

基于 BP 神经网络的手写体数字识别

张 充,史青宣,苗秀芬,杨 芳,田学东

(河北大学 数学与计算机学院,河北 保定 071002)

摘 要:手写体数字识别是多年来的研究热点,也是字符识别中的一个特别问题。由于手写体数字字体变化很大,传统的识别方法很难达到高的识别率。针对传统的数字识别方法的复杂性和局限性,提出了一种基于 BP 神经网络的手写体数字的识别方法。该方法在提取手写体数字点特征、笔划密度特征基础上,利用改进的 BP 神经网络进行训练识别。经实验,识别率达 94%。实验结果表明,该方法对手写体数字识别效果良好,不仅简化了传统识别的繁杂性,而且提高了识别的准确性。

关键词:模式识别;手写体数字;BP 算法;神经网络

中图分类号:TP391.4

文献标识码:A

文章编号:1673-629X(2008)06-0128-03

Handwritten Numeral Recognition Based on BP Neural Network

ZHANG Chong, SHI Qing-xuan, MIAO Xiu-fen, YANG Fang, TIAN Xue-dong

(College of Mathematics and Computer, Hebei University, Baoding 071002, China)

Abstract: Handwritten numeral recognition is a hotspot of study for years, and is an especial issue of character recognition. On account of great changes of handwritten font, it is very difficult for the traditional method of recognition to achieve high recognition rate. To counter the complexity and limitation of traditional digital recognition methods, a kind of handwritten numeral recognition method based on BP neural network is proposed. The point feature and stroke density feature for handwritten digits are extracted; then an improved BP neural network is applied to classify handwritten digits by those features. Via experiment, the recognition rate is 94%. Experiments show that the proposed approach has a good effect on handwritten numeral recognition. It not only simplifies the complexity of the traditional recognition, but also increases the accuracy of recognition.

Key words: pattern recognition; handwritten numeral; BP algorithm; neural network

0 引言

手写体字符,由于字体变化很大,传统的识别方法^[1]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近些年提出的新方法,为手写体字符识别研究提供了一种新手段,它具有一些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习能力。

神经网络模式识别过程如图 1 所示。预处理的目的是去处干扰、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,这是识别处理中所必

要的工作^[2]。

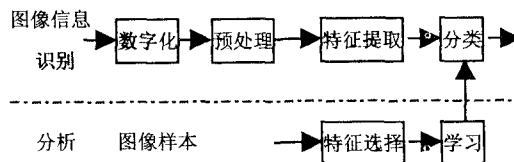


图 1 神经网络模式识别方法框图

1 预处理

字符处理前的首要工作是将视觉图像转化为可由计算机识别的二值图像。由于样本数字的灰度范围有较大差别,无法采用统一的二值化阈值。文中采用了最大方差法^[3]决定阈值。二值化后的字符可能包含孤立的黑白像素,或在应当是黑的地方出现白点,平滑化就是减少这些黑白噪声的技术。经过平滑处理后数字图像便于后继的细化操作。文中采用了区域平滑方法去掉了这些噪声。

字符的细化是为了获取字符的骨架,在细化图中利用检测点(黑色像素)的连接数可容易地求出点特

收稿日期:2007-09-06

基金项目:河北省科学技术研究与发展计划项目(06213598)

作者简介:张 充(1977-),女,硕士,研究方向为模式识别与图像处理、中文信息处理;田学东,教授,硕士生导师,研究方向为模式识别与图像处理、中文信息处理。

征,也容易获得字符的笔划描述。文中采用了综合型算法^[3]。它的优点是几乎不会产生毛刺现象,即算法具有去噪声的能力。

为了便于对任意字符的特征提取,还要对字符进行规范化处理。所谓规范化,就是把字符的尺寸变换成同一大小,字符位置(旋转、平移)纠正。规范化分为位置规范化和尺寸(大小)规范化^[4]。

2 特征提取

字符特征的选择和提取是构成识别系统的关键和基础。特征选择的主要原则一般包括:

①稳定性:即提取的特征稳定可靠,这是构成正确识别的基础;

②有效性:指所选择的特征使得字域中各个不同字符在特征空间的距离最大;

③可提取性;

④尽量减小特征维数,这对采用神经网络分类器的系统意义更大,因为它能在很大程度上降低系统的复杂程度。

由于手写体字符千差万别,因人而异,目前较为成熟的字符统计特征是很难对其统计分类的,而且稳定的、最能表现字符间差别的结构特征也不存在,所以必须提取若干特征进行组合才能达到要求。

2.1 点特征

文中选用了一组点特征,即端点数、分支点数、交叉点数。特征点是以检测点(黑色像素)的连接数定义的。字符的特征点反映了字符结构的本质,具有较高的稳定性。以二值图像的八连接为例:

连接数:

$$N = \sum_{k \in c} (\tilde{f}(x_k) - \tilde{f}(x_k) \tilde{f}(x_{k+1}) \tilde{f}(x_{k+2})) \quad (1)$$

这里, x_k 的位置如图 2 所示,

$$C = \{1, 3, 5, 7\}, \tilde{f} = 1 - f, x_9 = x_{10}$$

$m-1, n-1$ x_4	$m, n-1$ x_3	$m+1, n-1$ x_2
$m-1, n$ x_5	m, n x_0	$m+1, n$ x_1
$m-1, n+1$ x_6	$m, n+1$ x_7	$m+1, n+1$ x_8

图 2 x_k 的位置示意图

检测点的特征如表 1 所示。

2.2 区域笔划方向特征

还采用了区域笔划方向特征作为系统的识别特征,即把输入的待识别字符划分为四个区域,统计区域的水平、垂直、 $\pm 45^\circ$ 四方向的笔划密度特征。值得注意的是,文中是对细化后的字符抽取笔划密度特征。笔划

密度特征的具体获取办法就是对字符进行水平、垂直、左右、斜角分别扫描,统计穿过笔划的次数,反映了字符的四个方向的笔划分布密度,是种很有效的特征。

表 1 检测点的特征表

连接数	像素的特征
0	孤立点或内部点
1	端点或边界点
2	连接点
3	分支点
4	交叉点

现将所抽取的特征罗列如下:

端点数;分支点数;交叉点数;四区域各自的水平、垂直、 $\pm 45^\circ$ 四方向的笔划密度特征。

3 利用改进的 BP 神经网络分类与学习

3.1 BP 网络结构

由 Rumelhart 提出的多层前馈神经网络模型,由于采用了误差反传的 BP 学习算法,被称为 BP 网络^[5]。因其高度非线性映射能力,BP 网络的应用极广。一个 3 层的 BP 网络可以实现任意精确近似任何连续的非线性函数。文中采用了一个隐层 BP 网络作为训练网络,如图 3 所示。

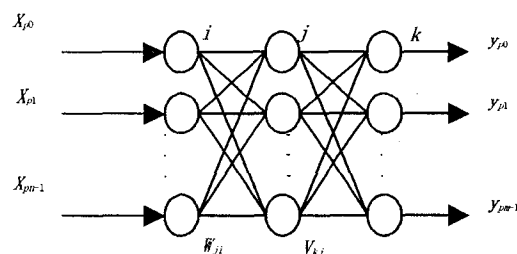


图 3 隐层 BP 网络示意图

1) 输入与输出层的设计。

在手写体数字的特征提取中,文中选取了 19 个特征组成特征向量空间,归一化后作为神经网络的输入向量(输入样本)。因此输入层结点数即为输入特征向量的维数(19)。输出层的结点数 10,分别代表 0, 1, ..., 9 十个类别。

