

中图法分类号: TP391.41 文献标志码: A 文章编号: 1006-8961(2011)05-0844-06

论文索引信息: 束鑫, 吴小俊, 潘磊. 利用 CPDH 和动态规划算法的形状检索 [J]. 中国图象图形学报, 2011, 16(5): 844-849

利用 CPDH 和动态规划算法的形状检索

束鑫^{1),2)}, 吴小俊¹⁾, 潘磊^{1),2)}

¹⁾(江南大学信息工程学院, 无锡 214122) ²⁾(江苏科技大学计算机科学与工程学院, 镇江 212003)

摘要: 用形状轮廓上点的坐标位置相对于形状重心位置的分布关系描述形状, 提出一种极坐标下形状轮廓点分布直方图描述符(contour points distribution histogram), 该描述符不仅符合人眼的视觉直观感受、计算简单, 而且其本质上具有缩放和平移不变性。用动态规划算法(dynamic programming algorithm)来度量轮廓点分布直方图之间的距离, 部分解决了轮廓点分布直方图对于旋转不变性的要求。在多个形状图像数据库中的实验结果表明, 该方法在单目标封闭轮廓的形状图像检索中取得了良好效果。

关键词: 形状; 轮廓点分布直方图; 图像检索; 动态规划算法

Shape retrieval based on CPDH and dynamic programming algorithm

Shu Xin^{1),2)}, Wu Xiaojun¹⁾, Pan Lei^{1),2)}

¹⁾(School of Information Engineering, Jiangnan University, Wuxi 214122 China)

²⁾(School of Computer Science & Engineering, Jiangsu University of Science and Technology, Zhenjiang 212003 China)

Abstract: Using the distribution relations between the coordinates of the points on shape contour and the centroid of the shape to describe a shape, a novel shape descriptor CPDH (contour points distribution histogram) is proposed under polar coordinate. This descriptor not only satisfies the human's visual perception and easy to be calculated, but also it essentially has the properties of invariant to scaling and translation. We also suggest using the Dynamic Programming Algorithm to measure the distance between CPDHs, and that the DP algorithm can partly solve the need of the CPDH's invariant to rotation. With a great deal of experiments in several shape databases, it is shown that this algorithm, used in image retrieval of shape with a single closed contour, can get favorable results.

Keywords: shape; contour points distribution histogram; image retrieval; dynamic programming algorithm

0 引言

近年来, 由于多媒体数据库的大量使用, 基于内容的图像检索成为一个活跃的研究领域。形状是图像可视内容的重要组成部分, 是基于内容图像检索的重要研究方向之一。在 2 维坐标平面中, 形状通常被认为是一条封闭轮廓线所包围的区域。对物体

形状表示和描述已经有了许多方法^[1-2], 通常可分为两大类: 基于区域的方法和基于轮廓的方法。基于区域的表示方法使用形状区域中所有像素来描述形状, 典型的方法有基于矩(moment based)的描述方法, 由 Hu 于 20 世纪 60 年代首先提出^[3]。常用于形状描述的矩还包括 Legendre 矩、Zernike 矩和 Fourier-mellin 矩^[4-6]等。但基于矩的方法通常缺乏明确的物理意义, 且计算量过大, 因此在实际应用中

收稿日期: 2010-02-05; 修回日期: 2010-05-18

基金项目: 国家自然科学基金项目(60973094, 60572034); 江苏省自然科学基金项目(BK2006081); 教育部新世纪优秀人才支持计划基金项目(NCET-06-0487)。

第一作者简介: 束鑫(1979—), 男, 讲师。江南大学信息工程学院轻工信息技术与工程专业博士研究生, 主要研究方向为模式识别、计算机视觉、基于内容的信息检索。E-mail: ecsishu@yahoo.com.cn。

