

基于轮廓重构和特征点弦长的图像检索*

师文^{1,2}, 朱学芳¹

¹(南京大学 信息管理学院 多媒体信息处理研究所, 江苏 南京 210093)

²(河南财经政法大学 计算机与信息工程学院, 河南 郑州 450000)

通讯作者: 师文, E-mail: wens163@163.com

摘要: 轮廓描述法作为形状检索中最为关键的步骤, 应体现目标的整体形状信息和重要特征点信息, 并具备对噪声干扰的鲁棒性. 提出一种基于轮廓重构和特征点弦长的图像检索算法, 首先在目标轮廓提取的基础上分析轮廓的能量保持率, 并进行轮廓的降维重构处理, 从而减少了随机噪声造成的轮廓畸变. 然后, 通过新定义的支持域来计算轮廓点的特征强度, 并分析了支持域半径与特征点提取结果的关系, 从而筛选出有效的轮廓特征点. 最后, 根据轮廓点和相应特征点间的弦长关系构造轮廓特征函数, 经相应处理后, 最终得到的形状描述子满足不变性要求. 大量实验结果表明, 该算法无论是在常规样本库中, 还是在噪声样本库中都具有更优的检索性能.

关键词: 图像检索; 轮廓重构; 傅立叶描述子; 特征点弦长函数; 不变性

中图法分类号: TP391

中文引用格式: 师文, 朱学芳. 基于轮廓重构和特征点弦长的图像检索. 软件学报, 2014, 25(7): 1557-1569. <http://www.jos.org.cn/1000-9825/4443.htm>

英文引用格式: Shi W, Zhu XF. Image retrieval based on contour reconstruction and feature point chord length. Ruan Jian Xue Bao/Journal of Software, 2014, 25(7): 1557-1569 (in Chinese). <http://www.jos.org.cn/1000-9825/4443.htm>

Image Retrieval Based on Contour Reconstruction and Feature Point Chord Length

SHI Wen^{1,2}, ZHU Xue-Fang¹

¹(Institute of Multimedia Information Processing, School of Information Management, Nanjing University, Nanjing 210093, China)

²(College of Computer and Information Engineering, He'nan University of Economics and Law, Zhengzhou 450000, China)

Corresponding author: SHI Wen, E-mail: wens163@163.com

Abstract: As the most important step in shape-based image retrieval, the description of image contour should reflect the information of global shape and key points, and be robust to random noise. This paper proposes a new image retrieval method based on contour reconstruction and feature point chord length. First, the contour of the shape is extracted, and in order to reduce the distortion caused by random noise, the contour is reconstructed by analyzing the energy retention rate. Then, based on the new defined supportive region, the feature intensity is calculated at each point of the contour to extract the valid feature points. After that, the contour feature function is structured by using the chord length between contour points and corresponding feature points. Finally, the shape descriptors are processed to meet the invariance property. A significant amount of experiments show that, in both normal and noisy sample sets, the proposed method demonstrates better performance compared with other seven techniques.

Key words: image retrieval; contour reconstruction; Fourier descriptors; feature point chord length function; invariance

随着数字信息技术的快速发展, 图像数据库的规模不断扩大. 为了有效管理和利用海量图像数据, 基于内容的图像检索(content-based image retrieval, 简称 CBIR)^[1]技术得到广泛的应用. 形状是重要的图像底层特征, 人类

* 基金项目: 国家自然科学基金重大招标项目(10&ZD134)

收稿时间: 2012-12-10; 修改时间: 2013-05-03; 定稿时间: 2013-06-21; jos 在线出版时间: 2014-03-27

CNKI 网络优先出版: 2014-03-27 13:55, <http://www.cnki.net/kcms/doi/10.13328/j.cnki.jos.004443.html>

视觉系统易于通过目标的形状来实现目标信息的感知和识别^[2],因此,基于形状的图像检索方法一直是CBIR研究领域中的活跃分支之一.目标形状的描述作为基于形状的图像检索技术的核心步骤,一直是研究的热点和难点.合理地描述目标形状,能够保证后续形状特征提取的有效性,并提高形状特征匹配的准确率.现有的形状描述方法主要包括基于形状区域和基于形状轮廓两大类.

基于形状区域的描述方法通过将整个图像区域作为考虑对象,来进行形状特征的提取.常见的基于形状区域的描述方法有Hu距^[3]、Legendre矩^[4]、Zernike矩^[5]、Tchebichef距^[6]和Fourier-mellin距^[7]等方法,该类方法计算复杂度较高,且没有明确的物理意义,因而难以在工程实践中得到较好的应用.此外,广义傅立叶变换^[8]、骨架^[9]和Hilbert曲线^[10]等也是较为常见的形状区域描述方法,该类方法多具备对噪声不敏感,且满足不变性要求等特点,但通常存在计算量过大或精确度不高的问题.

基于形状轮廓的描述方法通过目标轮廓上的像素点来提取特征,该类方法能够在描述形状总体特征的同时,增强对形状细节特征的描述,因而更符合人类视觉系统识别目标的习惯.该类方法主要包括傅立叶描述法^[11,12]、曲率尺度空间法^[13,14]和链码法^[15]等,其中,傅立叶描述方法在研究中通常表现出较好的检索精度和较低的算法复杂度^[12,16].在最近提出的形状轮廓描述法中,文献[17]提出了一种极坐标下形状轮廓点分布直方图描述符,该方法利用形状轮廓上点的坐标位置相对于形状中心位置的分布关系来描述形状,并用动态规划算法来度量轮廓点分布直方图之间的距离.文献[18]提出了一种基于高斯函数和骨架化的方法,该方法充分考虑了形状整体特征,并通过计算形状骨架与进化轮廓的距离直方图对目标形状进行描述.然而,动态规划算法在特征匹配过程中往往带来较大的计算量,使其无法很好地应用于大型图像数据库;从目标整体上进行形状分析,则忽视了对局部轮廓特征的准确描述,降低了检索精度.

傅立叶描述法作为一种有效的形状描述方法,在与其他轮廓描述方法的大量对比研究中显示出了优良的性能^[12].在傅立叶描述子的生成过程中,二维图像被转化为一维轮廓特征,通过建立对一维离散轮廓特征的函数表述来实现图像形状特征的聚集和提取,在此基础上,应用经典的傅立叶变换方法生成用于轮廓匹配的傅立叶描述子.得到的傅立叶描述子易于计算和表述,并且归一化过程简单,能够满足各种不变性要求.此外,傅立叶描述子所具有的高计算效率和向量维度的紧致性,使其能够很好地适用于在线图像检索等对实时性要求较高的应用系统.

