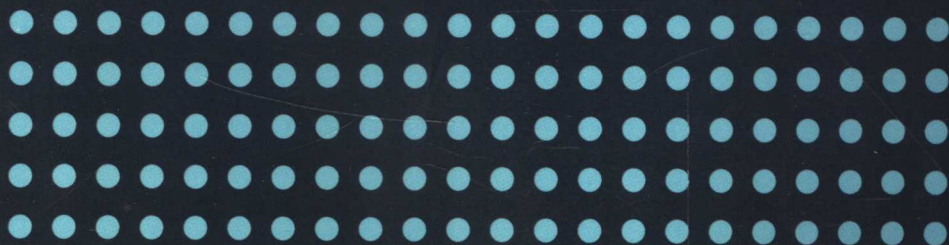


JYU SHUJUKU FANGSHI DE YAOGAN TUXIANGKU NEIRONG JIANSUO

基于数据库方式的 遥感图像库内容检索

陆丽珍 著



气象出版社

China Meteorological Press

基于数据库方式的 遥感图像库内容检索

陆丽珍 著



气象出版社

内 容 简 介

本书针对遥感图像库内容检索研究中存在的主要问题,采用理论联系实际应用的方法进行研究。理论上充分借鉴国内外相关研究成果,合理界定遥感图像内涵,提出可重用的遥感图像概念模型 URSICM,将遥感图像的元数据、原始像元信息、视觉特征、图像对象以及语义内容等信息纳入统一框架;探索遥感图像库内容检索新方法,发展融合颜色纹理特征和基于 GIS 语义的遥感图像检索新方法,并针对 CTFFBIR 设计优化方案。实际应用中通过基于 URSICM 的遥感图像数据组织与存储方案、遥感图像检索引擎 RSIRE 的设计,完成遥感图像库内容检索原型系统 RSIQuery 的开发。

本书适宜于从事图像处理、遥感信息管理与检索、3S 集成等方面研究的学者,也可作为地理信息系统、遥感科学与技术等专业本科学生和研究生的教学参考书。

图书在版编目(CIP)数据

基于数据库方式的遥感图像库内容检索/陆丽珍著. —北京:气象出版社,2005.9
ISBN 7-5029-4028-6

I. 基… II. 陆… III. 遥感图像-图像数据库-检索方法 IV. G354.4

中国版本图书馆 CIP 数据核字(2005)第 111881 号

Jiyu Shujukuo Fangshi de Yaogan Tuxiangku Neirong Jiansuo

基于数据库方式的遥感图像库内容检索

陆丽珍 著

责任编辑:丁问微 陶国庆 终审:朱文琴

封面设计:陈璐

版式设计:谷清 责任校对:王丽梅

出版发行:气象出版社

出版社地址:北京市海淀区中关村南大街 46 号

邮政编码:100081

出版社电话:68407112

传真号码:62176428

电子邮箱:cmp01@263.net

出版社网址: <http://cmp.cma.gov.cn/>

印刷:北京昌平环球印刷厂

版次:2005 年 9 月第 1 版

开本:787mm×960mm 1/16

印次:2005 年 9 月第 1 次印刷

印张:13.5

印数:1~4000

字数:272 千字

定价:30.00 元

本书如存在文字不清、漏印以及缺页、倒页、脱页等,请与本社发行部联系调换

前 言

遥感技术、计算机技术、网络技术的快速发展,使得各领域研究者获取所需要的高精度、高分辨率、多时相遥感图像成为可能,但与之对应的却是遥感图像检索理论和技术的严重滞后。如何从海量遥感图像库中快速准确地检索到所需要的信息具有十分重要的意义。本书从遥感图像检索研究现状和存在的问题出发,发展了通用遥感图像概念模型 URSICM,设计了面向对象的逻辑组织与数据存储方案,提出了融合颜色纹理特征 CTFFBIR 和基于 GIS 语义的遥感图像检索新方法,探讨了特征相似性检索的优化方案,设计并实现了支持内容检索的遥感图像数据库引擎 RasSDE,在此基础上开发了原型系统 RSIQuery,为遥感图像库的检索与管理提供新的思路。

