

轮廓矩不变量及其在物体形状识别中的应用

刘亦书 杨力华 孙倩

(中山大学科学计算与计算机应用系, 广州 510275)

摘要 为了有效地刻画物体的形状特征, 在基于区域的 Hu 矩不变量的基础上, 构造了一种基于物体轮廓曲线的新矩不变量, 即轮廓矩不变量。该不变量不仅独立于物体本身的颜色和灰度级, 而且具有平移、旋转和尺度不变性, 因此可将轮廓矩不变量应用于物体形状的识别, 为了能快速地进行物体形状识别, 还讨论了小波边缘检测和轮廓的获取问题及其算法。实验表明, 基于这种轮廓矩的识别算法具有很好的识别率。

关键词 Hu 矩不变量 轮廓矩不变量 模式识别 小波变换

中图分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2004)03-0308-06

Contour-based Moment Invariants and Their Application to the Recognition of Object Shapes

LIU Yi-shu, YANG Li-hua, SUN Qian

(Department of Scientific Computing and Computer Application, Sun Yat-sen University, Guangzhou 510275)

Abstract Object recognition is a challenging problem in the field of pattern recognition and computer vision. Hu's moments are classical tool in the field, which are defined based on the colors or gray levels of objects. This paper is an improvement of Hu's moments. A series of novel moments, which are called contour moments, are constructed based on object contours and their applications to object shape recognition are given in this paper. Some properties of these new moments including the invariance on shift, rotation and scale transforms are studied and proved. A central advantage of the new moments over Hu's moments is that they are independent of the colors or gray levels of objects. They are defined completely by the contours of objects, namely, that they are completely the shape features of objects. To support our new theory, an algorithm for object shape recognition is designed based on the new moments and experiments are conducted. In our experiments, wavelet transforms are employed to extract the contours of objects, therefore, a brief introduction on the theory of wavelet transform as a multi-scale edge detector is introduced. Considering that an object may have more than one contour, each of which is a close curve, this paper also gives detailed discussion on how to deal with several contours. Experiments give an encouraging high recognition rates.

Keywords Hu's moment invariants, contour-based moment invariants, pattern recognition, wavelet transform

1 引言

不变矩理论是模式识别和计算机视觉中的一个重要内容。常用的基于区域的矩不变量理论^[1]最早由 Hu 提出来, 并获得不断发展和广泛应用。为了获得与区域灰度无关的刻画, 本文对其进行了改进, 提出并研究了一种基于轮廓的矩不变量。下面先对 Hu 矩做

简单介绍。

设 $f(x, y)$ 是一个有界二维函数, 其 $(p+q)$ 阶矩定义为

$$m_{p,q} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy \quad (1)$$

$(p, q = 0, 1, 2, \dots)$

当 $f(x, y)$ 在 $x-y$ 平面的有限区域内分段连续时, $m_{p,q}$ 序列与 $f(x, y)$ 可互相唯一确定。

$(p+q)$ 阶中心矩定义为

$$v_{p,q} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x-x_0)^p (y-y_0)^q f(x,y) dx dy \quad (2)$$

$(p, q = 0, 1, 2, \dots)$

其中, (x_0, y_0) 是质心坐标

$$x_0 = \frac{m_{1,0}}{m_{0,0}}, y_0 = \frac{m_{0,1}}{m_{0,0}} \quad (3)$$

对于数字图像,有

$$m_{p,q} = \sum_x \sum_y x^p y^q f(x,y) \quad (4)$$

$$v_{p,q} = \sum_x \sum_y (x-x_0)^p (y-y_0)^q f(x,y) \quad (5)$$

其中, x_0, y_0 依然由式(3)定义。

容易证明,中心矩 $v_{p,q}$ 是平移不变的。为了得到尺度不变性,可对其进行规范化,以便得到如下的尺度规范化矩

$$\mu_{p,q} = \frac{v_{p,q}}{v_{0,0}^{1+(p+q)/2}} = \frac{v_{p,q}}{m_{0,0}^{1+(p+q)/2}}, (p, q = 0, 1, 2, \dots) \quad (6)$$

