

文章编号:1005-0523(2005)05-0147-04

# 基于内容的图像检索技术综述

范自柱

(华东交通大学 基础科学学院,江西 南昌 330013)

**摘要:**近年来,随着多媒体和网络技术的迅速发展,基于内容的图像检索技术(CBIR: Content-Based Image Retrieval)成为一个研究热点.本文介绍CBIR研究中的一些关键问题,并对其主要方法如基于颜色、形状、纹理等图像检索技术进行了讨论.

**关键词:**基于内容的图像检索;颜色空间;多媒体技术

**中图分类号:**TP391.4

**文献标识码:**A

## 1 引言

传统的图像数据库检索采用基于关键词或描述性文本的检索方式,查询时需要指明文本特征,要求用户对文本特征的描述具有一定的准确性和规范性.由于图像内容的丰富内涵以及人们对图像内容进行抽象时的主观性,使得这种方法往往不能满足人们的需要.

基于内容的检索(Content-Based Image Retrieval, CBIR)不同于传统的检索手段,它用于检索的是反映图像内容并与图像存储在一起的各种量化特征,使用的是基于相似性度量的示例查询(Query By Example)方法.目前,国外许多机构都在进行相关的研究,并推出了以IBM的QBIC<sup>[1]</sup>和MIT的PhotoBook<sup>[2]</sup>为代表的一系列成功的产品.

## 2 基于内容的图像检索的主要索引技术

基于内容的图像检索就是在图像库中找到查询者需要的图像.虽然使用目的和在书籍库查找某本书是完全类似的.但是就本质而言,书籍文献的查找可以实现精确定位,如知道书的名字就可以从

书库中查找到.而基于图像的检索就目前计算机和人工智能发展的水平来讲,完全是一种基于相似性的度量.虽然查询者的要求千差万别,但是查找的层次基本可以分为三层<sup>[25]</sup>:基于原始数据的查找、基于特征的查找和基于语义的查找.

目前研究和应用的层次主要是基于特征的查找.图像的特征就是一幅图像与其他图像不同的原始特性或根本属性.有的特征与图像的视觉外观相对应,具有原始性,如亮度,形状描述子,灰度纹理;还有一些特征缺少自然的对应性,如灰度直方图,颜色直方图,空间频谱图.不同的特征被划分在不同的特征矢量组.图像的区域划分基本上是根据特征矢量的不同特性.同一个区域内的特征矢量具有相同的属性.在n维特征空间的查找,实质就是模板特征矢量与存储在数据库内的特征矢量的相似性度量.数据库内的图像文件的特征矢量集以及抽取的方法被事先抽取并保存.查询时,使用者给出的模板图像按照库内保存的方法抽取目标特征矢量.基于特征的查找就是比较由图像中自动抽取的特征之间的相似性.最常用的特征包括颜色,纹理和形状,已有的软件系统都是基于这三种特征的检索.

### 2.1 颜色检索

收稿日期:2004-12-18

基金项目:华东交大科学基金项目(基于内容的图像检索技术研究)

作者简介:范自柱(1975-),男,安徽庐江人,讲师.

用颜色特征进行图像检索可以追溯到 Swain 和 Ballard<sup>[3]</sup>提出的颜色直方图的方法. 主要可以分为两类<sup>[24]</sup>: 全局颜色特征检索和局部颜色特征检索.

### 1) 全局色彩特征检索

全局色彩特征检索方法目前采用最多的是色彩直方图的方法, Swain 的主要思想是根据色彩直方图统计每种色彩在图像中出现的概率, 然后采用色彩直方图的交来度量两幅图像色彩的相似性, 其最大的缺点是完全丢失了图像色彩的空间信息. Pass<sup>[4]</sup>等人提出以图像的色彩聚合矢量 CCV (color coherence vector) 来作为图像的索引, 它是图像直方图的一种演变, 聚合矢量中的聚合信息在某种程度上保留了图像色彩的空间信息. Stricker 和 Orengo<sup>[5]</sup>提出了累计色彩直方图方法, 并且提出了色彩矩的方法, 认为色彩信息集中在图像色彩的低阶矩中. 他们主要对每种色彩分量的一阶、二阶和三阶矩进行统计.

