

Information

全国高职高专应用型规划教材
信息技术类



多媒体技术与应用

韩最蛟 主编



北京大学出版社
PEKING UNIVERSITY PRESS

全国高职高专应用型规划教材·信息技术类

多媒体技术与应用

主 编 韩最蛟

副主编 巩晓秋 樊广峰 张 永

参 编 汤 丽 吴险峰 汪 洋



北京大学出版社
PEKING UNIVERSITY PRESS

内 容 简 介

随着计算机技术、通信技术的发展,多媒体技术成为现代计算机技术的重要发展方向和快速发展的领域之一。本书编者在结合多年计算机教学经验的基础上,对多媒体技术进行了探讨。本书共7章,包括多媒体技术基础、多媒体音频技术、图形与图像处理、计算机动画技术、多媒体视频技术、网络多媒体技术和多媒体著作工具等内容。在章节之后配以小结和相关习题,以帮助学生归纳总结所学知识,加深学生对知识的理解,同时提高其分析问题和解决问题的能力。

本书内容新颖,选材丰富,叙述简明,逻辑清晰,可作为高职高专计算机应用技术、软件技术等专业的教材,也可作为从事相关软件设计工作人员的学习指导用书。

图书在版编目(CIP)数据

多媒体技术与应用/韩最蛟主编. —北京:北京大学出版社, 2009.9

(全国高职高专应用型规划教材·信息技术类)

ISBN 978-7-301-15819-7

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

中国版本图书馆 CIP 数据核字(2009)第 167449 号

书 名: 多媒体技术与应用

著作责任者: 韩最蛟 主编

责任编辑: 葛昊晗

标准书号: ISBN 978-7-301-15819-7

出版者: 北京大学出版社

地址: 北京市海淀区成府路 205 号 100871

网 址: <http://www.pup.cn>

电 话: 邮购部 62752015 发行部 62750672 编辑部 62765126 出版部 62754962

电子信箱: xxjs@pup.pku.edu.cn

印刷者: 涿州市星河印刷有限公司

发 行 者: 北京大学出版社

经 销 者: 新华书店

787 毫米×980 毫米 16 开本 21.5 印张 519 千字

2009 年 9 月第 1 版 2009 年 9 月第 1 次印刷

定 价: 39.00 元

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究

举报电话: 010-62752024; 电子信箱: fd@pup.pku.edu.cn

目 录

第 1 章 多媒体技术基础	1	2.5 MIDI 与音乐合成	66
1.1 多媒体技术概述	1	2.5.1 MIDI 技术概述	66
1.1.1 多媒体技术的基本概念	1	2.5.2 MIDI 合成方式	70
1.1.2 多媒体技术的特点	3	2.5.2 MIDI 音乐制作系统	72
1.1.3 多媒体技术的分类	4	本章实训	74
1.1.4 多媒体技术的发展	9	本章小结	77
1.1.5 多媒体技术的应用	11	习题	77
1.2 多媒体系统组成与分类	13	第 3 章 图形与图像处理	79
1.2.1 多媒体系统的层次结构	14	3.1 图形与图像概述	79
1.2.2 多媒体系统的硬件组成	15	3.1.1 光和颜色	79
1.2.3 多媒体系统的软件组成	16	3.1.2 图形与图像	81
1.2.4 多媒体系统的分类	17	3.1.3 图像数字化	83
1.3 多媒体艺术基础	18	3.1.4 图像的文件格式	84
1.3.1 艺术基础	18	3.2 静止图像压缩标准	87
1.3.2 构图	19	3.2.1 图像压缩方法概述	87
1.3.3 色彩设计	24	3.2.2 JPEG 图像压缩标准	89
本章小结	31	3.2.3 JPEG 图像压缩工具	90
习题	31	3.3 图形图像信息的采集	92
第 2 章 多媒体音频技术	33	3.3.1 数码相机	92
2.1 数字音频基础	33	3.3.2 扫描仪	93
2.1.1 声音的基本概念	33	3.4 图像处理	96
2.1.2 声音的数字化	36	3.4.1 ACDSee 8	96
2.1.3 数字音频的文件格式	38	3.4.2 Photoshop 11 图像处理	100
2.2 数字音频压缩标准	40	本章实训	111
2.2.1 音频压缩方法概述	40	本章小结	117
2.2.2 音频压缩技术标准	43	习题	117
2.2.3 音频压缩工具	46	第 4 章 计算机动画技术	119
2.3 音频信息的采集	49	4.1 计算机动画概述	119
2.3.1 数字音频的录制	49	4.1.1 计算机动画的基本概念	119
2.3.2 语音输入	50	4.1.2 计算机动画与传统动画的 差异	120
2.4 音频处理	54	4.1.3 计算机动画的分类	120
2.4.1 常用音频处理软件	54		
2.4.2 Adobe Audition 3.0 音频处理 ..	56		

