

文章编号:1001-9081(2006)03-0622-02

## 基于数学形态学的脱机手写体汉字识别方法

吴媛, 杨扬, 颀斌, 王宏

(北京科技大学信息工程学院, 北京 100083)

(s\_tart@163.com)

**摘要:**笔画特征是一种有效的脱机手写体汉字的识别特征,但是笔画细化往往会造成字体的变形,提出一种无需细化预处理的笔画特征提取方法,通过数学形态学中的腐蚀、膨胀等运算,采用不同的、具有自适应性的结构元素对汉字图像进行笔画分解,并利用弹性网格提取其方向特征,最后使用广义 K-L 变换对特征向量的维数进行压缩,去除冗余信息。实验结果验证了本方法的有效性。

**关键词:**数学形态学;脱机手写体汉字识别;特征提取;细化预处理;广义 K-L 变换

**中图分类号:** TP391.41 **文献标识码:** A

### Off-line recognition method for handwritten Chinese character based on mathematical morphology

WU Yuan, YANG Yang, XIE Bin, WANG Hong

(School of Information Engineering, University of Science and Technology Beijing, Beijing 100083, China)

**Abstract:** The stroke feature is efficient for off-line handwritten Chinese character recognition. But the process for thinning strokes always makes the characters distorted. A stroke feature extraction approach that doesn't need the thinning preprocess was proposed. By erosion and dilation operations of the mathematical morphology and using different adaptive structure elements, the characters were decomposed into strokes. Then the elastic meshes were applied to extract the directional feature. Finally, generalized K-L transform was used to decrease the dimensions of the feature vectors so as to get rid of redundancies. Experiment results show the effectiveness of the proposed approach.

**Key words:** mathematical morphology; off-line handwritten Chinese character recognition; feature extraction; thinning preprocess; generalized K-L transform

## 0 引言

特征提取是汉字识别过程中最重要的环节,快速地提取稳定、可靠并能充分描述汉字模式的特征,是手写体汉字识别的关键。

每个汉字区别于其他汉字的主要特征是笔画及其所在的位置,几乎所有汉字都可以看作是由“横、竖、撇、捺”四种基本笔画构成,四种笔画的数量及其相对位置不同,表示的汉字也不同。<sup>[1]</sup>因此,笔画的方向分解特征可以作为一种有效的识别特征。但现有的方法大多是基于细化预处理的笔画抽取,这种方法由于会在细化过程中产生伪笔画而无法达到理想的识别效果。因此,本文采用了一种新的无需细化预处理的方法,利用数学形态学的运算对手写体汉字图像进行方向分解,提取其笔画特征。

数学形态学是分析几何形状和结构的数学方法,是建立在集合代数基础上,用集合论方法定量描述集合结构的科学。它具有完备的数学理论基础,在运算速度方面比其他特征提取算法有更大的优势。

## 1 特征提取

### 1.1 数学形态学基本算法

形态学应用于图像处理的基本思想是利用结构元素收集图像的信息,采用不同的结构元素对同一幅图像进行同一种形态学运算,其结果是不一样的。当不同形状、大小的结构元素在图像中不断移动时,便可提取图像的各种结构特征。

腐蚀是数学形态学最基本的运算,它的作用是消除物体边界点。设  $X$  是待处理图像,  $S$  是结构元素。  $S$  在图像  $X$  上运动,在  $x$  点的状态记为  $S[x]$ 。则  $S$  对  $X$  的腐蚀运算记为:

$$X \ominus S = \{x \mid S[x] \subseteq X\} \quad (1)$$

如果结构元素取  $3 \times 3$  的黑点块,腐蚀将使物体的边界沿周边减少一个像素,如图 1(b) 所示。



(a)原始图像 (b) $3 \times 3$ 结构元素腐蚀结果 (c) $3 \times 3$ 结构元素膨胀结果  
图 1 腐蚀和膨胀运算结果

数学形态学的第二个基本运算是膨胀。膨胀是腐蚀的对偶运算,它是将与物体接触的所有背景点合并到该物体中,使边界向外扩张的过程。集合  $X$  被结构元素  $S$  膨胀,表示为:

$$X \oplus S = \{x \mid S[x] \cap x \neq \emptyset\} \quad (2)$$

膨胀运算相当于在原有  $X$  图像基础上向右方和上方各扩充一个单位,如图 1(c) 所示。在腐蚀和膨胀的基础上,可以

收稿日期:2005-09-15 修订日期:2005-12-02

作者简介:吴媛(1980-),女,辽宁沈阳人,硕士研究生,主要研究方向:模式识别、图像处理; 杨扬(1955-),男,河北承德人,教授,博士生导师,主要研究方向:图像处理、多媒体应用; 颀斌(1970-),男,陕西岐山人,讲师,硕士,主要研究方向:图像处理; 王宏(1973-),男,湖南衡阳人,讲师,博士研究生,主要研究方向:图像处理。

