几种手写体汉字网格方向特征提取法的比较研究*

金连文，高 学
(华南理工大学 电子信息学院，广东 广州 510641)

摘 要：近年来，网格方向特征已广泛应用于许多手写体汉字识别系统中，并认为是目前较成熟的手写体汉字特征之一。网格技术和方向分解是网格方向特征的两个关键技术。方向特征提取方法有许多种并各有优劣。对几种方向特征提取方法进行了比较研究，对其中一些方法进行了改进，并将我们提出的局部网格的划分方法应用到这几种方向分解特征的提取上，取得了较好的识别效果。

关键词：特征提取；手写体汉字识别；局部弹性网格；方向特征
中图法分类号：TP391 **文献标识码**：A **文章编号**：1001-3695(2004)11-0038-03

Study of Several Handwritten Chinese Character Directional Feature Extraction Approaches

JIN Lian-wen, GAO Xue
(College of Electronic Information, South China University of Technology, Guangzhou Guangdong 510641, China)

Abstract: Recently, it is found that the directional feature is considered suitable for Chinese character recognition, and directional feature has been widely used as one of the mainstream feature extraction approach. Directional decomposition algorithm and meshing method are two key factors in extraction of directional feature. This paper presents several directional features with different meshing methods for Handwritten Chinese Character Recognition (HCCR). Local elastic meshing technology is introduced in this paper, and it is found that local elastic meshing method is much better than global elastic meshing method.

Key words: Feature Extraction; Handwritten Chinese Character Recognition (HCCR); Local Elastic Meshing; Directional Feature

1 引言

特征提取是一个手写体汉字识别系统最为关键的环节之一。良好的特征必须能反映汉字的本质特征，能容忍手写体各种书写风格的变形和随意性，同时还应简洁并易于硬件实现。自 20 世纪 80 年代以来，特征提取一直是手写体识别中的一个研究重点[1,2]，已经提出许多特征提取方法。近年来，大量的研究实验发现，方向特征是一种较好的手写体汉字特征，有许多方向特征并已成功应用于许多手写体汉字识别系统中[3~11]，成为手写体汉字识别的主流特征提取方法[5]。一般而言，方向特征提取方法可用图 1 所示的流程图来表示。

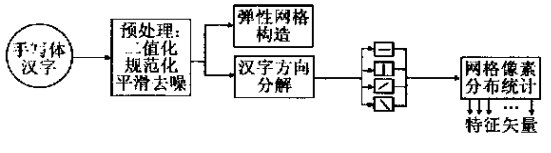


图 1 弹性网格方向分解特征的提取框图

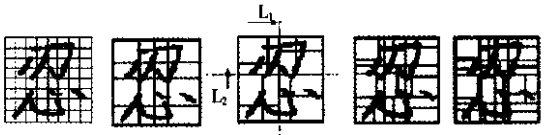
从图 1 中我们看到，手写体汉字图像经过预处理(去噪声、归一化、细化、轮廓提取等)，然后按照一定的规则构造网格，同时，经过预处理的汉字按一定的算法分解为横、竖、撇、捺四个方向，然后将网格应用到四个方向的子分量图像上，统计每个小网格内黑像素点的分布作为该汉字的统计特征。由于不同的网格构造方法及不同的方向分解方法，可以得到不同的方向特征(如轮廓方向线素特征[4,8]、模糊方向线素特征[7]、骨架方向特征[4,9,10]、边缘方向特征[9]等)，我们统称其为网格方向特征(Meshing Directional Feature)。不难看到，提取网格方向特征的两个关键技术是：网格的构造及方向分解方法。

收稿日期：2003-12-14
基金项目：国家自然科学基金资助项目(60275005)；广东省自然科学基金资助项目(011611)；Motorola 国际合作基金项目

2 弹性网格构造技术[5]

