



普通高等教育“十一五”国家级规划教材

21世纪高等学校计算机规划教材

21st Century University Planned Textbooks of Computer Science

多媒体技术 教程 (第3版)

Multimedia Technology (3rd Edition)

胡晓峰 吴玲达 老松杨 司光亚 编著

- 体现了作者多年的教学经验
- 吸收了多种国外著名教材的优点
- 强调基本原理，概念准确、深入浅出、理论适中



名家系列



人民邮电出版社
POSTS & TELECOM PRESS



普通高等教育“十一五”国家级规划教材

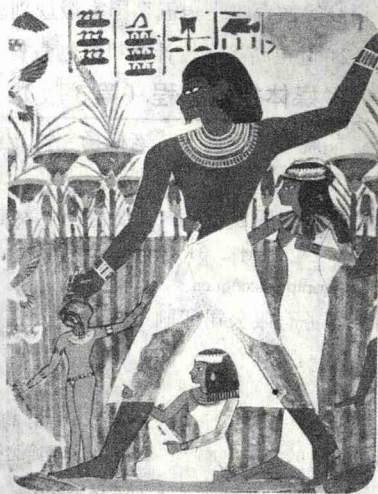
21世纪高等学校计算机规划教材

21st Century University Planned Textbooks of Computer Science

多媒体技术 教程 (第3版)

Multimedia Technology (3rd Edition)

胡晓峰 吴玲达 老松杨 司光亚 编著



人民邮电出版社

样书

专用章



名家系列

人民邮电出版社

北京

图书在版编目 (C I P) 数据

多媒体技术教程 / 胡晓峰等编著. —3版. —北京: 人民
邮电出版社, 2009. 4
21世纪高等学校计算机规划教材
ISBN 978-7-115-18940-0

I. 多… II. 胡… III. 多媒体技术—高等学校—教材
IV. TP37

中国版本图书馆CIP数据核字 (2008) 第155124号

内 容 提 要

本书对多媒体技术的基本概念、技术与系统等全面的论述。全书共分 11 章, 主要内容包括多媒体基本概念、媒体处理方法与技术、多媒体软硬件平台、多媒体信息管理及处理、网络多媒体技术, 以及多媒体分布应用等。本书既注重介绍多媒体技术的基础知识, 也适当介绍一些基本理论和方法, 难易适中。各章之后附有习题与思考题。

本书可作为计算机及其相关专业本科生和研究生的教材, 也可供从事多媒体技术研究的工程技术人员参考。

普通高等教育“十一五”国家级规划教材

21 世纪高等学校计算机规划教材

多媒体技术教程 (第 3 版)

-
- ◆ 编 著 胡晓峰 吴玲达 老松杨 司光亚
责任编辑 邹文波
 - ◆ 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>
三河市海波印务有限公司印刷
 - ◆ 开本: 787×1092 1/16
印张: 18.25
字数: 475 千字 2009 年 4 月第 3 版
印数: 75 001—78 000 册 2009 年 4 月河北第 1 次印刷

ISBN 978-7-115-18940-0/TP

定价: 29.80 元

读者服务热线: (010)67170985 印装质量热线: (010)67129223

反盗版热线: (010)67171154

计算机科学与技术日新月异的发展,对我国高校计算机人才的培养提出了更高的要求。许多高校主动研究和调整学科内部结构、人才培养目标,提高学科水平和教学质量,精炼教学内容,拓宽专业基础,优化课程结构,改进教学方法,逐步形成了“基础课程精深,专业课程宽新”的良性格局。作为大学计算机教材建设的生力军,人民邮电出版社始终坚持服务高校教学、致力教育资源建设的出版理念,在总结前期教材建设的成功经验的同时,深入调研和分析课程体系,并充分结合我国高校计算机教育现状和改革成果,推出“推介名师好书,共享教育资源”的教材建设项目,出版了“21世纪高等学校计算机规划教材”名家系列。

本套教材的突出特点如下:

(1) 作者权威 本套教材的作者均为国内计算机学科中的学术泰斗或高校教学一线的教学名师,他们有着深厚的科研功底和丰富的教学经验。可以说,这套教材汇聚了众师之精华,充分显示了这套教材的格调和品位。无论是刚入杏坛的年轻教师,还是象牙塔内的莘莘学子,细细品读其中的章节文字,定会收益匪浅。

(2) 定位准确 本套教材是为普通高等院校的学生量身定做的精品教材。具体体现在:一是本套教材的作者长期从事一线科研和教学工作,对高校教学有着深刻而独到的见解;二是本套教材在选题策划阶段便多次召开调研会,对普通高校的教学需求和教材建设情况进行充分摸底,从而保证教材在内容组织和结构安排上更加贴近实际教学;三是组织有关作者到较为典型的普通高等院校讲授课程教学方法,深入了解教师的教学需求,充分把握学生的理解能力,以教材内容引导授课教师严格按照科学方法实施教学。

(3) 教材内容与时俱进 本套教材在充分吸收国内外最新计算机教学理念和教育体系的同时,更加注重基础理论、基本知识和基本技能的培养,集思想性、科学性、启发性、先进性和适应性于一身。

(4) 一纲多本,合理配套 根据不同的教学法,同一门课程可以有多本不同的教材,教材内容各具特色,实现教材系列资源配套。

总之,本套教材中的每一本精品教材都切实体现了各位教学名师的教学水平,充分折射出名师的教学思想,淋漓尽致地表达着名师的教学风格。我们相信,这套教材的出版发行一定能够启发年轻教师们真正领悟教学精髓,教会学生科学地掌握计算机专业的基本理论和知识,并通过实践深化对理论的理解,学以致用。

我们相信,这套教材的策划和出版,无论在形式上还是在内容上都能够显著提高我国高校计算机专业教材的整体水平,为培养符合时代发展要求的具有较强国际竞争力的高素质创新型计算机人才,为我国普通高等教育的计算机教材建设工作做出新的贡献。欢迎各位老师和读者给我们的工作提出宝贵意见。

