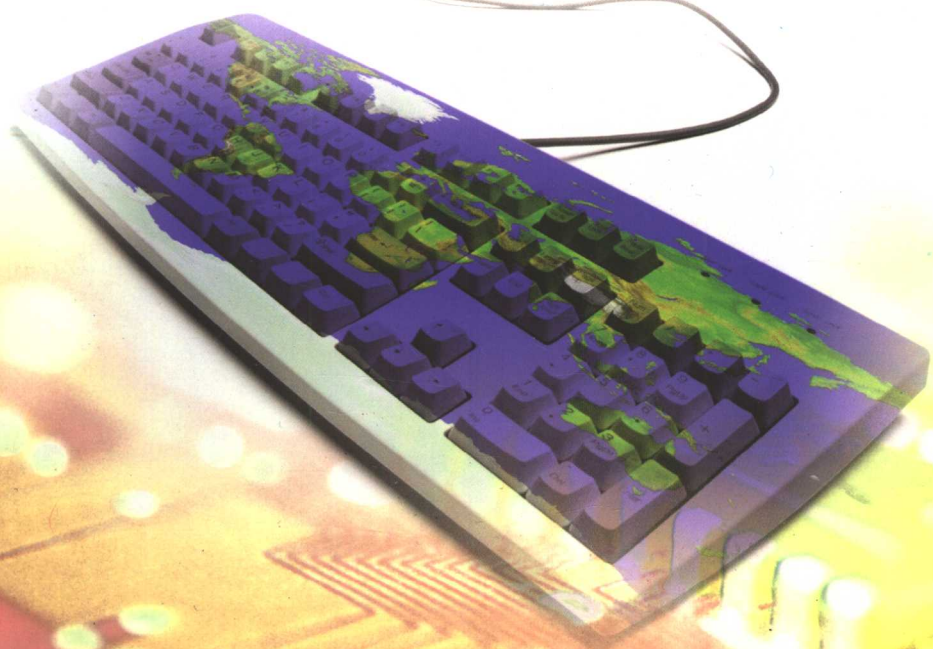


21 世纪计算机课程系列教材

多媒体

技术概论

主 编 朱范德



东南大学出版社

21 世纪计算机课程系列教材

多媒体技术概论

主 编 朱范德
副主编 李 峰 葛桂萍

东南大学出版社

内 容 提 要

本书较全面地介绍了多媒体技术主要的基本概念和相关技术,内容包括多媒体技术的基本概念、音视频数据的压缩编码技术、光存储媒体、MPC 及常见设备、多媒体通信系统、多媒体技术应用实例、多媒体开发应用技术和多媒体技术发展的热门新技术。

本书可作为计算机、电子、电气、通信和信息管理专业本科生、专科生或高职高专的教材,也可以供从事多媒体计算机技术研制、开发和应用的工程技术人员参考使用。

图书在版编目(CIP)数据

多媒体技术概论 / 朱范德主编. —南京:东南大学出版社,2006.2

(21 世纪计算机课程系列教材)

ISBN 7-5641-0182-2

I. 多... II. 朱... 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(2006)第 005047 号

多媒体技术概论

出版发行 东南大学出版社

出版人 宋增民

社 址 南京市四牌楼 2 号

邮 编 210096

电 话 025—83792954

电 邮 zhu_min_seu@163.com

经 销 江苏省新华书店

印 刷 溧阳市晨明印刷有限公司

开 本 787 mm×1092 mm 1/16

印 张 20.25

字 数 505 千字

版 次 2006 年 2 月第 1 版

印 次 2006 年 2 月第 1 次印刷

印 数 1—4000 册

定 价 32.00 元

(凡因印装质量问题,请与我社读者服务部联系。电话:025—83792328)

前 言

多媒体技术的出现是人类社会信息技术的又一次重要的飞跃。多媒体技术集成了多种技术,同时又是发展最快的技术之一。随着人类对信息处理和利用需求的不断增长,对信息形式和质量要求的不断提高,多媒体技术的应用也是突飞猛进。在越来越多的场合已经离不开多媒体技术的支持。学习和掌握多媒体技术知识对于现代科技人员和相关专业的学生显得越来越重要。

