

多媒体技术与应用

DUOMEITI JISHU YU YINGYONG

叶 绿 编著



浙江大学出版社

多媒体技术与应用

叶 绿 编著

浙江大學出版社

内 容 提 要

本书第1章、第2章介绍多媒体技术相关基础知识、数字图形图像处理基础和算法;第3章介绍图像识别技术及实现;第4章介绍各种压缩编码;第5章介绍小波变换压缩算法;第6章、第7章、第8章主要讨论图像和视频数据压缩技术及实现,并分别详细介绍多媒体数据压缩标准 JPEG、MPEG,以及 H. 261 标准和视频会议系统;第9章介绍多媒体数据的数字水印技术;第10章讨论多媒体超文本和超媒体文档结构,以及介绍超文本标记链接语言 HTML 语言。最后第11章是课后习题和部分参考答案。

本书可作为大学本科高年级或研究生多媒体课程的教材,也适用于作为对多媒体技术及应用感兴趣的研发人员的参考书。

图书在版编目(CIP)数据

多媒体技术与应用 / 叶绿编著. —杭州:浙江大学出版社, 2004. 11

ISBN 7-308-03959-5

I. 多... II. 叶... III. 多媒体技术 IV. TP37

中国版本图书馆 CIP 数据核字 (2004) 第 102989 号

责任编辑 周卫群

封面设计 俞亚彤

出版发行 浙江大学出版社

(杭州浙大路 38 号 邮政编码 310027)

(E-mail: zupress@mail. hz. zj. cn)

(网址: <http://www.zjupress.com>)

排 版 浙江大学出版社电脑排版中心

印 刷 浙江大学印刷厂

开 本 787mm×1092mm 1/16

印 张 20

字 数 512 千

版 次 2004 年 11 月第 1 版 2006 年 3 月第 2 次印刷

印 数 3001—4500

书 号 ISBN 7-308-03959-5/TP·275

定 价 28.00 元

前言

多媒体技术以及应用领域是当前的研究热点,早在 20 世纪 80 年代末 90 年代初全世界范围内就形成了一股研究多媒体计算、通信和应用技术的热潮,并且取得了举世瞩目的成果。随着计算机和网络技术的发展,它们处理和传输的已不仅仅是文字和图像,视频、音频和其他连续媒体已成为计算机应用系统的一部分。毫无疑问,多媒体技术是一项综合技术,涉及多个学科。随着多媒体技术的飞速发展、网络信息能力的提高和计算机处理速度的不断增长,通过 Internet 使大量的图像、视频、音频、动画和图形等多媒体信息传输到世界各个角落。多媒体把文字、图形、图像、动画、音频和视频集成到计算机中,人们能更加自然,更加“人性化”地利用信息。人机交互的这种变化极大地拓展了我们的信息空间,满足了人们对多种媒体信息进行统一处理的需要。

作者通过多年的研究,并结合自己多年在多媒体技术以及应用等方面的计算机本科专业教学、指导毕业设计的经验编著了这本书。本书主要是针对数字图像处理方法、数字图像识别技术、多媒体关键技术压缩编码,以及介绍在相关应用领域中多媒体应用技术的书籍,可作为大学本科或研究生多媒体课程的教材,也适于作为对多媒体技术及应用感兴趣的研究开发人员的参考书。由于多媒体技术发展迅速,一些多媒体技术有了很大变化,因而在本书中增加了很多新的内容。相信读者在理解本书所阐述的技术的基础上,就可以进行多媒体技术应用的开发与开发。

本书在以下几方面有自己的特色:

(1)内容全面,包括了数字图形图像处理基础和算法、多媒体技术的多项研究内容,涉及多媒体信息的压缩、多媒体技术的应用等。

(2)叙述清楚,通俗易懂,给出了详细的参考文献目录,从事这一研究方向的科技人员可以根据它们检索自己感兴趣的内容。

(3)全书配以大量的图表,方便读者的理解。另外,本书配有练习题。

(4)内容新颖,并且有一定的技术深度,可供从事这一研究方向的科技人员参考。

全书共分为 11 章,每一章都有丰富的信息量并且叙述深入浅出,通俗易懂,前 5 章对多媒体技术相关基础知识、数字图形图像处理基础和图像识别的应用技术,各种压缩编码技术及小波变换压缩算法进行了介绍;第 6 章、第 7 章、第 8 章讨论图像和视频数据压缩技术及实现,并分别介绍了常用的多媒体数据压缩标准 JPEG、MPEG,以及 H. 261 标准和视频会议系统。第 9 章介绍多媒体数据的数字水印技术,第 10 章讨论多媒体文档系统,介绍了超文本和超媒体文档结构,以及超文本标记链接语言 HTML。第 11 章是各种题型的课后习题和答案。