效果不太理想。此外,基于区域的方法还有 Hilbert 曲线^[7]、中轴^[8]和通用傅里叶描述子^[9]等。基于轮廓的表示方法利用边界信息来描述形状,常用方法有简单几何特征(如边界长度、边界半径、曲率、离心率、狭长度、挠度能量等^[10])、弹性匹配^[11]、曲率尺度空间 CSS^[12](curvature scale space)、傅里叶描述子^[13]、小波描述子^[14],形状上下文 SC^[15](shape contexts)等。其中,CSS 方法已被 MPEG-7 指定为描述形状轮廓的标准方法。SC 是近年来提出的性能较为优秀的轮廓描述符,在 2 维形状和字符识别中取得了不俗表现。其关键是要选择一个合适的半径来为每个轮廓点构建局部直方图,若半径过短,则局部直方图不具备良好的鉴别能力;若半径过长,则局部直方图会包含较多的噪声,都会在匹配阶段产生错误的点对应。

为了克服 SC^[15]方法的不足,提出一种新的 2 维形状描述符——轮廓点分布直方图(CPDH)。CPDH 以目标形状的最小外接圆(以目标形状的质心为圆心,以轮廓上的点到质心的最长距离为半径)为特征提取区域,以轮廓特征点分布为统计对象,以径向等分的同心圆和等分的圆周方式分割该最小外接圆而建立统计意义上的轮廓点分布直方图 $H(\rho, \theta, n)$,并以 DP(dynamic programming)算法来度量直方图之间的距离。与 SC 相比,本文只需要用一个点分布直方图 CPDH 就可描述一幅形状图像,不必为轮廓上的每个采样点都构建形状上下文,大大减少了特征提取阶段的计算量。此外,CPDH 的特点还在于形状轮廓反映了图像的显著结构信息,与人的视觉感知特性一致;轮廓点分布直方图不仅很好地反映了图像形状的全局信息也考虑了区域局部特性。检索阶段用 DP 算法来度量两幅图像之间的相似性(直方图之间的距离越小,两幅图像被认为越相似),省去了文献[15]中寻找对应点这一过程,提高了检索效率,而且 DP 算法还在某种程度上弥补了 CPDH 在旋转不变性上的要求。

1 特征提取

1.1 轮廓提取

首先采用标准的 Canny 算子对图像目标轮廓进行提取,但由于轮廓上点的数量过多,为了减少计算

量,采用等间隔距离方法对轮廓点进行采样,得到轮廓上的部分点集合

$$P = \{p_1, p_2, \dots, p_n\} \quad p_i \in \mathbb{R}^2$$

n 表示点的数量,来描述物体的形状(见图 1)。图 1 (a)原图像;(b)用 Canny 算子提取的轮廓;(c)对轮廓采样 100 个点的效果图(黑色圆圈为其最小外接圆,星点为其质心)。

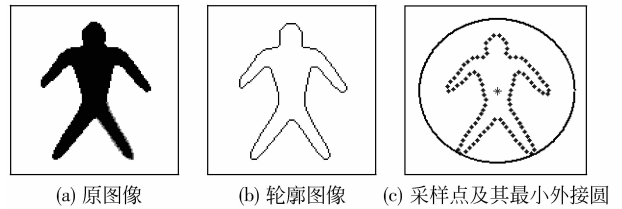


图 1 形状图像及其轮廓

Fig. 1 Shape image and its contour

1.2 构建 CPDH

以图 1(c)轮廓上的采样点集 P 为目标区域,以质心为原点建立极坐标系,并在该最小外接圆内以等分半径的同心圆和等分圆周方式建立如图 2 所示网格模型来统计轮廓点的分布特征。对每一个网格区域 R_i ,得到一个三元组: $H_i = (\rho_i, \theta_i, n_i)$,其中 ρ_i 为第 i 个同心圆的极半径, θ_i 为网格区域 R_i 的一条扇形边与水平方向(顺时针旋转)所成夹角, n_i 为落在网格区域 R_i 中的采样点的个数(Count)。这样,一幅图像 I 就可以描述成 R 个三元组所构成的直方图 H 。