易证,尺度规范化矩 $\mu_{p,q}$ 是平移、尺度不变的。

1962年, Hu 将代数不变量理论用于上述尺度规范化矩,构造出如下7个著名的 Hu 矩不变量^[1]:

$$\varphi_1 = \mu_{2,0} + \mu_{0,2} \quad (7)$$

$$\varphi_2 = (\mu_{2,0} - \mu_{0,2})^2 + 4\mu_{1,1}^2 \quad (8)$$

$$\varphi_3 = (\mu_{3,0} - 3\mu_{1,2})^2 + (3\mu_{2,1} - \mu_{0,3})^2 \quad (9)$$

$$\varphi_4 = (\mu_{3,0} + \mu_{1,2})^2 + (\mu_{2,1} + \mu_{0,3})^2 \quad (10)$$

$$\begin{aligned} \varphi_5 = & (\mu_{3,0} - 3\mu_{1,2})(\mu_{3,0} + \mu_{1,2}) \times \\ & [(\mu_{3,0} + \mu_{1,2})^2 - 3(\mu_{2,1} + \mu_{0,3})] + \\ & (3\mu_{2,1} - \mu_{0,3})(\mu_{2,1} + \mu_{0,3}) \times \\ & [3(\mu_{3,0} + \mu_{1,2})^2 - (\mu_{2,1} + \mu_{0,3})^2] \quad (11) \end{aligned}$$

$$\begin{aligned} \varphi_6 = & (\mu_{2,0} - \mu_{0,2})[(\mu_{3,0} + \mu_{1,2})^2 - (\mu_{2,1} + \mu_{0,3})^2] + \\ & 4\mu_{1,1}(\mu_{3,0} + \mu_{1,2})(\mu_{2,1} + \mu_{0,3}) \quad (12) \end{aligned}$$

$$\begin{aligned} \varphi_7 = & (3\mu_{2,1} - \mu_{0,3})(\mu_{3,0} + \mu_{1,2}) \times \\ & [(\mu_{3,0} + \mu_{1,2})^2 - 3(\mu_{2,1} + \mu_{0,3})^2] - \\ & (\mu_{3,0} - 3\mu_{1,2})(\mu_{2,1} + \mu_{0,3}) \\ & [3(\mu_{3,0} + \mu_{1,2})^2 - (\mu_{2,1} + \mu_{0,3})^2] \quad (13) \end{aligned}$$

这些 Hu 矩不变量具有平移、旋转和尺度不变性。

这7个矩不变量提出以后,已被应用于图像、字符识别以及工业质量检测等许多领域(参见文献[2, 3, 4])。但是,由于 Hu 矩不变量是基于区域的,且与图像的灰度值密切相关,因而计算量大。在图像识别的许多问题中,因为物体的本质特征不是其灰度,而是其形状,所以,基于灰度的 Hu 矩方法有时并不能有效和正确地反映物体的形状特征,并用之于识别。基于这样的思考,本文在 Hu 矩不变量理论的基础

上,提出了一种基于物体轮廓的新的矩不变量——轮廓矩不变量。相对于 Hu 矩不变量,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、尺度和旋转不变性。