作者于 1993 年至 1994 年在德国布伦瑞克—沃尔芬比特应用科学大学(University of Applied Science Braunschweig / Wolfenbuettel)留学进修期间,参与了 Friedhelm Seutter 教授和 Peter Franke 高级工程师主持的相关技术方向的课题的研究工作,对我今后从事这方面的研究打下了基础,也为本书的写作建立了良好的铺垫。在本书完成之即,谨向他们两位教授以

及德国同事表示衷心的感谢。

衷心感谢浙江省自然科学基金对本项目(项目编号:M603178)的资助。可以说没有该基金项目的资助,书中的有些研究成果就难以实现,也就没有本书的出版。

对浙江科技学院计算机系 99 级学生叶明球等历届计算机系学生对“多媒体技术”课程的建设 and 研究付出的努力,以及对作者在浙江大学计算机学院攻读博士研究生学习阶段各位导师的指导和帮助,在此一并表示衷心的感谢。

由于作者水平有限,时间紧迫,再加上多媒体技术是当前计算机科学的前沿技术,发展迅速,书中错误和不足之处,还请各位专家、学者、读者不吝指教,多提宝贵意见,希望在再版时予以修正。

叶 绿

浙江科技学院计算机与电子工程学系

2004 年 9 月

目 录

第 1 章 多媒体:媒体和数据流	1
1.1 媒体有不同表现形式	1
1.1.1 视觉类媒体	1
1.1.2 听觉类媒体	2
1.1.3 其他类媒体	3
1.2 媒体之间的相对性质	3
1.2.1 多媒体的交互性	3
1.2.2 多媒体数据自身特点	3
1.3 多媒体系统主要特性	4
1.3.1 多媒体系统定义	4
1.3.2 媒体的组合	4
1.3.3 独立性	4
1.3.4 计算机支持的媒体集成特性	4
1.3.5 数字化特性	5
1.3.6 通信系统	5
1.4 媒体	5
1.4.1 感知媒体	5
1.4.2 表示媒体	5
1.4.3 表现媒体	6
1.4.4 存储媒体	6
1.4.5 传输媒体	6
1.4.6 信息交换媒体	6
1.4.7 表示值和表示空间	6
1.4.8 表示维数	7
1.5 多媒体技术定义	7
第 2 章 数字图像和图形	8
2.1 基本概念	8
2.1.1 数字图像表示	8
2.1.2 图像格式	10
2.2 图像数学模型	13
2.2.1 色彩模式	14
2.2.2 平面图像表示	16

2.2.3	数字图像表示	17
2.3	计算机图像处理	17
2.3.1	图像合成	17
2.3.2	图像处理基本算法	24
2.3.3	从空间到频域正交变换(Orthogonal Transform)	37
第3章	图像识别技术	46
3.1	图像特征	46
3.2	图像识别步骤	47
3.2.1	调整	47
3.2.2	标记	47
3.2.3	归类	48
3.2.4	特征提取	48
3.2.5	匹配	48
3.3	传输图像的数据量	48
3.4	图像识别与实现	49
3.5	图像识别实例:人脸识别	51
3.5.1	人脸检测方法	52
3.5.2	肤色在人脸检测中的应用	52
3.5.3	肤色模型	53
3.5.4	人脸区域的分割	54
3.5.5	基于区域的物体分割	54
3.5.6	基于肤色的人脸检测算法	55
3.5.7	人脸区域分割算法	59
3.5.8	本例中使用的人脸区域分割算法	60
3.5.9	五官特征检测	63
3.5.10	双眼和嘴巴轮廓的提取	64
3.5.11	算法实现	66
第4章	数据压缩编码	76
4.1	编码需求	76
4.2	编码分类	78
4.3	基本压缩编码技术	80
4.3.1	行程编码	81
4.3.2	矢量量化	81
4.3.3	双元编码	82
4.3.4	香农—范诺编码	82
4.3.5	霍夫曼编码	83
4.3.6	算术编码	85
4.3.7	词典编码	88
4.3.8	线性预测编码	97