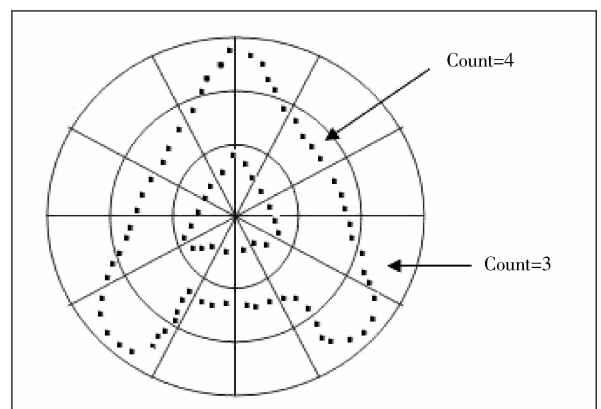


图 2 极坐标下轮廓点分布

Fig. 2 Contour points distribution under polar coordinate

1.3 CPDH 提取实例

依据 1.2 节所述方法对图像进行极坐标下的目

标轮廓点分布直方图提取,效果如图 3 所示。(a)(d)(g)为原始图像((d)是由(a)旋转后的图像), (b)(e)(h)为其对应的轮廓点分布, (c)(f)(i)为对应的轮廓点分布直方图, (c)(f)(i)中不同颜色的小格子表示分布在不同网格区域 R_i 中点的个数差异。从图 3 中还可以看出,由于(d)是由图(a)旋转一定角度后得到的,它们的 CPDH 也只是在 θ 方向上有循环错位,这一错位可以在后面的距离度量中使用 DP 算法加以解决。若图(g)中的形状与图(a)(d)中的存在较大差异,则提取到的 CPDH(i)与(c)(f)的差异也较大,所以可以通过 CPDH 来识别不同的形状。

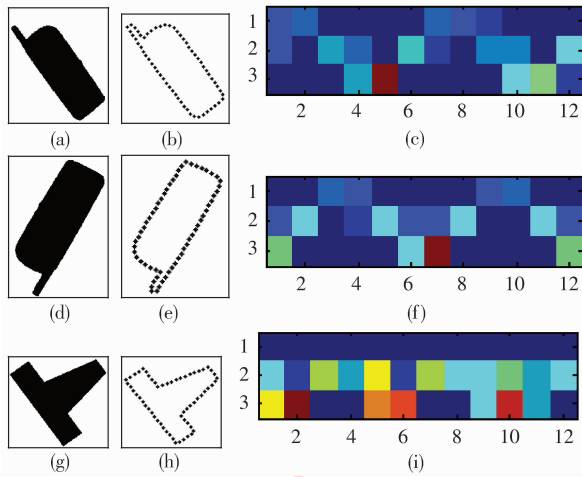


图 3 轮廓点分布直方图

Fig. 3 The contour points distribution histograms

2 CPDHs 距离度量

图像之间的相似性度量是直接影响识别与检索效果的重要因素,人们也一直在致力于研究如何用数值来有效地表示图像在特征上的相似程度。本文采用被广泛用于模式识别技术的 DP^[16-17] 算法来度量目标形状的 CPDH 之间的距离。在 DP 求解过程中,使用一张 DP 表来寻找两个 CPDH 之间的最佳匹配,并把搜索路径限定在对角线两侧固定宽度的窗口内,这样可以提高 DP 算法求解的效率,具体见图 4。把 CPDH 的一列看做是一个向量 r_i (对应于从图 2 的一个扇形区域中提取出来的点分布特征), CPDH1 的一个列向量 r_i 对应于 DP 表的一行, CPDH2 的一个列向量 r_j 对应于 DP 表的一列。DP 表中的初始值为

$$T_{\text{initial}}(n, m) = \begin{cases} 0 & \max(1, n - w + 1) \leq m \leq \min(N, n + w - 1) \\ \infty & \text{其他} \end{cases} \quad (1)$$

式中, $n, m \in \langle 1, N \rangle$, N 为 DP 表的行数或列数, w 为 DP 表对角线的窗口宽度(实验中, $w = 3$)。在 DP 算法求解过程中,只有 DP 表特定区域中(图 4 中的灰色对角线区域)的值得到更新,更新后的值为

$$T(n, m) = D(n, m) + \min \begin{cases} T(n - 1, m) \\ T(n - 1, m - 1) \\ T(n, m - 1) \end{cases} \quad (2)$$

