具代表性的网络之一 ,具有广泛的适应性和有效性 ,但存在速度慢、容易陷入局部极小点、算法不易收敛等缺点. 在此 ,将模值信赖域算法与共轭梯度算法结合<sup>[12]</sup> ,采用比例共轭梯度算法 ,以提高网络的稳定性和收敛性.

通过多次实验 ,确定了 4 层的网络结构 ,输入层节点数为 24 ,即连续 24 个印刷质量抽样检测数据; 目前对隐含层节点的确定方法还没有一个解析的关系式<sup>[12]</sup> ,实验过程中 ,依据网络性能 ,隐含层节点分别定为 18、16; 输出层节点为 4 ,即为印刷质量控制图的 4 种模式 ,相应模式对应的节点输出为 1 ,其他为 -1 ,基本模式与期望输出对应关系如表 2 所示.

表 2 基本模式与输出节点对应关系  
Table 2 Relationship of basic pattern and output node

基本模式	网络输出层 4 个节点				基本模式	网络输出层 4 个节点			
正常模式	1	-1	-1	-1	周期模式	-1	-1	1	-1
阶跃模式	-1	1	-1	-1	趋势模式	-1	-1	-1	1

网络各参数设置为: 收敛精度 0.001 ,最大训练次数 20 000 ,各层之间的传递函数分别为 tansig、tansig、purelin ,训练函数采用比例共轭梯度算法.

## 3 实验方案及结果

### 3.1 仿真结果分析

在型号 Intel Core2、CPU 2.33 GHz 的主机环境下 ,针对正常、阶跃、趋势、周期 4 种控制图模式 ,采用从小容量到大容量训练样本过渡的实验方案 ,选取  $4 \times 50$ 、 $4 \times 200$ 、 $4 \times 400$ 、 $4 \times 600$ 、 $4 \times 800$  计 5 种容量的训练样本组 ,为了不失一般性 ,测试样本容量选取为  $4 \times 200$  ,训练识别结果如表 3 所示 ,同时给出识别率为 95.87% 的  $4 \times 800$  样本组的每种模式的具体分布情况 ,如表 4 所示.

表 3 改进的 BP 网络对不同容量的识别结果  
Table 3 Recognition results of improved backpropagation network for different capacities

训练容量/个	各模式识别率/%				总体识别率/%
	正常模式	阶跃模式	周期模式	趋势模式	
$4 \times 50$	78.5	93.0	63.5	94.5	82.37
$4 \times 200$	86.0	96.0	80.0	97.5	89.87
$4 \times 400$	93.5	97.0	82.0	97.0	92.37
$4 \times 600$	93.0	99.0	87.5	98.0	94.37
$4 \times 800$	93.5	99.5	91.5	99.0	95.87

表 4 训练样本容量为  $4 \times 800$  时网络的识别结果  
Table 4 Recognition results of neural network of capacity of  $4 \times 800$

模式	正常模式结果/个	阶跃模式结果/个	周期模式结果/个	趋势模式结果/个	识别率 /%
正常模式	187	0	11	2	93.5
阶跃模式	0	199	0	1	99.5
周期模式	14	1	183	2	91.5
趋势模式	0	2	0	198	99.0

随着样本数量的增大,网络对于控制图各种模式的识别率逐渐提高,且逐渐趋于稳定,达到较高的水平.进一步分析表 4 中模式识别的具体分布情况可知,噪声较大的正常型与幅值较小的周期型不易区分,幅值较小的阶跃型与斜率较小的趋势型很接近,使识别结果出现小的错误.在实际中这些特征不明显的模式造成的影响是很小的.将上述样本容量  $4 \times 800$  训练好的网络保存,网络训练完成.

### 3.2 实例应用

实验选用北人集团 BR624A 型胶印机,美国 Intellitrax 自动测色扫描系统,江苏紫兴  $128 \text{ g/m}^2$  铜版纸,日本 HONG FAN 油墨,印刷速度  $7500 \text{ 张/h}$ ,抽样时间间隔为  $2 \text{ min}$ (约 250 张抽样 1 次).分别对印张的单一墨区和整幅画面检测,得到实地密度、网点增大、印刷反差和色差等质量特性参数<sup>[13]</sup>.本文针对 2 类控制图,以青色实地密度第 9 墨区和红色实地密度整幅画面为例证,检测数据如表 5、6 所示.

