

# 基于层次分类的脱机手写字符识别

王云鹏 苗夺谦 岳晓东

(同济大学电子与信息工程学院计算机科学与技术系 上海 201804)

(同济大学嵌入式系统与服务计算教育部重点实验室 上海 201804)

(国家高性能计算机工程中心同济分中心 上海 201804)

**摘要** 人类在进行字符识别活动时,会根据对象复杂度的不同,采用不同的识别方法。对于结构简单的字符,利用宏观整体信息识别;对于易混淆的形近字,利用微观具体信息区分。为了模拟人类智能进行字符识别活动的过程,设计了一种基于层次分类的脱机手写字符识别算法。该算法将分类器划分为宏观层和微观层,宏观层模拟简单字符识别过程,利用基于梯度的统计特征描述整体信息,完成识别;微观层模拟形近字识别过程,利用基于主曲线的结构特征描述具体信息,完成区分。算法还引入了可信度概念,用以量度推理过程及识别结果的不确定性程度。给出了形近字的定义及区分规则。实验表明,提出的算法有效地提高了脱机手写字符的识别率,对形近字的区分效果尤佳。

**关键词** 层次分类,手写字符识别,可信度,形近字,主曲线,梯度

## Off-line Handwritten Character Recognition Based on Hierarchical Classification

WANG Yun-peng MIAO Duo-qian YUE Xiao-dong

(Department of Computer Science and Technology, Tongji University, Shanghai 201804, China)

(The Key Laboratory of "Embedded System and Service Computing" Ministry of Education, Tongji University, Shanghai 201804, China)

(Tongji Branch, National Engineering & Technology Center of High Performance Computer, Shanghai 201804, China)

**Abstract** The paper proposed a method of off-line handwritten character recognition based on hierarchical classification. The method simulates the produce of character recognition of human. When a man wants to recognize a character, he uses different strategy in different situation. If the character has a simple structure, he uses global features; if it looks similar with other character, he uses local features. We divided the classifier into macro layer and micro layer. The macro layer uses gradient feature to represent global feature, it simulates the simple target recognition produce; the micro layer uses principal curve feature to represent local feature, it simulates the similar form character recognition produce. We used confidence value to measure indeterminacy of the produce and result. We gave the definition of similar form character, and rules to telling them. The experimental results indicate that the method can effectively improve the recognition rate of off-line handwritten character, especially well in telling similar form character.

**Keywords** Hierarchical classification, Handwritten character recognition, Confidence value, Similar form character, Principal curve, Gradient

## 1 引言

脱机手写字符识别技术在票据处理、签名识别、计算机自动阅读等领域有着广泛的应用<sup>[1,2]</sup>,多年来一直是模式识别领域的研究热点。经过数十年研究发展,研究者们提出了众多特征选取和分类器设计的方法。特征选取的方法主要分为基于统计特征和基于结构特征两类。基于统计特征的方法,如模板匹配<sup>[3,4]</sup>、傅立叶算子<sup>[5]</sup>、不变矩<sup>[6,7]</sup>等,算法复杂度低,训练简单,但抗形变能力差,不适合区分形近字。基于结构的方法,如骨架<sup>[8,9]</sup>、轮廓<sup>[10,11]</sup>、语义信息<sup>[12,13]</sup>等,对形近字区分能力强,但算法复杂度高。分类器设计方面,人工神经网络<sup>[13-15]</sup>、支撑向量机<sup>[16,17]</sup>等技术被应用到了字符识别研究

中,有效地提高了识别精度,但仍缺少基于特征的优化设计。

对人类视觉活动的研究表明,当人眼进行字符识别工作时,会根据对象结构复杂度的不同,采取两种不同的策略。对于结构简单的字符,人类将获取字符的整体信息,并在脑中搜索,完成与已存字符模板的匹配工作,形成认知结果;对于结构复杂或易混淆的字符,人类需要将它分解为更简单的结构或选取更具体的信息,进行搜索匹配,再将信息整合,形成认知结果<sup>[18]</sup>。根据人类视觉活动特点,将统计特征与结构特征相结合,是增强分类能力,降低算法复杂度的有效途径。现有的多分类器结合的方法,研究内容侧重于信息格式统一、分类器连接、可信度传递等方面,缺少分类器与图像语义信息相结合的研究<sup>[19,20]</sup>。

本文根据人类字符识别活动的原理,选取基于梯度的统计特征描述字符的宏观整体信息,选取基于主曲线的结构特征描述字符的微观具体信息,设计了基于层次分类的脱机手写字符识别算法,引入可信度概念、量度推理过程及识别结果的不确定性程度,并给出了形近字的定义及区分规则。分类器的宏观层利用统计特征,降低了识别算法的复杂度,而微观层利用结构特征,增强了对形近字的区分能力。

第2节将描述本文提出的用于脱机手写字符识别的层次分类器的算法框架,并给出形近字的定义及可信度的概念;第3节介绍宏观层的特征选取,分类器设计及可信度推理;第4节说明微观层的特征选取,分类器设计及形近字的区分规则;第5节将给出实验数据及结果分析;最后对本文工作进行总结和展望。

## 2 层次分类器

### 2.1 算法框架

本文提出的分类器算法模拟人类智能,利用分层的思想处理字符识别问题。分类器由处理整体信息的宏观层和处理具体信息的微观层组成,每层的识别结果由可信度机制评判。字符首先进入分类器的宏观层。分类器计算字符的梯度地图作为统计特征,调用最短距离分类算法,输出分类结果,并计算结果的可信度。分类器将宏观层识别结果的可信度与预先设置好的用于判别形近字的可信度阈值相比较,如果可信度大于阈值,分类器判断字符为非形近字,并将结果送往筛选器。否则,分类器判断字符为形近字,并根据宏观层分类结果,计算字符所属的形近字类别,将字符送入微观层。对于进入分类器微观层的字符,分类器提取其主曲线,并根据主曲线计算字符的结构特征,利用决策表中的形近字区分规则,得到识别结果,并更新可信度,将结果送入筛选器。筛选器中存有分类器用于判别识别结果是否为可以信任的可信度阈值,分类器将筛选器中识别结果的可信度与阈值相比较,对于大于阈值的结果,分类器判断结果可信,并输出结果;对于不大于阈值的结果,分类器判断结果不可信,拒绝接受识别结果。图1为算法流程图。