我们于1995年开设了“多媒体技术概论”课程,作为计算机专业学生的选修课,先后又在电子、电气、信息管理等专业作为选修课。在十年的教学过程中我们使用了多本教材,给了我们很大的帮助和启发。随着多媒体技术不断的快速发展以及社会对人才要求的变化,特别是不断更新的教学计划与本课程内容的衔接,促使我们开始编写本教材。为此,我们收集了国内外技术杂志、会议论文、产品手册、国际标准等,使用了十年经过不断修改的备课笔记作为本书的基本材料。在内容的选取上注意基本概念、最新技术发展和实践操作兼顾。对一些关键技术理论上进行了深入讨论,对相关技术发展的方向进行了详细的介绍,利用大量的实例介绍了相关技术的应用和实现。

本教材共分10章,其中朱范德编写了第1章多媒体技术概论、第3章图像信号处理技术、第4章图像和视频信号压缩编码技术与相关国际标准和第8章多媒体通信技术;葛桂萍编写了第2章音频信号处理技术、第5章光存储媒体技术和第9章多媒体技术的应用;李峰编写了第6章多媒体计算机系统及常见硬件设备、第7章多媒体应用系统开发技术和第10章多媒体新技术展望。从内容上看本教材包括了多媒体技术中几乎所有的基本概念和成熟的技术,同时,在第10章中简要阐述了数据压缩新技术、虚拟现实技术、智能交互技术、MPEG-21标准现实技术和信息高速公路等最新技术的内容。

本教材在编写过程中参考了有关教材、文献和一些网站的内容,并引用了部分材料,在此我们对这些作者表示衷心的感谢!同时向关心和帮助我们的同事和家人表示衷心的感谢!

由于多媒体技术涉及面太广,发展较快,虽然我们尽了自己最大的努力,难免会出现一些问题,在此谨向读者表示歉意,并希望各位指正,帮助我们修改再版。

编 者

2005年12月30日

目 录

1	多媒体技术概论	(1)
1.1	引言	(1)
1.2	多媒体技术	(1)
1.2.1	多媒体信息	(1)
1.2.2	多媒体信息处理	(3)
1.2.3	多媒体信息系统	(3)
1.2.4	多媒体技术的特点	(3)
1.3	多媒体系统的发展	(3)
1.4	多媒体技术研究的主要内容	(4)
1.4.1	数字化技术	(5)
1.4.2	数据压缩编码技术	(5)
1.4.3	计算机平台和硬件支持芯片	(8)
1.4.4	实时多任务操作系统	(8)
1.4.5	高速通信网络	(9)
1.4.6	多媒体著作系统	(11)
1.4.7	超文本和超媒体技术	(13)
1.4.8	多媒体信息的组织和管理技术	(15)
1.4.9	多媒体分布应用技术	(17)
1.4.10	标准化	(18)
1.5	多媒体技术的应用领域	(19)
1.5.1	多媒体技术在教育中的应用	(19)
1.5.2	多媒体技术在商业中的应用	(21)
1.5.3	多媒体技术在通信中的应用	(22)
1.5.4	电子出版物	(23)
1.5.5	家庭娱乐	(23)
1.5.6	办公自动化	(23)
1.5.7	国防和军事领域	(24)
1.6	多媒体计算机的发展趋势	(24)
2	音频信号处理技术	(26)
2.1	音频处理基础	(26)
2.1.1	声音的特性及分类	(26)
2.1.2	声音信息的数字化	(27)

2.1.3	数字音频的文件格式	(29)
2.2	声卡	(31)
2.2.1	声卡的发展历史	(31)
2.2.2	声卡的声道	(33)
2.2.3	声卡的功能和分类	(34)
2.2.4	声卡的工作原理	(34)
2.2.5	声卡的外接插口	(35)
2.3	音频压缩技术和标准	(36)
2.3.1	概述	(36)
2.3.2	声音压缩标准	(37)
2.4	MIDI	(40)
2.4.1	MIDI 标准的内容	(40)
2.4.2	MIDI 标准的优点	(40)
2.4.3	产生 MIDI 音乐的方法	(40)
2.4.4	MIDI 系统	(41)
2.5	语音识别技术	(42)
2.5.1	语音识别的发展历史	(42)
2.5.2	语音识别的分类	(43)
2.5.3	语音识别的工作原理	(44)
2.5.4	语音识别的困难与对策	(45)
2.5.5	语音识别的前景和应用	(46)
3	图像信号处理技术	(48)
3.1	数字图像基础	(48)
3.1.1	视觉系统对颜色的感知	(48)
3.1.2	图像的颜色模型	(49)
3.1.3	彩色空间的线性变换标准	(52)
3.1.4	图像的基本属性	(54)
3.2	图像的分类和格式	(57)
3.2.1	图像的分类	(57)
3.2.2	常用图像文件格式	(59)
3.2.3	图像处理中的常用名词	(62)
3.3	图像输入/输出设备	(62)
3.3.1	笔输入	(62)
3.3.2	触摸屏	(64)
3.3.3	扫描仪	(66)
3.3.4	数码相机	(67)
3.3.5	虚拟现实的三维交互工具	(70)
3.4	动态图像输入设备	(71)

