

# 基于 Gabor 变换的高鲁棒汉字识别新方法

王学文,丁晓青,刘长松

(清华大学电子工程系智能技术与系统国家重点实验室,北京 100084)

**摘要:** 本文提出了针对字符图像的基于 Gabor 变换的汉字识别新方法.在对 Gabor 变换深入分析的基础上,本文针对汉字图像的统计信息,提出了一种有效的 Gabor 滤波器组参数优化方法;同时,对 Gabor 滤波器组的输出进行非线性变换,使其适应不同亮度和低质量灰度字符图像的识别.本文还改进了分块特征的抽取算法,提高了对字符细节的分辨能力.实验表明,这种特征抽取方法大大加强了识别系统抵御图像噪声、干扰、亮度变化、笔画模糊、笔画断裂以及字符形变的能力,在应用于各种低质量的二值或者灰度的印刷和脱机手写字字符图像识别时,能获得较其他算法更良好的识别性能.

**关键词:** Gabor 滤波器; 字符识别

**中图分类号:** TP391.43 **文献标识码:** A **文章编号:** 0372-2112 (2002) 09-01317-06

## Gabor Filters Based Feature Extraction for Robust Chinese Character Recognition

WANG Xue-wen, DING Xiao-qing, LIU Chang-song

(State Key Laboratory of Intelligent Technology and Systems, Dept. of Electronic Engineering, Tsinghua University, Beijing 100084, China)

**Abstract:** This paper proposed a new feature extraction method for Chinese character recognition by using Gabor filters. Based on the theory of Gabor filters and the statistical information of Chinese character images, an effective method to design Gabor filters was developed. Moreover, to improve the performances for low quality images, non-linear functions were designed to regulate the outputs of Gabor filters adaptively. This paper also improved the feature extraction method to enhance the discriminability of histogram features. Experiments show that our method performs excellently for images with noises, backgrounds or stroke distortions and can be applied to printed or handwritten character recognition tasks in low quality greyscale or binary images.

**Key words:** character recognition, Gabor filter

### 1 引言

汉字识别的研究近年来取得重要进展,但适应于低质量灰度图像的高鲁棒性汉字识别问题并未获得满意的解决方法,目前仍然是重要的研究目标.已有的研究集中在下面几个方向上:(1)抽取结构特征,例如地形特征<sup>[3]</sup>,梯度边缘特征等;(2)自适应二值化处理<sup>[6-8]</sup>;(3)使用 FFT, DCT<sup>[2]</sup>等简单的正交变换方法.这些方法或者抗干扰能力很差,或者不能提取局部的笔画方向、结构信息,所以实用效果不佳.

Gabor 变换在分析数字图像中局部区域的频率和方向信息方面具有优异的性能,在计算机视觉和纹理分析等领域已经得到了广泛的应用.在 90 年代中期,Hamamoto<sup>[9]</sup>, Da Deng<sup>[10]</sup>等对脱机手写字字符识别中基于 Gabor 滤波器的特征抽取方法进行了研究.上述研究在抗噪声方面取得了一定的效果,但是仍然存在几个很重大的缺陷:(1)设计 Gabor 滤波器组时,根据识别率进行参数的筛选,过程烦琐,计算量很大,而且不能保证获得最好的参数;(2)这些方法都是基于二值图像的,对将 Gabor 滤波器应用于低质量灰度字符图像识别的问

题没有进行讨论.同已有的特征抽取方法相比,这些方法在识别性能上并不具有明显的优势.

本文利用 Gabor 滤波器抽取字符图像局部空间中最重要和稳定的笔画方向信息,根据字符图像的统计信息,解决了 Gabor 滤波器组参数的优化设计问题,保证了较优的识别性能;同时使用修正的 Sigmoid 函数进行非线性后处理,使笔画方向信息对图像亮度变化,笔画模糊、断裂,背景干扰都具有很强的抑制能力,大大增强了识别特征的鲁棒性;在分块特征提取时,分别对 Gabor 滤波器组实部输出中的正值和负值计算特征,提高了对细节的分辨能力.

在第 2 节和第 3 节中,分别对 Gabor 滤波器进行一定的理论分析并给出优化设计方法,第 4 节和第 5 节说明特征的提取算法,第 6 节给出对比实验及结果,第 7 节对全文作出总结.