4.2 计算机动画制作的硬件和软件.....	121	5.4.4 编辑视频素材	202
4.2.1 计算机动画制作硬件平台	121	5.4.5 编辑、处理视频	208
4.2.2 二维动画制作软件的功能	122	5.4.6 分享视频	222
4.2.3 动画制作软件介绍	124	本章实训	231
4.3 Flash 动画制作	125	本章小结	236
4.3.1 Flash 动画的工作环境和 基本概念	125	习题	236
4.3.2 绘制图形	127	第 6 章 网络多媒体技术	239
4.3.3 时间轴动画	129	6.1 网络多媒体概述	239
4.3.4 路径补间动画	133	6.1.1 多媒体信息对网络的需求	239
4.3.5 时间轴特效	134	6.1.2 多媒体网络环境	240
4.3.6 使用遮罩层	136	6.1.3 网络多媒体应用	243
4.3.7 元件、实例和库资源	138	6.2 多媒体通信协议及标准	245
4.3.8 使用视频	142	6.2.1 多媒体通信协议	245
4.3.9 声音处理	144	6.2.2 多媒体通信标准 (H.323 标准)	249
4.3.10 发布和导出	146	6.3 流媒体	253
4.3.11 ActionScript 简介	147	6.3.1 流媒体的基本概念	254
本章实训	148	6.3.2 流媒体传输协议	256
本章小结	153	6.3.3 常见流媒体文件格式	257
习题	153	6.3.4 流媒体系统	259
第 5 章 多媒体视频技术	156	本章实训	262
5.1 视频基础	156	本章小结	268
5.1.1 模拟视频	156	习题	268
5.1.2 数字视频	160	第 7 章 多媒体著作工具	270
5.1.3 视频的文件格式	162	7.1 多媒体著作工具概述	270
5.2 数字视频压缩标准	165	7.1.1 多媒体著作工具的功能	270
5.2.1 视频压缩的基本原理	165	7.1.2 多媒体著作工具的分类	271
5.2.2 视频编码标准	167	7.2 多媒体著作工具的评价和选择	271
5.2.3 视频格式转换工具	176	7.3 演示文稿制作工具	273
5.3 数字视频信息的采集	179	7.3.1 PowerPoint 2003 概述	273
5.3.1 视频卡	180	7.3.2 演示文稿的基本操作	273
5.3.2 IEEE 1394 接口	182	7.3.3 图形与表格	277
5.3.3 数码摄像机	185	7.3.4 多媒体幻灯片	278
5.4 会声会影 X2 编辑与处理视频	187	7.3.5 幻	279
5.4.1 非线性编辑概述	187	7.4 Authorware 多媒体著作工具	283
5.4.2 会声会影 X2 简介	190	7.4.1 Authorware 7.0 概述	283
5.4.3 捕获视频	198	7.4.2 图标的使用	284

7.4.3 多媒体素材管理	298	本章实训	327
7.4.4 交互功能的使用	304	本章小结	333
7.4.5 打包与发布	322	习题	333

第1章 多媒体技术基础

教学提示：本章介绍多媒体及多媒体技术的有关概念、多媒体技术的特点、多媒体计算机的基本硬件配置和软件环境以及多媒体技术的应用与发展趋势等知识。通过基础理论的学习，帮助大家了解多媒体有关知识，为进一步学习后续章节做好准备。

教学要求：理解媒体及多媒体的基本概念，掌握多媒体的特征和多媒体系统的构成，了解多媒体技术中的关键技术及其应用。

1.1 多媒体技术概述

随着计算机技术、通信技术的发展，人类获得信息的途径越来越多，获得信息的形式越来越丰富，信息的获得也越来越方便、快捷，人们对多媒体这个名词越来越熟悉。多媒体技术是现代计算机技术的重要发展方向，也是现代计算机技术发展最快的领域之一。多媒体计算机技术与通信技术、网络技术的融合与发展打破了时空和环境的限制，涉及了计算机出版业、远程通信、家用音像电子产品以及电影与广播等主要工业范畴，从根本上改变了人们的生活方式和现代社会的信息传播方式，是社会信息化高速公路的基础。

1.1.1 多媒体技术的基本概念

1. 媒体

媒体可以理解为是人与人或人与外部世界之间进行信息沟通与交流传递的载体。

国际电信联盟电信标准协会（ITU-TSS）对多媒体进行了定义，并制定了 ITU-TI.374 建议。在 ITU-TI.374 建议中，把媒体分为以下五大类。

① 感觉媒体：指能够直接刺激人的感觉器官，使人产生直观感觉的各种媒体。或者说，人类感觉器官能够感觉到的所有刺激都是感觉媒体。比如：人的耳朵能够听到的话音、音乐、噪声等各种声音；人的眼睛能够感受到的光线、颜色、文字、图片、图像等各种有形有色的物体等。感觉媒体包罗万象，存在于人类感觉到的整个世界。

② 显示媒体：指感觉媒体与电磁信号之间的转换媒体。显示媒体分为输入显示媒体和输出显示媒体。输入显示媒体主要负责将感觉媒体转换成电磁信号，比如：话筒、键盘、光笔、扫描仪、摄像机等。输出显示媒体主要负责将电磁信号转换成感觉媒体，比如：显示器、打印机、投影仪、音响等。

③ 表示媒体：对感觉媒体的抽象描述形成表示媒体。比如声音编码、图像编码等。通过表示媒体，人类的感覺媒体转换成能够利用计算机进行处理、保存、传输的信息载体形式。因此，对表示媒体的研究是多媒体技术的重要内容。

