

高新技术专著系列

Image Retrieval And Compressed Domain Processing

图像检索与 压缩域处理技术的研究

沈兰荪 张菁 李晓光 著

- 图像检索与压缩域图像处理技术的学术专著
- 相关领域的最新研究进展
- 理论和系统概念的平衡
- 先进性和实用性的完美统一



人民邮电出版社
POSTS & TELECOM PRESS

高新技术专著系列

Image Retrieval And
Compressed Domain
Processing

图像检索与
压缩域处理技术的研究

沈兰荪 张菁 李晓光 著

人民邮电出版社

北京

图书在版编目 (C I P) 数据

图像检索与压缩域处理技术的研究 / 沈兰荪, 张菁,
李晓光著. —北京: 人民邮电出版社, 2008. 12
ISBN 978-7-115-18994-3

I. 图… II. ①沈…②张…③李… III. ①图像数据库—
情报检索—研究②图像数据库—数据压缩—研究 IV.
G354.49 TP391.4

中国版本图书馆CIP数据核字 (2008) 第158379号

内 容 提 要

本书是关于图像检索与压缩域处理技术的一本学术专著, 反映了相关技术领域近年来的最新研究进展。

全书共分 4 篇 20 章。第一篇为基础知识, 介绍数字图像与编码的基础知识、图像编码的基本方法与有关国际标准, 基于内容的图像检索技术; 第二篇为基于视觉感知的个性化图像检索技术, 分析论述视觉注意机制及注意模型、视点跟踪技术、客观化的视觉感知技术、个性图像检索技术以及用户兴趣模型等内容; 第三篇为基于压缩域的图像处理技术, 阐述压缩域纹理、图像分类、基于轮廓特征检索、DCT 域处理、小波域处理以及 JPEG2000 等内容; 第四篇为全书的结束语, 主要介绍当前图像检索与压缩域处理方面的一些研究热点。

本书可供通信与电子系统、信号与信息处理、计算机应用等相关专业的研究人员、工程技术人员、高校教师、研究生和高年级本科生学习参考。

图像检索与压缩域处理技术的研究

- ◆ 著 沈兰荪 张菁 李晓光
责任编辑 姚予疆
执行编辑 杨凌
- ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子函件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京铭成印刷有限公司印刷
- ◆ 开本: 700×1000 1/16
印张: 32.5 彩插: 2
字数: 613 千字 2008 年 12 月第 1 版
印数: 1—2 500 册 2008 年 12 月北京第 1 次印刷

ISBN 978-7-115-18994-3/TN

定价: 98.00 元

读者服务热线: (010)67129264 印装质量热线: (010)67129223
反盗版热线: (010)67171154

本书的研究工作曾受到国家自然科学基金、国家“863”计划、北京市自然科学基金以及国家 211 工程的资助。

前 言

长期以来，人们在自然界感受到的信息主要是视觉信息。与此不同，传统的计算机与通信领域，能够处理和传输的主要是文字和语音信息。近年来，全球信息化和经济全球化已成为时代的发展潮流，研究新的宽带业务，开发网络多媒体应用，提高人类的生活质量，已成为各国科技界与产业界共同关注的问题。由于宽带业务与多媒体信息的主体是图像/视频，因此，研究开发以图像/视频为主体的多媒体业务已成为当今信息科学与技术的重要研究领域。

如何对急速膨胀的海量图像数据进行有效的检索处理，是人们研究开发以图像/视频为主体的多媒体业务首先要面对的一个问题。针对目前广泛使用的基于关键字检索（Keywords-Based Retrieval）的图像检索方法的种种不足，人们提出了基于内容的图像检索（CBIR, Content-Based Image Retrieval）技术。CBIR 技术通过提取图像的低层视觉特征（如颜色、纹理和形状等）为图像建立索引；用户通过提供示例图、草图或描述图像特征等进行检索；系统进行相似度匹配，返回与特征相似的结果。CBIR 技术的优点是，它在一定程度上体现了图像的内容，具有较强的客观性。