本书作者陆丽珍博士从事 GIS 和遥感信息处理的教学和科研工作已有十多年,从中积累了丰富的 GIS 和 RS 理论知识和实践经验。三年前开始的博士学位论文研究,为本书的编写积累了丰富的素材。作者旨在通过本书与各位读者探讨遥感信息处理领域的一些前沿问题。

本书第一章通过对 CBIR 和遥感图像库内容检索(Remote Sensing Image Base Content Retrieval,简称 RSIBCR)研究历史的回顾,以及对 RSIBCR 研究存在问题的分析,引出了本书的研究内容。第二章是本书相关理论和关键技术的文献综述,为后续章节的讨论做好理论与技术的铺垫。第三章提出通用遥感图像概念数据模型 URSICM,并探讨了基于 URSICM 的面向对象逻辑模型以及数据组织与存储方案。第四章讨论栅格空间数据库引擎的部分关键技术,并针对第三章提出的支持内容检索的遥感图像数据模型,详细阐述了面向网络环境下 RasSDE 原型系统的体系结构设计,以及主要功能模块的实现方法。第五章提出一种融合颜色与纹理特征的遥感图像检索 CTFFBIR 新方法。主要对 2D Gabor 纹理特征的提取、二叉树图像分解方法以及优化检索算法进行研究,并给出算法的核心思想,设计算法的

处理流程。第六章提出基于GIS语义的遥感图像检索方法GISSBIR,该方法充分利用了与遥感图像具有相同坐标框架的GIS数据描述空间对象语义和空间关系的能力,实现快速有效检索图像的目的。第七章基于前几章的研究,对遥感图像数据库内容检索原型系统RSIQuery的设计与实现展开论述。第八章介绍了遥感图像库内容检索RSIBCR的研究热点和趋势。

本书在编写过程中受到刘南教授、刘仁义教授的悉心指导,潘德炉院士、杨树峰教授提供了宝贵意见,在此表示衷心的感谢!李伟博士、谢炯博士、严志民博士、洪紫萍老师、吴宏海老师、童兆国、沈林芳、胡伟忠、滕龙妹、王洁、傅俊等学弟妹给予了作者相应的支持,在此一并表示感谢!万分感谢家人的牵挂和关心!

由于时间仓促,水平有限,书中难免有错误和疏漏之处,敬请各位专家、老师、读者给予批评指正。

陆丽珍

2005年8月于浙江大学

目 录

前言

第一章 绪论	(1)
§ 1.1 研究背景	(1)
§ 1.2 图像内容检索的发展与研究现状	(2)
1.2.1 图像检索技术的发展	(2)
1.2.2 研究内容与国内外研究现状	(6)
§ 1.3 遥感图像库内容检索研究现状与存在的问题	(9)
1.3.1 国内外研究现状	(9)
1.3.2 存在的主要问题	(12)
§ 1.4 研究内容与章节安排	(13)
1.4.1 研究内容	(13)
1.4.2 本书组织	(15)
第二章 遥感图像库内容检索关键技术综述	(17)
§ 2.1 遥感图像数据库关键技术	(17)
2.1.1 遥感图像数据存储与管理方式	(17)
2.1.2 索引机制	(20)
2.1.3 数据压缩技术	(22)
2.1.4 多分辨率分块金字塔构建	(28)
§ 2.2 视觉特征提取与相似性度量	(33)
2.2.1 视觉特征描述与提取	(33)
2.2.2 相似性度量方法	(45)
§ 2.3 相关反馈与检索算法评价	(47)
2.3.1 相关反馈机制	(47)
2.3.2 图像检索算法的评价	(51)
本章小结	(52)
第三章 支持内容检索的遥感图像数据模型	(53)
§ 3.1 空间数据模型分析	(53)
3.1.1 空间概念模型	(54)
3.1.2 空间逻辑模型	(59)

