

# 一种基于统计特征的手写体特殊数学符号识别方法

吴平平 杨戈 刘宏

北京大学深圳研究生院 集成微系统重点实验室 518055

北京大学 机器感知与智能重点实验室 100871

E-mail: wupingping33@yahoo.cn

**摘要:** 本文提出了一种采用 BP 神经网络来对无限制手写数学符号识别的方法。在选取基于统计特征值的基础上,对三层的 BP 网络进行训练。通过 200 个训练字符样本及 100 个测试样本,识别率达 93.67%。实验结果表明,该方法对手写体数学符号识别效果良好,不仅简化了传统识别的复杂度,而且提高了识别的准确性。

**关键词:** 神经网络 误差反传 字符识别

## A Statistical Feature Based Method for Recognizing Handwritten Special Mathematic Characters

Pingping Wu, Ge Yang, Hong Liu

Key Laboratory of Integrated Microsystem, Shenzhen Graduate School, Peking University, 518055

Key Laboratory of Machine Perception and Intelligence, Peking University, 100871

E-mail: wupingping33@yahoo.cn

**Abstract:** This paper introduces a method based on BP neural network to recognize unconstrained handwritten mathematic characters. On basis of statistic feature extraction, a three-layer BP network is trained. Via the experiment of 200 training and 100 testing character patterns, the recognition rate is 93.67%. The experiments show that the proposed approach has a good effect on handwritten mathematic characters. It not only simplifies the complexity of the traditional recognition, but also increases the recognition accuracy.

**Key Words:** neural network; back propagation; character recognition

### 1. 引言

手写体字符,由于字体的变化很大,传统的识别方法要达到高的识别率,有较大的困难,人工神经网络模式识别方法是这些年提出的新方法,为手写体字符识别提供了一种新手段,它具有传统技术所没有的优点:良好的容错能力、分类能力、并行处理能力和自学能力。

神经网络的模式识别过程如图 1 所示。预处理的目的是去除干扰、噪声及差异,将原图像变成适合于特征提取的形式,然后对预处理的图像进行特征抽取,最后进行判决分类,得到识别结果。判决依据是对图像样本进行特征选择和学习得到的,这是识别处理中必要工作。

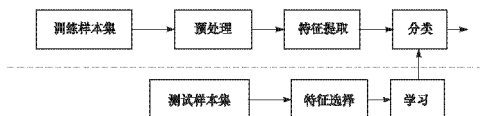


图 1 神经网络模式识别的方框图

## 2. 预处理和特征提取

### 2.1 样本集的获取与预处理

本实验系统训练和识别  $\emptyset$ ,  $\oplus$ ,  $\otimes$ ,  $\ominus$ ,  $\nabla$ ,  $\exists$ ,  $\partial$ ,  $\Delta$ ,  $\epsilon$ ,  $\zeta$  这 10 个数学符号, 它们是在屏幕上随意书写后, 由像素点提取程序提取数据并存入磁盘文件, 并对其进行尺度的归一化处理。图 2 给出部分原始数学符号图像和处理后的图像样本。

### 2.2 特征的选择与提取

在模式识别中, 特征的选择是一个关键点问题。任何识别过程的第一步, 无论是用计算机还是由人去识别, 都要首先分析各种特征的有效性并选出最有代表性的特征。

实验识别的对象是手写体数学符号图像, 属于字符识别。字符的特征可以分为结构特征和统计特征。

结构特征的提取重点在于确定以基元表示的结构信息, 主要有骨架、轮廓、笔画等得到的结构特征。骨架特征的提取极大的依赖于图像细化质量。由于现有的细化算法都或多或少地出现一些拓扑结构的改变, 如 Y 形分叉、毛刺、断线等。这就要求后继识别有较大的规则灵活性, 以骨架为主要特征的识别系统主要反映了对字符的二维分布情况判断。

轮廓也可以反映字符图像的结构, 轮廓大致有两种描述方式: 其一, 同骨架一样, 用结构点和弧构成图的方式来描述; 其二, 用标准化后的轮廓距边框的距离描述。对骨架引入更精确的描述将节省细节化的运算量, 但容易受到笔画宽度和断线的影响, 它较适合于图像质量较好和规范的字体。