第3版前言

从应用意义上讲,多媒体不是哪一种设备的升级换代,也不是什么新的发明,它标志着数字化全面主导信息领域各种技术的一个崭新时代。从20世纪90年代初开始,多媒体技术进入到了计算机、家用电子、通信、出版、娱乐、网络等几乎所有的信息领域,大家都在谈论多媒体,我们作为参与者,当然也会因为有些“先知先觉”而自我陶醉。有学生问我,多媒体技术发展到未来,会是什么样子?我当时的回答可能连我自己都很吃惊:多媒体技术没有未来,多媒体技术发展的目标,就是要消灭自己。到了今天,这个预言正在成为现实,多媒体技术已经融合进入了各个领域,成为必不可少的组成部分,但实际上也逐步地在回归信息技术的本身,而不用专门的突出多媒体技术自己。所有的计算机都超出了原先定义的多媒体计算机的范畴,所有的信息技术也都要考虑多媒体信息的需要了。也正是因为如此,多媒体技术也成为信息技术相关专业的必修课程,反而得到了更多的关注和普及。

本书的起源和完善已经过了很长的时间。1993年,我们根据自己教学的需要,编写了《多媒体系统原理与应用》一书作为教材内部使用,并于1995年由人民邮电出版社正式出版。该书第1版印刷6次,被许多学校选为教材,受到了读者广泛的欢迎。1997年,在多年的多媒体研究与教学的基础上,我们又撰写了《多媒体系统》一书,着重介绍了多媒体系统的设计原理和技术,成为了前书的姊妹篇,并成为国家“九五”规划重点教材。该书也被许多大学选为研究生教材。随着时间的推移,多媒体系统与技术已经有了很大的发展,我们感到原先书中介绍的许多内容已经有些陈旧,不太适合教学的要求。2002年,在许多朋友、专家和教师的鼓励下,我们又编写这本适合于高等院校本科生或低年级研究生教学使用的多媒体技术教材。该书又得到了广大读者的好评,成为了更多大学的教材。2005年,该书进行了第一次修订,并被评选为“十一五”国家级规划教材,再版发行并且多次印刷,得到教师和学生的欢迎。当然,也有许多读者特别是在第一线从事多媒体技术教学的教师给我们提出了很多很好的建议和修改意见,并希望能够在再次修订的时候进行充实和完善。非常感谢人民邮电出版社安排该书出第3版,这让我们有了做到这一点的机会。

本书在原修订版的基础上增加了一些新的内容,更新了部分章节,也删减了一些过时的内容,并且借这个机会改正了原书中的一些错误。全书现在缩减为11章,但增加了一个附录性质的多媒体实验参考,供教学使用。主题内容为4大部分:第一部分是多媒体基本概念,介绍媒体、数据压缩等基本内容;第二部分介绍多媒体的软件、硬件平台,以及基于时间的媒体和同步等;第三部分是多媒体信息管理及处理,介绍超媒体与Web技术、多媒体数据库技术,多媒体信息分析与处理等内容;第四部分是多媒体通信网络与分布应用,介绍网络多媒体技术,以及多媒体分布应用系统等新内容。每章都附有习题与思考题,供读者练习使用。考虑到本书读者主要是大学本科生、硕士研究生这样的层次,以及教师在讲授时的方便,书中的

主要内容既要具有一定的系统性,以基础知识为主,又适当介绍了一些理论方面的内容,但难度比较适中。本书以介绍多媒体的主要技术内容为主,同时也吸收了一些新的研究成果,供学生了解最新的研究进展。教师在讲授时,可以根据需要全部讲授或部分讲授。本教材是按40~60学时编写的,教师在实际授课时可根据选择的内容确定学时。为保证有较长的教材使用周期,本书着重于原理的讲解,而不对某种特定的产品进行介绍。

本书是在原作基础上仍由几位作者分工完成的,具体分工是:胡晓峰负责修订第1章、第2章、第6章、第7章和第8章;吴玲达负责修订第3章、第10章和第11章;老松杨负责修订第4章、第5章并编写附录;司光亚负责修订第11章。最后由胡晓峰、吴玲达进行统稿。许多同志特别是我们所带的博士、硕士研究生,在研究工作中及本书的写作过程中给予了我们很多的帮助,在这里我们表示衷心的感谢!