### 2 Gabor 变换的理论分析<sup>[1]</sup>

二维 Gabor 滤波器的冲激响应函数一般取如下的形式:

收稿日期:2002-03-20;修回日期:2002-07-03

基金项目:国家 863 高技术计划(No.2001AA114081);国家自然科学基金(No.69972024);专利(No.02117865.8)

$$h(x, y, \lambda, \phi, \sigma_x, \sigma_y) = \exp\left\{-\frac{1}{2}\left[\frac{R_1^2}{\sigma_x^2} + \frac{R_2^2}{\sigma_y^2}\right]\right\} \cdot \exp\left[i \cdot \frac{2\pi R_1}{\lambda}\right],$$

$$\begin{aligned} R_1 &= x \cdot \cos\phi + y \sin\phi, \\ R_2 &= -x \cdot \sin\phi + y \cdot \cos\phi \end{aligned} \quad (1)$$

Gabor 滤波器具有很优秀的时频聚集性,这里,重点考察其空间局域性、空间抽样间隔以及频率选择性(线条宽度和方向的选择性)<sup>[1]</sup>.

1 空间局域性及空间抽样间隔:

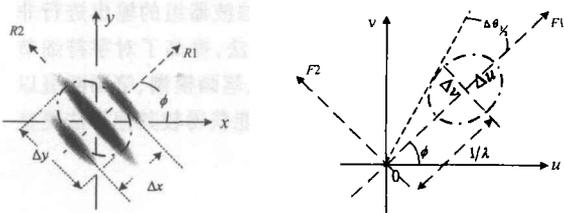


图 1 二维 Gabor 滤波器\*在(a)空域上(b)频域上(对称于原点的另一部分没有画出)的俯视图

如图 1(a)所示,设 Gabor 滤波器\*在空域中的标准有效带宽为 Δx 和 Δy(分别对应方向 R<sub>1</sub> 和 R<sub>2</sub> 上),则:

$$(\Delta x)^2 = \frac{\int_{-\infty}^{+\infty} h \cdot h^* \cdot (R_1)^2 \cdot d(R_1)}{\int_{-\infty}^{+\infty} h \cdot h^* \cdot d(R_1)},$$

$$(\Delta y)^2 = \frac{\int_{-\infty}^{+\infty} h \cdot h^* \cdot (R_2)^2 \cdot d(R_2)}{\int_{-\infty}^{+\infty} h \cdot h^* \cdot d(R_2)},$$

这里,  $h(x, y; \lambda, \phi, \sigma_x, \sigma_y)$  简记为  $h$ , 其共扼为  $h^*$ . 计算可得:

$$\Delta x = \sigma_x / \sqrt{2}, \Delta y = \sigma_y / \sqrt{2} \quad (3)$$

Δx 与 Δy 是 Gabor 滤波器在两个对称轴方向上的最小分析尺

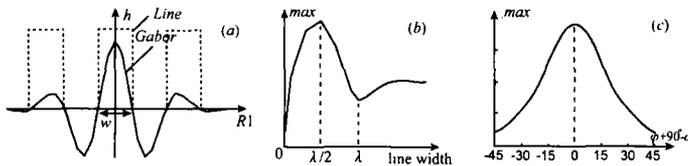


图 3 Gabor 滤波器\*对线条宽度、方向的选择性示意图.

(a) Gabor 滤波器的冲激响应函数在 R<sub>1</sub> 轴上的侧视图.

(b) 对只包含单个线条(方向为  $\alpha = \phi + 90^\circ$ )的图像,当线条宽度 W 变化时, Gabor 滤波器的输出变化;

(c) 宽度为  $\lambda/2$  的线条其方向  $\alpha$  变化时, Gabor 滤波器的输出响应\*\*.

而字符图像,尤其是汉字图像,具有下列突出的特点(如图 4 所示):

- (1) 大小规一化后的字符图像,其笔画的宽度基本一致;对于相同的字符,其笔画位置和间隔也相对稳定;
- (2) 字符笔画方向集中分布在几个特定的方向附近;
- (3) 笔画的模糊、断裂以及噪声干扰在空间尺度上一般小于字符笔画;背景亮度变化的空间尺度一般大于字符笔画. 字符笔画与这些干扰在频域上具有可分性.

这样,由于字符图像在空间上具有局域性和方向性;在频域上,字符笔画与干扰具有可分性,所以可以利用 Gabor 滤波

度.因此,可以按照一定的空间间隔  $D_x, D_y$  对图像进行抽样,而不损失有用的信息.从图 2 可见,空间抽样间隔  $D_x, D_y$  对特征提取性能有决定性的影响.但是,这个问题在过去的研究中<sup>[9,10]</sup>没有得到足够的重视,空间上的欠抽样严重降低了 Gabor 滤波器的性能.

