

多媒体技术

MULTIMEDIA TECHNOLOGY

黄孝建 / 主编



北京邮电大学出版社

多媒体技术

黄孝建 主编
杨 波 审校

北京邮电大学出版社
·北京·

内 容 提 要

本书的主要内容有:多媒体技术的基本概念;各种媒体信息的表示和编码方法;多媒体终端技术、多媒体数据库技术、多媒体通信技术的基本原理;有关多媒体技术的相关标准以及多媒体通信技术中各主要应用系统的组成和工作原理。通过阅读本书读者可掌握多媒体技术的基本概念和主要技术原理,了解多媒体通信系统的体系结构、相关技术标准和状况。本书的重点是有关多媒体的基本概念与各种媒体表示方法、多媒体通信系统工作原理,难点是有关多媒体信息的压缩、媒体同步概念与同步机制。

本书的读者对象主要是大专院校通信、电子和计算机专业的学生,也可供相关专业的工程技术人员作为培训教材使用。

图书在版编目(CIP)数据

多媒体技术/黄孝建主编. —北京:北京邮电大学出版社,2000.10

ISBN 7-5635-0436-2

I. 多... II. 黄... III. 多媒体技术 IV. TP37

中国版本图书馆 CIP 数据核字(2000)第 52991 号

出 版 者	北京邮电大学出版社(北京市海淀区西土城路 10 号)
	邮编:100876 电话:62282185 62283578
	网址: http://www.buptpress.com
经 销	新华书店北京发行所发行 各地新华书店经售
印 刷	北京源海印刷厂印刷
开 本	787 mm × 1 092 mm 1/16 印张 17.5 字数 402 千字
版 次	2000 年 10 月第 1 版 2000 年 10 月第 1 次印刷
印 数	1—3 000 册
书 号	ISBN 7 - 5635 - 0436 - 2/TN·193
定 价	38.00 元

编者的话

多媒体技术是一门综合的跨学科的交叉技术,它综合了计算机通信以及多种信息科学领域的技术成果,它的研究涉及到计算机硬件、软件和体系结构、图像处理、语音处理、数字信号处理、通信技术等诸多方面技术。

有关多媒体方面的书籍目前具有较大的需求,而且已有许多著作出现,在已出版的各类多媒体方面的书籍中,有的侧重于多媒体的体系结构,理论方面的意义较大;有的偏重于多媒体技术在通信方面的应用;有的主要讲述多媒体计算机及其硬件组成;还有的则重点讲授多媒体应用软件的使用方法。本书编写的主要目的是为北京邮电大学远程教育学院通信专业所开设的多媒体技术课程提供一本适当的教材。本书的特点在于对多媒体技术的各个方面进行了比较全面、系统地介绍,特别是根据先修课程安排的特点加强了对各类媒体信息基础知识的介绍,并努力使该书的内容和教学形式适合于远程教学的需要,尽量做到简单易懂;此外,本书还利用多媒体技术开发一套 CAI 课件,借助多媒体表现形式讲述多媒体技术自身。相信该书对于在大学本科和大专中开展多媒体技术方面的教学有一定的积极作用,特别是具备网络支持功能的 CAI 课件的开发对于该课程的网上教学更是必不可少的组成部分。

本书的主要内容有:多媒体技术的基本概念;各种媒体信息的表示和编码方法;多媒体终端技术、多媒体数据库技术、多媒体通信技术的基本原理;有关多媒体技术的相关标准以及多媒体通信技术中各主要应用系统的组成和工作原理。通过阅读本书读者可掌握多媒体技术的基本概念和主要技术原理,了解多媒体通信系统的体系结构、相关技术标准和状况。本书的重点是有关多媒体的基本概念、各种媒体表示方法以及多媒体通信系统工作原理,难点是有关多媒体信息的压缩、媒体同步概念与同步机制。本书的读者对象主要是大专院校通信、电子和计算机专业的学生,也可供相关专业的工程技术人员作为培训教材使用。

在本书中,王雷编写了第3章第1节、第7章大部分内容,王海英编写了第1章第1节、第2章第1节部分内容,其余内容由黄孝建编写并统编全书,其中张瑞琴为第4、5、6章的编写搜集了大量资料。在本书的内容安排等方面,杨波老师提出了许多好的建议,并对全书进行了审校。在本书的编写过程中,作者参阅了大量的论文和图书资料,在此对上述文献资料的原作者表示衷心的感谢。

由于多媒体技术本身尚在发展过程中,且作者水平有限,不当之处在所难免,敬请指正。

编者

2000年9月

目 录

第 1 章 概论	1
1.1 媒体与多媒体技术的基本概念	1
1.2 多媒体技术的回顾与展望	2
第 2 章 媒体技术基础	9
2.1 视觉媒体技术	9
2.1.1 人类视觉特性与电视技术基础	9
2.1.2 视频信息与视频信息数字化表示	16
2.1.3 视频数字压缩编码技术	20
2.1.4 视频编码技术标准	37
2.2 听觉媒体技术	55
2.2.1 人类听觉特性与电声技术基础	55
2.2.2 音频信息与音频信息数字化表示	61
2.2.3 音频数字压缩编码技术	62
2.2.4 音频编码技术标准	63
2.2.5 MPEG 数字声音压缩编码标准	64
2.2.6 音乐合成	73
2.3 触觉媒体与形体信息技术	75
2.3.1 概述	75
2.3.2 触觉反馈	75
2.3.3 触觉与动作信息输入设备	76
第 3 章 多媒体数据关系	80
3.1 多媒体数据的约束关系	80
3.2 超媒体技术与多媒体信息交换	83
3.2.1 超媒体的概念	83
3.2.2 超媒体系统的组成和原理	86
3.2.3 多媒体信息交换及其标准	90
第 4 章 多媒体终端技术	92
4.1 多媒体终端	92
4.2 多媒体终端软件平台	95
4.3 多媒体终端外部设备接口	104