3.1.3	空间物理模型	(60)
§ 3.2	通用遥感图像概念模型	(61)
3.2.1	遥感图像的特点和内容	(61)
3.2.2	通用遥感图像概念模型	(64)
§ 3.3	面向对象遥感图像数据模型	(68)
3.3.1	面向对象数据模型及特征	(69)
3.3.2	面向对象遥感图像数据模型	(70)
3.3.3	OORSIDM 详细结构	(72)
§ 3.4	支持内容检索的遥感图像数据组织	(75)
3.4.1	栅格数据组织方式	(75)
3.4.2	主流 SDE 栅格存储机理分析	(77)
3.4.3	关系数据库系统的选择	(83)
3.4.4	遥感图像数据的组织与存储	(88)
	本章小结	(94)
第四章	支持内容检索的遥感图像数据库引擎设计与实现	(95)
§ 4.1	遥感图像数据管理的发展趋势	(95)
4.1.1	空间数据库引擎技术	(95)
4.1.2	基于对象-关系数据库的栅格空间数据库引擎	(97)
§ 4.2	栅格空间数据库引擎关键理论和技术分析	(98)
4.2.1	分布式 GIS 技术	(98)
4.2.2	海量影像数据可视化与异步双缓冲技术	(104)
4.2.3	多源空间数据集成	(108)
4.2.4	栅格空间数据库性能优化	(109)
§ 4.3	RasSDE 原型设计	(111)
4.3.1	RasSDE 设计思想	(111)
4.3.2	RasSDE 四层分布式结构设计	(112)
4.3.3	开放多数据源空间数据访问组件设计与实现	(114)
4.3.4	RasSDE 数据通信设计与实现	(116)
§ 4.4	应用实例与性能测试分析	(119)
4.4.1	测试环境	(119)
4.4.2	中国海监飞机扫描仪图像处理及三维仿真平台	(119)
4.4.3	性能测试实验与结果分析	(122)
	本章小结	(126)

第五章 融合颜色与纹理特征的遥感图像检索	(127)
§ 5.1 综合特征检索的必要性	(127)
§ 5.2 遥感图像特征提取及相似性测度	(129)
5.2.1 多通道 Gabor 纹理特征提取	(129)
5.2.2 颜色特征提取	(132)
5.2.3 特征相似性度量	(132)
§ 5.3 Quin-Tree 遥感图像分解方法	(133)
5.3.1 图像分解的目的和意义	(133)
5.3.2 图像分解方法分析	(134)
5.3.3 Quin-Tree 分解法	(137)
§ 5.4 融合颜色与纹理特征的遥感图像检索	(140)
5.4.1 检索流程描述	(140)
5.4.2 优化检索方法	(141)
本章小结	(148)
第六章 基于 GIS 语义的遥感图像检索	(149)
§ 6.1 图像语义检索存在的主要问题	(149)
6.1.1 图像语义检索研究内容	(149)
6.1.2 存在的问题和研究方向	(153)
§ 6.2 基于 GIS 语义的遥感图像检索方法	(154)
6.2.1 GISSBIR 基本思想	(154)
6.2.2 GISSBIR 中空间对象的语义表达	(156)
6.2.3 GISSBIR 预处理	(168)
§ 6.3 GISSBIR 关键技术	(170)
6.3.1 GIS 数据在 Oracle Spatial 中的存储与管理	(170)
6.3.2 Oracle Spatial 空间关系的扩展	(177)
6.3.3 GIS 语义调解器	(179)
本章小结	(180)
第七章 图像内容检索原型系统的设计与实现	(181)
§ 7.1 系统设计	(181)
7.1.1 系统总体架构设计	(181)
7.1.2 RSIRE 设计	(183)
§ 7.2 系统的实现	(186)
7.2.1 CTFFBIR 的实现	(187)

7.2.2 GISSBIR 的实现	(193)
本章小结.....	(196)
第八章 总结与展望	(197)
§ 8.1 总结	(197)
§ 8.2 展望	(199)
参考文献	(201)

第一章 绪 论

21 世纪遥感技术、计算机技术、网络技术的快速发展,使得各领域研究者获取所需要的高精度、高分辨率、多时相遥感图像成为可能,但与之对应的却是遥感图像检索理论和技术的严重滞后。因此,如何从海量遥感图像库中快速准确地检索到需要的信息具有十分重要的意义。本章从日益增长的遥感图像数据需要高效、有效的管理和检索方法出发,论述基于内容的图像检索(Content-Based Image Retrieval,CBIR)的研究内容和技术发展现状,分析遥感图像库内容检索的研究现状以及存在的主要问题,介绍本书的主要研究内容和章节安排。

