

何立民 万跃华

# 数字图书馆中基于内容的图像检索关键技术

**摘要** 数字图书馆基于内容的图像检索数据库研究中,要研究的关键技术包括:建立数据模型,基于内容的编码,基于图像颜色特征、纹理特征、形状特征、空间关系特征的索引,以及检索效果的评价标准和高维索引等。参考文献10。

**关键词** 数字图书馆 多媒体信息检索 基于内容的图像检索

**分类号** G250.76

**ABSTRACT** In the studies of content-based image retrieval databases in digital library, key techniques include data modeling, content-based encoding, indexing based on image characteristics, evaluation criteria of search results, and high-dimensional indexing. 10 refs.

**KEY WORDS** Digital library. Multimedia information retrieval. Content-based image retrieval.

**CLASS NUMBER** G250.76

## 1 数字图书馆基于内容的图像检索主要关键技术

数字图书馆基于内容的图像检索(Content-Based Image Retrieval, CBIR)是指直接根据描述媒体对象内容的各种特征进行检索,它能从数据库中查找到具有指定特征或含有特定内容的图像(包括视频片段)<sup>[1]</sup>。它区别于传统的基于关键字的检索手段,融合了图像处理、图像理解、模式识别、计算机视觉、数据库等技术。

基于内容检索时根据媒体特征进行相似性匹配检索的媒体特征有:颜色、纹理、轮廓、形状、空间约束、动态、领域概念、结构描述及其他的图像信息。在过去的几年中基于内容的图像检索已经成为非常活跃的研究领域,并开发出许多具有可视化特征的系统<sup>[2]</sup>。其中建立图像数据库是一个重要环节,它是图像处理技术和传统数据库技术相结合的产物。实现数字图书馆基于内容的图像检索的主要关键技术包括以下几个方面。

### 1.1 基于内容的图像数据库数据模型结构

支持基于内容检索的图像数据库系统的数据模型结构,由于研究的角度不同,可以分为图像数据库的体系结构和框架结构。

较完善的图像数据库的体系结构是 Chang SK 提出的五层结构:用户视图、语义特征视图、图像特征视图、特征表达、特征的组织 and 检索<sup>[3]</sup>。在用户视图

层,用户信息主要是面向高层的事件,其主要任务是进行图像的空间推理。语义特征层主要对某一图像特征在特定领域中的语义信息进行描述。在特征视图层,主要是分析图像的内容,如图像的空间关系、形状、色彩等,这一层用户主要关心图像所包含的视觉内容。对于一种视觉内容特征来说,由于其特征的表达不是惟一的,如“圆”可以用其轮廓来表示,也可以用圆心和半径来表示,因此在特征层应支持同一特征的多种表示。图像的特征组织层等价于传统数据库中的物理存储结构(如B-树,Hash表等)。可以说这是对一个图像数据库发展整体框架的总体概括,从图像数据库发展的阶段来看,目前正处在第3阶段,即支持图像内容的检索阶段。目前虽然有一些图像数据库以关键词方式对图像语义特征进行检索,但这种检索是脱离图像的具体内容的。

在 Web 环境下实现一个基于内容检索的图像数据库系统的框架结构可分为两部分:图像特征索引的离线创建和在线图像检索。一般来说,图像特征的分析 and 索引的创建是在服务器端离线完成的,其主要功能包括图像入库前的预处理、图像的内容特征的提取与分析、图像特征描述的编码和存储。对于图像的在线检索,其主要的任务是将用户感兴趣的图像实例提交给服务器对其内容的特征进行提取和表示,然后调用图像检索引擎按一定的相似检索方法进行图像之间的相似度计算,对查询得到的相

似图像按它们的相似距离从小到大进行排序,并将

### 1.2 基于颜色特征的检索

这是 CBIR 中最基本的方法。颜色是物体表面的一种视觉特征,每种物体都有其颜色特征。颜色是图像内容组成的基本要素,是人识别图像的主要感知特征之一。相对于其他特征,颜色特征非常稳定,对于旋转、平移、尺度变化,甚至各种形变都不敏感,而且颜色特征计算简单,因此在现有检索系统中应用最广泛。