④ 存储媒体：指存储表示媒体的物理设备，比如磁盘、光盘、磁带等。

⑤ 传输媒体：指传输表示媒体的物理介质，比如电缆、光缆、电磁波等都是传输媒体。

ITU-TI.374 建议将感觉媒体传播存储的各种形式都定义成媒体，人类获得和传递信息的过程就是各种媒体转换的过程。

以上各种媒体的描述中其核心是表示媒体，即信息的存在形式和表现形式，如日常生活中的报纸、电视、广播、广告和杂志等信息，借助于这些载体得以交流传播。如果对这些媒体的本质进行详细分析，就可以找到媒体传递信息的基本元素，如声音、图片、视频、影像、动画和文字等，它们都是媒体的组成部分。

媒体在计算机领域中有两种含义：一个是用以存储信息的实体，如磁带、磁盘、光盘和半导体存储器；另一个是信息的载体，如数字、文字、声音、图形和图像。多媒体技术中的媒体是指后者。

2. 多媒体

多媒体来自英文 Multimedia，该词由 Multiple（多）和 Media（媒体）复合而成。简单理解，多媒体是指两个或两个以上的单媒体的有机组合，意味着“多媒介”或“多方法”。日常生活中媒体传递信息的基本元素是声音、文字、图像、动画、视频和影像等，这些基本元素的组合就构成了人们平常接触的各种信息。计算机中的多媒体就是指将文字、图形、图像、音频、视频和动画等基本媒体元素以不同形式组合以传递信息的有机综合。

3. 多媒体技术

多媒体技术的含义和范围很广泛，一般认为多媒体技术是指能同时获取、处理、编辑、存储和展示两个以上不同类型信息媒体（如文字、声音、图形、图像、动画、视频等）的技术。

多媒体技术是指利用计算机交互式综合处理多种媒体信息——文本、图形、图像和声音等，使多种媒体之间建立逻辑连接，集成为一个整体系统并具有一定的交互性。

多媒体技术是将多种媒体信息通过计算机进行数字化采集、编码、存储、传输、处理和再现等，使多媒体信息建立逻辑连接，并集成一个具有交互性的系统。

多媒体技术是一种崭新的、跨学科的综合技术，涉及计算机硬件、软件、计算机体系结构；编码学，数值处理方法；图像处理；计算机图形学；声音和信号处理；人工智能；计算机网络和高速通讯技术等。

多媒体技术的应用对信息管理、办公自动化、教育和学习带来巨大而深刻的变革，在上世纪 90 年代以前，人机交互方式主要是通过基于文字或简单图形的界面来实现，各种媒体往往是单独使用，有文字而无声音；有静止画面而无活动图像；交互活动枯燥，信息的媒体表现形式单调。多媒体技术为人机之间提供了全新的信息交流手段和信息的多种表现形式，可以把多种媒体巧妙地组合在一起，多维度地协同表达同一事物。例如，同时具有图像、动画、声音、视频组合的教学课件，对教育者和学习者形成了巨大的吸引力。

多媒体技术在许多领域都得到了广泛的应用：科学研究部门用来存储、显示、传送信息；文化、娱乐部门的应用；各级各类学校教育、培训；办公自动化；多媒体电子出版物、信息管理与咨询、多媒体通信等。

1.1.2 多媒体技术的特点

多媒体之所以能够迅速发展和广泛应用,是由于计算机技术和数字信息处理技术的突破性进展。多媒体技术就是指处理和应用多媒体的,包括硬件和软件在一起的一整套技术。多媒体技术主要具有以下特点。

1. 交互性

交互性是多媒体技术的关键特征。它可以更有效地控制和使用信息,增加对信息的理解。所谓交互性是指人的行为与计算机的行为互为交流沟通的关系。这也是多媒体与传统媒体最大的不同。电视教学系统虽然也具有“声、图、文”并茂的多种信息媒体,但电视节目的内容是事先安排好的,人们只能被动地接受播放的节目,而不能随意选择感兴趣的内容,这个过程是单方向的,而不是双向交互性的。如果用多媒体技术制作教学系统,学生可根据自己的需要选择章节不同、难易各异的内容进行学习。对于重点的内容,一次未搞明白,还可重复播放。学生可参与练习、测验和实际操作等。如果学生有错,多媒体教学系统能及时评判、提示和纠正。

在多媒体远程计算机辅助教学系统中,学习者可以人为地改变教学过程,研究感兴趣的问题,从而得到新的知识,激发学习者的主动性、自觉性和积极性。再如在多媒体远程信息检索系统中,初级交互性可帮助用户找出想读的书,从数据库中检录声音、图像或文字材料等,并且能快速跳过不感兴趣的部分。中级交互性则可使用户介入到信息的提取和处理过程中,如对关心的内容进行编排、插入文字说明及解说等。当采用虚拟现实或灵境技术时,多媒体系统可提供高级的交互性。