$$D(n, m) = \| r_n - r_m \| \quad (3)$$

式中, $D(n, m)$ 为向量 r_n 和 r_m 之间的欧氏距离(r_n 和 r_m 分别来自 CPDH1 和 CPDH2)。不失一般性, DP 算法求解过程从图 4 中标记为(1,1)的单元开始,沿着指定的对角线从左到右、从下到上根据式(2)依次更新, $T(N, N)$ 中的值即为该次动态规划所求结果。为了弥补 CPDH 对于旋转不变性要求的不足,可采用重复求解而部分地解决。即固定 CPDH1 不动,把 CPDH2 中的各列向量循环左移一列,重复一次 DP 算法,共 N 次,取其最小值为图像之间的相似性度量值,记为

$$S_{\text{original}} = \min_{i=1, \dots, N} \{ T_i(N, N) \} \quad (4)$$

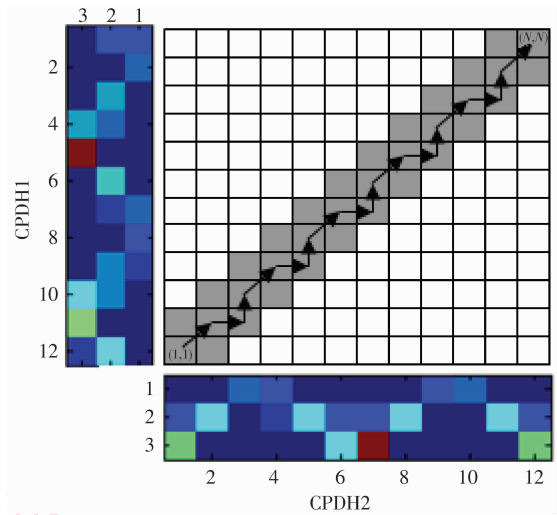


图 4 用于匹配 CPDHs 的 DP 表

Fig. 4 DP table used for matching CPDHs

同样,固定 CPDH1 不动,把 CPDH2 对称翻转一次,重复上述步骤,可得到具有对称不变性的相似性度量值,记为

$$S_{\text{symmetric}} = \min_{i=1, \dots, N} \{ T_i(N, N) \} \quad (5)$$

两幅图像的 CPDH 之间的最终相似度定义为

$$S = \min(S_{\text{original}}, S_{\text{symmetric}}) \quad (6)$$

3 实验结果与分析

3.1 图像数据库

本文所有实验均在国际标准形状图像库中进行,图5给出了部分图像。(a)包含9类不同形状,每类11幅共99幅图像;(b)的图像库包含3个子库 Set A1, Set A2, Set B。其中 Set A1 和 Set A2 都包含420幅图像,分为70个类别,每类6幅,Set A1 用于测试平移和缩放不变性,Set A2用于测试旋转不变

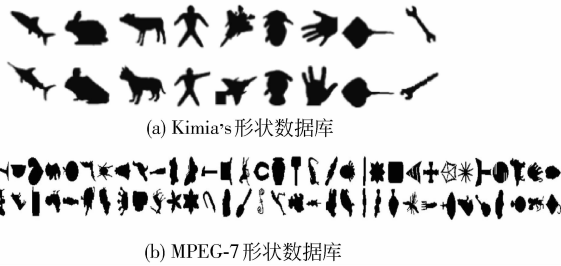


图5 形状图像库

Fig.5 Shape image data sets

性。Set B 包含1400幅图像,分为70个类别,每类20幅,该库用于相似性检索实验,测试形状描述符的全局鲁棒性。

3.2 实验结果与分析

1) CPDH 与 SC 的比较

根据上述算法,实验采用国际通用做法,把图像库中的每幅图像都作为查询图像,分别统计返回的前第*i* ($i \leq N$, N 为每个类别的图像数目)幅图像中与查询图像是同一类别的图像数目,表1和表2给出了CPDH与SC方法的比较,其中表2的数据在MPEG7 Set B的子图像库中进行,该子库包含216幅形状图像,分为18个类别,每类12幅。从表1、2中的数据来看,本文算法要优于SC^[15]算法,因为本文提出的CPDH既表达了图像的全局轮廓特征,通过等分半径的同心圆和旋转角度的方式又体现了形状的局部特征,而SC仅利用了局部轮廓点分布的特征,并且SC方法需要寻找两个形状之间的对应点,这一步骤不但费时而且容易造成全局意义上的错误对应。表1中数据由等间隔采样轮廓上150个点,4等分半径,12等分圆周所构建的CPDH得到;表2中数据由等间隔采样轮廓上200个点,5等分半径,12等分圆周所构建的CPDH得到。