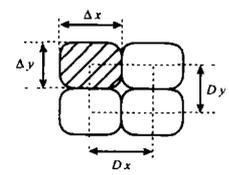


图 2, 空间抽样间隔

2 频率选择性

同样可以定义 Gabor 滤波器在频域中的标准有效带宽为 Δu 和 Δv:

$$\Delta u = 1/(2\sqrt{2}\pi\sigma_x), \Delta v = 1/(2\sqrt{2}\pi\sigma_y), \quad (4)$$

Gabor 滤波器的频率选择性在图像中表现为对线条宽度和方向的选择.

在图 3(a)中,我们对图 1(a)所示的 Gabor 滤波器在 R<sub>1</sub> 方向投影,显然,对于图像中的线条信号,只有当它们的方向与 Gabor 滤波器的方向相差 90°,且周期为 λ,线条的宽度相应的为 λ/2 时, Gabor 滤波器给出最大输出\*\*.

如图 3(b)和(c)所示的那样, Gabor 滤波器对于线条的宽度和方向具有很强的选择性.如图 1(b)所示,我们可以定义方向带宽 Δθ<sub>1/2</sub>:

$$\Delta\theta_{1/2} \approx \arcsin\left(\frac{\Delta v/2}{1/\lambda}\right) = \arcsin\left(\frac{\lambda}{4\sqrt{2}\pi\sigma_y}\right) \quad (5)$$

所以,滤波器能够分析方向约在  $(90^\circ + \phi - \Delta\theta_{1/2}, 90^\circ + \phi + \Delta\theta_{1/2})$  范围内的线条信息.

从上分析可见, Gabor 滤波器具有空间局域性和方向选择性的特点;与梯度或梯度过零点的笔划边缘提取方法相比,具有对线条宽度和方向的选择性,抗干扰能力强.

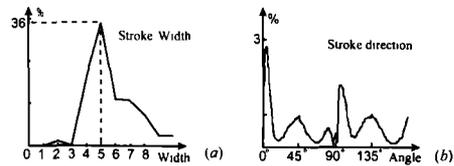


图 4 对手写汉字(64pixel × 64pixel)的笔画宽度(a)以及方向(b)分布的统计结果

器组进行特征抽取,以解决好劣质字符图像的识别问题.

3 适用于汉字识别的 Gabor 滤波器组设计及其实验验证

对于大小规一化的字符图像,可以获得字符图像的下列统计信息:(1)字符笔画的平均宽度 W;(2)字符笔画的最小分辨率 S;(3)笔画方向的分布  $\{\theta_k\}_{k=1,2..K}$ . 如果 Gabor 滤波器组

\* 这里只考察 Gabor 滤波器的实部

\*\* 这里用 Gabor 滤波器实部输出中的最大值代表 Gabor 滤波器的输出

的参数与字符图像的上述统计信息相匹配,将能获得字符图像稳定特征的最佳提取.

### 3.1 根据字符图像的统计信息确定 Gabor 滤波器组的初始参数

(1)空间局域性:Gabor 滤波器组应该具有分辨字符图像中最靠近的笔画的能力,即:

$$\Delta x \leq S, \Delta y \leq S, \quad (6)$$

因为各个方向的笔画分布相近,所以可以设定  $\Delta x = \Delta y$ . 代入式 3, 即有

$$\sigma_x = \sigma_y = \sigma \leq \sqrt{2} \cdot S \quad (7)$$

为了避免图像信息损失, Gabor 滤波器组在空间上的抽样间隔  $D_x, D_y$  也应该满足:

$$D = D_x = D_y \leq \Delta x = \Delta y = \sigma / \sqrt{2}, \quad (8)$$

(2)线条宽度选择性:Gabor 滤波器组对宽度为  $\lambda/2$  的线条最敏感, 所以有:

$$W = \lambda/2, \text{ 即 } \lambda = 2W \quad (9)$$

(3)线条方向选择性:参数  $\varphi$  与滤波器所能抽取线条的方向有  $90^\circ$  的相差, 所以有:

$$\{\varphi_k\}_{k=1,2,3,K} = \{\theta_k - 90^\circ\}_{k=1,2,K}, \quad (10)$$

同时,还应该保证每个 Gabor 滤波器在方向上互不重叠, 则方向带宽应该有:

$$2 \cdot (2 \cdot \Delta\theta_{1/2}) \leq 360^\circ / K, \text{ 即 } \Delta\theta_{1/2} \leq 90^\circ / K \quad (11)$$