2. 复合性

信息媒体的复合性是相对于计算机而言的,也可称为媒体的多样化或多维化,它把计算机所能处理的信息媒体的种类或范围扩大,不仅仅局限于原来的数据、文本或单一的语音、图像。众所周知,人类具有五大感觉,即视、听、嗅、味与触觉。前三种感觉占了总信息量的95%以上,而计算机远没有达到人类处理复合信息媒体的水平。计算机一般只能按单一方式处理信息。信息的复合化或多样化不仅是指输入信息,称为信息的获取,而且还指信息的输出,称为表现。输入和输出并不一定相同,若输入与输出相同,就称为记录或重放。如果对输入进行加工、组合与变换,则称为创作。创作可以更好地表现信息,丰富其表现力,使用户更准确、更生动地接收信息。这种形式过去在影视制作过程中大量采用,在当今多媒体技术中也经常采用这种方法。

3. 集成性

多媒体的集成性包括两方面,一是多媒体信息媒体的集成;二是处理这些媒体的设备和系统的集成。在多媒体系统中,各种信息媒体不是像过去那样,采用单一方式进行采集与处理,而是多通道同时统一采集、存储与加工处理,更加强调各种媒体之间的协同关系及利用它所包含的大量信息。此外,多媒体系统应该包括能处理多媒体信息的高速及并行的CPU、多通道的输入/输出接口及外设、宽带通信网络接口与大容量的存储器,并将这些

硬件设备集成为统一的系统。在软件方面,则应有多媒体操作系统,满足多媒体信息管理的软件系统、高效的多媒体应用软件和创作软件等。在网络的支持下,这些多媒体系统的硬件和软件被集成为处理各种复合信息媒体的信息系统。

4. 实时性

由于多媒体系统需要处理各种复合的信息媒体,决定了多媒体技术必然要支持实时处理。接收到的各种信息媒体在时间上必须是同步的,比如语音和活动的视频图像必须严格同步,因此要求实时性,甚至是强实时。例如电视会议系统的声音和图像不允许存在停顿,必须严格同步,包括“唇音同步”,否则传输的声音和图像就失去意义。

5. 控制性

多媒体技术是以计算机为中心,综合处理和控制在多媒体信息,并按人的要求以多种媒体形式表现出来,同时作用于人的多种感官。

6. 非线性

多媒体技术的非线性特点将改变人们传统循序性的读写模式。以往人们读写方式大都采用章、节、页的框架,循序渐进地获取知识,而多媒体技术将借助超文本链接的方法,把内容以一种更灵活、更具变化的方式呈现给读者。

1.1.3 多媒体技术的分类

多媒体基础技术包括多媒体操作系统技术、功能芯片技术、输入输出技术、数据压缩技术、光存储技术等。

1. 多媒体操作系统技术

多媒体操作系统技术也称为多媒体核心系统,具有实时任务调度、多媒体数据转换和同步控制对多媒体设备的驱动和控制,以及图形用户界面管理等。多媒体操作系统是多媒体应用程序的运行平台。苹果公司早在 20 世纪 80 年代就已推出专为处理多媒体数据而设计的操作系统。微软公司吸收了苹果系统的多媒体功能特点,设计开发了在个人计算机上应用的 Windows 多媒体操作系统。目前应用的版本有 Windows 2000、Windows XP 和 Windows Vista 等。为使用户享受到更出色的多媒体效果,Windows 操作系统支持 Microsoft DirectX 技术、数字化视盘(DVD)以及高速 IEEE 1394 接口,还有增强的色彩管理和多监视器工作的能力。

2. 多媒体功能芯片技术

多媒体技术的发展和超大规模集成电路技术的发展有着密不可分的关系。由于多媒体数据量极大,要实现视频、音频信号的实时压缩、解压缩和多媒体信息的播放处理,需要对大量的数据进行快速计算,必须具有多媒体功能的快速运算硬件支持。实现动态视频的



实时采集、变形、叠加和合成等特殊效果处理,也必须采用专用的视频处理芯片才能取得满意的效果。支持多媒体功能的 CPU 芯片 (MMX) 和专用的音频、视频处理芯片的研制都是在大规模集成电路技术的支持下实现的。

多媒体计算机的专用芯片可分为两类:一类是固定功能的芯片,另一类是可编程数字信号处理器 DSP 芯片。

最早推出的固定功能的专用芯片是图像处理的压缩处理芯片,即将实现静态图像的数据压缩/解压缩/算法做在一个专用芯片上,从而大大提高其处理速度。以后,许多半导体厂商或公司又推出执行国际标准压缩编码的专用芯片,由于压缩编码的国际标准较多,一些厂家和公司还推出多功能视频压缩芯片,而 Intel 公司开发的新一代 750 芯片,不仅为多媒体应用提供了足够的计算能力,而且已达到 1BIPS (Billion Instructions Per Second) 的运算速度。

由于采用多处理器并行技术,计算能力可望达到 2BIPS,这些高档的专用多媒体处理器芯片,不仅大大提高了音、视频信号处理速度,而且在音频、视频数据编码时增加特技效果。

除专用处理器芯片外,多媒体系统还需要其他集成电路芯片支持,如数/模 (D/A) 和模/数 (A/D) 转换器、音频、视频芯片,彩色空间变换器及时钟信号产生器等等。

