

图象图形科学丛书

章毓晋 著

基于内容的

视觉信息检索



科学出版社
www.sciencep.com

图象图形科学丛书

基于内容的视觉信息检索

章毓晋 著

科学出版社

北京

内 容 简 介

本书结合对基于内容视觉信息检索进行的研究和工作,全面系统地介绍其主要概念、基本原理、典型方法、实用技术以及国际上有关研究的新成果和新动向。本书分为4部分。第1部分概括介绍基于内容视觉信息检索的主要概念和基本原理,以及典型系统的特点;第2部分介绍基于特征的图象检索方法和系统,讨论了采用不同图象特征的技术;第3部分介绍基于内容的视频检索方法,涉及到表达、分析、组织等不同技术;第4部分介绍一些当前研究动态和方向,并讨论最新发展趋势。

考虑到基于内容的视觉信息检索涉及的学科多,范围广,本书选取了一些比较有特色的技术方法进行介绍,并结合科研成果给出形象的实例,以使本书既能较好地反映该领域的全貌,也有一定的层次,方便读者学习和使用。

本书可供信息工程、电子工程、计算机科学与技术、数据库管理、媒体制作和生产、远程教育 and 医疗、公安、遥感,以及军事侦察等领域的科技工作者参考,也可作为信号与信息处理、通信与电子系统、模式识别、计算机视觉、生物医学工程等学科专业课的教材或教学参考书。

图书在版编目(CIP)数据

基于内容的视觉信息检索/章毓晋著. —北京:科学出版社,2003

(图象图形科学丛书)

ISBN 7-03-011311-X

I. 基… II. 章… III. 图象数据库-情报检索 IV. G354.49

中国版本图书馆 CIP 数据核字(2003)第 022749 号

责任编辑:王淑兰/责任校对:包志虹

责任印制:吕春珉/封面设计:王浩

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

双青印刷厂 印刷

科学出版社发行 各地新华书店经销

*

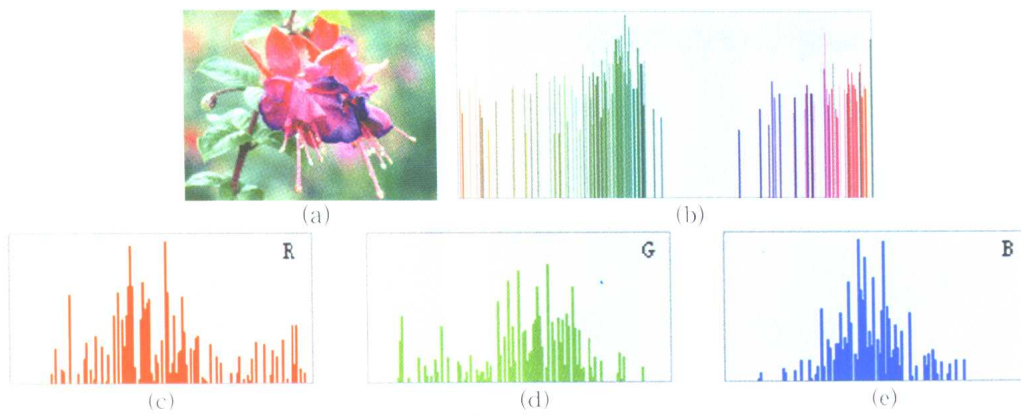
2003年5月第一版 开本:787×1092 1/16

2003年5月第一次印刷 印张:32 1/2 插页:1

印数:1—4 000 字数:770 000

定价:49.00元

(如有印装质量问题,我社负责调换〈路通〉)



彩图1 (图4.2.2)



彩图2 (图4.5.5)



彩图3 (图8.3.2)



彩图4 (图8.3.3)



彩图5 (图8.3.4)