### 2) 局部色彩特征检索

由于全局色彩特征检索捕获了整幅图像色彩分布的信息, 丢失了许多局部的色彩空间信息. 目前从划分局部区域的角度来说可分为: 基于固定块的图像分割、基于手工的区域分割、采用交互的半自动的区域分割以及一些自动的色彩分割方法. 局部区域中的色彩信息可以表示为平均色彩、主色彩、色彩直方图和二进制色彩集. Hsu<sup>[6]</sup>等人试图结合图像的色彩信息和图像色彩的部分空间信息对色彩的直方图进行检索. Smith 和 Chang<sup>[7]</sup>采用色彩的自动分割方法, 形成一个二进制的色彩索引集, 在图像匹配中, 比较这些图像色彩集的距离和色彩区域的空间信息. 色彩的空间关系主要有色彩区域的分离、包含、交, 每种关系对应一定的评分, 查询的空间距离是所有这些色彩区域所对应的空间关系的评分和.

关于图像色彩直方图之间的距离, 目前主要有直方图的交、直方图的绝对值距离, Niblack 和 Barber<sup>[8]</sup>等人采用直方图的欧几里德距离的度量方法, Hafner<sup>[9]</sup>等人提出了色彩直方图之间的加权距离作为直方图之间的度量方法.

综上所述, 人们对色彩索引方法的研究已经相当多, 它们都分别从不同的方面对色彩检索问题进行了研究. 但目前主要存在的问题是色彩空间选择并不完全统一, 有人使用 RGB 色彩空间, 更多的人采用 HSV, MUSHELL 等视觉意义上的色彩空间等, 而且这些方法之间关于色彩之间的度量方法也不

统一, 给各种方法检索效果的评价带来了一定的难度. 色彩检索方法仅仅是基于内容检索的一种方法, 在实际应用中和其它检索方法相结合, 效果可能更佳.

## 2.2 形状检索

形状描述是通过一些方法生成数值的描述子来描述形状, 描述子应在尽可能区别不同目标的基础上对目标的平移、旋转和尺度变化不敏感. 它的具体实现方法可以分为两类: 基于轮廓和基于区域方法.

对形状轮廓特征的描述主要有: 链码、样条和多边形逼近等方法. 链码是一种常见的形状表示方式, 它能有效的表示形状, 但是不能简化形状. 样条有最小化曲率的优点, 也就是用最小平均曲率的曲线近似给定的函数曲线, 其缺点是局部函数值的修改会影响整个样条表示. B 样条的提出就是为了不将局部函数值的改变传播到其它间隔中去. 它也可以用作由参数方程确定的平面曲线间的插值, 这样每一条参数方程都可以独立插值. Cohen<sup>[10]</sup>等人提出了一种基于 B 样条的曲线表示和匹配法. Saint-Marc<sup>[11]</sup>等人使用四次 B 样条表示形状, 并且由此判断形状的对称性. HannuOlkkonen<sup>[12]</sup>提出了离散 B 样条.

多边形逼近是用多边形线段来近似形状边缘, 即是以最小误差、最小多边形周长、最小多边形内部面积, 或最小多边形外部面积作为近似准则. 这类方法中最常用的是分裂和合并法, 在这个方法中, 曲线分裂由几个线段来表示直到误差达到可以接受. 同时分裂的线段又有可能融合, 如果融合后的线段同原始曲线的误差在允许的最大误差范围内, 线段即融合. Wu 和 Leou 提出了另一种不同的准则来获取多边形逼近, 他们使用的多边形逼近准则是最大内部面积、最小外部面积、最小面积偏差. Bengtsson 和 Eklundh<sup>[13]</sup>提出了一种层次化的多边形逼近方法. 分裂合并方法经常用于多边形逼近, 尺度空间的方法则常常用于跟踪曲线上的特征点. Chung<sup>[14]</sup>等人开发了一种基于 Hopfield 神经网络的形状多边形逼近方法. 这种方法是将多边形逼近定义为对神经网络能量函数的最小化问题, 就是最小化曲线和多边形的弧与弦之间的偏差.

对于形状的区域特征主要有形状的无关矩、区域的面积、形状的纵横比、主轴方向、凹凸面积<sup>[22]</sup>等, 这些特征得到了广泛的应用<sup>[15]</sup>. Flicker 等人在 IBM 的 OBIC 图像检索系统中采用了形状面积、圆

度、离心率、主轴惯量以及高阶无关矩,取得了满意的效果.