多媒体技术是一门综合性很强的技术,学科面非常宽,发展又快,而我们的水平和能力有限,书中存在缺点和错误是难免的,如蒙指正将不胜感谢。

作 者

目 录

第 1 章 绪论..... 1

1.1 多媒体的基本概念..... 1

1.1.1 媒体与多媒体..... 1

1.1.2 多媒体的关键特性..... 2

1.2 多媒体技术的产生与发展..... 3

1.2.1 多媒体是技术与应用发展的必然..... 3

1.2.2 多媒体改善了人类信息的交流..... 5

1.2.3 多媒体缩短了人类传递信息的路径..... 6

1.3 多媒体技术研究的主要内容..... 6

1.3.1 多媒体技术的基础..... 7

1.3.2 多媒体软硬件平台技术..... 7

1.3.3 多媒体信息管理与处理技术..... 8

1.3.4 网络多媒体与分布应用技术..... 8

1.4 小结..... 8

习题与思考题..... 9

第 2 章 媒体及媒体技术..... 10

2.1 媒体的种类和特点..... 10

2.1.1 常见的媒体元素..... 10

2.1.2 媒体的种类..... 13

2.1.3 媒体的性质和特点..... 14

2.2 听觉媒体技术..... 16

2.2.1 声音心理学..... 16

2.2.2 音频的数字化和符号化..... 19

2.2.3 音频媒体的三维化处理..... 22

2.3 视觉媒体技术..... 24

2.3.1 视觉心理学..... 24

2.3.2 模拟视频原理..... 26

2.3.3 视觉媒体数字化..... 26

2.3.4 视觉媒体的三维立体显示..... 28

2.4 触觉媒体技术..... 31

2.4.1 触觉媒体概述..... 31

2.4.2 简单指点设备与技术..... 32

2.4.3 位置跟踪..... 33

2.4.4 力反馈与触觉反馈..... 35

2.5 小结..... 36

习题与思考题..... 36

第 3 章 多媒体数据压缩..... 38

3.1 多媒体数据压缩技术概述..... 38

3.1.1 数据冗余的类型..... 38

3.1.2 数据压缩方法的分类..... 39

3.1.3 数据压缩技术的性能指标..... 39

3.2 常用的数据压缩编码方法..... 40

3.2.1 DPCM 和 ADPCM 编码..... 40

3.2.2 离散余弦变换编码..... 43

3.2.3 小波变换编码..... 44

3.2.4 行程编码和 Huffman 编码..... 45

3.3 音频压缩标准..... 46

3.3.1 音频压缩编码的基本方法..... 46

3.3.2 电话质量的语音压缩标准..... 47

3.3.3 调幅广播质量的音频压缩标准..... 48

3.3.4 高保真立体声音频压缩标准..... 48

3.4 图像和视频压缩标准..... 49

3.4.1 图像和视频压缩编码的基本方法..... 49

3.4.2 静止图像压缩标准..... 51

3.4.3 视频压缩标准..... 58

3.5 小结..... 69

习题与思考题..... 70

第 4 章 多媒体硬件环境..... 71

4.1 多媒体存储设备..... 71

4.1.1 光存储的类型..... 71

4.1.2 CD-ROM 光存储系统..... 72

4.1.3 CD-R 光存储系统..... 73

4.1.4 DVD 光存储系统..... 74

4.1.5 闪存与闪存卡..... 75

4.2 音频接口..... 76

4.2.1 音频卡的工作原理..... 76

4.2.2 音乐合成和 MIDI 接口规范..... 79

4.2.3 语音合成	80	6.1.2 同步的种类	128
4.2.4 语音识别	80	6.1.3 同步的分层服务模型	129
4.2.5 MP3 播放器	81	6.2 时间模型	130
4.3 视频接口	82	6.2.1 时间依赖的定义	130
4.3.1 视频图像显示	82	6.2.2 时间的概念模型	131
4.3.2 视频卡/盒	87	6.3 时间的规范与表示	132
4.4 多媒体 I/O 设备	88	6.3.1 时间规范	132
4.4.1 笔输入	88	6.3.2 相对时间规范与绝对时间规范	132
4.4.2 触摸屏	90	6.4 多媒体同步的表示方法	134
4.4.3 扫描仪	92	6.4.1 基于图形的表示法	135
4.4.4 数码相机	94	6.4.2 基于脚本的表示法	137
4.4.5 数码摄像机	96	6.4.3 交互和同步	140
4.4.6 虚拟现实的三维交互工具	97	6.5 时间同步与系统支持	141
4.4.7 输入/输出接口	99	6.5.1 概述	141
4.5 小结	102	6.5.2 系统支持的有关问题	142
习题与思考题	102	6.6 小结	143
习题与思考题	102	习题与思考题	143
第 5 章 多媒体软件基础	103	第 7 章 超媒体与 Web 系统	144
5.1 多媒体软件系统层次	103	7.1 超媒体系统的组成	144
5.2 多媒体素材制作软件	104	7.1.1 超媒体的概念和发展历史	144
5.2.1 文本编辑与录入软件	104	7.1.2 超媒体的组成要素	147
5.2.2 图形图像编辑与处理软件	105	7.1.3 超媒体系统的 Dexter 模型	152
5.2.3 音频编辑与处理软件	107	7.2 Web 超媒体系统	154
5.2.4 视频编辑	109	7.2.1 分布式超媒体系统 WWW	154
5.2.5 动画编辑	111	7.2.2 WWW 中的超媒体协议与标记语言	155
5.3 多媒体著作工具	113	7.3 智能超媒体系统	161
5.3.1 多媒体著作工具概述	113	7.3.1 智能超媒体的概念	161
5.3.2 Authorware 概述	115	7.3.2 智能超媒体的节点和链	162
5.3.3 Director 概述	117	7.3.3 智能超媒体的推理模型	162
5.4 多媒体软件编程技术	118	7.4 Web 搜索技术	164
5.4.1 媒体控制接口(MCI)	118	7.4.1 Web 搜索的概念	164
5.4.2 多媒体控件	121	7.4.2 Web 搜索技术及算法	165
5.4.3 多媒体软件开发工具包	122	7.5 小结	168
5.4.4 多媒体应用设计过程	124	习题与思考题	168
5.5 小结	125		
习题与思考题	126	第 8 章 多媒体数据库	169
第 6 章 多媒体的时间表示与同步	127	8.1 多媒体数据管理的问题	169
6.1 多媒体同步的基本概念	127	8.1.1 概述	169
6.1.1 同步的基本概念	127		