表 5 青色实地密度第 9 墨区的检测数据  
Table 5 Experimental data of the density of cyan for the 9th area

抽样序号	1	2	3	4	5	6	7	8	9	10	11	12
测量值	1.39	1.42	1.38	1.37	1.40	1.38	1.39	1.40	1.39	1.42	1.41	1.39
抽样序号	13	14	15	16	17	18	19	20	21	22	23	24
测量值	1.40	1.41	1.40	1.40	1.41	1.42	1.41	1.42	1.43	1.43	1.44	1.43

表 6 红色实地密度整幅画面的检测数据  
Table 6 Experimental data of the density of magenta for all areas

抽样序号	1	2	3	4	5	6	7	8	9	10	11	12
测量值	1.36	1.39	1.38	1.41	1.43	1.42	1.39	1.39	1.40	1.37	1.40	1.37
抽样序号	13	14	15	16	17	18	19	20	21	22	23	24
测量值	1.39	1.41	1.39	1.42	1.38	1.38	1.40	1.42	1.40	1.36	1.38	1.39

表 5 数据对应的控制图如图 2 所示,表 6 数据对应的控制图如图 3 所示.

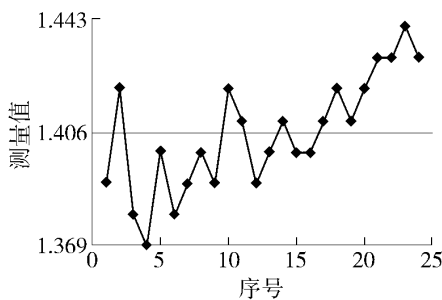


图 2 青色实地密度第 9 墨区质量控制图

Fig. 2 Control chart of density of cyan for the 9th area

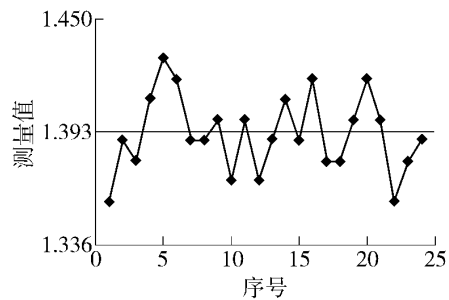


图 3 红色实地密度整幅画面质量控制图

Fig. 3 Control chart of density of magenta for all areas

为保持网络输入层数据的一致性,提高网络泛化能力,将表 5、6 中的实验数据按照 2.2 节方法进行预

处理,作为网络的输入,调用已保存的BP网络,得到表5的训练结果为 $[-0.9306, -1.0740, -1.0021, 1.0062]$ ;表6的训练结果为 $[0.9900, -1.0023, -0.9997, -0.9993]$ 。

依据表2,对应的模式分别为趋势异常模式、正常模式。结合图2、3以及印刷的实际状况,识别结果正确,且达到较高的精确度。其他的质量特性,比如网点增大、印刷反差和色差等,按照同样方法,分别对单一墨区质量控制图和整幅画面质量控制图训练识别,均获得了较为理想的结果。

## 4 结论

1) 确立了印刷质量控制图的4种模式分类,较好地涵盖了印刷过程中出现的异常情况,并建立相应模式的数学模型,在此基础上基于MC方法模拟印刷质量数据,生成网络样本。

2) 对于输入层数据,采用标准变换和线性编码方法进行预处理,在不失数据内在特征的基础上大大降低了数据复杂度,提高了网络的稳定性和收敛性能。

3) 实验确立了4层的BP网络模型,即输入层24节点,隐含层分别18、16个节点,输出层4个节点,并采用改进的BP训练算法,即比例共轭梯度算法,大大提高了网络的收敛速度。通过5种容量样本训练实验证明,大的训练样本能显著提高模式识别的正确率。当训练样本为 $4 \times 800$ ,测试样本为 $4 \times 200$ 时,得到模式识别率为95.87%,达到较高水平。

4) 选取青色实地密度第9墨区的质量控制图和红色实地密度整幅画面的质量控制图为例,利用BP网络训练,得到的识别结果与期望值非常接近。实例结果表明,该方法可以提高印刷企业的质量控制水平和自动化程度。