网格技术是提取网格方向特征的关键技术之一[5,14]。网格是一组假想的网线对汉字图像的区域划分，如图 2(a)所示。图中水平和垂直方向分别用八条网线对汉字进行划分，从而将该汉字图像分为 8×8 = 64 个小区域，每一个区域称为一个网格。由于图中网线是在垂直方向和水平方向均匀分布的，这样所得到的网格我们称之为均匀网格。根据汉字图像的笔画分布用非均匀的网线划分汉字得到的网格，就是非均匀网格，部分文献中称之为动态网格，我们又称之为弹性网格。通常，非均匀网线是根据汉字图像在水平、垂直两个方向上的直方图投影来确定的，对直方图的均匀等分实际上就是对汉字图像的非均匀等分，如图 2(b)所示。对弹性网格而言，一般是从汉字整体上来考虑而确定网格，我们统称其为全局网格。如果先对一个汉字图像构造全局网格将汉字划分为子图像 $\Omega_1, \Omega_2, \cdots, \Omega_n$，然后再对每个子图像 Ω_i 进行一次弹性网格划分，这样经过两次划分得到的网格称为局部弹性网格，如图 2(c)所示。L1 及 L2 两条线先将汉字非均匀分为四个区域，然后在每个小区域

中进行第二次非均匀划分,这样最终得到 16 个局部弹性网格。



(a) 8×8 固定网格　(b) 10×10 全局弹性网格　(c) 2×2 局部弹性网格　(d) 3×2 局部弹性网　(e) 4×2 局部弹性网格

图 2　网格

对手写体汉字进行弹性网格的划分实际上是对汉字图像进行的一种非线性变换,不同书写风格的汉字根据其汉字笔画分布所进行的变换是不同的,而这种非线性变换试图将同类型的不同书写风格的汉字图像变换到相同概率分布的网格上,从而能有效地将手写体汉字的共同特征反映出来。全局网格能反映汉字整体的结构信息,而经过两次划分得到的局部网格则同时能更精细地反映汉字的局部结构信息。此外,使用弹性网格的一个好处是不必对手写体汉字进行非线性规一化处理[14],从而避免了因非线性规一化处理引起的字形失真。

3　五种方向分解算法

3.1　骨架方向分解

手写体汉字首先经过骨架提取(细化),设 p 是细化后汉字图像中的一黑像素点,其八邻域如图 3 所示,则基本的骨架方向分解算法可描述如下:

如果 p_1 或 p_5 为黑像素点,则 p 属于横分量;

如果 p_2 或 p_6 为黑像素点,则 p 属于撇分量;

如果 p_3 或 p_7 为黑像素点,则 p 属于竖分量;

如果 p_4 或 p_8 为黑像素点,则 p 属于捺分量。

3.2　轮廓方向分解

轮廓方向分解与骨架方向分解类似,所不同的是分解是在汉字图像的轮廓上进行(首先要进行汉字图像的轮廓提取),而非在骨架上进行。根据轮廓进行分解得到的网格方向特征又有学者称之为方向线数特征。

以上两种分解方法是目前文献上使用得较多的方向特征分解方法。在此基础上有一些改进措施,例如使用重叠网格、模糊网格[7]等。此外,在分解方法上,一种主要的改进措施是采用加权技术。例如对横方向的分解,加权的方法是:如果 p_1 和 p_5 均为黑像素点,则 p 以权系数 1.0 属于横方向;如果 p_1 和 p_5 自有其中一个点是黑像素点,则 p 以权系数 0.5 属于横方向。其他方向的加权规则同理。

3.3　边缘方向分解

此外,可以用图像处理中边缘检测算子的办法来进行汉字的方向分解。在文献[9]中,我们使用如图 4 所示的四个方向算子作用于汉字图像上,检测汉字二值图像边缘点四个方向上梯度急剧变化的点,可以大体上将汉字图像四个方向的分量提取出来。