构造出形态学的其他各种组合运算,其中最重要的就是开运算和闭运算,分别定义为(3)式和(4)式。

$$X \circ S = (X \ominus S) \oplus S \quad (3)$$

$$X \cdot S = (X \oplus S) \ominus S \quad (4)$$

利用开运算可以消除图像中的孤立点和毛刺,即对图像进行平滑去噪;闭运算则用于填平小沟,选择适当的结构元素可以将邻近的两个部分连接起来。开运算与闭运算的一个重要性质是幂等性<sup>[2]</sup>,即:

$$(X \circ S) \circ S = X \circ S \quad (5)$$

$$(X \cdot S) \cdot S = X \cdot S \quad (6)$$

这意味着一次滤波就能把所有特定结构元素的噪声滤除干净,做重复运算不会再有效果。这与以往的中值滤波、线性卷积等经典方法有极大的不同。

### 1.2 数学形态学提取笔画方向特征

结构元素的选择在形态学运算中非常关键。通过选择不同的结构元素,对图像做侧重于不同方向的腐蚀运算,就可以得到不同方向的分解图像,再通过弹性网格对每个分解图像进行计算就可以提取汉字图像的局部特征。现有的方向分解方法大多是基于细化处理的笔画抽取<sup>[3]</sup>,这种方法相对简单,处理速度也很快,但是细化处理常常因为将交叉点(“十”形)分成两个分叉点(“T”形)而产生伪笔画或字体变形,而最终导致笔画抽取的可信度降低,如图 2 所示。

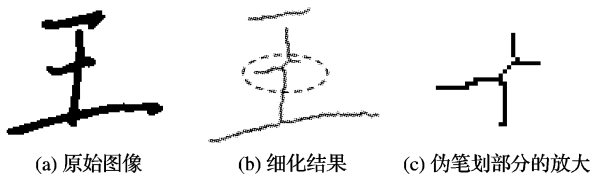


图 2 手写体汉字图像细化产生伪笔画示例

一些研究人员提出了无需细化处理的笔画抽取方法<sup>[4]</sup>,但是固定的结构元素在处理笔画宽度不同的字符图像时,会对某些方向的笔画分量不敏感而造成笔画抽取失败。

本文提出一种基于数学形态学的笔画特征提取方法,该方法不需进行细化预处理,通过自适应的结构元素对笔画宽度不同的字符图像提取方向特征:

1) 计算汉字笔画的平均宽度和分布密度,根据笔画密度在原始图像上划分弹性网格。

2) 分别用四种结构元素对汉字图像进行腐蚀运算,图 3 给出了 3×3 大小的结构元素的示意图,其中(a)~(d)分别用来提取横、竖、撇、捺四个方向的基本笔画。但在处理实际字符图像时,结构元素的大小要由笔画的平均宽度决定。这是因为,如果结构元素太小,会对笔画方向不敏感;如果结构元素过大,又会腐蚀掉太多的字符信息,所以结构元素要比笔画宽度大两个像素。

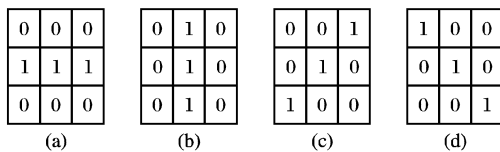


图 3 提取笔画方向分解特征的结构元素

3) 采用与 2) 腐蚀运算相同的结构元素,分别对所得到的含有四种基本笔画的子图像做开运算,滤除掉各图像中的孤立点和毛刺。

4) 为了补偿因腐蚀而造成的笔画信息丢失,分别对四个子图像进行膨胀运算,根据笔画方向决定膨胀的结构元素,每

一种笔画可选用多种结构元素连续膨胀,膨胀后的笔画宽度需大于原汉字图像的笔画宽度。

5) 将膨胀后的图像与原图进行交运算,滤除因膨胀而产生的多余部分,得到的就是该汉字图像的四个基本笔画的子图像。

6) 将 1) 生成的弹性网格分别作用于四幅笔画分解图像,假设得到的网格为  $\Omega_1, \Omega_2, \dots, \Omega_n$ ,得到的分解后“横”方向的汉字子图像为  $K_H(x, y)$ ,则“横”方向上,第  $i$  个网格内的特性可以由下式给出:

$$m_H^i = \frac{\iint_{\Omega_i} K_H(x, y) dx dy}{\iint f(x, y) dx dy} \quad (7)$$