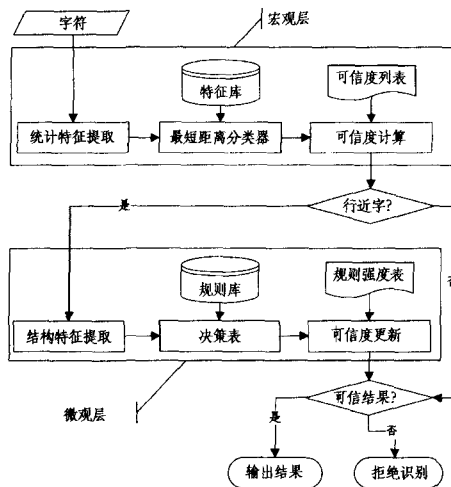


图1 层次分类器算法流程图

#### 算法1 层次分类器算法

输入:待识别字符  $C$

输出:识别结果  $R$

Step1 计算  $C$  的梯度地图特征  $F_1$ 。

Step2 调用最短距离分类器,根据  $F_1$ ,计算  $C$  的宏观层分类结

果  $R_h$ ,置  $R = R_h$ 。

Step3 计算识别结果可信度  $CF_1(H)$ ,置  $CF(H) = CF_1(H)$ 。

Step4 设形近字判别可信度阈值为  $CF_{sim}$ ,如果  $CF_1(H) > CF_{sim}$ ,转至 Step9。

Step5 根据  $R_h$ ,计算  $C$  所属的形近字类别  $S$ 。

Step6 计算  $C$  的主曲线结构特征  $F_s$ 。

Step7 查找决策表,根据  $F_s$ ,获得微观层分类结果  $R_w$ ,置  $R = R_w$ 。

Step8 更新识别结果可信度  $CF_2(H)$ ,置  $CF(H) = CF_2(H)$ 。

Step9 设结果信任可信度阈值为  $CF_{tr}$ ,如果  $CF(H) > CF_{tr}$ ,输出结果;否则,拒绝识别。

### 2.2 形近字

形近字区分是本文要解决的主要问题。下文根据宏观层分类结果,定义了形近字,并给出在宏观层选用梯度作为统计特征的条件,形成的形近字类别。本文所做的研究工作,都是基于这组形近字类别。

使用同一特定的分类器对一组字符进行识别,如果类别为  $C_m$  的字符有超过 5% 的概率被分类器识别为类别  $C_n$ ,则称类别  $C_n$  为类别  $C_m$  的形近字;如果类别  $C_m$  与类别  $C_n$  互为形近字,则称  $C_m$  与  $C_n$  属于同一形近字簇。形式化定义如下:

**定义1(形近字)** 已知数据集  $D$ ,分类器  $F$ , $C_m, C_n$  是  $D$  上的数据类别, $s$  是数据样本,设  $c(s)$  表示样本  $s$  的数据类别, $f(s)$  表示分类器  $F$  对样本  $s$  的分类结果的数据类别,对于  $D$  上的任意样本  $s$ ,若  $P(f(s) = C_n | c(s) = C_m) > 5\%$ ,则称  $C_n$  是  $C_m$  的形近字,记作  $C_n \approx C_m$ 。

**定义2(形近字簇)** 若  $C_n \approx C_m$  且  $C_m \approx C_n$ ,则称  $C_m$  与  $C_n$  属于同一形近字簇,记作  $C_m = C_n$ 。

**性质1(传递性)** 形近字簇具有传递性,若  $C_m = C_n, C_n = C_k$ ,则  $C_m = C_k$ 。

本文研究的数据集由大写英语字母  $A-Z$ ,小写英语字母  $a-z$ ,数字  $0-9$  组成,选用由基于梯度地图特征的最短距离分类器产生的一组形近字及形近字簇。另外,对于形态基本一致且人类智能需要依靠上下文对比信息才能区分的字符,如大写字母  $S$  与小写字母  $s$ ,分类器认为它们所属的类别相同,不作区分。表1给出了本文的形近字库。

表1 形近字库

形近字	形近字簇	不作区分字符
$G \approx 6F \approx T$	$g = q = y = 9$	$C \equiv cP \equiv p$
$J \approx TY \approx T$	$a = Q = 0$	$S \equiv sU \equiv u$
$t \approx l e \approx 2$		$V \equiv vW \equiv w$
$k \approx r$		$X \equiv xZ \equiv z$
		$0 \equiv O \equiv o1 \equiv I$

### 2.3 可信度

可信度是不确定性推理中用于量度证据、规则、结论不确定性的一种方法<sup>[21]</sup>。人类的字符识别活动,受到客观成像、主观知识、生理机能等多方面条件影响,本身就是一种不确定性推理的过程。本文引入可信度概念,用以量度字符识别过程中的不确定性,用可信度推理的方法,模拟人类智能的判别活动。

可信度在本文设计的层次分类算法中,具体的作用有以下3方面:

第一,量度宏观层分类结果的不确定性程度,判别字符是否统计特征不易区分的形近字。

第二,量度微观层形近字区分规则的不确定性程度,即规

则强度,以区分形近字,更新形近字可信度。

第三,量度最终分类结果的不确定性程度,判别分类结果是否可以被信任。

下文给出可信度的概念及可信度推理的方法。

设一个不精确推理过程的证据为  $E$ , 结论为  $H$ , 推理规则为:

IF  $E$  THEN  $H$

可以用可信度  $CF$  量度证据, 结论与规则的不确定性程度,  $CF$  的作用域为  $[-1, 1]$ 。

$CF(E)$  表示证据的可信度, 当证据以某种程度为真时,  $CF(E) > 0$ ,  $CF(E)$  的值越大, 证据越真, 当证据肯定为真时,  $CF(E) = 1$ ; 当证据以某种程度为假时,  $CF(E) < 0$ ,  $CF(E)$  的值越小, 证据越假, 当证据肯定为假时,  $CF(E) = -1$ ; 当对证据一无所知时,  $CF(E) = 0$ 。  $CF(E)$  的值是经过对大量样本的观察实验, 由实验数据给出的。

$CF(H, E)$  表示规则的可信度, 称为规则强度。  $CF(H, E) > 0$ , 表示证据增加了结论为真的程度, 且  $CF(H, E)$  的值越大, 结论  $H$  越真, 若  $CF(H, E) = 1$ , 则表示该证据使结论为真; 反之,  $CF(H, E) < 0$ , 表示证据增加了结论为假的程度, 且  $CF(H, E)$  的值越小, 结论  $H$  越假, 若  $CF(H, E) = -1$ , 则标识该证据使结论为假。  $CF(H, E) = 0$ , 表示证据与结论没有关系。

$CF(H, E)$  的值由规则强度的定义式计算得到。

定义 3(规则强度)

$$CF(H, E) = MB(H, E) - MD(H, E) \quad (1)$$

定义 4(信任增长度)

$$MB(H, E) = \begin{cases} 1, & \text{若 } P(H) = 1 \\ \frac{\max\{P(H|E), P(H)\} - P(H)}{1 - P(H)}, & \text{其他} \end{cases} \quad (2)$$

定义 5(非信任增长度)

$$MD(H, E) = \begin{cases} 1, & \text{若 } P(H) = 0 \\ \frac{\min\{P(H|E), P(H)\} - P(H)}{-P(H)}, & \text{其他} \end{cases} \quad (3)$$

$CF(H)$  表示结论的可信度, 当结论以某种程度为真时,  $CF(H) > 0$ ,  $CF(H)$  的值越大, 结论越真, 当结论肯定为真时,  $CF(H) = 1$ ; 当结论以某种程度为假时,  $CF(H) < 0$ ,  $CF(H)$  的值越小, 结论越假, 当结论肯定为假时,  $CF(H) = -1$ ; 当对结论一无所知时,  $CF(H) = 0$ 。

$CF(H)$  的值由证据和规则推理得到。

$$CF(H) = CF(H, E) \cdot \max\{0, CF(E)\} \quad (4)$$

若两条规则推出相同结论, 可信度各不相同, 则不确定性的合成算法为:

$$CF_{1,2}(H) = \begin{cases} CF_1(H) + CF_2(H) - CF_1(H) \cdot CF_2(H) & \text{if } CF_1(H) \geq 0, CF_2(H) \geq 0 \\ CF_1(H) + CF_2(H) + CF_1(H) \cdot CF_2(H) & \text{if } CF_1(H) < 0, CF_2(H) < 0 \\ CF_1(H) + CF_2(H) & \text{if } CF_1(H) \cdot CF_2(H) < 0 \end{cases} \quad (5)$$

### 3 宏观层

分类器设计宏观层的目的是模拟人类识别结构简单的字

符的过程, 利用整体信息, 快速准确地对形态差别较大的字符进行识别, 并计算识别结果的可信度。对于不超过可信度阈值的结果, 分类器将字符送交微观层处理。

宏观层具有提取统计特征、分类、计算可信度 3 部分功能。

#### 3.1 特征提取

宏观层选取的特征应具备以下特点: 对字符整体结构的描绘能力强; 计算复杂度低; 特征不易区分的形近字数量少, 且符合人类认知。这些正是统计特征的优点。梯度地图特征对轮廓变化比较敏感, 描述能力强, 信息量大, 且算法复杂度较低<sup>[22-24]</sup>, 满足宏观层对特征的要求, 是本文选用的统计特征。

下面给出梯度地图特征的概念, 本文将在实验部分进一步验证选用梯度地图特征的合理性。

梯度地图特征是利用分区思想, 通过计算每个区域中各梯度值的像素点个数而得到的图像统计特征。像素点梯度值的计算是提取梯度地图特征的基础。

设图像  $I$  中点  $(x, y)$  的像素值为  $I(x, y)$ , 首先, 分别求取点  $(x, y)$  的梯度水平分量  $S_x$  和梯度垂直分量  $S_y$ , 方法如下:

$$S_x(i, j) = I(i-1, j+1) + 2I(i, j+1) + I(i+1, j+1) - I(i-1, j-1) - 2I(i, j-1) - I(i+1, j-1) \quad (6)$$

$$S_y(i, j) = I(i-1, j-1) + 2I(i-1, j) + I(i-1, j+1) - I(i+1, j-1) - 2I(i+1, j) - I(i+1, j+1) \quad (7)$$

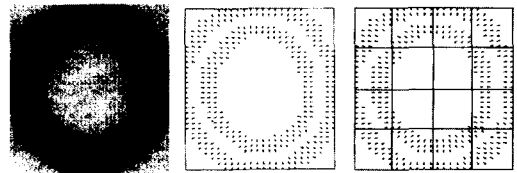
然后, 利用  $S_x$  和  $S_y$  的比值, 求出梯度方向  $\theta$ , 方法如下:

$$\theta(i, j) = \tan^{-1} \frac{S_y(i, j)}{S_x(i, j)} \quad (8)$$

最后, 通过量化, 得到取值范围为  $[1, 8]$  内整数的梯度值  $G$ , 方法如下:

$$G(i, j) = \theta(i, j) \bmod 45^\circ \quad (9)$$

由图像每个像素的梯度值构成的数据集合, 即为梯度地图。将梯度地图平分为  $4 \times 4$  个子区域, 分别计算每个子区域内  $1 \sim 8$  各梯度值的数量  $f_1 \sim f_8$ , 并将各子区域求得的特征  $f$  顺序组合, 得到 128 维的梯度地图特征  $F = (f_1, f_2, \dots, f_{128})$ 。图 2 给出了字符 0 的梯度地图及分区梯度地图特征。