图 3　像素点 p 的八邻域

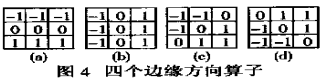


图 4　四个边缘方向算子

3.4　改进的轮廓方向角分解

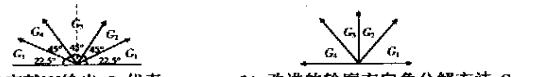
手写体汉字经过轮廓提取后,对每一个黑像素 p,定义如图 3 所示一个 3×3 的窗口,如果 p 是字符的轮廓点,那么按式(1)来计算该轮廓点的方向角 $\theta(p)$:

$$\theta(p) = \tan^{-1}\left(\frac{D_x}{D_y}\right) \tag{1}$$

式(1)中 D_x, D_y 是 p 点在 x 轴和 y 轴上的梯度函数,根据 Sobel 算子,D_x, D_y 定义为:

$$\begin{aligned} D_x &= (p_6 + 2p_7 + p_8) - (p_1 + 2p_2 + p_3) \\ D_y &= (p_3 + 2p_5 + p_8) - (p_1 + 2p_4 + p_6) \end{aligned} \tag{2}$$

方向角的取值范围为 0°～180°,按图 5(a)将方向角度分为 G_1, G_2, G_3, G_4 四类,分别对应汉字的横、撇、竖、捺四个方向[3]。经过实验,我们对原始的方法进行了适当改进,按照图 5(b)来进行分解。实验结果表明,这更加有效(对 1 034 类汉字实验识别率提高 1.02 %)。



(a) 文献[3]给出 G_1 代表横方向,G_2 代表撇方向,G_3 代表竖方向,G_4 代表捺方向　(b) 改进的轮廓方向角分解方法 G_1 同时代表横撇方向,G_2 同时代表撇竖主向,G_3 同时代表竖捺方向,G_4 同时代表横捺方向

图 5　两种轮廓方向角分解方法

3.5　笔画方向分解算法

设 $DN^l(l = 1,2,3,4)$ 表示二值图像中某黑像素点的四方向线数长度,l 取值为 1,2,3,4,分别代表横、竖、撇、捺四个方向。对黑像素点 (i, j), DN^l 定义为 l_{th}方向与该点相邻的轮廓点间的距离,如图 6 所示。此外,我们定义某像素点 (m, n) 的 α- 邻域集为:

$$L_\alpha(m,n) = \{(i,j) \mid \max[abs(i-m), abs(j-n)] \le \alpha\} \tag{3}$$

这里 α 为一常数。

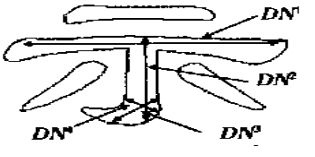


图 6　四方向线数长 DN^l 的定义

对汉字图像中的某点 (m, n),其方向属性数 $DN^l_\alpha(m, n)$ 定义为:

$$DN^l_\alpha(m,n) = \max_{i,j}\{DN^l(i,j), (i,j) \in L_\alpha(m,n)\} \quad l = 1,2,3,4 \tag{4}$$

基于如上定义,给出一种新的汉字四方向分解算法为:

对黑像素点 (i, j):

如果 $DN^k_\alpha(i,j) = \max_l\{DN^l_\alpha(i,j)\}, l = 1,2,3,4$ 或 $DN^k_\alpha(i,j) > W$,并且 $DN^k_\alpha(i,j) \le \max_l\{DN^l_\alpha(i,j)\}$,则该点被分解到第 k_{th}方向。

这里 W 是一个反映汉字笔画宽度的参数。

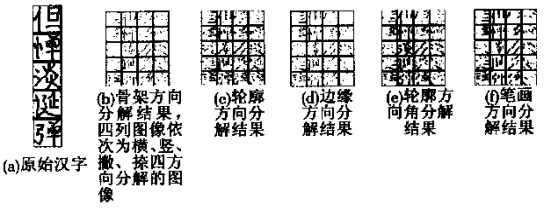


图 7　五种不同方向分解结果

由于这种分解方法直接在汉字笔画上进行分解,我们称之为笔画方向分解算法[13]。图 7 给出了使用如上五种不同分解

算法对汉字进行分解的结果。

4 网格方向特征提取

当汉字样本按一定的算法进行四方向分解之后,我们将第 2 节所述的虚拟网格作用于每个子方向的汉字图像上,统计各个小网格内黑像素点的统计分布,即得到该汉字的特征。由于网格数目的不同,可以得到不同维数的特征矢量。根据前述五种分解算法,我们可以得到五种不同类型的特征:

(1) 骨架方向特征(Thinning Directional Feature,TDF);

(2) 轮廓方向特征(Contour Directional Feature,CDF);

(3) 边缘方向特征(EDGE Directional Feature,EDF);

(4) 轮廓方向角特征(Contour Directional Angle Feature,CDAF);

(5) 笔画方向特征(Stroke Directional Feature,SDF)。

由上所述我们可以看到:骨架方向特征需要对汉字进行细化处理,对笔画的粗细不敏感,但如果应用到联机手写体汉字识别中,则可直接进行方向特征提取而无需进行细化运算(因为联机汉字无笔画宽度);轮廓方向特征在国内文献中报道较多,使用也比较广泛,但需要求汉字的轮廓,同时对汉字的笔画粗细敏感;边缘方向特征及笔画方向特征均可直接在原始汉字上进行方向提取,无需进行细化或轮廓提取;而轮廓方向角虽然需要提取轮廓信息,但由于是根据轮廓点的方向角度进行特征提取,因此对笔画粗细没有标准的轮廓特性那样敏感。此外,TDF,CDF,EDF 算法均比较简单,易于实现,特别是 EDF,无需细化和提取轮廓,直接在原始汉字图像上进行特征提取,最容易进行硬件实现。在实际系统中,这五种特征是具有一定互补性的,因此可根据不同的场合选择不同的特征,也可利用多分类器集成的方法将它们集成在一起从而提供整体系统的识别性能。

5 分类器设计

距离分类器作为一种简单有效的分类方法,在实际应用中受到广泛重视[1~3]。距离测度是距离分类器的关键,根据问题的不同可以有多种选择,常用的一些距离测度有欧氏距离、城市块距离、加权距离、马氏(Mahalanobis)距离等[11]。考虑到特征矢量的方差对分类的影响,文献[11]给出了一种加权距离(误差均衡距离)的定义并分析了其识别性能,我们在其基础上,进行了适当改进,定义两特征矢量 X 和 Y 的距离函数为:

$$d_j(X,Y)=\sum_{i=1}^{N}[w_i(x_i-y_i)^2+\varepsilon w_i^2] \tag{5}$$

式(5)中,$w_i=\dfrac{N}{(s_i+\alpha)\sum_{k=1}^{N}\dfrac{1}{s_i+\alpha}}$ (6)

$s=(s_1,s_2,\cdots,s_N)$ 是某类字的标准方差,N 为特征矢量的维数,ε 及 α 是两个小常数。不难看到,加权系数 w_i 具有如下性质:

$$\sum_{i=1}^{N}w_i=N \tag{7}$$

我们称由式(5)定义的距离为改进的误差均衡距离。实验表明,该距离测度比原始误差均衡距离识别率高出约0.8%,并且要明显优于欧氏距离(识别率高出 2.2%)、城市块距离(识别率高出 1.5%)等几种常用的距离测度。

6 实验

6.1 实验数据

我们使用的实验数据为国家"863"标准手写体汉字样本数据库——HCL2000[12]。HCL2000 包含国标一级字库共 1 000 套书写样本。在实验中,我们随机选择 80 套样本,每套样本用 16 区~26 区共 1 034 类字作为训练样本,另外选取 20 套样本作为测试数据。