#### 1.2.1 颜色模型

根据不同的应用场合,图像的颜色表示方式各不一样。CRT 显示屏使用的是 RGB 颜色模型,利用红、绿、蓝 3 种基本颜色以加色法的混色原理来混合。Lab 模型使用亮度 Light 和 2 个彩色分量 a 和 b 来表示颜色。HSV 是一种符合人类视觉感知特征的颜色模型,它把彩色信号表示为 3 种属性:色调、饱和度及亮度,这种颜色模型用 Munsell 三维空间坐标系统表示,具有与人观察颜色的方式相一致,有利于图像处理特点<sup>[4]</sup>。

为了得到与人的感觉类似的相似性度量,一些研究者还提出采用其他的颜色空间模型和距离测度。Zhang 等采用了符合人眼感觉的 HSI 模型。Hafner 等在进行直方图匹配时引入了二次型距离。另外,在提高检索对于光照的稳定性、空间分布信息的引入等方面,也出现了很多算法。存在的问题是颜色空间模型选择并不完全统一,关于颜色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一定难度。

#### 1.2.2 颜色特征索引

用颜色特征进行检索可以追溯到 Swain 提出的基于颜色直方图检索方法,它计算两幅图像的三维颜色直方图的每一个颜色单位,并进行细致的比较<sup>[5]</sup>。由于颜色直方图具有简单且随图像的大小、旋转变化不敏感等特点,得到研究人员的广泛关注,目前几乎所有的 CBIR 数据库系统都把颜色检索方法作为检索的重要手段,并提出了许多改进方法。归纳起来可以分为两类:全局颜色特征索引和局部颜色特征索引。

其他的基于颜色特征索引的方法有:颜色相关图法、累计颜色直方图方法、颜色矩的方法、颜色元组直方图的方法、颜色集方法、颜色聚类方法等。

由于全局颜色特征索引捕获了整幅图像颜色分布的信息,丢失了许多局部的颜色空间信息,许多人提出了颜色索引的改进方法——局部颜色特征索

结果返回给用户。

引。从划分局部区域的角度可分为: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颜色分割方法。局部区域中的颜色信息可以表示为平均颜色、主颜色、颜色直方图和二进制颜色集。

### 1.3 基于纹理特征的索引

纹理通常定义为图像的某种局部性质,或是对局部区域中像素之间关系的一种度量,其基本单位是纹理元。纹理特征可用来对图像中的空间信息进行一定程度的定量描述。

纹理分析的方法基本可以分为统计法、结构法、模型法和空间法/频率域联合分析法等四类。基于统计的方法是对图像中的颜色强度的空间分布信息进行统计,包括共生矩阵法、Laws 纹理能量法等,主要用于分析像木纹、沙地、草坪等细致而不规则的物体。基于结构的方法将重点放在分析纹理元之间的相互关系和排列规则上,适用于像布料或砖瓦等一类元素组成的纹理以及排列比较规则的东西。基于模型的方法假设纹理按某种类型分布,如 Markov 随机场模型、分形模型等。基于空间/频率域联合分析法主要包括 Gabor 变换法和小波变换法等<sup>[6]</sup>。这些方法大多都是查询特定的图像,而基于纹理元灰度的图像分析方法适合于大多数图像对象的检索。纹理特征提取的一种有效方法是以灰度级的空间相关矩阵即共生矩阵为基础的。Tamura 以人的主观心理度量为标准,提出了 6 个基本的纹理特征:粗糙度、对比度、方向度、线像度、规整度和粗略度<sup>[7]</sup>。其中最重要的特征是粗糙度、对比度和方向度,这些纹理特征集很好地对应于人类视觉感知,在许多图像检索系统中得到应用。随后又有不少人提出了纹理的特征集。目前还没有一个统一的标准来精确地表示纹理的特征,因为人对纹理的视觉特征的认识非常主观。

