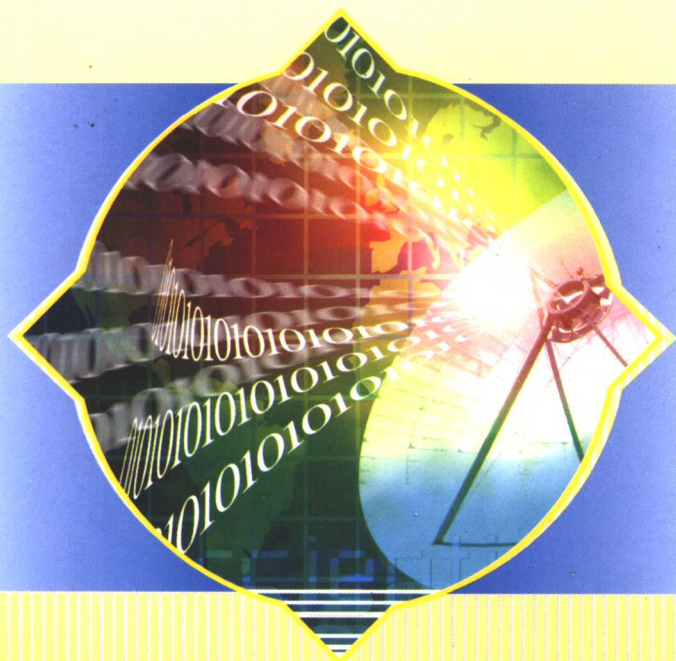


多媒体数据库技术

DUOMEITI SHUJUKU JISHU

黄志军 曾 斌 编著



國防工業出版社

<http://www.ndip.cn>

图书在版编目(CIP)数据

多媒体数据库技术 / 黄志军, 曾斌编著. —北京: 国防工业出版社, 2005. 1

ISBN 7-118-03570-X

I. 多... II. ①黄... ②曾... III. 多媒体—数据库系统 IV. TP311.134.3

中国版本图书馆 CIP 数据核字(2004)第 086286 号

国防工业出版社出版发行

(北京市海淀区紫竹院南路 23 号)

(邮政编码 100044)

国防工业出版社印刷厂印刷

新华书店经售

*

开本 850 × 1168 1/32 印张 11 1/4 279 千字

2005 年 1 月第 1 版 2005 年 1 月北京第 1 次印刷

印数: 1—3000 册 定价: 36.00 元

(本书如有印装错误, 我社负责调换)

国防书店: (010) 68428422

发行邮购: (010) 68414474

发行传真: (010) 68411535

发行业务: (010) 68472764

致 读 者

本书由国防科技图书出版基金资助出版。

国防科技图书出版工作是国防科技事业的一个重要方面。优秀的国防科技图书既是国防科技成果的一部分,又是国防科技水平的重要标志。为了促进国防科技和武器装备建设事业的发展,加强社会主义物质文明和精神文明建设,培养优秀科技人才,确保国防科技优秀图书的出版,原国防科工委于1988年初决定每年拨出专款,设立国防科技图书出版基金,成立评审委员会,扶持、审定出版国防科技优秀图书。

国防科技图书出版基金资助的对象是:

1. 在国防科学技术领域中,学术水平高,内容有创见,在学科上居领先地位的基础科学理论图书;在工程技术理论方面有突破的应用科学专著。

2. 学术思想新颖,内容具体、实用,对国防科技和武器装备发展具有较大推动作用的专著;密切结合国防现代化和武器装备现代化需要的高新技术内容的专著。

3. 有重要发展前景和有重大开拓使用价值,密切结合国防现代化和武器装备现代化需要的新工艺、新材料内容的专著。

4. 填补目前我国科技领域空白并具有军事应用前景的薄弱学科和边缘学科的科技图书。

国防科技图书出版基金评审委员会在总装备部的领导下开展工作,负责掌握出版基金的使用方向,评审受理的图书选题,决定资助的图书选题和资助金额,以及决定中断或取消资助等。经评审给予资助的图书,由总装备部国防工业出版社列选出版。