8.1.2 多媒体数据管理的问题·····170	9.5.1 多媒体特征融合·····215
8.1.3 多媒体数据与数据库管理·····172	9.5.2 单媒体交叉索引·····215
8.2 多媒体数据库体系结构·····173	9.5.3 单媒体结果融合·····216
8.2.1 多媒体数据库的一般结构形式·····173	9.5.4 融合多特征的新闻故事单元 探测方法·····217
8.2.2 多媒体数据库的层次结构·····175	9.6 小结·····218
8.3 多媒体数据模型·····176	习题与思考题·····218
8.3.1 NF ² 数据模型·····176	
8.3.2 面向对象数据模型·····177	
8.3.3 其他数据模型·····179	
8.4 多媒体数据库的用户接口·····180	第10章 网络多媒体技术 ·····219
8.4.1 字符数值型接口·····181	10.1 多媒体网络·····219
8.4.2 示例型接口·····182	10.1.1 多媒体网络的通信需求·····219
8.4.3 用户表现接口·····184	10.1.2 多媒体网络的服务质量·····222
8.5 小结·····186	10.2 流媒体技术·····225
习题与思考题·····186	10.2.1 流媒体技术概述·····225
	10.2.2 流媒体播送技术·····229
	10.2.3 流媒体技术原理·····230
	10.2.4 流媒体技术应用·····231
第9章 多媒体信息分析与处理 ·····187	10.3 无线多媒体通信技术·····232
9.1 基于内容检索的系统结构和方法·····187	10.3.1 无线多媒体通信网的系统结构·····232
9.1.1 基于内容检索的概念·····187	10.3.2 无线多媒体通信的关键技术·····232
9.1.2 基于内容检索系统的一般结构·····188	10.3.3 无线多媒体通信新技术·····233
9.1.3 基于内容检索的过程和指标·····189	10.4 多媒体通信协议·····236
9.2 图像内容分析及检索方法·····193	10.4.1 IPv6·····237
9.2.1 图像特征的提取与表达·····193	10.4.2 RSVP·····238
9.2.2 图像相似性检索与匹配方法·····197	10.4.3 RTP·····239
9.2.3 图像检索中的相关反馈机制·····199	10.4.4 RTSP·····241
9.3 视频结构化与视频摘要·····200	10.5 小结·····241
9.3.1 视频媒体基本特性·····200	习题与思考题·····242
9.3.2 镜头边界探测·····201	
9.3.3 关键帧提取·····203	
9.3.4 故事单元边界探测·····204	第11章 多媒体应用系统 ·····243
9.3.5 视频语义对象提取·····206	11.1 多媒体会议系统·····243
9.3.6 视频摘要·····207	11.1.1 多媒体会议系统概述·····243
9.4 基于内容的音频处理与检索·····208	11.1.2 多媒体会议系统结构与关键 技术·····245
9.4.1 音频媒体基于内容检索的概念·····208	11.1.3 多媒体会议系统的国际标准·····248
9.4.2 音频结构化·····209	11.1.4 虚拟空间会议系统·····251
9.4.3 音频特征提取·····210	11.2 交互视频服务系统·····252
9.4.4 音频分类·····211	11.2.1 交互视频服务系统的概念·····252
9.4.5 音频分割·····212	11.2.2 交互视频服务系统的组成与 结构·····254
9.4.6 基于内容的音频检索·····213	
9.5 多媒体融合分析与检索·····214	

11.2.3 用户接入网技术·····	257	A1.3 Micromedia Authorware 上机 实践·····	268
11.3 CSCW 和群件·····	259	A2 多媒体程序设计实验·····	271
11.3.1 CSCW 概述·····	259	A2.1 媒体控制接口（MCI）编程 实验·····	271
11.3.2 群件系统·····	260	A2.2 多媒体控件编程实验·····	274
11.4 小结·····	263	A2.3 多媒体软件开发工具包编程 实验·····	277
习题与思考题·····	264		
附录 A 多媒体技术实验 ·····	265	参考文献 ·····	280
A1 多媒体著作与编辑软件上机实践·····	265		
A1.1 Adobe Photoshop 上机实践·····	265		
A1.2 Adobe Premiere 上机实践·····	267		

第 1 章

绪论

多媒体化是信息化发展的一个必然阶段，是一个崭新的技术时代。多媒体引起了诸多信息技术的集成与融合的革命，它将计算机、家用电器、通信网络、大众媒体、人机交互和娱乐机器等原先界限分明的东西，融合成了新的系统、新的应用。多媒体与因特网（Internet）一起，成为推动 20 世纪末、21 世纪初信息化社会发展两个最重要的技术动力之一。多媒体技术与系统的产生和发展，是社会信息化发展的必然结果。

1.1 多媒体的基本概念

1.1.1 媒体与多媒体

所谓媒体（Medium）是指承载信息的载体。按照 ITU-T 建议的定义，媒体有以下 5 种：感觉媒体、表示媒体、显示媒体、存储媒体和传输媒体。感觉媒体指的是用户接触信息的感觉形式，如视觉、听觉、触觉等。表示媒体则指的是信息的表示形式，如图像、声音、视频、运动模式等。显示媒体（又称表现媒体）是表现和获取信息的物理设备，如显示器、打印机、扬声器、键盘、摄像机、运动平台等。存储媒体是存储数据的物理设备，如磁盘、光盘等。传输媒体是传输数据的物理设备，如光缆、电缆、电磁波、交换设备等。这些媒体形式在多媒体领域中都是密切相关的，但一般说来，如不特别强调，我们所说的媒体主要是指表示媒体，因为作为多媒体技术来说，主要研究的还是各种各样的媒体表示和表现技术。

“多媒体”（Multimedia），从字面上理解就是“多种媒体的综合”，相关的技术也就是“怎样进行多种媒体综合的技术”。多媒体技术概括起来说，就是一种能够对多种媒体信息进行综合处理的技术。略为全面一点，多媒体技术可以定义为：以数字化为基础，能够对多种媒体信息进行采集、编码、存储、传输、处理和表现，综合处理多种媒体信息并使之建立起有机的逻辑联系，集成为一个系统并能具有良好交互性的技术。

特别需要指出的是，很多人将“多媒体”看作是计算机技术的一个分支，这是不太准确的。多媒体技术以数字化为基础，注定其与计算技术要密切结合，甚至可以说要以计算技术为基础。但多媒体技术中还有许多东西原先并不属于计算技术的范畴，例如，电视技术、广播通信技术、印刷出版技术等。当然可以有多媒体计算机技术，但也可以有多媒体电视技术、多媒体通信技术。一般说来，“多媒体”指的是一个很大的领域，指的是和信息有关的所有技术与方法进一步发展的领域。所以说，要对多媒体有更准确的理解，更多的是要从它的关键特性上去考虑。