代入式 5, 可得:

$$\sigma \geq \frac{\lambda}{4 \cdot \sqrt{2} \cdot \sin\beta} = a, \quad (12)$$

再考虑式 7, 则有:

$$a \leq \sigma \leq \sqrt{2} \cdot S = b \quad (13)$$

### 3.2 使用最小平均熵相关系数 ECC 准则确定 Gabor 滤波器组 $\sigma$ 的取值

ECC (Entropy Correlation Coefficient)<sup>[11]</sup> 是一种使用互信息来刻画两幅图像匹配相似度的方法, 其仅依赖于两幅图像对应像素点的亮度分布的统计量, 与具体的亮度差别无关, 所以对于图像匹配具有优越的性能. 对于两幅图像  $A$  和  $B$ , 其熵分别为  $H(A), H(B)$ , 互信息为  $I(A, B) = H(A) + H(B) - H(A, B)$ , 则

$$ECC(A, B) = 2 \cdot I(A, B) / (H(A) + H(B)) \quad (14)$$

本文利用 Gabor 滤波器组将字符图像分解为多个方向上不相关的子图像, 在  $\lambda, D, \{\varphi_k\}$  已经确定时, 在上述  $\sigma$  的取值范围内, 一定存在一个合适的  $\sigma^*$ , 使得滤波器组输出  $\{F_k\}_{k=1,2,3,4}$  之间的平均熵相关系数最小或者小于一个给定的阈值:

$$\overline{ECC} = \frac{1}{C_k^2} \sum_{i=1}^k \sum_{j=1, j \neq i}^k ECC(F_i, F_j) \quad (15)$$

我们对脱机手写汉字样本的  $\overline{ECC}$  随  $\sigma$  的变化进行了统计, 图 5 给出了平均熵相关系数  $\overline{ECC}$  随  $\sigma$  的变化曲线. 可以看出, 在式 13 中给定的范围内, 存在一个  $\sigma^*$ , 当  $\sigma = \sigma^*$  时,  $\overline{ECC}$  取极小值; 而且在  $\sigma \geq \sigma^*$  时,  $\overline{ECC}$  的变化非常小. 所以, 由此可以确定最优的 Gabor 滤波器组参数  $\sigma$  和空间抽样间隔  $D$ .

### 3.3 Gabor 滤波器组参数选择的实验验证

对于规一化的汉字图像, 通过上述方法可以求得最优的 Gabor 滤波器组参数为  $\lambda = 10, K = 4, \{\varphi_k\}_{k=1,2,3,4} = \{-90^\circ, -45^\circ, 0^\circ, 45^\circ\}, \sigma = 5.6, D \leq 4$ . 下面的实验可验证其性能.

(1)印刷汉字样本的识别实验(实验条件见 6.1 节): 字符集仅取国标一级汉字的前 1000 字. 图 7 和图 8 给出了测试识别率随  $\lambda$  和  $\sigma$  的变化曲线(选定  $D = 1$ , 分别限定  $\sigma = 5.6, \lambda = 10$ ), 从中可以发现, 在选定参数附近, 识别系统稳定在一个最高的识别率上.

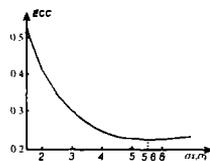


图 5 平均的熵相关系数随  $\sigma$  的变化曲线

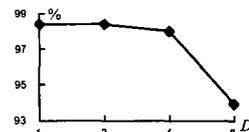


图 6 识别率随空间抽样间隔的变化曲线

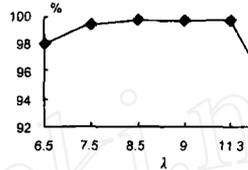


图 7 测试识别率随  $\lambda$  的变化曲线

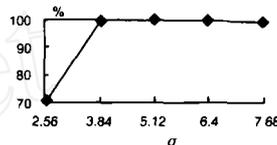


图 8 测试识别率随  $\sigma$  的变化曲线

(2)空间抽样间隔  $D$  对识别系统的性能影响: 使用 500 套书写较规范的脱机手写汉字样本(具体条件见 6.3 节)进行了识别实验. 测试识别率随  $D$  的变化曲线如图 6 所示, 在  $D \leq 4$  时, 识别率稳定在最优值附近, 而当  $D > 4$  后, 识别率剧烈下降. 上述实验结果以及 6.3 节中对 Hamamoto 方法<sup>[9]</sup> 的对比实验证明了本文的分析, 即根据字符图像统计信息来设计 Gabor 滤波器组参数, 可以获得优异的识别性能.

如图 9 所示, 当上述参数选定后, Gabor 滤波器组的实部输出已经准确的刻画了字符图像中笔画的位置和方向信息. 所以在本文的讨论中, 仅取 Gabor 滤波器组的实部输出来计算识别特征.