国防科技事业已经取得了举世瞩目的成就。国防科技图书承

IV

担着记载和弘扬这些成就,积累和传播科技知识的使命。在改革开放的新形势下,原国防科工委率先设立出版基金,扶持出版科技图书,这是一项具有深远意义的创举。此举势必促使国防科技图书的出版随着国防科技事业的发展更加兴旺。

设立出版基金是一件新生事物,是对出版工作的一项改革。因而,评审工作需要不断地摸索、认真地总结和及时地改进,这样,才能使有限的基金发挥出巨大的效能。评审工作更需要国防科技和武器装备建设战线广大科技工作者、专家、教授,以及社会各界朋友的热情支持。

让我们携起手来,为祖国昌盛、科技腾飞、出版繁荣而共同奋斗!

国防科技图书出版基金

评审委员会

前 言

几年前,多媒体研究和开发的主要焦点集中在多媒体通信和多媒体创作和显示。由于越来越多的图像、视频和音频等形式的数字多媒体数据正在被捕获和存储,并随着计算机网络的全面普及,近几年来焦点已经转移到高效和充分地存储和检索多媒体信息。这与 30 年前当越来越多的字母数字数据以计算机可读的形式存储时所发生的情况相似,当时的态势导致了数据库管理系统(DBMS)的发展。由于字母数字数据和多媒体数据的特征和要求之间有着很大的差别,因此传统的 DBMS 不能有效地处理多媒体数据,因而要求使用新的多媒体索引和检索技术。在这种情况下,多媒体数据库和基于多媒体内容检索的技术无疑将扮演一个十分重要的角色。为了构造出符合应用需要的多媒体数据库,不仅只是把多媒体数据加入到数据库中,还必须解决从体系结构到用户接口等一系列问题,这正是本书的研究重点。

本书综合分析了多媒体数据库管理的问题和当前的技术。首先介绍了多媒体的数据压缩、数据存储等相关技术,在此基础上,重点分析讨论了多媒体数据的主要特征和要求,然后描述了满足这些特征和要求的多媒体数据库管理系统的一般设计问题。我们讨论了索引和检索文本文档、音频、图像和视频的技术,这些技术的共性是它们都设法从原始媒体数据中抽取主要特征,然后根据存储项和查询的抽取特征向量之间的相似度或距离,设法检索与用户的查询有关的项目。由于多个媒体类型通常一同出现在多媒体对象或文档中,而且采用了不同的技术捕获这些对象或文

档的不同的特征或内容,因此,我们讨论了如何综合各种索引和检索技术有效地检索多媒体文档。本书重点说明的主要技术包括:

- 基于特征的多媒体文本、声音、图像、视频检索。多媒体信息具有丰富的内涵,给人以直观、形象、生动的感受,本书重点按照语音、颜色、形状、纹理、视频镜头等特征对多媒体信息进行检索。

- 信息全文检索。详细分析了信息全文检索方法,包括布尔检索、向量空间检索、概率检索、基于集群的检索和人工智能技术检索等多种检索机制,并以 Google 为例进行了综合说明。

- 内容的创建和获取。多媒体数据库需要将物理形式的资料转化为数字信息,并进行压缩和转化。多媒体信息载体由于采用了大量的图像、视频,其数据量比传统以文字为主的单媒体要大数百倍。鉴于数据的压缩及还原已经成了多媒体发展的一项关键技术,本书也对此进行了详细介绍。

- 信息存储管理。多媒体存储对象可以是文本、图形、图像等数字化信息。本书重点分析了大容量信息组织管理方法,探索多维空间索引方法以提高信息管理的有效性及检索的高效率。

- 多媒体搜索技术。多媒体数据库存储对象的数量和特征向量的维数非常巨大,要线性搜索所有存储的特征向量来找到查询目标会导致速度过慢。因此本书专门讨论了一些技术和数据结构以便快速定位与查询相关的特征向量。

本书适用于下列读者群体:

- 高年级的大学生和研究生。本书可作为多媒体数据库课程的一本教材。

- 学习开发多媒体数据库管理系统的系统设计人员和开发人员。

- 关心多媒体数据库新发展和最新研究方向的研究人员。
- 希望了解多媒体数据库技术问题和当前状态的其他学者。

本书由黄志军、徐红贤、彭建明、高志华、桂松、吴敏、潘杰、张晶、魏娜负责第一章、第四章~第七章的编写,曾斌、赵皑、潘祖金、姚路、陈海涛、戴沁芸、袁正刚、朱倩、钱兰华负责其余四章的编写,最后由黄志军、曾斌统稿、审校。由于时间仓促,书中不足和错误之处,恳请广大读者批评指正。

作者

国防科技图书出版基金

第四届评审委员会组成人员

名誉主任委员 陈达植

顾问 黄 宁

主任委员 刘成海

副主任委员 王 峰 张涵信 张又栋

秘书长 张又栋

副秘书长 彭华良 蔡 镭

委员 于景元 王小谟 甘茂治 冯允成

(按姓名笔画排序) 刘世参 杨星豪 李德毅 吴有生

何新贵 佟玉民 宋家树 张立同

张鸿元 陈火旺 侯正明 常显奇

崔尔杰 韩祖南 舒长胜

目 录

第一章 多媒体数据压缩技术.....	1
1.1 信息表示与编码	1
1.1.1 概述	1
1.1.2 数字图像编码技术	3
1.1.3 多媒体数据转换	4
1.2 常用的数据压缩技术	5
1.2.1 概述	5
1.2.2 预测编码	6
1.2.3 变换编码	6
1.2.4 信息熵编码	8
1.3 静态图像压缩标准 JPEG	11
1.3.1 JPEG 标准的主要内容	11
1.3.2 JPEG 静态图像压缩算法	12
1.4 运动图像压缩标准 MPEG	22
1.4.1 MPEG 标准简介	22
1.4.2 MPEG 音频	24
1.4.3 MPEG 视频数据流的结构	25
1.4.4 MPEG-1 视频编码技术	26
1.4.5 MPEG-2 标准	30
1.4.6 MPEG 标准的应用	34
1.5 视听通信编码解码标准 H.261	35
1.5.1 H.261 标准简介	35
1.5.2 $P \times 64\text{kb/s}$ 视频压缩编码算法	35

1.5.3	视频层次数据结构	37
1.5.4	视听业务视频压缩标准的发展	38
1.6	声音编码	39
1.6.1	声音压缩标准	39
1.6.2	CCITT 语音标准化方案	42
1.7	多媒体数据和应用的主要特征和要求	47
1.7.1	存储和带宽要求	48
1.7.2	多媒体信息的语义结构	49
1.7.3	延时和延时抖动要求	49
1.7.4	相关媒体的时间和空间关系	50
1.7.5	多媒体数据含义的主观性和模糊性	50
1.7.6	多媒体应用程序的数据库组织	50
1.7.7	多媒体系统的事务管理	52
1.8	小结	52
第二章	多媒体存储技术	54
2.1	磁介质技术	55
2.1.1	磁盘存储在多媒体上的使用	55
2.1.2	适用于 AV(音频/视频)的磁盘驱动器	58
2.2	分层存储管理	60
2.2.1	永久存储与瞬时存储	61
2.2.2	光盘库	62
2.2.3	分层存储的应用	64
2.3	存储系统的高速缓冲存储器管理	64
2.3.1	低层次的磁盘高速缓冲存储	66
2.3.2	分层存储系统的高速缓冲存储器结构	71
2.3.3	分布式客户机——服务器系统的高速缓冲存储器	75
2.3.4	多媒体应用软件如何使用高速缓冲存储器	78
2.4	磁盘调度及许可控制	80

2.4.1	传统的磁盘调度算法	81
2.4.2	最早截止时间优先算法	82
2.4.3	扫描最早截止时间优先算法	82
2.4.4	循环调度算法(Round - Robin)	83
2.4.5	组扫描调度(GSS)	84
2.5	用户交互准备	84
2.5.1	暂停和继续	85
2.5.2	快进和后退	86
2.5.3	与用户交互有关的 QoS 问题	88
2.6	服务器配置和网络连接	88
2.7	小结	89
第三章	多媒体数据库技术	91
3.1	一些重要的定义	92
3.1.1	媒体类型和多媒体	92
3.1.2	数据库和 DBMS(数据库管理系统)	93
3.1.3	文本文件信息检索	94
3.1.4	多媒体索引和检索	94
3.1.5	特征抽取、内容表示和索引	97
3.2	需要采用 MIRS 的原因	97
3.2.1	多媒体数据的扩张(Proliferation)及其特征	98
3.2.2	DBMS 及其在处理多媒体数据中的作用	98
3.2.3	IR 系统以及它们的多媒体检索中的作用	101
3.2.4	多媒体信息索引和检索的综合方法	102
3.3	MIRS 系统结构	102
3.3.1	概述	102
3.3.2	MIRS 基本结构	103
3.4	MIRS 的预期功能和常见应用	105
3.5	多媒体数据管理	107