4.3.9	脉冲编码调制(PCM)	98
4.3.10	变换编码	101
第5章	小波变换压缩算法	103
5.1	小波分析理论发展史	103
5.2	离散余弦变换	103
5.2.1	傅立叶(Fourier)分析	104
5.2.2	短时傅立叶变换 STFT	105
5.2.3	小波分析理论	106
5.3	小波变换	107
5.3.1	连续小波变换(CWT)	107
5.3.2	小波分析和傅立叶分析的主要比较	108
5.3.3	离散小波变换	108
5.4	多分辨率分析与 Mallat 塔式变换算法	109
5.4.1	多分辨率分析概念	109
5.4.2	Mallat 塔式算法	110
5.5	常用小波基简介	111
5.5.1	Haar 小波	112
5.5.2	Daubechies 小波系	112
5.5.3	Biorthogonal 函数系	112
5.5.4	Coiflet 函数	113
5.5.5	Symlets 函数系	113
5.5.6	常用小波基比较	115
5.6	零树法编解码	115
5.6.1	图像小波变换后系数的分布特点	115
5.6.2	零树法编码过程	116
5.6.3	小波编码实例	119
5.6.4	零树法译码过程	122
5.6.5	解码实例	122
第6章	JPEG 压缩技术	124
6.1	JPEG 技术	124
6.2	电视图像数字化	125
6.2.1	彩色电视制式	125
6.2.2	电视扫描和同步	126
6.3	数字化标准	129
6.4	图像子采样	130
6.5	图像准备	131
6.6	基于 DCT 的有损顺序模式	134
6.6.1	图像处理	134
6.6.2	量化	135

6.6.3 熵编码	136
6.7 扩展的基于 DCT 的有损模式	138
6.8 基于预测的无损模式	138
6.9 层次模式	139
6.10 JPEG 2000 压缩算法	140
6.10.1 JPEG2000 的新特征及其应用领域	140
6.10.2 JPEG2000 的基本框架和实现	142
6.10.3 JPEG2000 中的核心算法	143
6.11 一个简单 JPEG 算法程序举例	143
第 7 章 MPEG 压缩技术	174
7.1 视频编码	174
7.1.1 4:2:0 YCbCr 格式	174
7.1.2 MPEG 四种不同的图像帧	176
7.2 寻找最佳宏块搜索法	179
7.2.1 二维对数搜索法(2D—logarithmic search)	179
7.2.2 三步搜索法(three-step search)	180
7.2.3 对偶搜索法(conjugate search)	180
7.3 编码	181
7.4 电视图像的结构	181
7.5 音频编码	182
7.5.1 音频流	182
7.5.2 视频流	182
7.6 MPEG-2 技术	183
7.7 MPEG-4 技术	186
7.7.1 MPEG-4 电视图像编码	186
7.7.2 电视图像对象区的概念	186
7.7.3 电视图像编码方案	187
7.7.4 电视图像分辨率可变编码	187
7.8 数字视频交互技术(DVI)	189
7.8.1 音频与图像编码	189
7.8.2 视频编码	190
7.9 MPEG-7 多媒体内容描述接口	191
7.10 网络视频(流媒体)技术	192
7.10.1 流媒体的传统意义	192
7.10.2 流媒体关键技术	193
7.10.3 流媒体发展现状	194
7.11 小结	194
第 8 章 视频编码标准	196
8.1 H. 261 视频编码标准	196