但是，人们主要是根据图像的含义去理解图像，而不是仅根据图像的低层视觉特征观察图像。图像的含义就是图像的高层语义，它包含着人对图像内容的理解。因此，对于基于内容的图像检索（CBIR）技术，一个严峻的问题就是图像低层视觉特征与高层语义之间存在着巨大的语义鸿沟（Semantic Gap），用户对图像的理解或检索意图无法用图像处理算法提取的低层视觉特征来完全表达。如何克服语义鸿沟，是当今信息科学与技术的一个研究热点。

如今，图像/视频编码的研究已从传统的面向存储转向面向传输。除了通常的压缩性能与重建质量外，人们还要求编码算法能提供关于质量、分辨率、信噪比等多种可扩展编码的结构，能实现多分辨率编码、嵌入式编码以及提供抗误码的能力，从而能在互联网与无线移动环境向用户提供个性化的服务。

然而，当前有关多媒体数据，尤其对图像数据的压缩和检索的研究是相互分离的：压缩算法的设计和实现仅仅考虑在保证一定重构质量的前提下，以尽量少的比特数来表征图像信息，并未考虑对后续的操作（比如检索）提供有效的支持；而基于内容的检索技术的研究也主要是针对原始图像数据进行的，对于普遍存在的各种压缩形式的多媒体数据，必须先解压缩才能进行相应的分析、处理。这无疑极大地影响各类多媒体应用系统的实用性和灵活性，特别是对一些大型的图像检索应用系统和要求实时性的应用场合，这种传统的处理模式往往难以满足要求。正是在这样的背景下，压缩域处理技术的研究受到了人们的重视。

本书是一本关于图像检索与压缩域处理技术的学术专著。本书内容的第一个重点是讨论基于视觉感知的个性化图像检索技术的研究。针对 CBIR 的语义鸿沟，希望通过直接感知用户视觉信息来克服语义鸿沟，实现一种能够客观反映用户语义的个性化图像检索，从而有效解决目前 CBIR 研究中存在的一些问题。

本书讨论的第二个重点是压缩域处理技术。与传统的图像压缩和检索相互分离的方式不同，该技术从相互结合、相互支持的观点出发进行图像压缩和检索算法的研究，即图像检索的实现应充分利用图像压缩算法及其所形成压缩数据的特点，直接在不解码或尽量少解码的压缩数据上进行；而图像压缩算法的实现应使其不仅具有较高的压缩效率和重建质量，还应能更好地支持图像信息的基于内容的检索操作。

本书是在北京市信号与信息处理研究室多年的研究工作的基础上编写的。编写本书时，采用或引用了研究室的一些科研成果。先后参与这些工作的有魏海博士、黄祥林博士、李晓华博士、刘党辉博士、卓力博士、王素玉博士、张菁博士、李晓光博士以及孔潇硕士、宋磊硕士、朱旭娟硕士、田巍硕士、孙慧平硕士、李嵩硕士、贺琳硕士、赵孟凯硕士、高静静硕士等，在此一并致谢。

.....

应着重指出的是魏海、黄祥林、李晓华等的工作，特别是魏海博士这位极有才气的年轻人，他在 20 世纪末在压缩域图像处理方面的一些开拓性工作，今天看来仍然很有价值。

感谢北京工业大学数理学院李云章教授审阅了本书第 17 章有关小波分析的数学内容。

人民邮电出版社对本书的出版给予了大力支持，出版社对学术著作的这种精益求精的精神实属难能可贵。

由于作者水平、时间和精力所限，本书还有不少不尽如人意的地方，尽管我们做了努力，但缺点、错误在所难免，希望得到读者的批评指正。

沈星茹

于科学楼 830 室

2008 年 7 月 27 日

目 录

第一篇 基础知识

第1章 数字图像

1.1	图像与视频	3
1.2	数字图像的特点	4
1.2.1	图像数据的海量性	4
1.2.2	图像的压缩格式具有普遍性	5
1.2.3	图像的内容表现具有非结构性	6
1.2.4	应用环境的多元化	6
1.3	图像的彩色空间	7
1.3.1	彩色空间的视觉属性	7
1.3.2	RGB 彩色空间	7
1.3.3	YCbCr 彩色空间	8
1.4	视频格式	9
1.5	感兴趣区的人眼视觉系统特性	10
1.5.1	视觉感兴趣区	10
1.5.2	视觉掩盖效应	11
1.6	图像质量的评价	15
1.6.1	概述	15
1.6.2	客观质量的评价	16
1.6.3	主观质量的评价	16
1.6.4	感兴趣区质量评价	17
1.6.5	感兴趣区质量评价的实验结果	19