3.4.1	图像捕捉卡	(71)
3.4.2	摄像头和摄像机	(72)
4	图像和视频信号压缩编码技术与相关国际标准	(74)
4.1	引言	(74)
4.1.1	压缩的重要性和可行性	(74)
4.1.2	数据压缩技术的分类	(76)
4.1.3	常用编码方法介绍	(76)
4.2	静止图像压缩编码标准	(85)
4.2.1	基于 DCT 的编码器框图	(85)
4.2.2	JPEG 编码算法和实现	(85)
4.3	H.261 标准	(89)
4.3.1	概述	(89)
4.3.2	图像格式	(89)
4.3.3	H.261 编码器框图	(90)
4.3.4	信源编码	(90)
4.3.5	视频图像复接编码器	(93)
4.3.6	信道编码器	(94)
4.3.7	其他 H 系列的视频标准	(94)
4.4	MPEG-1 标准	(96)
4.4.1	MPEG 简介	(96)
4.4.2	MPEG-1 视频编/解码器	(97)
4.4.3	MPEG-1 图像复接编码器	(99)
4.5	通用活动图像编码标准 MPEG-2	(101)
4.5.1	概述	(101)
4.5.2	系统部分	(101)
4.5.3	MPEG-2 的编解码器	(104)
4.6	MPEG-4 视频	(105)
4.6.1	概述	(105)
4.6.2	视频编码	(107)
4.6.3	视频轮廓	(113)
4.6.4	MPEG-4 文件格式	(114)
4.7	MPEG-7 多媒体内容描述接口	(115)
4.8	其他压缩编码标准	(117)
4.8.1	M-JPEG	(117)
4.8.2	MPEG-21 标准	(118)
4.8.3	其他压缩编码标准	(119)

5	光存储媒体技术	(120)
5.1	概述	(120)
5.1.1	光盘的发展简史	(120)
5.1.2	光盘的特点	(122)
5.1.3	光盘的分类	(122)
5.1.4	光盘系统的主要参数	(123)
5.2	光盘的标准	(124)
5.2.1	红皮书——CD-DA	(124)
5.2.2	黄皮书——CD-ROM	(126)
5.2.3	ISO 9660	(128)
5.2.4	绿皮书——CD-I	(129)
5.2.5	橙皮书——CD-R	(130)
5.2.6	白皮书——Video CD	(131)
5.3	CD-ROM 系统	(133)
5.3.1	CD-ROM 光盘结构	(133)
5.3.2	CD-ROM 扇区的数据结构	(136)
5.3.3	CD-ROM 光驱	(137)
5.4	CD-R 和 CD-R/W 光盘	(139)
5.4.1	CD-R 光盘	(139)
5.4.2	CD-R/W 光盘	(140)
5.5	VCD 与 DVD	(143)
5.5.1	VCD	(143)
5.5.2	DVD 的分类	(144)
5.5.3	DVD 的特点	(148)
5.5.4	DVD 为增大存储容量采取的措施	(149)
5.5.5	DVD 播放机	(151)
5.5.6	DVCD 光盘技术	(152)
6	多媒体计算机系统及常见硬件设备	(154)
6.1	多媒体计算机系统	(154)
6.1.1	多媒体个人计算机	(155)
6.1.2	专用多媒体系统	(156)
6.1.3	多媒体工作站系统	(156)
6.2	输入设备	(157)
6.2.1	键盘	(157)
6.2.2	鼠标	(158)
6.2.3	手写输入设备	(160)
6.2.4	触摸屏	(161)