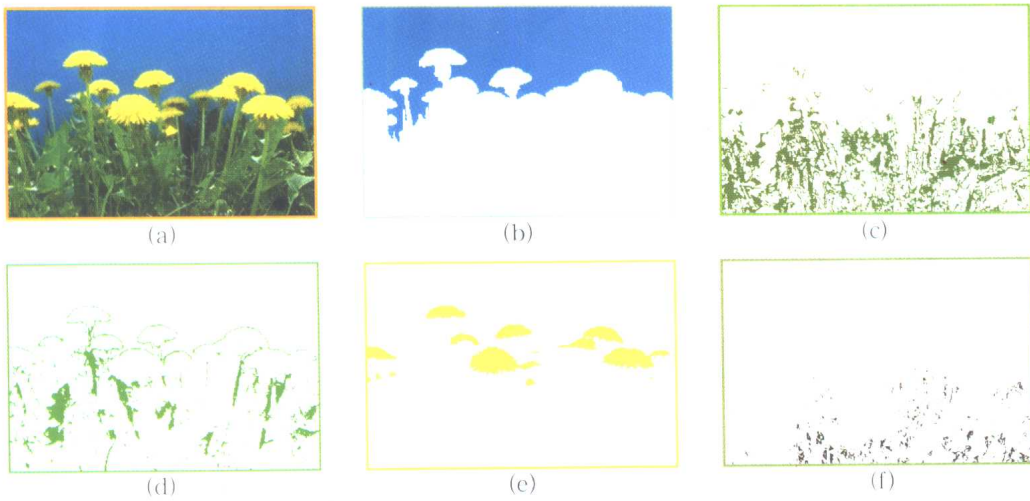
彩图6 (图9.2.7)



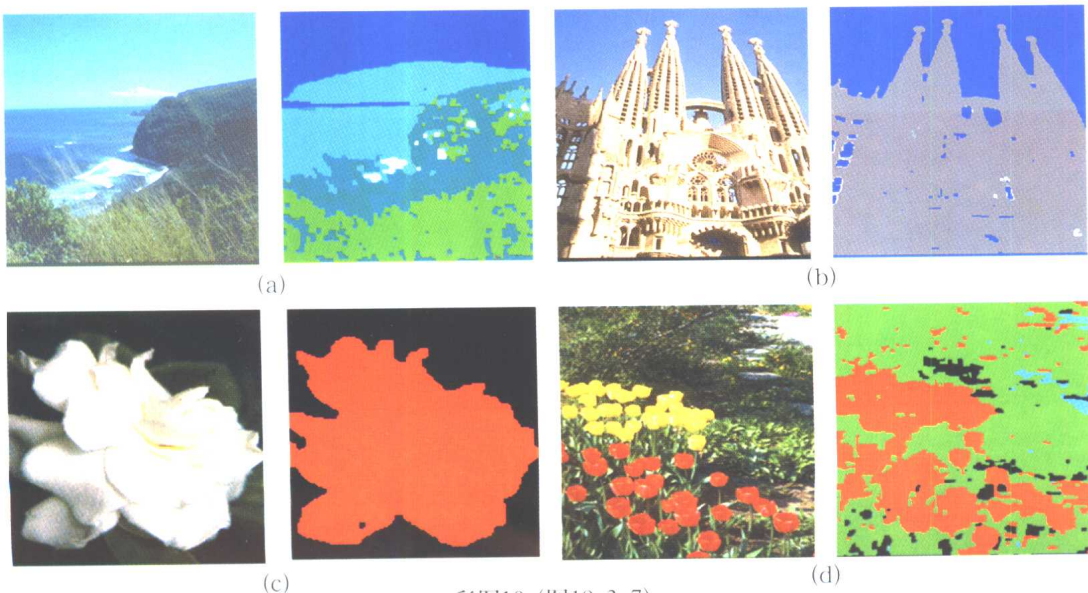
彩图7 (图9.2.8)



彩图8 (图9.4.3)



彩图9 (图18.5.1)



彩图10 (图19.3.7)

图象图形科学丛书编委会名单

主任委员

潘云鹤

副主任委员

(按姓氏笔划为序)

王宝兴 王淑兰 罗志安 章毓晋

委 员

(按姓氏笔划为序)

