

基于混合特征的孟加拉手写体数字识别

刘春丽,吕淑静

LIU Chun-li, LV Shu-jing

上海工商外国语学院,上海 201300

Shanghai Industry & Commerce Foreign Language College, Shanghai 201300, China

E-mail: lclsl@163.com

LIU Chun-li, LV Shu-jing. Bangla handwritten numeral recognition based on blend features. Computer Engineering and Applications, 2007, 43(20): 214-215.

Abstract: A recognition system based on blend features is proposed for Bangla handwritten numeral. The direction features, including horizontal vector, vertical vector, right diagonal vector and left diagonal vector, which combine with density features, are extracted by Kirsch operator. Then the numerals are recognized by means of BP network. Experimental results show that the proposed method obtains 96.1% correct recognition rate.

Key words: feature extraction; Kirsch operator; Bangla handwritten numeral; BP network

摘要: 根据孟加拉手写体数字的特点,用 Kirsch 算子提取象素的水平、垂直、右对角线以及左对角线的特征矢量,并与字符图像的密度特征相结合,采用 BP 算法训练的 MLP 网络作分类器进行识别。最后,用从实际孟加拉信封图像中采集到的手写体数字作样本进行实验,达到了 96.1% 的识别率。

关键词: 特征提取; Kirsch 算子; 孟加拉手写体数字; BP 神经网络

文章编号: 1002-8331(2007)20-0214-02 **文献标识码:** A **中图分类号:** TP391.1

1 引言

几十年来, 手写体阿拉伯数字的识别一直是一个研究热点, 并且已在商业、邮政及办公自动化等方面得到了广泛应用。随着时代的发展和信息的流通, 对其它一些地域语言数字识别的需求也越来越迫切。顺应这种需求, 本文针对手写体孟加拉数字, 提出了基于方向特征和密度特征相混合的识别方法。

孟加拉语是南亚次大陆的第二通用语言, 是世界第五通用语言。图 1 是手写体孟加拉数字的一个子集, 描述了它的书写特点和多变性。



图 1 手写体孟加拉数字

字符的方向特征是指数字图像的每一线段的方向信息, 它在阿拉伯数字的识别中已经得到了成功的应用。例如, 文献[1]中 S.Ray 通过计算黑象素点的角度来描述字符的方向特征, 文

献[2]中通过 Sobel 算子来提取方向特征, 文献[3]通过 Kirsch 算子来提取方向特征等等。经过实验比较, Kirsch 算子比较稳定, 对孟加拉数字的识别有较好的识别率和较快的识别速度。因此, 本文选择用 Kirsch 算子来提取孟加拉数字的方向特征。但是, 仅选用方向特征存在一定的局限性, 抗噪声干扰能力有限。而密度特征是基于字符图象中笔划点阵分布情况考虑的, 对于噪声有较强的抗干扰能力。因此, 本文采用方向特征和密度特征相混合的方法, 并选用分类能力较强的 MLP 神经网络分类器对样本进行了分类识别, 经实验, 达到了较好的识别效果。

2 预处理

本方法的特征提取是基于 32x32 的二值化图像进行的, 其主要的预处理包括二值化、去噪、定位和大小规格化。用 OTSU 算法对字符图像进行二值化处理, OTSU 算法是一种适应性较强的二值化算法, 它通过统计分析图像的灰度分布特点来确定阈值, 从而对字符图像进行二值化处理, 用黑象素表示字符, 白象素表示背景, 然后通过统计连通的黑色象素的个数来去除字符图像中夹带的杂点噪声。再根据水平和垂直两个方向的扫描, 确定最小的包围矩形, 并将得到的最小包围矩形规格化为 32x32 的大小。

3 特征提取

由图 1 可以看出, 孟加拉数字字符之间笔画方向相异性较大, 所以它们的边缘信息可以很好的表征其类属。因此检测孟加拉数字笔画的边缘局部方向是一种好的特征提取方法。本文

以 Kirsch 表征码(Mask)提取方向特征^[4]。

Kirsch 定义了一个非线性边缘测量算子:

$$G(i,j)=\max\{1, \max_{k=0}^7 [15S_k-3T_k]\} \quad (1)$$

其中:

$$S_k=A_k+A_{k+1}+A_{k+2} \quad (2)$$

$$T_k=A_{k+3}+A_{k+4}+A_{k+5}+A_{k+6}+A_{k+7} \quad (3)$$

式中的 $G(i,j)$ 为象素 (i,j) 的梯度, A 的下标是模 8 后的余数, $A_k(k=0,1,\dots,7)$ 为象素 (i,j) 的八个相邻单元,图 2 为它们的定义。

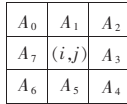


图 2 象素的 8 个相邻单元的定义

公式(4)~(7)分别描述了象素 (i,j) 的水平特征矢量 H , 垂直特征矢量 V , 右对角线特征矢量 R 和左对角线特征矢量 L :