式中:  $i = 1, 2, \dots, n$ ,  $n$  为网格数目。其他几个方向的特征同理可以定义,这样对于每个汉字图像就得到了一个  $n \times 4$  维特征向量。

由于这样得到的特征维数比较高,会使分类器的训练更复杂,计算量更大。因此我们采用广义 K-L 变换<sup>[5]</sup>来降低特征的维数。根据 K-L 变换理论,特征值越大,样本在其所对应的特征向量上的投影就越真实地反映了该样本的主要成分,据此可以实现对特征的降维。

## 2 实验结果

为了检验本文方法的有效性,在国标 GB 2312-80 的 16 区到 55 区中选择 120 类常用汉字进行识别实验,每类汉字采集 200 个样本,其中 150 个做特征提取,另外 50 个做测试。这些汉字以 400dpi 的分辨率被扫描成二值图像,并经过了切分去噪声等预处理。分类器的好坏对整个系统的识别性能有较大的影响,而本实验的目的只为对本文所提出的无需细化处理的手写体汉字方向分解和特征提取方法进行验证,因此采用了比较简单的欧氏距离分类器。

首先根据汉字图像在水平和垂直两个方向上的直方图投影对其进行  $8 \times 8$  的纵横弹性网格划分,如图 4 所示。

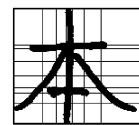


图 4 对汉字图像进行  $8 \times 8$  弹性网格划分

然后利用本文提出的基于数学形态学的方法进行笔画分解得到四幅子图像,如图 5 所示,由于数学形态学的运算速度非常快,笔画抽取的平均速度仅为 35.65ms。将弹性网格分别作用于 4 幅子图像,利用公式(7)就可以对每个汉字得到  $8 \times 8 \times 4 = 256$  维的特征向量。

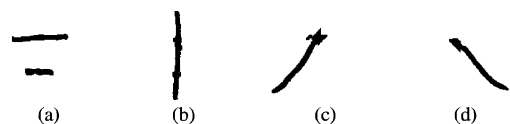


图 5 笔画分解得到的 4 幅子图像

采用 K-L 变换对原始特征进行降维处理后,可将特征向量从原来的 256 维特征空间变换到 72 维空间,这样运算速度会加快很多。

最后根据这 72 维特征向量采用欧氏距离法对 4800 个测试样本进行识别,同时与基于细化预处理的笔画方向特征提取方法做比较,得到实验结果,如表 1 所示。

12.25% 为 ROI 的左上角坐标为 (200,200)、大小为  $128 \times 128$  的区域,  $\eta = 6.125\%$  为 ROI 的左上角坐标为 (240,240)、大小为  $64 \times 64$  的区域。

从表 2 中可以看出, ROI 区域占整幅图像的百分比越高, ROI 区域的信噪比越高。这是因为 ROI 区域占整幅图像的百分比越高, 背景区域损失的高频信号就越少。当  $\eta = 25\%$  时, 随着  $R$  的增加, ROI 区域的图像得到比较好的恢复。

表 3 为采用 Scaling-based shift 方法和本算法 PSNR 之间的对比, 其中 Scaling-based shift 算法采用背景右移 3 位, 本算法采用  $n = 4$ 。

表 2 本算法中 ROI 取不同大小时的 PSNR 值

$\eta$ (%)	区域	$R$ /bpp					
		0.1	0.2	0.5	1.0	1.5	2.0
56.250	ROI	16.22	18.16	22.46	27.42	30.15	31.76
	BG	12.02	17.15	22.95	22.73	22.60	22.52
25.000	ROI	15.37	17.60	22.59	28.58	31.22	31.69
	BG	13.48	17.72	22.51	22.40	22.24	22.19
12.250	ROI	12.82	16.74	22.66	29.25	29.61	29.61
	BG	13.96	17.76	22.46	22.23	22.17	22.17
6.125	ROI	12.82	16.74	22.66	29.25	29.61	29.61
	BG	14.17	17.97	22.67	22.44	22.39	22.38

表 3 本算法和 Scaling-based shift 方法之间的 PSNR 的对比

算法	区域	$R$ /bpp					
		0.1	0.2	0.5	1.0	1.5	2.0
本算法 $n = 4$	ROI	15.37	17.60	22.59	28.58	31.22	31.69
	BG	13.48	17.72	22.51	22.40	22.24	22.19
Scaling-based shift 方法, 背景右移 3 位	ROI	15.13	17.15	21.70	28.55	32.57	35.81
	BG	10.44	10.44	17.73	21.90	24.40	26.82