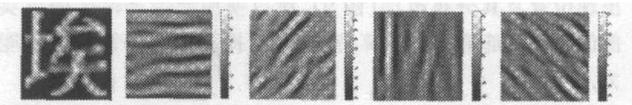


图 9 灰度字符图像经 Gabor 滤波器组分析后的结果. 最左边是低分辨率字符图像插值的结果, 随后是 Gabor 滤波器组实部输出的结果( $\varphi$  依次取  $-90^\circ, -45^\circ, 0^\circ, 45^\circ$ ).

## 4 对 Gabor 滤波器组输出的非线性变换

因为不同的光照环境或者图像摄取过程, 图像的亮度、对比度都是不同的, 即使在同一字符图像中, 各笔画也可能因不均匀光照而亮度不同; 而且, 图像中的干扰仍然会产生一定的干扰输出. 必须通过对滤波器组输出的后处理进一步抑制亮度变化和干扰输出.

在 Jain 的研究<sup>[12]</sup> 中, 使用了 Sigmoid 函数  $\Phi(t)$  (式 16) 来对 Gabor 滤波器组输出的幅值进行非线性变换. Sigmoid 函数

(图 10) 对大输入的抑制作用可以补偿图像各部分亮度的差异, 但对小输入的放大作用会造成问题: (1) 放大干扰输出, 破坏 Gabor 滤波器对线条宽度的选择性; (2) 增强由于 Gabor 滤波器的冲激响应函数的旁瓣造成的“伪”条状输出。

为了解决上述缺点, 本文对 Sigmoid 函数进行了修正  $\theta(t)$  (式 17), 维持在大输入时的饱和特性, 而在小输入时表现出抑制特性。下面以对滤波器输出中的正值的处理为例, 对这种自适应变换进行分析; 对于负值, 可以取  $\theta(t)$  的奇对称形式。  $\Phi(t)$  和  $\theta(t)$  定义为:

$$\phi(t) = \tanh(\alpha t) = (e^{2\alpha t} - 1) / (e^{2\alpha t} + 1), \quad (16)$$

$$\theta(t) = \tanh(\alpha(t - \chi)) + \beta \quad (17)$$

式 17 中,  $t$  是归一化后的输入, 其范围为  $[0, 1]$ , 当输入  $t < \nu$  时  $\theta(t)$  的输出被抑制, 则可以限定:

$$\left. \frac{d\theta}{dt} \right|_{t=\nu} = 1, \nu < \chi, \text{ 且 } \theta(0) = 0, \quad (18)$$

因为获得背景干扰的分布统计非常困难, 所以这里以消除滤波器组输出中的“伪”条状物来进行讨论, 得到的结果对降低背景干扰同样有效。式 18 中,  $\nu$  ( $\nu < \chi$ ) 必须对区分背景

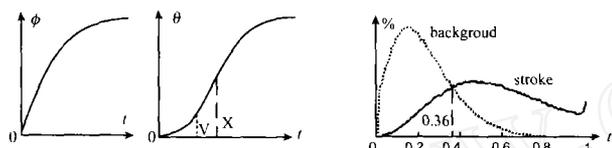


图 10 Sigmoid 函数及修正后函数图示

图 11 在字符图像的背景和笔画位置上, Gabor 滤波器组输出中局部极大值的分布

## 5 分块特征的抽取

我们对传统的分块特征抽取方法<sup>[4,9]</sup>进行了改进, 以提高对字符细节的分辨能力。

如图 3(a) 所示, 在 Gabor 滤波器冲激响应函数的中心两旁存在两个较大的负旁瓣, 这样在笔画位置上输出规则的负-正-负“三条纹”分布, 这种分布比较准确的描述了笔画的出现位置及其分布范围(图 9), 所以使用滤波器输出的正实部值和负实部值分别计算特征, 可以提高识别特征的分辨能力。

经过自适应变换后的 Gabor 滤波器组的输出, 可以按照图 13 所示的方式, 将 4 个  $N \times N$  的输出平面分割为  $M \times M$  个互相交叠的区域(边长为  $L$ ), 在每个区域  $r(x, y)$  中, 分别求实部输出的正值和负值的高斯加权和  $Sum^+(x, y)$ 、 $Sum^-(x, y)$ 。这里简记方向为  $\varphi_k$  的 Gabor 滤波器的输出为  $F_k(x, y)$ :

$$Sum^+(x, y) = \sum_{(m,n) \in r(x,y)} G(m-x, n-y) \times \max(0, F_k(m, n)), \quad (20)$$

$$Sum^-(x, y) = \sum_{(m,n) \in r(x,y)} G(m-x, n-y) \times \min(0, F_k(m, n)), \quad (21)$$