### 1.4 基于轮廓和形状特征的检索

轮廓的特征量能够较好地区别不同内容的图像。在二维图像空间中,形状通常被认为是一条封闭的轮廓曲线所包围的区域,所以对形状的描述涉及到对轮廓边界的描述以及对这个边界所包围区域的描述。利用形状来检索可以提高检索的准确性和效率。寻找符合人眼感知特性的形状特征不是一件简单的工作,首要的困难是要将不同物体从图像中分割出来,这也是计算机视觉的困难问题之一,至今没有很好解

决。目前的基于形状检索方法大多围绕着从形状的轮廓特征和形状的区域特征建立图像索引。

对形状轮廓特征的描述主要有:直线段描述、样条拟合曲线、傅立叶描述以及高斯参数曲线等。关于形状的区域特征主要有形状的无关矩、区域的面积、形状的纵横比等。

### 1.5 基于对象空间关系特征的检索

早在1976年,Tanimoto提出了用像元方法来表示图像中的实体,并提出了用像元来作为图像对象索引。随后被Chang SK等采纳,并提出用二维符号串(2D String)的表示方法来进行图像空间关系的检索<sup>[8]</sup>。由于该方法简单,并且对于部分图像来说可以从2D String重构它们的符号图,因此被许多人采用和改进。该方法的缺点是:仅用对象的质心表示空间位置是不够的;对于一些图像来说人们不能根据其2D String完全重构其符号图;上述的空间关系太简单,实际中的空间关系要复杂得多。针对这些问题,许多人提出了改进方法。如广义2D String(2DG String)的方法、2DC String的方法、2DC<sup>+</sup> String的方法、2DB String的方法、RS String表示方法、2D PIR检索方法、高空间关系特征的图像检索方法等。

### 1.6 多特征的综合检索方法

一幅图像往往具有多种特征,如颜色、纹理、形状等。即使是同一特征,也有多种不同的表示方法。例如颜色特征,可以用直方图、颜色距、颜色集、主色调等多种特征来表示。这些表示方法从不同角度反映了媒体的特征,如何组织这些特征,使应用能够调用合成的特征和特征表示来支持查询,按照用户的查询要求合并各种特征的查询结果,来解决多种检索手段相结合的问题,以提高图像检索的效率,是需要研究的一个问题。

图像多特征的综合检索方法目前还没有得到解决,有许多学者已经开展了这方面的研究。如Mehre等提出了在一些技术上结合形状和颜色特征的合成特征标准,已经开发了一个给定一对图像后计算配合度的一个相似标准原型系统,该系统被用于基于形状或颜色以及两者结合的图像检索中<sup>[9]</sup>。

### 1.7 图像多特征的相关反馈检索技术

目前,数字图书馆基于内容检索中存在的一个主要问题是现有的检索方法都是以计算机为中心,使得一些查询结果计算机认为是相似的,而人却认为是不相似的。要解决这个矛盾,一方面对上面的检索方法进行改进;另一方面就是增进人机之间的通信,计算机将查询的信息反馈给人,人对查询结果

的评判信息反馈给计算机,这种反馈的结构是让计算机能够具有一定的学习能力。

Rui等将相关反馈方法用于图像的检索中<sup>[10]</sup>。对于查询结果用户可以作出5种选择:很相关、相关、默认、不相关、很不相关,这些选择对应于一定的评分,系统然后重新计算每个权值,并进行动态更新。

### 1.8 基于内容的图像检索效果的评价

其目的是在相同的条件下找出最佳算法,使不同的检索方法能改进和提高,使基于内容的检索方法朝更好的方向发展。检索效果通常用查全率、查准率和检索时间这3个尺度来衡量。从现在的情况看,对于系统的响应时间和吞吐率的评价论述得较少,对检索效果评价更多地放在对检索结果的正确与否,主要使用的是查全率和查准率两个指标。