2 轮廓矩不变量

这里先介绍曲线的矩的定义。

平面曲线 l 的 $(p+q)$ 阶矩定义为

$$m_{p,q} = \int x^p y^q ds, (p, q = 0, 1, 2, \dots) \quad (14)$$

其中, ds 为曲线 l 的弧微分。由此容易看出,零阶矩 $m_{0,0}$ 是曲线的弧长。

$(p+q)$ 阶中心矩定义为

$$v_{p,q} = \int (x-x_0)^p (y-y_0)^q ds \quad (15)$$

$(p, q = 0, 1, 2, \dots)$

其中, (x_0, y_0) 是质心坐标

$$x_0 = \frac{m_{1,0}}{m_{0,0}}, y_0 = \frac{m_{0,1}}{m_{0,0}} \quad (16)$$

对于数字图像的边缘,有

$$m_{p,q} = \sum_{x,y} x^p y^q \Delta s \quad (17)$$

$$v_{p,q} = \sum_{x,y} (x-x_0)^p (y-y_0)^q \Delta s \quad (18)$$

其中, Δs 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,那么, $\Delta s = 1$; 如果两个相邻像素点的连线与水平方向成 45° 或 135° 角,那么 $\Delta s = \sqrt{2}$ 。式(17)和式(18)中的 x_0 和 y_0 依然由式(16)定义。

与上节类似,中心矩 $v_{p,q}$ 是平移不变的。为了得到尺度不变性,同样也要引进尺度规范化矩的定义。与上节基于区域的情况不同的是,这时分母中的 $v_{0,0}$ (或 $m_{0,0}$) 的指数是 $1+p+q$ 而非 $1+(p+q)/2$

$$\mu_{p,q} = \frac{v_{p,q}}{v_{0,0}^{1+p+q}} = \frac{v_{p,q}}{m_{0,0}^{1+p+q}}, (p, q = 0, 1, 2, \dots) \quad (19)$$

尺度规范化矩 $\mu_{p,q}$ 的平移不变性是易证的。至于尺度不变性,可先假设曲线是光滑的(这在实际中是可行的,因为任何曲线都可由一条光滑曲线来逼近),然后利用第1类曲线积分的计算公式加以证明。

最后,为了使轮廓矩不变量具有旋转不变性,只需将尺度规范化矩式(19)代入式(7)~式(13)。证明细节类似于文献[1]中基于区域的矩的相应证明。

这样,就可得到对应于 Hu 矩不变量的7个轮

廓矩不变量(式(7)~式(13))。但需要强调的是,这时,式(7)~式(13)中的 $\mu_{p,q}$ ($p, q=0, 1, 2, 3$) 是式(19)所定义的轮廓曲线的尺度规范化矩,而非式(6)所定义的基于区域的尺度规范化矩。

3 数值实验

因为上述理论是基于轮廓曲线的,所以,必须提取图像的边缘。在物体形状识别的实际应用中,由于图像通常是二值的或 256 灰度级的,因此可利用小波变换对其进行多尺度边缘提取。该方法的好处是,由于它是一种基于多灰度级图像的多尺度边缘提取方法,且有较强的抗噪能力,因此通常能够满足应用需要。但对于噪声很强或质量太差的图像,边缘提取以及边缘校正本身就是一个正在研究的困难问题。

3.1 用小波变换提取图像边缘

图像的边缘提取是图像处理领域一个经典的问题,迄今为止已形成了许多行之有效的方法,但对于质量不好的图像,自适应边缘提取依然是一个复杂的问题。本文不深入这一问题的一般性讨论,仅采用小波方法进行边缘提取,以支持本文所提出的轮廓矩理论的应用。有关小波边缘的一般理论请参阅文献[5,6]。

设 $\psi(r)$ 是 $[0, \infty)$ 上一紧支函数,那么

$$\psi^{(1)}(u, v) = \psi(r) \frac{u}{r}, \quad \psi^{(2)}(u, v) = \varphi(r) \frac{v}{r} \quad (20)$$

$$(r = \sqrt{u^2 + v^2})$$

为两个二维小波。如果尺度参数 $s > 0$, 记 $\psi_s^{(i)}(u, v) = \psi^{(i)}(u/s, v/s)/s^2$ ($i=1, 2$), 那么二维图像 $f(x, y)$ 的小波变换为