这里,  $G(x, y) = \exp\{- (x^2 + y^2) / (2\tau^2)\} / (2\pi\tau^2)$ , 其中  $\tau^2$  是高斯抽样函数  $G(x, y)$  的方差, 它控制抽样函数在空间上的

和笔画位置上 Gabor 滤波器组的输出具有最佳的效果, 这样才能实现背景(背景干扰)的抑制。在图 11 中, 我们对 Gabor 滤波器组输出中的局部极大值(分布在条状输出的中线上)进行了统计, 可见, 在  $t = 0.36$  时, 区分背景和笔画具有最小的错误率。所以, 取  $\nu = 0.36$ 。不同的  $\alpha$  值给出不同的  $\theta(t)$  函数, 对 Gabor 滤波器的线条宽度选择性(图 3(b))有不同的影响。而为了实现干扰抑制, 这种影响应尽可能小。本文定义了一个指标  $error(\alpha)$  来刻画这种影响。设原始的线条宽度选择性曲线为  $S(w)$ , 变换后的选择性曲线为  $S_\alpha(w)$ , 则:

$$error(\alpha) = \sum_{w=0}^{\lambda} [S(w) - S_\alpha(w)]^2 \quad (19)$$

图 12(a) 中显示, 在  $[4, 10]$  区间内,  $\alpha = 7$  ( $\nu = 0.36$ , 再代入式 18 中可以确定  $\chi = 0.59, \beta = 1.00$ ) 时,  $error(\alpha)$  取最小值。图 12(b) 和 (c) 分别给出了此时的  $\theta(t)$  函数和 Gabor 滤波器对线条宽度的选择性响应曲线。

上述的非线性变换, 对于高质量的图像, 对识别率的影响很小; 而对于质量较差的字符图像, 却明显抑制了干扰或者亮度变化的影响, 可以提高系统的识别性能。

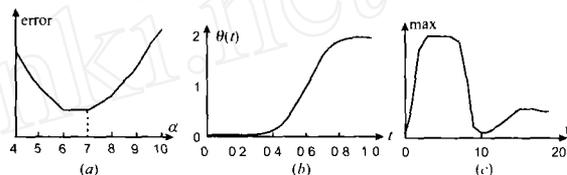


图 12 (a)  $error(\alpha)$  随参数  $\alpha$  的变化; (b)  $\alpha = 7$  时的  $\theta(t)$  函数; (c)  $\alpha = 7$  时的 Gabor 滤波器的线条宽度选择性响应曲线。

作用范围, 一般可以取  $\tau = L/2$ 。

这样, 对于每个方向的 Gabor 滤波器输出, 可得到  $2M^2$  维特征矢量,  $M$  一般可以取为 8。将四个方向上的特征矢量合并成一个维数为  $8M^2$  的矢量, 即特征矢量  $V$ 。特征矢量  $V$  可以直接输入统计分类器进行分类, 也可以对其使用线性鉴别分析或者主分量分析等方法进行特征压缩<sup>[5]</sup>, 使其在统计上更符合高斯分布的假设, 提高分类器的性能。

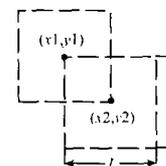


图 13 分块示意图

## 6 实验及其结果

我们分别对不同劣化条件下的印刷汉字图像样本、脱机手写汉字样本以及实际的身份证样本进行了识别实验, 实验中使用了下面几种识别特征:

- (1) 本文的方法得到的 Gabor 特征, 简记为 GAB;
- (2) GAB<sup>1</sup>: 取 Gabor 滤波器组输出实部的绝对值计算特征, 其它同 GAB;
- (3) GAB<sup>2</sup>: 取 Gabor 滤波器组输出的能量计算特征;
- (4) GAB<sup>3</sup>: 分别使用 Gabor 滤波器组输出的实部和虚部的正、负值计算特征;
- (5) DCT 特征<sup>[2]</sup>: 基于灰度字符图像, 简记为 DCT;
- (6) Hamamoto 采用的 Gabor 特征<sup>[9]</sup>: 基于二值图像, 简记

为 HGAB;

(7)四方向线素特征<sup>[4]</sup>:广泛使用的基于二值图像的特征,简记为 DH.

在实验中,都使用文献<sup>[4]</sup>中给出的改进的马氏距离分类器进行训练和识别.

