

目标边界点集的层次化描述及其形状检索应用*

刘 锋¹, 王 斌^{1,2}

¹(南京财经大学 信息工程学院, 江苏 南京 210023)

²(江苏省电子商务重点实验室(南京财经大学), 江苏 南京 210023)

通讯作者: 王斌, E-mail: wangbin@njue.edu.cn



摘 要: 提出用于轮廓线形状和区域形状图像检索的形状描述方法,该方法将目标形状的边界(包括内边界)表示为一个无序的点集,沿各方向对点集的迭代分割,建立层次化的边界点集描述模型.通过对各层形状边界的分割比和分散度的几何特征度量,产生各层的形状特征描述,对它们进行组合,建立对目标形状的层次化描述.两个目标形状的差异性度量定义为它们的层次化描述子的 L-1 距离.该方法具有:(1) 通用性.能够描述轮廓线形状和区域形状这两种不同类型的形状;(2) 可扩展性.基于所提出的分层描述框架,可以将分割比和分散度这两种几何度量进行扩展,纳入更多其他几何特征度量,以进一步提高形状描述的精度;(3) 多尺度描述特性.提出的分层的描述机制,使得描述子具有内在的由粗到细的形状表征能力;(4) 较低的计算复杂性.由于仅仅计算目标图像的边界像素点,使得算法具有较高的计算效率.用 MPEG-7 CE-2 区域形状图像库和 MPEG-7 CE-1 轮廓线形状图像库这两个标准测试集对该方法进行评估,并与同类的其他形状描述方法进行比较,实验结果表明:提出的方法在综合考虑检索精确率、检索效率和一般应用能力等指标的情况下,其性能上要优于各种参与比较的方法.

关键词: 图像分析;边界点集;层次化描述;形状检索

中图法分类号: TP391

中文引用格式: 刘锋,王斌.目标边界点集的层次化描述及其形状检索应用.软件学报,2019,30(9):2886–2903. <http://www.jos.org.cn/1000-9825/5535.htm>

英文引用格式: Liu F, Wang B. Hierarchical point-set description of object edge and its application in shape retrieval. Ruan Jian Xue Bao/Journal of Software, 2019, 30(9): 2886–2903 (in Chinese). <http://www.jos.org.cn/1000-9825/5535.htm>

Hierarchical Point-set Description of Object Edge and Its Application in Shape Retrieval

LIU Feng¹, WANG Bin^{1,2}

¹(School of Information Engineering, Nanjing University of Finance and Economics, Nanjing 210023, China)

²(Jiangsu Provincial Key Laboratory of Electronic Business (Nanjing University of Finance and Economics), Nanjing 210023, China)

Abstract: A novel shape description which can be generally applied to both contour shape and region shape recognition is proposed in this study. This method treats the edge (including the inner edge) of the object as an unordered point-set, a hierarchical description model is built by iteratively partitioning the edge of the object into progressively smaller parts along different directions. At each level of the hierarchy structure, the geometrical features of the object edge are characterized by two measurements, partition ratio, and dispersion degree. Combining them, a hierarchical description of the object shape can then be constructed. The dissimilarity of two shapes can be

* 基金项目: 国家重点研发计划(2017YFD0700501); 国家自然科学基金(61372158); 江苏省自然科学基金(BK20181414); 江苏省高校优秀科技创新团队项目(2017-15)

Foundation item: National Key R&D Program of China (2017YFD0700501); National Natural Science Foundation of China (61372158); Natural Science Foundation of Jiangsu Province of China (BK20181414); Program for Outstanding Science and Technology Innovation Team of Jiangsu Higher Education Institutions (2017-15)

收稿时间: 2017-09-26; 修改时间: 2017-10-26; 采用时间: 2017-12-15; jos 在线出版时间: 2019-01-21

CNKI 网络优先出版: 2019-01-22 13:49:00, <http://kns.cnki.net/kcms/detail/11.2560.TP.20190122.1348.008.html>

measured by computing the L-1 distance between their hierarchical shape descriptors. The merits of the proposed method can be summarized as follows. (1) Both contour shape and region shape can be effectively described by this method, thus it has the ability for general use. (2) Based on the proposed hierarchical description framework, besides the proposed two measures, partition ratio, and dispersion degree, many other measures can be included for meeting various accuracy requirements on shape recognition, so the proposed method has extendibility. (3) The proposed hierarchical description scheme make the available descriptors characterize the shape from coarse to fine, so the proposed descriptor is multi-scale. (4) Instead of using all the pixel points of the object, the proposed method only take the edge points of the object into account, for this reason, it has a relative low computational complexity. Two standard test sets of MPEG-7 CE-2 region shape database and MPEG-7 CE-1 contour shape database are used to evaluate the performance of the proposed method. The experimental results indicate that the proposed method outperforms the state-of-the-art approaches in terms of a comprehensive consideration on the retrieval rates, retrieval efficiency, and general application ability.

Key words: image analysis; edge point-set; hierarchical description; shape retrieval

近些年来,随着数字信息技术的快速发展、图像数据库规模不断扩大,为有效管理和利用海量的图像数据,基于内容的图像检索(CBIR)^[1]技术在计算机视觉领域得到了广泛的研究与应用.通过计算机软件对图像的特征进行自动识别和提取,以用于图像的检索和分类^[2].由于形状是大多数物体固有的视觉属性,是人类视觉系统对物体进行识别和分类的重要依据,因此,形状分析成为CBIR研究领域十分活跃的分支,并且已经在文本分析、神经科学、农业、生物医药、工程技术等领域产生了大量的应用^[3,4].

在图像处理中,对于从场景中分割出来的目标形状,我们可以将其表示为分布在目标区域的所有像素点的集合,如果可以将其简化成一条仅由边界像素点构成的曲线,而且形状信息能够完全保持(该闭合曲线所围成的区域即为目标区域),我们称该类形状为轮廓线形状.如果形状存在不连通的目标区域,或区域是连通的,但目标区域内有孔洞,使其具有复杂的内部结构,则该类形状为区域形状.我们在图1中给出了形状分类框图以及两类形状的示例图像和对应的边界图,以便于观察这两类形状的特点,其中,轮廓线形状和区域形状的示例分别取自MPEG-7 CE-1 轮廓线形状图像库和 MPEG-7 CE-2 区域形状图像库.

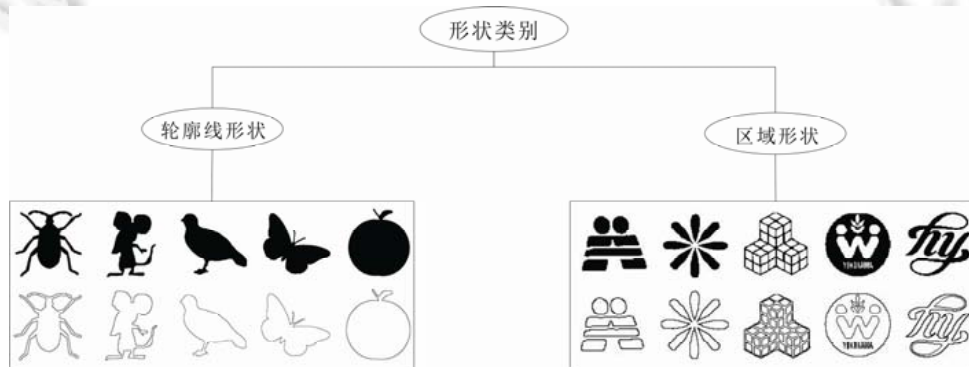


Fig.1 Classification of shapes and its example images and edge maps

图1 形状的分类及其示例图像和边界图

形状描述是形状分析的核心步骤,旨在提取目标形状的有效特征,以用于后续目标形状的匹配、检索和分类等形状分析任务^[4].现有的形状描述方法可分为基于曲线的描述方法和基于点集的描述方法两大类.

基于曲线的描述方法根据目标边界像素点的邻接关系进行轮廓跟踪,从而得到一条轮廓线曲线,即轮廓的有序点集.这类方法近些年得到广泛和深入的研究,研究成果^[5-23]如傅里叶描述子^[13,14]和曲率尺度空间^[15,16],二者是经典的轮廓线形状描述子,后者被MPEG-7推荐为标准的轮廓线形状描述子,其在MPEG-7 CE-1形状测试集上获得了75.44%的检索准确率^[15].又如多尺度凹凸性^[5]、内部距离^[6]、三角形面积^[7]、轮廓柔韧性^[8]、形状树^[17]、分段多项式描述子^[18]、局部仿射不变描述子^[19]以及一些其他轮廓线描述子^[9-12,20-23].这类方法可借助曲线分析的数学工具,如傅里叶变换、小波变换、微分几何等,产生许多具有强辨识力的形状描述子,还可以利用

深度学习的方法对特征进行有效的选取^[24,25].而在形状匹配阶段,可以借助像素点的有序信息进行轮廓线的精确匹配,如采用动态规划的匹配方法,又如利用距离度量学习的方法进行形状相似度的度量^[26].该类方法形式优美,性能优越,虽然这些方法在 MPEG-7 CE-1 轮廓线形状测试集上报告的检索精确率大都超过了 85%,但这些方法是专门为轮廓线形状设计的,需要把目标形状的轮廓表示为一条曲线才能进行特征抽取,因而只能处理图 1 中的轮廓线形状.但对于图 1 中的区域形状,由于其轮廓不能用一条曲线来进行表示,该类方法则无能为力.综上所述,我们可以看出基于曲线的描述方法因对目标形状轮廓的提取质量有着严重的依赖性,不能处理具有复杂内部结构的区域形状,因此具有很大的局限性,不能用于一般的形状识别任务.