$$W_s^{(1)}f(x, y) = f * \psi_s^{(1)}(u, v) \quad (21)$$

$$W_s^{(2)}f(x, y) = f * \psi_s^{(2)}(u, v)$$

其中,“*”表示卷积运算。相应的小波变换模定义为

$$M_s f(x, y) = \sqrt{|W_s^{(1)}f(x, y)|^2 + |W_s^{(2)}f(x, y)|^2} \quad (22)$$

为了检测图像的边缘,需要计算小波变换模沿着梯度方向的局部最大值点。对于数字图像中的一个像素点,只考虑与之最邻近的 8 个像素点,从而有 8 个梯度方向。如果上述的尺度参数 $s > 0$, 则图像基于小波变换的梯度方向的数学定义为

$$\alpha_s = \arctan\left(\frac{W_s^{(2)}f(x, y)}{W_s^{(1)}f(x, y)}\right)$$

当 α_s 落入一个区域时,则其将被量化为某一区域中心线的角度。由于每个梯度的相反方向结果是

相同的,所以用 4 个代码就可以表示梯度方向了,例如,可用数字 0, 1, 2, 3 这 4 个代码表示 4 个不同的梯度方向。而每个方向的正切值 $\tan\alpha_s$, 则将分别落入下面的某个区间中:

$$[-1 - \sqrt{2}, 1 - \sqrt{2}); [1 - \sqrt{2}, \sqrt{2} - 1);$$

$$[\sqrt{2} - 1, \sqrt{2} + 1); [\sqrt{2} + 1, +\infty) \cup (-\infty, -1 - \sqrt{2})]$$

当 $\tan\alpha_s$ 落入上面的区间中,则可分别用 0, 1, 2, 3 这几个代码来表示,记为: $Code_{\alpha_s}$ 。给定一个输入数字图像 $\{f(k, l) | k=0, 1, \dots, K; l=0, 1, \dots, L\}$, 则小波变换局部最大模边缘检测算法的步骤如下:

(1) 计算其小波变换的模 $M_s f(k, l)$ 和代码 $Code_{\alpha_s}(k, l)$ (其中 $k=0, 1, \dots, K; l=0, 1, \dots, L$);

(2) 取一个阈值 $T > 0$, 对 $k=0, 1, \dots, K; l=0, 1, \dots, L$, 如果 $|M_s f(k, l)| \geq T$, 且 $|M_s f(k, l)|$ 沿梯度方向达到局部最大,那么 (k, l) 就是一个边缘像素;否则, (k, l) 是一个非边缘像素,而所有边缘像素即构成图像的边缘。

本文所使用的小波 $\psi(r)$ 为如下的二次样条函数

$$\psi(r) = \begin{cases} 8r(3r-2) & 0 \leq r \leq 1/2 \\ -8(r-1)^2 & 1/2 < r \leq 1 \\ 0 & r > 1 \end{cases}$$

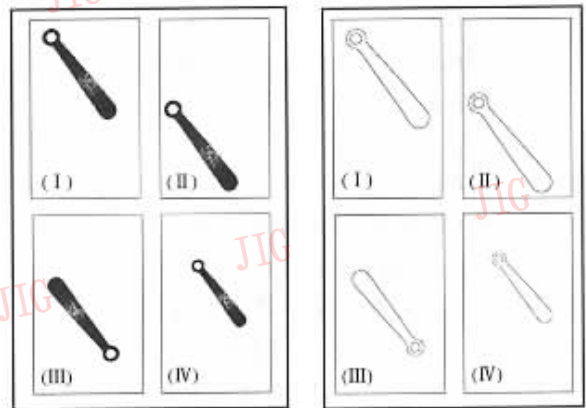
而相应的尺度参数则通常取 $s=4$ 。

3.2 实验结果

下面给出一些实验结果。

实验 1 机械零件识别

图 1(a) 中有 4 幅扳手图像, 其中图(I)、图(II)、图(III)、图(IV)分别是图(I)做平移、旋转和伸缩后的图像, 图 1(b) 是对应的边缘图。