§ 1.1 研究背景

自 1957 年前苏联发射了人类第一颗人造地球卫星以来,遥感技术经历 40 多年的发展,现已广泛应用于环境监测、区域规划、气象预报、水环境治理与规划、林业等资源清查与监测、通讯网络规划、数字地球等领域。随着新一代高分辨率遥感卫星的发射成功,遥感图像数据量正以每天数 TB 的速度在增长(据美国航空航天局的数字行星项目每天要产生 1000GB 新数据)。这些从不同遥感平台获得的不同空间分辨率、光谱分辨率和时间分辨率的遥感图像数据,给遥感用户提供了从粗到精、从多光谱到细分光谱段的海量对地观测数据源。

各领域研究者能否从海量遥感图像库中找到有效信息,将成为遥感图像是得到有效利用,还是被束之高阁的关键。换句话说,用户的困扰在于有效信息在哪里,而不是信息在哪里。解决这一问题的关键是发展有效的遥感图像管理和检索方法,而其中图像检索技术是近年来海量图像处理所面临的瓶颈,已成为图像数据库研究的核心技术之一。

图像检索 IR(Image Retrieval,或称图像查询)。本书中如无特殊说明,图像检索等价于图像查询),简单地说,就是从视觉数据库中提取与一个查询内容相关的图像或图像序列。传统的图像检索方法产生于 20 世纪 70 年代末,它借鉴文本检索的技术经验,通过对图像进行手工注解,利用文本检索技术进行关键字的检索,因此又称为基于文本的图像检索。传统的文本信息检索已经相当成熟,如目前图像和视频的著名搜索引擎 Yahoo, AltaVista 和 Google 等,均采用这种方式进行查询。但是,“百闻不如一见”,人类通过视觉获得的信息量约占人类从外界获得总信息量的 75%,文本无法替代视网膜上映象为人类提供的便捷直观的交流方式;况且,传统的文本检索虽然取得了成功的应

用,但也存在手式注解工作量太大和手工注解的主观性与不精确性等不可回避的问题。为了克服文本检索存在的问题,基于内容的图像检索 CBIR 应运而生。

CBIR 产生于 20 世纪 90 年代,它借助对图像从低层到高层进行处理、分析和理解的过程,来获取图像的客观视觉特征,如颜色、纹理、对象的形状等;并根据这些特征内容进行检索。CBIR 涉及到对图像进行有效的查询、索引、浏览、搜索和提取等处理,其主要特点是利用图像本身包含的客观视觉特征,通过这些特征的相似性进行检索,它不需要人来解释,也就不需要或者仅需要少量的人工干预。CBIR 现已成为视觉信息研究中的新热点,在医学图像、多媒体等领域已取得了大量研究成果,在一些原型和商业系统,如 IBM 的 QBIC(Niblack et al,1993)、MIT 的 Photobook(Pentland,1993)、UC Berkeley 的 Chabot(Ogle et al,1995)、哥伦比亚大学的 VisualSeek(Smith et al,1996)、Illinois 大学的 MARS 系统(Ortega et al,1997)和 Virage 图像搜索引擎(Bach et al,1996)等中得到了成功应用,SPIE 和 IEEE 每年都有会议或专刊探讨基于内容检索技术研究的最新进展。

虽然 CBIR 研究已经取得了很大的进展,但在特征选取、相似性比较、数据组织与存储、用户交互等方面还存在许多问题,还需要进一步研究。遥感图像作为图像的一种,其检索可充分借鉴医学和多媒体等一般图像的 CBIR 研究成果。然而,遥感图像的尺度大(一幅图像可以跨越上百平方千米地表范围)、主题不明确(一幅图像一般有多于一个的主题)、多时相(同一区域可有多个观测时间的图像)、语义丰富(表现在两方面:一方面,遥感图像的大尺度特点决定了图像的主题不是单一的,也即语义信息相当丰富;另一方面,不同领域专家因考虑问题的角度不同,对同一图像的理解可以截然不同)等特点,决定了遥感图像内容检索 RSICR(Remote Sensing Image Content Retrieval),从特征选取、相似性比较、查询机制设计,到系统结构设计等诸多方面都不同于一般图像的内容检索,因而 CBIR 的研究成果不能直接应用于 RSICR,RSICR 研究面临着许多困难与不足。