基于点集的方法是另一类形状描述方法,该类方法将目标形状区域的像素点看成一般的点集,抽取点集的几何特性,产生形状描述子.该类方法又可进一步地分为基于区域点集的描述方法和基于边界点集的描述方法:前者在抽取形状的特征时,考虑目标形状区域的所有像素点,即将所有的目标像素点看成一个点集;而后者则仅考虑目标形状的边界像素点,将边界像素点构成一个点集来进行特征抽取.我们在图 2 中给出了形状描述方法的分类框图.由基于点集的形状描述机制可知,基于区域点集的描述方法和基于边界点集的描述方法都能够描述轮廓线形状和区域形状.

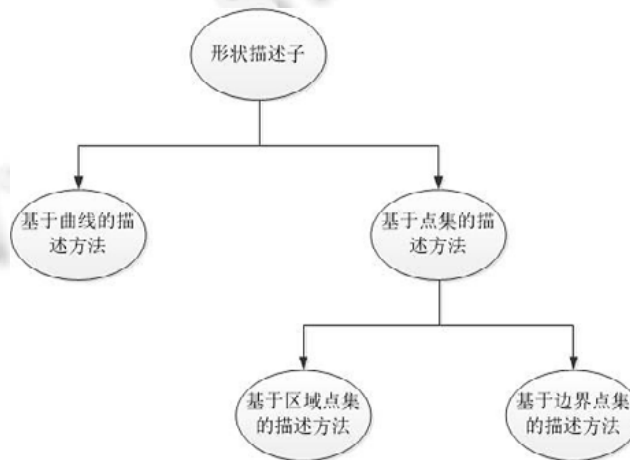


Fig.2 Classification of shape description methods

图 2 形状描述方法的分类

人类的视觉系统对目标的边界信息非常敏感.从图 1 中形状的边界图可以看出,无论是轮廓线形状还是区域形状,在移去目标区域内的所有像素点,仅保留边界像素点的情况下,人的视觉系统仍然能对目标形状进行准确辨识.因此,目标形状的边界像素点为我们抽取具有强辨识力的形状特征提供了重要的线索.相较于基于曲线的描述方法,基于边界点集的描述方法克服了前者仅能处理轮廓线形状的限制性,能够一般地处理两类目标形状,且由于不需要对边界点进行序化操作,从而避免了因序化操作所带来的不稳定性.而相较于基于区域点集的描述方法,基于边界点集的描述方法由于仅需处理边界像素点,因此降低了点集的基数,从而具有非常高的计算效率;且其直接从边界提取特征,使得描述子更具辨识能力.因此,本文着眼于基于边界点集的形状描述方法.其主要贡献在于:提出了一种目标边界的分层描述模型,通过对目标边界从多个方向的分层度量,多尺度地抽取了目标形状几何特征,无论是在标准的 MPEG-7 CE-2 区域形状测试集上,还是在 MPEG-7 CE-1 轮廓线形状测试集上,该方法都以较低的计算成本,取得了比同类方法(基于点集的描述方法)更高的形状检索精确率.

1 相关工作

本文提出的是一种新的基于边界点集的形状描述方法,所以,与本文联系最紧密的工作是近年来一些基于点集的形状描述方法.本节中,我们对一些常见的基于点集的方法按基于区域点集和基于边界点集的分类展开

综述.

常用的基于区域点集的形状描述方法有几何不变矩^[27]、Zernike 矩(Zernike moments,简称 ZM)^[28]、伪 Zernike 矩^[29]等各种基于矩的形状描述方法和通用傅里叶描述子(generic Fourier descriptor,简称 GFD)^[30]等基于傅里叶变换的形状描述方法.在众多基于矩的形状描述方法中,Zernike 矩^[28]是一种经典的基于区域点集的形状描述方法.它将图像投影到一组定义在单位圆上的基于 Zernike 多项式的正交化函数,矩的大小用以生成对图像进行描述的特征向量.Zernike 矩能够很容易地构造图像的任意高阶矩,并能够使用较少的矩来重建图像.虽然其计算比较复杂,但是 Zernike 矩在图像旋转和噪声敏感度方面具有较大的优越性.由于 Zernike 矩具有图像旋转不变性和较低的噪声敏感度,且可以构造任意高阶矩,所以被广泛地应用于模式识别等领域中,MPEG-7 标准中已将 Zernike 矩列为一种标准的区域描述符^[31].基于 Zernike 矩,文献[32]提出一种 Zernike 矩边缘梯度方法(Zernike moments & edge gradient,简称 ZMEG)用于商标图像检索,这种方法将 Zernike 矩作为目标形状的全局特征,提取形状的边缘梯度共生矩阵作为局部特征,结合全局特征和局部特征对形状进行描述.文献[30]提出一种通用傅里叶描述子,该方法首先对图像进行预处理,将原始图像变换为极坐标图像,再对其进行二维傅里叶变换,用其变换系数的模作为描述子.该方法满足对目标形状的旋转、缩放和平移的不变性,适合于一般的形状图像检索应用.近年来,Yang 等人^[33]提出一种自适应分层质心的算法,这种方法递归地对形状区域进行分割,通过若干次分割,将形状区域分割成越来越小的区块,每一次递归,计算当前的形状区块的质心,根据得到的质心,将当前区块分割成 4 个更小的区块,将每次递归产生的区块质心坐标构成的集合,作为描述形状的特征向量.基于这种对图像递归地进行分块的技术,Sidiropoulos 等人^[34]提出了一种自适应分层密度直方图(adaptive hierarchical density histogram,简称 AHDH)的方法,该方法用每一个形状区块的密度特征代替质心坐标来产生特征向量,以此来表示形状区块像素点的分布特性.这两种方法在对形状进行分割时,分割的方向依赖于目标形状所处的坐标系,因此不满足旋转不变性.此外,随着递归次数的增加,区块数量会呈指数级增长,导致计算的复杂度很高,难以满足实时性需求.

形状上下文(shape contexts,简称 SC)^[35]是经典的基于边界点集的形状分析方法,该方法将目标形状的边界像素点重新采样成指定个数的点集(一般 100 个~300 个点),将这个点集中的每一个点分别作为参考点,通过在该点放置一个极坐标栅格来统计点集中的其他各点相对于该点的空间分布,并产生直方图作为该点的描述子,也称为该点的形状上下文.该方法有效地抽取了边界上点的空间分布信息和相对位置关系,抽取的形状特征具有较强的辨识力,在 MPEG-7 CE-1 形状测试集上取得了 76.51%的识别准确率.距离集(distance set,简称 DS)描述方法^[36]是另一种基于边界点集的形状描述方法,该方法对目标形状边界上的点重新采样得到 N 个点的点集,对点集中的每一个点,计算该点与其最邻近的 $n(n \leq N)$ 个点的距离,构成一个距离集.经过归一化处理的距离集可以用来描述该点与其临近各点的空间关系,而由各点的距离集构成的集合,即距离集的集合,构成了描述整个点集的空间安排.该方法在 MPEG-7 CE-1 形状测试集上取得了 78.38%的检索精确率.形状上下文和距离集在形状匹配阶段都需要计算点到点的优化问题,该优化问题可归结为经典的二次指派问题,而求解该问题的匈牙利算法^[37]的时间复杂度为 $O(N^3)$,这里, N 表示点集的规模.根据文献[35]的报告,形状上下文方法在执行一次两个形状的匹配时,在对每个形状仅采样 100 个点的情况下,便需要耗费 0.2s.而根据文献[36]的报告,距离集方法在执行形状匹配时,在采样 250 个点的情况下,需要耗费 0.7s.很显然,考虑到匹配算法的计算复杂度,在使用该类方法时,对轮廓点进行重新采样得到的点集规模不能太大(一般不超过 300 个点).而对于具有复杂内部结构的区域形状,如果对边界点重采样构成的点集规模太小,显然不足以描述形状的复杂结构信息.形状上下文和距离集虽然能处理区域形状,但是由于二者计算效率低,所以有着很大的局限性.文献[38]提出一种称为轮廓点分布直方图(contour points distribution histogram,简称 CPDH)的方法,该方法将极坐标栅格放置于整个形状的质心,从而得到描述整个形状轮廓像素点空间关系的轮廓点分布直方图.形状的相似性用 EMD(earth mover's distance)距离进行度量.该方法实质上是一种全局描述子,因不需要为点集中的每一个点进行特征描述,因此无论在形状描述还是匹配阶段,都减少了计算量.该方法在 MPEG-7 CE-1 形状测试集上取得了 76.56%的检索精确率.