表 1 不同分辨率的字符图像的识别率(%)

		GAB	DCT	DH
宋体 (仿宋) (黑体)	75 - 100dpi	99.28	98.96	97.71
	150 - 200dpi	99.87	99.87	99.59
	250 - 300dpi	99.82	99.91	99.82
楷体	75 - 100dpi	96.71	86.61	89.47
	250 - 300dpi	99.40	98.31	99.75
隶书	75 - 100dpi	74.41	55.87	47.45
	150 - 200dpi	95.60	85.37	86.55
	250 - 300dpi	98.18	93.74	95.74

表 2 加噪图像的识别率(%)

噪声参数	原图	高斯噪声(均值=0)				椒盐噪声				散斑噪声(均值=0)			
		2.5	12.5	25	38.3	5%	10%	15%	20%	2.5	12.5	25	38.3
GAB	99.44	98.92	98.38	96.56	92.53	99.26	98.81	97.93	95.07	99.07	98.69	98.21	97.17
DCT	99.33	98.00	95.54	86.11	71.65	99.09	98.57	96.53	89.99	98.14	96.80	91.11	84.05
DH	99.55	92.05	72.90	30.85	8.28	99.32	98.33	89.04	43.16	94.59	84.65	55.91	31.39

实验表明,GAB 特征在识别被噪声劣化的字符图像时,具有极佳的鲁棒性.

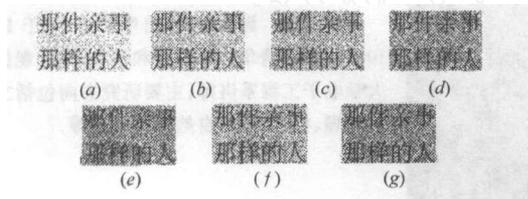


图 14 加噪图像示例。(a)原始图像;(b)~(e)高斯噪声;(b)  $N(0,2.5)$ ;(c)  $N(0,12.5)$ ;(d)  $N(0,25)$ ;(e)  $N(0,38.3)$ ;(f) 比率为 20% 的椒盐噪声;(g)均值为 0,方差为 38.3 的散斑噪声.

### 6.3 识别脱机手写汉字的实验

对 500 套书写较规范的脱机手写汉字样本进行识别实验.从其中随机选出 370 套作为训练样本,剩下的 130 套作为测试样本.使用上述特征进行识别的结果如表 3 所示.

可以看出:(1)分别使用 Gabor 滤波器实部输出的正值和负值计算特征,具有更好的识别性能;(2)在识别脱机手写字符时,GAB 特征的识别性能与传统方法中性能最好的四方向线素特征(DH)相当,大大优于 Hamamoto 提出的 Gabor 特征(HGAB)以及 DCT 特征.这说明,GAB 特征具有优秀的抗字符变形的性能.

表 3 对脱机手写汉字测试样本的识别率(%)

	首位	前十位
GAB	98.50	99.85
GAB <sup>1</sup>	97.63	99.58
GAB <sup>2</sup>	96.53	99.51
GAB <sup>3</sup>	98.30	99.80
DH	98.39	99.79
HGAB	95.27	99.43
DCT	94.89	98.02

### 6.4 识别身份证的实验

如图 16 所示,实验中采集了 41 张各种质量的印刷身

### 6.1 识别低分辨率字符图像的实验

训练样本是 200 套包含国标一级汉字的常用字体,如宋、仿、黑、楷、隶及圆等样本.测试样本包含了 582 套宋、仿、黑、楷和隶书样本,图像是以 75 - 300dpi,不同的亮度、对比度分别扫描得到的.在以 75dpi 扫描图像中,汉字字符的大小约为 15pixel × 15pixel.

从表 1 可以看出,在低分辨率下,GAB 特征都取得了最高的识别率.这说明,GAB 特征在抵抗低分辨率扫描产生的笔画模糊、断裂方面,具有非常突出的性能.

### 6.2 识别加噪字符图像的实验

对包含 11,688 个汉字的印刷文本,以 300dpi 扫描为灰度图像,并加入不同的噪声:(1)形式为  $N(0, \sigma^2)$  的加性高斯噪声;(2)加性椒盐噪声;(3)均值为 0 的乘性散斑噪声.图 14 给出了加噪图像的例子.这里使用 6.1 节中产生的识别库进行识别,结果显示在表 2 中.

份证图像,采用 GAB 特征对其中的姓名区域和证号区域进行了识别,识别库中包含姓名中常见的 4399 个汉字以及 10 个数字.实验中,如果一张身份证的姓名或者证号识别错,则整张身份证识别错误;使用两种方法:(1)对 Gabor 滤波器的输出进行简单归一化处理;(2)采用修正的 Sigmoid 函数  $\theta(t)$  进行处理.识别结果表明,采用  $\theta(t)$  进行处理,可以降低识别错误数.