8.1.1	图像准备	196
8.1.2	编码算法	197
8.1.3	数据流	197
8.2	网络视频会议的两种框架结构	198
8.3	网络视频会议系统的关键技术	198
8.4	H. 323 的协议框架结构	199
8.4.1	H. 323 协议体系	199
8.4.2	网络视频会议系统框架结构	200
8.4.3	视频会议系统与 CSCW	200
8.5	会议系统 C/S 体系结构设计	201
8.5.1	视频会议软件模块分割	201
8.5.2	以 MCU 为中心的星形结构	202
8.6	紧耦合模式的管理	203
8.6.1	紧耦合与松耦合的会议体系	203
8.6.2	多播与体系结构	203
8.7	Windows 平台下客户端模块功能实现	204
8.7.1	客户端软件模块流程	204
8.7.2	客户端多媒体信息采集与压缩	205
8.8	过程设计	205
8.8.1	音频、视频传输模块的设计实现	206
8.8.2	RTP 与 RTCP 拆/封包设计	207
8.9	紧耦合模式下的会议控制与管理	210
8.9.1	会议控制的功能划分	210
8.9.2	成员控制	210
8.9.3	元会议管理(Meta-Conference)	211
8.9.4	实现模型	214
8.9.5	Server 端的多媒体通信	217
8.10	面向两种连接方式的数据传输	218
8.10.1	信令传输策略	218
8.10.2	H. 323 的传输协议	219
8.10.3	H. 225 呼叫信令	219
8.10.4	通信实体的连接及实现	219
8.10.5	不同系统终端的多媒体数据同步	224
8.11	MCU 对多点的传输——组播技术及实现	224
8.11.1	IP 组播技术	224
8.11.2	多播在系统中的实现	225
8.12	视频会议的安全问题	226
8.12.1	视频会议的安全保密要求	226
8.12.2	用户级的安全措施	226
8.12.3	算法的选择	226

第 9 章 多媒体数据的数字水印技术	228
9.1 数字水印技术基本理论	229
9.2 数字水印的分类及特性	229
9.2.1 数字水印的分类	230
9.2.2 水印特性	231
9.2.3 数字图像水印模型	232
9.3 数字水印典型算法和攻击分析	233
9.3.1 数字水印典型算法	233
9.3.2 数字水印系统攻击分析	235
9.4 三维运动盲水印检测算法	236
9.4.1 相关技术	237
9.4.2 统计检测理论	238
9.4.3 DCT 变换系统的统计分析	239
9.4.4 最优相关性检测	239
9.4.5 水印机制	240
9.4.6 嵌入算法	240
9.4.7 检测算法	241
9.4.8 检测阈值的确定	241
9.4.9 实验结果	242
9.5 三维运动盲水印提取算法	243
9.5.1 相关技术	244
9.5.2 量化	244
9.5.3 扩展频谱	245
9.5.4 运动信号自由度排序	246
9.5.5 小结	248
第 10 章 超文本和超媒体	249
10.1 文档	249
10.1.1 文档结构	249
10.1.2 多媒体数据处理	249
10.2 超文本和超媒体	250
10.2.1 超文本、超媒体和多媒体	251
10.3 超文本标记链接语言(HTML)	252
10.3.1 HTML 简介	252
10.3.2 文档元素和标签的概念	253
10.4 HTML 语言	253
第 11 章 课后习题	283
参考文献	305

第1章

多媒体：媒体和数据流

多媒体(Multimedium),从字面上分析,有以下两方面的含义:

Multi[词根:多]:很多

Medium[词根:中间,介质]:媒体,介质

这种解释是纯字面的,并不能说明多媒体的确切含义。

不同的公司有不同的认为:如:多媒体是文字、图像、动画、图像和音响的结合,而该公司的Mac机则是把它们结合起来的“胶水”;如:多媒体是传统的计算机媒体——文字、图像、图形、以及分析等与视频、音频以及为了知识创建和表达的交互应用的结合体。

有人把多媒体看成三种“革命”因素的总和;其一是在通信中结合了电视中的音像能力;其二是出版发行的能力;其三是计算机交互处理的能力。

多媒体的定义可以是:多媒体是指能够同时获取、处理、编辑、存储和展示两个以上不同类型信息媒体的技术。这些信息媒体包括文字、图像、图形、动画和活动图像等。

更一般的说,多媒体是指独立的机器可处理的以各种媒体表示的信息的生成、表示、处理、存储、检索和分布。这些信息是以多种依赖于时间或与时间无关的诸如数据、文本、声音、图表、图像、音响和电视等多种媒体来表示的。