6.2.5	条形码	(163)
6.2.6	读卡器、磁卡与 IC 卡	(165)
6.2.7	光学字符识别	(166)
6.2.8	语音输入系统	(167)
6.2.9	数字摄像头	(169)
6.2.10	其他输入设备	(171)
6.3	输出设备	(171)
6.3.1	显示器	(171)
6.3.2	投影机	(172)
6.3.3	打印设备	(174)
6.4	通信设备	(176)
6.4.1	调制解调器	(176)
6.4.2	ISDN 和 DSL	(177)
6.4.3	PC 传真卡	(179)
6.5	多媒体 I/O 总线和接口标准	(179)
6.5.1	计算机传输总线	(179)
6.5.2	SCSI 接口标准	(180)
6.5.3	USB 串行总线接口标准	(180)
6.5.4	IEEE1394 高速串行总线接口标准	(181)
7	多媒体应用系统开发技术	(182)
7.1	多媒体应用系统的开发过程	(182)
7.2	多媒体素材制作与加工	(184)
7.2.1	文字素材的获取	(185)
7.2.2	音频数据的获取与加工	(186)
7.2.3	图形与图像素材的获取与加工	(190)
7.2.4	视频素材的获取与加工	(194)
7.2.5	动画的制作	(198)
7.3	多媒体编著工具	(202)
7.3.1	多媒体编著工具的功能	(202)
7.3.2	多媒体编著工具的创作模式	(202)
7.3.3	多媒体编著工具的类型	(203)
7.3.4	基于编著工具的多媒体应用制作实例	(204)
7.4	Windows 多媒体程序设计	(212)
7.4.1	Windows MCI 编程接口介绍	(212)
7.4.2	基于多媒体组件的多媒体应用编程	(215)

8	多媒体通信技术	(227)
8.1	概述	(227)
8.1.1	多媒体网络应用	(227)
8.1.2	应用分类	(228)
8.1.3	应用开发面临的问题	(228)
8.1.4	改善服务质量	(229)
8.1.5	多媒体网络应用的发展	(229)
8.2	多媒体通信基础	(229)
8.2.1	多媒体通信的特点和对网络的要求	(229)
8.2.2	当前网络对多媒体通信的支持	(232)
8.3	多媒体应用的网络需求	(233)
8.3.1	音频信息的网络需求	(234)
8.3.2	视频信息的网络需求	(238)
8.3.3	音频和视频网络传送的需求	(242)
8.3.4	其他需求	(243)
8.4	网络多媒体应用系统和相关标准	(244)
8.4.1	概述	(244)
8.4.2	计算机支持的协同工作(CSCW)	(245)
8.4.3	点播服务系统	(247)
8.4.4	相关标准	(249)
9	多媒体技术的应用	(253)
9.1	典型的网络多媒体应用系统	(253)
9.1.1	可视电话	(253)
9.1.2	视频会议系统	(255)
9.1.3	多媒体远程教学系统	(260)
9.1.4	多媒体远程医疗系统	(263)
9.1.5	视频点播系统	(265)
9.1.6	IP电话	(269)
9.2	超文本和超媒体	(272)
9.2.1	超文本的发展历史	(272)
9.2.2	超文本与超媒体的定义	(273)
9.2.3	超文本与超媒体的体系结构	(276)
9.2.4	超媒体的应用	(280)
9.2.5	超文本、超媒体发展前景	(281)
9.3	多媒体数据库系统	(282)
9.3.1	多媒体数据库的基本概念	(282)
9.3.2	多媒体数据库管理系统的基本功能与体系结构	(285)

9.3.3 多媒体数据库的基于内容的检索	(288)
10 多媒体新技术展望	(292)
10.1 数据压缩新技术	(292)
10.1.1 矢量量化编码	(292)
10.1.2 结构编码	(292)
10.1.3 图像编码	(293)
10.1.4 小波变换编码	(293)
10.1.5 基于模型的编码	(294)
10.1.6 分形编码	(294)
10.2 虚拟现实技术	(295)
10.2.1 概述	(295)
10.2.2 虚拟现实系统的分类	(296)
10.2.3 虚拟现实技术的应用	(296)
10.2.4 未来的发展趋势	(297)
10.3 智能交互技术	(298)
10.3.1 人机交互技术概述	(298)
10.3.2 智能交互技术的进展	(299)
10.4 MPEG-21 标准现实技术	(302)
10.5 信息高速公路及其影响	(305)
10.5.1 概述	(305)
10.5.2 信息高速公路对社会的影响	(306)
参考文献	(310)