3. 多媒体输入输出技术

输入输出技术是处理多媒体信息传输接口的界面,主要包括媒体转换技术、媒体识别与理解技术(如语音识别)等。其中既包括硬件技术,又包括软件技术。

(1) 媒体变换技术:是指改变媒体的表现形式,如当前广泛使用的视频卡、音频卡(声卡)都属于媒体变换技术。

(2) 媒体识别技术:是对信息进行一对一的映像过程。例如,语音识别是将语音映像为一串字、词或句子;触摸屏是根据触摸屏上的位置识别其操作。

(3) 媒体理解技术:是对信息进行更进一步的分析处理和理解信息内容,如,自然、语言理解,图像理解,模式识别这类技术。

(4) 媒体综合技术:是把低维信息表示映像成高维的模式空间的过程,如语音合成器就可以把语音的内部表示综合为声音输入。

(5) 人工智能输入/输出技术:主要包括语音识别、语音合成、语言翻译、语言和文本间转换;图像识别和处理,图/文/表分离技术;笔式输入技术和智能推理技术等。围绕实用过程均需进一步解决压缩、集成和交互、同步等处理。

(6) 外围设备控制技术:主要包括多媒体文件存储、数据格式转换、控制界面、外围设备驱动程序、调色板控制、高分辨率全彩色显示、三维彩色、声音效果处理、通讯效果处理、多媒体窗口程序等。

(7) 多媒体网络传输技术:主要包括网络管理技术、高速网络协议、开放式文件结构、视像会议、不同网络间的传输技术、ISDN 通信技术、电子邮件传送等。

4. 多媒体数据压缩技术

多媒体数据压缩是多媒体技术的主要特征。未经压缩的视频和音频数据占用大量的空

间,例如未经压缩的影像和立体声音乐数据量分别是 1680MB/min 和 10MB/min。如此庞大的数据量不仅难于用普通计算机处理,而且存储和传输都有问题。因此,视频、音频和图像数据的编码和压缩算法在多媒体技术中占有非常重要的地位。视频、音频数字信号的编码和压缩算法成为一个重要的研究课题。

编码理论研究已有 40 多年的历史,技术已日趋成熟,在研究和选用编码时,主要有两个问题:一是该编码方法能用计算机软件或集成电路芯片快速实现。二是一定要符合压缩编码/解压缩编码的国际标准。

信息理论认为,若信源编码的熵大于信源的实际熵,该信源中一定存在冗余度。去掉冗余不会减少信息量,仍可原样恢复数据;但若减少了熵,数据则不能完全恢复。不过在允许的范围内损失一定的熵,数据可以近似地恢复。根据压缩过程中是否减少了熵,目前常用的压缩编码方法可以分为两大类:一类是无损压缩编码法,也称冗余压缩法或熵编码法;另一类是有损压缩编码法,也称为熵压缩法。

无损压缩法去掉或减少了数据中的冗余,但这些冗余值是可以重新插入到数据中的,因此,这种压缩是可逆的,也称为无失真压缩。为了去除数据中的冗余度,常常要考虑信源的统计特性,或建立信源的统计模型,因此许多适用的冗余度压缩技术均可归结于统计编码方法。此外,统计编码技术在各种熵压缩方法中也经常会用到。

统计编码方法有霍夫曼编码、算术编码、游程编码等。冗余压缩法由于不会产生失真,因此在多媒体技术中一般用于文本、数据以及应用软件的压缩,它能保证完全地恢复原始数据。但这种方法压缩比较低,如 LZ 编码、游程编码、霍夫曼编码的压缩比一般在 2:1~5:1。无损压缩法广泛用于文本数据、程序和特殊应用场合的图像数据(如指纹图像、医学图像等)的压缩。由于压缩比的限制,仅使用无损压缩方法不可能解决图像和数字视频的存储和传输问题。

有损压缩法压缩了熵,会减少信息量。因为熵定义为平均信息量,而损失的信息是不能再恢复的,因此这种压缩法是不可逆的。熵压缩主要有两大类:特征抽取和量化。特征抽取的编码方法如基于模型的编码、分形编码等。对于实际应用而言,量化是更为通用的熵压缩技术,包括特征提取、零记忆量化、预测编码、直接映射、变换编码等,其中预测编码和变换编码是最常见的实用压缩编码方法。熵压缩法由于允许一定程度的失真,可用于对图像、声音、动态视频等数据的压缩。如采用混合编码的 JPEG、MPEG 等标准,它对自然景物的灰度图像,一般可压缩几倍到几十倍,而对于自然景物的彩色图像,压缩比将达到几十倍甚至上百倍;采用自适应差分脉冲编码调制的声音数据,压缩比通常能到 4:1~8:1;动态视频数据的压缩比最为可观,采用混合编码的多媒体系统,压缩比通常可达 100:1~400:1。有损压缩广泛应用于语音、图像和视频数据的压缩。

1948 年,Oliver 提出了第一个编码理论——脉冲编码调制(Pulse Coding Modulation,简称 PCM);同年,Shannon 的经典论文——“通信的数学原理”首次提出并建立了信息率失真函数概念;1959 年,Shannon 进一步确立了码率失真理论,以上工作奠定了信息编码的理论基础。

同时在数据编码上,又有第一代编码和第二代编码之称。第一代编码的方法主要有预测编码、变换编码和统计编码,也称为 3 大经典编码方法。

(1) 预测编码: 基本思想是根据数据的统计特性得到预测值, 然后传输图像像素与其预测值的差值信号, 使传输的码率降低, 达到压缩的目的。预测编码方法简单经济, 编码效率较高。