下面结合计算机处理这一应用背景,来详细讨论各种媒体及其特征。

1.1 媒体有不同表现形式

在计算机中为了简化信息处理方式,便于人机交互,常按人的感觉来划分媒体。一般将媒体划分为视觉类媒体、听觉类媒体、触觉类媒体、活动媒体、抽象事实媒体和混合媒体等,其中视觉类媒体和听觉类媒体所承载的信息占人们获取信息的绝大部分。对于这些媒体,根据不同的抽象程度可分为若干层次,每一层次对应不同的表现形式。譬如,视觉类媒体中的图像可抽象为图形,图形又可抽象为文字,而听觉类媒体中的声音可抽象为音乐等。

1.1.1 视觉类媒体

人们主要是通过视觉来获取信息的。凡是通过视觉传递信息的媒体都属于视觉媒体。通常,视觉类媒体主要包括以下几类:

位图图像(bitmap):将所观察的图像按行列方式进行数字化,对图像的每一点(称为像素点)数字化为一个值(称为灰度值),所有这些值就组成了位图图像。这是一种对视觉信号进行直接量化的媒体形式,反映了信号的原始形式。根据量化的颜色深度的不同,又可分为二值和

灰度(彩色)图像两大类。位图图像是所有视觉表示方法的基础。目前在计算机上有关位图图像的存储、显示、变换等技术已比较成熟,在多媒体计算机(MPC)上得到了广泛的应用。

矢量图形:矢量图形一般也称为图形(graphics),它是对图像进行抽象化的结果,反映了图像中实体最重要的特征,如点、线、面等。它不直接描述图像数据的每一个点,而是描述产生这些点的过程和方法。抽象化(矢量化)过程可以由计算机自动进行,也可以由人工进行。在计算机中,矢量图形以一组指令的形式存在,而不用像位图那样对图像的每一个点进行量化存放,因此,它所需要的存储量很小。在计算机图形学上,图形一般分为二维图形和三维图形两大类。二维图形的基本单位称为图元,如一条曲线、一个矩阵、一个圆、一个填充的封闭区域、一个地图符号或电路符号、一个字符串等。若干个有联系的图元的集合就构成了一个图段。三维图形的显示和变换在理论和技术上都比较成熟,但如何对其再增加光照效果、质感等,使之逼近真实的图形效果,目前还在不断改进和深入研究中。

动态图像:这是利用人眼的视觉惰性,将若干连续的静态图像或图形在时间轴上不断变化所产生的结果。如果单帧图像是真实图像,则为动态影视视频(Video);若单帧图像是由计算机生成的真实感图像,则为三维真实感动画(Animation);如果在连续过程中变化的是图形,则是二维或三维动画。视频和动画序列都是由节段构成的,它们都是时基类媒体。与传统的影视视频和动画不同的是,在多媒体计算机中,这些动态图像不仅要存储和播放,而且要能进行变换、压缩、传输和编辑等处理。

符号:符号是人类对信息进行抽象的结果。符号可以表示数值,表示事物或事件,也可以表示语言和文字。文本是具有上下文相关特性的符号流。由于符号是人类创造出来表示某种含义的符号,所以它与使用者的知识有关,是比图形更高一级的抽象。必须具备特定的知识,才能解释特定的符号,才能解释特定的文本(如语言)。符号用特定值表示,例如 ASCII 码、中文图标码等。

1.1.2 听觉类媒体

凡是通过声音形式以听觉传递信息的媒体,都属于听觉类媒体。这类媒体是连续的时基类媒体,具有过程性和时间性,它的范围比视觉类媒体小得多,主要有以下几类:

波形声音(Wave Voice):所谓波形声音,实际上已经包含了所有的声音形式,因为任何声音都可以进行采样量化,并适当恢复出来。事实上,波形声音就是自然界中所有声音的“数字化声音”,它的获取是通过声音的数字化接口进行的。一般来说,对声音的处理主要集中在声音的数据压缩、数据编辑和声音效果的处理上。

语音(Voice):语音主要是指人类说话的各种声音。它与波形声音不同,因为人的说话声不仅表现为一种波形,而且还具有内在的语言、语音学内涵,可以经由特殊的方法提取,即进行一次抽象,所以常把它作为一种特殊的媒体。语音中包含了许多音节及其相互间过渡过程的连接体,而且语音也与具体的语言有关,如汉语中,一个字就是一个音节,字是一个独立的发音单位,而这些字又都是由 23 个声母、39 个韵母及 4 个音调变化组合而成,有 1000 多种不同发音。因此,了解并区分这些音节及相互关系,对于媒体转换,内容的处理都是有用的。