如果从统计特征来定义手写字符, 能较大程度上反映字符的固有几何分布特性。统计特征是从原始数据中提取与分类最相关的信息, 使类内差距极小化, 类间差距极大化。选择的特征应对同一类字符的形变变化较小。统计特征可以分为全局特征和局部特征。

全局特征是对整个字符图像进行变换, 将变换的系数作为图像的一种特征, 主要包括 KL (Karhunen-Leeve)、Fourier 变换、Hadamard 变换、Cosine、Sine、Gabor 变换、RaPid 变换、Hough 变换、矩阵变换、代数特征、小波变换、粗外围特征等。局部特征是在特定的位置对特定大小窗口内的图像进行变换, 主要包括局部灰度特征、投影特征、方向线像素特征等。

特征提取的方法取决于实际的应用状态, 实验表明任何一种特征都很难完整地表达手写字符的所有一般特征和特异特征。

粗网格特征提取方法是先把待识别字符进行大小和位置的归一化, 再等分成  $M \times N$  个网格, 然后依次统计各网格内黑像素的数量, 从而取得一个以数值表示的  $M \times N$  维的网格特征。粗网格特征属于统计特征中的局部特征, 又称为局部灰度特征。它通过把字符分成  $M \times N$  个网格, 统计每个网格中的像素数量, 而每个网格各自反映字符的某一部分特征, 在识别阶段, 把各个网格组合起来作为字符的统计特征, 以此来识别字符。粗网格特征反映了字符的整体形状分布, 但字符的错切和上下左右部分的关系变化对特征的影响较大, 由于笔画、结构的相互位置的变化, 如字符的倾斜、偏移等导致了对应网格之间的错位, 大大降低了识别的正确率。

这里采用粗网格的特征提取方法, 并且对粗网格特征提取方法做了一些简化, 不但程序实现比较简单, 且具有一定的抗畸变干扰能力, 将归一化之后的字符图像划分成  $8 \times 8$  个网格 (如图 3 所示), 然后统计每个网格的黑像素的个数, 及网格内含有字符笔画像素的个数, 当含有笔画的个数没有达到网格总像素的 20% 时, 将这些像素当成干扰信号, 将这个网格设为

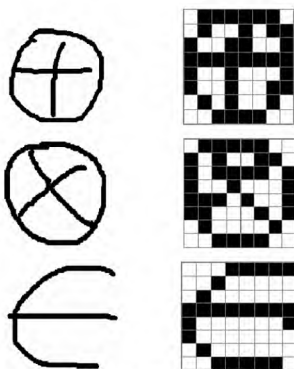


图 2 部分原始图像和处理后的图像样本

背景信息 0，否则将其设为字符信息 1，得出的 0、1 组成的矩阵即为字符的特征向量。

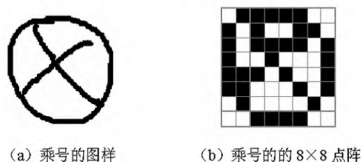


图 3 “X” 的点阵提取

本系统选择手写体数学符号 8×8 点阵的水平投影、垂直投影和左周边特征作为系统输入层的节点，图 4 (a) 所示是数字 “X” 的 8×8 点阵，它在水平方向的直方图如图 4 (b) 所示，它在垂直方向的直方图如图 4 (c) 所示，它在左周边遇到第一个黑点的序号如图 4 (d) 所示，每个数字的这 24 个特征的组合是不一样的，这样就可以将它们区别开来。由这 24 个特征训练的 BP 网络的数字识别率为 93.6%，使用者还可以添加输入自己的习惯字体并让系统重新训练，其识别率会进一步提高，将手写数字的点阵从 8×8 提高到 32×32，如果将手写数字 32×32 的点阵也作为特征输入，并引进更多的诸如游程统计特征、矩特征和 Gabor 小波特征等特征，系统的识别率能达到 95% 以上。