1.7 讨论	21
参考文献	21

第2章

图像压缩编码技术的发展

2.1 概述	23
2.2 熵编码	24
2.3 传统的图像/视频编码技术	25
2.3.1 预测编码	25
2.3.2 变换编码	26
2.3.3 矢量量化	28
2.4 新型图像/视频编码技术	28
2.4.1 第二代图像编码方法	29
2.4.2 分形编码	29
2.4.3 模型编码	30
2.4.4 小波编码	31
2.5 图像/视频编码标准	32
2.5.1 静止图像编码的标准	35
2.5.2 视频编码标准	35
2.6 JPEG2000 标准	40
2.6.1 JPEG2000 体系结构与关键技术	40
2.6.2 JPEG2000 的主要特点与性能描述	45
2.7 讨论	46
参考文献	49

第3章

基于内容的图像检索技术

3.1 概述	52
3.2 图像的内容特征	54
3.2.1 颜色特征	55
3.2.2 纹理特征	57
3.2.3 轮廓特征	60
3.2.4 形状特征	60

3.2.5	对象特征	61
3.2.6	空间特征	63
3.2.7	字符特征	64
3.2.8	语义特征	64
3.2.9	图像的语义层次模型	66
3.3	图像相似度计算	67
3.3.1	距离度量	67
3.3.2	相关计算	68
3.3.3	关联系数计算	69
3.3.4	多特征相似度计算	70
3.3.5	具有光照不变性的图像检索	70
3.4	人机接口及相关反馈	71
3.4.1	友好的人机交互接口	71
3.4.2	相关反馈的新进展	74
3.5	图像的检索	76
3.5.1	低层视觉特征	76
3.5.2	局部图像	76
3.5.3	自定义特征检索	77
3.5.4	示例图与草图检索	77
3.5.5	浏览检索	77
3.6	图像检索系统	78
3.7	图像检索性能的评价	79
3.8	讨论	80
	参考文献	82

第二篇 基于视觉感知的个性化图像检索

第4章

视觉注意机制及注意模型

4.1	概述	89
4.2	人眼视觉系统的生理特性	90
4.2.1	视觉感官	92
4.2.2	视觉通路	93

4.2.3	视觉中枢	94
4.2.4	视觉感知	97
4.3	视觉注意机制	98
4.3.1	What 和 Where 视觉通路	98
4.3.2	预注意和注意阶段	99
4.3.3	自底向上和自顶向下的注意模式	99
4.3.4	选择性注意机制	100
4.3.5	注意焦点的选择与转移	102
4.4	视觉注意模型	102
4.4.1	基于空间的注意模型	103
4.4.2	基于特征的注意模型	111
4.4.3	基于空间和特征整合的注意模型	112
4.4.4	基于对象的注意模型	112
4.5	讨论	113
	参考文献	114

第 5 章

基于视觉注意模型的感兴趣区检测

5.1	概述	118
5.2	感兴趣区检测的研究进展	119
5.3	一种基于改进 Itti 模型与进化规划的感兴趣区检测方法	121
5.3.1	视觉显著度量	122
5.3.2	注意焦点的选择与转移	124
5.3.3	感兴趣区的生成	126
5.3.4	一种基于改进 Itti 模型与进化规划的感兴趣区检测方法 方法的实验	128
5.4	基于相似距离的 Top-Down 模型的初步研究	135
5.4.1	基于相似距离的 Top-Down 模型结构	135
5.4.2	相似区域的判定方法	136
5.4.3	基于相似距离的视觉期望	137
5.4.4	基于相似距离的 Top-Down 模型的初步实验	138
5.5	讨论	139
	参考文献	140