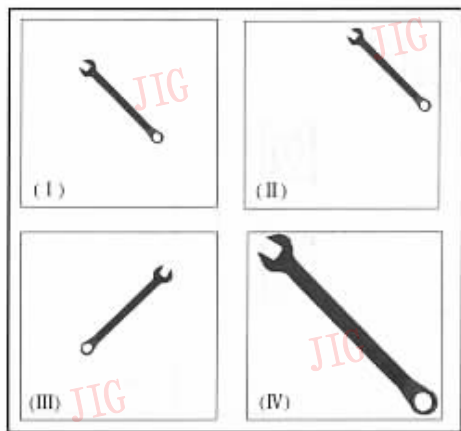


(a) 第 1 组扳手图像

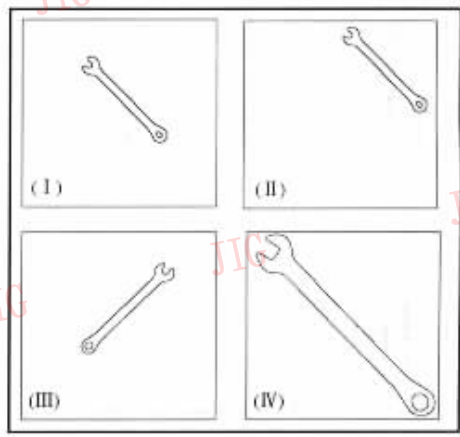
(b) 第 1 组扳手图像的边缘

图 1 第 1 组扳手图像及其边缘图像

类似地, 第 2 组扳手图像如图 2 所示。



(a) 第 2 组扳手图像



(b) 第 2 组扳手图像的边缘

图 2 第 2 组扳手图像及其边缘图像

由图 1(b)和图 2(b)计算出第 1 组和第 2 组扳手图中每幅图像的 7 个轮廓矩不变量分别如表 1 和表 2 所示。

表 1 第 1 组扳手图的轮廓矩不变量

	φ_1	$\varphi_2 \times 10^5$	$\varphi_3 \times 10^8$	$\varphi_4 \times 10^8$	$\varphi_5 \times 10^{16}$	$\varphi_6 \times 10^{10}$	$\varphi_7 \times 10^{17}$
图(I)	0.009 5	7.69	1.25	2.22	3.70	1.94	-1.27
图(II)	0.009 5	7.68	1.26	2.24	3.76	1.96	-1.28
图(III)	0.009 5	7.69	1.25	2.22	3.70	1.94	-1.27
图(IV)	0.009 5	7.69	1.24	2.24	3.74	1.96	-1.29

表 2 第 2 组扳手图的轮廓矩不变量

	φ_1	$\varphi_2 \times 10^5$	$\varphi_3 \times 10^{10}$	$\varphi_4 \times 10^{10}$	$\varphi_5 \times 10^{19}$	$\varphi_6 \times 10^{12}$	$\varphi_7 \times 10^{21}$
图(I)	0.009 2	7.99	2.52	7.66	3.36	6.84	-7.54
图(II)	0.009 2	7.99	2.52	7.66	3.36	6.84	-7.54
图(III)	0.009 2	8.00	2.99	8.40	4.21	7.52	-6.71
图(IV)	0.009 1	7.73	1.80	6.39	2.17	5.62	-8.75

为了利用轮廓矩进行物体形状的识别,下面的实验中选择图 1(b)中的图(I)作为模板 1,选取图 2(b)中的图(I)作为模板 2,然后利用如下的 1 阶 Minkowski 相对度量方法来计算两幅图像之间的距离

$$d_{i,j} = \sum_{k=1}^n \frac{|x_{i,k} - x_{j,k}|}{\min(|x_{i,k}|, |x_{j,k}|)}$$