第 5 章 多媒体信息存储	113
5.1 磁盘存储系统	113
5.2 光盘存储系统	124
5.3 多媒体数据存储策略与 SAN	129
5.4 其他存储设备	133
第 6 章 多媒体信息获取与输出技术	136
6.1 视频信息采集与显示	136
6.2 音频信息采集与播放	141
6.3 图形图像信息输入设备	145
6.4 文字信息输入技术	150
6.5 语音输入输出技术	153
6.6 其他多媒体输入技术	161
第 7 章 多媒体数据库与内容查询	167
7.1 多媒体数据库	167
7.2 基于内容的查询	176
第 8 章 多媒体通信	187
8.1 多媒体通信业务与需求	187
8.2 多媒体通信系统中的同步技术	189
8.3 多媒体通信网与信息传输技术	195
8.3.1 多媒体宽带通信网	195
8.3.2 宽带干线传输技术	197
8.3.3 宽带交换技术	202
8.3.4 宽带接入技术	216
第 9 章 多媒体通信应用系统	237
9.1 多媒体会议系统	237
9.1.1 多媒体会议与协同工作	240
9.1.2 多媒体会议系统结构	241
9.1.3 多媒体会议国际标准	243
9.1.4 各种网络环境下的多媒体会议系统	249
9.2 交互多媒体信息点播系统	257
9.2.1 多媒体信息点播系统的基本概念与分类	257
9.2.2 VOD 系统结构与参考模型	258
9.2.3 视频服务器	263
9.2.4 用户点播终端	265

第1章 概 论

自 80 年代以来,随着电子技术和大规模集成电路技术的发展,计算机、广播电视和通信这三大原来各自独立的领域,相互渗透相互融合,进而形成了一门崭新的技术——多媒体技术,并日益成为人们关注的热点之一。多媒体技术的不断发展给人们的工作、生活和娱乐带来了深刻的革命。

经过十几年的不断摸索和研究,人们对多媒体的认识进一步加深,在有关多媒体的概念、定义、媒体类型、多媒体系统的特征等方面逐渐达成了共识,并推出了大量多媒体应用系统,使之渗透到了人们生活的各个方面,以至于如今的多媒体计算机变成了家用计算机的代名词。在多媒体技术的发展和对其客观规律认识不断加深的过程中,人们同时也发现人体本身就是一个绝妙的多媒体信息处理系统。人类通过视觉、听觉、触觉、嗅觉和味觉来感知外部世界,并获取有关图像、图形、声音、文本、数据等信息,通过神经网络系统传输至大脑进行存储、处理和加工,做出判断后再通过神经网络系统返传到全身各个部位去控制发声器官发出声音、控制肢体做出各种动作等等。多媒体系统在协助人类获取外部信息和对外部世界进行控制的过程中,充分利用了人类的各种感觉、语言和肢体动作,使人与多媒体系统的界面变得越来越友好。本章将对有关媒体、多媒体和多媒体技术等基本概念进行简要介绍,对多媒体技术的发展进行回顾和展望。

1.1 媒体与多媒体技术的基本概念

所谓“媒体”指信息传递和存储的最基本的技术和手段,即信息的载体。由于信息在被人们感觉、加以表示、使之呈现、实现存储或传输的载体上各有不同,因此媒体可分为以下六类。

- 感觉媒体 人们的感觉器官所能感觉到的信息的自然种类。如声音、图形、图像和文本等。
- 表示媒体 被交换的数据类型,它们定义了信息的特性。如用其编码形式加以描述,相应地有语音编码、图像编码和文本编码等表示媒体。
- 呈现媒体 为人们再现信息的物理手段的类型(输出设备),或者指获取信息的物理手段的类型(输入设备)。如显示器、扬声器和打印机等输出类呈现媒体,以及键盘、鼠标器和扫描器等输入类呈现媒体。
- 存储媒体 存储数据物理媒介的类型。如磁盘、光盘和磁带等存储媒体。
- 传输媒体 传输数据物理媒介的类型。如同轴电缆、光缆、双绞线和无线电链路等传输媒体。
- 交换媒体 在系统之间交换数据的手段与类型。它们可以是存储媒体、传输媒体,或者是两者的某种结合。

一般所说的“多媒体”,不仅指多种媒体信息本身,而且指处理和应用多媒体信息的相应技术,因此“多媒体”常被当作“多媒体技术”的同义词。可将其描述为:多媒体技术是一种能同时综合处理多种信息,在这些信息之间建立逻辑联系,使其集成为一个交互式系统的技术。多媒体技术主要用于实时地综合处理声音、文字、图形、图像和视频等信息,是将这多种媒体信息用计算机集成在一起同时进行综合处理,并把它们融合在一起的技术。