多媒体技术概论

1.1 引言

多媒体技术(Multimedia Technology)是20世纪80年代末期兴起的一门新技术。简单地说,多媒体技术就是利用计算机综合处理多种媒体信息(数据、文本、图形、声音、图像和视频),使其建立逻辑连接,集成为一个系统并具有交互性。多媒体技术最先出现于计算机领域,随着信息技术、通信技术、超大规模集成电路(VLSI)技术、网络技术的发展以及多媒体技术应用领域的不断开拓,如今多媒体技术不仅是计算机领域的热点之一,而且也是通信技术、信息技术等领域的热门课题。多媒体技术必将渗透到与信息有关的各个领域,可以说90年代是以应用多媒体技术(多媒体信息、多媒体信息处理、多媒体信息系统)为特征的信息时代。多媒体技术不仅引起了计算机产业的第二次革命,而且把通信技术推向崭新的全方位的多媒体通信时代,并改变着人们的工作、学习和生活,改变着90年代人类社会面貌。当前人类正向“多媒体社会”迈进。

1.2 多媒体技术

1.2.1 多媒体信息

人类进行信息交流最常用的手段是通过视觉和听觉进行的。视觉信息和听觉信息约占人类获取信息总量的85%。视觉信息包括文字、图形、图像、视频、动画等,视觉信息是最直接最生动的信息。听觉信息包括语言、音乐等。我们把数据、文本、声音、图像等承载信息的载体称为媒体(Media)。信息都是依附于一定载体上才能够进行传播、存储、接收、处理和使用的。根据国际电报电话咨询委员会 CCITT 定义,媒体有以下五种:

1) 感觉媒体(Perception Medium)

感觉媒体是直接作用于人的感觉器官,能使人直接产生感觉的媒体。感觉媒体有语音、音乐、数据、文字、图形、图像、视频、动画、自然界的各种声响等。

2) 表示媒体(Representation Medium)

表示媒体是为了传输感觉媒体而人为研究出来的媒体,以便能更有效地将感觉媒体从一地传向另一地。表示媒体有各种语音编码、音乐编码、文本编码、图形编码、图像(静止图像、运动图像)编码等。

3) 表现媒体(Presentation Medium)

表现媒体指的是电信号和感觉媒体之间转换用的媒体。表现媒体有两种,一种是输入表

现媒体,它将感觉媒体转换为电信号,如键盘、鼠标器、摄像机、扫描仪、光笔、话筒等;另一种是输出表现媒体,它将电信号转换为感觉媒体,如喇叭、打印机、显示器等。

4) 存储媒体(Storage Medium)

存储媒体用于存放表示媒体,以便随时调用。存储媒体有硬盘、软盘、光盘、磁带、存储器等。

5) 传输媒体(Transmission Medium)

传输媒体是用来将表示媒体从一处传送到另一处的物理实体,传输媒体有双绞线、电话线、同轴电缆、光纤、无线电链路等。

当信息载体不仅是数据、文字,还包括图形、图像和声音等多种感觉媒体时,我们就称为多媒体信息,或简称多媒体。因此,多媒体信息不仅包含了数据、文字,还包括各种视觉信息和听觉信息,是全方位信息。

多媒体信息具有以下属性:

(1) 模拟形式和数字形式兼有

真实世界的视觉信息和听觉信息是模拟形式的,作用于人眼睛和耳朵的信息也是模拟的,而计算机处理和集成、存储和传输的则是数字化的信息。

(2) 静止的和运动的属性

多媒体信息可以是静止的也可以是运动的。文字、图形、静止图像都属于静止信息形式,而视频、广播、动画、语音、音乐则属于运动信息。在人类的信息交互过程中,运动信息占有很大比重。运动信息能以一种带有过程的形式表达出特别的内容,除了信息本身是运动的外,运动的含义还包括过程本身,我们称包括过程的运动为活动。活动包括学习过程、表达和变换过程。学习过程是接受信息的过程,变换过程就是通过交互掌握和利用信息的过程。所以,多媒体信息本身的活动性就是重要的信息交互过程。

(3) 空间和时间属性

多媒体信息在空间和时间上都有意义,例如文本、静止图像需要一定的显示空间,声音则需要表现时间,而空间和时间属性又往往结合在一起。例如动画是图形加运动,视频是图像、声音加运动,它们既需要显示空间,在时间上也需要延续,活动和表达也是如此。