本书旨在通过对遥感图像库内容检索若干理论和技术问题的研究,为推动该领域的发展尽绵薄之力。

§ 1.2 图像内容检索的发展与研究现状

1.2.1 图像检索技术的发展

1.2.1.1 基于文本的图像检索阶段

信息检索 IR(Information Retrieval)一词早在 1951 年就已出现(Moores,1951),它用来描述用户通过提交一个查询信息的请求,然后借助系统的某种转换或计算(模糊),得到与请求相似或相关信息的过程。早期的信息检索主要是文本检索,随着互联网

的发展,允许人们发布和获取除文本以外的包括图、声、触、嗅等信息,但信息检索仍然保留了其本意,并对“图”发展了图像检索、对“声”发展了声音检索(Audio Retrieval)等研究方向。

图像检索是信息技术的一个重要研究内容,实际上是对传统信息检索的扩展,其目的是从图像数据库中快速地提取出与一个查询相关的图像或图像序列,最早的图像查询概念在 20 多年前由 Chang 提出(Chang et al,1980)。

传统的图像检索方法是基于文本特征的检索。图像的文本特征包括:文件名、标题、关键词等,它们被用来注解和检索图像数据库。从 20 世纪 70 年代末期开始,基于文本的图像检索技术 TBIR(Text-Based Image Retrieval)(Chang et al,1980,1987)随着数据库技术的进步而在数据建模、多维数据索引、查询优化和查询评估等方面取得了一定的成果(Tamura et al,1978,1984)。在 Internet(因特网)普及之前,TBIR 检索方法是有效的,也得到了不少应用。

然而,随着 Internet 的出现和网络技术的发展,这种检索方法的不足和局限性日益凸出,主要存在如下几个方面的问题:

(1)人工注解工作量太大

在这种检索方法中,人必须介入图像的注解过程,用一组预先选择好的关键词对越来越多的图像一一进行手工描述和标记。巨大的工作量和复杂的标注过程使得这一方法陷入了困境。在大多数图像中都有若干个物体,每个物体都有自己的属性集合可用于检索时的索引,物体之间又有相互关联,因此注解并不是个简单操作,需要花费脑力进行分析,对如此海量数据的分析并不是人脑所能承受的。

(2)人工注解的主观性和不确定性

用于注解图像内容的关键词难以应付冷僻词等非常规的检索:目前这些关键词是由观察者通过对图像的理解添加上去的,受观察者知识背景、感受等因素影响,不同的观察者或同一观察者在不同条件下对同一幅图像的描述会有不同,因而手工注解不够客观,难以完备描述图像内容。

(3)系统无法正确解译用户自然语言

受人工智能等技术研究水平的限制,目前自然语言理解技术的研究还不成熟,系统无法正确解译用户的请求,因而也限制了基于文本描述的检索的应用。

(4)关键词无法完全概括图像内容

图像获取的日期、关键词等可提供有用的检索线索,但不能完全概括图像的内容,而且图像中需要描述的对象往往不止一个,对象彼此之间的空间布局关系复杂,当图像库容增大时,关键词集合变得庞大、复杂,影响到系统的检索速度和检索质量。

由于上述问题的存在,TBIR 在很多情况下不能满足实际需要。例如:公安部门如何通过罪犯留在作案现场的指纹从指纹库中查找与之相同或最相似指纹,以便找到犯

罪嫌疑人;医生如何从大容量的图片库中找出具有某种器官发生病变的抽取图片;地质学者如何从大量遥感图像中找到某一种地质类型的所有图像等等。这些问题都难以用 TBIR 解决,但如果在提取出图像内容基础上,采用图像内容进行检索(如在指纹查找时用内容信息进行匹配),就可以提高检索结果的精度。

由此可见,图像检索需要根据图像所表达的内容来完成,即采用基于内容的检索方法。

1.2.1.2 基于内容的图像检索阶段

CBIR 一词于 1992 年开始在国际上使用(Kato,1992),至今还没有严格的、明确的定义。本书采用章毓晋的定义:基于内容的图像检索是借助对图像从低层到高层进行处理、分析和理解的过程获取其内容并根据内容进行检索(章毓晋,2003)。它涉及到对图像进行有效的查询、索引、浏览、搜索和提取等。