多媒体的关键特性在于信息载体的多样性、交互性和集成性。这也是多媒体技术研究中必须解决的主要问题。信息载体的多样性体现在信息采集、传输、处理和呈现的过程中,要涉及到多种表示媒体、传输媒体、存储媒体或呈现媒体,或者多个信源或信宿的交互作用。集成性和交互性在于,所处理的文字、数据、声音、图像、图形等媒体数据是一个有机的整体,而不是一个个“分立”的信息类的简单堆积,多种媒体间无论在时间上还是在空间上都存在着紧密的联系,是具有同步性和协调性的群体。同时,使用者能对信息处理的全过程进行完全有效的控制,并把结果综合地表现出来,而不是单一数据、文字、图形、图像或声音的处理。多媒体系统一般具有如下功能:捕捉、操作、编辑、存储、呈现和通信,用户能够随意控制声音、影像,实现用户和用户之间、用户和计算机之间的数据双向交流的操作环境以及多样性、多变性的学习和展示环境。

1.2 多媒体技术的回顾与展望

多媒体技术是一门综合的跨学科交叉技术,它综合了计算机通信以及多种信息科学领域的技术成果,它的研究涉及到计算机硬件、软件和体系结构、图像处理、语音处理、数字信号处理、通信技术等诸多方面技术。以 Commodore 公司将声音、图像处理能力集于一台计算机之中并构成 Amiga 系统和 Apple 公司用窗口技术设计了多媒体计算机为标志,多媒体技术经历了近 15 年的发展。

促进多媒体技术走向成熟的因素很多,但其中的关键技术涉及多媒体信息的获取与输出技术、多媒体信息压缩技术、多媒体信息的存储技术、多媒体信息的检索技术、多媒体网络技术以及多媒体技术的软硬件支撑平台等。

1. 多媒体信息的获取与输出技术

一般来讲,多媒体信息的获取与输出主要指人机交互的界面,它介于用户和多媒体系统之间,是人与多媒体系统之间传递、交换信息的媒介,是用户使用系统的综合操作环境。通过人机界面,用户向系统提供命令、数据等输入信息,这些信息经系统处理后,又通过人机界面,把产生的输出信息回送给用户。可见人机界面的核心内容包括了信息的显示风格和用户的操作方式,它集中体现了多媒体系统的输入输出功能,以及用户对系统的各个部件进行操作的控制功能。多媒体技术可使用户用十分自然的方式与系统交互,从而改善了人机界面。

在多媒体技术推出之前人机界面并不友好,普遍采用字符界面,用户在屏幕上看到的是各种字符信息,而用户输入计算机的也是字符。用户接收信息的装置主要是字符终端,主要的输入工具是键盘。现在,多媒体终端普遍采用图形图像以及活动视频和声音作为信息输出手段,采用鼠标、跟踪球、电子笔、触摸屏、扫描仪、数字相机、数字摄像头和视音

频采集卡作为输入设备,大大改善了人机接口,使多媒体系统获取信息的手段越来越丰富。随着技术进步,采用语音识别、图形识别和图像理解等先进技术,人机接口将越来越方便人们对多媒体终端设备的使用。

使电脑具有类似于人一样的说话能力和听懂人说话的能力是多媒体技术的一个重要发展方向。语音识别和语音合成技术是实现人机语音通信,建立一个具有听、说能力的智能系统所必需的两项关键技术。目前这两项关键技术均取得了重大进展,推出了大量实用系统。在语音识别方面,自从1997年9月IBM语音识别技术产品、汉语普通话连接语音识别系统ViaVoice问世,在很短的时间内,就有大量语音识别输入产品进入了市场。目前,语音识别输入产品的识别率相当高,有6万条以上的词汇表,覆盖了日常生活的各个方面;此外,20余万条词汇的备份字典提供了更广泛的选择空间。针对行业的特殊需要,语音识别输入产品还可提供加装专业语库的接口,如用于法律和医学的专业语库,大大提高了专业用语识别的正确性。与进行语音输入所用的语音识别技术相对应,利用语音合成技术可实现语音的输出,为人们提供了全新的服务。在银行、邮局、劳保、证券、专利、信访声讯服务系统中,借助文-语转换技术把数据库中的文字变成声音,用户可利用电话收听即时变化的信息。目前,语音技术已成为多媒体领域的研究热点,合成语音的可懂度、清晰度已基本解决,但自然度还不尽如人意,表现力的差距较大,技术的成熟度、应用的广泛性还需进一步提高。

在多媒体信息输入技术中,另一项取得较大进步的是为了克服使用键盘输入的不便而独立发展但又相互关联的一组技术——联机手写识别技术、脱机手写文稿与脱机印刷文稿识别技术,这些技术的核心是文字识别。从用途上讲,脱机手写文稿与脱机印刷文稿识别解决的是手写或印刷的已存在文献的重新输入问题,即纸介质文献的电子化或数字化,目前几乎所有的扫描仪都随机配备了脱机文稿识别软件,识别率完全达到了实用水平。当前手写识别核心算法也已经得到突破,各种手写识别软件结合书写板和笔输入技术基本上解决了手写输入问题。在衡量手写识别系统时的最终标准是书写自由度和识别正确率,在同样的硬件配置下软件占用资源的多寡会反映到识别速度上,识别速度又直接影响到书写的自由度。目前,手写识别软件中连笔字识别率有了很大的提高,放宽了对书写的许多限制,不限笔顺和连笔识别技术开始融合。但识别正确率似乎正在达到一个极限,需要进一步的突破,对用户的书写还是有一定的限制,距真正意义上的自由书写还有一段距离。尤其是在全屏幕连续书写方面,对书写速度、节奏和字间距的限制仍然较为严格。