### 2.3 纹理检索

纹理是图像的又一重要特征,虽然尚无精确的定义,但却受到广泛研究.一般认为纹理就是纹理元素有规律的排列组合,而把具有重复性、形状简单、强度一致的区域看作纹理元素<sup>[23]</sup>.视觉纹理特性主要有:粗糙度、对比度、方向度、粗略度等,图像检索用到的纹理特征表示方法主要有<sup>[16]</sup>:Tamura法(反映了粗糙度、对比度、方向度等)、MRSAR(multi-resolution simultaneous auto-regressive model)、Canny角直方图法、Gabor变换.Tamura法是对整幅图像进行处理,效果较差,改进方法是采用直方图.和其他方法(纹理的主分量分析、小波变换特征等)的比较<sup>[17]</sup>可知,MRSAR<sup>[18]</sup>区分不同纹理模式的能力较强,但计算开销也较大.MRSAR与Gabor方法虽然性能较好,但其复杂度较高、处理时间也较长,在图像检索中使用基于TWT、PWT的方法比较合适.

## 3 基于内容图像检索技术发展趋势

目前的基于内容图像检索技术主要是采用前一节所述的一种或多种方法,若采用多种方法,则给每种方法赋以一定的权值,再结合用户检索时的反馈信息<sup>[19,20,21]</sup>,不断调整这些参数以达到最优效果.但是无论是使用颜色,形状还是纹理特征,每种特征都有各自的不足之处.仅用颜色会使图像的空间信息损失,形状特征的提取本身就是一个难题,纹理又有其局限性,一个纹理特征集一般只对某个图像数据库可行,而对其他图像库不用定有效.比较可取的方法是把这三种特征综合起来使用,但仍不能从根本上解决问题,因为这些都是图像的低层特征,它们和人眼的主观视觉特征存在着不一致的地方.

综上所述,未来的图像检索技术发展趋势主要有以下几个方面:

1) 图像检索的快速算法.目前流行的大多算法,不论是基于那种特征都有计算量大这一缺点,图像自身数据量大是其一个原因.但是,若能够采用更简洁而又精确的方法来描述图像,其计算复杂度就会降低.因此,在未来相当长的一段时间内,图像的快速检索仍是一个重要的研究课题,这也是图像检索技术实用化的必然要求.

2) 图像的智能检索.现今的很多图像检索技术

或多或少的采用了智能方法,但还远没有达到人们的要求.当然,模式识别和人工智能等领域的理论技术还有待突破,但这并不妨碍我们借鉴这些领域的理论,如数据挖掘、神经网络、智能搜索以及蚁群算法、人工免疫等优化方法.

3) 图像数据库模型的建立.传统的关系数据库模型在处理图像时会遇到很多问题,例如,怎样快速的读写图像数据,查询优化,抽取和存储图像特征,相似度计算等等.这些问题的较好解决方法是建立面向对象数据库,但也有不少问题,如面向对象数据库技术还需要完善.即使完善了,面向对象数据库和关系数据库还会共存,它们之间相互存取操作问题还得解决.所以,这也是今后一个重要的研究方向.

4) 反馈技术.图像检索的效果很大程度上需要由用户来评价,因此,用户的反馈信息在检索计算中不可忽视,它可以提高检索的效率和效果.当前的很多检索算法也用到反馈技术,但大部分算法都是将其作为辅助手段,这样,反馈技术就不能很好地优化检索算法.怎样设计用户的反馈方式和充分利用反馈信息是一个值得关注的课题.

## 4 总结

本文阐述了基于内容的图像检索技术原理方法,图像检索综合了图像处理、模式识别、数据库和人的认知心理等多种技术.它是计算机领域又一前沿技术,并越来越受到人们的关注,它的发展也越来越快,本文只对其发展作了概括介绍.可以预见,在不久的将来图像检索技术将会取得更大成果.

### 参考文献:

- [1] Flickner M et al. Query by image and video content[J]. The QBIC System. IEEE Computer, 1995, 28(9):23-32.
- [2] Pentland A P, Picard R, Sclaroff S. Photobook: Content-based manipulation of image databases[J]. International Journal of Computer Vision, 1996, 18(3):233-254.
- [3] Swain M, Ballard D. Color indexing[J]. International journal of Computer Vision, 1991, 7(1):11-32.
- [4] G, Zabih R, Miller J. Comparing images using color coherence vectors. In: Proc of ACM Intern Conf Multimedia. Boston, MA, 1996.
- [5] Stricker M A, Oren M. Similarity of color images. In: Proc of SPIE: Storage and Retrieval for Image and Video Databases III. San Jose, CA, 1995, 2420: 381~392.