计算结果如表 3 所示。

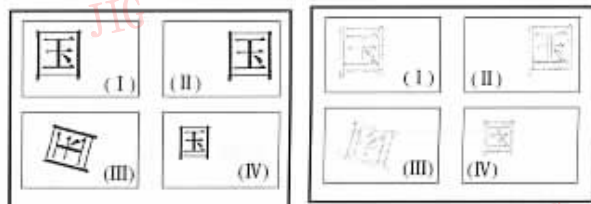
表 3 待识别的图像与两个模板之间的距离

	图 1(b)			图 2(b)		
	(I)	(II)	(IV)	(I)	(II)	(IV)
模板 1	0.050	0.022	0.053	2 885	2 859	3 291
模板 2	2 921	2 852	2 929	0	0.76	1.58

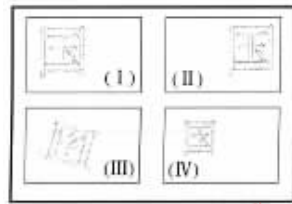
本实验的阈值取 10,也就是说,当待识别的图像与模板之间的距离小于或等于 10 时,就将它归入该模板所属的类。由表 3 可知,实验识别率为 100%。

实验 2 字符识别

在图 3~图 6 中,图(II)、图(III)、图(IV)均是图(I)做平移、旋转和伸缩后的字符图像。



(a) 经平移、旋转、伸缩的“国”字符



(b) “国”的边缘

图 3 “国”字符及其边缘

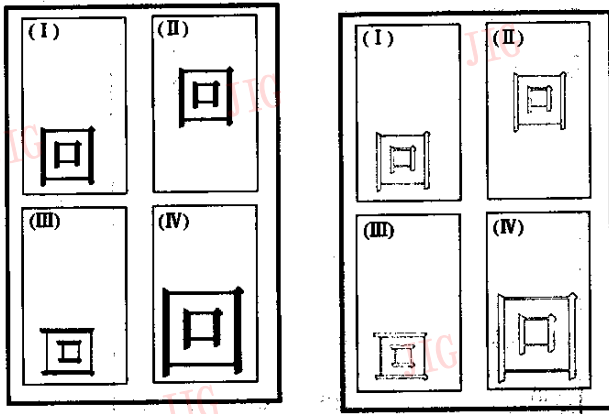


(a) 经平移、旋转、伸缩的“员”字符



(b) “员”的边缘

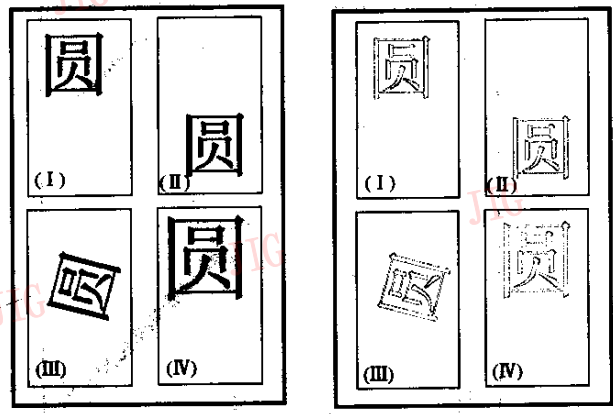
图 4 “员”字符及其边缘



(a) 经平移、旋转、伸缩的“回”字符

(b) “回”的边缘

图 5 “回”字符及其边缘



(a) 经平移、旋转、伸缩的“圆”字符

(b) “圆”的边缘

图 6 “圆”字符及其边缘

计算得到的轮廓矩不变量如表 4 所示。

表 4 汉字字符的轮廓矩不变量

轮廓矩不变量	字符				
	国	员	回	圆	
$\rho_1 \times 10^4$	(I)	10	17	17	8
	(II)	10	17	17	8
	(III)	9	16	17	8
	(IV)	10	17	17	9
$\rho_2 \times 10^9$	(I)	16	202	8	13
	(II)	16	202	8	13
	(III)	13	187	8	10
	(IV)	16	215	7	13
$\rho_3 \times 10^{13}$	(I)	11	3 221	16	8
	(II)	11	3 221	16	8
	(III)	9	3 510	16	5
	(IV)	10	3 433	15	11
$\rho_4 \times 10^{13}$	(I)	31	815	56	4
	(II)	31	815	56	4
	(III)	19	1 046	54	4
	(IV)	31	826	56	3
$\rho_5 \times 10^{26}$	(I)	25	-663 250	-1 419	-12
	(II)	25	-663 250	-1 419	-12
	(III)	28	-649 005	-1 352	-10
	(IV)	13.3	-700 873	-1 362	-9
$\rho_6 \times 10^{17}$	(I)	-36	3 545	40	4
	(II)	-36	3 545	40	4
	(III)	-20	4 493	39	4
	(IV)	-33	3 714	35	3
$\rho_7 \times 10^{25}$	(I)	59	-114 064	-84	-2
	(II)	59	-114 064	-84	-2
	(III)	55	-189 729	-84	-2
	(IV)	51	-120 254	-83	-2

识别时,先将标识为“(I)”的字符作为模板,再利用 1 阶 Minkowski 相对度量方法计算距离,其计算结果如表 5~表 7 所示。