除了图像、语音和手写文字识别等技术以外,触觉、味觉、手势和面部表情等输入和识别技术也正在起步,多媒体信息的获取方式将更加智能化与人性化。

2. 多媒体数据压缩技术

数据压缩技术作为多媒体技术中最为关键的核心技术,在技术本身和应用方面近年来都取得了引人注目的进展,而其中图像压缩算法更是如此。一般来说,图像编码方法可以分成三类:第一类是考虑图像信源的统计特性,采用预测编码方法、变换编码方法、矢量量化编码方法、子带-小波编码法以及神经网络编码法等,以仙农信息论为基础;第二类根据人眼视觉特性,采用基于方向滤波的图像编码法和基于图像轮廓-纹理的编码法,充

分考虑到了信息接收者的主观特性和主观意义。第三类则考虑到了图像传递的景物特征,采用分形编码方法和基于模型的编码方法。

基于模型的压缩编码方法代表着新一代的压缩编码方向,也是目前最活跃的研究领域,其研究主要沿语义基和物体基两个方向发展。语义基方法基于限定场景,景物中物体的三维模型严格已知,这样只需对一些有限的参数进行编码,压缩比很大。但受目前视觉建模等领域的研究水平所限,语义基编码的研究主要局限于人的头肩像场景,应用于视频电话、视频会议等领域。该方法最大的问题在于可以编码的图像场景单一,在实际应用中极不灵活,只要情况稍有变化,模型就会失效。所以,这种方法只有当计算机视觉研究取得突破性进展之后,才具有普遍意义和实用价值。物体基方法针对未知场景,需要先对图像进行低层分析,将场景中物体和背景分割出来分层描述。常用的描述模型有形状参数、运动参数和内部纹理参数三要素,其中最大的难点在于景物的分割。

目前,国际标准化组织(ISO)、国际电工委员会(IEC)和国际电信联盟(ITU)陆续完成了各种数据压缩与通信的标准和建议。如面向静止图像压缩的 T. 81 及 ISO 10918 (JPEG),在运动图像方面用于视频会议的 H. 261 (Px64)和用于可视电话的 H. 263,针对 CD-ROM 应用传输率为 1.5 Mbit/s 的 MPEG-1,以及针对 DVD、数字电视和高清晰度电视应用,传输率为 2~8 Mbit/s 的 MPEG-2。MPEG 已经制定完成 MPEG-4 的第一、二版,目前正在制定 MPEG-4 的第三、四版和 MPEG-7,并在酝酿 MPEG-21。

在 JPEG, H. 261, H. 263, MPEG-1, MPEG-2 等多媒体压缩编码标准中,主要采用的是以仙农信息论为基础的第一代的数据压缩编码技术,而 MPEG-4 则采用了面向对象、基于内容的压缩编码。MPEG-4 的初衷是针对视频会议、可视电话的超低比特率编码,但现在制定的 MPEG-4 则将目标确定为支持多种多媒体应用,主要侧重于对多媒体信息内容的访问,并可根据应用的不同要求现场配置解码器。这意味着需要将基于内容的检索与编码结合起来考虑,在压缩数据中应有描述视频内容的信息,从而使对多媒体信息内容的访问可以直接针对压缩数据进行,这种压缩编码方法被称作基于内容的压缩方法。新的编码方法可以基于通用芯片,打破了原来压缩编码方法基于专用硬件的限制,可以引入涉及到图像分析的较复杂的算法。因而,MPEG-4 新的编码系统是开放的,可以随时加入新的有效的算法模块。

为了支持对动态视频内容的访问,MPEG-4 中引入了对象(Object)的概念,实际上就是用物体基方法得到的分层区域。因此,物体基方法是基于内容压缩编码方法中的一类主要算法。按照 MPEG-4 的开放性原则,它对具体的算法将不作规定,可以认为有一个工具箱,只要符合某种描述规则的新的好的算法都可以加入。在已知的有关 MPEG-4 的产品方面,已有夏普公司于 1999 年底推出的采用 MPEG-4 编码压缩活动图像、使用 64 MB 智能媒体(Smart Media)存储的第二代网络摄像机;微软公司也在其新近推出的媒体编码器和媒体播放器中采用了基于 MPEG-4 标准的编解码器。这些产品的推出,将会对 MPEG-4 标准的推广利用产生十分重要的影响。

目前,世界各国和各大电子公司均十分重视对视频压缩编码的研究,在已定稿的各种视频压缩编码标准 MPEG-1, MPEG-2, H. 261, H. 263 中,已注册的专利就多达上千个,从目前 MPEG 组织的活动情况看,基于内容的压缩编码方法将是未来压缩编码的主要发展