“基于内容”说明 CBIR 技术已经不仅仅局限于以往的、基于对图像语义注释(Annotation)的数据库查询,而是充分考虑图像的“内容”。其中“内容”包括基于点级的低层视觉特征(Low-Level Feature),如颜色(Color)、纹理(Texture)等,基于线、面级的中级视觉特征(Middle-Level Feature),如形状(Shape)、空间位置(Layout)等,以及基于对全图理解的高级特征(High-Level Feature),如对象语义(Semantics)等。

CBIR 经历十多年的发展,聚集了大批研究力量,并且随着人工智能、数字信号处理、统计学、自然语言理解、数据库技术、心理学、计算机视觉、模式识别和 Internet 技术水平的不断发展而取得了长足的进步。这可以从以下几方面得到证明:

一是 QBIC, Photobook, Visualeek 和 MARS 等诸多图像检索系统的问世;

二是 IEEE Computer Magazine, Journal of Multimedia Tools and Applications, Image and Vision Computing, Pattern Recognition, International Journal of Image and Graphics 等许多国际刊物组织了有关基于内容视觉信息检索的专集;

三是有关网站和网页的不断涌现(Maryland, California, Boston)。

综观 CBIR 的发展可以看出,1999 年以前, CBIR 领域的研究力量主要集中在基于低层视觉特征的检索上。但是,实际应用中,用户主要根据图像的语义,而不是颜色、纹理、形状等特征,直观地进行分类并判别图像满足自己的需要程度,即用户希望检索如蓝天白云的图像,而不是检索如 Wavelet 系数如何如何、HSV 直方图如何如何等的图像。

图像的高层语义包含了人对图像内容的理解,需要根据人的知识来判断得到。但计算机与人的本质区别在于:人在观察图像时能根据自己的经验及知识推理图像的语义信息,而目前的计算机视觉技术还不能很好地解决这个问题,这也决定了图像低层视觉特征与高层语义之间存在着难以逾越的鸿沟。

例如,图 1-1 所示的两幅图像,人的视觉感知分别为水和石头,即两者语义截然不同,但它们的颜色、纹理向量距离却很接近。

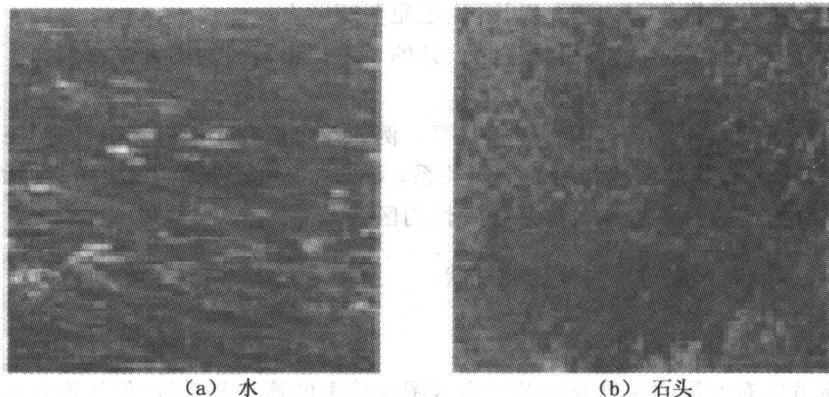


图 1-1 相同的特征向量、不同的语义

自研究者们逐渐认识到基于低层视觉特征难以满足用户的检索需求后,填补低层特征与高层语义间的鸿沟已成为 CBIR 领域的新热点,有很多研究将交互学习的方法引入图像检索的过程中,试图通过人的介入和帮助,来弥补这方面的差距。例如,QBIC 系统中采用了交互式图像分割方法(Daneels, 1993);MIT 将研究重点从“自动”的 PhotoBook 系统转换到“交互”的 FourEyes 系统(Minka, 1997);MARS 系统中形式化地提出了相关反馈技术 RF (Relevance Feedback)(Rui et al, 1997, 1998, 1999)。自此,相关反馈技术集聚了大量的研究力量。

基于相关反馈的图像检索方法的基本思路是:用户针对系统返回的第一轮查询的结果进行评价和标记,并将这些相关与否的信息反馈给系统,系统在进行下一轮的检索时,通过这些反馈的信息进行学习,从而提高检索质量。

CBIR 不断与相关反馈、机器语言领域的成果结合,开创了 CBIR 的一个崭新的发展点。

1.2.1.3 CBIR 的应用领域

经历十多年的发展,CBIR 已在许多领域得到了应用,主要有:

①电子图书馆、艺术博物馆、远程教育等。用户可在艺术博物馆中检索出与给定查询图像具有某些相似性的所有油画,以发现不同画家间的共同特性以及他们油画中的色彩搭配和主题表现方面的相互影响。

②医学分析、研究和远程医疗系统。例如医生需要通过检查医学数据库中某种病例的相似性或对一个新病例的确证。

③专利检索、商标注册管理。例如商标注册人员需要检查申请的商标是否与商标库中已经登记注册的商标具有相同或相似之处(Jain et al, 1998)。

④人口户籍管理、档案查询,以及公安、安全监视系统。例如,公安人员在办案时,将

现场留下的指纹与指纹库中进行对比,确定犯罪嫌疑人。

⑤服装设计:服装设计师可根据其设计的要求提出需要的图案进行选择。

⑥建筑设计、内部装潢等。

⑦遥感等地理图像信息的管理和共享。例如在土地利用图像图中确定某种土地利用的位置,或它与其他土地利用的空间关系;又如在卫星图像库中,战场指挥人员可从中确定某种特定目标或提取具体某种特性的区域。

1.2.2 研究内容与国内外研究现状

1.2.2.1 研究内容

目前虽然有大量的文献介绍基于内容图像检索的算法和系统,但还没有统一的、完整的、准确的基于内容的定义,对CBIR所研究的内容也没有明确的界定。查阅大量文献,基于内容图像检索的研究内容主要包括以下几个方面:

(1)图像数据模型

图像所承载的现实世界信息极为复杂,为了能用一种高效的数据组织方式对其进行表达、存储,使之尽可能地包含图像所承载的信息,方便用户快速地查询到所需信息,就必须合理地界定其内涵,并在此基础上设计通用(或称可重用)的图像数据模型UIDM(Universal Image Data Model)。

(2)图像内容特征描述

特征描述是内容检索的核心问题,其描述质量对检索效果有直接影响,主要涉及的问题包括:

①特征的选择和提取。特征的选择不仅仅局限于目前使用最多的颜色、纹理和形状等低层视觉特征,还有对象空间关系以及对象语义等中高层特征;而且针对一种类型的特征,如纹理,可采用不同的提取方法,如基于统计方法、几何方法、基于模型的方法和信息处理方法等;

②特征的组织 and 表达;

③特征描述间的相似性度量。

(3)图像数据库技术

图像数据库技术,是实现图像快速检索的关键技术之一。大多数情况下,图像的检索任务大多是面向子图的检索,因此图像内容的组织与存储主要涉及以下几个问题:

①图像数据库系统模型;

②图像分解(或称图像分块。在不作说明时,本书中两者是同意的)方法。基于颜色、纹理、形状等全局统计特征,难以实现面向子图、区域或感兴趣对象的检索。若以区域作为描述基元,则可以克服上述缺点。基于区域的内容检索已成为当前研究的热点(樊昀等,2002),它主要有两个研究方向:基于图像分割和基于图像分解;

③数据库索引。数据库索引是使检索系统适应海量数据应用需求的必要手段,主要目的是通过建立索引减少系统的响应时间。

(4)内容检索系统的设计

研究内容主要涉及:

①友好的人机交互界面。要求具有多种模式的查询提交方式,要求检索结果的表达是直观的、明了的;

②接口技术。即人与系统的交互方式,包括查询者以何种方式提出查询的内容,检索系统如何解释接受,并给出查找结果,查询者针对检索结果如何给出正确的反馈,进行再次查询,并优化结果等。

(5)检索性能的评价

主要研究建立公认的测试图像数据库以及选择行之有效的评价标准。

1.2.2.2 国内外研究现状

下面针对以上五项研究内容,评述目前国内外研究现状。

(1)图像数据模型

图像数据模型方面的研究已有一些成果,如 Lawrence 等提出了三层面向对象图像数据模型(Lawrence,1997);Kitamoto 等在研究台风云图检索时,提出了五层图像模型——层次属性关系图(Kitamoto et al,1995)等。但因目前对图像内容还没有明确的界定,因此这些模型缺乏通用性。建立通用图像数据模型是检索系统之间信息共享(或称互操作,在不作说明时,本书中两者是同意的)的前提。