1.1.2 多媒体的关键特性

多媒体的关键特性主要包括信息载体的多样性、交互性和集成性这3个方面,这既是多媒体的主要特征,也是在多媒体研究中必须解决的主要问题。

1. 信息载体多样性

信息载体的多样性是相对于计算机而言的,指的就是信息媒体的多样化,有人称之为信息多维化。把计算机所能处理的信息空间范围扩展和放大,而不再局限于数值、文本或是被特别对待的图形或图像,这是计算机变得更加人性化所必须具备的条件。

人类对于信息的接收和产生主要在5个感觉空间内,即视觉、听觉、触觉、嗅觉和味觉,其中前三者占了95%以上的信息量。借助于这些多感觉形式的信息交流,人类对于信息的处理可以说是得心应手。但是,计算机以及与之相类似的一系列设备,都远远没有达到人类处理信息能力的水平。在信息处理的传统过程中不得不忍受着种种变态:信息只能按照单一的形态才能被加工处理,只能按照单一的形态才能被理解。计算机在许多方面需要把人类的信息进行变形之后才可以使用,例如将中文转换成某种代码才能输入计算机。可以说,在信息交互方面计算机还处于初级水平。

多媒体就是要把机器处理的信息多样化或多维化,使之在信息交互的过程中,具有更加广阔和自由的空间。多媒体的信息多维化不仅仅指输入,还指输出。但输入和输出并不一定都是一样的。对于应用而言,前者称为获取(Capture),后者称为表现(Presentation)。如果两者完全一样,这只能称为记录和重放,从效果上来说并不是很好。如果对其进行变换、组合和加工,亦即我们所说的创作或综合,就可以大大丰富信息的表现力和增强效果。这些创作与综合也不仅仅局限在对信息数据方面,还包括对设备、系统、网络等多种要素的重组和综合,目的都是能够更好地组织信息、处理信息和表现信息,从而使用户更全面、更准确地接受信息。

2. 交互性

多媒体的第2个关键特性是交互性。长久以来,人们在很多情况下已经习惯于被动地接收信息,例如看电视、听广播。多媒体系统将向用户提供交互式使用、加工和控制信息的手段,为应用开辟更加广阔的领域,也为用户提供更加自然的信息存取手段。

交互可以增加对信息的注意力和理解力,延长信息在头脑中保留的时间。但在单向的信息空间中,这种接收的效果和作用就很差,只能“使用”所给的信息,很难做到自由地控制和干预信息的获取和处理过程。多媒体信息在人机交互中的巨大潜力,主要来自于它能提高人对信息表现形式的选择和控制能力,同时也能提高信息表现形式与人的逻辑和创造能力结合的程度。多媒体信息比单一信息对人具有更大的吸引力,它有利于人对信息的主动探索而不是被动地接收。在动态信号与静态信号之间,人更倾向于前者。多媒体信息所提供的种类丰富的信息源恰好能够满足人在这个方面的需要。

当交互性引入时,“活动”本身作为一种媒体便介入到了数据转变为信息、信息转变为知识的过程之中。因为数据能否转变为信息取决于数据的接收者是否需要这些数据,而信息能否转变为知识则取决于信息的接收者能否理解。借助于交互活动,可以获得我们所关心的内容,获取更多的信息,例如,对某些事物进行选择,有条件地找出事物之间的相关性,从而获得新的信息内容。对某些事物的运动过程进行控制可以获得某种奇特的效果,例如,倒放、慢放、快放、变形、虚拟等,从而激发学生的想象力、创造力,制造出各种讨论的主题。在某些娱乐性应用中,用户可以改变故事的结局,从而使用户介入到故事的发展过程之中。即使是最普遍的信息检索应用,用户也可以找出想读的书籍、想看的电视节目,可以快速跳过不感兴趣的部分,可以对某些所关心

的内容插入书评、进行编排等，从而改变现在使用信息的方法。

可以想象，交互性一旦被引入到用户的活动之中，将会带来多大的作用。从数据库中检录出某人的照片、声音及文字材料，这是多媒体的初级交互应用；通过交互特性使用户介入到信息过程中（不仅仅是提取信息），才达到了中级交互应用水平。当我们完全地进入到一个与信息环境一体化的虚拟信息空间自由遨游时，这才是交互式应用的高级阶段，这就是虚拟现实（Virtual Reality）。人机交互不仅仅是一个人机界面的问题，对于媒体的理解和人机通信过程可以看成是一种智能的行为，它与人类的智能活动有着密切的关系。

3. 集成性

多媒体系统充分体现了集成性的巨大作用。事实上，多媒体中的许多技术在早期都可以单独使用，但作用十分有限。这是因为它们是单一的、零散的，如单一的图像处理技术、声音处理技术、交互技术、电视技术、通信技术等。但当它们在多媒体的旗帜下集合时，一方面意味着技术已经发展到了相当成熟的程度；另一方面，也意味着各种技术独自发展不再能满足应用的需要。信息空间的不完整，例如仅有静态图像而无动态视频，仅有语音而无图像等，都将限制信息空间的信息组织，限制信息的有效使用。同样，信息交互手段的单调性、通信能力的不足、多种设备和应用的人为分离，也会制约应用的发展。因此，多媒体系统的产生与发展，既体现了应用的强烈需求，也顺应了全球网络的一体化、互通互连的要求。