趋势。

3. 多媒体信息的检索技术

多媒体信息检索就是根据用户的要求,对图形、图像、文本、声音、动画等多媒体信息进行检索,从而得到用户所需的信息。它广泛用于诸多领域,并有着广阔的应用前景,引起了人们的普遍关注。如今,越来越多的信息以数字形式存储和传输,这为人们更灵活地使用这些信息提供了可能性。但随着信息的爆炸性增长,人们获取感兴趣信息的难度却越来越大,而传统的基于关键字或文件名的检索方法显然不适于数据量庞大且不具有天然结构特征的声像数据。因此,近年来多媒体研究的一个热点是声像数据基于内容的检索。实现这种基于内容检索的一个关键性的步骤是要定义一种描述声像信息内容的格式,而这与声像信息的编码密切相关。MPEG 组织注意到了这方面的需求和潜在的市场,在尚未完成 MPEG-4 最后定稿的情况下就已开始着手制定专门支持基于内容检索的多媒体信息编码方案 MPEG-7。

MPEG-7 作为 MPEG 系列标准中的一个新成员,正式名称叫作“多媒体内容描述接口”,其目的是为各种类型的多媒体信息规定一种标准化的描述,这种描述与多媒体信息的内容本身一起,支持用户对其感兴趣的各種信息进行快速有效的检索。在这些信息当中包括静止图像、图形、音频、动态视频,以及如何将这些元素组合在一起的合成信息。

这种对信息的标准化描述可以加到任何类型的多媒体数据上。不管多媒体数据的表达格式或压缩形式如何,只要加上了这种标准化描述的多媒体数据就可以被索引和检索。各种类型信息的标准化描述可以分成一些语义上的层次。以视频资料为例,较低层次的描述是颜色、形状、纹理、空间结构等信息;而最高层次的语义描述信息可以是“画面中有一栋白色的小楼在左边,而一棵苹果树在小楼的右前方”。也可以有介于上面两种层次之间的中层语义描述信息。同样的内容根据不同的应用领域要求,可以携带不同类型的描述信息。MPEG-7 标准的制定将有助于数字化图书馆、多媒体目录服务、广播式媒体选择以及个人电子新闻服务、媒体著作等应用领域的发展。

目前正在酝酿的 MPEG-21,是一个多媒体的应用框架,而并不是一个单纯的视音频编码标准。考虑到目前标准越来越多,媒体的应用越来越广泛,为了规范这些标准与内容传输间的关系(如底层部件与基本结构的关系、它们之间的集成方式等),并达到一个关于多媒体框架的共识,支持电子内容的传输,就需要一个更大的框架来描述与内容传输有关的所有已制定的和正在开发中的标准化活动。实际上要真正地完成不同标准之间的集成,要从用户的角度来建立一个可协同操作的多媒体内容传输服务机制,就要求用新的综合的解决方法来管理这些不同类型内容的传输过程。媒体的服务对用户应是全部透明的,如查找并获得内容及服务质量保证等。

MPEG-21 正是在这种需求下提出的一个多媒体框架,用以支持新的网络化的数字市场与商业模型。这样的框架要求有共享的视觉内容,并保证系统传递电子内容时能尽可能是自动地、简单地协同操作。这一多媒体框架将用以满足对内容传递、内容安全、权限管理、安全付款以及技术授予权等下层结构的要求。MPEG-21 的覆盖范围可描述为两项技术的集成,即用户如何通过个人或智能代理来找到并获得内容,以及内容如何根据用户与内容的权限来呈现给用户。一旦该标准形成,未来的多媒体应用将能非常方便地协同

操作不同标准间的媒体信息。

4. 多媒体信息存储技术

多媒体信息有两个显著的特点,一是信息的存在和表现有多种形式,且信息量很大,尤其对动态的声音视频图像更为明显;二是多媒体信息传输具有实时性,声音和视频必须严格地同步。多媒体的这两个特点给存储系统提出了很高的要求,即存储设备的存储容量必须足够大,以满足多媒体信息的存储要求;存储设备速度要快,要有足够的带宽,以便高速传输数据,使得多媒体信息能够实时地传输和显示。

多媒体信息存储可采用的介质多种多样,有硬盘、光盘、磁带等。目前新型增强 IDE 接口已具有 9~66 MB/s 的传输速率,硬盘最大容量达到 37 GB;SCSI 硬盘的传输速率已达 160 MB/s。CD-ROM, CD-R, CD-WR, DVD, DVD-RAM 等各类光盘的推出,更为大容量多媒体数据的存储提供了极为丰富的手段。随着多媒体与计算机技术的发展,多媒体数据库的存储容量越来越大,对存储设备的要求越来越高。适合于多媒体数据存储的方式主要有硬盘、RAID(磁盘冗余阵列)、光盘库以及存储区域网络等。

5. 多媒体网络技术

在多媒体技术的发展过程中,网络技术占有极为重要的地位。正是由于宽带网络技术的快速发展,才使多媒体技术从独立的终端应用模式,走向大规模的互联网应用模式,使人们可以在世界范围内不仅仅通过语音,而且可以通过多媒体方式进行沟通。在多媒体网络技术中涉及信息的宽带干线传输技术、宽带交换技术以及用户的宽带接入技术。

(1) 宽带干线传输技术

伴随着光纤通信技术的发展,宽带网络传输的主要物理介质由电缆和微波逐渐过渡到了光纤。利用光纤组建宽带传输网的优势在于通过复用技术和数据压缩技术完全能够适应多媒体通信对传输速率与容量的各种要求。光纤通信向大容量、高速发展,为多媒体宽带网提供了坚实的基础。在宽带干线传输技术中,取得较大成功的技术当属同步数字系列传输体制 SDH 的确立,而最有发展前景的技术则为密集波分复用(DWDM, Dense Wave Division Multiplexing)。

