

文章编号:1001-9081(2006)11-2778-03

## 利用高斯描绘子进行字符识别

刘亦书

(华南师范大学 地理科学学院, 广东 广州 510631)

(yishuliu@21cn.com)

**摘要:**高斯描绘子是一种基于边缘的形状特征,具有识别/匹配率高、相对于平移、旋转、尺度和反射不变、计算量小、对适度的边缘变动和噪声不敏感以及适用范围广等优点。将高斯描绘子用于字符识别,并与另一种基于边缘的特征轮廓矩不变量 Hu 矩的推广和改进进行比较。实验结果表明,高斯描绘子有很好的识别效果。

**关键词:**字符识别;高斯描绘子;Hu 矩;轮廓矩不变量

**中图分类号:** TP391.41 **文献标识码:** A

## Character recognition by Gaussian descriptors

LIU Yi-shu

(School of Geography, South China Normal University, Guangzhou Guangdong 510631, China)

**Abstract:** Gaussian descriptors are contour-based shape features. They are invariant to translation, scaling changes, rotation and reflection. Compared to the existing shape features, they are more robust against noise and slight edge variations, and have lower computation complexity and higher recognition/retrieval rate. In addition, they are application-independent. In this paper, Gaussian descriptors were used as features for character recognition. A comparison with another contour-based moment invariants, which is an improvement and extension of classical Hu moments, was also given. Numerical experimental results show that Gaussian descriptors are an attractive tool for character recognition.

**Key words:** character recognition; Gaussian descriptors; Hu's moments; contour-based moment invariants

### 0 引言

光学字符识别(Optical Character Recognition, OCR)是模式识别领域一个重要的课题,而选择有效的特征提取方法是提高 OCR 系统性能最重要的因素。用于光学字符识别的特征提取方法主要有:1)基于边缘的技术,如傅里叶描绘子<sup>[1]</sup>、样条曲线逼近<sup>[2]</sup>、分区方法<sup>[3]</sup>和边缘剖面<sup>[4]</sup>;2)基于区域的技术,如 Hu 矩<sup>[5,6]</sup>、Zernike 矩<sup>[7]</sup>、投影直方图<sup>[8]</sup>和单元变换<sup>[9]</sup>。用上述方法提取出来的特征,大多具有平移、旋转、尺度和反射不变性,这对于提高识别能力是很重要的。

这些特征中,最经典、最常用的是基于边缘的傅里叶描绘子和基于区域的 Hu 矩。傅里叶描绘子要求边缘是封闭的,这对很多边缘提取方法来说都是很苛刻的要求;而且,傅里叶描绘子只使用物体的外边缘,这就大大地削弱了它的识别能力,比如,无法区分“国”和“圆”。至于 Hu 矩,由于是基于区域的,和图像的灰度值密切相关,因而计算量很大。在图像识别的许多问题中,物体的本质特征不是灰度,而是形状,所以, Hu 矩有时并不能有效和正确地反映物体的形状特征。基于这样的思考,文献[10]中对 Hu 矩进行改进,构造出基于边缘的轮廓矩不变量,并将其用于字符识别和其他物体的识别,获得良好的实验效果。相对于 Hu 矩,轮廓矩不变量的计算量大为减少,更重要的是,它们仅与物体的形状(轮廓)有关,而与图像的颜色、对比度和光照度无关,而且它们同样具有平移、旋转、尺度和反射不变性。

作者最近还提出另一种新的基于边缘的特征——高斯描

绘子,并从理论和实验结果上证明它具有识别/匹配率高、相对于平移、旋转、尺度和反射不变、计算量小、对适度的边缘变动(如边缘断裂、粘连、缺失等)和噪声不敏感以及适用范围广等优点。

本文将高斯描绘子用于字符识别,并与轮廓矩不变量的实验结果作比较。

### 1 高斯描绘子

设曲线  $C$  由曲线  $C_1, C_2, \dots, C_n (n \geq 1)$  组成,其质心坐标可用第一型曲线积分表示为:

$$\bar{x} = \frac{\int_C x dx}{\int_C ds}, \quad \bar{y} = \frac{\int_C y dy}{\int_C ds} \quad (1)$$

