

网络多媒体中的 关键技术及应用研究

WANGLUO DUOMEITI ZHONG DE
GUANJIAN JISHU JI YINGYONG YANJIU

冯明 亓慧 黎瑞成 编著



中国水利水电出版社
www.waterpub.com.cn

网络多媒体中的 关键技术及应用研究

冯明 亓慧 黎瑞成 编著



中国水利水电出版社
www.waterpub.com.cn

内 容 提 要

本书主要内容包括网络多媒体技术概论、多媒体数据压缩技术、多媒体通信技术、数字音频处理技术、图形图像处理技术、动画处理技术、数字视频处理技术、流媒体技术、网络多媒体技术综合应用等。

本书在内容编排上既包括基本理论部分,又融合了多媒体实用技术,同时对某些领域的前沿技术、热点问题也有所涉及,对在多媒体领域从事研究工作的研究人员具有很重要的参考价值。

图书在版编目(CIP)数据

网络多媒体中的关键技术及应用研究/冯明,亓慧,
黎瑞成编著. —北京:中国水利水电出版社,2015.7
ISBN 978-7-5170-3168-0

I. ①网… II. ①冯… ②亓… ③黎… III. ①计算机
网络—多媒体技术—研究 IV. ①TP37

中国版本图书馆 CIP 数据核字(2015)第 101681 号

策划编辑:杨庆川 责任编辑:陈 洁 封面设计:崔 蕾

书 名	网络多媒体中的关键技术及应用研究
作 者	冯 明 亓 慧 黎瑞成 编著
出版发行	中国水利水电出版社 (北京市海淀区玉渊潭南路 1 号 D 座 100038) 网址:www.waterpub.com.cn E-mail:mchannel@263.net(万水) sales@waterpub.com.cn
经 售	电话:(010)68367658(发行部)、82562819(万水) 北京科水图书销售中心(零售) 电话:(010)88383994、63202643、68545874 全国各地新华书店和相关出版物销售网点
排 版	北京鑫海胜蓝数码科技有限公司
印 刷	三河市天润建兴印务有限公司
规 格	184mm×260mm 16 开本 17.25 印张 420 千字
版 次	2015 年 8 月第 1 版 2015 年 8 月第 1 次印刷
印 数	0001—3000 册
定 价	60.00 元

凡购买我社图书,如有缺页、倒页、脱页的,本社发行部负责调换

版权所有·侵权必究

前 言

多媒体技术是指利用网络把文字、音频、视频、图形、图像、动画等多媒体信息通过计算机进行数字化采集、获取、压缩/解压缩、编辑、存储等加工处理,再以单独或合成形式表现出来的一体化技术。网络多媒体技术则是一门综合的、跨学科的技术,它综合了通信技术、计算机技术、多媒体技术等信息科学领域的技术成果,目前已成为发展最快且最富有活力的高新技术之一。我国网络多媒体技术出现于 20 世纪 90 年代,属于一个新生事物。近年来,随着信息技术的迅速发展,网络多媒体已逐渐被人们所关注,成为当代计算机网络发展的新成就。

网络多媒体技术的出现标志着计算机将不仅仅作为办公室和实验室的专用品,而将进入家庭、商业、旅游、娱乐、教育乃至艺术等几乎所有的社会与生活领域。由于网络多媒体技术具有很强的实用价值,因而在计算机科学及其相关学科领域得到发展和融合,并开拓了计算机在国民经济领域的广泛应用,而且渗透到生活的各个方面,对社会、经济产生了重大的影响,并已成为国际学术界的一个研究热点。本书作者始终高度关注网络多媒体技术的发展,并一直致力于该领域的教学与研究工作。

本书共分 9 章,从相关应用技术的角度对网络多媒体技术及应用展开讨论。第 1 章网络多媒体技术概论,涉及内容:媒体元素、发展历程、应用领域、社会意义;第 2 章多媒体数据压缩技术,涉及内容:压缩性能指标、无损压缩、有损压缩、音频压缩、图像压缩;第 3 章多媒体通信技术,涉及内容:多媒体通信服务质量 QoS、通信网络、通信协议、无线多媒体技术;第 4 章数字音频处理技术,涉及内容:数字音频编码、获取与处理技术、识别与合成技术,以及常用的数字音频处理软件——GoldWave;第 5 章图形图像处理技术,涉及内容:图像的文件格式、数值描述、图形图像的获取与处理,以及典型的图像处理软件——Photoshop;第 6 章动画处理技术,涉及内容:动画的文件格式、动画的制作,以及常用动画制作软件——Flash;第 7 章数字视频处理技术,涉及内容:数字视频的文件格式、数字视频的获取与处理,以及常用的数字视频编辑软件——Adobe Premiere;第 8 章流媒体技术,涉及内容:流媒体的文件格式、系统组成、实现的关键技术、传输协议、应用领域及应用系统;第 9 章网络多媒体技术综合应用,涉及内容:网络多媒体应用系统的设计、创作、应用,典型应用案例包括多媒体视频会议系统、视频点播系统、交互电视系统、多媒体 IP 电话系统、多媒体远程教育系统、远程医疗系统、视频监控系统。

全书由冯明、亓慧、黎瑞成撰写,具体分工如下:

第 3 章、第 4 章第 1 节~第 2 节、第 5 章、第 8 章:冯明(雅安职业技术学院);

第 6 章、第 7 章、第 9 章:亓慧(太原师范学院);

第 1 章、第 2 章、第 4 章第 3 节~第 5 节:黎瑞成(海南软件职业技术学院)。

本书是一个完整的系统,内容上既包括原理与技术基础,又融合了网络多媒体方面的实用技术,同时对该领域的一些前沿技术、热点问题也有所涉及。可以说是作者多年来从事网络多媒体技术教学和实践的总结。本书的编撰得到了许多同行专家的支持和帮助,在此表示深深的谢意。同时,在本书的编撰过程中,参考了大量有价值的文献与资料,吸取了许多人的宝贵经验,在此向这些文献的作者表示敬意。

网络多媒体技术是一门综合性很强的技术,其发展速度也相当迅速,相关研究资料更是浩如烟海,但是由于作者视野所限,加之时间仓促,书中难免存在不足与错误之处,恳请广大读者进行批评指正,不胜感激!

