

文章编号: 2095-4980(2013)04-0624-05

一种结合边缘与区域信息的图像特征提取算法

张瑜慧, 王海燕, 郑步芹, 李 洁

(宿迁学院, 江苏 宿迁 223800)

摘 要: 为了能够较准确地表示图像的形状特征, 提出了一种融合边界和区域信息的新的形状描述子。首先对图像进行二维离散余弦变换, 获得低频系数作为区域特征。之后提取图像的轮廓并进行采样, 形成描述形状轮廓的有序点列表, 对每个采样点分别顺时针和逆时针等距离跟踪获得 2 个邻点, 计算拱高以及质心距离。然后获取由拱高和质心距离组成的复函数的频域描述子, 并组合区域特征与轮廓特征。对 MPEG-7 标准图形库的检索实验显示, 该描述子的检索性能显著优于三角形面积函数、质心距离函数与拱高半径复函数等同类描述子。

关键词: 轮廓曲线; 区域特征; 形状特征; 图像检索

中图分类号: TN911.73; TP391.3 **文献标识码:** A **doi:** 10.11805/TKYDA201304.0624

An image feature extraction algorithm based on edge and regional information

ZHANG Yu-hui, WANG Hai-yan, ZHENG Bu-qin, LI Jie

(Suqian College, Suqian Jiangsu 223800, China)

Abstract: A new shape descriptor with fusion contour and region information is proposed in order to represent the shape characteristics of the image more accurately. Apply two-dimensional Discrete Cosine Transform(DCT) to obtain the low-frequency coefficients as regional characteristics. Then extract and sample the contours of image, and get the ordered list of points. Track each sampling point equidistantly clockwise and counterclockwise to get two adjacent points, and calculate the arch-height and centroid distance. Construct a complex function of arch-height and centroid distance, and obtain the DCT coefficients, then combine regional and contour features. Retrieval experiments on MPEG-7 shape database show that, the retrieval performance of this descriptor is better than that of the triangle area function, the centroid distance function, and arch-height radius complex function significantly.

Key words: contour curve; region-based features; shape features; image retrieval

基于内容的图像检索研究重点多为尽量全面地提取图像的视觉特征, 同时对所提取的视觉特征寻求最好的表达方式^[1]。在图像的视觉特征中, 形状是一种主要的低层特征, 它是人对图像理解的重要内容之一, 相比于颜色和纹理特征, 它更能从语义上描述目标图像。文献[2]将现有的形状描述子分为两大类, 一类是基于轮廓的描述子, 另一类是基于区域的描述子。轮廓形状描述了图像的边界信息, 具有易于提取的特点, 因此在很多场合得以应用。轮廓法又分为两个子类, 一类将边界作为整体, 从中提取表示形状的参数(轮廓曲线坐标、质心距离^[2]、三角形面积函数^[3]、拱高^[4]等等), 再进行后续处理, 形成诸如曲率尺度空间^[5]、傅里叶描述子^[2,6]、小波描述子^[7]等形状描述法。另一类根据一定准则, 将边界用原语表示, 最终的形状特征通常用字符串或树来表示。在基于区域的形状特征的提取过程中, 区域内的所有像素都参与形状描述。其中, 矩是一种有效的描述法, 常见的矩有 Hu 不变矩、Zernike 矩等。除此之外形状的分区描述法也是一类有效的区域描述法。

形状的两类描述各有利弊。前者主要对目标边界上的像素进行处理, 从视觉生理的角度讲, 这类方法更接近人眼识别形状的机理, 而且该类方法参与计算特征的数据量小。后者是对整个区域像素进行处理, 其参与计算的数据量大, 相比前者, 该类方法在抵抗噪声等外部干扰上具有较强优势。文献[8]给出一个评价形状特征优劣的准则: 较高的检索准确率, 紧凑的描述性, 具有通用性, 计算复杂度低, 并且对图形的检索具有鲁棒性等。其中

收稿日期: 2012-08-03; 修回日期: 2012-09-21

基金项目: 宿迁市科技创新专项基金资助项目(Z201206); 宿迁市科技创新专项基金资助项目(Z2010017); 宿迁学院科研基金资助项目(2011KY21)

较高的准确率是最基本也是最重要的准则。本文提出一种综合轮廓边缘与区域描述优点的形状特征。该算法从图像的边界曲线中提取表示形状内容的频域描述，为了克服轮廓法过于依赖边缘检测结果的弊端，算法结合区域描述，组合形成具有较高检索性能的混合形状描述子。