王宝兴	王淑兰	刘健勤	朱述龙	江 早
石教英	何江华	鲍虎军	罗志安	张永生
章毓晋	崔 屹	潘云鹤	潘志庚	

丛书序言

图象图形是人类相互交流和认识客观世界的主要媒体。科学研究和统计表明，视觉系统帮助人类从外界获得 3/4 以上的信息，而图象图形带给我们的正是视觉世界中的所有信息。视觉信息所获得的客观作用是其他信息不能替代的，百闻不如一见就是一个非常形象的例子。图象图形是现代信息化社会的重要支柱。

图象图形科学是一门理论与现代高科技相结合来系统地研究各种视觉原理、技术和应用的综合性很强交叉学科。图象图形技术在广义上是各种与视觉有关技术的总称。人类基于视觉的活动，是一个广阔、复杂、富有挑战性的研究领域。图象图形科学和技术是这个领域的有力工具。该学科包括利用计算机和其他电子设备观察世界而获得的数据及按产生数据处理并且显示这些数据的理论和技术的研究。

图象图形科学具有涉及面广，内容丰富，跨行业、跨学科的特点。从它的研究方法来看，它与数学、物理学、生理学、心理学、电子学、计算机科学等许多学科可以相互借鉴；从它的研究范围来看，它与人工智能、神经网络、遗传算法、模糊逻辑等理论和技术都有密切的联系；它的发展应用与医学、遥感、通信、影视、文档处理和工业自动化等许多领域也是不可分割的。在科学史上，它代表了最活跃和令人振奋的边缘学科之一。

图象图形技术已经迅速渗透到人类生活和社会发展的各个方面。图象图形技术无论在科学研究、工业生产、文化娱乐、管理部门都得到越来越多的重视。图象图形技术在工业检测、高空侦察、制导、文化处理、动画、虚拟现实、生物医学、人体科学、艺术、远程教育、科学可视化、计算机辅助设计、遥感、航天等方面都得到越来越多的应用。进入 21 世纪后，其发展将更加迅速。

“图象图形科学丛书”正是在这种形势下组织出版的。中国图象图形学会和科学出版社为该套书的出版付出了很多的努力。这套丛书比较全面地覆盖了图象图形科学的各个分支，是广泛了解图象图形领域基本理论、技术应用和发展动态的最好读物；也是从事图象图形领域研究、技术开发和实际应用人员的工具书。

“图象图形科学丛书”由我国该领域的专家编写，这些专家既对图象图形领域的发展有全面的把握，又分别在其中的某一方向上有深入的研究和独到的见解，充分反映了当前图象图形科技研究的前沿、进展和水平。希望该套书能为发展图象图形科学技术，活跃学术气氛，交流研究成果，促进科技发展，为迎接信息技术的挑战，为我国图象图形事业做出应有贡献。

潘雪鹤

作者简历



章毓晋，清华大学电子工程系教授，博士生导师，图象图形研究所副所长。1989年获比利时列日大学应用科学博士学位，从1989年至1993年在荷兰德尔夫特大学作博士后及研究工作。1993年到清华大学工作，1997年被聘为教授，1998年成为博士生导师。

在清华大学，先后开出和讲授课程：数字图象处理，图象处理和分析，图象理解与计算机视觉，图象工程专题（英语），基于内容的视觉信息检索；组织开出和讲授课程：图象新技术，参与讲授课程：电子学学科前沿。编写了大学生和研究生教材：《图象工程（上册）——图象处理和分析》，《图象工程（下册）——图象理解与计算机视觉》和《图象工程（附册）——教学参考及习题解答》（清华大学出版社）；编写了远程教育研究生教材：《图象处理和分析基础》（高等教育出版社）。积极从事教学研究，已在国内外发表约20篇教学研究论文。承担国家“九五”重点科技攻关项目，研制出版了“图象处理和分析多媒体计算机辅助教学课件”（高等教育出版社）。