音乐(Music):音乐与语音相比形式就更为规范一些。事实上,音乐就是符号化了的声音,这种符号就是乐曲,乐谱则是转变为符号媒体形式的声音。就计算机媒体而言,基于乐器数字接口 MIDI(Musical Instrument Digital Interface)的音乐,就是十分规范的一种形式。与波形声音不同,MIDI 数据是指令集合,因此,MIDI 的数据量很小,对其进行编辑处理很灵活方便,

在音序器的帮助下,用户可自由地改变音调、音色等属性,直到达到理想的效果。

1.1.3 其他类媒体

其他类媒体是指除视觉类、听觉类媒体之外,在信息传递过程中许多其他种类的信息媒体(如触觉类媒体等)。有的使用其他感觉空间,如触觉、味觉和嗅觉等,有的则使用更为复杂的形式,如活动、自然规律、抽象数据等。从理论上说,凡是能够承载并且传递信息的都是信息媒体,但由于人类的感觉空间是有限的,所以任何媒体必须转变为可被人类接受的媒体形式才能被接受,故不好再用其他形式来描述那些承载信息的特殊媒体。现在多媒体系统中,已经把触觉媒体作为一种重要的媒体引入到实际系统中,有的甚至包括了嗅觉。特别是模拟类应用。

1.2 媒体之间的相对性质

空间和时间:多媒体的时间意义是指在表现上需要时间和媒体的时间属性;而多媒体信息的空间意义,第一种是指表现空间,这主要是指显示空间的安排。第二种是强调媒体之间相互的空间关系,这主要是指将多媒体信息在空间上进行有序地组织,这样有序的组织可以有助于反映信息的整体或全貌。

媒体的单位和量级:依照媒体的特性,各种媒体所采用的单位可划分为元素级、中间级、帧级和复合级,而不同的量级对应不同的单位。

同步和异步(SYNC & ASYN):所谓异步,是指时间上不能预知何时发生,没有严格的限制;而同步则要求在时间上必须满足各种媒体的实际需要,如电影中的像与音,口型与声音在时间上的同步,误差应小于150ms等。在多媒体信息处理中,媒体的同步和异步问题实质上就是媒体之间的协调关系问题,而且这种协调往往是必须考虑的关键问题。

1.2.1 多媒体的交互性

“交互”是人与信息系统交换信息的全过程,具有多层含义。对数据的交互是最低层次的交互形式,只是将数据进行交换转移。若交互以信息为主,如对媒体的选择、形式的变化、查询、命令和请求的转换以及用户本身的各种因素,则需要对数据进行解释,那就需要知识的辅助。如对每一种媒体进行程度不同的抽象,就是对媒体内容程度不同的理解。对对象信息的内容进行处理,以求达到更新层次的理解,这是更高层次的交互,目前尚在不断探索和研究中。

1.2.2 多媒体数据自身特点

数据量巨大:在多媒体环境下,有许多媒体形式(如声音、图像、视频等)的数据量是相当巨大的,如一幅中等分辨率的24位真彩色的位图图像(640×480 , 24位/像素)的数据量就有0.9MB。又如,数字音频磁带(DAT),采样精度16bit/样本,1秒钟时间内采样位数为 $48 \times 103 \times 16 = 768\text{KB}$,这样的数据量对于数据的处理、存储、传输都是个难题。因此,系统就必须能够适应这种数据的量级。

数据类型繁多:从媒体的种类来说,有文字、声音、图形、图像、动画、视频等多种形式。即使同属于图像一类,也还可以分为黑白、彩色、高分辨率等多种格式。这样,无论在媒体的输入上,还是在媒体的表现上,尤其是多媒体综合上,都会带来一系列的问题。同时,媒体的种类还在随着系统的不断进展而继续增多。

数据类型之间的差别大:这种差别首先体现在量上,有的媒体存储量很小,而有的媒体存储量却大得惊人。其次体现在内容上,不同类型的媒体由于格式、内容的不同,相应的类型管理、处理方法及内容的结实方法等也就不同,很难由某一种方法来统一处理这种差别。再次,这种类型的差别不仅仅体现在空间上,而且还体现在时间上,时基类媒体(如声音、动态影像视频等)的引入,与原先建立在空间数据基础上的信息组织方法会有很大的不同。