表1 Kimia-99数据中的检索结果比较

Tab.1 Comparison of retrieval rates for different algorithms tested on the Kimia-99 database

方法	<i>i</i>										总数
	1	2	3	4	5	6	7	8	9	10	
SC ^[15]	97	91	88	85	84	77	75	66	56	37	756
CPDH + DP	99	92	91	88	86	83	82	74	63	46	804

表2 MPEG-7子库中的检索结果比较

Tab.2 Comparison of retrieval rates for different algorithms tested on the subset of MPEG-7 shape database

方法	<i>i</i>											总数
	1	2	3	4	5	6	7	8	9	10	11	
SC ^[15]	214	209	205	197	191	178	161	144	131	101	78	1809
CPDH + DP	214	211	208	206	200	196	184	167	152	133	120	1991

2) 查询示例

表3给出了在MPEG7 Set B子库(216幅)中的检索示例,每行第一列为查询图像,其余为查询结果图像,按照相似性大小排序,图像下的数字表示该图像与查询图像之间的距离度量值(值越小越相似)。

3) 不同参数条件下性能比较

表4给出了不同参数条件下构建的CPDH在检索实验中性能比较(MPEG7 Set B图像库中)。从

表4中可以看出,半径等分数取6时,所得到的检索率(采用Bulls-eye^[11]方法进行统计)和识别率都相对较好;在半径等分相同情况下,轮廓上采样点的数目较多,则检索和识别性能也相对较好。

4) 与典型方法的比较

通常,人们使用检索精度和回召率来评价检索算法的性能^[18]。检索精度定义为返回的相关图像数与检索返回的图像总数的比率,检索回召率定义

表 3 在 MPEG-7 子库中的查询结果

Tab.3 Several query results on the subset of MPEG-7 shape dataset

查询图像	最相似的前11幅图像										
	1	2	3	4	5	6	7	8	9	10	11
	9.61	9.86	10.31	10.62	12.49	12.69	13.22	16.89	25.10	28.93	30.35
	18.67	19.32	19.86	21.06	22.75	23.37	23.52	24.09	25.82	26.27	27.26
	4.83	6.24	8.24	8.83	8.97	9.66	9.66	10.24	10.39	10.80	11.66
	6.65	10.06	21.52	23.12	23.72	26.96	27.77	27.80	28.54	30.53	32.02
	7.41	8.24	9.89	10.24	10.80	11.07	12.89	13.54	14.08	14.39	15.26
	15.49	16.03	16.97	17.23	19.03	19.49	19.99	20.26	20.50	22.26	24.52
	10.39	16.59	19.56	19.61	20.48	20.61	21.24	21.48	22.29	33.23	39.61
	16.06	18.23	19.44	20.60	22.02	30.34	32.17	32.78	34.63	36.49	36.57
	15.37	15.46	19.23	21.46	22.04	22.35	22.52	29.23	30.55	32.84	33.29
	10.21	10.24	28.13	30.29	30.50	33.40	38.85	41.17	46.91	47.52	48.94

表 4 不同参数条件下的检索性能比较表

Tab.4 Performances compare with different parameters

等分半径数	采样 100 个轮廓点		采样 200 个轮廓点	
	检索率/%	识别率/%	检索率/%	识别率/%
4	66.42	91.07	67.35	91.57
5	67.91	89.79	68.91	91.36
6	69.01	93.21	69.74	93.43
7	67.20	91.93	68.08	92.86

为返回的相关图像数与图像库中所有相关图像总数的比率。根据 PVR 指标的定义,实验中把数据库中每幅图像都作为查询图像,统计平均检索精度与回召率。本实验中,CPDH 构建的参数为 6 等分半径,12 等分圆周,轮廓等间隔采样 100 个点。图 6—8 给出了本文算法与其他典型的基于轮廓特征的检索算法比较,包括基于轮廓矩^[2](BM)算法;基于曲率尺度空间^[12](CSS)算法;基于轮廓点与质心距离的傅里叶描述^[13](FD)算法的对比。其中,曲率尺度空