(2) 变换编码: 基本思想是由于数字图像像素间存在高度相关性, 因此可以进行某种变换来消除这种相关性。目前, 国际上已经制订了基于离散余弦变换的静止图像压缩标准 JPEG 和运动图像压缩标准 MPEG 等一系列标准。

(3) 统计编码: 基本思想是针对无记忆信源, 根据信息码字出现概率的分布特征而进行压缩编码, 寻找概率与码字长度间的最优匹配。常用的统计编码有游程编码、Huffman 编码和算术编码 3 种。

这些编码技术都是非常优秀的纹理编码方案, 它们能够在中等压缩率的情况下, 提供非常好的图像质量, 但在非常低的位率情况下, 无法为一般的序列提供令人满意的质量。20 世纪 80 年代初期, 第一代编码技术已经达到了顶峰, 这类技术去除客观和视觉冗余信息的能力已接近极限。究其原因是一些技术都没有利用图像的结构特点, 因此它们也就只能以像素或块作为编码的对象, 另外, 这些技术在设计编码器时也没有考虑人类视觉系统的特性。

为了克服第一代编码技术的局限性, Kunt 等人于 1985 年提出了第二代编码技术。他们认为, 第一代编码技术只是以信息论和数字信号处理技术为理论基础, 旨在去除图像数据中的线性相关性的一类编码技术, 其压缩比不高。而第二代编码技术不局限于信息论的框架, 要充分利用人的视觉生理、心理和图像信源的各种特征, 实现从波形编码到模型编码的转变, 以便获得更高的压缩比。第二代编码方法主要有: 基于分形的编码、基于模型的编码、基于区域分割的编码和基于神经网络的编码等。

(1) 基于分形的编码: 一种不对称的编码技术, 适用于相似性较强的自然景物图像。

(2) 基于模型的编码: 在编码端通过各种分析手段, 提取所建模型的特征与状态参数。在解码端依据这些参数, 通过模型及相关知识生成所建模的信源。这类方法把计算机视觉和计算机图形学中的方法应用到视频(图像)编码。

(3) 基于区域分割的编码: 根据图像的空域特征将图像分成纹理和轮廓两部分, 然后分别对它们进行编码。该方法一般可分为 3 步来完成, 即预处理、编码和滤波。预处理将图像分割成纹理和轮廓两部分。选取分割方法是关键, 它直接影响图像编码的效果。分割之后图像成为一系列相连的小区域。对纹理可采用预测编码和变换编码, 对轮廓则采用链码方法进行编码。这种方法较好地保存了对人眼十分重要的边缘轮廓信息, 因此在压缩比很高时解码图像质量仍然很好。

(4) 基于神经网络的编码: 模仿人脑处理问题的方法, 通过各种人工神经网络模型对数据进行非线性压缩。人工神经网络是一个非线性动态网络, 工作过程一般分训练和工作两个阶段。训练阶段就是使用一些训练图像和训练算法, 调整网络的权重, 使重建图像的误差最小。目前直接用于图像压缩编码的神经网络主要有反向误差传播(BP)型和自组织映射型。

第二代编码方法充分利用了计算机图形学、计算机视觉、人工智能与模式识别等相关学科的研究成果, 为视频(图像)压缩编码开拓了广阔的前景。但是由于第二代编码方法

增加了分析的难度,所以大大增加了实现的复杂性。从当前发展情况来看,第二代编码方法仍处于深入研究的阶段。例如,分形法由于图像分割、迭代函数系统代码的获得是非常困难的,因而实现起来时间长,算法非常复杂。模型法则仅限于人头肩像等基本的视频(图像)上,进一步的发展有赖于新的数学方法和其他相关学科的发展。神经网络的工作机理至今仍不清楚,硬件研制不成功,所以在视频(图像)编码中的应用研究中进展缓慢,目前多与其他方法结合使用。但由于其巨大压缩性能的潜力,人们都在致力于这些新方法的研究之中。

(1) 过渡编码技术:近年来,出现了一类充分利用人类视觉特性的多分辨率编码方法,如子带编码、塔形编码和基于小波变换的编码。这类方法使用不同类型的一维或二维线性数字滤波器,对视频(图像)进行整体的分解,然后根据人类视觉特性对不同频段的数据进行粗细不同的量化处理,以达到更好的压缩效果。这类方法原理上仍属于线性处理,属于波形编码,可归入经典编码方法,但它们又充分利用了人类视觉系统的特性,因此可以被看做是第一代编码技术向第二代编码技术过渡的桥梁。

(2) 子带编码技术:子带编码是一种高质量、高压缩比的图像编码方法,它早已在语音信号压缩编码中获得了广泛的应用。其基本依据是:语音和图像信号可以划分为不同的频域段,人眼对不同频域段的敏感程度不同,例如图像信号的主要能量集中在低频区域,它反映图像的平均亮度,而细节、边缘信息则集中在高频区域。子带编码的基本思想是利用一滤波器组,通过重复卷积的方法,经取样将输入信号分解为高频分量和低频分量,然后分别对高频和低频分量进行量化和编码。解码时,高频分量和低频分量经过插值和共轭滤波器而合成原信号。进行子带编码的一个关键问题,是如何设计共轭滤波器组,除去混叠频谱分量。