(a) 原始图像; (b) 梯度地图; (c) 分区梯度地图特征

图 2

#### 3.2 分类算法

宏观层所选用的分类算法, 应具备的最主要特征是执行速度快。最短距离分类器是算法复杂度最低的分类器之一, 并且训练方法简单<sup>[26]</sup>, 能够很好地满足宏观层分类算法的要求。本文选用最短距离分类器实现宏观层的分类功能。

下面给出最短距离分类器的实现原理。

设样本  $s$  的特征向量为  $x$ , 类别  $C_i$  的样本数为  $N_i$ , 平均向量为  $m_i$ ,  $x$  到  $m_i$  的欧式距离为  $D_i(x)$ , 则:

$$m_i = \frac{1}{N_i} \sum_{x \in C_i} x_i \quad i = 1, 2, \dots, W \quad (10)$$

$$D_i(x) = \|x - m_i\| \quad i = 1, 2, \dots, W \quad (11)$$

样本  $s$  的类别是使  $D_i(x)$  达到最小值的类别  $C_i$ 。

### 3.3 可信度计算

分类器在宏观层的最后,计算识别结果的可信度,用以判断字符是否为形近字。下文将给出宏观层分类结果的可信度计算方法。

设字符类别总数为  $N$ ,则识别结果的先验概率为  $P(H)=1/N$ ;条件概率  $P(H|E)$  为实验得到的统计数据,利用最短距离分类器进行推理的规则强度  $CF_1(H,E)$  为:

$$CF_1(H,E) = \frac{P(H|E) - P(H)}{1 - P(H)} \quad (12)$$

设宏观层证据可信度  $CF_1(E)=1$ ;则宏观层结论可信度  $CF_1(H)$  为:

$$CF_1(H) = CF_1(H,E) \cdot CF_1(E) = \frac{P(H|E) - P(H)}{1 - P(H)} \quad (13)$$

对于  $CF_1(H)$  的值不大于形近字判别可信度阈值  $CF_{sim}$  的字符,分类器查找形近字的均值向量列表,通过最短距离,确定字符的形近字类别,并将字符送入微观层。

## 4 微观层

分类器设计微观层的目的是模拟人类识别结构复杂、易混淆的字符的过程,利用更具体、更本质的信息,准确区分结构差别小、统计特征不易区分的形近字,并更新识别结果的可信度。

微观层具有提取结构特征、区分形近字、更新可信度 3 部分功能。

### 4.1 特征提取

微观层选取的特征应该能够描绘更具体的信息,反映字符更本质的特征,刻画形近字间更细微的差别。基于主曲线的结构特征,能够很好地满足微观层的要求。本文选用推广的多边形主曲线算法,生成字符骨架,并提取笔画数、回路数等作为微观层选用的结构特征。

下面将介绍主曲线的概念、特点,并给出基于推广的多边形主曲线算法得到的结构特征的定义。

#### 4.1.1 主曲线

主曲线概念是 Hastie 和 Stuetzle 于 1984 年提出的<sup>[26]</sup>。定义如下:

**定义 6(主曲线)** 假设随机向量  $X=(X_1, X_2, \dots, X_p)$  的概率密度为  $g_x(x)$ ,则通过  $X$  数据分布中间的一条曲线  $f(\lambda)$  如果满足:

$$f(\lambda) = E(X | \lambda_f(x) = \lambda) \quad (14)$$

则称  $f(\lambda)$  是  $X$  的一条主曲线。其中  $\lambda_f(x)$  是数据点  $x$  投影到曲线  $f(\lambda)$  上  $\lambda$  点的值,即:

$$\lambda_f(x) = \sup\{\lambda: \|x - f(\lambda)\| = \inf\|x - f(\tau)\|\} \quad (15)$$

形象地说,主曲线就是一条通过数据分布“中间”并满足“自相合”的光滑曲线。它具有两个明显的优点:一是对数据的信息保持性好,二是它与数据间的距离均方差小,较好地勾画出了原始信息的轮廓。在最初的 HS 主曲线之后,研究者们又提出了 T 主曲线、D 主曲线、K 主曲线等多种主曲线<sup>[27]</sup>。其中,推广的多边形主曲线对复杂的结构进行了优化,并且算法复杂度低<sup>[28]</sup>,在相关的数字识别研究中,取得了不错的识别效果<sup>[29,30]</sup>。本文利用推广的多边形主曲线提取字符骨架,并计算基于主曲线定义的结构特征,用以区分形近字。

#### 4.1.2 结构特征选取

本文通过对大量形近字样本的观察实验,挑选了以下 10 种结构特征,用以区分形近字。

**定义 7(笔画数 NS)** 字符中光滑曲线段的个数称为笔画数。

**定义 8(回路数 NC)** 字符中闭合光滑曲线段的个数称为回路数。

**定义 9(直线数 NL)** 在这里定义的直线是近似直线。设曲线的两个端点为  $v_i$  和  $v_j$ ,如果曲线满足两端点间距离  $d(v_i, v_j)$  除以曲线长度  $\sum_{k=i}^{j-1} d(v_k, v_{k+1})$  大于某一参数  $parameter1$ ,即:

$$\frac{d(v_i, v_j)}{\sum_{k=i}^{j-1} d(v_k, v_{k+1})} > parameter1 \quad (16)$$

则认为这条曲线是直线,否则称曲线不是直线。经大量训练样本训练可知参数  $parameter1=0.871$  时效果最好,组成字母的曲线中直线的总数称为直线数。