间方法是 MPEG-7 指定的形状描述方法。

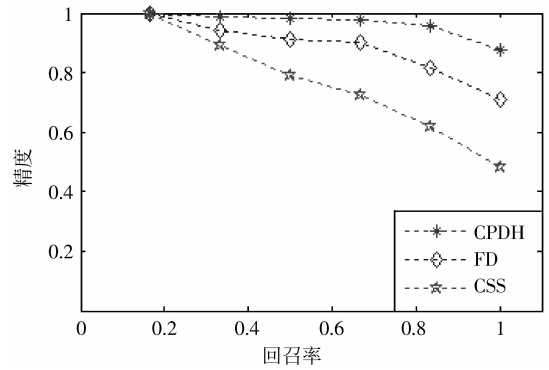


图 6 PVR 曲线比较 (MPEG7 Set A1 图像库中)

Fig.6 Compare of PVR Curves (MPEG7 Set A1 Database)

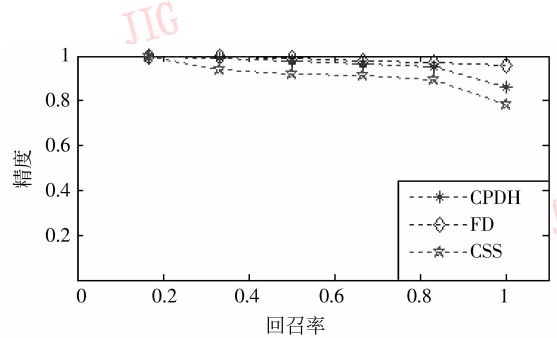


图 7 PVR 曲线比较 (MPEG7 Set A2 图像库中)

Fig.7 Compare of PVR Curves (MPEG7 Set A2 Database)

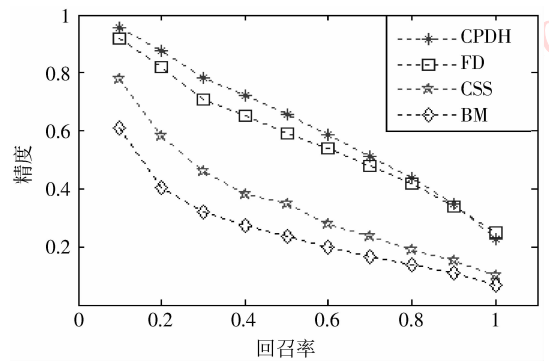


图 8 PVR 曲线比较 (MPEG7 Set B 图像库中)

Fig.8 Compare of PVR Curves (MPEG7 Set B Database)

图 6 中的实验结果来自 MPEG7 Set A1 图像库,从图中 PVR 曲线来看,本文算法较 CSS 和 FD 算法更具有平移和缩放不变性;图 7 中的实验结果来自 Set A2 图像库,从图中 PVR 曲线来看,本文算法对旋转不变性的要求略逊于傅里叶描述子算法,但明显优于曲率尺度空间算法。图 8、9 中的实验结果来自 Set B 图像库,从图中 PVR 曲线来看,本文算法要

优于其他 3 种典型的基于轮廓特征的形状图像检索方法,能取得更好的检索精度与回召率,且较基于轮廓矩和曲率尺度空间方法的优势尤为明显。为了进一步说明本文方法性能的优势,图 9 比较了 CPDH 和 Zernike Moment (ZM) 算法。从图中 PVR 曲线可以看出本文方法也明显优于 MPEG-7 指定的形状区域标准描述方法 Zernike 矩。

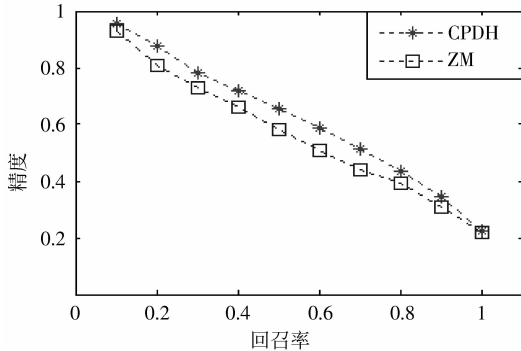


图 9 PVR 曲线比较 (MPEG7 Set B 图像库中)