6.2 局部弹性网格与全局弹性网格的性能比较

使用不同的弹性网格对五种特征提取方法进行的实验结果如表 1 所示。从表 1 中我们可以看到,当网格数较低时(64 维特征矢量),全局网格性能略优于局部网格,但此时识别率均不高。随着网格数的增加,识别率也在增加,但局部弹性网格明显比全局弹性网格识别率高,该结论对 TDF,CDF,EDF,CDAFD 等四种特征提取方法都是正确的,而对 SDF,局部网格与全局网格识别率相近。此外,我们还看到,识别率并非能随网格数目的增加(特征维数也相应增加)而无限制地增加,当特征维数达到 256 维时,识别率已基本饱和,400 维的特征对识别率并无多大提高。此外,我们看到这五种特征基本上都能达到比较不错的识别性能,根据第 4 节的分析,各种特征各有所长。识别率的高低并非是衡量特征好坏的唯一标准,在实际应用中,可根据算法复杂度、识别性能、硬件实现难易程度等因素选择合适的特征来构造分类系统。

表 1 不同网格及不同方向特征的实验性能对比

网格	特征维数(网格数)	识别率(%)				
		TDF	CDF	EDF	CDAF	SDF
局部 2×2	64	84.38	78.10	84.42	87.49	76.27
局部 3×2	144	91.69	89.28	92.56	93.65	88.77
局部 4×2	256	92.92	91.26	93.82	94.89	92.24
局部 5×2	400	92.51	91.05	93.80	94.89	93.05
全局 4×4	64	84.09	78.55	85.07	88.23	78.45
全局 6×6	144	91.56	88.71	91.68	93.52	89.70
全局 8×8	256	92.71	90.82	92.71	93.82	92.35
全局 10×10	400	92.36	90.60	92.16	93.36	92.76

7 结论

本文对目前手写体汉字识别中的网格方向特征提取方法进行了研究,对一些特征提取方法和分类方法进行了改进,分析比较了五种方向特征的性能,并将局部弹性网格应用到五种方向特种上,取得了较好的识别效果,表明局部弹性网格优于全局弹性网格。本文工作重点在比较五种方向特征的性能,并未使用到较先进的分类方法,如果能使用较好的分类器(如二次判决函数分类器、SVM 神经网络分类器、多分类器集成等),并结合本文所述的方向特征方法构造实际的识别系统,识别率有望还能得到较大提高。

参考文献:

[1] Qivind Due Trier, et al. Feature Extraction Methods for Character Recognition: A Survey[J]. Patter Recognition, 1996, 29(4): 641-662.

[2] R Plamondon, S N Srihari. On-line and Off-line Handwriting Recognition: A Comprehensive Survey[J]. IEEE Trans. on PAMI, 2000, 22(1): 63-81.

[3] Yi-Hong Tseng, Chi-Chang Kuo, Hsi-Jian Lee. Speeding Up Chinese Character Recognition in an Automatic Document Reading System[J]. Pattern Recognition, 1998, 31(11): 1601-1612.

(下转第 90 页)

http :/ / www. javareport. com

表 1 查询 Java 的结果

	NDDS	Google	Yahoo
1	http :/ / java. sun. com/	http :/ / java. sun. com/	http :/ / java. sun. com
2	http :/ / www. gamelan. com	http :/ / www. sun. com/	http :/ / www. sun. com/
3	http :/ / www. jcp. org/	http :/ / javaboutique. internet. com/	http :/ / javaboutique. internet. com
4	http :/ / www. borland. com/ jbuilder	http :/ / javascript. internet. com/	http :/ / softwaredev. earthWeb. com/ java
5	http :/ / developer. java. sun. com	http :/ / www. javaworld. com/	http :/ / javascript. internet. com
6	http :/ / www. javaworld. com	http :/ / www. blackdown. org/ java-linux. html	http :/ / www. javaworld. com/
7	http :/ / www. jars. com	http :/ / www. javalobby. org/ index. jsp	http :/ / www. blackdown. org/ java-linux. html
8	http :/ / javaboutique. internet. com/	http :/ / www. anfyteam. com/	http :/ / www. javalobby. com
9	http :/ / www. javasoft. com/	http :/ / developer. java. sun. com/	http :/ / www. microsoft. com/ java/
10	http :/ / www. uddi. org	http :/ / www6. sun. com/ software/ java/	http :/ / www. jars. com