**定义 10(长直线数 NLL)** 设曲线的两个端点为  $v_i$  和  $v_j$ ,曲线所属字符的最高点为  $v_h$ ,最低点为  $v_l$ ,如果曲线满足两端点间距离  $d(v_i, v_j)$  大于字符高度  $d(v_h, v_l)$  的一半,即:

$$d(v_i, v_j) > d(v_h, v_l) / 2 \quad (17)$$

则认为这条曲线是长直线,否则称曲线不是长直线。组成字母的曲线中长直线的总数称为长直线数。

**定义 11(水平线数 NHL)** 在这里定义的水平线是近似水平线。设曲线的两个端点为  $v_i$  和  $v_j$ ,其横坐标分别为  $x_i$  和  $x_j$ ,两端点横坐标间距离  $d(v_i, v_j) = |x_i - x_j|$ ,如果曲线满足  $d(v_i, v_j)$  除以曲线长度  $\sum_{k=i}^{j-1} d(v_k, v_{k+1})$  大于某一参数  $parameter2$ ,即:

$$\frac{d(v_i, v_j)}{\sum_{k=i}^{j-1} d(v_k, v_{k+1})} > parameter2 \quad (18)$$

则认为这条曲线是水平线,否则称曲线不是水平线。经大量训练样本训练可知参数  $parameter2=0.897$  时效果最好,组成字母的曲线中水平线的总数称为水平线数。

**定义 12(竖直线数 NVL)** 在这里定义的竖直线是近似竖直线。设曲线的两个端点为  $v_i$  和  $v_j$ ,其纵坐标分别为  $y_i$  和  $y_j$ ,两端点纵坐标间距离  $d(v_i, v_j) = |y_i - y_j|$ ,如果曲线满足  $d(v_i, v_j)$  除以曲线长度  $\sum_{k=i}^{j-1} d(v_k, v_{k+1})$  大于某一参数  $parameter3$ ,即:

$$\frac{d(v_i, v_j)}{\sum_{k=i}^{j-1} d(v_k, v_{k+1})} > parameter3 \quad (19)$$

则认为这条曲线是竖直线,否则称曲线不是竖直线。经大量训练样本训练可知参数  $parameter3=0.934$  时效果最好,组成字母的曲线中竖直线的总数称为竖直线数。

**定义 13(交点数 NP)** 字符中光滑曲线段的交点个数称为交点数。

**定义 14(回路位置 PC)** 设回路中心坐标为  $(x_s, y_s)$ ,字符中心坐标为  $(x_c, y_c)$ ,回路中心相对于字符中心的位置,称为回路位置,值为:

$$PC = \begin{cases} 1 & \text{若 } y_s < y_c \\ 2 & \text{若 } x_s < x_c \text{ 且 } y_s > y_c \\ 3 & \text{若 } x_s > x_c \text{ 且 } y_s > y_c \end{cases} \quad (20)$$

对于有  $n$  个回路的字符,各回路位置分别表示为  $PC_1, PC_2, \dots, PC_n$ 。

**定义 15(终点位置 PP)** 设字符终点坐标为  $(x_t, y_t)$ , 字符中心坐标为  $(x_c, y_c)$ , 终点相对于字符中心的位置,称为终点位置,值为:

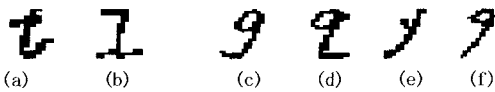
$$PP = \begin{cases} 1 & \text{若 } x_t < x_c \\ 2 & \text{若 } x_t > x_c \end{cases} \quad (21)$$

**定义 16(回路内端点数 NPC)** 对于仅有一个回路的字符,该回路内端点的个数,称为回路内端点数。

#### 4.2 形近字区分

微观层采用决策表作为分类器,利用上文定义的 10 种结构特征形成的规则,对形近字进行区分。本文利用基于互信息的属性约简算法<sup>[31]</sup>,去除决策表内的冗余属性,得到了最终的形近字区分规则。

对于进入微观层的形近字,分类器根据形近字类别信息,查找决策表中对应规则,得到分类结果。例如,图 3 中小写字母  $t$  与大写字母  $I$  具有相近的梯度地图特征,宏观层不易区分,但他们的结构特征差别很大,  $t$  必然由一条水平线与一条竖直线相交而成,交叉点两侧一定存在两段水平线,  $NHL=2$  的字符一定是  $t$ ;又如,图 3 中形近字簇  $g=q=y=9$ ,形态相近,且  $g, y$  都存在多种写法,需要利用回路数、回路位置、长直线数等结构特征的组合相互区分。表 2、表 3 给出了形近字及形近字簇的区分规则。



(a)~(b)为  $t \approx I$ ; (c)~(f)为  $g=q=y=9$

图 3

表 2 形近字区分规则

形近字	规则及规则强度
$e \approx 2$	IF NS=2, NC=1 THEN $R_w = 'e'$ $CF_2(H, E) = 1$
$G \approx 6$	IF NS>2, NC=0 THEN $R_w = '6'$ $CF_2(H, E) = -1$
$t \approx I$	IF NHL=2 THEN $R_w = 't'$ $CF_2(H, E) = 1$
$k \approx r$	IF NS=4, NL=4 THEN $R_w = 'k'$ $CF_2(H, E) = 1$
$J \approx T$	IF NVL=0 THEN $R_w = 'J'$ $CF_2(H, E) = 1$
$Y \approx T$	IF NHL=0 THEN $R_w = 'Y'$ $CF_2(H, E) = 1$
$F \approx T$	IF NHL>2 THEN $R_w = 'F'$ $CF_2(H, E) = 1$

表 3 形近字簇区分规则

形近字簇	规则及规则强度
$a=Q=0$	IF NS=1, NC=1 THEN $R_w = '0'$ $CF_2(H, E) = 0.95$ IF NS>3, NP>2 THEN $R_w = 'Q'$ $CF_2(H, E) = 0.9$ IF NS<4, NC=0 THEN $R_w = 'a'$ $CF_2(H, E) = 0.9$ IF NPC>0 THEN $R_w = 'Q'$ $CF_2(H, E) = 0.8$ IF NPC=0 THEN $R_w = 'a'$ $CF_2(H, E) = 0.8$
$g=q=y=9$	IF NLL>1 THEN $R_w = 'y'$ $CF_2(H, E) = 0.95$ IF PC=2 THEN $R_w = 'y'$ $CF_2(H, E) = 0.95$ IF NLL=1, PC=1 THEN $R_w = '9'$ $CF_2(H, E) = 0.7$ IF PC <sub>1</sub> =1, PC <sub>2</sub> =2 THEN $R_w = 'g'$ $CF_2(H, E) = 0.8$ IF PC=1, PP=1 THEN $R_w = 'g'$ $CF_2(H, E) = 0.8$ IF PC <sub>1</sub> =1, PC <sub>2</sub> =3 THEN $R_w = 'q'$ $CF_2(H, E) = 0.8$ IF PC=1, PP=2 THEN $R_w = 'q'$ $CF_2(H, E) = 0.8$