2) 隐单元数目的选择。

对于隐层单元数的选择是个十分复杂的问题,往往根据设计的经验和试验来选择。隐单元数太少可能不能训练出来,或网络不“强壮”,不能识别以前没有看到的样本,容错性差,但隐单元数太多又使学习时间过长,误差也不一定最佳,因此存在一个最佳的隐单元数,目前下面两个公式可作参考。

$$\textcircled{1} \quad n_1 = \sqrt{n + m} + a$$

其中 m 为输出神经元, n 为输入神经元, a 为 1 ~ 10

之间的常数。

$$\textcircled{2} \quad n_1 = \log_2 n$$

n 为输入神经元数。

按上述的两个公式并考虑到对测试样本的容差性,取隐单元数为 10。

3.2 改进的 BP 算法

标准的 BP 算法的主要缺点:收敛速度慢、容易陷入局部极小。在实际应用中,标准 BP 算法很难胜任。文中采用了动量法和学习速度自适应调整的策略,并与批处理的训练方式^[6]有机地结合起来,从而提高了学习速度并增加了算法的可能性。

BP 网络输入层结点数为 n ($n = 19$),隐层结点数为 r ($r = 10$),输出层结点数为 m ($m = 10$),第 i 个输入结点与第 j 个隐结点的连接权值为 w_{ji} ,第 j 个隐结点与第 k 个输出结点的连接权值为 v_{kj} ,隐结点阈值为 θ_j ,输出层结点的阈值为 γ_k ,并设有 N 个学习样本 (X_p, Y_p) ($p = 1, 2, \dots, N$),其中 $X_p = (x_{p0}, x_{p1}, \dots, x_{pm-1})^T$ 为第 p 个学习样本的输入向量, $o_p = (o_{p0}, o_{p1}, \dots, o_{pr-1})^T$ 为隐层的输出, $Y_p = (y_{p0}, y_{p1}, \dots, y_{pm-1})^T$ 为网络的实际输出, $T_p = (t_{p0}, t_{p1}, \dots, t_{pm-1})^T$ 为第 p 个学习样本的期望输出向量。

对于输入层结点,取其输入输出相同,即

$$S_{pi} = x_{pi} \quad (i = 0, 1, \dots, n-1) \quad (2)$$

隐层第 j 个结点的输入及输出分别为:

$$\text{net}_{pj} = \sum_{i=0}^{n-1} w_{ji} S_{pi} - \theta_j = \sum_{i=0}^{n-1} w_{ji} s_{pi} \quad (3)$$

$$o_{pj} = f(\text{net}_{pj}) = 1/[1 + \exp(-\text{net}_{pj})] \quad (4)$$

式中 $j = 0, 1, \dots, r-1$; $\theta_j = w_{jn}$; $s_{pn} = -1$ 。

输出层第 k 个结点的输入输出分别为:

$$\text{net}_{pk} = \sum_{j=0}^{r-1} v_{kj} o_{pj} - \gamma_k = \sum_{j=0}^{r-1} v_{kj} o_{pj} \quad (5)$$

$$y_{pk} = f(\text{net}_{pk}) = 1/[1 + \exp(-\text{net}_{pk})] \quad (6)$$

式中 $k = 0, 1, \dots, m-1$; $\gamma_k = v_{kr}$; $o_{pr} = -1$ 。

神经网络学习过程是通过调整权值和阈值使学习样本的期望值和实际输出值的均方误差趋于最小实现的。可是神经网络学习的是许多组样本,如果一组一组的训练,会出现学新忘旧的现象,学习效率不高。所以文中采用了所谓的批处理的训练方式^[6],即在整个训练样本集都提交网络后才更新参数。

选择误差函数:

$$E = \frac{1}{2} \sum_p \sum_k (t_{pk} - y_{pk})^2 \quad (7)$$

从而,输出层单元 k 的 j 的误差信号:

$$\delta_{jk} = (t_{jk} - y_{jk}) y_{jk} (1 - y_{jk}) \quad (8)$$

隐层单元 j 的误差信号:

$$\delta_{pj} = o_{pj} (1 - o_{pj}) \sum_k v_{kj} \delta_{pk} \quad (9)$$

神经网络的权值(包括阈值)调整公式分别如下:

$$v_{kj}(d+1) = v_{kj}(d) + \eta(d) \sum_p \delta_{pk} o_{pj} + \alpha [v_{kj}(d) - v_{kj}(d-1)] \quad (10)$$

$$w_{ij}(d+1) = w_{ij}(d) + \eta(d) \sum_p \delta_{pj} x_{pi} + \alpha [w_{ij}(d) - w_{ij}(d-1)] \quad (11)$$