第 6 章 视点跟踪技术

6.1	概述	142
6.2	视点跟踪技术的研究进展	143
6.2.1	人眼的运动特性	143
6.2.2	视点跟踪设备	144
6.2.3	人眼检测与定位	145
6.2.4	视点跟踪技术分类	147
6.3	改造的视点跟踪设备	150
6.3.1	视点跟踪设备的选择及参数设置	150
6.3.2	视点跟踪设备的改造方案	154
6.3.3	方案选择与对比	157
6.4	基于 Gabor 变换的人眼定位方法	157
6.4.1	归一化处理	158
6.4.2	Gabor 小波变换	159
6.4.3	投影分析	160
6.4.4	基于 Gabor 变换的人眼定位方法的实验	161
6.5	一种基于双阈值的低复杂度视点跟踪方法	163
6.5.1	双阈值人眼定位算法	163
6.5.2	双阈值人眼定位算法的实验	166
6.5.3	基于几何映射的视点跟踪方法	167
6.5.4	一种基于双阈值的低复杂度视点跟踪方法的实验	168
6.6	一种基于 Web Camera 的自适应模板视点跟踪方法	169
6.6.1	自适应模板人眼定位算法	169
6.6.2	自适应模板人眼定位算法的实验	171
6.6.3	基于神经网络法的视点跟踪方法	172
6.6.4	一种基于 Web Camera 的自适应模板视点跟踪方法的实验	174
6.7	讨论	175
	参考文献	176

第 7 章 客观化的视觉感知技术

7.1	概述	179
-----	----	-----

7.2 客观化的视觉感知信息180

7.3 基于视觉感知的感兴趣区检测181

7.3.1 注视点分布图181

7.3.2 注视兴趣度183

7.3.3 基于视觉感知的感兴趣区184

7.4 基于视觉感知的感兴趣区检测实验平台186

7.4.1 实验平台186

7.4.2 实验结果188

7.5 讨论198

参考文献199

第8章
个性化图像检索技术

8.1 概述200

8.2 图像搜索引擎202

8.3 用户兴趣模型205

8.3.1 Agent 推理机205

8.3.2 个性化文件209

8.3.3 用户兴趣模型的更新213

8.4 基于规则的个性化推荐214

8.5 基于内容的个性化推荐218

8.5.1 概述218

8.5.2 潜在变量分析219

8.5.3 广义概率主分量分析模型221

8.6 协作过滤的个性化推荐224

8.6.1 概述224

8.6.2 基于记忆的协作过滤225

8.6.3 基于模型的协作过滤229

8.6.4 基于协作过滤的图像检索230

8.7 混合过滤的个性化推荐233

8.8 个性化推荐系统的评价235

8.9 相关反馈技术235

8.10 讨论237

参考文献	238
------	-----

第9章

应用于个性化图像检索的用户兴趣模型的构建

9.1 概述	242
9.2 用户兴趣模型描述	243
9.2.1 用户兴趣的表示方式	243
9.2.2 用户兴趣模型的分类	244
9.2.3 用户兴趣模型的构建方法	245
9.2.4 用户兴趣模型的框图	246
9.3 用户短期兴趣的获取	247
9.3.1 短期视觉特征	248
9.3.2 短期视觉特征实验结果	253
9.3.3 短期语义特征	257
9.3.4 短期语义特征实验结果	262
9.4 用户长期兴趣的获取	263
9.4.1 推理机的设计与实现	264
9.4.2 长期视觉特征	265
9.4.3 长期视觉特征实验结果	266
9.4.4 长期语义特征	270
9.4.5 长期语义特征实验结果	273
9.5 基于用户兴趣模型的个性化图像检索系统	273
9.5.1 个性化图像检索系统组成	273
9.5.2 个性化图像检索实验结果与分析	274
9.6 讨论	281
参考文献	282

第三篇 基于压缩域的图像处理

第10章

压缩域图像处理技术综述

10.1 概述	287
---------	-----

10.2	压缩域图像处理系统的构成	288
10.2.1	基于压缩域的图像检索系统框图	289
10.2.2	压缩域处理可操作位置	289
10.3	压缩域图像处理的研究方法	290
10.3.1	寻求对等操作	291
10.3.2	寻求特有操作	291
10.4	压缩域图像检索技术的研究进展	292
10.4.1	变换压缩域检索技术	292
10.4.2	空间压缩域检索技术	298
10.4.3	混合压缩域检索技术	299
	参考文献	300