### 参考文献:

- [1] 王毓芳,郝凤. 过程控制与统计技术[M]. 北京: 中国计量出版社,2001.
- [2] 赵艳东. 统计过程控制(SPC)方法在印刷质量管理中的应用研究[D]. 武汉: 武汉大学印刷与包装系,2005.  
ZHAO Yan-dong. Study of statistical process control (SPC) application in printing quality management[D]. Wuhan: School of Printing and Packing, Wuhan University, 2005. (in Chinese)
- [3] 李刚,王霄,蔡兰. 基于神经网络的控制图模式识别技术研究[J]. 制造业自动化,2000,22(5): 31-34.  
LI Gang, WANG Xiao, CAI Lan. Research of control charts pattern Recognition using neural network [J]. Manufacturing Automation, 2000, 22(5): 31-34. (in Chinese)
- [4] 李孟清,陈志祥. 质量控制图中趋势模式的综合识别方法[J]. 华中理工大学学报,2000,28(5): 24-26.  
LI Meng-qing, CHEN Zhi-xiang. A synthetical approach of fuzzy logic and neural network for trend pattern recognition in control charts [J]. Journal of Huazhong University of Science and Technology, 2000, 28(5): 24-26. (in Chinese)
- [5] CHENG C S. A neural network approach for the analysis of control chart pattern [J]. International Journal of Production Research, 1997, 35(3): 667-697.
- [6] 管涛,费仁元,王民. 基于神经网络的控制图异常模式识别研究[J]. 北京: 北京工业大学学报,2006,32(8): 673-676.  
ZAN Tao, FEI Ren-yuan, WANG Min. Research on abnormal pattern recognition for control chart based on neural network [J]. Beijing: Journal of Beijing University of Technology, 2006, 32(8): 673-676. (in Chinese)
- [7] YANG J H, YANG M S. A control chart pattern recognition system using a statistical correlation coefficient method [J]. Computers & Industrial Engineering, 2005, 48(2): 205-221.
- [8] 乐清洪. 智能工序质量控制的理论与方法研究[D]. 西安: 西北工业大学机电工程学院,2002.  
LE Qing-hong. Research on the theory and methodology of intelligent quality control for machining operations [D]. Xi'an: School of Mechatronics Engineering, Northwestern Polytechnical University, 2002. (in Chinese)
- [9] 高清,陈平,马玉林. 基于神经网络的质量控制图自动识别技术的研究[J]. 哈尔滨工业大学学报,1995,27(6): 121-124.  
GAO Qing, CHEN Ping, MA Yu-lin. Artificial neural networks for the identification of quality control charts [J]. Journal of Harbin Institute of Technology, 1995, 27(6): 121-124. (in Chinese)

- [10] 徐钟济. 蒙特卡洛方法[M]. 上海: 上海科学技术出版社, 1985.
- [11] 殷显安. 试验模拟的蒙特卡洛方法[J]. 测试技术学报, 1994, 8(2): 49-51.  
YIN Xian-an. Monte Carlo method for test simulation[J]. Journal of Test and Measurement Technique, 1994, 8(2): 49-51. (in Chinese)
- [12] 闻新, 周露, 李翔, 等. Matlab神经网络仿真与应用[M]. 北京: 科学出版社, 2003: 258-284.
- [13] 初红艳, 蔡力钢, 郑刚, 等. 基于Intellitrac的印刷色彩检测与控制系统[J]. 包装工程, 2007, 28(10): 94-97.  
CHU Hong-yan, CAI Li-gang, ZHENG Gang, et al. Printing color detection and control system based on intellitrac[J]. Packaging Engineering, 2007, 28(10): 94-97. (in Chinese)

## Pattern Recognition for Printing Quality Control Chart Based on MC Method and BP Neural Network

CHU Hong-yan, LI Peng, CAI Li-gang, LI Feng-guang

(The Lab of Beijing Advanced Manufacturing Technology, Beijing University of Technology, Beijing 100124, China)

**Abstract:** A mathematical model for patterns of the printing quality control chart is established, and the data of printing quality is simulated based on Monte Carlo method. Then the complexity of the sample data is reduced by using the method of standard transformation and linear encoding. A 4-layer BP neural network model, as 24-18-16-4, is established through the experiments, and a scaled conjugated gradient training algorithm is adopted to enhance the stability and convergence of the network. The paper uses different capacity of training samples in pattern recognition for control chart, and the recognition accuracy achieves 95.87%. Results of experiments show that this method can improve the level of quality control and degree of automation for printing enterprise.

**Key words:** printing quality control chart; pattern recognition; Monte Carlo method; BP neural network

(责任编辑 吕小红)