多媒体的集成性主要表现在两个方面，一是多媒体信息媒体的集成，二是处理这些媒体的设备与设施的集成。首先，各种信息媒体应该能够同时地、统一地表示信息。尽管可能是多通道的输入或输出，但对用户来说，它们就应该是一体的。这种集成包括信息的多通道统一获取，多媒体信息的统一存储与组织，以及多媒体信息表现合成等各方面。因为多媒体信息带来了信息冗余性，可以通过媒体的重复、使用别的媒体或是并行地使用多种媒体的方法消除来自于通信双方及环境噪声对通信产生的干扰。由于多种媒体中的每一种媒体都会对另一种媒体所传递信号的多种解释产生某种限制作用，所以多种媒体的同时使用可以减少信息理解上的多义性。总之，不应再像早期那样，只能使用单一的形态对媒体进行获取、加工和理解，而应注意保留媒体之间的关系及其所蕴涵的大量信息。其次，多媒体系统是建立在一个大的信息环境之下的，系统的各种设备与设施应该成为一个整体。从硬件来说，应该具有能够处理各种媒体信息的高速及并行的处理系统、大容量的存储、适合多媒体多通道的输入输出能力及外设、宽带的通信网络接口，以及适合多媒体信息传输的多媒体通信网络。对于软件来说，应该有集成一体化的多媒体操作系统、各个系统之间的媒体交换格式、适合于多媒体信息管理的数据库系统，以及合适的创作工具和各类应用软件等。

多媒体中的集成性应该说是在系统级的一次飞跃。无论是信息、数据，还是系统、网络、软硬件设施，通过多媒体的集成性构造出支持广泛信息应用的信息系统， $1+1>2$ 的系统特性将在多媒体信息系统中得到充分的体现。

1.2 多媒体技术的产生与发展

1.2.1 多媒体是技术与应用发展的必然

多媒体技术的概念起源于 20 世纪 80 年代初期，但真正蓬蓬勃勃发展起来是在 90 年代。多媒体并不是新的发明，从某种意义上说，它是信息技术与应用发展的必然。多媒体是在计算技术、

通信网络技术、大众传播技术等现代信息技术不断进步的条件下,由多学科不断融合、相互促进而产生出来的。

计算机中信息的表达最初只能用二进制的 0、1 来表示,它的目的纯粹是为了计算。但在应用过程中,这种 0、1 的形式使用起来非常不方便,后来就产生了像 ASCII 这一类的字符代码。将字符处理过程引入到计算机中,不仅方便了用户,而且也使计算机不再局限于计算的范围,而进入了事务处理领域。中文标准代码的出现和使用很大程度上依赖于计算机图形技术和软件技术的发展,使之能够以一种图形的方法来表达信息。随后计算机开始处理图形、图像、语音、音乐,直至发展到能处理影像视频信息。这个过程就是计算机的多媒体化的过程,当然早期的集成度还相当不够。与此同时,在大众传播及娱乐界,从印刷技术开始了电子化、数字化的过程,逐步发展了广播、电影、电视、录像、有线电视直至推出的交互式光盘系统、高清晰度电视(HDTV),并且逐渐开始具有交互能力。在这个过程中,通信网络技术的发展,从邮政、电报电话,一直到计算机网络等,一方面不断地扩展了信息传递的范围和质量,另一方面又不断支持和促进了计算机信息处理和通信、大众信息传播的发展。从信息系统的角度来看,这些各自目的不同、技术不同而又相互促进和支持的领域之间,尽管未来很明显会全部统一起来,但由于技术发展、商业利益等各方面的原因,对于最终用户而言还存在着较大的差别,至今还没有完成完全会师的宏愿。

多媒体技术直接起源于计算机工业界、家用电器工业界和通信工业界对各自领域未来发展的预测。最早研究和提出多媒体系统的分别是计算机工业的代表 IBM、Intel、Apple 及 Commodore 公司,以及家用电器工业的代表 Philips、Sony 公司等。他们从两个方面提出的发展方向和目标可以说是不谋而合,都是要推出能够交互综合处理多媒体信息的设备或系统。IBM 和 Intel 公司联合推出的 DVI(Digital Video Interactive)可使计算机能够处理影像视频信息,这就使得计算机跨入了传统的电视领域。Microsoft 公司等一大批软件开发商以多媒体应用为契机,推出的各类多媒体软件,造就了一大批计算机的多媒体应用和用户。而以 Philips 和 Sony 公司为代表的家用电器工业,将电视技术进行改进,使其向智能化、有交互能力方向发展,交互式视频光盘 CD-I 是他们最早的尝试。现在又与通信网络普遍结合,开发出的电视机顶盒(Set Top Box)、大规模视频服务器,也显示出了交互式电视的潜在能力。通信工业也不甘落后,不仅在通信传输、电话终端等方面保持优势,而且在许多新的领域大力拓展,努力开发新一代的产品。可视电话、视频会议、远程服务、综合电话终端等都是通信工业在技术上的新的领地。

进入 20 世纪 90 年代以后,由于“信息高速公路”计划的兴起,Internet 的广泛使用,刺激了多媒体信息产业的发展和网络互连的需求,在全球掀起了一股家电行业、有线电视网、娱乐行业、计算机工业及通信工业相互兼并、联合建网的浪潮,从而使得 20 世纪 90 年代被称为“多媒体时代”。计算机、通信、家电和娱乐业的大规模联合,造就了新一代的信息领域,产生了崭新的信息社会概念,也创造了无穷的机遇和潜在的市场。从技术发展的历史来看,直到今天,各种在各自领域中独立发展的技术,终于要走到一起来了。这一发展导致了计算机工业、家用电子行业、通信网络业以及大众传媒业的融合和竞争,最终受益的将是广大用户。技术发展到这一步,往往意味着一个旧时代的结束,一个新时代的开始,而新时代的疆域又需要我们去不断开拓。

在多媒体产生过程中,数字化充当了极为重要的角色。早期的模拟系统起始于 20 世纪 70 年代,采用的都是模拟设备,如模拟光盘,每面光盘可存放 30min 录像。模拟光盘具有随机访问功能,便于计算机进行控制。在这一阶段开发的多是模拟教学系统,如领航学习系统、模拟旅游系统等。这些早期的工作显示了多媒体应用开发的潜力,然而多媒体系统只有向数字化方向发展,才能达到更高的技术水平,才能更好地支持应用。随着存储技术、计算技术、通信技术的发展,