第 11 章 压缩域纹理图像分类

11.1	概述	305
11.2	基于 DCT 压缩域区域能量方向性的纹理图像分类	306
11.2.1	基于 DCT 编码的基本框图	306
11.2.2	DCT 域区域能量的方向性	307
11.2.3	基于能量方向性的纹理图像分类算法设计	309
11.2.4	基于 DCT 压缩域区域能量方向性的纹理图像 分类算法实验	310
11.3	基于 DCT 压缩域的具有旋转不变性的纹理图像分类	312
11.3.1	DCT 域的多分辨率特性	312
11.3.2	具有抗旋转性的纹理图像分类方法	314
11.3.3	基于 DCT 压缩域的具有旋转不变性的纹理图像 分类算法实验	315
11.4	基于小波压缩域纹理图像分类	317
11.4.1	基于小波的纹理分析方法的发展	317
11.4.2	基于子带间相关性纹理特征提取	318
11.4.3	基于小波压缩域纹理图像分类算法	321
11.4.4	基于小波压缩域纹理图像分类算法实验	321
11.5	讨论	323
	参考文献	324

第 12 章

DCT 压缩域基于轮廓特征的图像检索

12.1	概述	327
12.2	基于图像检索的连通直方图方法	328
12.2.1	图像的连通直方图	328
12.2.2	基于连通直方图的图像检索	330
12.2.3	基于连通直方图的图像检索的实验	331
12.3	DCT 压缩域基于轮廓特征的图像检索方法	333
12.3.1	重组 DCT 系数	334
12.3.2	提取图像的轮廓	334
12.3.3	基于轮廓的连通直方图	336
12.3.4	DCT 压缩域基于轮廓特征的图像检索流程	337
12.3.5	DCT 压缩域基于轮廓特征的图像检索的实验	337
12.4	讨论	340
	参考文献	340

第 13 章

压缩域字符定位

13.1	概述	341
13.2	DCT 压缩域的字符定位方法	343
13.2.1	DCT 压缩域中字符特征的提取	343
13.2.2	DCT 压缩域的字符定位算法	344
13.2.3	DCT 压缩域字符定位算法的实验	345
13.3	基于加权频率的 DCT 压缩域的字符定位方法	347
13.3.1	DCT 压缩域字符区新特征——加权频率	347
13.3.2	基于加权频率的 DCT 压缩域字符定位算法	348
13.3.3	基于加权频率的 DCT 压缩域字符定位算法的实验	352
13.4	小波压缩域字符定位	355
13.4.1	字符特征在小波压缩域图像中的表示	355
13.4.2	基于小波压缩域的字符定位算法	357
13.4.3	基于小波压缩域字符定位算法的实验	362
13.5	讨论	363
	参考文献	364

第 14 章
压缩域的肤色分割

14.1 概述.....366

14.2 DCT 压缩域基于三维椭圆模型的自适应肤色分割.....367

14.2.1 三维椭圆肤色模型.....367

14.2.2 基于 DCT 直流系数的肤色初分割.....371

14.2.3 自适应修正分割结果.....372

14.2.4 DCT 压缩域中基于三维椭圆模型的自适应肤色分割实验.....373

14.3 小波压缩域双适形肤色分割模型.....375

14.3.1 双适形肤色模型.....375

14.3.2 小波压缩域基于双适形模型的肤色分割.....377

14.3.3 平滑滤波与连通域分析.....378

14.4 讨论.....378

参考文献.....379

第 15 章
压缩域的人脸检测

15.1 概述.....381

15.2 DCT 压缩域的人脸检测方法.....383

15.2.1 一种基于多级梯度能量特征的 DCT 压缩域人脸
检测算法框图.....383

15.2.2 多级梯度能量特征.....385

15.2.3 级联分类器.....387

15.2.4 搜索策略.....388

15.2.5 基于多级梯度能量描述的 DCT 压缩域人脸
检测算法的实验.....389

15.3 小波压缩域的人脸检测.....390

15.3.1 一种基于多级梯度能量描述的小波压缩域人
脸检测算法框图.....390

15.3.2 多级梯度能量描述.....391

15.3.3 级联检测器.....394

15.3.4 搜索策略.....396

15.3.5 结果仲裁.....397