主要研究领域为其积极倡导的图象工程（图象处理、图象分析、图象理解及其技术应用）学科。近年对图象处理、分析和理解的综合进行了专门研究，并将它们结合，首次完整地提出了图象工程学科的概念定义和内含外延并积极倡导。结合图象工程的提出，从1996年起已连续8年对国内图象工程的研究及主要文献进行了系统的年度分类总结综述，以推动我国图象工程的研究和应用工作。近年来作为负责人承担了国家自然科学基金、国家高技术计划（863）及国家教委博士点基金等研究项目，已在国内外发表约200篇图象工程研究论文，出版专著《图象分割》和《基于内容的视觉信息检索》（科学出版社）。

现为国际电气电子工程师协会（IEEE）高级会员（senior member），国际刊物“Pattern Recognition Letters”和“International Journal of Image and Graphics”编委（associate editor）；中国图象图形学会副理事长，学术委员会副主任；《中国图象图形学报》副主编，《电子与信息学报》和《计算机辅助设计与图形学学报》编委；曾任第一届和第二届国际图象图形学术会议（ICIG' 2000, ICIG' 2002）程序委员会主席。

前 言

基于内容的视觉信息（主要包括图象和视频信息）检索自 10 年前正式“冠名”以来，得到国内外信息领域科技人员的广泛重视和研究，迅速成为多媒体，特别是图象、视频和数据库技术中的研究热点之一。本书试图回顾相关的发展过程，介绍基本的原理和关键技术，总结已有的丰富成果，探索深入研究的方向，以推动基于内容视觉信息检索的研究和应用工作。

基于内容的视觉信息检索（主要包括基于内容的图象检索和基于内容的视频检索）是当前信息领域的重要研究热点。它是随着科学技术的进步、发展和推广应用，在数据快速增长、信息急剧膨胀情况下为满足人们快速提取有用视觉信息的需求而逐步被重视起来的。

基于内容的视觉信息检索是一个挑战性的研究，涉及许多学科和领域。从研究内容来看，它与数学、物理学、生理学、心理学等许多学科密切相关，从技术基础来看，主要包括图象技术、计算机技术、数据库技术等信息技术，而从发展应用来看，它与工业、通信、娱乐、医学、遥感等是密切相关的。

基于内容的视觉信息检索从正式提出至今应该说已有 10 个年头，简单回顾一下作者与本书写作内容有关的几个时间标记：

1992 年，基于内容的图象检索（CBIR）一词开始在国际上正式使用，同时基于内容的视频检索（CBVR）研究也已开始。当时人们将多媒体仅与存储和快速传输各种媒体信息的技术联系在一起时，基于内容的视觉信息检索从信息加工的角度指出，对多媒体信息进行自动查询和搜索也是非常重要的。

1994 年，在国内开始关注这方面的研究，并先后在一些刊物和杂志上概括通俗地介绍了有关的原理、方法和研究应用的特点，以唤起国内科技人员的注意和重视。同时，也结合在相关领域已有的工作基础，开始了基于内容视觉信息检索的基础研究，这个研究方向的选择现在看来是很值得的。

1996 年，在这方面的研究工作得到了国家自然科学基金项目（69672029）的支持，其后几年又陆续得到国家高技术发展计划（863）基金项目（863-317-9604-05）、清华大学创新基金项目（EE9906）、国家自然科学基金项目（60172025）等的支持。近年来，在这些基金项目的支持下，先后发表了相关研究论文约 80 篇，为本书的写作打下了坚实的基础。

1998 年，应邀在中国图象图形学会主办的第 9 届全国图象图形学术大会上做了有关基于内容视觉信息检索的特邀报告（其后还在多个大学和会议上介绍过相关内容），受到与会人员的欢迎。与同行的交流后，萌生了写一本有关基于内容视觉信息检索书籍

的念头。

2000年,在每年一次(从1996年开始)撰写的中国图象工程系列综述中增列了基于内容的图象和视频检索文献类(由于当年开始有相当数量的有关文献发表)。后来,该类的文献统计量逐年上升,已成为中国图象工程近年发展的主要热点之一。同年在清华大学为研究生开设了一门新的专业课——基于内容的视觉信息检索。