- [6] Hsu W, Chua T S, Pung H K. An integrated color-spatial approach to content-based image retrieval. In: Proc ACMM Multimedia '95 Conference. San Francisco, 1995. 305~313.
- [7] Smith J R, Chang S F. Tools and techniques for color image retrieval. In: Proc of SPIE: Storage and Retrieval for Image and Video Database. San Jose, CA, 1996, 2670; 426~437.
- [8] Niblack W, Barber R et al. The QBIC project: Querying images by content using color, texture and shape. In: Proc of SPIE: Storage and Retrieval for Image and Video Database. San Jose, CA, 1994.
- [9] Hafner J, Sawhney H, Equitz W et al. Efficient color histogram indexing for quadratic form distance function. IEEE Trans on Pattern Analysis and Machine Intelligence, 1995, 17(7); 129~256.
- [10] CohenFS, ZhuangZYang. Invariant matching and identification of curves using B-splines curve representation. IEEE Trans. Image Processing, 1995, 4(1); 1~10.
- [11] Saint-Marc Philippe, Rom Hillel, Medioni Gerard. B-spline contour representation and symmetry detection. IEEE Trans. PAMI, 1993, 15(11); 1191~11973.
- [12] Olkkonen Hannu. Discrete binomial splines. GMIP, 1995, 57(2); 101~106.
- [13] BengtssonA, EklundhJ. Shape representation by multiscale contour approximation. IEEE Trans. PAMI, 1991, 13(1); 85~944.
- [14] ChungP, TsaiC, ChenE, SunY. Polygonal approximation using competitive Hopfield neural network. Pattern Recognition, 1994, 27; 1505~1512.
- [15] AngYH, LiZhao, OngSH. Image retrieval based on multidimensional feature properties. SPIE, 1995, 2420; 47~57.
- [16] Ma Weiyang, Zhang Hongjiang. Benchmarking of image features for content-based retrieval [A]. The Thirty-Second Asilomar Conference on Signals, System & Computers [C], California, USA; 1998, 1; 253~257.
- [17] B S Manjunath, W Y Ma. Texture features for browsing and retrieval of image data[J]. IEEE T-PAMI, 1996, 18(8); 837~842.
- [18] J Mao, A K Jain. Texture classification and segmentation using multiresolution simultaneous autoregressive models [J]. Patter Recognition, 1992, 25(2); 173~188.
- [19] H Greenspan, G Dvir, and Y Rubner. Region correspondence for image matching via EMD flow. In CVPR 2000 Workshop on Content-Based Access of Image and Video Libraries, 2000, pp. 27~31.
- [20] ———, Support vector machines for region-based image retrieval. in Proc, IEEE Int. Conf. Multimedia & Expo, 2003.
- [21] Feng Jing, Mingjing Li, etc. An Efficient and Effective Region-Based Image Retrieval Framework [J]. IEEE Transactions on Images Processing, 2004, 13(5); 699~710.
- [22] 丁险峰等, 形状匹配综述[J]. 自动化学报, 2001, 27(5); 678~694.
- [23] 黄祥林, 沈兰荪. 基于内容的图像检索技术研究[J]. 电子学报, 2002, 30(7); 1066~1071.
- [24] 李向阳, 庄越挺, 潘云鹤. 基于内容的图像检索技术与系统[J]. 计算机研究与发展 2001, 38(3); 344~354.
- [25] 王文惠, 周良柱, 万建伟. 基于内容的图象检索技术的研究和发展[J]. 计算机工程与应用, 2001. 5; 54~56.

## A Survey of Content-based Image Retrieval

FAN Zi-zhu

(School of Natural Science, East China Jiaotong University, Nanchang 330013, China)

**Abstract:** In recent years, with the rapid development of the technology of multimedia and network, Content-based image retrieval (CBIR) has been an active research topic. In this paper, some key issues on CBIR are introduced. Then image retrieval methods based on color, shape and texture are discussed.

**Key words:** content-based image retrieval (CBIR); color space; multimedia technology