其中 d 是迭代次数, $\eta(d)$ 是学习速度, α 是动量因子 ($0 < \alpha < 1$)。

$\eta(d)$, α 调整规则如下:

$$\text{当 } \Delta E(d) < 0, \eta(d+1) = \eta(d) \cdot \phi \quad \alpha = \alpha$$

$$\text{当 } \Delta E(d) > 0, \eta(d+1) = \eta(d) \cdot \beta \quad \alpha = 0$$

其中 $\phi > 1, \beta < 1, \Delta E(d) = E(d) - E(d-1)$

4 实验结果

文中在实验中所用的手写体数字样本是由 30 个人随意书写,共 30 组样本。选择其中 20 组为训练样本,共 200 个训练样本(如图 4 所示);另外 10 组作为测试样本,共 100 个样本(如图 5 所示)。

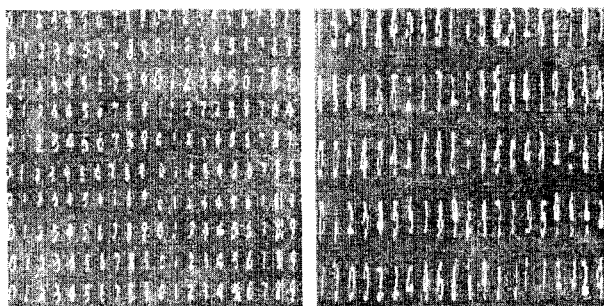


图 4 训练样本示意图 图 5 测试样本示意图

实验参数选择如下: $\alpha = 0.9, \phi = 1.05, \beta = 0.7$ 。另外,由于网络的传输函数采用 S 型压缩函数,其范围限制在 $0 \sim 1$ 内,所以将 $0 \sim 9$ 的值也压缩到 $0 \sim 0.9$ 范围内作为输出期望值,即 $0 \rightarrow 0.0, 1 \rightarrow 0.1, 2 \rightarrow 0.2, 3 \rightarrow 0.3, 4 \rightarrow 0.4, 5 \rightarrow 0.5, 6 \rightarrow 0.6, 7 \rightarrow 0.7, 8 \rightarrow 0.8, 9 \rightarrow 0.9$ 。训练迭代次数 2167 次,网络收敛。

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)	
错误数	识别率	错误数	识别率	错误数	识别率
0	100.00%	18	82.0%	18	94.0%

(下转第 163 页)

$$p_m = \begin{cases} k_3 \frac{(f_{\max} - f)}{f_{\max} - f_{\text{avg}}} & f' \geq f_{\text{avg}} \\ k_4 & f' < f_{\text{avg}} \end{cases} \quad (4)$$

在公式(3)和(4)中,一般取 $k_1 = 1.0, k_2 = 1.0, k_3 = 0.5, k_4 = 0.5$ 。

3 实验结果

假设某个批次的试验里,有8块96孔微孔板,每块微孔板要经过加样→孵育→洗板→加样→孵育→洗板→加样→孵育→加样→读板10个步骤。已知每个步骤占用每个模块处理时间矩阵 T 为:

$$T = \begin{bmatrix} 6 & 30 & 3.6 & 1.5 & 20 & 3.6 & 1.5 & 15 & 1.5 & 0.5 \\ 4.8 & 30 & 3.6 & 1.5 & 30 & 3.6 & 1.5 & 15 & 1.5 & 0.5 \\ 2.8 & 40 & 1.6 & 0.9 & 30 & 1.6 & 0.9 & 15 & 0.9 & 0.23 \\ 5.2 & 30 & 3 & 1.3 & 30 & 3 & 1.3 & 15 & 1.3 & 0.41 \\ 3.6 & 40 & 3.6 & 1.5 & 30 & 3.6 & 1.5 & 15 & 1.5 & 0.5 \\ 4.4 & 30 & 2.62 & 1.1 & 30 & 2.62 & 1.1 & 15 & 1.1 & 0.38 \\ 2 & 30 & 1.2 & 0.5 & 30 & 1.2 & 0.5 & 15 & 0.5 & 1.3 \\ 4 & 30 & 2.4 & 1.0 & 30 & 2.4 & 1.0 & 15 & 1.0 & 0.27 \end{bmatrix}$$

对于上述试验任务,用改进的自适应遗传算法和随机算法分别生成试验执行序列^[6]。在使用随机算法时,该批次试验执行完毕所需时间为225.3分钟。而在使用改进的自适应遗传算法时,取 $p_m = 0.1, p_c = 0.6, N = 80$,种群规模为20,该批次试验完成的时间是132.6分钟。效率提高了69.9%。同时,在医院的试用过程当中,血样检测的准确率也达到了99.6%以上。

4 结束语

根据实际工程需要,按软件工程的设计思想,归纳

出全自动酶免工作站计算机控制系统的对象,给出系统实现单文档多视和控制类的详细设计,并将改进的自适应遗传算法应用于系统执行效率的优化。通过模拟比较实验可见,改进后的系统执行效率显著提高,实验结果对全自动酶免工作站的设计具有指导性意义。预期将会应用于新的全自动酶免工作站的设计之中。每台将自适应遗传算法应用于系统执行效率的酶免仪可以提高将近70%的执行效率,每天可以多检测120个左右的血样,如果按照现在医院血液检测的指导价,每台酶免仪每年可为医院多创造150万人民币左右的利润。

参考文献:

- [1] 金伯泉.细胞和分子免疫学实验技术[M].西安:第四军医大学出版社,2002:61-65.
- [2] Douglass B P. Real-Time Design Patterns - Robust Scalable Architecture for Real-Time System[M].北京:北京航空航天大学出版社,2004:112-119.
- [3] 王万良,吴启迪,宋毅.求解作业车间调度问题的改进自适应遗传算法[J].系统工程理论与实践,2004,20(2):58-62.
- [4] Goncalves J F, de Magalhães Mendes J J, Resende M G C. A hybrid genetic algorithm for the job shop scheduling problem[J].European Journal of Operational Research,2005,167:77-95.
- [5] Cheng R, Gen M, Tsujimura Y. A tutorial survey of job-shop scheduling problems using genetic algorithms representation[J].Computers & Industrial Engineering,1996,30(4):983-997.
- [6] 许家玉,经亚枝.基于DSP+FPGA的遗传算法硬件实现[J].微计算机信息,2005(1):127-128.

(上接第130页)

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94%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95%~97%。

5 结束语

提出一种基于BP神经网络的手写体数字的识别方法,通过提取手写体数字的点特征和区域笔划方向特征以及通过改进的BP算法设计神经网络分类器,实现了手写体数字的识别。实验结果证明,此方法行之有效。利用基于BP神经网络的手写体数字的识别

方法,不仅简化了传统识别的繁杂性,而且提高了识别的准确性。因此人工神经网络作为模式识别的手段,具有很大的潜力和广泛的应用前景。

参考文献:

- [1] 边肇祺.模式识别[M].北京:清华大学出版社,1999.
- [2] 阮秋琦.数字图像处理学[M].北京:电子工业出版社,2001.
- [3] 谷口庆治.数字图形处理[M].北京:科学出版社,2002.
- [4] Castleman K R. Digital Image Processing[M]. [s.l.]: Prentice-Hall International, Inc, 1988.
- [5] Rumelhart D E, McClelland J L. Parallel Distributed Processing[M]. Cambridge: MIT Press, 1986.
- [6] 张立明.人工神经网络的模型及其应用[M].上海:复旦大学出版社,1993.