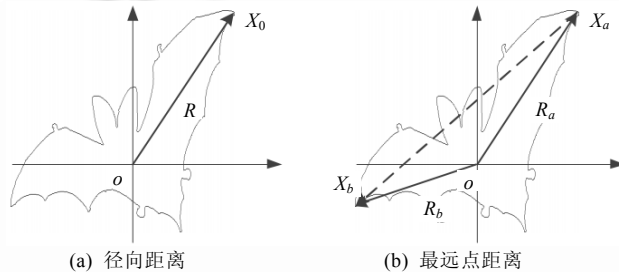


Fig.1 Contour feature description

图1 轮廓特征描述

该领域的学者提出了许多基于傅立叶描述法的轮廓特征函数,常见的有径向距离函数^[19]、角半径坐标函数^[20]、自适应曲率函数^[21]、极坐标函数^[20]以及三角形面积表述函数^[22]等.其中,径向距离函数(如图1(a)所示)通过轮廓上像素点 X_0 到形状几何中心的距离 R 来描述形状的特征,在函数构造中引入几何中心的概念,保证了所得到的形状特征具备对平移的不变性.在相关文献的研究中,径向距离函数在与其他函数的比较中表现出了优越的检索性能^[19].在径向距离函数的基础上,文献[23]提出了一种新颖

的轮廓特征函数:最远点距离函数(如图1(b)所示),在描述函数的构建过程中,首先计算轮廓点 X_a 的径向距离 R_a ,然后根据距离当前轮廓点最远的点 X_b 的位置,计算最远点的径向距离 R_b ,最后将两个距离的和定义为弧长的函数,即为最远点距离函数.由于最远点距离函数是在径向距离函数的基础上改进得到,因此这两种基于形状几何中心的方法都侧重描述轮廓的整体特征,而在轮廓细节信息的提取上存在不足.虽然最远点距离函数加入了最远轮廓点的径向距离信息,但是最远轮廓点并不是形状的点,更不是聚集轮廓主要特征信息的点,因此,该方法在有效体现轮廓的形状特征上仍存在不足.

为了增强轮廓特征函数的描述能力,并提高傅立叶描述方法在随机噪声环境下的检索效率,本文提出了一

种基于轮廓重构和特征点弦长的图像检索算法,本文主要工作如下:

- (1) 对轮廓进行降维重构,以减少轮廓畸变,在分析常规样本轮廓重构的基础上,实现了对噪声样本轮廓的有效重构;
- (2) 提出一种参考点的定位方法,通过改进的支持域,计算轮廓点的特征强度,并分析了特征强度与支持域半径取值的关系;
- (3) 使用轮廓点到最大及最小特征点的弦长构造了一种新的轮廓特征函数,通过离散傅立叶变换生成轮廓描述子,并针对其特点进行取模和归一化处理,以使其满足不变性要求。

1 轮廓重构

图像的轮廓常具有多样性,根据图像复杂度的差异,轮廓像素点的数量也各不相同.可以发现,过多的轮廓特征点会影响视觉感官对形状重要信息的获取,且当图像轮廓线存在随机噪声干扰时,原目标的真实形状也无法有效呈现.为了对目标形状进行统一的规范化,并减少随机噪声带来的曲线畸变,需要对轮廓像素点进行平滑采样处理.

在传统的等间距轮廓采样方法中,采样点数量的多少至关重要,采样点越多,则轮廓形状越逼真,算法复杂度也相应增大;减少采样点,可以有效提高算法的实时性,但是对目标轮廓细节的描述会随之减少,匹配精度也会受到影响.虽然该类方法对平滑轮廓具有较好的采样效果,但图像中往往存在随机噪声的干扰,采样结果会因此受到较大影响.高斯卷积方法^[24]能够较好地消除随机噪声带来的影响,得到平滑的轮廓,但该算法易使轮廓发生收缩形变.传统的傅立叶重构方法^[25]缺乏对重构系数提取量的合理评估和选择,过小的系数频率范围会使重构轮廓损失部分形状信息,而系数频率范围过大时又会造成算法实时性的降低.鉴于上述方法的不足,本文通过分析轮廓的频谱特征来提取适当频率的傅立叶重构系数,并以此对轮廓图像进行降维重构,最后,在重构基础上进行轮廓的标准化采样.

首先提取目标图像的轮廓.设边界轮廓包含的像素数为 K ,则通过边界跟踪算法^[26],可以得到所有的轮廓像素点.图 2 显示了 MPEG-7 CE Shape-1 Part B 标准图像库中 Bat 目标图像的轮廓提取效果.



Fig.2 Result of contour extraction

图 2 轮廓提取结果

轮廓像素点沿逆时针方向的坐标序列依次记为 $(x_0, y_0), \dots, (x_{K-1}, y_{K-1})$,将 $x(k), y(k)$ 分别记为 x_k 和 y_k ,则可用坐标序列 $p(k)=[x(k), y(k)]$ 表示目标轮廓,其中, $k=0, 1, 2, \dots, K-1$.将轮廓像素坐标表示复数化后,其相应的序列表示为 $p(k)=[x(k), jy(k)]$,从而完成了轮廓的参数化.在参数化目标轮廓的基础上,通过离散傅立叶变换将信号 $p(k)$ 映射到频域,如公式(1)所示:

$$z(\mu) = \frac{1}{K} \sum_{k=0}^{K-1} p(k) e^{-j \frac{2\pi k}{K} \mu}, \mu \in [0, K-1] \tag{1}$$

其中, $z(\mu)$ 为轮廓的描述系数.

使用全部复系数对目标轮廓进行傅立叶重构的表达式如下:

$$p(k) = \sum_{\mu=0}^{K-1} z(\mu) e^{j \frac{2\pi k}{K} \mu}, k \in [0, K-1] \quad (2)$$

离散傅立叶变换中的低频描述系数反映了目标的轮廓总体特征,高频描述系数则反映了目标的轮廓细节特征,选择适当数量的低频系数 $P(0 < P < K)$ 进行轮廓重构,能够突出重构轮廓的关键形状信息,并同时增强轮廓对噪声以及分割错误的鲁棒性.使用部分描述系数对目标轮廓的傅立叶重构计算如公式(3)所示:

$$p'(k) = \sum_{\mu=0}^{P-1} z(\mu) e^{j \frac{2\pi k}{K} \mu}, k \in [0, K-1] \quad (3)$$

轮廓重构过程中,可以将实验数据的频谱特征^[27]作为 P 值的选择依据.

设信号的能量表示为 E ,当使用前 P 个系数重构信号时,信号能量可表示为公式(4)的形式,式中排除了对目标形状没有实际描述作用的直流(DC)成分:

$$E(P) = \sum_{k=1}^{K-1} |z(k)|^2 \quad (4)$$

信号的能量保持率定义为 $H = E(P)/E(K)$.图3显示了 MPEG-7 图像库中 Bat 和 Device 两个图形类的平均 H 值变化趋势.实验中发现,当重构系数 P 取值小于 16 时,信号的能量保持率低于 85%,并不能很好地反映信号的特征.当 P 取值大于 64 时,能量保持率呈缓慢增长趋势,其中,当 P 取值范围为 [64, 512] 时,能量保持率的增幅均低于 5%,且此时轮廓的平滑效果将被削弱,算法的实时性也随之明显降低.可以看到,当 P 取值 64 时,相应的能量保持率达到 94.6%,能够有效体现轮廓的主要特征.图4显示了 Bat 和 Device 目标图像的轮廓,以及使用不同 P 值对其进行降维重构的结果.