基于数字化的多媒体系统在 20 世纪 80 年代不断涌现,将系统的交互能力、媒体质量、处理灵活性等方面提高到了一个新的水平。宽带数字网络的发展,使系统的集成性有了基础,不再局限于个人计算领域,而是向分布综合服务的发展方向发展。

无论从技术还是应用角度来看,多媒体都是发展的必然。这不仅仅是研究和生产某一类设备的问题,而是在信息系统级上的重新组合和调整,它将意味着更加剧烈的竞争和更加光辉灿烂的前景,这一点无论是对于研究界还是对于工商界,都是如此。

1.2.2 多媒体改善了人类信息的交流

计算机和其他信息机器的发展使得人类的信息处理手段得到加强,高速的计算能力扩展了对数据进行重复计算的能力,大规模的存储扩展了记忆信息的范围,高速通信网使得我们可以同远在地他乡和同事、朋友、亲人甚至陌生人进行快速的信息交换。这些机器成为我们与他人进行交流的中介。但是计算机缺乏类似于人类眼睛、耳朵等感官所得到的视觉、听觉以及触觉、嗅觉、味觉的能力,无法从现实世界中自由地收集信息和表达各种信息,又使得计算机成为了在信息交流中的一道深深的鸿沟。人类如果要借助于机器的能力,就必须忍受交互过程中信息的转换和变态,这不能不说是一大遗憾。多年来,人们一直都在致力于消除这个遗憾。多媒体的目标也是如此。

用户及计算机的信息交流采用 4 种形式:人—人(经由计算机)、人—计算机、机—人和计算机—计算机。其中每一种交流形式在信息的表示和传递方面都各有不同,可能仅仅是数据的转移(即无解释),也可能是在信息传递时形式被转换、数据被重新组织,表现也被改变。

从计算机科学的角度来看,计算机—计算机的通信是一门已经得到充分发展的技术,尽管多媒体系统的需求已向系统设计者提出了新的要求。计算机之间的通信包括通过网络传送数据、存储数据以备后续的检索等。虽然已经开发了许多关于图像、声音的信息交换标准,但是不同标准、不同程序和不同类型计算机之间的交流仍比较困难。动画和影像视频媒体的实时性对信息网络还具有十分严格的时间要求。应用对于图像、声音类媒体的需求,也要求在传输带宽上和存储空间上应大大高于现有系统的量级。多媒体信息系统的产生,为综合考虑多媒体信息的处理,统一数据格式化和相应的传输协议与标准奠定了基础。在未来,不用再人为地割裂为几个网或几个专用系统,统一的网络传输标准将使所有的信息机器得到有效的沟通。

在人与人通过计算机交流方面,计算机起着高效信息传递媒介的作用。其中一个很重要的原因是计算机不必理解人与人交流通信中的全部内容。例如,在电子邮件中只有收信者、目的地、日期等成分需要计算机去解释和执行,但对邮件本身计算机只把它当作一串字符或一堆数据。与之相类似,声音、图像、视频以及其他一些类型的媒体,基本上都可以被存储和传送,而不必被计算机理解。尽管如此,还是需要计算机帮助人们有效地组织和表达信息,多媒体系统将在这个方面发挥出更大的作用。计算机将成为人与人之间交流的“宽”通道,而不是像从前那样只能使用文本的“窄”通道;计算机将支持更多的人与人交流的应用,从医学会诊、学术讨论等协同性工作,到私人信函、可视电话等个人间信息交互。如果计算机能够理解信息的含义,就将显著地改变计算机所支持的人与人之间的交互方式,使之达到一个更高的水平,例如自动语言翻译等,但这在目前还处于十分初级的阶段。

人与计算机之间的交互必须考虑两者的局限性。为了使计算机发挥它所应有的作用,必须遵循“可计算”的 3 个条件,即:能用形式化的方法描述这个问题,能找到一个算法解决这个形式化的问题,以及能以一个合理的复杂程度在计算机上实现这个算法。这就是说,为了使用计算机,必须把人类头脑中大部分属于并发的、联想的、形象的、模糊的和多样化的思维强行地翻译成

冯·诺依曼计算机所能接受的串行的、刻板的、明确的、严格遵守形式逻辑规则的机器指令。这种翻译过程不仅仅是烦琐和机械的，而且技巧性很强，因机器而异。机器所能接受和处理的也仅仅是数字化了的信息。虽然几十年来，计算机已从庞大的玻璃房子里走出来放在了我们的桌上，计算机操作系统（它的目的就是要协助用户用好计算机）也从任务调度、资源管理逐步发展到了具有图形化的用户接口，但仍然很难为最终用户所接受。这之中很重要的原因便是我们更多考虑的仍是计算机，而非如何使用计算机，尤其在研究界，这种现象更为明显。过去用户接口工作常常被认为“缺乏理论深度”而遭到冷遇，直到工业界做出了出色的成绩，人们才猛醒过来。事实上，用户接口工作往往会成为一个系统或一次研究成败的关键部分。多媒体的出现，将会在这个方面起到至关重要的作用。很多人把多媒体技术归到了用户接口领域，虽不全面，但也有一定道理。

1.2.3 多媒体缩短了人类传递信息的路径

信息的巨大物化力量主要表现在信息的共享特性上。当人们真正认识到信息共享是开展信息技术研究的首要任务之后，就必须研究和探索什么是表示、传送和处理信息的较好途径。比较理想的途径应是能较完整地表示概念，能较迅速地传递概念，能以符合人类的认识过程的方式加工概念的方法，从而使得完成某个智力任务的行为过程得到较大的改善。所谓“较好”的途径，随着时间的推移会有不同的标准，但总的来说就是使得一个人头脑中的一个概念成为另一个人头脑中对于那个概念的理解的路径较短。