近年来,一些研究工作将复杂网络模型应用于形状分析当中.

文献[39]提出,将目标形状的边界建模为一个小世界复杂网络.在该模型中,形状边界上的像素点对应网络上的节点,像素点间的连线对应网络上的边,像素点间的归一化欧氏距离作为边的权值.通过网络的动态演化,提取各时刻网络的度特征和联合度特征,组合成一个特征向量作为描述子,在形状匹配阶段,则通过特征向量间的欧氏距离来确定形状间相似度.文献[40]中提出一种无序点集描述方法(unordered point-set description,简称 UPSD),该方法将目标形状边界看成无序点集,提出一种新的基于复杂网络模型的目标边界无序点集形状分析方法.该方法用自主的网络动态演化机制分层地提取形状特征,通过对网络的局部测量和全局测量,获取网络的无权特征和加权特征,构成形状的局部描述子和全局描述子.该描述子为目标形状的认识提供了具有强辨识能力的特征,在形状匹配阶段,用较低计算复杂度的 Hausdorff 距离和 L-1 距离分别作为局部距离和全局距离,从而节省了时间成本.该方法在 MPEG-7 CE-1 形状测试集上取得了 78.18% 的检索精确率.

2 边界点集的分层描述模型

对于一个目标形状的图像,我们首先抽取目标形状的边界像素点,得到边界的无序点集 B_0 ,该点集由边界像素点的坐标构成,用 (x,y) 表示像素点的坐标,则目标形状边界的中心点 (\bar{x}_0, \bar{y}_0) 可以定义为

$$\bar{x}_0 = \frac{\sum_{(x,y) \in B_0} x}{|B_0|}, \bar{y}_0 = \frac{\sum_{(x,y) \in B_0} y}{|B_0|} \quad (1)$$

其中, $| \cdot |$ 表示集合的基数.我们用中心点 (\bar{x}_0, \bar{y}_0) 对边界 B_0 进行位置归一,从而得到:

$$\bar{B}_0 = \{(x - \bar{x}_0, y - \bar{y}_0) | (x, y) \in B_0\} \quad (2)$$

再将边界 \bar{B}_0 顺时针旋转一个角度 θ , 这里 $\theta \in [0, 2\pi)$, 此时的边界可以表示为

$$\bar{B}_0^\theta = \{(x \cos \theta + y \sin \theta, y \cos \theta - x \sin \theta) | (x, y) \in \bar{B}_0\} \quad (3)$$

记边界 \bar{B}_0^θ 的中心点 $(\bar{x}_0^\theta, \bar{y}_0^\theta)$, 显然, 我们有 $\bar{x}_0^\theta = 0$ 和 $\bar{y}_0^\theta = 0$. 通过该中心点, 我们可以对整个形状边界 \bar{B}_0^θ 进行第 1 层分割, 得到如下集合:

$$\bar{B}_1^\theta = \{(x, y) | x \leq \bar{x}_0^\theta, y \geq \bar{y}_0^\theta, (x, y) \in \bar{B}_0^\theta\} \quad (4)$$

显然, \bar{B}_1^θ 是 \bar{B}_0^θ 的一个子集, 我们称其为子边界.

同样, 我们进一步计算子边界 \bar{B}_1^θ 的中心点坐标 $(\bar{x}_1^\theta, \bar{y}_1^\theta)$, 并通过该中心点对子边界 \bar{B}_1^θ 进行与公式(4)类似的分割, 得到子边界 \bar{B}_2^θ .

如此迭代地分割下去, 将经过 l 层分割得到的边界记为 \bar{B}_l^θ , 则该边界可以表示为

$$\bar{B}_l^\theta = \{(x, y) | x \leq \bar{x}_{l-1}^\theta, y \geq \bar{y}_{l-1}^\theta, (x, y) \in \bar{B}_{l-1}^\theta\} \quad (5)$$

其中, $(\bar{x}_{l-1}^\theta, \bar{y}_{l-1}^\theta)$ 是子边界 \bar{B}_{l-1}^θ 的中心点坐标. 显然有 $\bar{B}_0^\theta \supset \bar{B}_1^\theta \supset \bar{B}_2^\theta \supset \dots \supset \bar{B}_l^\theta$.

我们对 θ 在区间 $[0, 2\pi)$ 上均匀地采样 m 个角度, 用 $j=0, 1, \dots, m-1$ 表示它们的索引. 让变量 θ 依次取这 m 个值, 这样我们可以对目标形状边界 \bar{B}_0 沿 m 个方向进行 l 层分割, 得到 $m \times l$ 个子边界, 构成集合:

$$\Psi(m, l) = \{\bar{B}_l^j, i=1, 2, \dots, l, j=0, 1, \dots, m-1\} \quad (6)$$

我们称其为目标形状边界的分层描述模型, 其中, m 和 l 为模型的参数, 分别表示该描述模型的方向角的个数和分层的层级数.

图 3 和图 4 给出了目标形状边界的分层描述模型示意图, 其中, 图 3 为对区域形状边界的分层描述, 图 4 为对轮廓线形状边界的分层描述, 模型的参数取 $m=5, l=3$.

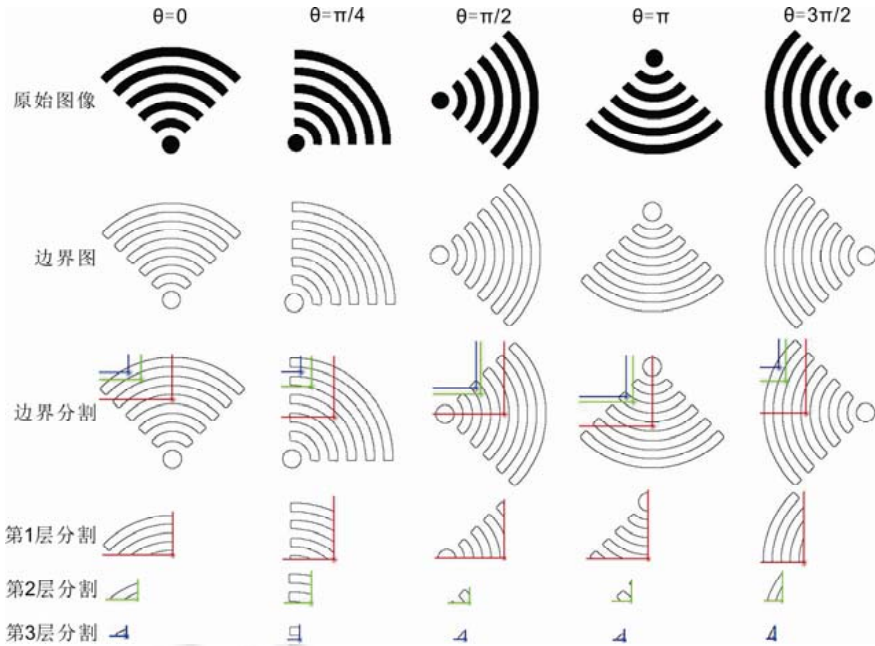


Fig.3 Visual illustration of iteratively partitioning the edge of region-based image in progressively smaller part along different directions

图3 沿不同方向对区域形状的边界迭代地分割示意图

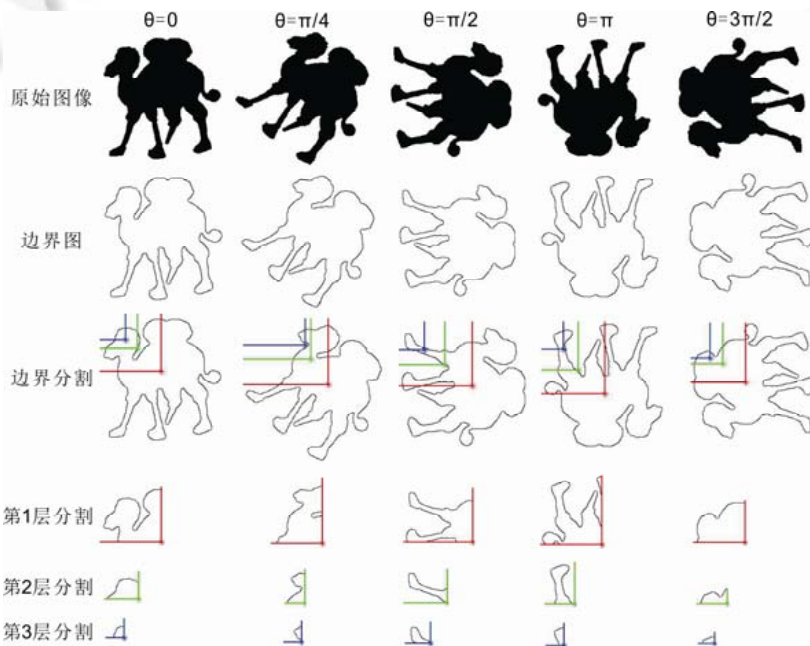


Fig.4 Visual illustration of iteratively partitioning the edge of contour-based image in progressively smaller part along different directions

图4 沿不同方向对轮廓线形状边界迭代地分割示意图

3 分层度量方法

第2节定义了描述目标形状边界的多层描述模型 $\Psi(m, l) = \{\bar{B}_i^j, i=1, 2, \dots, l, j=0, 1, \dots, m-1\}$. 在本节, 我们给出对子边界 \bar{B}_i^j 的几何度量方法, 通过该度量来抽取形状描述子.