SDH 传输体制作为一种完整严密的传送网技术体制,有着全世界统一的网络节点接口,简化了信号的互通以及信号的传输、复用和交叉连接过程;它安排有丰富的开销比特用于网络的管理和维护;有统一的标准光接口,能够在基本光缆段上实现横向兼容;采用 SDH 组网技术还可以构成具有高度可靠性的自愈环结构,确保了实现业务的透明性,这对多媒体业务应用十分重要。

波分复用 WDM 是在一根光纤中能同时传输多个波长光信号的一种技术,其基本原理是在发送端将不同波长的光信号组合(复用),在接收端又将组合的光信号分开(解复用)并送入不同的终端。采用波分复用技术后,原来只能采用一个光波长作为载波的单一光信道变为数个不同波长的光信道同时在光纤中传输,从而光纤通信容量成倍提高。此外,利用 WDM 技术还可以实现单纤全双工传输,在光纤用户网中增加组网的灵活性。目前采用密集波分复用技术的高速传输系统,其产品通信速率已达 400 Gbit/s,实验室的研究水平已超过 3 Tbit/s。WDM 技术对网络的扩容升级、发展宽带多媒体业务、充分挖掘和利用光纤带宽能力、实现超高速通信等具有十分重要的意义。

(2) 宽带交换技术

在多媒体宽带通信网中交换技术起着举足轻重的作用,它不仅可使多台通信终端共享传输媒体,而且完成网络中的任意两个或多个用户的相互连接。通信网中交换方式的确立,决定了网络的总体运行方式和网络性能,从而也就对用户终端的类型和接入方式提出了相应的要求。当传输技术逐渐向宽带迈进的时候,使各种不同业务可以充分地利用传输资源的交换设备也有了与之相适应的发展,以 IP 和 ATM 为代表的分组转发和交换技术是当前网络建设中的热点。IP 的灵活特性和 ATM 的快速交换能力互有千秋,必将在将来的网络技术中起到关键的作用。

ATM 技术是以分组交换模式为基础,同时融合了电路传送方式的优点发展而成的。它克服了电路交换方式不适应速率变化的缺点,并简化分组通信中的协议,由硬件完成对简化通信协议的处理,各交换节点不再对信息进行差错控制,提高了通信处理能力。它既继承了电路交换方式的速率独立性,又保留了高速分组交换方式的速率任意性,是目前唯一能够全面支持 QoS 的传输交换技术,成为下一代通信网的基础技术之一。

IP 技术是一种非面向连接的分组/包交换网络技术,它可以较容易地集成语音、数据、图像和视频多媒体业务。它对通信资源的利用率远远高于传统的基于电路交换的通信网络技术,通信费用也低得多。Internet 的迅猛发展使得 IP 网络目前已遍及到社会的每一个角落,有超越传统的以电路交换为主的电信网、成为未来信息高速公路基础的趋势。作为 IP 技术典型代表的以太网一直以来被作为局域网的代名词,但交换式以太网的出现、千兆以太网的发展以及 IP QoS 的提出已经把以太网技术和多媒体应用紧密地联系在一起。交换式以太网把多个终端共享 10 Mbit/s 带宽升级到独占 10 Mbit/s, 100 Mbit/s 甚至 Gbit/s 带宽。以太网和交换技术的结合已经彻底消除了传输带宽对多媒体应用的限制。

尽管现有的通信系统都采用电子交换方式,但随着通信网络逐渐向全光平台发展,未来的全光网络却需要由纯光交换机来完成信号路由功能以实现网络的高速率和协议透明性。由于光网络容量持续扩展,而电交换机不适应超过 Gbit 速率的要求,开发高速高性能的交换机就成为必然的趋势。光交换机可以保证网络的可靠性和提供灵活的信号路由平台。目前已经出现了许多种商用光交换机,如光电交换机、光机械交换机以及基于热学、液晶、声学 and 微机电(MEM: Micro-Electro-Mechanical)技术的光交换机。光交换机将在新兴的多通路、可重新配置的光子网络中发挥越来越重要的作用。当出现更有效的信号管理方式时,全光网络最终会变成事实。在未来的大容量光网络中,光交换机必将起到关键的作用。

(3) 宽带接入技术

宽带接入网是多媒体通信网的基础,是解决用户接入多媒体网络最重要、也最复杂的部分。目前,宽带接入技术主要有:基于电信网用户线的数字用户线(DSL)接入技术、基于 HFC 网络的电缆调制解调器(Cable Modem)接入技术、宽带光纤接入技术、基于以太网的局域网接入技术以及基于无线传输手段的无线接入技术。上述各种接入技术各有优缺点,它们互相竞争、互为补充,共同构成了用户访问多媒体网络的接入平台。

6. 多媒体技术的支撑环境

多媒体技术的发展,离不开软硬件技术的支撑。计算技术的迅猛发展为多媒体技术提供了良好的技术支持。以前以通用计算机为平台的多媒体系统大多采用专用硬件实现多媒体扩展,用以解决计算机在处理实时视音频数据时的速度和接口问题。如今,随着 Intel 公司和 AMD 公司相继推出主频在 1 GHz 以上的 CPU 芯片、高速串行总线 USB2.0 和 IEEE1394 的问世,使得完全用软件实现多媒体信息处理成为可能。在不远的时间内,任何一台通用计算机在不添加任何其他硬件插卡的情况下,即已具备多媒体信息的处理能力。声卡和视频采集卡这些多媒体计算机的标志性硬件设备,也许很快将成为历史。