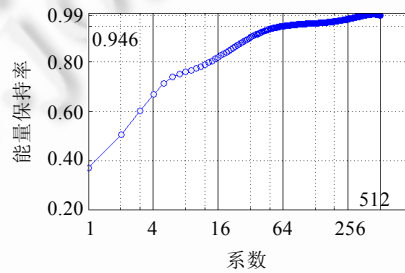


Fig.3 Energy retention curve

图3 能量保持曲线

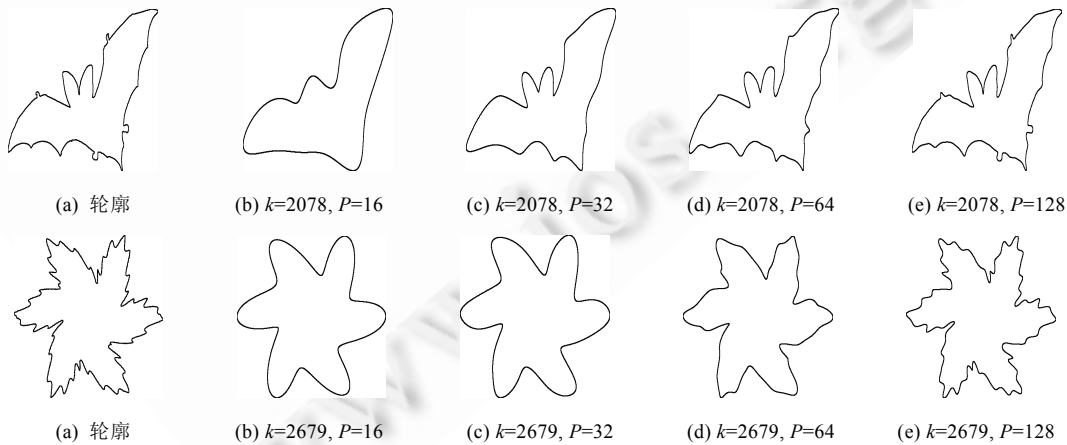


Fig.4 Results of contour reconstruction

图4 轮廓重构结果

高强度的随机噪声会造成目标轮廓线的严重畸变,因此,轮廓的畸变校正是完成轮廓有效描述的重要前提.轮廓重构算法通过削弱傅立叶高频系数对形状的描述作用,使轮廓更为平滑,同时也减少了噪声对目标细节的干扰.在上述算法分析的基础上,对目标轮廓追加 20dB 的随机高斯噪声,并对其进行傅立叶重构.图 5 显示了 Bat 和 Device 目标图像在随机高斯噪声干扰下的轮廓图像,以及使用不同 P 值对其进行降维重构的结果.

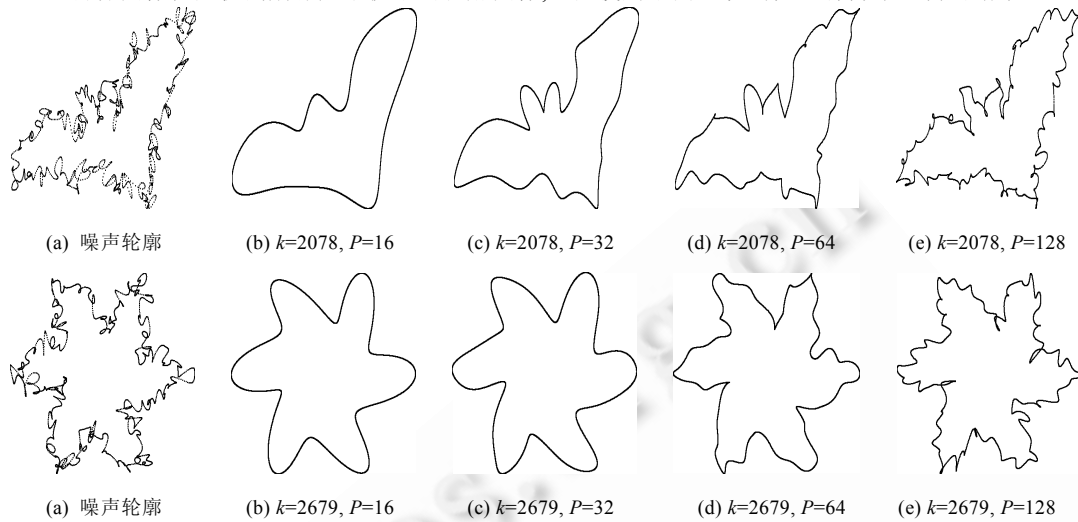


Fig.5 Results of noisy contour reconstruction

图 5 噪声轮廓重构结果

2 基于特征点弦长的轮廓描述

2.1 特征点检测

特征点是一种有效表征形状轮廓的元素,是人眼视觉系统识别形状的主要关注点.本文通过约束条件下的支持域来定位相应的参考点,将轮廓点与参考点之间的距离作为该轮廓点的特征强度,并依据每个轮廓点特征强度值的大小进行特征点筛选,以保证提取特征点的有效性.

目标的轮廓曲线经过重构和采样处理后可表示为平面坐标上的一个点列,将曲线的任意点作为起始点,沿一个方向追踪,可得到一组有序排列的点集 $C=\{p_i=(x_i,y_i),i=0,\dots,K-1\}$, K 表示轮廓点的个数, p_i 和 p_{i+1} 为相邻点,当研究对象为闭合曲线时,满足 $p_{i+K}=p_i$.

对于轮廓上任意一点 p_i ,其相关的支持域可定义为

$$U = \{p_{i-L_1}, p_{i-L_1+1}, \dots, p_{i-1}, p_i, p_{i+1}, \dots, p_{i+L_1-1}, p_{i+L_1}\},$$

其中, p_{i-L_1} 和 p_{i+L_1} 为支持域的两个端点, L_1 和 L_2 为约束条件下的支持域参数.

通过约束条件定义的改变,可以得到不同大小的支持域^[28].等弦长约束条件定义为 $|p_{i-L_1}p_i| = |p_i p_{i+L_1}| = R$,其中, $|p_{i-L_1}p_i|$ 和 $|p_i p_{i+L_1}|$ 表示点 p_i 到弧线两个端点的弦长, R 称为支持域半径.该方法将以点 p_i 为圆心、弦长 R 为半径的圆形区域作为约束条件,圆形区域内包含的轮廓线即为约束条件下的支持域,最后,将支持域确定的多边形的形心作为参考点.通过等弦长约束条件下的支持域来进行参考点的定位,往往存在较大局限性,在曲率半径小于邻域半径等特殊情况下,算法不能表现出足够的敏感度,需要根据支持区域的曲率半径对邻域半径进行自适应调节,计算复杂度较高.