其中,  $ds$  为弧微分,  $\int_C ds$  为曲线  $C$  的总弧长。

曲线  $C$  上所有点的分散程度可用方差  $\sigma^2$  来描述:

$$\sigma^2 = \frac{\int_C [(x - \bar{x})^2 + (y - \bar{y})^2] ds}{\int_C ds} \quad (2)$$

用  $r$  来表示曲线  $C$  上所有点与质心之间距离的平均值:

$$r = \frac{\int_C \sqrt{(x - \bar{x})^2 + (y - \bar{y})^2} ds}{\int_C ds} \quad (3)$$

称以  $(\bar{x}, \bar{y})$  为圆心、以  $r$  为半径的圆为曲线  $C$  的参照圆。

对任意的  $\lambda \in (0, +\infty)$ , 记曲线  $C^{(\lambda)} = \{(x, y) | (x,$

$y) \in C, (x - \bar{x})^2 + (y - \bar{y})^2 \leq (\lambda r)^2$ 。对任意的  $\theta \in [0, 2\pi]$  和  $\lambda \in (0, +\infty)$ , 定义函数:

$$f(\theta, \lambda) = \frac{\int_{C(\lambda)} e^{-\frac{(x-\bar{x}-\lambda r \cos\theta)^2 + (y-\bar{y}-\lambda r \sin\theta)^2}{2\sigma^2}} ds}{\int_C ds} \quad (4)$$

称函数  $f(\theta, \lambda)$  为高斯势函数 (Gaussian Potential Function, GPF)。对于给定的  $\theta$  和  $\lambda$ ,  $f(\theta, \lambda)$  是关于平移和尺度不变的。

最后, 我们定义函数:

$$I(\lambda) = \frac{\int_0^{2\pi} f(\theta, \lambda) d\theta}{2\pi} \quad (5)$$

称  $I(\lambda)$  为曲线  $C$  的高斯描绘子 (Gaussian Descriptor, GD)。可以证明, 对于给定的  $\lambda$ , 高斯描绘子  $I(\lambda)$  具有平移、旋转、尺度和反射不变性。

对于数字图像的边缘, 需将 (1) ~ (5) 式离散化。为此, 我们先给出计算曲线  $C$  总弧长  $arclength$  的离散化公式:

$$arclength = \sum_{i=1}^N \Delta s_i \quad (6)$$

(1) ~ (4) 式的离散化结果分别对应下列各式:

$$\bar{x} = \frac{\sum_{i=1}^N x_i \Delta s_i}{arclength}, \bar{y} = \frac{\sum_{i=1}^N y_i \Delta s_i}{arclength}, (x_i, y_i) \in C \quad (7)$$

$$\sigma^2 = \frac{\sum_{i=1}^N [(x_i - \bar{x})^2 + (y_i - \bar{y})^2] \Delta s_i}{arclength}, (x_i, y_i) \in C \quad (8)$$

$$r = \frac{\sum_{i=1}^N \sqrt{(x_i - \bar{x})^2 + (y_i - \bar{y})^2} \Delta s_i}{arclength}, (x_i, y_i) \in C \quad (9)$$

$$f(\theta, \lambda) = \frac{\sum_{i=1}^M e^{-\frac{(x_i - \bar{x} - \lambda r \cos\theta)^2 + (y_i - \bar{y} - \lambda r \sin\theta)^2}{2\sigma^2}} \Delta s_i}{\int_C ds} \quad (10)$$

在 (6) ~ (10) 式中,  $\Delta s_i$  是轮廓曲线上两个相邻像素点之间的距离: 如果两个相邻像素点的连线与水平方向平行或垂直, 则  $\Delta s_i = 1$ ; 如果两个相邻像素点的连线与水平方向成  $45^\circ$  或  $135^\circ$  角, 则  $\Delta s_i = \sqrt{2}$ 。

为了将 (5) 式离散化, 我们将区间  $[0, 2\pi]$  平分为  $K$  份。则:

$$I(\lambda) \approx \frac{\sum_{i=1}^K f(\frac{2\pi}{K}i, \lambda)}{K} \quad (11)$$