2002年,在多年积累的基础上,正式开始了本书的写作。10年弹指一挥间,这期间基于内容的视觉信息检索的研究和应用已经发生了巨大的变化并取得了瞩目的进展,这既为本书的写作提供了广泛的素材,也对本书的选材是一个挑战。

本书在编写上,基本上从专著和科研的角度出发,使理论性、实用性、系统性和实时性相结合,并在本书体现出来。本书既参考了许多他人的有关文献,更结合了作者近年在该方面的研究工作,不仅有较系统全面的原理介绍,还结合科研成果给出了许多实例。本书不仅选取了比较典型和成熟的技术方法,还注意了近年来国际上的最新研究动向。本书从内容上试图尽可能地覆盖基于内容视觉信息检索的方方面面,希望读者能借助本书独立地和全面地了解该领域的基本理论、实用技术、应用情况和发展前景,并可以投身到有关研究和应用工作中来。

本书分为4部分。第1部分(包含第1,2,3章和附录A,B)概括介绍主要概念和基本原理以及典型系统的特点,并介绍国际标准MPEG-7和提供进一步参考资源;第2部分(包含第4,5,6,7,8,9章)介绍基于特征的图象检索方法和系统,讨论了采用不同图象特征(包括颜色、纹理、形状和空间关系)的技术;第3部分(包含第10,11,12,13,14,15,16章)介绍基于内容的视频检索方法和应用,涉及到数据库管理和模型、镜头检测、聚类描述、组织索引、表达查询、分析检索、浏览显示等不同技术;第4部分(包含第17,18,19,20,21,22章)介绍一些当前研究动态和方向,如压缩域检索,网络搜索、语义检索、交互反馈、性能评价,并讨论最新研究和应用发展趋势。

本书从篇幅上看,共有22章正文和2个附录,在与此对应的24个一级标题下有108个二级标题(节),再之下还有260个三级标题(小节),从内容上大体覆盖了基于内容视觉信息检索的各个重要分支和基本技术。全册书共有文字(也包括图象、绘图、表格、公式等)约70多万字。本书共有编了号的图285个(包括几百幅图片)、表格55个、公式507个。另外书中列出了直接引用的近1000篇有代表性的参考文献目录和一些相关的刊物会议名称及网站、网页地址,希望对读者进一步深入学习有关内容,了解技术细节和开展科研工作有所帮助。

本书是作者和学生的共同研究成果。多年来,先后有刘忠伟、姚玉荣、陆海斌、徐寅、杨翔英、高永英、罗芸、李娟、俞天利、李勍、黄翔宇、戴声扬、姜帆、徐丹等直接参与了有关的研究工作,本书也使用了他们学位论文的一些内容和他们收集与实验得到的图片。另外,还有许多学生也为本书以不同方式做出了贡献,特别是3年来参与“基于内容的视觉信息检索”研究生专业课的学生,他们的提问、作业、项目训练,以及与他们的讨论都为本书的选材和写作提供了许多有益的启发。

本书在整个编写过程中得到王淑兰编审的热心支持，她认真组织了对本书的审阅修改，使全书大为增色也使作者受益匪浅。

最后要指出，作者妻子何芸、女儿章荷铭和其他家人在各方面的理解和支持，是作者决心写作并最终能在今年的最后一天定稿的重要动力和主要保证！

章毓晋

2002年12月31日于蓝旗营

通信:北京清华大学电子工程系(100084)