## 2 典型的基于内容检索原型系统与数字图书馆系统

因特网上已经有许多基于内容检索的商业系统、研究原型系统与数字图书馆系统。它们一般实现了如下功能:图像的随机浏览功能;按图例检索的功能;按草图检索的功能;按图像类别浏览的功能。

QBIC是IBM Almaden研究中心开发的第一个商用基于内容的图像及视频检索系统,使用64位颜色直方图,纹理检索采用改进的Tamura模型;采用KL变换降低维数,再利用R\*树建立索引。它同时支持关键词检索以及内容检索。QBIC系统已经在IBM数字图书馆得到应用,并应用到多个数字图书馆项目。

VisualSEEK图像查询系统和WebSEEK图像及视频搜索引擎是美国哥伦比亚大学开发的基于内容检索原型系统。该系统用到了图像区域的空间关系查询和直接从压缩数据中提取视觉特征。所用到的视觉特征有颜色集、纹理特征的小波变换,使用二进树进行索引,支持空间位置关系的查询。采用了形状识别和相似性计算来提高检索效率。为了加快检索过程,还开发了基于二叉数的索引算法。

麻省理工学院媒体实验室开发的PhotoBook人脸特征图像识别系统,包括轮廓、纹理、脸部特征查询;根据用户反馈采用不同的模型。

密歇根大学和麻省理工学院多媒体实验室开发的Virage图像系统,通过对不同单特征空间中的距离赋予权重,构成总特征空中的复合距离,通过调节权重和检索的特征值,表达不同的检索请求。

NETRA是加州大学圣巴巴拉分校亚历山大数字图书馆项目组开发的一个图像检索原型系统。NETRA利用图像分割区域中的颜色、纹理、形状和空间

位置关系来检索数据库中的相似区域。该系统的主要检索特征是基于纹理分析的 Gabor 滤波器。

DVLS 是美国堪萨斯大学的数字视频图书馆系统,此系统的目标是存储、索引及检索图像、音视频信息,已建立了一个称为 VISION 的原型系统及一个视频数据库。

Infomedia 是卡内基 梅隆大学的集成声音、图像和语言理解技术创建和探索的数字图像视频图书馆系统,结合语音识别、图像视频分析和文本检索技术支持 2000 小时的图像、视频广播的检索;实现全内容的、基于知识的查询和检索。

### 3 研究中存在的主要问题及进一步研究方向

#### 3.1 数据模型

目前,许多研究关注的是图像的特征提取及检索匹配算法,虽然方法很多,但没有一个统一的模型。数据模型是基于内容的图像检索系统的核心,模型决定了系统的查询和检索能力。建立的数据模型要能够充分反映媒体对象的内容,反映与领域无关、能够有效支持存储的物理特征和逻辑特征。要求能够表达无结构的图像实体,以层次结构组织图像的语义特征、图像特征和图像实体,可由基本图像特征构造复杂的图像特征,能根据特定领域知识从图像特征导出语义特征,支持空间推理。现有的数据模型还不能满足这些要求。图像数据检索模型的研究将设计图像内容描述的定义、图像内容描述的操作、图像内容的描述查询、图像描述的存储、组织和编码。

#### 3.2 基于内容的编码

存在的主要问题是使用的媒体的格式和编码没有考虑到内容,只是针对颜色、像素、样值来编码。从这些数据中抽取内容特征非常困难。如果我们在对媒体数据编码表示时就考虑到媒体的内容,那么,对这些数据的内容进行检索就会更有效更准确。国际上 MPEG 标准化组织制定的 MPEG 7 标准,其目标就是实现基于内容的编码。

#### 3.3 基于图像颜色特征的索引

存在的主要问题是人对颜色特征的视觉感知方面考虑得仍然不够。虽然目前大多数基于颜色特征的图像检索采用了和人对颜色感知相一致 HSV 颜色空间,但关于两种颜色之间的相似度的定义和视觉上人对相似颜色的判定仍有一定的差距。从颜色特征的表现来看,各种形式的颜色直方图是最常采用的表示方法,一般指定相同的颜色集,采用几十到上百维的高维直方图。实际上人对两图像画面的颜色