$K$  的取值范围为  $3 \leq K \leq 8$  ( $K$  为整数), 本文实验取  $K = 8$ 。

为了得到足够的信息来刻画物体的形状特征, 从而提高识别/检索率, 我们取了 8 个不同的  $\lambda$  值, 它们分别是  $\lambda_1 = 0.25r$ ,  $\lambda_2 = 0.50r$ ,  $\lambda_3 = 0.75r$ ,  $\lambda_4 = 1.00r$ ,  $\lambda_5 = 1.25r$ ,  $\lambda_6 = 1.50r$ ,  $\lambda_7 = 1.75r$  和  $\lambda_8 = 2.00r$ 。相应地, 可得到 8 个特征  $I_1 \sim I_8$ , 如前所言, 它们具有平移、旋转、尺度和反射不变性。我们用这 8 个不变量构成字符图像的特征向量并用于字符识别。

算法描述如下:

1) 提取字符的边缘 (使用小波进行多尺度边缘提取<sup>[12]</sup>);

2) 计算总弧长  $arclength$ , 质心坐标  $(\bar{x}, \bar{y})$ , 方差  $\sigma^2$  和参照圆半径  $r$ ;

3) 计算高斯描绘子的 8 个离散值以构成字符图像的特

征向量:

for  $j = 1:8$

$I_j = 0$

for  $i = 1:8$

$I_j = I_j + f(\frac{\pi}{4}i, \lambda_j)$

end

$I_j = I_j/8$

end

4) 分类决策或模板匹配。

## 2 数值实验

图 1 是一些字符图像, 其中, (a) 图中的字符是模板, (b) 和 (c) 分别是 (a) 的旋转和伸缩版本。

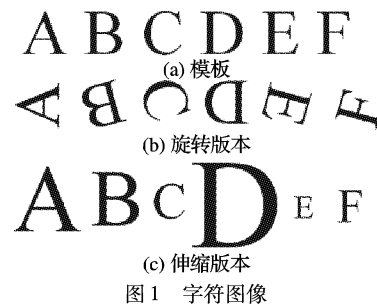


图 1 字符图像

利用小波进行边缘提取, 得到图 2。

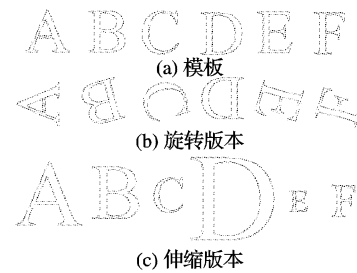


图 2 字符图像的边缘

计算图 2 中各边缘图像的 8 个高斯描绘子, 结果如表 1 所示 (扩大 100 倍)。

表 1 英文字母的高斯描绘子

字母	样式	$I_1$	$I_2$	$I_3$	$I_4$	$I_5$	$I_6$	$I_7$	$I_8$
A	模板	4.84	11.45	21.19	29.71	32.70	31.13	27.50	22.18
	旋转	4.75	11.42	21.02	29.80	32.61	31.09	27.52	22.16
	伸缩	4.86	11.37	21.07	29.92	32.79	31.06	27.48	22.19
B	模板	4.93	10.41	16.84	24.05	30.93	32.86	27.82	21.36
	旋转	4.96	10.64	16.53	23.73	30.97	32.95	27.81	21.36
	伸缩	4.89	10.32	16.90	24.00	30.90	32.90	27.82	21.36
C	模板	0.00	0.00	6.37	21.45	39.08	32.84	25.70	19.18
	旋转	0.00	0.00	6.53	21.65	39.12	32.85	25.71	19.19
	伸缩	0.00	0.00	6.28	21.63	39.09	32.84	25.70	19.18
D	模板	0.00	1.94	7.30	23.71	36.73	32.93	26.09	19.59
	旋转	0.00	1.94	7.30	23.77	36.68	32.95	26.10	19.59
	伸缩	0.00	1.88	7.23	23.83	36.70	32.94	26.09	19.59
E	模板	5.27	12.07	21.00	22.46	31.63	33.62	28.05	21.61
	旋转	5.23	12.00	20.94	22.46	31.37	33.66	28.05	21.61
	伸缩	5.25	12.02	21.13	22.66	31.73	33.61	28.08	21.64
F	模板	5.45	12.34	23.69	27.09	32.28	31.35	27.97	22.15
	旋转	5.42	12.32	23.55	27.29	32.26	31.34	27.94	22.16
	伸缩	5.62	12.17	23.50	27.03	32.22	31.37	28.01	22.16