(3) 基于小波变换的编码技术:基于小波变换的编码技术具有特别重要的意义。它不仅为多分辨率分析、时-频分析和子带编码建立了统一的分析方法,提供了更合理的表示框架,而且它体现着小波分析这一新型分析方法的优越性。目前小波变换在多媒体编码中的应用研究主要有:正交小波基的选择(小波包法)、小波变换与各种量化方式的结合、小波变换在分形法中实现初级分形、小波变换用于运动估值等方面。可以说,小波变换法处于多媒体编码当前首选方法的位置,一方面,它有快速算法,实现起来简单方便、速度快,可暂时弥补第二代编码技术的不足;另一方面,它有着先进的分析方法,可有效提高现有标准的水平,实现突破性进展。

5. 光存储技术

大量多媒体信息数据需要很大的存储空间,因此,多媒体技术的发展和必须有大容量存储技术的支持。多媒体的音频、视频、图像等信息虽经过压缩处理,但仍需相当大的存储空间,只有在大容量只读光盘存储器 CD-ROM 问世后才真正解决了多媒体信息存储空间问题。

近几年,光存储技术得到迅速发展,目前存储容量很大的 CD、DVD 光盘存储器已广泛使用。这些盘在形状、尺寸、面积和重量等方面基本一样,但 DVD 的存储容量和带宽大大高于 CD。单面单层 DVD 盘片能够存储 4.7GB 的数据,单面双层盘片的容量为 8.5GB。

由于存储在 PC 机服务器上的数据量越来越大,使得 PC 机服务器的硬盘容量需求提高很快。为了避免磁盘损坏而造成的数据丢失,采用了相应的磁盘管理技术,磁盘阵列 (Disk Array) 就是在这种情况下诞生的一种数据存储技术。这些大容量存储设备为多媒体应用提供了便利条件。

1.1.4 多媒体技术的发展

1. 多媒体技术发展简介

1964 年,美国 SRI 公司发明了鼠标,使计算机的输入操作方式产生了变革,为 20 世纪 70 年代的图形用户界面 (GUI) 等图形处理软件的诞生与应用,起了支撑的作用。

1971 年,Intel 公司推出世界上第一个微处理器 Intel 4004,这是计算机发展史上的一座里程碑。随着大规模、超大规模集成电路 (LSI、VLSI) 的出现,计算机的文字处理、图形处理功能走向实用,声像处理功能取得突破,多媒体计算机已然是呼之欲出了。

1982 年,Philips 和 Sony 公司联合推出数字激光唱盘 CD-DA,CD-DA 在当时只能记录数字化的音频信息,但是它必定也能记录计算机的数据信息。

1984 年 APPLE 公司的 Macintosh 个人计算机,首先引入了位图 (Bitmap) 的概念来描述和处理图形和图像,并使用窗口 (Window) 和图标 (Icon) 构筑图形用户界面 (GUI)。

1984 年 Apple 公司的 Macintosh 被誉为世界上最早的多媒体计算机 (MPC)。它的组成部分包括主机、多媒体插板、CD-ROM 驱动器,以及图像输入输出设备等。Macintosh 的主要贡献有:率先采用位映射和图符技术来处理图形;运用超级卡 (Hypercard),使高保真音响和动态图像处理功能融入计算机;运用了窗口、菜单、面向对象和超文本技术等。

1985 年,美国 Commodore 公司推出多媒体计算机系统 Amiga,后来形成系列产品。到目前为止,该公司已经推出 Amiga500,1000,2000,2500 以及 3000 等型号的产品。这套系统以其功能完备的视听处理能力、大量丰富的使用工具和性能优良的硬件,使世界看到了多媒体技术的美好未来。

1985 年,Philips 和 Sony 公司又联合推出可读光盘系统 (CD-ROM),它就是专为计算机使用的新一代存储系统。CD-ROM 盘片的直径为 12cm,容量 650MB,可储存 3 亿个汉字,相当于 15 万张 A4 纸的存储量。

1986 年,Philips 和 Sony 再次联合推出可读光盘交互系统 (CD-I),同时公布了一种新的 CD-ROM 存储格式,后来国际标准化组织 (ISO) 采纳该格式作为 CD-ROM GreenBook 标准。CD-I 系统把高质量的声音、文字、计算机程序、图形、动画以及静止图像等都以数字的形式存放在容量为 650MB 的 5 英寸只读光盘中。用户可通过 CD-ROM 驱动器来播放光盘中的内容。

1989 年,新加坡 Creative Labs 公司在世界上率先推出支持数字化录音、放音功能的 PC 机音效卡,号称“声霸卡”。1991 年推出 2.0 版,以后又推出 SoundBlasterPro 版和 SoundBlaster 16ASP 版等。此外,新加坡 Creative Labs 公司还推出了视频卡。在 PC 机上加接视频卡就可以存储、定格、处理和播放影视节目,在图像中叠加图形或文字,调节色度、亮度和对比度,可以使之与录像机、摄像机、有线电视、数字相机、激光视盘等设备相连,还可以