除此以外,操作系统、面向对象技术、并行处理和复杂结构的分布处理技术等软件平台的发展也为多媒体的发展提供了很好的基础。如今,几乎所有新推出的操作系统、www 服务器以及 www 浏览器等均具备了对多媒体信息包括对流式媒体等的支持能力。此外,还出现了大量音频、图形与图像、动画和视频等多媒体素材制作软件,如 WaveStudio, Corel-DRAW, Photoshop, Animator Studio, 3Dstudio MAX, Premiere 以及多媒体著作工具,如 Authorware, ToolBook 等。

目前,多媒体技术已经在日常生活的各个方面得到了广泛的应用,如各种电子出版物、多媒体家电、多媒体辅助工程、电子商务、多媒体数据库、数字化图书馆、多媒体会议系统、家庭影院、虚拟演播室、多媒体非线性编辑系统等。在未来若干年内,多媒体技术的研究将集中于巨量多媒体数据的存储与检索、宽带大容量多媒体数据的传输与交换、多媒体信息系统的服务质量保障、智能数据复制、更加友好和智能化的下一代用户界面以及多媒体系统集成框架的建立。

思考题

- 1.1 媒体共分几类? 多媒体主要有什么特点?
- 1.2 多媒体技术研究的主要课题是什么?
- 1.3 多媒体技术涉及哪些主要关键技术?

第2章 媒体技术基础

2.1 视觉媒体技术

2.1.1 人类视觉特性与电视技术基础

视觉媒体直接服务于人类的眼睛,为适应人类视觉的需求,需根据人的视觉特性制定合适的指标,并根据人类的视觉特性研究图像压缩编码的方法。因此,有必要首先研究有关人类视觉的相关特性。

1. 人类视觉特性

(1) 图像的对比度与视觉的对比度灵敏度特性

图像的对比度表示图像两相邻面积或相邻点之间的亮度差别。人眼区分某一给定空间频率的正弦光栅明暗差别所需要的最低对比度,称为辨别这一空间频率的临界对比度 C_r ,临界对比度的倒数 $1/C_r$,叫做人眼对于这一空间频率的对比度灵敏度。临界对比度表示人眼在给定的亮度环境下所能区分的景物中的最小亮度差别,这一最小亮度差别称为一个亮度级(灰度级)。

在什么样的亮度范围内人眼能区分景物的对比度变化,以及在某一给定的亮度下,眼睛能够区分多少个对比度等级和差别,这些确定了人眼的亮度适应特性。人眼能够区分对比度变化的背景亮度范围约为 $10^{-4} \sim 10^4 l_x$ 。

(2) 空间频率与视觉的空间频率响应

人在观看景物时,对细节的分辨能力称为人眼的分辨力,人眼的分辨力究竟有多高,用什么方式来表示人眼的分辨力,这是图像传输技术中的一个研究课题。如果图像系统传送的图像细节高于人眼的分辨力,这无疑是一种浪费;反之,如果传送的图像清晰度太低,则达不到人类视觉的满意度。因此有必要首先搞清人眼的分辨力。人们经过科学探索发现,将人眼等效为一个空间频率的滤波器,用滤波器频率特性可以较为客观全面地表示人眼的分辨力。因此对人眼分辨力的研究,就集中于求这一空间滤波器的空间频率特性。

我们知道,时间频率是用单位时间内某物理量(电压、电流、波动等)周期性变化的次数来定义的,单位为周每秒。类似的,在不考虑电磁辐射的传播方向时,空间频率是指物理量(如亮度)在单位空间距离内周期性变化的次数,单位为周每米。即

$$f_{xm} = \frac{d\varphi(x)}{dx}$$

f_{xm} ——表示亮度信号在 x 方向的空间频率, x 为空间变量, $\varphi(x)$ 代表亮度信号在 x 方向上的相位变化。

分析亮度信号的空间频率成份,与分析随时间变化的电信号的时间频率成份方法完

全相同,都是通过傅氏变换来进行的。人眼分辨特性的空间频率响应如图 2-1 所示。

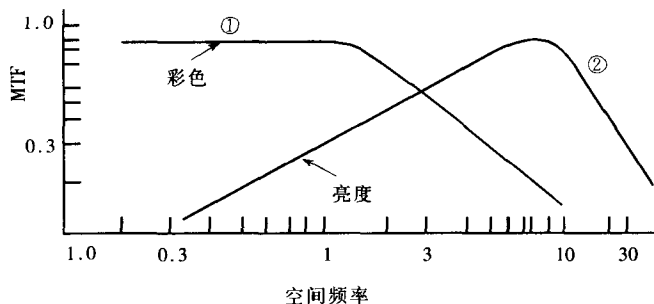


图 2-1 人眼的空间频率响应

图中曲线②和①分别表示人眼对于图像亮度细节和彩色细节的分辨能力。曲线的纵坐标为调制传输函数(MTF),调制传输函数的含意与电子网络中的幅频响应等效,只是MTF代表空间频率域的传输特性,以空间频率为自变量。