针对上述问题,本文提出一种改进的参考点定位方法,首先以 p_i 为起点,沿顺时针方向走过长度为 R 的弧线段,达到点 p_{i+L_2} ,再从 p_i 沿逆时针方向走过相同长度的弧线段,至点 p_{i-L_1} 处结束,将得到的弧线段 $p_{i-L_1}p_{i+L_2}$ 作为支持域,通过连接弧 $p_{i-L_1}p_{i+L_2}$ 内所有的相邻点就构成一个封闭的区域,可求得该封闭区域的形心为 G' .同理,当

约束条件是以 p_i 为圆心、 R 为半径的圆形区域时,可求得封闭区域的形心为 G'' ,将两个形心连线的中点 G 作为计算弧线上 p_i 点特征强度的参考点.改进算法在保留圆形区域内总体曲率特征的基础上,更精确地表述了中心点的曲率特征.本文参考点的定位方法如图 6 所示.

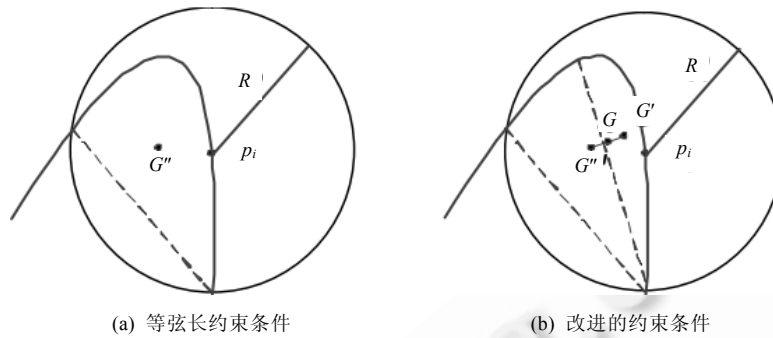


Fig.6 Reference point location

图 6 参考点定位

设起始点 p_i 的坐标为 (x_i, y_i) , 参考点 G 的坐标为 (x_c, y_c) , 通过计算两点间的距离得到点 p_i 的特征强度:

$$d(p_i) = ((x_i - x_c)^2 + (y_i - y_c)^2)^{1/2} \quad (5)$$

根据上述算法,依次计算出轮廓曲线上各点的特征强度,选取其中具备一定特征强度的轮廓点作为候选特征点,筛选阈值 T 的取值范围为 $[30\% \cdot R, 50\% \cdot R]$.最后,采用局部非最大化抑制,进一步排除局部具有一定特征强度的伪特征点的干扰,从而得到最终的特征点.

在特征点的提取算法中,支持域半径 R 的取值直接影响着特征强度计算的敏感度和算法的实时性.图 7 显示了使用 bat 类目标图像作为测试数据时,其平均特征强度的方差在归一化支持域半径 R 影响下的变化情况, R 的取值范围为 $[0, R_0]$,其中 R_0 为距形心最远轮廓点的径向距离.为了避免个别异常数据点造成的方差曲线大幅波动,对得到的特征强度数据进行相应的修正,去除最大和最小 5% 的数据点对方差产生的影响.修正后的方差曲线较为清晰地显示了算法敏感度的总体变化趋势,其中,较小的 R 值代表着较高的计算敏感度,增大取值则特征强度的计算趋于稳定.进一步分析发现,过小和过大的 R 值都不会带来明显的波动,随着取值逐渐增大,特征强度方差值快速增长,在 $R=0.221$ 时取得最大值,并在一定范围内保持较高的数值;随后,方差曲线呈波动下降趋势,逐渐趋于稳定.因此,在轮廓曲线较为平缓时,可以取较小的 R 值,以增强算法敏感度,避免轮廓特征点的漏取,并降低算法复杂度;当轮廓曲线细节变化较为复杂时,可通过取值的适当增大来提高算法稳定性,减少特征值较大的伪特征点的干扰.表 1 显示了支持域半径取值对特征点数以及算法实时性的影响.

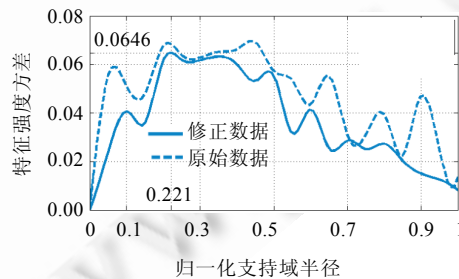


Fig.7 Variance curve of characteristic strength

图 7 特征强度方差曲线

Table 1 Comparison of feature point extraction results

表 1 特征点提取结果比较

归一化 支持域半径	T=30%·R		T=40%·R		T=50%·R	
	特征点数	时间(ms)	特征点数	时间(ms)	特征点数	时间(ms)
0.2	35	1.036	29	1.023	28	0.970
0.4	36	1.278	34	1.121	28	1.105
0.6	34	1.592	29	1.554	29	1.542
0.8	33	1.903	31	1.872	25	1.726

2.2 特征点弦长描述子

特征点作为图像轮廓上曲率存在显著变化的点,能够有效地聚集图像轮廓的形状信息.因此,可以使用轮廓点到特征点的弦长来构成适应视觉识别的轮廓描述函数.设轮廓点 X_i 的坐标为 (x_i, y_i) , 距离 X_i 最远特征点 a 的弦长记为 $L_a(X_i)$, 距离 X_i 最近特征点 b 的弦长记为 $L_b(X_i)$, 则特征点弦长函数可表示如下:

$$L'(X_i) = L_a(X_i) + L_b(X_i) = ((x_i - x_a)^2 + (y_i - y_a)^2)^{1/2} + ((x_i - x_b)^2 + (y_i - y_b)^2)^{1/2} \quad (6)$$

其中, (x_a, y_a) 和 (x_b, y_b) 分别为 a 点和 b 点的坐标.

特征点弦长函数反映了轮廓点与关键特征点之间的距离信息,也反映了二者之间的几何位置关系.公式(6)中,因变量 $L'(X_i)$ 随当前轮廓点 X_i 的变化而变化.若以 X_0 为起始点, X_0 沿逆时针到 X_i 的距离用轮廓点数量表示为 l_i , 则函数 $L'(X_i)$ 可用离散形式表示为 $z(l_0), z(l_1), \dots, z(l_{K-1})$, 其中, K 为轮廓包含的像素点数量.因此,可将 $z(l_i)$ 视作轮廓点 l_i 的函数,则 $L'(X_i)$ 可以表示为轮廓点的有限集离散函数 $z(l_i)$, 其中, $l_i \in [0, K]$.图 8 描述了特征点与轮廓点间弦长函数的生成过程.