(2)图像内容特征描述

目前有关特征描述的研究侧重于如何选择更符合人的感知特性的特征,例如:颜色特征除 RGB 外,还可以选择 HSV,Lab,Munshell 等,并采用直方图(Swain et al,1991)、累加直方图(Striker,1996)、颜色布局(Carson et al,1997)等方式进行相似性度量;比较常见的纹理表示方法有统计方法,如灰度共生矩阵(Haralick,1979)、几何方法,如 V 氏图剖分法(Tuceryan et al,1990)、基于模型的方法,如分形 Fractals 的方法(Pentland,1984)和信息处理方法如小波 Wavelet 域的方法(Usnsen,1995)等 4 大类;常见的形状描述方法大致分为两类:基于边缘(边缘曲线(Scassellati,1994)、边缘方向直方图(Jain et al,1996))和基于区域的形状(如不变矩法(Mehrtre et al,1997)、小波重要系数法(Jacobs,1995)等)方法等。这些研究侧重于某一类特征的描述与表达,而较少研究多类特征的组织形式。

事实上,单类特征只能表达图像的部分内容,在图像有较大变化时无法取得有效的检索结果。为此,人们研究了综合利用颜色、纹理、形状和空间关系等特征的综合检索方法。例如,Huang 等提出结合颜色和空间关系特征的颜色布局检索方法(Huang et al,1997a),Virage、QBIC 和 Photobook 等系统均提供了综合特征检索方法。与单类特征

的检索相比,综合特征检索可以利用不同特征的优势互补,达到提高检索的灵活性和检索质量的效果。

单类特征也好,综合特征也罢,与人对图像的认知即图像的高层语义之间都存在“语义鸿沟”(Semantic Gap)。图像检索的实际应用中,人们往往是通过高级语义来检索自己所需要的图像,即采用基于语义的图像检索方法来查询信息。图像语义检索主要涉及图像语义描述和图像语义提取两方面问题,其中图像语义描述方法主要有文本描述和人工智能等,具体地有:语义联想的词典(WordNet)技术(Voorhees,1993)、语义网络模型,以及知识表达模型(Meghini et al,1997;Cavazza et al,1998;Colombo et al,1999)等;图像语义提取方法主要有基于知识的语义提取、人工交互语义提取、基于外部信息的语义提取等方法(王惠锋等,2002)。

这些方法虽然在不同的场合分别显示了它们在语义表达或语义匹配方面的能力,但是,目前缺乏一种对不同的图像库检索都能收到较好效果的方法。这其中很重要的一个原因是:目前图像理解技术发展水平和对认知理解水平均有限,难以实现图像低层特征和高层语义之间的自动映射。

(3) 图像数据库技术

目前图像数据库模型主要有传统的关系数据库、对象数据库、对象关系数据库三种,它们各有特点。在设计图像数据库时,比较现实的选择仍是以关系数据库,并扩展非结构化的图像二进制大对象的存储方式。

图像分解是图像数据库的关键技术之一,它主要通过两种方法实现:即基于图像分割和基于图像分解。

图像分解是图像数据库的关键技术之一。其思路很早就体现在CBIR中(Smeulders et al,2000),例如,Stricker提出了一种静态5区域分块法(Stricker et al,1996),Lu提出了四叉树图像分解法(Lu et al,1994),Sheikholeslami提出了一种九叉树图像分解法(Sheikholeslami et al,1999)等。基于子块图像检索的准确度与其计算复杂度之间平衡关系,取决于所采用的子块方法以及子图像的尺寸大小两个因素。

索引是数据库,也是图像内容检索的核心技术之一。图像检索中主要采用的索引方法有:网格法(Samet,1989)、K-d树(Bentley,1975)、R-树及其改进(Sellis et al,1987)等传统数据库的高维索引方法,以及三角不等式索引等。随着基于语义图像检索的不断深入,基于语义分类的索引(Gorkani et al,1994;Szummer et al,1998;Yiu,1996)正成为研究热点。

(4) 内容检索系统的设计

许多有特色的内容检索系统,如QBIC,VisualSeek,MARS,VIRAGE,PhotoBook等等,具有很多值得借鉴的地方,但是统观这些系统就会发现,如何有效的实现多维特征索引和相关反馈,仍是系统设计中值得研究的重要问题。