(4) 分散性和集成性属性

多媒体数据作为数字形式的多媒体信息具有分散性,这是指各种媒体之间没有特定的统一形式,很难找到一种能把所有不同媒体信息组合在一起的统一的数据结构。因为各种媒体的数据(字符数据、语音数据、图形数据、图像数据、视频数据、动画数据等)在形式上、数量上和处理方法上都有很大的区别。所以,对多媒体数据的处理(包括存、取)不能像对其他数据那样用单一的事先指定的方法,而必须根据不同的媒体采用不同的方法。

多媒体信息的集成性是指信息表现时往往需要将多种媒体信息相互配合集成在一起。例如将图形、文字和图像集成在同一画面上,将图像、动画与解说、音乐相配合。

(5) 同步和异步的属性

同步和异步是多媒体信息的两个重要属性。异步是指时间上不能预知何时发生的,需要特别协议的信息,例如电子邮件等。同步则是指同时的、多通道的、实时的。例如多媒体信息中,图像和声音的同时出现。然而,同步和异步既有差别又可以统一,例如在一个通道中只能

异步传送声音和画面,但当数据到达目的地时,则可以由同步方式再现出来。

1.2.2 多媒体信息处理

对多媒体数据的处理除了有查找、检索、排序外,还有压缩、转换、识别、理解、合成、存储、传输和利用等特殊操作。

压缩是多媒体信息处理的关键,可以说没有实时压缩能力就没有多媒体信息,也不可能实现多媒体信息的存储和传输。

各种媒体信息间的转换是一个重要的处理能力。例如将文字变为语音或将语音识别为文字,都涉及到语音合成与识别问题;文字和图形、图像之间的转换则涉及到字符识别、图形理解等问题。目前有些媒体信息间可以互换,有的还不能互换。随着人工智能的发展,将会实现各种媒体信息间的随意互换,那时,多媒体技术就进入了一个新的智能化阶段。

合成和创作是多媒体信息处理的又一重要方面,它体现了多媒体信息的集成性和交互性。

多媒体信息的存储和传输也是多媒体信息处理不可缺少的方面。

此外,还涉及到各种媒体信息本身的处理技术,如语音处理技术,图形、图像处理技术,数据库技术等。

1.2.3 多媒体信息系统

多媒体信息系统就是将先进的计算机系统、通信系统和广播系统统一成一个综合化的多媒体信息服务系统。也就是说多媒体信息系统是集多媒体信息的多样性,计算机的交互性,通信的分布性和广播的真实性于一体的全方位信息服务系统。它的发展和完善对人类社会将产生重要影响。

1.2.4 多媒体技术的特点

多媒体信息技术就是利用多媒体信息系统处理和综合利用多媒体信息的技术。它具有以下几种特性:

(1) 多样性:包括媒体的多样性,数据格式的多样性,输入/输出设备的多样性,服务方式的多样性。

(2) 集成性:包括两种以上媒体信息的集成,数据种类的集成,表现方式和传播方式的集成,硬件系统和软件系统的集成。

(3) 交互性:是指用户能自主地控制和干预多媒体信息的处理、制作和利用的全过程。

1.3 多媒体系统的发展

人们对多媒体信息并不陌生,目前千家万户使用的电视就是接收由声、文、图等多种感觉媒体信息组成的系统。但是,这种家用电视是模拟的,只能被动地接收,不能实现交互式的收看,因而不能称之为多媒体系统。随着数字化技术、数据压缩技术、存储技术和计算机技术的发展,使具有交互功能的多媒体系统有了实现的基础。

1983年美国 RCA 公司的戴维·沙诺夫研究中心首先提出了数字化视频交互技术(Digital Video Interactive,DVI)的设想。第一代 DVI 产品 Action media 750-I 于 1989 年推

出,第二代产品 Action media - II 于 1991 年问世。

世界上第一个多媒体系统 Amiga 是 Commodore 公司于 1985 年推出的,该系统具有音响、视频信号处理和动画显示功能。1986 年 Philips 公司和 Sony 公司公布了基本的 CD - I (Compact Disc - Interactive) 系统,这是较早的交互式数字多媒体系统。但是直到 1991 年秋才出现商业产品。

Apple 公司的 Macintosh 具有良好的人机界面,使得它成为桌面印刷和演示系统的先驱。1991 年推出的 Mac 操作系统 7.0 增加了多媒体功能,同时 Macintosh 拥有大批功能卓越的多媒体应用软件。