查询 2 :Abortion

根集 :http :/ / www. gynpages. com

http :/ / www. naral. org/

http :/ / www. abortionfacts. com/

表 2 查询 Abortion 的结果

	NDDS	Google	Yahoo
1	http :/ / www. nrlc. org/	http :/ / www. gynpages. com	http :/ / www. gynpages. com
2	http :/ / www. gynpages. com/	http :/ / www. prochoice. org/	http :/ / www. prochoice. org/
3	http :/ / www. nnuf. org	http :/ / www. naral. org/	http :/ / www. naral. org
4	http :/ / www. naral. org/	http :/ / hometown. aol. com/ abtribng/ index. htm	http :/ / www. caral. org
5	http :/ / www. caral. org	http :/ / www. abortionfacts. com/	http :/ / www. abortionfacts. com
6	http :/ / www. plannedparenthood. org	http :/ / www. caral. org/	http :/ / www. religioustolerance. org/ abortion. htm
7	http :/ / www. NDDSonline. com	http :/ / www. religioustolerance. org/ abortion. htm	http :/ / ethics. acusd. edu/ abortion. html
8	http :/ / www. caral. org/ abortion. html	http :/ / ethics. acusd. edu/ Applied/ abortion/	http :/ / www. abortionrv. com
9	http :/ / www. isad. org	http :/ / www. overvoice. com/ aam/ main/	http :/ / www. afterabortion. org
10	http :/ / www. prochoice. org/	http :/ / www. afterabortion. org/	http :/ / www. overvoice. com/ aam/ main

NDDS 与搜索引擎 Google , Yahoo 的搜索结果比较 ,我们可以发现 ,在某个主题下一些重要的资源都出现在各自结果的最前面 ,同时每个系统结果的侧重点不相同。NDDS 系统是根据用户反馈的根集资源 ,自动学习分析后到 WWW 发现相关资源的 ,所以结果带有个性化的成分。Google 和 Yahoo 的结果是从它们本地的数据库中检索出来的 ,所以结果比较相似。另外 ,由于 NDDS 是实时地从 WWW 上分析获得资源的 ,所以搜索的时间稍长。

4 结束语

NDDS 系统使用经过改进的 VSM ,智能 Crawler ,链接分析三个关键技术 ,为用户提供了令人满意的搜索结果。通过用户的反馈 ,系统提供个性化的实时搜索服务。随着用户的增加 ,系统在每个主题下学习和积累的资源会越来越丰富 ,不同用户之间的资源可以得到共享 ,系统的搜索结果会越来越令人满意。因为系统的搜索结果主要是实时的从 WWW 上下载并且经过分析 ,系统现在的硬件配置也不高 ,所以搜索的时间比现有的搜索引擎稍长。随着网络带宽的增加和使用更高性能的机器 ,这个问题是能够解决的。另外 ,系统是简单地使用一个标准向量来标志一个主题 ,效率很高 ,如果结合使用分类聚类技术 ,精度还可以提高。

参考文献:

[1] R Baeza-Yates ,B Ribeiro-Neto. Moderm Information Retrieval[M]. Addison Wesley ,1999.

[2] G Salton. Automatic Text Processing: The Transformation ,Analysis ,and Retrieval of Information by Computer[M]. Addison-Wesley Series in Computer Science ,Addison-Wesley Longman Publ. Co. Inc ,1989.

[3] Sergey Brin ,Larry Page. The Anatomy of a Large-scale Hypertextual Web Search Engine[C]. Proceedings of the 7th International World Wide Web Conference ,1998.

[4] J Kleinberg. Authoritative Sources in a Hyperlinked Environment[R]. IBM Research Report RJ 9076 ,1997.