#### 4.3 可信度更新

分类器在微观层的最后,对识别结果的可信度进行更新。下文将给出微观层的可信度更新算法。

设微观层规则库的条件可信度  $CF_2(E) = 1$ ; 规则强度  $CF_2(H, E)$  是实验得到的统计数据,是存储在规则库中的

量。则微观层结论可信度  $CF_2(H)$  为:

$$CF_2(H) = CF_2(H, E) \cdot CF_2(E) = CF_2(H, E) \quad (22)$$

设形近字可信度阈值为  $CF_{sm}$ , 根据式(13), 式(22), 得到分类器对字符  $C$  分类结果的可信度  $CF(H)$  为:

$$CF(H) = \begin{cases} \frac{P(H|E) - P(H)}{1 - P(H)}, & \text{若 } CF_1(H) > CF_{sm} \\ CF_2(H, E), & \text{其他} \end{cases} \quad (23)$$

对于  $CF(H)$  的值大于结果信任阈值  $CF_{th}$  的字符, 分类器输出分类结果; 否则, 分类器认为字符无法识别。

## 5 实验与数据分析

本文根据人类智能进行字符识别活动的原理, 设计了一种基于层次分类的脱机手写字符识别算法。该算法的性能可以从 3 方面衡量。第一, 宏观层模拟人类识别简单字符过程的效果; 第二, 微观层模拟人类识别形近字过程的效果; 第三, 可信度机制模拟人类对认知结果判别过程的效果。本文通过实验, 对分类器上述 3 方面的性能表现进行了测试, 根据对实验结果的分析, 确定了更适合分类器使用的特征和可信度阈值。

本文的实验数据源于自建的脱机手写字符库。库中存储的字符为大写英语字母、小写英文字母和阿拉伯数字, 共 62 类, 每类训练样本 100 个, 测试样本 100 个, 其中, 分类器不作区分的字符类别标记为同一类别, 所以, 分类器可识别的字符类别数为 52。图 4 为一个实验数据样本。

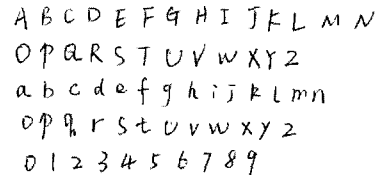


图 4 实验数据样本

### 5.1 统计特征测试

本文对分区特征<sup>[32]</sup>、轮廓方向向量<sup>[33]</sup>、骨架放射线<sup>[34]</sup>、梯度地图等 4 种常用的统计特征在宏观层的性能表现进行了比较试验。实验的评判标准有 3 个: 识别率、形近字类别数量、形近字类别与人类评判标准的相近程度。例如, 分类结果中若字符 6 与字符 b 为形近字, 则认为与人类判断标准近似层度高; 若字符 6 与字符 c 为形近字, 则认为与人类判断标准近似层度低。实验选用的分类器为最短距离分类器。

实验结果表明, 梯度地图特征识别率较高, 形成的形近字类别组数较少, 且该特征下的形近字类别与人类智能判别的形近字类别具有很高的相似度, 即对于分类器认为仅通过统计特征无法准确区分的字符, 人类智能同样认为不易区分, 需要利用更具体的结构特征才能准确区分。当选用梯度地图作为统计特征时, 宏观层的性能表现能够满足分类器的设计要求。表 4 为实验结果。

表 4 宏观层统计特征测试结果

统计特征	识别率	形近字组数	相近程度
分区特征	67.97%	42	中
轮廓方向向量	71.71%	35	中
骨架放射线	60.16%	51	中
梯度地图	84.56%	18	高

### 5.2 形近字区分测试

本文对微观层区分形近字的性能进行了测试实验。实验选用的形近字判别可信度阈值为 0.6, 结果信任可信度阈值为 0.8, 即对于  $CF_1(H)$  不大于 0.6 的字符, 都将进入微观层, 作为测试的形近字对象; 对于  $CF(H)$  大于 0.8 的字符, 都将输出有效结果, 用于判别测试结果。由于无效结果的产生与形近字判别无关, 实验仅统计进入微观层且输出有效结果的字符。

实验结果表明, 微观层的形近字区分规则能够很好地对形近字进行区分, 满足分类器设计要求。表 5 为实验结果。

表 5 形近字区分测试结果

形近字及形近字簇	识别率
$e \approx 2$	100%
$G \approx 6$	100%
$t \approx 1$	100%
$k \approx r$	100%
$J \approx T$	100%
$Y \approx T$	100%
$F \approx T$	100%
$a=Q=0$	96.71%
$g=q=y=9$	92.27%

### 5.3 可信度测试

本文对分类器在不同信任标准下的性能进行了测试。实验分别选取 0, 0.6, 0.8, 0.9, 0.95, 1.0 作为结果信任可信度阈值  $CF_{tr}$ , 测试分类器的识别正确率、形近字区分率、拒识率, 并观察  $CF_{tr}$  的变化对分类器性能的影响。

实验结果表明, 结果信任可信度阈值设置为 0.8 时, 分类器的识别正确率、形近字区分率均较高, 且拒识率较低, 分类器性能达到最佳。阈值设置太低, 识别结果不可信; 阈值设置太高, 导致较高的拒识率, 会浪费分类器的识别能力。形近字的识别率与可信度阈值的设置关系不大, 这是因为结构特征的区分能力很强。一般情况下, 只有对于书写过于潦草, 基本丧失了应有的结构特征的字符样本, 才会出现误识。可信度机制能够满足分类器的设计要求。表 6 为实验结果。