多媒体数据的输入输出复杂:在多媒体数据的输入过程中,有两种方式:一是多通道异步输入方式。这种方式允许在通道、时间都不相同的情况下输入各种媒体数据并存储,最后按合成效果在不同的设备上表现出来。二是多通道同步输入方式。这种方式要求系统具有多通道同时输入并分解媒体的能力,由于涉及的设备众多,多媒体数据的输入输出就要复杂得多。

1.3 多媒体系统主要特性

1.3.1 多媒体系统定义

如果根据字典中的解释,任何支持一种以上媒体的系统都可以称为多媒体系统,但这个定义并不全面,因为它只作出了一个量上的估计。比如,根据这种狭义的定义,一个系统,只要它同时处理了文本和图形,就可以将其看做多媒体系统,但这样的系统早在“多媒体”这个术语提出之前就存在了,因此“多媒体”这个词必然包含着计算机技术质的进步。

如果从质的角度判断一个系统是否多媒体系统,我们就会更看重系统所支持的媒体的类型而非数量。应该指出的是对于这种定义方式一直存在争论,因此即使在像 ISO 这样的标准化组织内,通常使用的也还是一个不很严格的定义。

多媒体系统区别于其他系统的特征有很多,我们着重说明的是其中一些最重要的特性,比如媒体的组合、媒体独立性及计算机控制与集成等。

1.3.2 媒体的组合

并不是任意的媒体组合都能满足术语“多媒体”的要求。例如,一般意义上可以把具有图像嵌入功能的文本处理程序称为多媒体应用,因为在一个程序中使用了两种媒体。但严格意义上的多媒体系统必须对这两种媒体都进行了处理,所以在这个角度上具有图像嵌入功能的文字处理程序不应被看成多媒体应用。

1.3.3 独立性

媒体的一个重要性质是它们之间的独立性。一般的应用也需要各种媒体具有一定的独立性,但多媒体系统需要的是多个层次的独立性。比如,一方面,计算机控制的视频记录仪存储的视频和音频信息是紧密耦合在共同的存储介质上的;但另一方面,又可能要将 DAT(Digital Audio Tape)声音信号取出并与文本等其他媒体组合起来以满足一定的表示需要。

1.3.4 计算机支持的媒体集成特性

媒体的独立性使得各种媒体的组合成为可能。计算机是进行媒体组合的理想工具。媒体组合系统必须具有媒体处理能力,同时系统也应具备二次开发的功能。一个系统中对不同媒体的简单输入输出(比如摄像机)并不是计算机控制的媒体集成,计算机集成是将独立的媒体数据集成起来以完成一定的任务。为了达到这个目的,必须引入时间、空间和语意同步信息。一

个支持文字、表、视频片段的文字处理程序如果不支持这些数据间的连接就不满足集成的要求,相反,如果对表的内容改动能引起对应的视频场景和文本的变化,我们就认为它具有较高的集成度。

即使在许多非常好的多媒体产品中,像上面一样对媒体的灵活操作也不多见。因此,我们用“集成多媒体系统”对这一方面的功能加以强调。简言之,在这样的系统中,每一种用来表示文本和图形的方法都可用于视频和音频。例如在传统系统中,可以向其他用户发送文字信息,那么在具有高度集成性的系统中就应该允许发送音频或音频文字混合信息。

以计算机为基础的多媒体系统第二个重要特征在于其集成化的概念。多媒体系统不仅有能力拥有各种手段以产生、存储、传递和显示信息,而且多媒体倾向于尽可能地用集成化的方法来完成这些功能。在分布式多媒体中,集成化的概念涉及两个方面:计算机手段的集成和网络手段的集成。

1.3.5 数字化特性

数字化描述的主要优点在于其描述的普遍性。由于任何媒体,无论是文本、图像或声音都以惟一的形式编码,直至最终产生一个二进制位串序列。因而各种类型的信息都能以同样的方式,由同一类型的设备处理、存储和传输。