1 形状特征的提取原理与过程

1.1 区域特征

离散余弦变换(Discrete Cosine Transform, DCT)是一种具有对称性、可分离性的正交变换，其具有将能量集中于低频系数，以及高频能量小且衰减快的特征。一维 DCT 定义如式(1)所示：

$$F(u) = C(u) \sum_{x=0}^{N-1} f(x) \cos \frac{\pi(2x+1)u}{2N} \quad (1)$$

式中： $F(u)$ 是第 u 个余弦系数， u 是广义频率变量， $u=1,2,\dots,N-1$ ； $f(x)$ 是 N 点序列， $x=1,2,\dots,N-1$ ； $C(u)$ 是归一化加权系数，定义为：

$$C(u) = \begin{cases} \sqrt{1/N}, & u=0 \\ \sqrt{2/N}, & 1 \leq u \leq N-1 \end{cases} \quad (2)$$

二维 DCT 是一维 DCT 在空域的扩展，定义为：

$$F(u,v) = C(u)C(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \cos \left[\frac{\pi}{M} u \left(x + \frac{1}{2} \right) \right] \cos \left[\frac{\pi}{N} v \left(y + \frac{1}{2} \right) \right] \quad (3)$$

式中： $F(u,v)$ 是 2 维 DCT 变换系数， u,v 分别代表水平空间和垂直空间的频率变量， $u=1,2,\dots,M-1$ ， $v=1,2,\dots,N-1$ ； $f(x,y)$ 是空间域二维变量， $x=1,2,\dots,M-1$ ， $y=1,2,\dots,N-1$ 。加权系数 $C(u),C(v)$ 分别定义为：

$$C(u) = \begin{cases} \sqrt{1/M}, & u=0 \\ \sqrt{2/M}, & 1 \leq u \leq M-1 \end{cases}, \quad C(v) = \begin{cases} \sqrt{1/N}, & v=0 \\ \sqrt{2/N}, & 1 \leq v \leq N-1 \end{cases} \quad (4)$$

图 1 给出了一幅二值图像(图 1(a))及其对应的 DCT 系数的灰度图(图 1(b))。图 1(b)左上角是频谱的低频部分，可以看出 DCT 具有很强的能量集中在频谱低频部分的特性。由于图像的能量主要集中在低频部分，而高频部分的分量很弱，仅仅揭示了图像的某些细节。通过离散余弦变换，将图像从空间域转换到频域，以获得表示图像特征信息的低频 DCT 系数作为形状描述子。

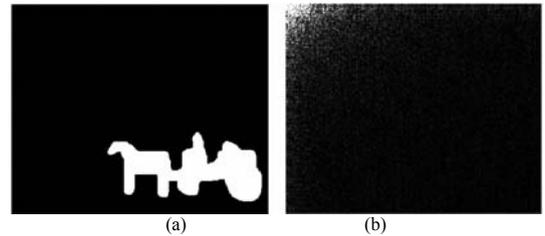


Fig.1 Image and its DCT spectrogram
图 1 图像及其 DCT 频谱图

1.2 轮廓特征

形状是把一个物体从它的周围物体区别出来的重要特征，一般描述形状的图像可分为目标区域与背景区域，因此往往可用 1 和 0 两个值来表示图像，即对图像进行二值化处理。为提取目标图形的边界描述，首先通过边缘检测获取二值图像的边界曲线，然后在曲线上进行 N 点采样。设曲线长度为 L ，在对轮廓曲线进行采样时，首先在曲线上任取一点，每隔 $[L/N]$ 个点采样一次(“ $[]$ ”表示取整数)，形成用于描述形状轮廓的有序采样点列表。

设区域形状图像 $I = \{f(x,y); 0 \leq x < M, 0 \leq y < N\}$ ，区域形状 I 的大小为 $M \times N$ ，其质心坐标 (x_c, y_c) 定义如下：

$$x_c = \frac{1}{A} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} xf(x,y) \quad (5)$$

$$y_c = \frac{1}{A} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} yf(x,y) \quad (6)$$