与以往的方法相比，计算机在数据处理方面有了很大的改善。计算机所提供的功能强大的数据组织和构造技术，如传统数据结构中的数组、向量、队列、堆栈、树、堆等，为动态地加工和处理数据提供了基础。高效的算法和高速的网络通信，大大地加强了用文字和数据表示概念的能力并加速了它的传递过程。但人类并不是仅仅依赖文本这一类单一的数据形式来传递所有的信息和接受概念的，图像、声音等多媒体信息都是人类获取和传递信息极为重要的渠道。图像的信息量最大，一幅画胜过千言万语，最直观、最能一目了然。而动态的影像视频和动画则更生动、更逼真、更接近客观世界的原型、更能反映事物的本质和内涵。声音和文字也是信息的重要媒体，综合应用不仅有利于接受，也有利于存储（记忆）和保留。这就意味着必须同时启动大脑的形象思维和逻辑思维，才能更好地获得更多更有用的信息。因此，通过多种感觉器官用多种信息媒体形式向人提供信息才算是更好的表达方法，它不仅加速和改善了理解，并且提高了信息接受的兴趣和注意力。多媒体正是利用各种信息媒体形式，集成地用声、图、文等来承载信息，也就是缩短信息传递的路径。

另外，多媒体技术也使得信息的包装可以随意改变，变被动为主动，根据自己的需求“量体裁衣”选择最好的信息“包装”形式，这也是改善信息交流的极为重要的方面。“最好的”方法就是“最合适”的方法，这是因人而异的，并不是所有的信息都成为影像视频才是最好，也并不是在程序中加入声音就能增进信息的理解，或许这些都是画蛇添足。如何运用和协调各种媒体，正是要重点加以研究的问题。

1.3 多媒体技术研究的主要内容

可以认为多媒体是一个技术时代，需要研究的内容几乎遍及所有与信息相关的领域。多媒体的研究一般分为两个主要的方面：一是多媒体技术，主要关心基本技术层面的内容；二是多媒体

系统,主要重心在多媒体系统的构成与实现。这两个方面也不是截然分开的,只是侧重点不同而已。另外,还有专门研究多媒体创作与表现的,则更多的属于艺术而不属于技术的范畴。本书将主要强调多媒体技术基础范畴内的内容。

1.3.1 多媒体技术的基础

研究多媒体首先要研究媒体。媒体是传播信息的载体,首先就要研究媒体的性质与相应的处理方法。传统的媒体形式对计算机来说主要是文字和数值,但人类更加熟悉的是图形、图像和声音。媒体对人类来说不仅仅是个表示与表达的问题,在很大程度上与人类的知觉有关,与心理学有关。例如,对媒体的研究就要弄清人类对视觉和听觉的依赖究竟达到什么样的程度,从心理学的角度搞清人类的视觉和听觉的特性将对媒体的处理会产生什么样的影响等。人类的另一个重要的感觉是触觉,在交互性达到更高的程度时,触觉就必不可少。人类的的不同感觉器官实际上是在同时工作的,每一种感觉之间也会相互影响,产生出“感觉相乘”的效应,加强感觉的效果。对每一种媒体的采集、存储、传输和处理就是多媒体系统要做的首要工作。

关于多媒体的另一个技术基础是数据压缩。众所周知,各种媒体数据之间具有很大的差别。文本数据即使带有非常复杂的格式说明,它的数据量按现在的标准也不算很大。而基于时间的媒体,特别是高质量的视频数据媒体,哪怕很短的时间都会占用很大的存储空间。尽管现在各种存储设备已经具有很大的容量,通信网络已经具有很大的带宽,但采用相应的压缩技术对媒体进行压缩,还是多媒体数据处理的必要基础。数据压缩技术,或者称为数据编码技术,不仅可以有效地减少媒体数据占用的空间,也可以减少传输占用的时间,例如 JPEG、MPEG1、MPEG2 和 H.26L 等;另一方面,这些编码还可以用于复杂的内容处理场合,增强对信息内容的处理能力,例如 MPEG4 等。

本书将用 2 章分别介绍媒体的基本特性和媒体数据压缩编码的基本方法和标准。

1.3.2 多媒体软硬件平台技术

软件及硬件平台是实现多媒体系统的物质基础。在过去的研究与开发中,每一项重要的技术突破都直接影响到多媒体的发展与应用的进程。大容量的光盘、数字视频交互卡 DVI、带有多媒体功能的软件 Windows 3.0 等都曾直接推动了多媒体技术向前迅速发展。在这个方面,输入、输出、处理、存储、管理、传输等都是需要研究的内容,包括各种技术和设备。

在硬件方面,各种多媒体的外部设备现在已经成为了标准配置,例如光盘驱动器、声音适配器、图形显示卡等,而现在的计算机 CPU 也都加入了多媒体与通信的指令体系,许多过去不敢想象的性能在现有的计算机上都成为了可能。扫描仪、彩色打印机、带振动感的鼠标、机顶盒、交互式键盘遥控器、数码相机等,都越来越普及到每个家庭。像 DVI、CD-I 这两个最早的典型视频多媒体接口虽具有里程碑意义,但当前确实已经没有任何应用价值。多媒体现在已经在向更复杂的应用体系发展,其硬件平台自然要更加复杂,例如视频点播系统、虚拟现实系统等。目前在基于网络的、集成一体化的多媒体设备上还在做更多的努力。

在软件方面,随着硬件的进步,也在快速发展。从操作系统、编辑创作软件,到更加复杂的专用软件,早已是遍地开花,形成了相应的工业标准,产生了一大批多媒体软件系统。特别是在 Internet 发展的大潮之中,一批基于 Internet 的多媒体应用软件更是独领风骚,还促进了网络的应用发展。

研究多媒体,必须首先研究媒体的时间和空间性质与相应的处理方法,必须要充分考虑基于