1988 年 10 月第一种从设计上贯彻多媒体思想的 Next 计算机问世,它在硬件、软件设计上都考虑了如何适应多媒体的思想。

随着人们对音频、视频数据压缩编码技术的深入研究,相继建立了图像数据压缩编码的各种国际标准:JPEG 标准(静止图像压缩编码标准)、H. 261(P×64)标准(可视电话、会议电视压缩编码标准)、MPEG 标准(活动图像压缩编码标准),使得多媒体技术中数据压缩这一关键问题的解决有了统一可行的国际标准。与此同时,许多 VLSI 制造公司推出了能实时实现这些标准算法的专用芯片和通用芯片,加上个人计算机性能价格比的不断提高,于是出现了许多多媒体板级产品,如多媒体信息采集卡、数据压缩编码卡,以及以 PC 机为平台配置的各种多媒体板级卡而形成的各种多媒体个人计算机(MPC)和多媒体系统等。目前形式多样的多媒体系统如雨后春笋般地涌现,根据功能不同大体上可以分为以下几类:

1) 开发系统

具有多媒体技术应用的开发功能,该系统配有功能强大的计算机,齐全的声、文、图信息的外部设备和多媒体演示的编著工具,典型的用户是多媒体技术应用的制作、电视编辑系统等。

2) 演示系统

是一个增强型的桌上系统,可完成各种多媒体的应用并与网络连接。典型的用户是专业技术工作者、大公司经理、高等教育工作者等。

3) 训练/教育系统

是一种用户多媒体信息播放系统,以计算机为基础配上 CD - ROM 驱动器、音响和图像接口控制卡连同相应的外部设备,通常用于家庭教育、小型商业销售点和教育培训等。

4) 家用系统

是一种家庭信息亭,通常配有 CD - ROM,采用家用电视机作为显示设备,可供 5 名以下观众使用。通常用于家庭学习、娱乐、一般信息的存储和处理。

1.4 多媒体技术研究的主要内容

多媒体技术是一门综合技术,它涉及到计算机技术、信号处理技术、通信技术、压缩编码技术、计算机平台和硬件支持芯片技术、实时多任务操作系统、窗口管理系统技术、高速信息网络技术、编著工具系统以及超媒体技术等。因此多媒体技术的研究涉及面非常广,在此我们只介绍主要的内容。

1.4.1 数字化技术

由于多媒体技术要利用计算机来综合处理文字、声音、图形、图像、视频等多种媒体信息,这些信息本身都是模拟量,只有数字化以后才能由计算机平台进行各种处理和综合。因此,数字化技术是多媒体技术的必要基础。对于不同媒体,信息数字化要求和实现方法均有所不同。

音频信号除 CD 音响和电子乐器已是数字信号外,现有的语音、广播(调幅、调频)和立体声音乐均是模拟信号,一般需经滤波器和模/数(A/D)转换器将上述各种模拟音频信号转换为数字信号。视频信号通常是由摄像机、录像机等视频图像输入设备获得模拟图像。这些信号大多数是标准的彩色全电视信号,必须经彩色解码电路将全电视信号分解为模拟彩色分量信号——R、C、B(或 Y、U、V)信号,再经 A/D 转换器转换为数字式信号。各种媒体信息的数字化通常是由各种多媒体信息的采集卡(图形卡、图像卡、音频卡、视频卡等)来实现的,它集中体现了多媒体信息的数字化技术,其主要指标是采样速度、精度和功能。

1.4.2 数据压缩编码技术

1) 进行数据压缩的必要性

为了能在计算机上实现对多媒体信息的交互处理,就必须对各种媒体信息进行数字化。而数字化信息的数据量是十分庞大的,1 秒钟的 CD 质量的立体声数字化音频信息,其数据量为 $1.411 \text{ Mbit}(2 \times 44.1 \text{ kHz} \times 16 \text{ bit})$,40 MB 存储媒体仅能存放大约 4 分钟的立体声声音,一张 650 MB 的光盘也只能存放 1 小时的立体声音乐,如以 NTSC 制播放 640×480 全彩色数字视频,则数据的传输速率为 $221 \text{ Mb/s}(640 \times 480 \times 24 \text{ bit} \times 30/\text{s})$,一张 650 MB 光盘仅能存储 20 秒左右的数字视频。由此可见,数字音频和视频庞大的数据量不仅造成存储和传输的困难,而且计算机的总线也难以承受。尽管有各种不同方法在不同程度上提高计算机的传输能力,但都不能彻底解决问题。彻底解决问题的方法是对多媒体信息数据进行压缩。一幅 $512 \times 512 \times 8 \text{ bit}$ 的静止图片在电话线上以 2.4 Kb/s 速率传送约需 15 分钟,若将数据压缩 15 倍,则传送时间降为 1 分钟。又如未压缩的 NTSC 制数字电视信号码率为 220 Mb/s ,若采用压缩比达 200 以上的数据压缩技术,则码率可降到 1 Mb/s 以下。可见数据压缩编码技术在多媒体技术中的重要作用。