式中 A 是区域面积，定义如式(6)：

$$A = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \quad (7)$$

图像的质心距离是像素到图像质心间的距离，质心距离由式(8)定义：

$$r = \sqrt{(x - x_c)^2 + (y - y_c)^2} \quad (8)$$

根据式(5)~式(8)对有序采样点列表进行处理后,获得图像轮廓的质心及质心距离,之后再对所有采样点进行跟踪,获取邻点。具体操作是:对某个采样点 a 分别顺时针和逆时针沿着图像边界跟踪距离 S (S 是该采样点处质心距离的 $1/2$), 获得邻点 b 和 c (如图 2 所示)。获得邻点后可以获得 a 到 bc 连线的垂直距离 ad 作为拱高的值, 拱高的符号由采样点 a 相对 bc 连线的位置确定, 将 bc 看作从 b 指向 c 的有向线段, 若 a 在 bc 的右边, 拱高的符号定义为正, 反之定义为负^[4]。拱高的符号可以反映图形边界的凹凸情况, 如图 2(a) 拱高符号为正, 对应点的边界是凸的, 而图 2(b) 符号为负, 对应边界是凹的。这样对于 N 个采样点, 每个点获得拱高以及质心距离 2 个参数, 用于表示该点处的形状信息。将描述形状相似性信息的质心距离作为实部, 将描述形状细节内容的拱高作为虚部, 构成拱高半径复函数。对复函数进行 DCT, 获得低频余弦变换系数作为形状描述子, 形成本文算子的第 2 个子特征。

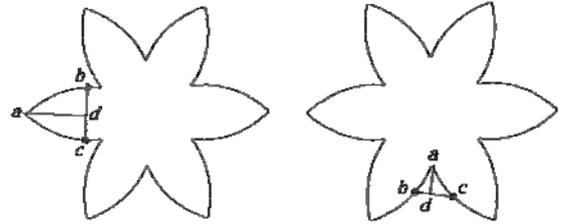


Fig.2 Contour feature at sampling points
图 2 采样点处的轮廓特征

上述子特征是针对于图像轮廓提取的, 因此存在轮廓特征的一些固有缺点, 比如易受噪声干扰, 以及依赖边缘检测的结果。为了增强算法的鲁棒性, 本文考虑将其与区域形状子特征进行融合, 形成一个能全面描述图像形状内容的特征。

1.3 特征组合

在进行特征组合时, 不同子特征可能具有不同的物理意义, 变化幅度也可能具有较大差异, 不具有直接可比性, 因而为了缩小分量取值之间的差距, 即为了使子特征在特征相似性比较中具有公平的贡献, 在进行特征组合前, 需要对子特征进行归一化。

设图像特征具有 K 维, 则第 i 幅图像的特征向量可以表示为: $[f_{i1}, f_{i2}, \dots, f_{iK}]$, 如果数据库的规模是 N , 则可以获得大小为 $N \times K$ 的特征矩阵。对矩阵采用公式(9)进行高斯归一化:

$$f'_{ij} = \left(\frac{f_{ij} - u_j}{3\sigma_j} + 1 \right) / 2 \quad (9)$$

式中: u_j 和 σ_j 分别是特征矩阵中第 j 列的均值和方差; f_{ij} 是第 i 幅图像第 j 列特征值, f'_{ij} 是高斯归一化后的特征值。归一化后的数值, 在区间 $[0, 1]$ 的概率达到 99% 以上。经过归一化处理后, 特征向量的取值可直接进行比较。

1.4 形状特征的提取流程

检索算法的基本思想是: 首先对图像库中的所有图像提取形状特征, 形成特征库, 在检索时, 将待检索图像的特征与特征库中的其余特征进行比较, 并按相似性大小返回检索结果。具体实现步骤如下:

- 1) 对图像进行 2 维离散余弦变换, 获得表示图像能量的 DCT 系数, 并以低频系数作为区域子特征;
- 2) 对图像进行边缘检测, 并等间隔采样 N 个点, 求取图像的质心 (x_c, y_c) , 以及采样点到质心的距离 r ;
- 3) 对采样点在边界顺时针和逆时针两个方向分别跟踪 $r/2$ 距离, 获得两个邻点, 计算其拱高函数;
- 4) 以质心距离为实部, 拱高为虚部构成复函数, 在频域中提取表示形状内容的子特征;
- 5) 对 2 个子特征进行归一化, 组合成结合图像边缘与区域特征的混合形状描述子。

2 实验结果与讨论

2.1 图像库

MPEG-7 多媒体内容描述接口的目标是产生一种通用的多媒体数据内容描述接口, 以满足各种多媒体信息系统对多媒体数据的管理和检索的需求。MPEG-7 图形库(set B)包含 1 400 幅图形, 70 大类, 每类图形包含 20 幅在形状上存在差异的图片, 它是用于形状特征描述算子的标准验证库, 主要用于进行图像相似性检索实验, 测试形状描述子对任意形状畸变的鲁棒性。