3.1 分割比

如前所述, 当前层边界是对上一层边界进行一次分割得到的, 分割后的边界是上一层边界的一个子集, 当前层边界上的像素点在上一层边界中占的比重, 构成了当前层边界对上一层边界的几何分割的一个度量, 这里我们将该特征称为分割比, 定义为

$$d_i^j = \frac{|\bar{B}_i^j|}{|\bar{B}_{i-1}^j|} \tag{7}$$

图5给出了边界分割比的示意图, 其中, 第1行和第2行分别表示图3和图4中的区域形状和轮廓线形状在 j 取 0、 i 分别取 1, 2, 3 时的分割比特征.

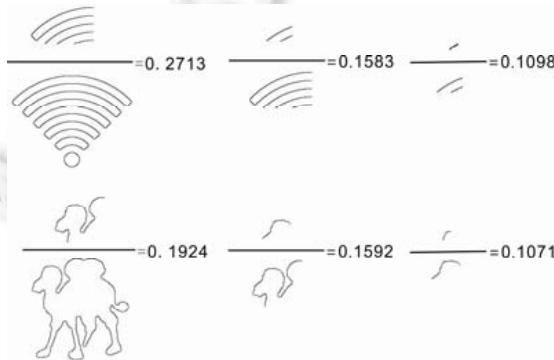


Fig.5 Visual illustration of partition ratio

图5 形状边界的分割比特征示意图

3.2 分散度

度量边界点在二维坐标平面上的分散程度, 是另一类刻画形状边界几何特性的有用特征. 为描述这一特性, 我们对子边界 \bar{B}_i^j 进行如下度量:

首先, 计算子边界 $\bar{B}_i^j : \{(x_k, y_k), k=1, 2, \dots, |\bar{B}_i^j|\}$ 中的每一个点到上一层子边界 \bar{B}_{i-1}^j 的中心点 $(\bar{x}_{i-1}^j, \bar{y}_{i-1}^j)$ 的距离, 构成距离序列:

$$\Omega = \{c_k = \sqrt{(x_k - \bar{x}_{i-1}^j)^2 + (y_k - \bar{y}_{i-1}^j)^2}, k=1, 2, \dots, |\bar{B}_i^j|\} \tag{8}$$

计算该序列的均值:

$$\rho_i^j = \frac{\sum_{k=1}^{|\Omega|} c_k}{|\Omega|} \tag{9}$$

作为对子边界 \bar{B}_i^j 的一个度量, 再进一步计算该序列的方差:

$$\sigma_i^j = \frac{\sum_{k=1}^{|\Omega|} \left(\frac{c_k}{\max(\Omega)} - \rho_i^j \right)^2}{|\Omega|} \tag{10}$$

得到对子边界 \bar{B}_i^j 的另一度量. 公式(9)和公式(10)中的距离 c_k , 我们用距离序列的最大值对其进行了归一化处理. 这两个度量是以上一层子边界的中心点为参照点, 描述 \bar{B}_i^j 中的点的空间分布的统计特征.

离心率是一种被广泛应用的形状特征,它反映了目标形状像素点绕着中心点的分散程度^[41].这里,我们定义子边界 \bar{B}_i^j 上点的离心率,以子边界 \bar{B}_i^j 的当前层的中心点为参考点来度量该边界上点的分散度.

首先计算子边界 \bar{B}_i^j 的中心矩:

$$\mu_{p,q} = \sum_{(x,y) \in \bar{B}_i^j} (x - \bar{x}_i^j)^p (y - \bar{y}_i^j)^q \quad (11)$$

这里, $(\bar{x}_i^j, \bar{y}_i^j)$ 表示 \bar{B}_i^j 的中心.由中心矩 $\mu_{p,q}$,我们可以得到矩阵:

$$U = \begin{bmatrix} \mu_{2,0} & \mu_{1,1} \\ \mu_{1,1} & \mu_{0,2} \end{bmatrix} \quad (12)$$

而边界的主惯性矩是 U 的特征值,这样,边界 \bar{B}_i^j 的离心率可表示为

$$e_{cc} = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} \quad (13)$$

这里, λ_{\max} 和 λ_{\min} 是 U 的特征值中的最大和最小值.这种描述方法仅仅取决于形状本身,无关于形状的大小和方向.从公式(13)可以看出 $e_{cc} > 1$,这里,我们取 \bar{B}_i^j 离心率的倒数来对 \bar{B}_i^j 进行度量,即:

$$e_i^j = \frac{1}{e_{cc}} \quad (14)$$

组合上述定义的两类分散度量,得到特征向量:

$$s_i^j = [\rho_i^j \quad \sigma_i^j \quad e_i^j] \quad (15)$$

作为描述子边界 \bar{B}_i^j 上点的分散特性的描述子.

总结给出的分层度量方法具有如下优点.

- (1) 具有形状描述的一般性.该度量方法既可以描述轮廓线形状,又可以描述区域形状;
- (2) 具有特征抽取的可拓展性.除了分割比和分散度,我们可以将更多其他的对目标边界的几何度量纳入到给出的分层描述模型,从而满足不同精度的检索需求;
- (3) 具有对目标形状的多尺度描述性.本文提出的分层描述机制使得该形状描述子具有内在的由粗到细的多尺度表征能力;
- (4) 较低的计算复杂性.由于我们在特征提取时仅仅考虑目标形状的边界像素点,使得给出的分层度量方法具有较高的计算效率.

4 形状相似度量方法

形状相似度量度的主要任务是测量一副检索图像的特征向量与图像数据集中图像特征向量之间的差异.本节基于定义的目标形状边界点集的分层描述方法,提出一种循环移位的特征匹配方法,以度量两个目标形状的相似度.

4.1 形状描述子的构造

用我们前面提出的形状边界分层描述模型 $\Psi(m,l) = \{\bar{B}_i^j, i=1,2,\dots,l, j=0,1,\dots,m-1\}$ 以及对每一个子边界 \bar{B}_i^j 的分割比度量 d_i^j 和分散度量 s_i^j ,我们可以构建矩阵 D 和矩阵 S :

$$D = \begin{bmatrix} d_1^0 & d_2^0 & \dots & d_l^0 \\ d_1^1 & d_2^1 & \dots & d_l^1 \\ \vdots & \vdots & \ddots & \vdots \\ d_1^{m-1} & d_2^{m-1} & \dots & d_l^{m-1} \end{bmatrix}, S = \begin{bmatrix} s_1^0 & s_2^0 & \dots & s_l^0 \\ s_1^1 & s_2^1 & \dots & s_l^1 \\ \vdots & \vdots & \ddots & \vdots \\ s_1^{m-1} & s_2^{m-1} & \dots & s_l^{m-1} \end{bmatrix} \quad (16)$$

显然,矩阵 D 的规模为 $m \times l$,矩阵 S 的规模为 $m \times 3l$ (因为 $s_i^j = [\rho_i^j \quad \sigma_i^j \quad e_i^j]$).这两个矩阵从 m 个方向和 l 个尺度,分别描述了目标形状边界的分割比特性和边界点的分散度.

组合矩阵 D 和矩阵 S ,产生一个对目标形状进行综合描述的规模为 $m \times 4l$ 的矩阵 F_0 :

$$F_0 = [\omega D \quad (1-\omega)S] \quad (17)$$

其中, $\omega \in [0, 1]$ 为权重变量, 用以调节描述矩阵 D 和 S 在形状相似性度量中的贡献.

4.2 循环移位特征匹配

我们将特征矩阵 F_0 的每一行表示成一个行向量 $V_i, i=0, 1, \dots, m-1$, 这样, 特征矩阵 F_0 就可以改写成一个列向量的形式, 即:

$$F_0 = \begin{bmatrix} V_0 \\ V_1 \\ \vdots \\ V_{m-1} \end{bmatrix} \quad (18)$$

需要指出的是, 当目标形状发生旋转时, 我们对目标进行分割的起始方向会发生改变, 使得方向序列 (长度为 m) 会发生循环移位. 为此, 对于一个待识别的目标形状 A , 在将其与数据集里的形状进行匹配之前, 我们先准备其特征矩阵 F_0 以及 F_0 的 $m-1$ 个循环移位后的版本:

$$F_1 = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_{m-1} \\ V_0 \end{bmatrix}, F_2 = \begin{bmatrix} V_2 \\ \vdots \\ V_{m-1} \\ V_0 \\ V_1 \end{bmatrix}, F_3 = \begin{bmatrix} V_3 \\ \vdots \\ V_0 \\ V_1 \\ V_2 \end{bmatrix}, \dots, F_{m-1} = \begin{bmatrix} V_{m-1} \\ V_0 \\ V_1 \\ \vdots \\ V_{m-2} \end{bmatrix} \quad (19)$$

这样, 形状 A 与数据集里面的形状 B 之间的差异度可以定义为

$$dis(A, B) = \min_{j=0, 1, \dots, m-1} \|F_j^A - F_0^B\| \quad (20)$$

这里, $\|\cdot\|$ 表示 L-1 距离.