将图像画面存储到硬盘中。目前市场上还有多种视频输出卡,常见的有 TV-Coder 卡、ProVGA / TV 卡、VideoPower-1000 卡等。

1990 年,由 Microsoft 公司和多家厂商成立了多媒体计算机市场协会,制定了著名的 MPC 标准。1991 年,多媒体计算机市场协会又制定了多媒体 PC 机的基本标准 MPC-1。

1992 年,“运动图像专家小组”正式公布 MPEG-1 标准。

1993 年,多媒体计算机市场协会又推出了 MPC 第二个标准。

2. 多媒体技术的发展特点

多媒体技术的飞速发展导致了计算机应用领域的一场革命,把信息社会推向了一个新的历史时期,使人类生活进入一个崭新的世界,对人类社会产生了深远的影响。多媒体技术的发展,显示出以下几个特点。

(1) 多学科交汇

多媒体技术的发展融合了计算机科学、微电子科学、声像技术、数字信号处理技术、网络与通信技术和人工智能等多门学科,而且有与其他科学联合的趋势。

(2) 顺应信息时代发展的需要

现代人类文明的发展与进步,要求提供全方位的综合信息处理技术,提供信息表示和显示的全新工具。多媒体技术改善了人机之间的界面,使计算机应用更有效,更接近于人类习惯的信息交流方式。信息空间走向多维化,使人们思想的表达不再局限于顺序的、单调的、狭窄的范围,而有了一个充分自由的空间,多媒体技术为这种自由提供了多维化空间的交互能力,人与信息、人与系统、信息与系统之间的交互方法发生了变革,顺应了信息时代的需要,必将推动信息社会的进一步发展。

(3) 多领域应用

多媒体技术已经在我们的生活中达到了应用,用多媒体计算机进行的家庭教育和个人娱乐已成时尚,多媒体应用逐渐进入千家万户,这一新兴的技术,必然会在社会上崛起一支新兴的产业大军,多媒体技术必将渗入到我们生活、工作的各个方面。

3. 多媒体技术的发展方向

目前,多媒体技术主要向以下几个方向发展。

(1) 多媒体通信网络的研究和建立

多媒体通信网络的研究和建立将使多媒体从单机、单点向分布、协同多媒体环境发展,在世界范围内建立一个可全球自由交互的通信网。对该网络及其设备的研究和网上分布应用与信息服务研究将是热点。未来的多媒体通信将朝着不受时间、空间和通信对象等方面的任何约束和限制的方向发展,其目标是“任何人,在任何时刻,与任何地点的任何人,进行任何形式的通信”。人们将通过多媒体通信迅速获取大量信息,反过来又以最有效的方式为社会创造更大的社会效益。

(2) 智能处理

利用图像理解、语音识别、全文检索等技术,研究多媒体基于内容的处理,开发能进



行基于内容处理的系统,是多媒体信息管理的重要方向。

(3) 多媒体标准的规范

各类标准的研究建立将有利于产品规范化。以多媒体为核心的信息产业突破了单一行业的限制,涉及诸多行业,而多媒体系统集成特性对标准化提出了更高的要求,所以必须开展标准化研究,它是实现多媒体信息交换和大规模产业化的关键所在。

(4) 多学科交互

多媒体技术与其他技术相结合,提供了完善的人机交互环境。同时多媒体技术将继续向其他领域扩展,并使其应用范围进一步扩大。多媒体仿真、智能多媒体等新技术层出不穷,扩大了原有技术领域的内涵,并不断创造出新的概念。

多媒体技术与外围技术构造的虚拟现实研究仍在继续进展。多媒体虚拟现实与可视化技术需要相互补充,并与语音、图像识别和智能接口等技术相结合,建立高层次虚拟现实系统。

未来多媒体技术将向着以下6个方向发展。

- ① 高分辨化,以提高显示质量。
- ② 高速度化,以缩短处理时间。
- ③ 简单化,便于使用操作。
- ④ 高维化,三维、四维或更高维发展。
- ⑤ 智能化,进一步提高信息识别能力。
- ⑥ 标准化,便于信息交换和资源共享。

多媒体技术正在向自动控制系统、人机交互系统、人工智能系统以及仿真系统等技术领域渗透,所有具有人机界面的技术领域都离不开多媒体技术的支持。这些相关技术在发展过程中创造出许多新的概念,产生了许多新的观点,正在被人们所接受,并成为研究课题之一。

1.1.5 多媒体技术的应用

多媒体技术具有多维性、集成性和交互性的特点,为计算机应用开拓了广阔的前景,目前多媒体技术已经成功应用于以下领域中。

1. 教育与培训

计算机辅助教学(Computer Assisted Instruction,简称CAI)是一种以学生为中心的新型教学模式,是对以教师为中心的传统教学模式的革命。在多媒体技术应用之前,CAI只能靠文字和简单图形来进行人机对话,没有语言、影像,界面单调,缺乏生动形象,限制了CAI优越性的发挥。

多媒体技术将声、文、图集成于一体,使传递的信息更丰富、更直观,这是一种合乎自然的交流环境和方式,人们在这种环境中通过多种感官来接受信息,加速了理解和接受知识信息的学习过程,并有助于接受者的联想和推理等思维活动。

将多媒体技术引入CAI中称为MCAI(Multimedia Computer Assisted Instruction),它是多媒体技术与CAI技术相结合的产物,是一种全新的现代化教学系统。随着多媒体技