表 6 可信度测试结果

可信度阈值	识别正确率	形近字区分率	拒识率
0	89.93%	96.43%	0%
0.6	92.03%	96.43%	3.75%
0.8	95.23%	97.86%	7.88%
0.9	96.01%	99.29%	23.17%
0.95	99.62%	98.59%	51.37%
1.0	100%	100%	75.00%

**结束语** 本文模拟人类智能进行字符识别活动的过程, 设计了一种基于层次分类的脱机手写字符识别算法。利用分类器的宏观层识别结构简单的字符, 提高了识别效率; 利用分类器的微观层识别形近字, 增强了形近字的区分能力; 利用可信度量度不确定性, 提升了分类器的灵活性和可靠性。下一步将进行的工作包括: 在算法中加入结果反馈回路, 使得分类器能利用识别结果更新规则可信度; 加入形近字动态生成算法, 以适应不同的字符库; 对算法的时间与空间复杂度进行深入的理论分析和实验验证。

### 参考文献

[1] Trier Ø D, Jain A K, Taxt T. Feature extraction methods for character recognition—a survey[R]. Michigan State Univ., to appear in Pattern Recognition, Dec. 1994

[2] Arica N, Yarman-Vural F T. An overview of character recognition focused on off-line handwriting[J]. IEEE Trans. Syst., Man, Cybern. C, 2001, 31: 216-233

[3] Bimbo A D, Santini S, Sanz J. OCR from poor quality images by deformation of elastic templates[C]//Proc. of the 12th IAPR International Conference on Pattern Recognition. Jerusalem, 1994, 2: 433-435

[4] Gader P D. Recognition of handwritten digits using template and model matching[J]. Pattern Recognition, 1991, 24(5): 421-431

[5] Taxt T, Bjerde K W. Classification of Handwritten Vector Symbols Using Elliptic Fourier Descriptors[C]//Proc. 12th IAPR International Conference on Pattern Recognition. Jerusalem, Vol. 2 Oct 1994: 123-128

[6] Hu M. Visual Pattern Recognition by Moment Invariants[J]. IEEE Trans. Information Theory, 1962, 8: 179-187

[7] Kim W-Y, Kim Y-S. A region-based shape descriptor using Zernike moments[J]. Signal Processing: Image Communication, 2000, 16: 95-102

[8] Suganthan P N, Yan H. Recognition of handprinted Chinese characters by constrained graph matching[J]. Image Vision Comput, 1998, 16: 191-201

[9] Rocha J, Pavlidis T. A Shape Analysis Model with Applications to a Character Recognition System[J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 1994, 16(4): 393-404

[10] Sekita I, Toraichi K, Mori R, et al. Feature extraction of handwritten Japanese characters by spline functions for relaxation matching[J]. Pattern Recognition, 1988, 21(1): 9-17

[11] Lam L, Suen C Y. Structural classification and relaxation matching of totally unconstrained handwritten zip-code numerals[J]. Pattern Recognition, 1998, 21(1): 19-31

[12] Pavlidis T. Recognition of printed text under realistic conditions[J]. Pattern Recognition Letter, 1993: 326

[13] Shridhar M, Badreldin A. High accuracy syntactic recognition algorithm for handwritten numerals[J]. IEEE Trans. Syst., Man, Cybern., 1985, SMC-15: 152-158

[14] Wettschereck D, Dietterich T. Improving the performance of radial basis function networks by learning center locations[M]//J. E. Moody, S. J. Hanson, R. P. Lippmann, eds. Advances in Neural Information Processing Systems 4. Morgan-Kaufmann, Los Altos, CA, 1992: 1133-1140

[15] Tarassenko L, Roberts S. Supervised and unsupervised learning in radial basis function classifiers[J]. IEE Proc. -Vis. Image Signal Process, 1994, 141(4): 210-216

[16] Platt J C. Fast training of support vector machines using sequential minimal optimization[M]//B. Scholkopf, C. J. C. Burges, A. J. Smola, eds. Advances in Kernel Methods: Support Vector Learning. Cambridge, MA: MIT Press, 1999: 185-208

[17] Mangasarian O L, Musicant D R. Successive overrelaxation for support vector machines[J]. IEEE Trans. Neural Networks, 1999, 10(5): 1032-1037

[18] Plamondon R, Srihari S N. On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey. IEEE Trans[J]. Pattern Analysis and Machine Intelligence, 2000, 22(1): 63-84

[19] Rahman A, Fairhurst M. Multiple classifier decision combination strategies for character recognition: a review[J]. Int J Doc Anal Recog, 2003, 5(4): 166-194