的相似性判定主要考虑少数几种主要的颜色。不同的图像有不同的颜色集,对包含不同颜色集两图像之间的相似性判定仍需要进一步研究。为进一步提高颜色特征的检索效果,本领域的研究将逐渐走向更接近人类心理学和人类视觉以及颜色物理几方面进行综合深入研究。

#### 3.4 基于图像纹理特征的索引

存在的主要问题是各种方法所选择的纹理特征集依赖于具体的纹理图像。往往是一种方法所选择的纹理特征集对表达一个纹理图像库比较有效,但对另一个纹理图像库来说不一定可行。对于不同的纹理图像库如何进行纹理特征集的自动切换仍需要较长时间进行研究。在结合纹理描述方面,如何借助与人的视觉有关的参数,及采用小波变换提取纹理特征等都有许多值得研究的问题。

#### 3.5 基于形状特征的索引

形状特征提取是非常繁重的工作。各种形状特征表达方法对形状信息的丢失非常严重;只有少量的形状特征表达方法和形状的几何变换无关;再则形状度量方法仍不具有很好的形状区分能力,不能有效表达形状之间的相似性。研究形状特征检索仍然是基于内容检索较具挑战性的研究课题。

#### 3.6 基于空间关系特征的索引

主要问题是如何保证各种空间关系的方法与图像的旋转无关;如何实现空间特征的相似度量从定性到定量的转变,目前仍无很好的解决办法。在图像多特征的相关反馈检索中,由于不同的特征其度量空间是不一样的,如何将这些距离转变为图像之间的相似度并能准确地表示人对图像之间的相似性认识是非常难的一件事情,仍需要开展大量的心理实验研究。

#### 3.7 相似性判断

图像检索的效果很大程度上取决于匹配算法的优劣,如何以一定的标准来判断图像内容是否相关仍是今后的研究方向。

#### 3.8 检索效果的评价标准

由于图像内容的丰富性和复杂性,以及人对图像内容评判的主观性,使得其检索性能的优劣很难有一个统一的标准。这也是 CBIR 技术研究的一个方向。目前基于内容检索结果的评价方法采用的是信息检索中的查全率和查准率方法,远非令人感到满意,定义一个很好的评价方法,困难主要在于人对图像内容认知上的主观性。检索效率将是今后研究中需要解决的问题。

#### 3.9 基于语义的图像检索

虽然基于语义的图像检索有许多缺点,但在某些情况下仍不失为一种主要方法。人工智能研究中,对图像语义特征的表示方面已经有许多成果,可以应用在CBIR中分析图像视觉特征和图像语义特征之间的转换,使得高层概念和低层视觉特征之间沟通成为可能。国际MPEG标准化组织2001年制定的MPEG7标准,其目标就是实现集高层语义特征和低层视觉特征的基于内容的多特征综合检索,今后研究的热点之一将是高层的基于语义内容的图像检索。

### 3.10 认知科学在图像检索中的应用

在研究基于内容的图像检索中,应以认知科学的研究成果分析图像内容的特征和人对图像的认知。图像信息在人脑中的长期记忆为心像,人对心像的记忆、检索等操作过程实际上是形象思维过程,形象思维科学中关于心像的表征和计算模型将对基于内容的图像检索提供一定指导。

### 3.11 多特征的综合检索方法

对于单一检索手段,由于其约束信息不足,在返回目标图像的同时往往会返回大量其他也满足此检索要求的图像。采用多个检索手段相结合的方法无疑可提供更多的约束而使得返回图像中目标图像的比率得到提高,但检索手段间的融合是所要解决的问题。此外,使系统在与操作者的检索交互中进行学习,更好地理解检索的内容,以及使检索性能更接近人类视觉的特性,也是研究中要解决的问题。