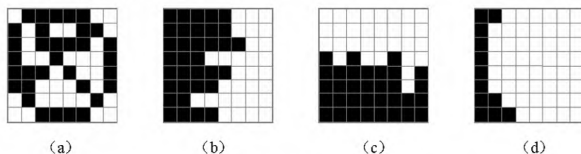


图 4 “X” 的各种特征的选取

提取水平、垂直和水平方向扫描线第一次遇到的黑点的序号数字特征，所得到的数据作为神经网络输入层的输入数据。在生成训练集时，不但要把上述数据存入磁盘文件，同时还要将目标值也分别写入磁盘文件，作为训练集数据。输入与期望输出值得对应关系如下表所示：

表 1 输入与输出的对应关系

输入	∅	⊕	⊗	⊖	∇	∃	∂	Δ	∈	∉
输出	0	1	2	3	4	5	6	7	8	9

### 3. 识别和训练

#### 3.1 拓扑结构

这里讲的 BP 神经网络的拓扑结构采用的是三层拓扑结构，即一个输入层、一个隐含、一个输出层。输入层的神经元包括水平方向的直方图、垂直方向的直方图和左周边轮廓特征，它选择为 8+8+8=24；隐含层的神经元个数为 20；对于输出层来说，因为识别的字符包括 10 个数学符号，因此输出端采用 4 个神经元作为输出。

#### 3.2 输出编码

关于输出编码问题，因为 BP 网络的训练时根据输出值与标准值之间的误差来调整权重，因此必须对目标期望输出进行编码，这里采用的是 8421 编码，即对于输出 “0”，采用 (0, 0, 0, 0)，对于 “3”，采用 (0, 0, 1, 1)，对于 “9”，采用 (1, 0, 0, 1)，但是，因为激

活的 S 型激活函数的输出永远不可能达到 0 或者 1，因此对于 (0, 0, 0, 0) 这样的标准输出，BP 网络无法收敛。这里为了避免这个问题，用 0.0001 代替 0，用 0.9999 代替 1，其他一次类推。

### 3.3 网络的训练

BP 训练过程如下：

1) 工作信号的正向传播：输入信号从输入层经过隐含层，传向输出层，在输出端产生输出信号。这是工作信号的正向传播。在信号传递的过程中网络的权值是不变的，每一层神经元的状态只影响下一层神经元的状态。如果在输出层得到的输出和期望输出的偏差较大，则转入误差信号的反向传播。

2) 误差信号的反向传播：网络的实际输出和期望输出的差值就是误差信号。误差信号的反向传播就是误差信号从输出端传向输入端。在这个过程当中，网络的权值由误差反馈进行调节。通过不断的修改网络的权值从而使得网络的输出不断的逼近期望值。在反向传播算法中通常采用梯度法修正权值，为此要求输出函数可微，通常采用 Sigmoid 函数作为输出函数。

BP 网络的训练方法有两种：一种是每输入一个样本修改一次权值；一种是批处理方式，即训练完全部样本之后再修改权值。这里采用的方法是前一种方法。样本采用图片方式，将训练样本图片进行特征提取之后送入 BP 网络进行训练。