电话:(010) 62781430

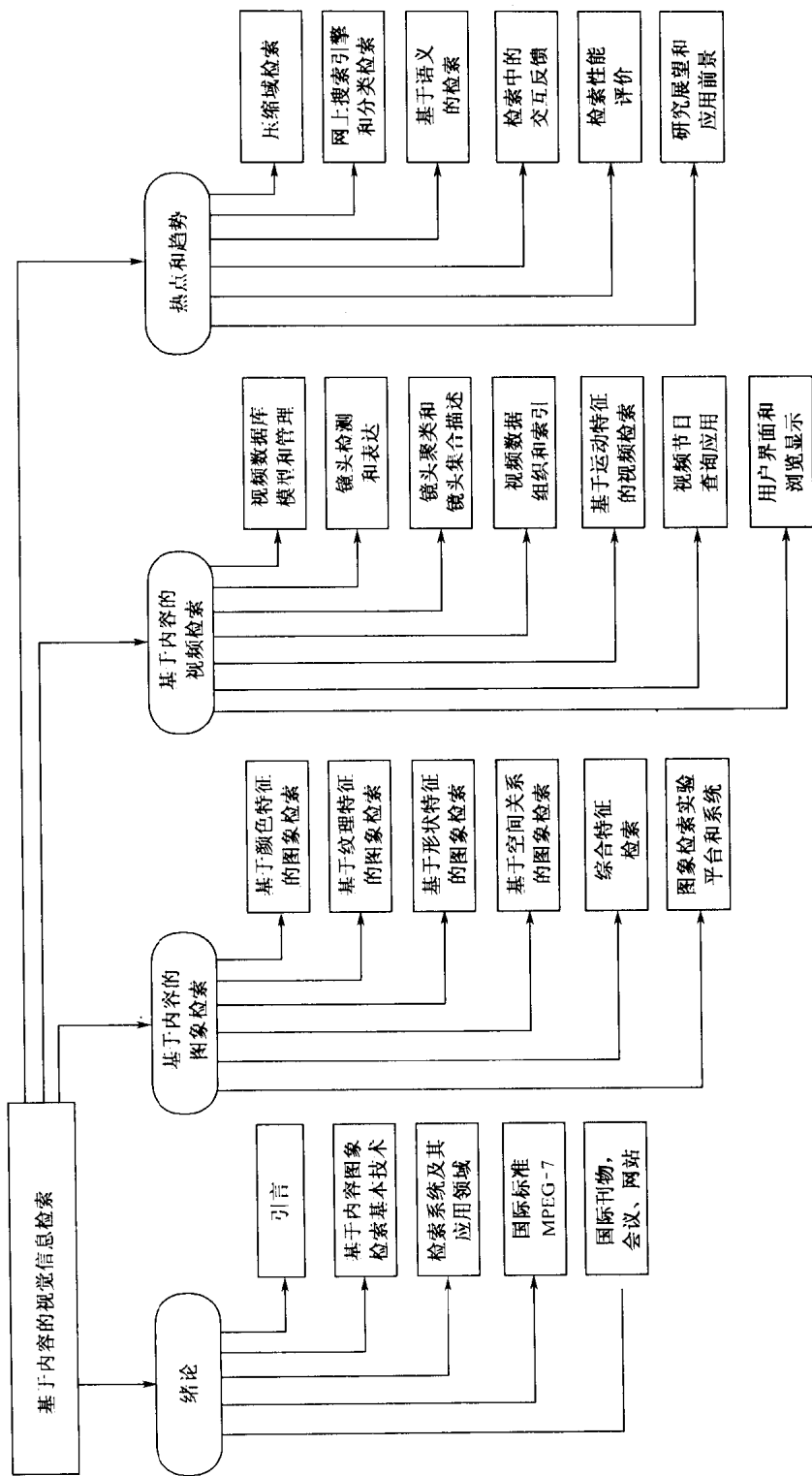
传真:(010) 62770317

电子邮件:zhangyj@ee.tsinghua.edu.cn

个人主页:www.ee.tsinghua.edu.cn/teachers/zhangyujin

研究室网:image.ee.tsinghua.edu.cn

全书框架图



目 录

第 1 篇 绪 论

第 1 章 引言	3
1.1 为什么需要基于内容的视觉信息检索	3
1.1.1 视觉信息的快速增长	3
1.1.2 视觉信息检索	4
1.1.3 传统视觉信息检索的局限	5
1.2 什么是基于内容的视觉信息检索	6
1.2.1 基于内容视觉信息检索的特点	6
1.2.2 基于内容的视觉信息检索示例	7
1.3 如何实现基于内容的视觉信息检索	9
1.3.1 归档和检索流程图	9
1.3.2 查询和检索需解决的问题	10
1.3.3 所需的基础知识	10
1.4 基于内容视觉信息检索的发展	11
1.4.1 第一代视觉信息检索系统	11
1.4.2 新一代视觉信息检索系统	12
1.4.3 目前状况和回顾	13
1.5 本书概要	14
1.5.1 整体框架安排	14
1.5.2 各章概要	15
参考文献	17
第 2 章 基于内容视觉信息检索的基本技术	21
2.1 基本工作框架和功能模块	21
2.1.1 基本框架	21
2.1.2 模块功能简介	22
2.2 图象检索技术	23
2.2.1 查询模块和技术	23
2.2.2 描述模块和技术	25
2.2.3 匹配模块和技术	27
2.2.4 提取模块和技术	30
2.2.5 验证模块和技术	31
2.3 视频检索技术	31
2.3.1 视频查询和访问	31
2.3.2 视频组织和描述	32
2.3.3 视频匹配要点	34

2.3.4	视频提取和特点	34
2.3.5	对返回视频的验证	35
2.4	检索涉及的领域和技术	36
	参考文献	37
第3章	检索系统及其应用领域	39
3.1	典型的图象检索系统	39
3.1.1	图象检索系统功能和概况	39
3.1.2	QBIC	40
3.1.3	计算机辅助人面像推断和检索系统	41
3.2	典型的视频检索系统	42
3.2.1	视频检索系统特点和概况	42
3.2.2	面向对象的视频信息数据库	43
3.2.3	视频索引和检索原型系统	44
3.2.4	代数视频系统	45
3.3	系统性能指标和评价准则	47
3.4	基于内容视觉信息检索的应用领域	49
	参考文献	50

第 2 篇 基于内容的图象检索