表 5 字符模板图像与其平移后图像的距离

图(II) (平移)	模板(I)			
	国	员	回	圆
国	0	29 216	66	56
员	29 215	0	2 145	122 347
回	66	2 145	0	193
圆	56	122 347	193	0

表 6 字符模板图像与其旋转后图像的距离

图(II) (旋转)	模板(I)			
	国	员	回	圆
国	3.48	28 606	63	29
员	29 987	1.46	3 076	164 071
回	63	2 168	0.10	187
圆	56	129 376	208	1.35

表 7 字符模板图像与其伸缩后图像的距离

图(IV) (伸缩)	模板(I)			
	国	员	回	圆
国	4.79	7 656	20	61
员	30 869	0.32	2 265	129 115
回	63	2 213	0.31	186
圆	66	150 408	245	1.38

然后利用上面 3 个距离矩阵进行分类,其识别率为 100%。

从上面两个实验可看出,因为轮廓矩不变量的确具有平移、旋转和尺度不变性,且能较好地反映出不同类之间的差异性,所以,轮廓矩不变量不失为一种良好的特征。

参 考 文 献

- 1 Hu M K. Visual pattern recognition by moment invariants[J]. IEEE Transactions on Information Theory, 1962, **8**(2):179~187.
- 2 Dudani S A, Breeding K J. Aircraft identification by moment invariants[J]. IEEE Transactions on Computers, 1977, **26**(1): 39~45.
- 3 Sluzek A. Identification and inspection of 2-D objects using new moments-based shape descriptors [J]. Pattern Recognition Letters, 1995, **16**(7):687~697.
- 4 El-Khaly F, Sid-Ahmed M A. Machine recognition of optically captured, machine printed arabic text [J]. Pattern Recognition, 1990, **23**(11):1207~1214.
- 5 Mallat S, Hwang W L. Singularity detection and processing with wavelets[J]. IEEE Transactions on Information Theory, 1992, **38**(2):617~643.
- 6 Mallat S. 信号处理的小波导引(杨力华等译)[M]. 北京:机械工业出版社, 2002.



刘亦书 1974 年生,助教,2002 年获中山大学计算数学专业硕士学位。主要研究方向为图像处理和模式识别。



杨力华 1962 年生,教授,博士生导师,1995 年获中山大学计算数学专业博士学位,现为中山大学计算科学与计算机应用研究所所长。主要研究领域为小波分析、模式识别与图像处理。发表论文 30 余篇,出版专著、译著各 1 部。



孙倩 1979 年生,2001 年获中山大学计算数学专业学士学位,现为中山大学计算数学专业硕士研究生。主要研究方向为小波分析与图像处理。