5 算法性能分析

本节对前面提出的形状描述方法和特征匹配算法的不变性和计算时间复杂度进行分析.

5.1 不变性分析

因在进行分层描述之前已将坐标原点移到了形状边界的中心点, 所以产生的描述子已满足对目标形状的全局平移不变性. 由分割比 d_i^j 的定义 (公式(7)), 该度量方法的计算用到了相邻层边界像素点的基数比, 当目标形状发生缩放时, 缩放因子在计算过程中会被消掉, 这就使得分割比 d_i^j 具有缩放不变性. 对于构成分散度度量的 ρ_i^j , σ_i^j (公式(8)~公式(10)), 计算这个两个度量时, 已用距离序列的最大值对距离序列中的距离进行了归一化处理, 从而度量 ρ_i^j, σ_i^j 具有缩放不变性. 而对于离心率 e_i^j (公式(14)), 我们在计算时要用到边界主惯性矩的特征值, 由公式(13)可知, 特征值 λ_{\max} 和 λ_{\min} 相除将会消除缩放因子. 因此, 离心率 e_i^j 也具有缩放不变性. 因为分散度度量在文中被定义为 $s_i^j = [\rho_i^j \quad \sigma_i^j \quad e_i^j]$ (公式(15)), 因此分散度度量 s_i^j 具有缩放不变性. 而对于目标形状的旋转变换, 在形状匹配时, 我们针对从形状中提取的特征矩阵给出了循环移位的形状距离度量方法, 从而消除了目标形状旋转的影响.

5.2 算法伪代码和时间复杂度分析

本文提出的形状描述算法的输入有: (1) 目标形状的边界点集 B_0 ; (2) 对角度区间 $[0, 2\pi)$ 的采样频率 m ; (3) 分层的层级数 l . 我们在算法 1 中给出了形状描述算法的伪代码.

算法 1. Calculating Hierarchical Descriptor of Unordered Edge Point-set.

Input:

B_0 : the edge-point set of a shape;

m : the number of sampling angles in range $[0, 2\pi]$;

l : the number of hierarchical level;

Output:

$(d_i^j)_{m \times l}, (s_i^j)_{m \times l}$: shape descriptors.

1: Initialize $(d_i^j)_{m \times l} = 0, (s_i^j)_{m \times l} = 0, \theta = \frac{2\pi}{m}$;

2: $\bar{x}_0 = \frac{\sum_{(x,y) \in B_0} x}{|B_0|}, \bar{y}_0 = \frac{\sum_{(x,y) \in B_0} y}{|B_0|}$; /*Calculate the centroid of the set B_0 */

3: $\bar{B}_0^0 = \{(x - \bar{x}_0, y - \bar{y}_0) | (x, y) \in B_0\}$; /*Move the centroid of the set B_0 to the origin of the coordinate*/

4: **for** $j=0$ to $m-1$ **do** /*The loop control variable j is the index of the sampled angles*/

5: $\bar{x}_0^j = 0, \bar{y}_0^j = 0$; /*The centroid of the set \bar{B}_0^j always takes value $(0,0)$ */

6: **for** $i=1$ to l **do** /*The loop control variable i is the index of the hierarchical levels*/

7: **if** $|\bar{B}_{i-1}^j| > 0$ **then**

8: $\bar{B}_i^j = \{(x, y) | x \leq \bar{x}_{i-1}^j, y \geq \bar{y}_{i-1}^j, (x, y) \in \bar{B}_{i-1}^j\}$; /*Partition the set \bar{B}_{i-1}^j into four parts along the direction of the axis of the coordinate system and take the second quadrant from them*/

9: $d_i^j = \frac{|\bar{B}_i^j|}{|\bar{B}_{i-1}^j|}$; /*Calculate the partition ratio*/

10: $\Omega = \{c_k = \sqrt{(x_k - \bar{x}_{i-1}^j)^2 + (y_k - \bar{y}_{i-1}^j)^2}, (x_k, y_k) \in \bar{B}_{i-1}^j\}$; /*Calculate the distances of all the points in the set \bar{B}_{i-1}^j to the centroid of the set \bar{B}_{i-1}^j */

11: $\rho_i^j = \frac{\sum_{k=1}^{|\Omega|} c_k}{|\Omega|}$; /*Calculate the mean value of distance sequence Ω */

12: $\sigma_i^j = \frac{\sum_{k=1}^{|\Omega|} \left(\frac{c_k}{\max(\Omega)} - \rho_i^j \right)^2}{|\Omega|}$; /*Calculate the variance of distance sequence Ω */

13: $\bar{x}_i^j = \frac{\sum_{(x,y) \in \bar{B}_i^j} x}{|\bar{B}_i^j|}, \bar{y}_i^j = \frac{\sum_{(x,y) \in \bar{B}_i^j} y}{|\bar{B}_i^j|}$; /*Calculate the centroid of the set \bar{B}_i^j */

14: $\mu_{2,0} = \sum_{(x,y) \in \bar{B}_i^j} (x - \bar{x}_i^j)^2, \mu_{1,1} = \sum_{(x,y) \in \bar{B}_i^j} (x - \bar{x}_i^j)(y - \bar{y}_i^j), \mu_{0,2} = \sum_{(x,y) \in \bar{B}_i^j} (y - \bar{y}_i^j)^2$;
/*Calculate the central moment of \bar{B}_i^j */

15: $X = \text{eigenvalue} \left(\begin{bmatrix} \mu_{2,0} & \mu_{1,1} \\ \mu_{1,1} & \mu_{0,2} \end{bmatrix} \right)$; /*Calculate the eigenvalue of matrix $\begin{bmatrix} \mu_{2,0} & \mu_{1,1} \\ \mu_{1,1} & \mu_{0,2} \end{bmatrix}$ */

16: $e_i^j = \sqrt{\min(X) / \max(X)}$; /*Calculate the reciprocal of eccentricity*/

17: $s_i^j = [\rho_i^j \quad \sigma_i^j \quad e_i^j]$; /*Combined the dispersity measures ρ_i^j, σ_i^j and e_i^j to form a descriptor for \bar{B}_i^j */

18: **else**

19: **break**;

20: **end if**

21: **end for**

22: $\bar{B}_0^{j+1} = \{(x \cos \theta + y \sin \theta, y \cos \theta - x \sin \theta) | (x, y) \in \bar{B}_0^j\}$; /*Rotate the set \bar{B}_0^j by angle θ */

23: **end for**

24: **Return** $(d_i^j)_{m \times l}, (s_i^j)_{m \times l}$;

该算法最耗时的计算是从第 4 步~第 23 步的外层循环(执行 m 次),其内层循环(第 6 步~第 21 步)共执行 l 次,执行第 i 次内层循环会将子边界 \bar{B}_{i-1}^j 分割成 4 个部分,并留下左上角的边界 \bar{B}_i^j 作为新的当前层子边界,然后对其进行特征描述.假设初始边界点集 B_0 有 n 个像素点,由于每次分割都是根据被分割的子边界的中心点,沿坐标横轴和坐标纵轴两个方向将其分割成 4 个部分,可以认为被分割的子边界的像素点在这 4 个部分分布的概率相等,因此第 i 次内层循环在对上一层子边界 \bar{B}_{i-1}^j 中的 $|\bar{B}_{i-1}^j|$ 个点进行分割后,依概率分布,有 $\frac{|\bar{B}_{i-1}^j|}{4}$ 个点将会被分到当前层子边界 \bar{B}_i^j ,根据递推关系,可推出 $|\bar{B}_{i-1}^j| = \frac{n}{4^{i-1}}$ 和 $|\bar{B}_i^j| = \frac{n}{4^i}$.所以在第 i 次内层循环,完成对子边界 \bar{B}_{i-1}^j 分割的算法第 8 步的时间复杂度为 $O\left(\frac{n}{4^{i-1}}\right)$,而完成对子边界 \bar{B}_i^j 进行度量的算法的第 9 步~第 17 步的时间复杂度为 $O\left(\frac{n}{4^i}\right)$.因此,执行第 i 次内层循环的时间复杂度为 $O\left(\frac{n}{4^i}\right)$,从而执行整个内层循环的时间复杂度为 $O\left(n + \frac{n}{4} + \frac{n}{16} + \dots + \frac{n}{4^{l-1}}\right) = O\left(\frac{4}{3}n\left(1 - \frac{1}{4^l}\right)\right) = O(n)$.又执行对边界的旋转操作(算法第 22 步)的时间复杂度为 $O(n)$,所以执行整个外层循环的平均时间复杂度为 $O(m(n+n))=O(mn)$,即该算法的时间复杂度为 $O(mn)$.