由图中可见,人眼对彩色的空间分辨率要低于对亮度的空间分辨率,这即是我们能够对图像的亮度分量和彩色分量采用不同带宽处理的依据。

(3) 视觉的时间域响应

人眼的视觉是有惰性的,也就是人眼的主观亮度感觉与光的作用时间有关。光像一旦在视网膜上形成,在它消失后,视觉将会对这个光像的感觉维持一个有限的时间,这也叫做视觉暂留。视觉惰性事实是近代电影和电视的基础,电影和电视正是利用眼睛的视觉暂留特性,使得本来在时间上和空间上都不连续的图像给人以真实的连续的感觉。

如果让观察者观察一幅按时间顺序重复的亮度脉冲,当脉冲重复频率不够高时,人眼就有一亮一暗的感觉,这称之为闪烁,当重复频率足够高时,闪烁感觉消失,看到的则是一个恒定的亮点。闪烁感觉刚好消失时的频率叫做临界闪烁频率,脉冲的亮度越高,临界闪烁频率也相应地提高。

在通常的电影银幕的亮度下,人眼的临界闪烁频率约为 46 Hz,因此在电影中,普遍采用每秒钟向银幕上投射 24 幅画面的标准,而在每幅画面停留的时间内,用一个机械遮光阀将投射光遮挡一次得到每秒 48 次的重复频率,使观众产生连续的,不闪烁的感觉。同时要保证图像中动作的连续性,则要求每秒钟不少于 15 ~ 20 幅画面。

(4) 彩色的计量与彩色视觉

根据解剖实验证实,人眼视网膜是由大量光敏细胞组成的,按其形状分为杆状细胞和锥状细胞。杆状细胞只能感光,而锥状细胞既能感光又能感色。在感光能力上,杆状细胞对弱光的灵敏度比锥状细胞大得多,而锥状细胞则在正常光照条件下才能引起视觉并产生色感。另外,锥状细胞又是由 3 种不同的细胞组成的,它们分别对红、绿、蓝光产生视觉,并具有各自的光谱灵敏度特性。

通过实验测试能够得到人眼 3 种锥状细胞的光谱灵敏度曲线如图 2-2 所示。若把图中对应于每一波长的 3 条曲线上的数值相加,就得到人眼的视觉灵敏度曲线(如图中虚线

所示),这证明了人眼的亮度感觉是由红、绿和蓝3种彩色视觉合成的综合结果。

自然界彩色成千上万,日常生活中我们有很多的名词来描绘它,如紫、红、青、蓝、绿等等,但都无法精确地描述某一种彩色,如红色则有浅红、深红、粉红、紫红等,即使这样也仍不能达到准确。对于自然界中这数以千计的彩色种类,要想准确地表达它们,只有借助于数学坐标,这也就是色度学所要做的工作。人们提出了多种彩色度量的坐标,如三维矢量空间、麦克斯韦三角形、RGB坐标、XYZ坐标、UCS坐标等多种不同的坐标系统,在利用这些系统进行色度计算时,不同的条件下各有其不同的方便之处。

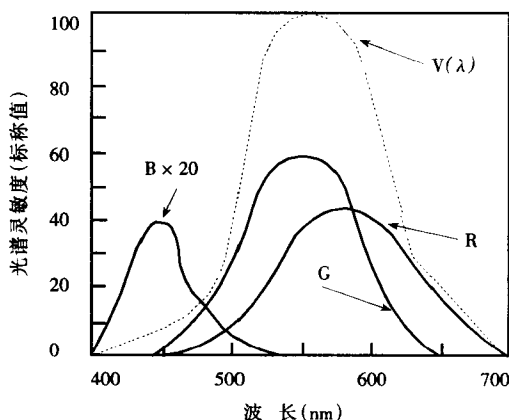


图 2-2 人眼 3 种锥状细胞的光谱灵敏度曲线

2. 电视技术基础

电视是利用光电和电光转换原理,将光学图像转换为电信号进行远距离传输,然后再还原为光图像的一门技术。整个电视系统主要由成像、电视信号形成、信号处理、传输系统、电视信号接收与显示等部分组成。电视技术涉及微电子学、光度和色度学、视觉心理学、通信理论、数字信号与数字图像处理、压缩编码理论等许多学科。

自从 1936 年英国首先开始电子式的黑白电视广播以来,电视技术经历了黑白电视、彩色电视、数字与高清晰度电视的发展过程。同时,电视也具有了无线广播、有线传播、卫星直播和双向通信等多种传输手段。广播电视经过几十年的发展逐渐形成了目前国际上的三大模拟彩色电视制式: PAL, NTSC 和 SECAM。

(1) 电视信号与图像扫描

电视技术是利用光电转换原理实现光学图像到电视信号变换的,这一转换过程通常是在摄像机中完成的。当被摄景物通过摄像机镜头成像在摄像管的光电导层时,光电靶上不同点随照度不同激励出数目不等的光电子,从而引起不同的附加光电导,产生不同的电位起伏,形成与光像对应的电图像。该电图像必须经过扫描才能形成可以被处理和传输的电视信号。

客观景物图像对于人眼的感觉来说,可以认为是由很多有限大小的像素组成的。对于观察者来说,每一个像素都有它的光学特性和空间位置,并且随时间变化而变化。景物图像中的任意一个像素 P 均可以用 8 个物理量来表达,即