从表 3 可以看出, 在  $R$  较低时, 两种算法的 ROI 区域 PSNR 值大小相差不大, 背景区域的 PSNR 本算法大于 Scaling-based shift 算法。这是因为在  $R$  比较小时, 本算法中, 高比特面不区分 ROI 区域和背景区域, 能够首先恢复出 ROI 区域和背景区域的大致轮廓; 在 Scaling-based shift 算法中, 背景区域右移 3 位, 首先对 ROI 区域编码, 然后对背景区域编



图3 Lena原图



图4 本算法在 0.2bpp 时的恢复图像



图5 Scaling-based shift方法在 0.2bpp 时的恢复图像

码, ROI 区域被先恢复出来, 而背景区域比较模糊。当然, 在高  $R$  时, 由于本算法对  $n > 4$  的背景区域进行清零操作, 对背

景区域的信号有很大的损伤, 在 Scaling-based shift 算法中, 背景区域右移 3 位, 对背景区域的信号损伤不大, 所以背景区域的 PSNR 本算法小于 Scaling-based shift 算法。

图 3 为 lena  $512 \times 512$  原图。图 4 为采用本算法  $n = 4$ ,  $\eta = 25\%$ ,  $R = 0.2$ bpp 时的恢复图像, 图 5 为采用 Scaling-based shift 方法背景右移 3 位、 $R = 0.2$ bpp 时的恢复图像。从图中可以看出, 用两种方法实现, ROI 区域的视觉效果差不多, 但是本算法的背景区域的效果好于 Scaling-based shift 方法。

#### 4 结语

在本篇论文中, 应用 SPIHT 编码算法对 ROI 编码进行改进, 此算法有如下优点: 1) ROI 区域不需要移位处理, 从而减少了位平面数; 2) 编码时, 在前几个比特面不分 ROI 区域与非 ROI 区域, 整个图像的基本信息能够优先被恢复出来。在解码端, 依靠原有的 SPIHT 解码程序就可以将其解码出来。可用于现有的 SPIHT 算法应用中, 并且能够很好的恢复感兴趣的图像。

#### 参考文献:

- [1] SAID A, PEARLMAN WA. A New Fast and Efficient Image Code Based on Set Partitioning in Hierarchical Trees[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1996, 6(3): 243 - 250.
- [2] SKODRAS A, CHRISTOPOULOS C, EBRAHIMI T. The JPEG2000 Still Image Compression Standard[J]. IEEE Signal Processing Magazine, 2001, 18(5): 36 - 58.
- [3] TAUBMAN D, MARCELLIN MW. JPEG2000 图像压缩基础、标准 and 实践[M]. 魏江力, 柏正尧, 译. 北京: 电子工业出版社, 2004.
- [4] 郭武, 梅丽, 罗建书. 一种基于 SPIHT 的 ROI 图像编码[J]. 中国空间科学技术, 2003, 2(1): 61 - 65.

(上接第 623 页)

表 1 不同笔画方向特征提取识别结果比较

处理方法	识别总数	正确识别数	识别率(%)
采用自适应结构元素, 无需细化预处理	4 800	4 377	91.19
采用固定大小结构元素, 无需细化预处理	4 000	3 384	84.60
基于细化预处理	3 755	3 210	85.48

#### 3 结语

从实验结果可以看出, 采用自适应结构元素, 无细化预处理的笔画特征提取方法不但能够迅速正确地提取方向特征, 也由于能够避免伪笔画的产生, 使得其识别率更高。笔画特征属于局部特征, 而多年来的实践研究表明, 采用单一的特征

提取方法不可避免会存在一些识别的“死角”。因此, 下一步可以考虑将本特征与全局特征进行融合, 并结合效果更好的分类器, 以提高识别率。

#### 参考文献:

- [1] 高彦宇. 脱机手写体汉字识别关键环节的研究[D]. 北京科技大学, 2004.
- [2] CASTLEMAN KR. Digital Image Processing(影印版)[M]. 北京: 清华大学出版社; PRENTICE HALL, 1998. 83 - 115.
- [3] LIU C - L, KIM R - J, KIM JH. Model - based stroke extraction and matching for handwritten Chinese character recognition[J]. Pattern Recognition, 2001, 34(12): 2339 - 2352.
- [4] 高彦宇, 杨扬. 基于正交特征的手写体汉字识别方法[J]. 仪器仪表学报, 2003, (4): 446 - 447.
- [5] YANG J, YANG J - Y. Generalized K-L transform based combined feature extraction[J]. Pattern Recognition, 2002, 35(1): 295 - 297.