以上是形状描述算法的时间复杂度分析,本文提出方法的另外一部分是形状匹配任务.在得到形状描述矩阵 F_0 (公式(17)和公式(18))后,我们要对其进行 $m-1$ 次循环移位(公式(19)),由于 F_0 的规模为 $m \times 4l$,所以循环移位操作的时间复杂度为 $O(ml)$.计算两个特征矩阵的 L-1 距离的时间复杂度为 $O(ml)$,而我们要将形状 A 的特征矩阵及其 $m-1$ 个循环移位版本与形状 B 的特征矩阵进行匹配(公式(20)),因此,计算两个形状距离的时间复杂度为 $O(lm^2)$,从而整个形状匹配阶段的时间复杂度为 $O(ml+lm^2)=O(lm^2)$.值指出的是, l 为本文提出的形状描述算法的一个输入参数,若 $l \leq \log_4 n = \frac{1}{2} \log_2 n$,即算法在执行 l 次内部循环后仍有 $\frac{n}{4^l} \geq 1$ 个点划分到了边界点集 \bar{B}_i^j 中,这种情况下,算法的内层循环执行了 l 次,因此,此时形状匹配阶段的时间复杂度小于等于 $O(m^2 \log_2 n)$;若 $l > \log_4 n$,因 $\frac{n}{4^l} < 1$,算法的内部循环最多执行到 $\log_4 n$ 步后便会跳出循环,此时,形状匹配阶段的时间复杂度为 $O(m^2 \log_2 n)$,因此,形状匹配阶段的时间复杂度在最坏情况下为 $O(m^2 \log_2 n)$.

上述分析表明,本文提出的算法无论在形状描述阶段,还是在形状匹配阶段,都有着可接受的计算复杂度.

6 实验结果和讨论

为评估本文提出的目标形状识别算法的性能,我们将 MPEG-7 CE-2 区域形状测试集和 MPEG-7 CE-1 轮廓线形状测试集这两个被广泛采用的形状测试集作为基准测试集.因本文提出的方法属于基于边界点集的描述方法,所以在实验中,我们选择的主要比较对象为 4 种基于边界点集的形状描述方法——形状上下文 SC^[35]、距离集 DS^[36]、无序点集描述方法 UPSD^[40]和轮廓点分布直方图 CPDH^[38].另外,两种基于区域点集的描述方法:Zernike 矩 ZM^[28]、自适应密度直方图 AHDH^[34]和一种区域点集和边界点集相结合的描述方法:Zernike 矩边缘梯度 ZMEG^[32],在实验中也纳入为比较对象.其中,ZM^[28]是 MPEG-7 推荐的标准的区域形状描述方法;而 AHDH^[34]同本文的算法一样,都属于分层的形状描述方法.我们的实验参数设置如下:角度采样频率 $m=72$,层级数 $l=3$;构建特征矩阵过程中,用以调节描述矩阵 D 和 S 在形状相似性度量中贡献的权重参数 $\omega=0.8$.实验平台为一台 CPU 为 Intel Core i5-4590 3.30GHz,内存 8G,操作系统为 Win10 的台式计算机,算法用 Matlab 编程实现.

6.1 MPEG-7 CE-2区域形状测试集

我们用于算法评估的第 1 组实验是在 MPEG-7 CE-2 区域形状测试集^[30]上进行的.MPEG-7 CE-2 测试集总共包含 3 621 幅区域形状图像,其中的部分样本如图 6 所示.该测试集中有 651 个样本,被分成 31 组,每组有 21 个样本,同一组中的样本属于同类形状,图 7 给出了第 8 组形状的 21 个样本.我们采用被广泛用于形状检索性能评估的方法——bull's-eye test 评估方法^[6-12,15-17,19-23,34-36,38,40,43,44]进行算法的性能评估.采用该方法,我们把数据

集中 651 个样本依次作为检索样本跟测试集中的 3 621 样本进行形状相似性比较,返回差异度最小的前 $2 \times 21 = 42$ 幅图像,统计返回的 42 幅图像中,与查询样本属于同一组的数目.假设返回 r 幅属于同一组的样本,显然 $r \leq 21$,计算 $r/21 \times 100\%$ 的值作为该查询样本的检索率,651 个样本的平均检索率作为整个测试的检索率 (bull's-eye test score).



Fig 6 Samples from MPEG-7 CE-2 dataset

图 6 MPEG-7 CE-2 数据集的样本

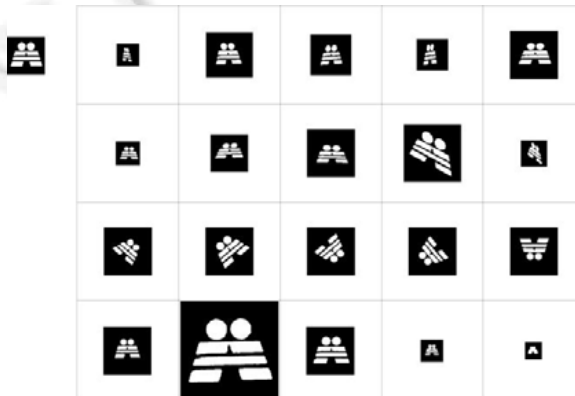


Fig 7 All the samples in the group 8 of MPEG-7 CE-2 dataset

图 7 MPEG-7 CE-2 数据集中第 8 组形状的所有样本

我们在图 8 中给出部分形状的前 21 个检索结果(框以红色矩形的形状是错误的检索结果),并在表 1 中给出了本文提出的方法和参与比较的两种基于边界点集的方法 SC^[35]和 UPSD^[40]、两种基于区域点集的方法 ZM^[28], AHDH^[34]和一种区域点集和边界点集相结合的方法 ZMEG^[32]在 MPEG-7 CE-2 区域形状测试集上的检索率.从表中给出的数据可以看出,与同类基于边界点集的方法 SC 和 UPSD 相比,本文提出方法的检索精确率分别高出了 15.48 个和 2.33 个百分点.与基于区域点集的方法 ZM 相比,本文的方法高出了 6.26 个百分点;与区域点集和边界点集相结合的方法 ZMEG 相比,本文的方法高出了 4.04 个百分点.值得指出的是,与本文的思想最为接近的是同样采用分层描述的 AHDH 方法,但该方法在 MPEG-7 CE-2 测试集上仅取得了 49.94% 的检索率.其主要原因是,该方法仅从单个方向上对目标形状进行描述,该方法的描述和匹配依赖于目标形状的方向;而本文提出的方法采用了一种旋转分层的描述框架,从多个方向上描述了形状的分层复杂性,因此其形状匹配结果不依赖于形状的方向.该组实验结果表明:本文提出的方法在 MPEG-7 CE-2 区域形状测试集上的测试性能要优于参与比较的同类方法,也好于参与比较的两种基于区域点集的形状描述方法和一种区域点集和边界点集相结合的方法,从而证明本文提出的方法比其他方法更能有效地处理包含复杂内部结构的形状图像,对目标形状检索具有一般性.需要指出的是,对于 DS^[36]和 CPDH^[38],由于报告这两种方法的文献没有给出在 MPEG-7 CE-2 测试集的

测试结果,也没有给出具体的实现细节,因此在本组实验中,我们略去了与这两种方法的比较,但在下一组实验中,我们将直接引用这两种方法报告的结果作为对比.