2.2 相似性比较

在进行检索时, 必须进行特征之间的相似性比较。若图像 A 与 B 的第 i 个子特征 f_{iA} 与 f_{iB} 之间的距离表示为 $D(f_{iA}, f_{iB})$, 则本文利用 2 个子特征进行相似性比较时引入权值, 采用公式(10)的方式进行:

$$D(A, B) = w_1 D(f_{1A}, f_{1B}) + w_2 D(f_{2A}, f_{2B}) \tag{10}$$

式中： w_1 和 w_2 依次是子特征 1,2 的权值，并且满足 $w_1 + w_2 = 1$ ，权值 w_1, w_2 的具体取值可以通过对图像库中的图像进行检索学习获得，分析各个子特征对检索结果的影响，对检索准确率贡献大的子特征赋予较大权值，贡献小的权值相应小些。在对图像进行检索学习时，引入平衡点(Break Even Point, BEP)参数。假设每次检索实际匹配数为 n_i ，则总的实际匹配数为 $\sum_{i=1}^S n_i$ ，其中 S 为图像库中图像的总数。每类图像的数目若为 M ，则总的理想正确匹配数为 $S \times M$ 。BEP 定义为实际的正确匹配数与理想正确匹配数的比例，如式(11)：

$$BEP = \frac{\sum_{i=1}^S n_i}{S \times M} \tag{11}$$

对 2 个子特征分别进行检索学习，计算其对应的 BEP 指数，根据指数确定其在相似性比较时的权值。

2.3 实验结果

信息检索中衡量一个算法性能的指标通常是查全率和查准率。其中，查全率表示从图像库中成功检索出相似图像的能力；查准率表示检索返回的图像中相似图像所占的比例。查全率和查准率之间具有互逆的关系，一般查全率较高时查准率较低，而当查准率较高时查全率又较低。

图 3 是 4 种不同算法的检索结果示例，左上角第一幅是待检索图像，每个算法返回最相似的 19 幅图像。其中图 3(a)是本文算法的检索结果，图 3(b)是拱高半径复函数^[4]的结果，图 3(c)和(d)分别是三角形函数^[3]和质心距离傅里叶描述子^[2]的检索结果。从图中可以看出，本文算法在针对这一形状时的检索结果中只有 2 幅错误检索图像，其他 3 种算法错误反馈的图像数目分别为 7,18 和 12。