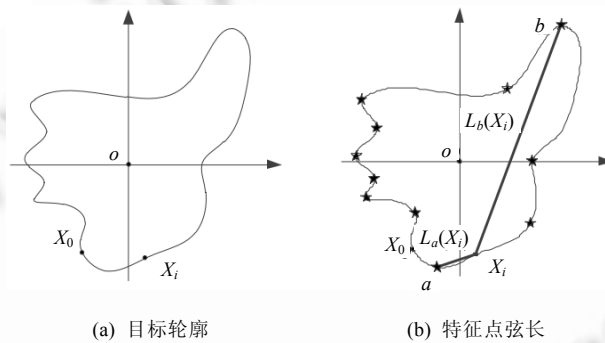


Fig.8 Generation of feature point chord length function
图 8 特征点弦长函数的生成

在空域直接使用特征点弦长函数作为特征向量,需要对函数进行复杂的归一化处理,这样会带来相当大的匹配计算量.因此,本文采用离散傅立叶变换方法,以减少检索过程中的特征匹配环节的计算量,并降低形状特征的噪声敏感度.

使用傅立叶变换方法处理函数 $z(l_i)$ 的时间复杂度为 $T(K^2)$, 为了提高算法的实时性,采用快速傅立叶变换(FFT)对函数进行处理^[20].由于 FFT 算法在 $K=2^n$ 时可以达到最优效率,本文对函数的自变量区间进行等间隔离散化处理^[29],将区间重采样为 N 个点,其中, $N=2^n$.经过上述处理后的特征点弦长函数为 $z(x_0), z(x_1), \dots, z(x_{K-1})$, 函数的时间复杂度为 $T(N \log_2^N)$, 特征点弦长描述子可计算如下:

$$Z(n) = \frac{1}{N} \sum_{i=0}^{N-1} z(x_i) e^{-j \frac{2\pi}{N} n i}, n \in [0, N-1] \quad (7)$$

2.3 不变性分析及处理

用于最终形状匹配的傅立叶描述子应满足各种不变性要求,对描述子的不变性分析及相应处理如下:

性质 1(平移不变性). 文中特征点的提取算法均是基于形状轮廓点与相应参考点间的特征强度,由于特征点在轮廓上的位置分布具有相对于轮廓平移的独立性,则通过特征点弦长函数得到的傅立叶系数在不考虑直流成分时具有平移不变性,此时的不变性傅立叶描述子可表示为 $\{Z(n), 0 < n \leq N-1\}$.

性质 2(尺度不变性). 设原始傅立叶系数和尺度变换后的傅立叶系数分别为 $Z_o(n)$ 和 $Z(n)$, 考虑公式:

$$\frac{Z(n)}{Z(0)} = \frac{s_1 \times Z_o(n)}{s_2 \times Z_o(0)}$$

其中, s_1, s_2 为系数的尺度变化项. 由于尺度变换对傅立叶系数产生无差别影响, 则有 $s_1 = s_2$; 又因为直流成分体现了信号的平均能量值, 且多为系数中的最大值, 所以可以使用直流成分 $Z(0)$ 对傅立叶系数进行尺度归一化处理, 从而得到具有尺度不变性的描述子.

性质 3(旋转不变性). 同性质 2, 傅立叶系数具有无差别的旋转变换项 $e^{j\theta}$, 于是通过消除相同相位信息带来的影响, 可以使傅立叶系数同时满足对旋转的不变性.

基于上述讨论, 将得到的傅立叶描述子表示如下:

$$FD = \left[\frac{Z(1)}{Z(0)}, \frac{Z(2)}{Z(0)}, \dots, \frac{Z(N-1)}{Z(0)} \right] \quad (8)$$

经过改造的傅立叶描述子 FD 具备对于平移、旋转以及尺度变化的不变性, 并且经过归一化处理后, 其取值范围为 $[0, 1]$.

性质 4(起始点不变性). 在目标轮廓为闭合曲线的情况下, 起始点 X_0 的变化会对函数 $L(l)$ 产生影响^[30]. 不同的起始点位置, 会使函数在水平轴方向发生位移, 从而导致傅立叶系数相位信息的改变. 由于傅立叶系数存在有差异起始点变化项, 则直流部分与其他系数的相位差异可表示为 $\frac{Z(n)}{Z(0)} = \frac{e^{jn\alpha} \times Z_o(n)}{e^{j\alpha} \times Z_o(0)} = \frac{Z_o(n)}{Z_o(0)} e^{j(n-1)\alpha}$. 因此, 在生成最终的形状描述子时, 通过提取傅立叶系数的幅度信息, 并同时消除相位信息改变带来的影响, 可以使形状描述子进一步具备对起始点变化的不变性.

为了使特征向量紧致而有效, 并降低其对随机噪声的敏感度, 取低频域的前 $M(0 < M < N)$ 个傅立叶系数, 就构成了本文最终得到的基于特征点弦长的形状描述子:

$$FD_0 = \left[\frac{|Z(1)|}{|Z(0)|}, \frac{|Z(2)|}{|Z(0)|}, \dots, \frac{|Z(M)|}{|Z(0)|} \right] \quad (9)$$

低频傅立叶系数体现了轮廓的总体形状, 具有较高的分布稳定性, 而高频傅立叶系数较多地体现了轮廓的细节信息, 对噪声敏感度较高. 因此, 选择适当数量的轮廓描述子作为特征向量能够保证算法最终的检索效率.

2.4 相似性度量

Hausdorff 距离表征了匹配特征间的极大、极小距离, 能够有效地度量两个形状间的差异, 因此, 本文将 Hausdorff 距离作为度量图像相似度的标准. 设检索图像 A 和库中图像 B 的特征点弦长描述子为 $A = \{a_1, a_2, \dots, a_p\}$ 和 $B = \{b_1, b_2, \dots, b_p\}$, 则 A, B 间的 Hausdorff 距离定义如下:

$$H(A, B) = \max[h(A, B), h(B, A)] \quad (10)$$

其中,

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|, \quad h(B, A) = \max_{a \in A} \min_{b \in B} \|b - a\|.$$

本文使用欧式距离来表示点集间的距离范数 $\|\cdot\|$. $h(A, B)$ 称为点集 A, B 间的有向 Hausdorff 距离, 其中, a 点代表了点集 A 中距离点集 B 最远的点, $h(A, B)$ 即为点 a 到点集 B 的最小距离, 同理可以解释 $h(B, A)$.

3 实验结果与讨论

3.1 形状测试集

为了有效评价不同形状描述算法的性能, 需要对测试图库进行合理的选择. 一些研究中使用的自建图库存

在一定的局限性,在图库大小以及实验图像选择的普适性上缺乏统一标准.MPEG-7 标准图像库在近年来的图像检索研究中得到了广泛的应用,该库中的图像均来源于实物,可应用于不同图像种类以及同类图像在不同变换下的图像检索实验.本文的检索算法测试部分均是基于 MPEG-7 中的 CE-1 形状库,在该库的子库 Set B 中包含有 1 400 种图形,这些图形被分为 70 个类,每类图形中又包含 20 种存在类内差异的样本图形.Set B 可用于测试基于相似性的检索算法的性能,以及测试形状描述子对各种随机形状变化的鲁棒性.图 9 显示了部分来自于 Set B 的图形样本.