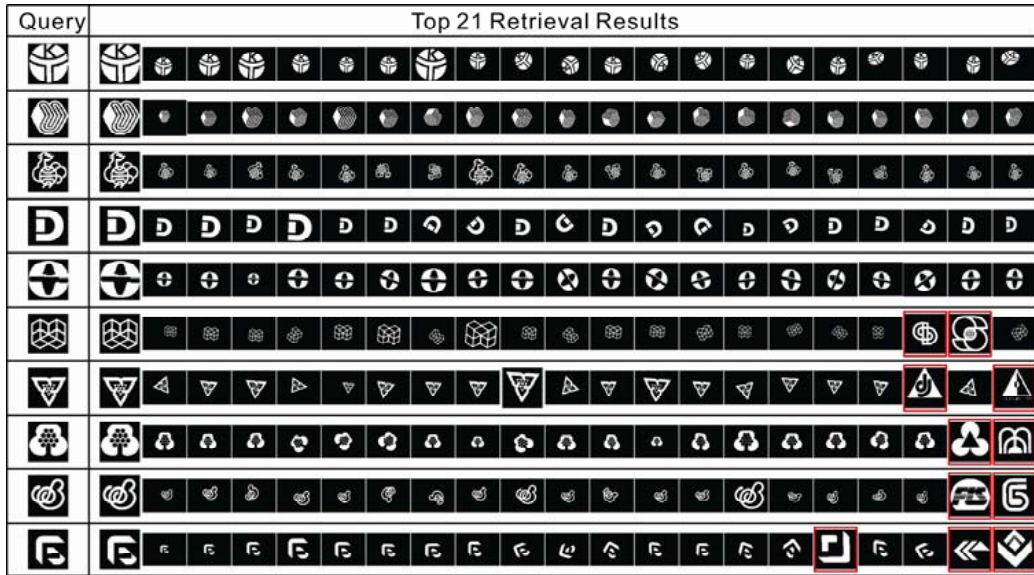


Fig.8 Top 21 retrieval results of partial shapes in MPEG-7 CE-2 dataset

图 8 MPEG-7 CE-2 数据集中部分形状的前 21 个检索结果

Table 1 Retrieval rates of the compared methods on the MPEG-7 CE-2 shape dataset

表 1 各种对比方法在 MPEG-7 CE-2 形状测试集上的检索率

Algorithm	Bulls-eye test score (%)
SC ^[35]	70.98
UPSD ^[40]	84.13
ZM ^[28]	80.20
ZMEG ^[32]	82.42
AHDH ^[34]	49.94
The proposed method	86.46

6.2 MPEG-7 CE-1轮廓线形状测试集

我们用于算法评估的第 2 组实验是在 MPEG-7 CE-1 轮廓线形状测试集^[42]上进行的.MPEG-7 CE-1 轮廓线形状测试集是被广泛采用的用于形状检索性能评估的标准测试集.该测试集共包含 1 400 幅轮廓线形状图像,它们被分成 70 种形状类型,不同类型形状的样本如图 9 所示,每种类型的形状包含 20 个样本,如图 10 所示.该组实验我们依然用 bulls-eye test 评估方法.与上一组实验不同的是,这里我们用测试集中的每一个样本作为一个查询样本,跟测试集中的所有样本进行相似性比较;然后,根据形状的差异度进行排序,返回差异度最小的前 2×20=40 幅图像,统计返回的这 40 幅图像中有多少幅与查询样本属于同一类.假设有 r 幅返回图像跟查询图像属于同一类,显然 $r \leq 20$,计算 $r/20 \times 100\%$ 的值作为该查询样本的检索率,取 1 400 个检索样本的检索率的平均值作为整个测试的检索率(bulls-eye test score).

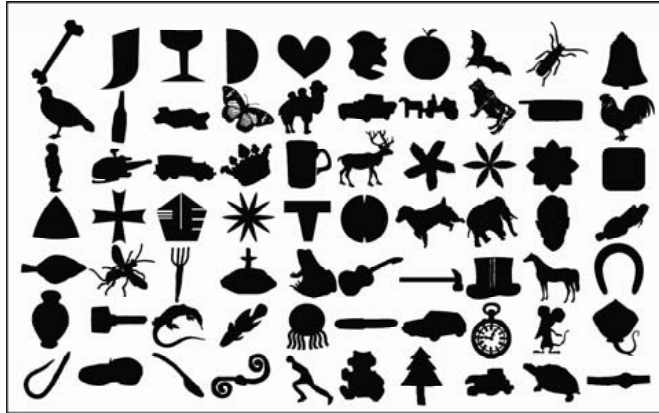


Fig.9 All kinds of shapes from the MPEG-7 CE-1 dataset

图9 MPEG-7 CE-1 数据集的所有类型形状样本

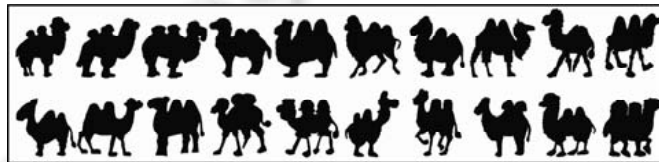


Fig.10 All the samples of the camel shape in the MPEG-7 CE-1 dataset

图10 MPEG-7 CE-1 数据集中骆驼形状的所有样本

我们在图 11 中给出了部分形状的前 20 个检索结果(框以红色矩形的形状是错误的检索结果),并在表 2 中给出了本文提出的方法和参与比较的 4 种基于边界点集的方法(带“*”标识的结果直接引自参考文献)、两种基于区域点集的方法和一种区域点集和边界点集相结合的方法在 MPEG-7 CE-1 轮廓线形状测试集上的检索率.

Query	Top 20 Retrieval Results																			

Fig.11 Top 20 retrieval results of partial shapes in MPEG-7 CE-1 dataset

图11 MPEG-7 CE-1 数据集中部分形状的前 20 个检索结果

Table 2 Retrieval rates of the compared methods on the MPEG-7 CE-1 shape dataset**表 2** 各种对比方法在 MPEG-7 CE-1 形状测试集上的检索率

Algorithm	Bulls-eye test score (%)
SC ^[35]	76.51*
DS ^[36]	78.38*
UPSD ^[40]	78.18*
ZM ^[28]	68.90
ZMEG ^[32]	70.64
AHDH ^[34]	63.95*
CPDH ^[38]	76.56*
The proposed method	82.61

从表中给出的数据可以看出,与同类基于边界点集的方法 SC^[35],DS^[36],UPSD^[40]和 CPDH^[38]相比,本文提出方法的检索精确率分别高出了 6.1,4.23,4.43 和 6.05 个百分点;与基于区域点集的方法 ZM^[28]和 AHDH^[34]相比,本文的方法分别高出了 13.71 和 18.66 个百分点;与区域点集和边界点集相结合的方法 ZMEG^[32]相比,本文的方法高出了 11.97 个百分点.在本组实验中,我们还比较了各类算法执行的效率,需要指出的是:对于数据集中的形状,用于表征这些形状的形状描述子可以在线下进行提取,所以这里我们比较的是各算法执行线上形状匹配的效率.根据文献[35]报告的结果,采用 Shape Contexts 方法,执行一次两个形状的匹配计算(采样 100 个点),在一台 Pentium III/ 500MHZ 的计算机上需要耗时 2×10^{-1} s;根据文献[36]报告的结果,采用 Distance Sets 方法,执行一次两个形状的匹配计算(采样 250 个点),在一台 Pentium III/667MHZ 的计算机上大约耗时 7×10^{-1} s;根据文献[40]报告的结果,采用 Unordered Point-set Description 方法,在一台 CPU 为 Intel Core i7-4770/3.4GHZ 的计算机上,执行一次两个形状的匹配计算(采样 300 个点),耗时为 9.4×10^{-4} s;而采用本文提出的方法,在一台 CPU 为 Intel Core i5-4590/ 3.30GHZ 的计算机上,执行一次两个形状的匹配计算,仅用时 2.7×10^{-4} s.我们在表 3 中给出了几种经典方法和本文提出的方法在 MPEG-7 CE-1 形状测试集上的匹配时间复杂度和平均匹配时间对比(带“*”标识的结果直接引自参考文献).该组实验结果表明:相较于参与比较的其他方法,本文提出的方法不仅具有更好的检索精确度,而且具有更高的计算效率.

Table 3 Time complexity and average matching time of several remarkable methods and the proposed method on the MPEG-7 CE-1 shape dataset**表 3** 几种经典方法和本文提出的方法在 MPEG-7 CE-1 形状测试集上的匹配时间复杂度和平均匹配时间

Algorithm	Time complexity	Average matching time (ms)
SC ^[35]	$O(N^3)$	200*
DS ^[36]	$O(N^3)$	700*
UPSD ^[40]	$O(N \log N)$	0.94*
The proposed method	$O(M^2 \log N)$	0.27

对于一些基于曲线的方法,如多尺度凹凸性^[5]、内部距离^[6]、三角形面积^[7]、轮廓柔韧性^[8]、形状树^[17]、局部仿射不变描述子^[19],虽然这些方法在 MPEG-7 CE-1 测试集上报告的结果大都超过 85%,但这些方法是专门为轮廓线形状设计的,对于 MPEG-7 CE-2 中的区域形状,这些方法则不能对其进行有效地处理.值得指出的是,本文提出的方法虽然没有用到边界点的有序信息,但其在 MPEG-7 CE-1 上的识别率已超过了一些基于曲线的方法,如生物序列(biological code,简称 BC)^[43]、距离内缘比率(distance interior ratio)^[44]等方法.

7 结 论

本文提出了基于目标边界无序点集的形状描述方法,以应用于一般的形状图像检索任务.该方法通过对目标形状的边界点集沿各个方向的逐层分割,建立了一种对目标形状边界的多层描述模型,对各层子边界的分割比和分散度的几何度量,多尺度地描述了目标形状边界点的空间分布特性.通过在 MPEG-7 CE-2 区域形状测试集和 MPEG-7 CE-1 轮廓线形状测试集上的两组图像检索实验和其他同类方法的对比,验证了该方法的有效性.