3.5.1	概述	107
3.5.2	多媒体数据的管理环境	107
3.5.3	多媒体数据库管理系统特点	110
3.5.4	多媒体数据库的功能要求	114
3.5.5	数据库语言 SQL 及其进展	115
3.6	多媒体系统的设计目标	117
3.6.1	多媒体应用程序的数据库组织	119
3.6.2	多媒体数据库的特点	123
3.6.3	多媒体数据库的事务管理结构	126
3.6.4	将超媒体记录当作对象来管理	128
3.7	多媒体数据库的关键技术问题	131
3.7.1	多媒体数据的性质	131
3.7.2	数据模型	131
3.7.3	物理存储模型	139
3.7.4	体系结构	139
3.7.5	时空编组与数据模拟	143
3.7.6	查询处理与索引机制	153
3.7.7	用户接口技术	156
3.8	特征抽取、索引和相似性度量	160
3.8.1	特征抽取概念	161
3.8.2	索引结构	162
3.8.3	相似性度量	163
3.8.4	基于内容的检索技术	163
3.8.5	实例介绍	165
3.9	面向对象的数据库模型	167
3.9.1	优点	167
3.9.2	面向对象的数据库系统结构	168
3.9.3	面向对象的数据库系统的存储结构和 存取方法	170
3.10	小结	171

第四章 多媒体的索引和检索技术	172
4.1 引言	172
4.2 IR 系统和 DBMS 之间的区别	173
4.3 自动文本文档索引和布尔检索模型	175
4.3.1 基本的布尔检索模型	175
4.3.2 文件结构	176
4.3.3 条目操作和自动索引	178
4.3.4 自动文档索引小结	181
4.4 向量空间检索模型	182
4.4.1 基本的向量空间检索模型	182
4.4.2 相关反馈技术	183
4.5 概率检索模型	184
4.6 基于集群的检索模型	185
4.6.1 集群的产生	185
4.6.2 基于集群的检索	186
4.7 非传统的 IR 方法	186
4.8 性能度量	187
4.9 不同的 IR 技术之间的性能比较	189
4.10 WWW 搜索引擎	190
4.10.1 WWW 简介	191
4.10.2 资源发现	193
4.10.3 IR 系统和 WWW 搜索引擎的主要区别	195
4.10.4 WWW 搜索引擎的一般结构	199
4.10.5 搜索引擎的一个例子	199
4.11 小结	203
第五章 音频的索引和检索	204
5.1 引言	204
5.2 主要音频性能和特征	207
5.2.1 从时域中推导出的特征	209
5.2.2 从频域中推导出的特征	210

5.2.3	频谱图	212
5.2.4	主观特征	213
5.3	音频分类	213
5.3.1	不同类型声音的主要特征	213
5.3.2	音频分类框架	214
5.4	语音识别和检索	216
5.4.1	语音识别	216
5.4.2	发音者识别	222
5.4.3	小结	222
5.5	音乐索引和检索	222
5.5.1	结构化音乐和声音效果的索引和检索	223
5.5.2	基于样本的音乐的索引和检索	224
5.6	使用音频和其它媒体之间的关系对多媒体信息 进行索引和检索	226
5.7	小结	227
第六章	图像索引和检索	230
6.1	引言	230
6.2	图像索引和检索的各种不同的方法	231
6.3	基于文本的图像检索	233
6.4	基于颜色的图像索引和检索技术	235
6.4.1	基于颜色的图像检索技术	237
6.4.2	基本技术的改进	238
6.5	基于形状的图像检索	245
6.5.1	常用术语的定义和一些简单的形状度量 方法	246
6.5.2	不变矩量	247
6.5.3	傅里叶描述符方法	248
6.5.4	重要边界的直方图	249
6.5.5	感兴趣点的排序列表	250
6.5.6	弹性模板匹配	250