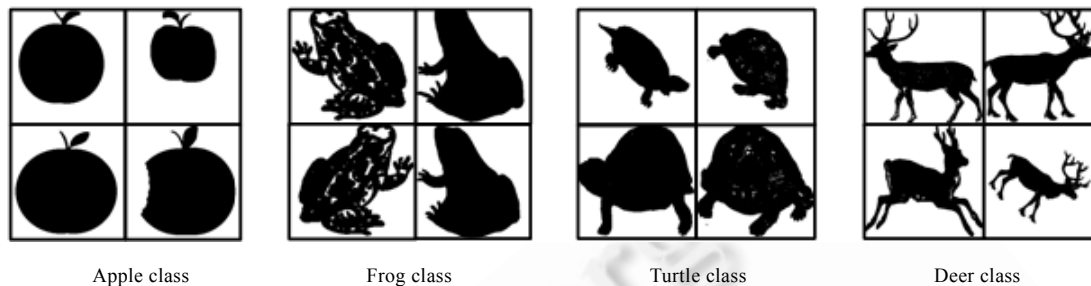


Fig.9 Some graphic samples from Set B

图 9 Set B 中的部分图形样本

3.2 对比实验及参数设置

为了测试不同算法在常规图库以及噪声图库中的检索效果,在实验中,选择最近提出以及常见的几种检索算法与本文算法进行比较.由于本文算法的主要思想是基于形状轮廓的描述,因此将同属于轮廓描述方法的常见算法作为比较对象,其中包括径向距离(radial distance,简称 RD)描述方法^[19]、角半径坐标(angular radial coordinates,简称 ARC)描述方法^[20]、自适应曲率(adaptive curvature,简称 AC)描述方法^[21]、极坐标(polar coordinates,简称 PC)描述方法^[20]以及三角形面积表述(triangular area representation,简称 TAR)方法^[22].在新近提出的算法中,选择最远点距离(farthest point distance,简称 FPD)算法^[23]以及一种基于曲率(curvature-based,简称 CB)的算法^[31]作为比较对象.

本节的实验均是在 Intel Core2 Duo 2.26 GHz 处理器、4 GB 内存的 PC 机以及 MATLAB 7.1 开发环境下完成的.根据傅立叶系数对轮廓能量保持率的影响,在轮廓平滑重构过程中,将 P 值设置为 64,并将重构轮廓统一采样为 256 个像素点.通过文中对支持域半径与特征点数量及算法实时性关系的分析,取 $R=0.4, T=40\% \cdot R$.轮廓的平滑重构算法能够有效降低随机噪声带来的轮廓畸变,也相应地减少了傅立叶高频系数受到的噪声干扰,因此在特征向量的维度控制上,可以将 M 值设置为 32,通过较大的 M 值引入适量的高频系数来丰富轮廓细节的描述.

3.3 检索结果与分析

目前较为通用的图像检索性能评估标准是查全率和查准率.查全率的定义为 $R=r/(r+m)$,查准率则定义为 $P=r/(r+n)$.其中, r 代表检索结果返回图中与查询图相关的图像数目; m 为查询到的不相关图像的数目; n 为图库中所有与检测图像相关,但未能检出的图像数目.

图 10 显示了各种方法在 MPEG-7 Set B 标准图像库中的查全率、查准率曲线,由于 AC 算法在实验中得到的平均检索率较低,当返回图像数统一设置为 40 幅时,该算法的平均检索率(所有图像查全率的均值)仅为 27.82%,因此图中给出了检索性能较好的 6 种算法与本文算法的比较结果.实验结果显示,当查全率相同时,FPCL 能够得到更高的查准率,具有较为明显的优势.在其他比较算法中,RD 算法的检索性能与 FPCL 算法最为接近,但其查准率、查全率曲线整体低于 FPCL 算法,在查全率大于 70%时,二者的查准率均相差 10%以上.图中显示的其他比较算法在检索效果上均与本文算法存在明显差距.

不同图像类别间的形状特征存在较大差别,这就要求形状描述子具有较强的区分类间差异的能力.为了检测各种算法对不同类型形状的适应性,实验中,从 MPEG-7 SetB 标准图像库中随机抽取 15 类、共 300 幅图像作为检索对象,计算某类中每幅图像的查全率,进而获得该类图像的平均检索率.实验中将返回图像数设置为 40 幅,并选择平均检索率较高的 3 种算法与 FPCL 算法进行比较,结果如图 11 所示.

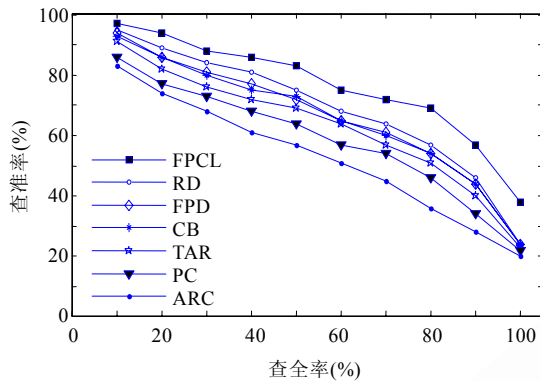


Fig.10 Comparison of P-R Curves in MPEG-7 database

图 10 MPEG-7 图像库中的 P-R 曲线比较

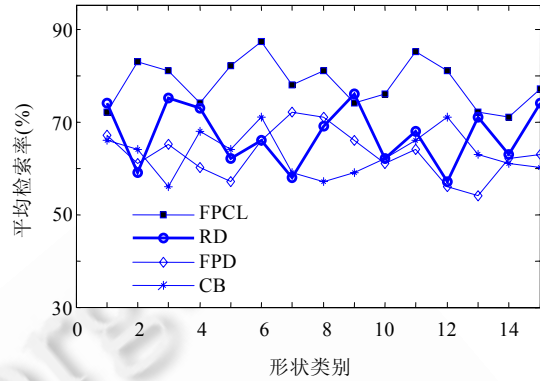


Fig.11 Comparison of retrieval adaptability for images from different classes

图 11 类间图像的检索适应性比较

在实验结果中,FPCL 算法在第 1 类(beetle class)、第 4 类(chicken class)、第 9 类(guitar class)、第 13 类(teddy class)以及第 15 类(horseshoe class)中的表现与 RD 算法相当,并显示出了明显高于 FPD 和 CB 算法的检索性能.从其余形状类的实验结果曲线可知,FPCL 算法较之其他 3 种比较算法均表现出明显优势,并在第 6 类(device class)的检索实验中得到了 87.3%的平均检索率.

为了检测算法在噪声环境下的检索效率,本文在 Set B 的基础上建立了相应的噪声图库.由于检测的目的是衡量形状描述子的噪声鲁棒性,所以噪声图库在 Set B 的 70 种形状类中随机抽取 20 类,对初始形状的轮廓图像追加随机高斯噪声,将加入噪声的信噪比设定为 40dB,35dB,30dB,25dB 和 20dB 这 5 个级别,从而构成由 2 000 幅噪声形状组成的噪声图库.噪声图库的生成过程以及图库中的部分形状样本如图 12 所示.