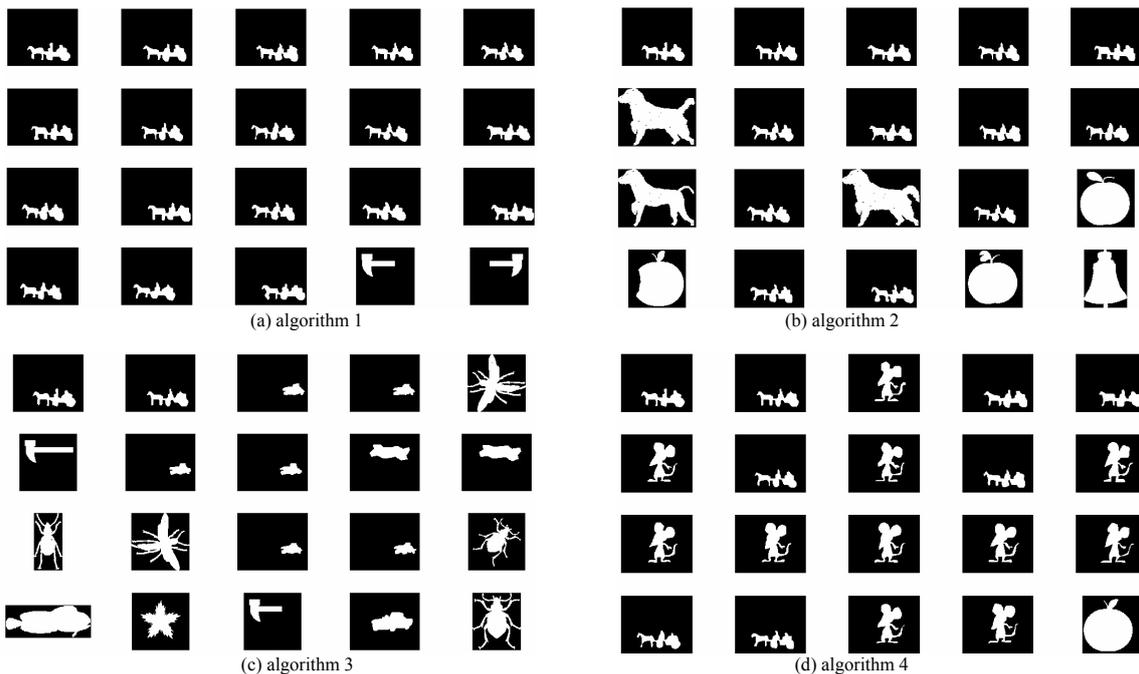


Fig.3 Results of image retrieval of different algorithms
图 3 不同算法的检索结果示例

为了能够客观全面地测试算法的性能，将 4 种算法在 MPEG-7 图形库上进行对比实验。对图形库中的每一幅图片都进行检索匹配。即将图形库中的某一幅作为待查询图形，在剩余的图形中检索同类图形，总共需进行 1 400 次匹配。其中每次检索都记录查全率分别为 10%,20%,...,100%时对应的查准率，绘制检索精度—召回率(Precision Versus Recall, PVR)曲线，用 PVR 指数评价检索结果。从表 1 可看出 4 种算法中，算法 3 的检索性能相对最差，算法 4 其次，算法 2 的检索平均准确率较前两者分别高出 4 个百分点和 9 个百分点。而算法 1(本文算法)又比算法 2 的平均准确率高出近 23 个百分点。实验数据显示，本文算法的检索性能显著优于参与比较的其余 3 种算法。

表 1 不同算法的查全率-准确率
Table 1 Recall-precision of different algorithms

algorithm	precision of different recalls									
	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
algorithm 1	0.928 6	0.809 1	0.709 0	0.634 0	0.576 7	0.514 7	0.462 5	0.406 8	0.338 1	0.229 4
algorithm 2	0.719 4	0.518 2	0.432 0	0.374 3	0.320 1	0.268 8	0.233 4	0.198 0	0.161 1	0.114 0
algorithm 3	0.647 1	0.399 7	0.309 9	0.258 4	0.210 7	0.173 0	0.145 8	0.114 5	0.086 9	0.060 9
algorithm 4	0.714 3	0.482 8	0.368 5	0.306 0	0.257 3	0.215 9	0.189 1	0.160 4	0.134 3	0.089 8

3 结论

现有的形状特征一般分为基于轮廓和基于区域两类特征。为克服轮廓描述法易受噪声等因素干扰的缺点, 本文将两种描述法结合, 提出一种融合轮廓和区域信息的新的形状描述算法。该算法将包含形状信息的轮廓采样点的拱高函数和质心距离组成的复函数, 通过 DCT 转换到频域中, 并以其低频系数构成轮廓形状特征。在对图形区域进行处理时, 同样获得表示区域形状信息的 DCT 系数。最后对两种形状特征进行组合, 构成形状描述子。在对 MPEG-7 图形库的检索结果显示, 该算法具有较高的检索准确率。今后将结合其他特征及相关反馈技术来获取更好的检索性能。

参考文献:

- [1] 梁洪, 李金, 鲍佩华. 基于组合特征双重加权的相关反馈算法[J]. 信息与电子工程, 2011, 9(4): 491-496. (LIANG Hong, LI Jin, BAO Peihua. Relevance feedback algorithm based on dual weighted combined features[J]. Information and Electronic Engineering, 2011, 9(4): 491-496.)
- [2] Zhang D S, Lu G J. Study and evaluation of different Fourier methods for image retrieval[J]. Image and Vision Computing, 2005, 23(1): 33-49.
- [3] Alajlan N, Kamel M S, Freeman G. Multi-object image retrieval base on shape and topology[J]. Signal Processing: Image Communication, 2006, 21: 904-918.
- [4] 王斌. 一种用于形状描述的拱高半径复函数[J]. 电子学报, 2011, 39(4): 831-836. (WANG Bin. Shape Description Using Arc-Height Radius Complex Function[J]. Acta Electronica Sinica, 2011, 39(4): 831-836.)
- [5] Mokhtarian F, Ung Y K, Wang Z T. Automatic fitting of digitized contours at multiple scales through the curvature scale space technique[J]. Computers & Graphics, 2005, 29(6): 961-971.
- [6] Kunttu I, Lepisto L. Shape-based retrieval of industrial surface defects using angular radius Fourier descriptor[J]. IET Image Processing, 2007, 1(2): 231-236.
- [7] Yadav R B, Nishchal N K, Gupta A K. 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.
- [8] Datta R, Toshi D, Li Jia. Image retrieval ideas influences and trends of the new age[J]. ACM Computing Surveys, 2008, 40(2).

作者简介:



张瑜慧(1979-), 女, 江苏省盐城市人, 硕士, 讲师, 主要研究方向为图像处理与模式识别.
email: zyh.zawh@126.com.

王海燕(1975-), 女, 江苏省宿迁市人, 硕士, 讲师, 主要研究方向为数据库、图像处理。

郑步芹(1981-), 女, 江苏省盐城市人, 硕士, 讲师, 主要研究方向为网络技术。

李 洁(1980-), 女, 江苏省宿迁市人, 硕士, 讲师, 主要研究方向为本体与信息检索。