6.5.7 基于区域的形状表示和相似性度量	251
6.6 基于纹理的图像检索	255
6.7 基于压缩图像数据的图像索引和检索	256
6.7.1 基于 DCT 系数的图像索引和检索	257
6.7.2 基于小波系数的图像索引和检索	258
6.7.3 基于 VQ 压缩数据的图像索引和检索	258
6.8 其它图像索引和检索技术	260
6.8.1 基于模型压缩的图像检索	260
6.8.2 基于空间关系的图像检索	261
6.9 综合图像索引和检索技术	261
6.9.1 QBIC	261
6.9.2 Virage 图像搜索引擎	262
6.9.3 WebSEEK	262
6.9.4 ImageRover WWW 搜索引擎	263
6.10 小结	263
第七章 视频索引和检索	266
7.1 引言	266
7.2 基于镜头的视频索引和检索综述	267
7.3 视频镜头检测或分割	268
7.3.1 基本的视频分割技术	269
7.3.2 用渐进变化检测镜头边界	270
7.3.3 防止虚假的镜头检测	271
7.3.4 其它镜头检测技术	272
7.3.5 压缩视频的分割	273
7.4 视频索引和检索	274
7.4.1 基于 r 帧的索引和检索	275
7.4.2 基于动态信息的索引和检索	277
7.4.3 基于对象的索引和检索	278
7.4.4 基于元数据的索引和检索	279
7.4.5 基于注释的索引和检索	279

7.4.6 视频索引和检索的综合方法	280
7.5 有效的视频表示和抽象	280
7.5.1 题目或主题分类	281
7.5.2 动态图标或视频图标	282
7.5.3 视频横幅(Streamer)	283
7.5.4 剪辑图	283
7.5.5 分级视频浏览器	283
7.5.6 故事板(Storyboard)	284
7.5.7 马赛克	284
7.5.8 情景转换图	285
7.5.9 视频撤除	285
7.6 小结	285
第八章 综合多媒体索引和检索	287
8.1 引言	287
8.2 综合索引和检索技术	288
8.2.1 综合音频的索引和检索	289
8.2.2 综合图像的索引和检索	290
8.2.3 综合视频的索引和检索	290
8.2.4 基于单个特征所获得的结果的合并	291
8.2.5 媒体变换	292
8.3 多媒体信息管理的一般体系结构	292
8.4 用户界面	295
8.5 系统例子	296
8.5.1 QBIC	297
8.5.2 Monash 大学开发的综合 WWW 图像搜索 引擎	298
8.5.3 元搜索(MetaSearch)引擎	301
8.6 小结	304
第九章 有效的多媒体相似性搜索的技术和数据结构	305
9.1 引言	305

9.2 降低搜索空间的过滤过程	307
9.2.1 使用分类、结构化属性和关键词的过滤	307
9.2.2 基于三角不等式的方法	308
9.2.3 基于颜色直方图检索的方法	309
9.2.4 基于向量空间 IR 的潜在语义索引	311
9.3 B ⁺ 树和 B 树	312
9.3.1 B ⁺ 树	313
9.3.2 B 树	315
9.4 集群	315
9.5 多维 B ⁺ 树	316
9.5.1 二维空间中 MB ⁺ 树的综述	317
9.5.2 构建一个二维 MB ⁺ 树	318
9.5.3 MB ⁺ 树中的搜索	318
9.5.4 高维 MB ⁺ 树	320
9.6 k-d 树	320
9.7 网格文件	321
9.8 R 树族	322
9.8.1 R 树结构综述	322
9.8.2 区域对象的搜索、插入和删除	323
9.8.3 点数据的搜索、插入和删除	324
9.8.4 R 树的搜索效率	324
9.8.5 R [*] 树、R ⁺ 树和 VAMSplit R 树	325
9.8.6 SS 树和 SS ⁺ 树	325
9.9 TV 树	326
9.10 小结	326
参考文献	328