在表 1.1 和表 1.2 中分别列出了 1 分钟数字声音信号和视频信号的有关参数:

表 1.1 1 分钟数字声音信号所需的存储空间

数字音频格式	频带(Hz)	带宽(kHz)	取样率 (kHz)	量化位数	存储容量(MB)
电 话	200~3 400	3.2	8	8	0.48
会议电视伴音	50~7 000	7	16	14	1.68
CD-DA	20~20 000	20	44.1	16	5.292×2
数字音频广播	20~20 000	20	48	16	5.76×2

表 1.2 1 分钟数字视频信号所需的存储空间

数字电视格式	空间×时间分辨率(帧/s)	取样率(MHz)	量化位数	存储容量(MB)
CIF	352 × 288 × 30	亮度 4:1:1	12	270
CCIR 601 建议	NTSC 制 720 × 480 × 30	亮度 4:2:2	16	1 620
	PAL 制 720 × 576 × 25			1 620
HDTV	1 280 × 720 × 60	60	8	3 600

2) 数据压缩的可能性

研究发现,图像数据表示中存在着大量的冗余。通过去除这些冗余数据可以使原始图像数据极大地减少,从而解决图像数据量巨大的问题。图像数据压缩技术就是研究如何利用图像数据的冗余性来减少图像数据量的方法。因此,进行图像压缩研究的起点是研究图像数据的冗余性。

下面我们介绍常见的一些图像数据冗余的情况。

(1) 空间冗余

这是静态图像存在的最主要的一种数据冗余。一幅图像记录了画面上可见景物的颜色。同一景物表面上各采样点的颜色之间往往存在着空间连贯性,但是基于离散像素采样来表示物体颜色的方式通常没有利用景物表面颜色的空间连贯性,从而产生了空间冗余。我们可以通过改变物体表面颜色的像素存储方式来利用空间连贯性,达到减少数据量的目的。例如,在静态图像中有一块表面颜色均匀的区域,在此区域中所有点的光强、色彩以及饱和度都是相同的,因此数据有很大的空间冗余。

(2) 时间冗余

这是序列图像(电视图像、运动图像)表示中经常包含的冗余。序列图像一般位于一个时间轴区间内的一组连续画面,其中的相邻帧往往包含相同的移动物体,只不过移动物体的空间位置略有不同,所以后一帧的数据与前一帧的数据有许多共同的地方,这种共同性是由于相邻帧记录了相邻时刻的同一场景画面,所以称为时间冗余。

(3) 结构冗余

在有些图像的纹理区,图像的像素值存在着明显的分布模式,例如方格状的地板图案等,我们称此为结构冗余。已知分布模式,可以通过某一过程生成图像。

(4) 知识冗余

有些图像的理解与某些知识有相当大的相关性。例如,人脸的图像有固定的结构,嘴的上方有鼻子,鼻子的上方有眼睛,鼻子位于正脸图像的中线上等等。这类规律性的结构可由经验知识和背景知识得到,我们称此类冗余为知识冗余。根据已有的知识,对某些图像中所包含的物体,我们可以构造其基本模型,并创建对应各种特征的图像库,进而图像的存储只需要保存一些参数,从而可以大大减少数据量。模型编码主要利用知识冗余的特性。

(5) 视觉冗余

事实表明,人类的视觉系统对图像场的敏感性是非均匀和非线性的。例如,在记录原始的图像数据时,通常规定视觉系统是线性和均匀的,对视觉敏感和不敏感的部分同等对待,从而产生了比理想编码(即把视觉敏感和不敏感的部分区分开来编码)更多的数据,这就是视觉冗余。通过对人类视觉进行的大量实验,发现了以下的视觉非均匀特性。