第4章	基于颜色特征的图象检索	57
4.1	颜色模型	57
4.1.1	面向硬设备的颜色模型	58
4.1.2	面向视觉感知的颜色模型	60
4.1.3	均匀颜色空间模型	63
4.2	颜色特征的表达	64
4.2.1	统计直方图	64
4.2.2	累积直方图	65
4.2.3	颜色布局	66
4.3	颜色匹配算法	67
4.3.1	简单直方图匹配方法	67
4.3.2	累积直方图及优越性	69
4.4	对 10 种基于颜色特征检索方法的比较	71
4.4.1	10 种检索算法及编号	71
4.4.2	实验结果	72
4.4.3	结果分析和讨论	74
4.4.4	HVC 和 HSI 彩色空间的比较	75
4.5	局部累加直方图方法	76
4.5.1	局部累加直方图	76
4.5.2	相似色区间的划分	77
4.5.3	基于局部累加直方图的 4 种检索算法	80
4.5.4	算法性能检验及结果	81

参考文献	82
第 5 章 基于纹理特征的图象检索	84
5.1 纹理描述模型	84
5.1.1 纹理概述	84
5.1.2 基于空间性质的纹理模型	85
5.1.3 基于频域性质的纹理模型	87
5.1.4 基于结构感知性质的纹理模型	88
5.2 统计法纹理描述	88
5.2.1 基本纹理描述符	88
5.2.2 分形模型	91
5.3 频谱法纹理描述	93
5.3.1 基于傅里叶变换的纹理描述	93
5.3.2 基于小波变换的纹理描述	94
5.3.3 国际标准 MPEG-7 推荐的纹理描述符	94
5.4 结构法纹理描述	95
5.5 基于纹理的检索方案	96
5.6 利用纹理特征的检索算法示例	98
参考文献	100
第 6 章 基于形状特征的图象检索	102
6.1 形状描述	102
6.1.1 形状描述概述	102
6.1.2 基于外部参数的形状描述	103
6.1.3 基于内部参数的形状描述	106
6.1.4 基于变形的形状描述	108
6.2 基于形状的检索	109
6.2.1 基于特征的方法	110
6.2.2 基于变换域的方法	113
6.2.3 基于变形的的方法	114
6.2.4 草图查询	115
6.3 小波轮廓描述符	116
6.3.1 小波轮廓描述符的定义	117
6.3.2 小波轮廓描述符的性质	118
6.3.3 相对傅里叶轮廓描述符的一些优越性	120
6.4 小波模极大值和多尺度不变矩方法	123
6.4.1 小波模极大值	123
6.4.2 算法要点	125
6.4.3 实验及结果	126
参考文献	130
第 7 章 基于空间关系的图象检索	134
7.1 空间关系概述	134
7.2 基于目标结构的表达	136