总结该算法具有如下优点.是一种通用的形状图像描述算法,既能描述轮廓线形状,又能描述区域形状;具

有特征抽取的开放性,虽然基于本文提出的描述框架,我们仅用了分割比和分散度这两种几何度量对每一层子边界进行了度量,但该描述框架具有一般性,每一层子边界还可以引入其他几何度量,如矩特征、傅里叶特征等,以满足对不同精度的形状图像识别任务的要求;具有内在的多尺度特性和对旋转、缩放和平移的不变性;具有比同类方法更好的检索精确率和更高的计算效率。

在未来的研究工作中,在目标形状的特征抽取阶段,我们会将一些其他的边界度量方法纳入到给出的分层描述框架中,结合机器学习的方法进行有效的特征选取,从而获取更加精确的形状描述子;在形状匹配阶段,我们将考虑结合动态规划和距离度量学习的方法,来进行更加准确的形状相似度的度量。

References:

- [1] Smeulders AWM, Worring M, Santini S, Gupta A, Jain R. Content-Based image retrieval at the end of the early years. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2000,22(12):1349–1380.
- [2] Cao J, Wu Z, Wu J, Liu W. Towards information-theoretic K -means clustering for image indexing. *Signal Processing*, 2013,93(7):2026–2037.
- [3] Cao J, Wang B, Brown D. Similarity based leaf image retrieval using multiscale R -angle description. *Information Sciences*, 2016, 374:51–64.
- [4] Costa LDF, Cesar RM. *Shape Analysis and Classification: Theory and Practice*. 2nd ed., Florida: CRC Press LLC, 2001. 1–25.
- [5] Adamek T, O'Connor NE. A multiscale representation method for nonrigid shapes with a single closed contour. *IEEE Trans. on Circuits and Systems for Video Technology*, 2004,14(5):742–753.
- [6] Ling H, Jacobs DW. Shape classification using the inner-diatance. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2007,29(2): 286–299.
- [7] Alajlan N, Rube IE, Kamel MS, Freeman G. Shape retrieval using triangle-area representation and dynamic space warping. *Pattern Recognition*, 2007,40(7):1911–1920.
- [8] Xu C, Liu J, Tang X. 2D shape matching by contour flexibility. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2009,31(1):180–186.
- [9] Biswas S, Aggarwal G, Chellappa R. An efficient and robust algorithm for shape indexing and retrieval. *IEEE Trans. on Multimedia*, 2010,12(5):372–384.
- [10] Bai X, Yang X, Latecki LJ, Liu W, Tu Z. Learning context-sensitive shape similarity by graph transduction. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2010,32(5):861–874.
- [11] Wang J, Bai X, You X, Liu W, Latecki LJ. Shape matching and classification using height functions. *Pattern Recognition Letters*, 2012,33(2):134–143.
- [12] Wang B, Gao Y. Hierarchical string cuts: A translation, rotation, scale and mirror invariant descriptor for fast shape retrieval. *IEEE Trans. on Image Processing*, 2014,23(9):4101–4111.
- [13] Zahn CT, Roskies RZ. Fourier descriptors for plane closed curves. *IEEE Trans. on Computers*, 1972,C-21(3):269–281.
- [14] Persoon E, Fu KS. Shape discrimination using Fourier descriptors. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1986,3:388–397.
- [15] Mokhtarian F, Abbasi S, Kittler J. Efficient and robust retrieval by shape content through curvature scale space. In: *Proc. of the Int'l Workshop on Image Databases and Multimedia Search*. Amsterdam, 1996. 51–58.
- [16] Mokhtarian F, Bober M. *Curvature Scale Space Representation: Theory, Application, and MPEG-7 Standardization*. Kluwer Academic Publishers, 2003.
- [17] Felzenszwalb PF, Schwartz JD. Hierarchical matching of deformable shapes. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*. Minneapolis, 2007. 1–8.
- [18] Xiong GQ, Qi DX, Guo FH. A class of orthonormal complete piecewise polygonal systems and applications. *Science in China (Information Sciences)*, 2012,42(1):70–82 (in Chinese with English abstract).
- [19] Wang Z, Liang M. Locally affine invariant descriptors for shape matching and retrieval. *IEEE Signal Processing Letters*, 2010, 17(9):803–806.

- [20] Hu RX, Jia W, Ling H, Zhao Y, Gui J. Angular pattern and binary angular pattern for shape retrieval. *IEEE Trans. on Image Processing*, 2014,23(3):1118–1127.
- [21] Daliri MR, Torre V. Robust symbolic representation for shape recognition and retrieval. *Pattern Recognition*, 2008,41(5):1782–1798.
- [22] Attalla E, Siy P. Robust shape similarity based on contour segmentation polygonal multiresolution and elastic matching. *Pattern Recognition*, 2005,38(12):2229–2241.
- [23] Super BJ. Retrieval from shape databases using chance probability function and fixed correspondence. *Int'l Journal of Pattern Recognition and Artificial Intelligence*, 2006,20(8):1117–1137.
- [24] Li Z, Liu J, Tang J, Lu H. Robust structured subspace learning for data representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2015,37(10):2085–2098.
- [25] Li Z, Liu J, Yang Y, Zhou X, Lu H. Clustering-Guided sparse structural learning for unsupervised feature selection. *IEEE Trans. on Knowledge and Data Engineering*, 2014,26(9):2138–2150.
- [26] Li Z, Tang J. Weakly supervised deep metric learning for community-contributed image retrieval. *IEEE Trans. on Multimedia*, 2015,17(11):1989–1999.
- [27] Hu MK. Visual pattern recognition by moment invariants. *IRE Trans. on Information Theory*, 1962,IT-8:179–187.
- [28] Khotanzad A, Hong YH. Invariant image recognition in Zernike moments. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1990, 12(5):250–261.
- [29] Nassery P, Faez K. Signature pattern recognition using pseudo Zernike moments and a fuzzy logic classifier. In: *Proc. of the IEEE Int'l Conf. on Image Processing*. 1996. 197–200.
- [30] Zhang D, Lu G. Shape based image retrieval using generic Fourier descriptor. *Signal Processing: Image Communication*, 2002,17(10):825–848.
- [31] Zhang D, Lu G. Review of shape representation and description techniques. *Pattern Recognition*, 2004,37(1):1–19.
- [32] Anuar FM, Setchi R, Lai YK. Trademark image retrieval using an integrated shape descriptor. *Expert Systems with Applications*, 2013,40(1):105–121.
- [33] Yang M, Qiu G, Huang J, Elliman D. Near-duplicate image recognition and content-based image retrieval using adaptive hierarchical geometric centroids. In: *Proc. of the 18th Int'l Conf. on Pattern Recognition (ICPR 2006)*. 2006. 958–961.
- [34] Sidiropoulos P, Vrochidis S, Kompatsiaris I. Content-Baese binary image retrieval using the adaptive hierarchical density histogram. *Pattern Recognition*, 2011,44(4):739–750.
- [35] Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2002,24(4):509–522.
- [36] Crigorescu C, Petkov N. Distance sets for shape filters and shape recognition. *IEEE Trans. on Image Processing*, 2003,12(10):1274–1286.
- [37] Papadimitriou CH, Steiglitz K. *Combinatorial Optimization: Algorithms and Complexity*. New York: Dover Publications, INC, 1998. 248–254.
- [38] 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.
- [39] Backes AR, Casanova D, Bruno OM. A complex network-based approach for boundary shape analysis. *Pattern Recognition*, 2009, 42(1):54–67.
- [40] Wang B. Shape recognition using unordered point-set description and matching of object contour. *Ruan Jian Xue Bao/Journal of Software*, 2016,27(12):3131–3142 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5101.htm> [doi: 10.13328/j.cnki.jos.005101]
- [41] Mokhtarian F, Abbasi S. Matching shape with self-intersections: Application on leaf classification. *IEEE Trans. on Image Processing*, 2004,13(5):653–661.
- [42] Latecki LJ, Lakamper R, Eckhardt U. Shape descriptors for non-rigid shapes with a single closed contour. In: *Proc. of the IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*. 2000. 424–429.

- [43] Bicego M, Lovato P. A bioinformatics approach to 2D shape classification. *Computer Vision and Image Understanding*, 2016,145: 59–69.
- [44] Kaothanthong N, Chun J, Tokuyama T. Distance interior ratio: A new shape signature for 2D shape retrieval. *Pattern Recognition Letters*, 2016,78:14–21.

附中文参考文献:

- [18] 熊刚强,齐东旭,郭芬红.一类完备的正交分段多项式函数系及其应用. *中国科学:信息科学*,2012,42(1):70–82.
- [40] 王斌.用于形状识别的目标轮廓无序点集描述与匹配. *软件学报*,2016,27(12):3131–3142. <http://www.jos.org.cn/1000-9825/5101.htm> [doi: 10.13328/j.cnki.jos.005101]



刘锋(1992—),男,安徽六安人,硕士生,主要研究领域为计算机视觉,图像处理,模式识别.



王斌(1969—),男,博士,教授,CCF 高级会员,主要研究领域为计算机视觉,图像处理,模式识别.