利用如下的 Chebychev 度量计算两幅图像之间的距离:

$$d_{i,j} = \max_k |x_{ik} - x_{jk}| \quad (12)$$

可得到如表 2 所示的距离矩阵。

表 2 以高斯描绘子为特征的距离矩阵

模 板	旋转后的字母						伸缩后的字母					
	A	B	C	D	E	F	A	B	C	D	E	F
A	0.17	5.99	14.66	13.89	7.25	2.43	0.20	5.72	14.90	13.96	7.05	2.68
B	5.75	0.33	10.41	9.54	4.10	6.71	5.86	0.08	10.56	9.61	4.29	6.66
C	14.64	10.64	0.20	2.40	14.57	17.18	14.70	10.52	0.19	2.39	14.76	17.12
D	13.72	9.24	2.39	0.06	13.64	16.26	13.77	9.60	2.35	0.13	13.84	16.20
E	7.34	4.46	14.47	13.70	0.27	4.83	7.46	4.10	14.72	13.77	0.20	4.57
F	2.71	7.15	17.16	16.39	4.62	0.20	2.83	6.79	17.41	16.46	4.42	0.19

为方便对比,我们再计算轮廓矩不变量,结果如表 3 所示(因为原值太小,我们对其绝对值取常用对数)。

表 3 英文字母的轮廓矩不变量

字母	样式	$\phi_1$	$\phi_2$	$\phi_3$	$\phi_4$	$\phi_5$	$\phi_6$	$\phi_7$
A	模板	-2.44	-6.56	-7.44	-9.26	-18.72	-12.83	-17.61
	旋转	-2.44	-6.54	-7.44	-9.27	-19.08	-12.82	-17.63
	伸缩	-2.44	-6.58	-7.43	-9.24	-18.52	-12.86	-17.58
B	模板	-2.49	-6.11	-8.69	-9.55	-18.69	-12.60	-19.16
	旋转	-2.52	-6.13	-8.77	-9.77	-19.05	-12.84	-19.77
	伸缩	-2.49	-6.12	-8.70	-9.56	-18.72	-12.62	-19.15
C	模板	-2.22	-5.56	-7.93	-8.63	-16.93	-11.42	-17.47
	旋转	-2.22	-5.56	-7.91	-8.64	-16.93	-11.42	-17.47
	伸缩	-2.22	-5.57	-7.95	-8.63	-16.95	-11.42	-17.40
D	模板	-2.33	-6.23	-8.13	-9.52	-18.34	-12.63	-20.02
	旋转	-2.33	-6.24	-8.13	-9.50	-18.32	-12.62	-19.21
	伸缩	-2.34	-6.24	-8.14	-9.49	-18.31	-12.61	-19.25
E	模板	-2.40	-5.63	-9.26	-10.53	-21.50	-13.42	-20.43
	旋转	-2.45	-5.73	-9.33	-10.92	-21.85	-13.84	-21.05
	伸缩	-2.40	-5.63	-9.30	-10.50	-20.84	-13.46	-20.43
F	模板	-2.28	-5.12	-7.44	-8.35	-16.30	-10.94	-16.59
	旋转	-2.33	-5.22	-7.58	-8.49	-16.56	-11.12	-16.91
	伸缩	-2.28	-5.12	-7.43	-8.34	-16.28	-10.93	-16.56