Fig. 9 Compare of PVR Curves (MPEG7 Set B Database)

4 结 论

提出一种新的形状特征描述符——轮廓点分布直方图 (CPDH), 该特征描述符不仅符合人眼视觉感受, 计算简单, 而且本质上具有对目标的平移与缩放不变性; 而其对旋转不变性的要求可以通过匹配阶段的 DP 算法部分解决。实验结果表明, 本文算法较其他典型算法更能满足平移、缩放和旋转不变性要求。此外, CPDH 描述符可以推广并应用于描述 3 维图像, 这也是作者进一步的研究方向。

参考文献 (References)

- [1] Ding Xianfeng, Wu Hong, Zhang Hongjiang, et al. Review on shape matching [J]. Acta Automatica Sinica, 2001, 27(5): 678-694. [丁险峰, 吴洪, 张宏江, 等. 形状匹配综述 [J]. 自动化学报, 2001, 27(5): 678-694.]
- [2] Zhang D S, Lu G J. Review of shape representation and description techniques [J]. Pattern Recognition, 2004, 37(1): 1-19.
- [3] Hu M K. Visual pattern recognition by moment invariants [J]. IRE Transactions on Information Theory, 1962, 8: 179-187.
- [4] Khalid M H. Exact Legendre moment computation for gray level images [J]. Pattern Recognition, 2007, 40(12): 3597-3605.

- [5] Papakostas G A, Boutalis Y S, Karras D A, et al. A new class of Zernike moments for computer vision applications [J]. Information Sciences, 2007, 177(13): 2802-2819.
- [6] Zhang H, Shu H Z, Haigron P, et al. Construction of a complete set of orthogonal Fourier-Mellin moment invariants for pattern recognition applications [J]. Image and Vision Computing, 2010, 28(1): 38-44.
- [7] Ebrahim Y, Ahmed M, Abdelsalam W, et al. Shape representation and description using the Hilbert curve [J]. Pattern Recognition Letters, 2009, 30(4): 348-358.
- [8] Goh W B. Strategies for shape matching using skeletons [J]. Computer Vision and Image Understanding, 2008, 110(8): 326-345.
- [9] Zhang D S, Lu G J. Shape-based image retrieval using generic Fourier descriptor [J]. Signal Processing: Image Communication, 2002, 17(10): 825-848.
- [10] Castleman K R. Digital Image Processing [M]. Translated by Zhu Zhigang, et al. Beijing: Publishing House of Electronic Industry, 1998 (in Chinese). [Castleman K R. 数字图像处理 [M]. 朱志刚, 等译, 北京: 电子工业出版社, 1998.]
- [11] Attalla E, Siy P. Robust shape similarity retrieval based on contour segmentation polygonal multiresolution and elastic matching [J]. Pattern Recognition, 2005, 38(12): 2229-2241.
- [12] Mokhtarian F, Ung Y K, Wang Z T. Automatic fitting of digitised contours at multiple scales through the curvature scale space technique [J]. Computers & Graphics, 2005, 29(6): 961-971.
- [13] Zhang D S, Lu G J. A comparative study of curvature scale space and Fourier descriptors for shape-based image retrieval [J]. Journal of Visual Communication and Image Representation, 2003, 14(1): 39-57.
- [14] Yadav R B, Nishchal N K, Gupta A K, et al. Retrieval and classification of shape-based objects using Fourier, generic Fourier, and wavelet-Fourier descriptors technique: a comparative study [J]. Optics and Lasers in Engineering, 2007, 45(6): 695-708.
- [15] Belongie S, Malik J. Shape matching and object recognition using shape contexts [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(4): 509-522.
- [16] Alajlan N, Elrube I, Kamel M S, et al. Shape retrieval using triangle-area representation and dynamic space warping [J]. Pattern Recognition, 2007, 40(7): 1911-1920.
- [17] Alajlan N, Kamel M S, Freeman G H. Geometry-based image retrieval in binary image databases [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(6): 1003-1013.
- [18] Zhang Yu Jin. Content-Based Visual Information Retrieval [M]. Beijing: Science Press, 2003 (in Chinese). [章毓晋. 基于内容的视觉信息检索 [M]. 北京: 科学出版社, 2003.]