$$G(i,j)_H=\max(|5S_0-3T_0|, |5S_4-3T_4|) \quad (4)$$

$$G(i,j)_V=\max(|5S_2-3T_2|, |5S_6-3T_6|) \quad (5)$$

$$G(i,j)_R=\max(|5S_1-3T_1|, |5S_5-3T_5|) \quad (6)$$

$$G(i,j)_L=\max(|5S_3-3T_3|, |5S_7-3T_7|) \quad (7)$$

上面的方向特征矢量用 Kirsch 表征码描述如图 3。

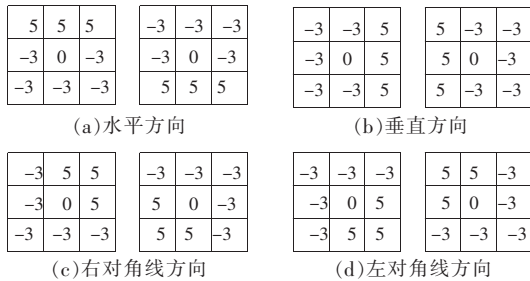


图 3 Kirsch 表征码用于抽取四个方向特征

首先将规格化的数字图像分成 4×4 个 8×8 大小的子块, 然后用上面 Kirsch 算子分别计算各个子块的每个象素的水平、垂直、右对角线和左对角线方向特征, 如果计算出的方向特征 $G(i,j)$ 大于所设定的阈值, 则所对应的方向特征为 1, 否则为 0. 统计子块中象素的方向特征的频率作为子块的方向特征, 这样就得到了四个 4×4 方向特征。然后提取数字图像的密度特征, 这样在一定程度上能增强特征对噪声的抗干扰能力同时也降低了由于规格化产生的毛刺的影响。密度特征的提取就是统计每个子块内的黑象素数量并与子块内象素的总数量相比, 即得到此子块的密度特征。由此得到五个 4×4 特征(四个 4×4 方向特征与一个 4×4 统计特征), 共 80 个特征。图 4 表示了提取特征的过程。

4 实验

本方法的分类器采用三层 MLP 网络^[5], 用经典的 BP 算法训练 MLP 网络, 输入层和输出层的神经元个数分别为 80 个(特征的维数)和 10 个(样本的类别数), 隐含层的神经元个数经过多次实验后确定为 50 个。选用 5 900 个手写体孟加拉数字样本进行实验, 这些样本都是从孟加拉信封上采集的邮政编码, 具有很强的代表性和应用性。从中随机抽取 2 000 个(每类 200 个)作为训练样本, 剩下的 3 900 个(每类 390 个)作为测试样本。表 1 描述了测试结果, 并与文献[1]和文献[2]中介绍的提取方向特征的的算法进行了比较。

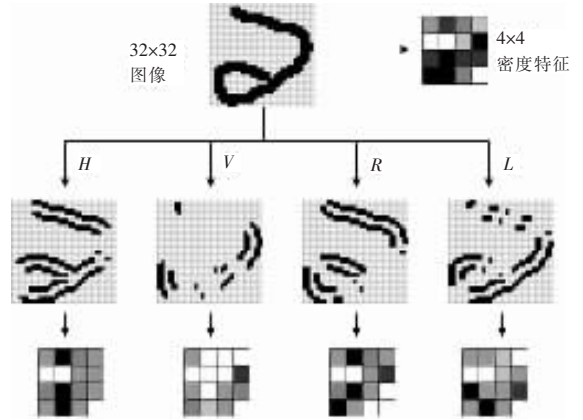


图 4 字符的一个样本抽取特征的全过程

表 1 实验结果

| 测试样本 | 文献[1]中的算法/% | 文献[2]中的算法/% | 本文算法/% | |
|------|-------------|-------------|--------|------|
| 0 | 390 | 92.8 | 93.6 | 95.6 |
| 1 | 390 | 89.9 | 90.2 | 92.8 |
| 2 | 390 | 93.1 | 92.8 | 96.2 |
| 3 | 390 | 92.8 | 94.1 | 95.6 |
| 4 | 390 | 93.6 | 94.1 | 96.4 |
| 5 | 390 | 95.4 | 96.1 | 97.2 |
| 6 | 390 | 94.6 | 94.6 | 96.2 |
| 7 | 390 | 96.2 | 96.2 | 98.5 |
| 8 | 390 | 92.3 | 95.6 | 97.9 |
| 9 | 390 | 91.3 | 93.3 | 94.1 |
| 总体 | 3 900 | 93.2 | 94.1 | 96.1 |

从表 1 可以看出对于所采集的测试样本, 本文的方法的识别率达到了 96.1%, 与文献[1]和文献[2]中的算法相比, 得到了很大提高。

5 总结

本文采用 Kirsch 算子提取孟加拉数字图像的方向特征, 并与密度特征相混合, 以三层的 MLP 神经网络作分类器, 对孟加拉数字进行识别研究。以实际采集的信封图像上的邮政编码作样本库, 实验结果证明, 识别率达到了 96.1%, 识别效果较好, 如果与其它算法组合成多专家识别系统, 可望从整体上进一步提高识别率, 对孟加拉的信函分拣系统有很大的应用价值。(收稿日期: 2006 年 11 月)

参考文献:

- [1] Ray S.A heuristic noise reduction algorithm applied to handwritten numeric characters[J].Pattern Recognition Letters, 1988, 7:9-12.
- [2] Keyzers D, Gollan C, Ney H. Local context in non-linear deformation models for handwritten character recognition[J]. ICPR, 2004 (4): 511-514.
- [3] Cho Sung-Bae. Neural network classifiers for recognizing totally unconstrained handwritten numerals[J]. IEEE Trans on Neural Networks, 1997, 8(1): 43-53.
- [4] Kim Y J, Lee S W. Off-line recognition of unconstrained handwritten digits using multilayer backpropagation neural network combined with genetic algorithm (in Korean)[C]/Proc 6th Wkshp Image Processing Understanding, 1994: 186-193.
- [5] 边肇祺, 张学工. 模式识别[M]. 2 版. 北京: 清华大学出版社, 250-271.