7.2.1	网格	137
7.2.2	四叉树	137
7.2.3	二叉树	138
7.2.4	<i>K-d</i> 树	139
7.2.5	<i>R</i> -树	139
7.3	基于关系的表达	140
7.3.1	符号投影和 2-D 串表达	140
7.3.2	2-D 串族	142
7.3.3	变换下不变的表达	145
7.4	度量关系	148
7.4.1	空间距离测量	148
7.4.2	空间朝向测量	149
7.5	利用空间关系查询和检索	151
7.5.1	使用文字查询和检索	151
7.5.2	利用范例的图标查询和检索	152
7.5.3	借助草图查询和检索	155
7.6	利用子区域向量匹配的空间关系检索	155
7.6.1	算法原理和步骤	155
7.6.2	算法应用示例——篮球站位分析系统	157
7.6.3	算法性能实验及结果分析	158
	参考文献	160
第 8 章	综合特征检索	162
8.1	不同特征的特点	162
8.1.1	不同特征的自身特点	162
8.1.2	不同特征的相对特点	164
8.2	综合特征	164
8.2.1	特征的综合使用	164
8.2.2	特征归一化	165
8.3	结合颜色和纹理的检索	167
8.3.1	颜色的空间布局	167
8.3.2	综合利用颜色和纹理特征进行检索	168
8.4	结合颜色和形状的检索	169
8.4.1	方法介绍	169
8.4.2	试验结果	170
8.5	结合纹理和形状的检索	171
8.5.1	边缘点数直方图	171
8.5.2	边缘直方图描述符	172
8.5.3	使用纹理和形状的组合测度	174
8.6	结合空间关系的检索	175
	参考文献	178
第 9 章	图象检索实验平台和系统	180

9.1 基于特征的图象查询和检索系统	180
9.1.1 系统模块和功能	180
9.1.2 系统查询实验结果	182
9.2 基于特征的算法试验平台	184
9.2.1 设计原则和功能要求	184
9.2.2 平台主要模块	185
9.2.3 平台有特色的检索功能	188
9.3 基于服务器和客户端的网上图象查询系统	191
9.3.1 系统设计	191
9.3.2 系统流程和分析	193
9.3.3 系统界面	195
9.4 WWW 上图象浏览检索系统	196
9.4.1 系统组成及功能	196
9.4.2 系统实现	197
9.4.3 系统用户界面特点	198
参考文献	199

第 3 篇 基于内容的视频检索

第 10 章 视频数据库模型和管理	203
10.1 视频和视频数据库特点	203
10.2 数据模型	205
10.2.1 实体-联系模型	205
10.2.2 语义对象模型	206
10.3 视频数据模型	207
10.3.1 时间线模型	207
10.3.2 时间层次模型	209
10.3.3 代数模型	210
10.3.4 视频对象数据模型	211
10.4 数据库结构和模型	212
10.4.1 数据库结构	212
10.4.2 数据库模型	212
10.5 对视频库的管理和检索	214
10.5.1 视频数据库管理	214
10.5.2 视频库检索的一种结构化框架	215
10.5.3 关于视频库检索的讨论	217
参考文献	218
第 11 章 镜头检测和表达	221
11.1 视频序列的时域分割——镜头检测	221
11.2 切变检测	223
11.2.1 基于全局特征的变化检测	224
11.2.2 基于局部特征的变化检测	227