- [20] Ho T K, Hull J J, Srihari S N. Decision combination in multiple classifier systems[J]. IEEE Transactions PAMI, 1994, 16(1): 66-75
- [21] 蔡自兴, 徐光祐. 人工智能及其应用[M]. 北京: 清华大学出版社, 2004
- [22] Fujisawa H, Liu C L. Directional pattern matching for character recognition revisited[C]//Proc. 7th Int. Conf. Document Analysis and Recognition, Edinburgh, Scotland, 2003; 794-798
- [23] Srikantan G, Lam S W, Srihari S N. Gradient-based contour encoding for character recognition[J]. Pattern Recognition, 1996, 29(7): 117-1160
- [24] Liu H, Ding X. Handwritten character recognition using gradient feature and quadratic classifier with multiple discrimination schemes[C]// Proceedings of the 8th ICDAR. Seoul, Korea, 2005; 19-23
- [25] Duda R O, Hart P E, Stork D G. Pattern Classification[M]. New York; Wiley-Interscience, 2001
- [26] Hastie T. Principal curves and surfaces[R]. Laboratory for computational Statistics Stanford University, Dept. of Statistics Technical Report, 1984
- [27] Zhang Junping, Wang Jue. An overview of principal curves[J] Chinese Journal of Computers, 2003, 26(2): 129-146
- [28] Kégl B, Krzyzak A. Piecewise linear skeletonization using principal curves[J]. IEEE Trans. Pattern Anal. Machine Intell, 2002, 24: 59-74
- [29] 苗夺谦, 张红云, 李道国, 等. 基于主曲线的脱机手写数字识别[J]. 电子学报, 2005, 33(9): 1639
- [30] 张红云, 苗夺谦, 张东星. 基于主曲线的脱机手写数字结构特征分析及选取[J]. 计算机研究与发展, 2005, 42(8): 1344
- [31] 苗夺谦, 胡桂荣. 知识约简的一种启发式算法[J]. 计算机研究与发展, 1999, 36(6): 681-684
- [32] Jin Jian-ming. Research on typeset mathematical expression image processing[D]. Nankai University, 2003
- [33] Kimura F, Shridhar M. Handwritten numerical recognition based on multiple algorithms[J]. Pattern Recognition, 1991, 24(10): 969-983
- [34] Lee Hsi-Jian, Wang Jiumn-Shin. Design of a mathematical expression understanding system[J]. Pattern Recognition Letters, 1997, 18: 289-298

(上接第 202 页)

抗反演中应用的进一步研究方向包括: 1) 本文仅通过理论模型对算法的有效性进行了验证, 但实际地震资料会受到更多干扰, 因此还需要探索算法在更为困难的实际资料反演中的应用; 2) 结合具体的实际资料反演实例, 分析问题本身的特性, 充分利用问题本身的经验以加快算法的收敛速度, 提高解的质量。

### 参 考 文 献

- [1] Yang Li-qiang, Song Hai-bin, Hao Tian-yao. Wave Impedence Inversion and Application Based on BP Neural Network[J]. Progress in Geophysics, 2005, 20(1): 34-37
- [2] Yang Wen-cai. Trends of Geophysics Inversion [J]. Earth Science Frontiers, 2002, 9(4): 389-396
- [3] Yao Yao. SA Improvement of Geophysics Nonlinear Inversion [J]. Chinese Journal of Geophysics, 1995, 38(5): 643-650
- [4] Zhang Hong-bing, Shang Zuo-ping, Tan Sheng-zhang. Fast SA for Wave Impedence Inversion[J]. Journal of Hehai university, 2005, 33(4): 434-437
- [5] Lu Peng-fei, Yang Chang-chun. Improvement of SA and Implication in Parameters Inversion[J]. 2008, 23(1): 104-109
- [6] Jiang Long-cong, Liu Jiang-ping. Improvement of SA[J]. Chinese Journal of Engineer Geophysics, 2007, 4(2): 135-140
- [7] Yang Wen-cai. Preview and Trends of Geophysics Inversion[J]. Progress in Geophysics, 2002, 17(2): 255-261
- [8] Yao Zhen-xing, Zhang Lin-bin. Mix Optimization of Wave Impedance[J]. Progress in Geophysics, 1999, 14(2): 1-6
- [9] Zhang Hai-Yan, Pan Cheng-xin. Mix Optimization Based on GA [J]. Computer Engineering and Application, 2006, 17: 56-61
- [10] Cai Han-Peng, He Zheng-hua. Wave Impedance Inversion Based on PSO[J]. Petroleum Exploration and Development, 2008, 43(5): 535-541
- [11] Zhao Xian-sheng. Wave Impedance Mix Inversion Based on GA [J]. Natural Gas Industries, 2005, 25(7): 32-34
- [12] Zhu Shi-Jun. Mix Optimization of Wave Impedance Inversion [J]. Natural Gas Industries, 2006, 26(7): 46-49
- [13] Gao Ying. Immune Particle Swarm Optimization[J]. Computer Engineering and Application, 2004, 23(6): 30-33
- [14] Chen Ying. Improved Particle Swarm Optimization Based on Immune Swarm Optimization[J]. Journal of System Simulation, 2008, 20(6): 1471-1474
- [15] Zhang Lei, lu Zhen-su. Multi-user Detection Based on Improved Adaptive Cloning Selection PSO [J]. Journal on Communications, 2007, 10(1): 37-42
- [16] Lin Cheng-jian. An efficient immune - based symbiotic particle swarm optimization learning algorithm for TSK-type neuro-fuzzy networks design[J]. Fuzzy Sets and Systems, 2008, 159(21): 2890-2909
- [17] Li Xuyuan, Xu Hualong, Cheng Zhao-gang. One Immune Simplex Particle Swarm Optimization and it's Application[C]// Fourth International Conference on Natural Computation; 331-335
- [18] Ge Hong-wei, Sun Liang, Liang Yan-chun. An Effective PSO and AIS-Based Hybrid Intelligent Algorithm for Job-Shop Scheduling[J]. IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, 2008, 38(2): 358-368
- [19] Kwok N M, Liu D K, Dissanayake G. Evolutionary computing based mobile robot localization[J]. Engineering Applications of Artificial Intelligence, 2006, 19(8): 857-868
- [20] Yin Peng-yeng, Wang Jing-yu. A particle swarm optimization approach to the nonlinear resource allocation problem[J]. Applied Mathematics and Computation, 2006, 183(1): 232-242
- [21] Du Ji-xiang, Huang De-shuang, Zhang Guo-jun, et al. A novel full structure optimization algorithm for radial basis probabilistic neural networks[J]. Short Communication Neurocomputing, 2006, 70(1): 592-596
- [22] Gao L, Gao H B. Particle swarm optimization based algorithm for cutting parameters selection [C] // Proceedings of IEEE World Congress on Intelligent Control and Automation, Hangzhou, 2004, 1(4): 2847-2851
- [23] de Castro L N, Timmis J. An artificial immune network for multimodal function optimization[A] // 2002 Congress on Evolutionary computation[C]. Honolulu, Hawaii, USA, 2002; 699-704