(1) 初始化权值，设定最大迭代次数、最小均方误差 Min\_ex、训练步长 eta、输入样本个数 Num、BP 网络的相关系数 mutmodulus 等参数。

(2) 将第 K 个样本的特征向量输送到输入层，同时将预订的输出结果输送到理想输出单元。

(3) 调用前向激活函数 AheadCalculate 分别计算隐含层、输出层每个节点的输出。

(4) 调用函数 OutPutError 计算输出层每个神经元的输出值与标准值的误差。

(5) 根据输出层的误差，调用函数 HiddenError 计算隐含层每个神经元的输出误差。

(6) 根据输出层的误差，调用函数 AdjustWeights 调整隐含层与输出层之间的权值。

(7) 根据隐含层每个神经元上的误差，调用函数 AdjustWeights 调整输入层与隐含层之间的误差。

(8) 统计输出层的平方差，并进行叠加。

(9) 判断样本是否训练完，是则转如第 10 步；否则输下一个样本的特征向量，转入第 3 步，继续执行。

(10) 计算均方误差，判断是否小于最小值 Min\_ex，是则中断循环，转入 12 步执行；否则转入第 11 步继续执行。

(11) 判断迭代次数是否达到最大值，是则中断循环，转入第 12 步执行；否则将均方误差清零，转入地 2 步执行。

(12) 保存各层的权值，判断均方误差是否达到最小值，是则显示“训练成功”；否则显示“训练失败，需要重新训练”。

整个 BP 网络训练和识别的流程图如图 5 所示。

## 4. 实验与分析

BP 网络对训练样本和测试样本的识别结果如表 2 所示。

表 2 识别结果

训练样本 (200 个)		测试样本 (100 个)		所有样本 (300 个)	
错误个数	识别率	错误个数	识别率	错误个数	识别率
0	100.00%	19	81.00%	19	93.67%

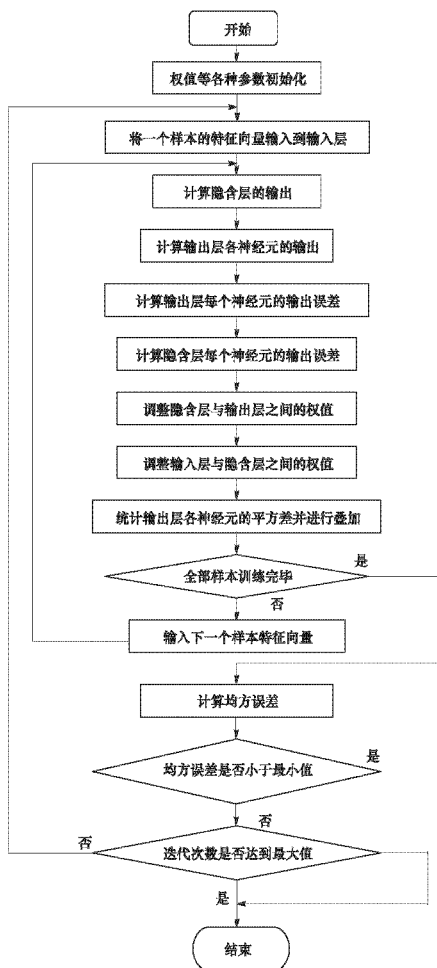


图 5 BP 网络训练流程图

本系统采用 delphi 7 编写，在 Windows XP 环境下调试通过。通过较小的样本量（训练样本 200 个；测试样本 100 个）的训练与测试，识别程序已能达到 93.7% 的识别率。因限于各种因素，只对网络进行了小样本的训练，按照理论分析，随着训练样本的增加，识别率会大幅上升。可以预见，若学习 1000 个样本以上，识别率应能达到 94%~96%。

## 5. 结束语

本文提出了一种基于 BP 神经网络来对无限制手写体数学符号识别的方法。实验结果表

明, 本系统具有运算速度快, 识别率较高的特点, 并具有良好的容错能力。如果让系统学习更多的样本, 则识别率将会进一步提高。

## 6. 致谢

本文研究工作得到国家自然科学基金项目 (60675025, 60875050)、国家 863 课题 (2006AA04Z247) 和广东省自然科学基金项目 (9151806001000025)、深圳市科技计划项目及基础研究计划项目 (JC200903160369A) 支持。

## 参 考 文 献

- [1] J.P.Marques de Sa. Pattern Recognition Concepts Methods and Applications [M]. Springer 2001:171-181.
- [2] Theodoridis S, Koutroumbas K. Pattern Recogniton [M], 2nd ed. 北京: 机械工业出版社.2003:1-33.
- [3] 边肇祺, 张学工, 等. 模式识别 [M], 2 版. 北京: 清华大学出版社.2000:23-48.
- [4] NG S C, Cheung C C, Leung S H. Fast convergence for back-propagation network with magnified gradient fuction [J]. IEEE Trans. on Neural Network 2003, 9 (3) :1903-1908.
- [5] 戴群, 陈松九, 王喆. 一个基于自组织特征映射网络的混合神经网络结构 [J]. 软件学报. 2009, 05:1329-1336.
- [6] Z Wang, S Chen, J Liu, D Zhang. Pattern representation in feature extraction and classifier design:matrix versus vector [J]. IEEE Trans. on Neural Network. 2008, 19 (5) :758-769.