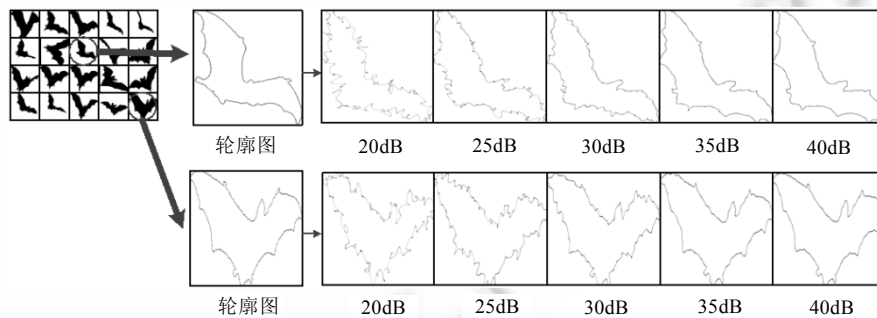


Fig.12 Generation of noisy database

图 12 噪声图库的生成

表 2 显示了各种算法在随机高斯噪声干扰环境下的平均检索率.可以看到,FPD 算法具有一定的噪声适应性,但整体检索效率低于 FPCL 和 RD 算法.RD 算法在噪声强度增长至 20dB 时,具有 65.40%的平均检索率.当噪声强度从 40dB 增加至 20dB 时,FPCL 算法的平均检索率仅下降了 1.74 个百分点,检索效率明显高于其他算法,显示出了较好的噪声适应能力.图 13 中显示了噪声强度为 20dB 时,4 种检索性能较好算法的查全率、查准率曲

线.从图中可知,在相同的查全率下,FPCL 算法的查准率均高于其他算法,表明了该算法在噪声环境下具有良好的鲁棒性.

Table 2 Comparison of experimental results

表 2 实验结果比较

噪声强度		40dB	35dB	30dB	25dB	20dB
比较 方法	FPCL	76.12	75.98	75.81	75.39	74.38
	RD	68.56	68.01	67.29	66.63	65.40
	FPD	64.83	64.49	64.07	63.58	63.02
	CB	63.26	62.65	62.14	61.67	60.19
	TAR	60.80	59.44	57.36	55.42	53.03
	PC	55.12	54.76	54.59	54.12	53.97
	ARC	48.05	46.61	45.19	43.22	41.30
	AC	27.10	26.74	26.33	26.09	25.87

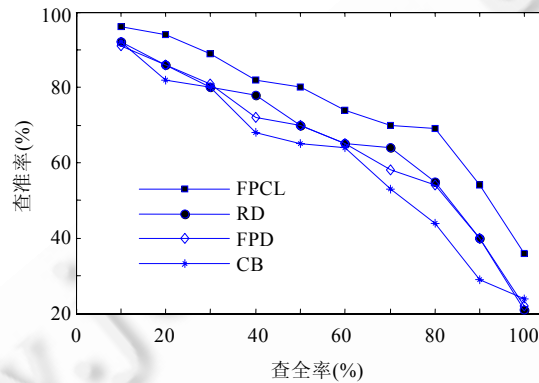


Fig.13 Comparison of P-R curves in noise database

图 13 噪声图库中的 P-R 曲线比较

为了对算法复杂度进行分析,我们统计了 FPCL,RD,FPD 和 CB 在噪声图库上运行时每次检索的 CPU 平均耗时情况,其中,形状描述子提取时间分别为 114.3ms,82.8ms,110.9ms 和 318.1ms,匹配时间分别为 1.539ms,1.451ms,1.533ms 和 2.081ms.该结果显示,CB 的计算复杂度明显高于其他算法,而 RD 由于只考虑了径向距离这一因素,算法耗时较少.FPCL 和 FPD 的算法耗时较为接近,其中,FPCL 需要进行轮廓重构和最优特征点定位,而 FPD 需要在形状全部轮廓点的基础上对最远点进行动态更新,这都导致了算法复杂度的相应增加.从最终的检索结果来看,FPCL 在获得更优检索精度的同时会带来小幅的耗时增加,但该耗时增幅在可以接受的范围之内.

4 结束语

在基于形状的图像检索领域,传统的傅立叶形状描述方法在体现图像的特征信息上存在不足,且噪声干扰带来的图像畸变会明显地降低检索效率.针对以上制约该类技术发展的瓶颈,本文提出一种通过轮廓降维重构和特征点弦长描述函数进行图像检索的策略,在目标轮廓提取的基础上对其进行降维重构,平滑处理后的轮廓更好地体现了目标的形状信息.在此基础上,实现了对噪声轮廓的重构,减少了轮廓畸变对形状描述算法的影响.通过新定义的支持域来确定轮廓点的特征强度,并对特征强度和支持域半径间的关系进行了分析,保证了特征点提取的有效性.提出的特征点弦长函数,有效地描述了目标轮廓的特征,对生成的傅立叶描述子进行归一化处理,得到了满足各种不变性要求的轮廓描述子.最后,本文从样本库整体检索、样本库分类检索以及噪声样本库检索这 3 个角度对算法性能进行了测试评估,本文算法在常规样本实验中,显示出了明显高于传统算法的检索效率.在噪声样本实验中,本文算法在 20dB 随机噪声干扰下的平均检索率为 74.38%,高出其他比较算法至少 8.98 个百分点.大量实验结果显示,本文算法无论是在常规样本环境还是在噪声样本环境中,都具有明显高于

传统算法的检索性能.

References:

- [1] Datta R, Joshi D, Li J, Wang JZ. Image retrieval: Ideas, influences, and trends of the new age. *ACM Computing Surveys*, 2008, 40(2):1–60. [doi: 10.1145/1348246.1348248]
- [2] Berretti S, Del Bimbo A, Pala P. Retrieval by shape similarity with perceptual distance and effective indexing. *IEEE Trans. on Multimedia*, 2000,2(4):225–239. [doi: 10.1109/6046.890058]
- [3] Flusser J. On the independence of rotation moment invariants. *Pattern Recognition*, 2000,33(9):1405–1410. [doi: 10.1016/S0031-3203(99)00127-2]
- [4] Khalid MH. Exact Legendre moment computation for gray level images. *Pattern Recognition*, 2007,40(12):3597–3605. [doi: 10.1016/j.patcog.2007.04.014]
- [5] Papakostas GA, Boutalis YS, Karras DA, Mertzios BG. A new class of Zernike moments for computer vision applications. *Information Sciences*, 2007,177(13):2802–2819. [doi: 10.1016/j.ins.2007.01.010]
- [6] Mukundan R, Ong SH, Lee PA. Image analysis by Tchebichef moments. *IEEE Trans. on Image Processing*, 2001,10(9):1357–1364. [doi: 10.1109/83.941859]
- [7] Zhang H, Shu HZ, Haigron P, Li BS, Luo LM. Construction of a complete set of orthogonal Fourier-Mellin moment invariants for pattern recognition applications. *Image and Vision Computing*, 2010,28(1):38–44. [doi: 10.1016/j.imavis.2009.04.004]
- [8] Zhang DS, Lu GJ. Shape-Based image retrieval using generic Fourier descriptor. *Signal Processing: Image Communication*, 2002, 17(10):825–848. [doi: 10.1016/S0923-5965(02)00084-X]
- [9] Goh WB. Strategies for shape matching using skeletons. *Computer Vision and Image Understanding*, 2008,110(8):326–345. [doi: 10.1016/j.cviu.2007.09.013]
- [10] Ebrahim Y, Ahmed M, Abdelsalam W, Chau SC. Shape representation and description using the Hilbert curve. *Pattern Recognition Letters*, 2009,30(4):348–358. [doi: 10.1016/j.patrec.2008.09.013]
- [11] Kauppinen H, Seppanen T, Pietikainen M. An experimental comparison of autoregressive and Fourier-based descriptors in 2D shape classification. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1995,17(2):201–207. [doi: 10.1109/34.368168]
- [12] Kunttu I, Lepisto L, Rauhamaa J, Visa A. Multiscale Fourier descriptors for defect image retrieval. *Pattern Recognition Letters*, 2006,27(2):123–132. [doi: 10.1016/j.patrec.2005.08.022]
- [13] Abbasi S, Mokhtarian F, Kittler J. Enhancing CSS-based shape retrieval for objects with shallow concavities. *Image and Vision Computing*, 2000,18(3):199–211. [doi: 10.1016/S0262-8856(99)00019-0]
- [14] Abbasi S, Mokhtarian F, Kittler J. Curvature scale space image in shape similarity retrieval. *Multimedia Systems*, 1999,7(6): 467–476. [doi: 10.1007/s005300050147]
- [15] Li ZM, Lu TB, Sang XY, Qin BS. Shape description based on axis of least inertia and chain. *Journal on Communications*, 2009, 30(4):1–5 (in Chinese with English abstract). [doi: 10.3321/j.issn:1000-436X.2009.04.001]
- [16] Zhang DS, Lu G. A comparative study of curvature scale space and Fourier descriptors for shape-based image retrieval. *Journal of Visual Communication and Image Representation*, 2003,14(1):41–60. [doi: 10.1016/S1047-3203(03)00003-8]
- [17] Shu X, Wu XJ. A novel contour descriptor for 2D shape matching and its application to image retrieval. *Image and Vision Computing*, 2011,29(4):286–294. [doi: 10.1016/j.imavis.2010.11.001]
- [18] Xie BW, Wang JJ. A contour-based image retrieval algorithm. *Journal of Image and Graphics*, 2008,13(7):1367–1373 (in Chinese with English abstract).
- [19] Zhang DS, Lu GJ. Study and evaluation of different Fourier methods for image retrieval. *Image and Vision Computing*, 2005,23(1): 33–49. [doi: 10.1016/j.imavis.2004.09.001]
- [20] Kunttu I, Lepisto L. Shape-Based retrieval of industrial surface defects using angular radius Fourier descriptor. *IET Image Processing*, 2007,1(2):231–236. [doi: 10.1049/iet-ipr:20060113]
- [21] Urdiales C, Bandera A, Sandoval F. Non-Parametric planar shape representation based on adaptive curvature functions. *Pattern Recognition*, 2002,35(1):45–53. [doi: 10.1016/S0031-3203(01)00041-3]

- [22] El-Rube I, Alajlan N, Kamel MS, Ahmed M, Freeman GH. MTAR: A robust 2D shape representation. *Int'l Journal of Image and Graphics*, 2006,6(3):421-443. [doi: 10.1142/S021946780600232X]
- [23] El-Ghazal A, Basir O, Belkasim S. Farthest point distance: A new shape signature for Fourier descriptors. *Signal Processing: Image Communication*, 2009,24(7):572-586. [doi: 10.1016/j.image.2009.04.001]
- [24] Mackworth SK, Mokhtarian F. The renormalized curvature scale space and the evolution properties of planar curves. In: *Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition*. Vancouver: University of British Columbia, 1988. 318-326. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=196255 [doi: 10.1109/CVPR.1988.196255]
- [25] Zhang ZY, Shi ZP, Shi ZZ. Image retrieval based on contour. *Ruan Jian Xue Bao/Journal of Software*, 2008,19(9):2461-2470 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/19/2461.htm> [doi: 10.3724/SP.J.1001.2008.02461]
- [26] Chan DY, Hsu C. Robust shape-preserving contour tracing with synchronous redundancy pruning. *Pattern Recognition Letters*, 2008,29(5):569-579. [doi: 10.1016/j.patrec.2007.11.005]
- [27] Llaría B, Paolo C, Marco P. WARP: Accurate retrieval of shapes using phase of Fourier descriptors and time warping distance. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2005,27(1):142-147. [doi: 10.1109/TPAMI.2005.21]
- [28] Nguyen TP, Debled I. A discrete geometry approach for dominant point detection. *Pattern Recognition*, 2011,44(1):32-44. [doi: 10.1016/j.patcog.2010.06.022]
- [29] Huang CL, Huang DH. A content-based image retrieval system. *Image and Vision Computing*, 1998,16(3):149-163. [doi: 10.1016/S0262-8856(97)00062-0]
- [30] Wang B. Shape description using arc-height radius complex function. *Acta Electronica Sinica*, 2011,39(4):831-836 (in Chinese with English abstract).
- [31] El-ghazal A, Basir O, Belkasim S. Invariant curvature-based Fourier shape descriptors. *Journal of Visual Communication and Image Representation*, 2012,23:622-633. [doi: 10.1016/j.jvcir.2012.01.011]

附中文参考文献:

- [15] 李宗民,陆天波,桑鑫焱,秦宝山.基于最小惯量轴及链码的图像形状描述方法. *通信学报*,2009,30(4):1-5. [doi: 10.3321/j.issn:1000-436X.2009.04.001]
- [18] 谢邦旺,王加俊.一种基于轮廓的图像检索算法. *中国图像图形报*,2008,13(7):1367-1373.
- [25] 张志勇,施智平,石志勇,史忠植.基于轮廓的图像检索. *软件学报*,2008,19(9):2461-2470. <http://www.jos.org.cn/1000-9825/19/2461.htm> [doi: 10.3724/SP.J.1001.2008.02461]
- [30] 王斌.一种用于形状描述的拱高半径复函数. *电子学报*,2011,39(4):831-836.



师文(1984—),男,河南郑州人,硕士,CCF 学生会员,主要研究领域为计算机视觉,模式识别,多媒体信息处理.
E-mail: wens163@163.com



朱学芳(1962—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为模式识别,多媒体信息处理,信息检索.
E-mail: xfzhu@nju.edu.cn