图 15 脱机手写汉字样本示例(汉字“埃”的不同样本)

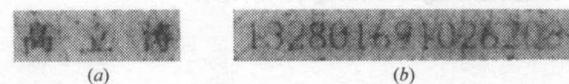


图 16 典型的身份证图像的姓名(a)和证号(b)区域.

从上面的一系列实验可以得到这样的结论,本文的方法能同时对二值和灰度字符图像进行识别,既具有优秀的抗噪声、干扰以及其它图像劣化的能力,又具有优秀的抗字符变形的能力,这些性能是已有的特征抽取方法<sup>[2,3,4,7,9,10]</sup>不可能同时具有的.

表 4 身份证识别结果

	证	汉字	数字
总数	41	104	633
识别规一化	6	2	4
错误 $\theta(t)$	3	3	0

## 7 结论

理论分析和实验结果表明,本文提出的基于 Gabor 变换的识别特征充分反映了笔画结构在空间上的局域性,笔画的方向性以及频域上笔画与干扰的可分性等重要的特性,从而使抽取得到的分块特征在抵抗图像噪声、干扰、图像亮度变

化以及字符形变等诸方面都具有优秀的性能,较好解决了实际应用中各种低质量字符图像的识别问题,具有极大的通用性。

#### 参考文献:

- [1] J G Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters [J]. J Opt Soc Am A, 1985, 2(7): 1160 - 1169.
- [2] Wang X W, et al. A gray-scale image based character recognition algorithm to low quality and low resolution Images [A]. Document Recognition and Retrieval VIII, Electronic Imaging 2001 [C]. San Jose: IS&T/SPIE, 2001.
- [3] L Wang, et al. Pavlidis. Direct gray-scale extraction of features for character recognition [J]. IEEE Trans on PAMI, 1993, 15(10): 1053 - 1066.
- [4] Zhang J Y, et al. Multi-scale feature extraction and nested-subset classifier design for high accuracy handwritten character recognition [A]. Proc ICPR' 2000 [C]. Barcelona: IAPR, 2000.
- [5] J Mao, et al. Artificial neural networks for feature extraction and multivariate data projection [J]. IEEE Trans on Neural Networks, 1995, 6(2): 296 - 317.
- [6] O D Trier, et al. Goal-directed evaluation of binarization methods [J]. IEEE Trans on PAMI, 1995, 17(12): 1191 - 1201.
- [7] H Kamada, et al. High speed, high-accuracy binarization method for recognizing text in images of low spatial resolutions [A]. Proc ICDAR' 99 [C]. Bangalore: IAPR, 1999. 139 - 142.
- [8] W Niblack. An Introduction to Digital Image Processing [M]. New Jersey: Prentice Hall, 1986. 115 - 116.
- [9] Y Hamamoto, et al. A Gabor filter-based method for recognizing handwritten numbers [J]. Pattern Recognition, 1998, 31(4): 395 - 400.
- [10] D Deng, et al. Handwritten Chinese character recognition using spatial Gabor filters and self-organizing feature maps [A]. Image Processing, Proceedings. ICIP-94, IEEE International Conference, Vol. 3 [C]. Austin: IEEE Signal Processing Society, 1994.
- [11] J P W Pluim, et al. Mutual information matching in multiresolution contexts [J]. Image and Vision Computing, 2001, 19: 45 - 52.
- [12] A K Jain, et al. Unsupervised texture segmentation using Gabor filters [J]. Pattern Recognition, 1991, 24(12): 1167 - 1186.

#### 作者简介:



王学文 男, 1973 年出生于湖北, 1992 年进入清华大学电子工程系无线电技术与信息系统专业学习, 1997 年获工学学士学位。现在清华大学电子工程系信号与信息处理系统专业攻读博士。研究方向为图像处理、模式识别等。E-mail: wxw@ocrserv. ee. tsinghua. edu. cn Tel: 62772369



丁晓青 女, 生于 1939 年 7 月, 贵州人。1962 年毕业于清华大学无线电电子学系。现为清华大学电子工程系教授、博士生导师, 智能图文信息处理责任教授, 电子学会高级会员, 中国通信学会会士。主要研究方向为模式识别、图像处理、汉字识别、生物特征身份鉴别、多媒体信息处理以及视频智能监测等。获国家科技进步奖多次, 发表论文 180 余篇, 合作专著 2 本。



刘长松 男, 1969 年出生于辽宁, 于 1992 和 1995 年获得清华大学学士和硕士学位, 现任清华大学电子工程系讲师。主要研究方向包括文本图像处理、模式识别、自然语言处理等。