[5] R Lempel ,S Moran. The Stochastic Approach for Link-Structure Ana-lysis (SALSA) and the TKC Effect[C]. Proc. 9th International World Wide Web Conferrence ,2000.

[6] Chakrabarti ,M van den Berg ,B Dom. Focused Crawling :A New Approach Topic-specific Web Resource Discovery[C]. The 8th International WWW Conference ,1999.

[7] CC Aggarwal ,F Al-Garawi ,PS Yu. Intelligent Crawling on the World Wide Web with Arbitrary Predicates[C]. Proc. 10th International WWW Conference ,2001.

[8] 韩家炜 ,孟小峰 ,王静 ,等. Web 挖掘研究[J]. 计算机研究与发展 ,2001 ,38(4).

[9] Soumen Chakrabarti ,Byron Dom ,Automatic Resource Compilation by Analyzing Hyperlink Structure and Associated Text[C]. Proc. 7th International WWW Conference ,1998.

[10] http :/ / www. google. com[EB/ OL].

[11] http :/ / www. yahoo. com[EB/ OL].

作者简介:

朱炜(1978-) ,男 ,硕士研究生 ,主要研究领域为 Web 信息检索、Agent 技术;李俊(1977-) ,男 ,硕士研究生 ,主要研究领域为 Agent 技术、信息检索;王超(1977-) ,男 ,硕士 ,主要研究领域为 Agent 技术、信息检索;潘金贵(1952-) ,男 ,教授 ,博士生导师 ,主要研究领域为中间件、Agent 技术、多媒体远程教育。

(上接第 40 页)

[4] C L Liu ,Y J Liu ,R W Dai. Preprocessing and Statistical/ Structural Feature Extraction for Handwritten Numeral Recognition[C]. Progress of Handwriting Recognition ,Downton & Impedovo ,World Scientific ,1997. 162-168.

[5] 金连文 ,徐秉铮. 手写体汉字识别的一种新的特征提取方法 —弹性网格方向分解特征[J]. 电路与系统学报 ,1997 ,2(3) :7-12.

[6] 陈友斌 ,丁晓青. 一种手写特征提取新方法[J]. 信号处理 ,1998 ,14(2) :117-122.

[7] 马少平 ,夏莹 ,朱小燕. 基于模糊方向线素特征的手写体汉字识别[J]. 清华大学学报 ,1997 ,37(3) :42-45.

[8] H W Hao ,X H Xiao ,R W Dai. Handwritten Chinese Character Recognition by Metasynthetic Approach[J]. Pattern Recognition ,1997 ,30(8) :1321-1328.

[9] Lianwen J IN ,Gang Wei. Handwritten Chinese Character Recognition with Directional Decomposition Cellular Features[J]. Journal of Circuit ,System and Computer ,1998 ,8(4) :517-524.

[10] Tze Fen Li ,Shiaw Shian Yu. Handprinted Chinese Character Recognition Using the Probability Distribution Feature[J]. International Journal of Pattern Recognition and Artificial Intelligence ,1994 ,(8) :1241-1258.

[11] 金连文 ,梁羽杰. 一种新的距离分类方法及其应用[J]. 计算机工程 ,1999 ,25(8) :30-32.

[12] 郭军 ,蔺志青 ,张洪刚. 一个新的手写汉字数据库模型及其应用[J]. 电子学报 ,2000 ,28(5) :115-116.

[13] Xue Gao ,et al. A New Stroke-based Directional Feature Extraction Approach for Handwritten Chinese Character Recognition[C]. Proceedings ICDAR2001 ,USA ,2001. 635-639.

[14] 高学 ,金连文 ,尹俊勋. 基于笔画密度的弹性网格特征提取方法[J]. 模式识别与人工智能 ,2002 ,15(3) :351-354.

作者简介:

金连文 ,副教授 ,博士 ,研究方向为汉字识别、图像处理、计算机视觉、模式识别等;高学 ,博士研究生 ,研究方向为汉字识别、图像处理、遗传算法。