依然用 Chebychev 度量计算两幅图像之间的距离,得到表 4。

表 4 以轮廓矩不变量为特征的距离矩阵

模 板	旋转后的字母						伸缩后的字母					
	A	B	C	D	E	F	A	B	C	D	E	F
A	0.37	2.16	1.79	1.61	3.44	2.15	0.20	1.54	1.77	1.64	2.81	2.43
B	1.54	0.61	1.76	0.56	3.17	2.25	1.59	0.03	1.76	0.55	2.15	2.61
C	2.15	2.31	0.02	1.75	4.92	0.56	1.59	1.79	0.07	1.79	3.91	0.91
D	2.40	0.71	2.56	0.80	3.52	3.12	2.45	0.87	2.63	0.77	2.50	3.47
E	2.80	2.45	4.57	3.18	0.63	4.93	2.98	2.78	4.55	3.19	0.66	5.22
F	2.41	3.18	0.88	2.63	5.55	0.32	2.22	2.56	0.81	2.67	4.54	0.03

(上接第 2761 页)

## 5 结语

$\mu\text{C}/\text{OS}-\text{II}$  移植到 TMS320LF2407A 上,实现了电压电流信号的采集、频率的测量、CAN 通信等多个功能,为系统进一步功能扩展和应用到电力系统测控中打下基础。

参考文献:

按最小距离分类,由表 2 可知,高斯描绘子的识别率为 100%;由表 4 可知,旋转和伸缩后的字母 D 都被误识为 B,所以轮廓矩不变量的识别率仅为  $10/12 = 83\%$ 。而且,从这两个表还可看出,相对于轮廓矩不变量,高斯描绘子的类间距离更大,类内距离更小。

我们还用很多其他字符做过实验,结果表明:高斯描绘子的识别率通常都大于 95%,而轮廓矩不变量的识别率接近 87%;至于类间和类内距离,结论与上述相同。

## 3 结语

本文将基于边缘的形状特征高斯描绘子用于字符识别,并与轮廓矩不变量(也是一种基于边缘的特征,是经典的 Hu 矩的推广和改进)进行比较。实验结果表明,较之后者,高斯描绘子具有高得多的识别率,且类内距离很小,类间距离很大。

参考文献:

- [1] CHELLAPPA R, BAGDAZIAN R. Fourier coding of image boundaries[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1984, 6(1): 102-105.
- [2] 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.
- [3] TAKAHASHI H. A neural net OCR using geometrical and zonal pattern features[A]. Proceedings of First International Conference on Document Analysis and Recognition [C]. Saint-Malo, France, 1991. 821-828.
- [4] KIMURA F, SHRIDHAR M. Handwritten numerical recognition based on multiple algorithms[J]. Pattern Recognition, 1991, 24(10): 969-983.
- [5] EL-KHALY F, SID-AHMED MA. Machine recognition of optically captured machine printed arabic text[J]. Pattern Recognition, 1990, 23(11): 1207-1214.
- [6] TSIRIKOLIAS K, MERTZIOS BC. Statistical pattern recognition using efficient two-dimensional moments with applications to character recognition[J]. Pattern Recognition, 1993, 26(6): 877-882.
- [7] KHOTANZAD A, HONG YH. Invariant image recognition by Zernike moments[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, 12(5): 489-497.
- [8] KASTURI R, TRIVEDI MM. Image Analysis Applications[M]. New York: Marcel Dekker, 1990.
- [9] ANDREWS HC. Multidimensional rotations in feature selection[J]. IEEE Transactions on Computers, 1971, 20(9): 1045-1051.
- [10] 刘亦书,杨力华,孙倩. 轮廓矩不变量及其在物体识别中的应用[J]. 中国图像图形学报, 2004, 9(3): 308-313.
- [11] MALLAT S, HWANG WL. Singularity detection and processing with wavelets[J]. IEEE Transactions on Information Theory, 1992, 38(2): 617-643.

- [1] 刘和平,王维俊,江渝,等. TMS320LF240x DSP C 语言开发应用[M]. 北京:北京航空航天大学出版社, 2003.
- [2] TI. TMS320LF/LC240XA DSP Controllers System and Peripherals Reference Guide[Z], 2001.
- [3] LABROSSE JL. 嵌入式实时操作系统  $\mu\text{C}/\text{OS}-\text{II}$  [M]. 第 2 版. 邵贝贝,等译. 北京:北京航空航天大学出版社, 2003.
- [4] 尹勇,欧光军,关荣锋. DSP 集成开发环境 CCS 开发指南[M]. 北京:北京航空航天大学出版社, 2003.