1.3.6 通信系统

多媒体系统必须具备通信能力。因为目前大多数的计算机都是网络互联的,如果只从本地处理的角度看待多媒体系统将带来很大的局限性。另一个原因是很多有趣的多媒体应用是专为分布式环境而设计的,单个的计算机不足以产生、处理、表示和存储这些媒体信息。

1.4 媒体

一般说来,媒体是一种信息发布和表现的方法(如报纸、电视、杂志等),但从计算机处理的角度讲,媒体的含义要更进一步,指信息表示的具体方法,如文字、图形、语音和音乐等。

媒体可按不同的标准分类[ISO 93a],我们把媒体分为感知媒体、表示媒体、表现媒体、存储媒体、传输媒体和信息交换媒体。

1.4.1 感知媒体

感知媒体帮助人来感知环境,要解决的主要问题是:人在计算机环境中如何感知信息。就目前而言,人类主要靠视觉和听觉来感知环境的信息,触觉作为一种感知方式也慢慢引入到计算机系统中。

在使用计算机时,聆听信息和观看信息是有区别的。在通过视觉来感知信息时,可视媒体是文本、图像和视频;在通过听觉来感知信息时,可听媒体有音乐、噪声和语音。

媒体的差别还可以进一步区分,例如,视频可分解成不同的视频场景,而视频场景又由单个图像组成。

1.4.2 表示媒体

表示媒体的特征由信息的计算机内部表示刻画,要解决的主要问题是:计算机信息如何编

码。在计算机中使用不同的格式来表示媒体信息,例如:

- 文本字符用 ASCII 或 EBCDIC 码来表示
- 图形根据 CEPT 或 CAPTAIN 视频正文标准编码,图形标准 GKS 也用作编码基础
- 音频流可以使用简单的 PCM 来表示,每个样本采用 16 位线性量化值表示
- 图像可以用 JPEG 格式或其他格式编码
- 组合音频/视频序列可以用不同的 TV 标准格式(PAL, SECAM, NTSC)编码,并用 MPEG 格式保存在计算机中。

1.4.3 表现媒体

表现媒体是指信息输入/输出的工具和设备,要解决的问题是:信息通过何种媒体输入到计算机中或从计算机中输出。输出媒体的设备有纸、屏幕和喇叭,而输入媒体的设备有键盘、鼠标、摄像机和麦克风等。

1.4.4 存储媒体

存储媒体是指能支持信息存储的数据载体,数据存储并不局限于计算机的组件,因此,纸也是一种存储媒体。要解决的问题是:信息存储在什么地方。存储媒体的设备有微型胶片、软盘、硬盘和 CD-ROM 等。

1.4.5 传输媒体

传输媒体用于传输数据信息,它们支持连续媒体数据传输。要解决的问题是:在什么上面传输信息。目前主要使用网络来进行信息传输,网络介质有同轴电缆、光纤等,也可以使用无线方式进行通信。

1.4.6 信息交换媒体

信息交换媒体包括所有用于传输和存储的信息载体。要回答的主要问题是:在不同地点之间使用哪种信息载体来进行信息交换。这个问题有三个答案:信息通过中间存储媒体流动;使用计算机网络进行直接传输;组合使用存储媒体和传输媒体(如电子邮件系统)。

1.4.7 表示值和表示空间

上面的媒体分类方法可用作描述信息处理领域中媒体的特征,这里所描述的感知媒体与我们要讨论的媒体的含义类似,每种媒体定义了表示值和表示空间。

可视表示空间是纸和计算机屏幕,在用计算机来控制幻灯片的放映,并对计算机屏幕的内容进行投影时,整个电影屏幕就是一个表示空间。立体声和四轨录音/放音确定声音表示空间,表示空间也能被当成是上述用于信息输出的表现媒体的一部分。

表示值决定不同媒体的信息表示:文本媒体以可视方式表示由一串字符组成的句子,这个句子也可以用语音媒体表示。某些表示值是自包含的,也就是说,它们可以由接收者作适当解释,这样的例子有温度、味道和气味。其他媒体需要一预定义符号集,用户必须遵从这样的符号集定义。文本、语音和手势就是这种媒体的例子。

表示值可以是离散型的,也可以是连续型的。声音波形变化确定了声音信号的特征,它并不以离散型值出现。同样,电磁波也不是离散型的值,它们是连续型值。文本字符和音频样本