### 3.12 内容描述标准

对于多媒体数据,需要一个标准来描述其内容,以便地进行基于内容的检索。MPEG7国际标准描述的特征包括标注信息、生产信息、统计信息、合成信息、客观特征、主观特征以及概念等。它的范围不包括特征提取和搜索引擎,目的是为了保留竞争余地。可进一步开展存取接口、特征抽取、搜索引擎以及多媒体应用研究方面的工作。

### 3.13 相关反馈

不断修正用户的查询和相似度匹配算法,从而使数据库中的图像分类更接近使用者的主观愿望,使检索结果符合使用者的个性化要求。目前基于相关反馈的方法存在的主要问题是局部收敛问题。

### 3.14 高维索引

数字图书馆中,图像的内容特征非常复杂,其特征矢量是高维的,尤其是在集成的检索中,特征矢量高达102级,常规数据库的索引方法已经不适用于多媒体数据的检索。因此,迫切需要研究新的索引结构和算法,以支持快速检索。目前一般采用降低维数,再采用多维索引的办法来构造索引,但维数的降低不可避免地会造成信息丢失。为了更好地支持多特征、异构特征、权重、主键特征等方面的要求,研究

和探索高维索引方法将会成为今后的研究热点。

## 4 结束语

数字图书馆基于内容的图像检索数据库系统,投资少,见效快,效益大,适合从互联网上源源不断地采集图像数据,进行分析、归类与重组,发现新线索、新现象和新规律,用以指导数字图书馆实验工作的设计。这是一条既快又省的科研路线。可避免不必要的重复,少走弯路,提高我国基于内容的检索技术水平。目前的研究中有一点特别要引起重视:尽快研究开发出具有我国自主知识产权的数字图书馆图像检索数据库。随着电子出版物、网络化信息资源的普及利用,CBIR技术的应用将会越来越普遍。目前CBIR仍处于初步研究探索阶段,我国急需跟踪这方面的前沿研究和实际应用系统,引进国外先进技术与系统,早日将我们自己的有自主知识产权的数字图书馆基于内容的图像检索数据库系统推上互联网。

## 参考文献

- 1 卢汉清,孔维新,廖明等.基于内容的视频信号与图像库检索中的图像技术.自动化学报,2001,27(1)
- 2 Rui, Yong et al. Relevance feedback: A power tool for interactive content-based image retrieval. IEEE Transactions on Circuits and Systems for Video Technology. 1998,8(5)
- 3 Chang S K. Image information systems: Where do we go from here. IEEE Transon Knowledge and Data Engineering, 1992, 4(5)
- 4 李国辉,柳伟,曹莉华.一种基于颜色特征的图像检索方法.中国图像图形学报,1999,4(3)A
- 5 Swain MJ, Ballard D H. Color indexing. International Journal of Computer Vision. 1991, 7(1)
- 6 盛文,杨江平,柳健,吴新建.一种基于纹理元灰度模式统计的图像纹理分析方法.电子学报,2000,28(4)
- 7 Tamura, H et al. Textural features corresponding to visual perception. IEEE Transactions on Systems, Man and Cybernetics, 1978, 8(6)
- 8 Chang S K, Shi Q Y, Yan C W. Iconic indexing by 2-D strings. IEEE Trans on Pattern Analysis and Machine Intelligence, 1987,9(3)
- 9 Mehre, Babu M. et al. Content-based image retrieval using a composite color-shape approach. Information Processing & Management, 1998, 34(1)
- 10 Rui, Yong; Huang, Thomas, S.; Mehrotra, Sharad. Content-based image retrieval with relevance feedback in MARS. IEEE International Conference on Image Processing v 2 1997, IEEE Comp Soc, Los Alamitos, CA, USA, 97 CB36144

何立民 浙江工业大学图书馆馆长,副研究馆员。通讯地址:杭州市。邮编 310032。

万跃华 浙江工业大学图书馆信息咨询部主任,馆员。通讯地址同上。

(来稿时间:2002-04-12)