作者

2015年3月

目 录

前 言

第 1 章 网络多媒体技术概论	1
1.1 网络多媒体技术概述	1
1.2 网络多媒体技术的发展	6
1.3 网络多媒体处理技术的应用领域	9
第 2 章 多媒体数据压缩技术	13
2.1 数据压缩概述	13
2.2 无损压缩和有损压缩	15
2.3 多媒体数据的压缩及其标准	29
第 3 章 多媒体通信技术	54
3.1 多媒体通信概述	54
3.2 多媒体通信的服务质量 QoS	59
3.3 多媒体通信网络	63
3.4 多媒体通信协议	80
3.5 无线多媒体技术	88
第 4 章 数字音频处理技术	90
4.1 数字音频的基础知识	90
4.2 数字音频编码与文件格式	93
4.3 数字音频的获取与处理	101
4.4 数字语音的识别与合成技术	107
4.5 数字音频处理常用软件	110
第 5 章 图形图像处理技术	120
5.1 图形图像的基础知识	120
5.2 图像的文件格式	126
5.3 图形图像的获取与处理	131
5.4 图形图像处理常用软件	138

第 6 章 动画处理技术	148
6.1 动画的基础知识	148
6.2 动画的文件格式	155
6.3 动画的制作	157
6.4 动画制作常用软件	164
第 7 章 数字视频处理技术	172
7.1 视频的基础知识	172
7.2 数字视频的文件格式	179
7.3 数字视频的获取与处理	181
7.4 数字视频编辑常用软件	189
第 8 章 流媒体技术	198
8.1 流媒体概述	198
8.2 流媒体系统的组成	202
8.3 流媒体实现的关键技术	204
8.4 流媒体传输协议	207
8.5 流媒体技术的应用	218
第 9 章 网络多媒体技术综合应用	230
9.1 网络多媒体应用概述	230
9.2 多媒体视频会议系统	240
9.3 视频点播和交互电视系统	243
9.4 其他多媒体应用系统	249
参考文献	268

第1章 网络多媒体技术概论

多媒体技术是一项新兴技术,是涉及计算机、电子、通信网络、数字压缩技术等多门学科的综合学科,也是信息领域发展极为迅速、时代特征极其鲜明的一门多学科交叉的技术,还是新世纪信息产业的一项重大工程技术。网络多媒体技术的高速发展,给人们的生活和工作方式带来了巨大的变化,在信息社会中具有十分重要的地位。多媒体技术发展前景非常广阔,从普及计算机应用、拓宽计算机处理信息的类型看,利用多媒体是计算机技术发展的必然趋势。

1.1 网络多媒体技术概述

“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”。到目前为止,多媒体还没有一个统一而严格的定义。不同的研究目的从不同角度对多媒体给出了不同的解释。目前,多媒体一般被看作是人类运用先进的计算机技术交互处理多媒体信息(其中特别指传统计算机无法处理的音频、视频信息)的方法和手段,例如获取、编辑、存储、传输、展示等。从更加广义的角度看,“多媒体”指的是一个领域,一个和信息处理有关的,包括家用电器、通信、出版、娱乐等在内的所有技术与方法进一步发展的领域。

1.1.1 多媒体中的媒体元素

多媒体中的媒体元素是指多媒体应用中可显示给用户的媒体形式,主要有文本、图形、图像、声音、视频和动画等。这些媒体元素有各自的特点和性质,不同类型的媒体元素有机地结合与互补,才能充分发挥多媒体集成的优势。

(1) 文本

文本是指各种文字,包括各种字体、尺寸、格式及色彩的文本。在多媒体应用系统中适当地组织使用文字可以使显示的信息更容易理解。多媒体应用中使用较多的是带有段落格式、字体格式、边框等格式信息的文字。这些文字可以先使用文本编辑软件(如 Word、WPS 等),或使用图形图像制作软件将文字编辑处理成图片,再输入到多媒体应用程序中,也可以直接在多媒体创作软件中进行制作。

(2) 图形

图形是指用计算机绘图软件绘制的从点、线、面到三维空间的各种有规则的图形,如直线、矩形、圆、多边形,以及其他可用角度、坐标和距离来表示的几何图形。

在图形文件中只记录生成图的算法和图上的某些特征点,因此也称矢量图。通过读取这些指令并将其转换为屏幕上所显示的形状和颜色而生成图形的软件通常称为绘图程序。在计算机还原输出时,相邻的特征点之间用特定的诸多段小直线连接就形成曲线。若曲线是一条封闭的图形,也可靠着色算法来填充颜色。

图形的最大优点在于可以分别控制处理图中的各个部分,如在屏幕上移动、旋转、放大、缩小、扭曲而不失真,不同的物体还可在屏幕上重叠并保持各自的特性,必要时仍可分开。因此,图形主要用于表示线框型的图画、工程制图、美术字等。绝大多数 CAD 和 3D 造型软件使用矢量图形来作为基本图形存储格式。

微机上常用的矢量图形有 .3ds(用于 3D 造型)、.dxf(用于 CAD)、.wmf(用于桌面出版)等。图形技术的关键是图形的制作和再现,图形只保存算法和特征点,所以,相对于图像的大数据量来说,它占用的存储空间也就较小,但在屏幕每次显示时,它都需要经过重新计算。另外,在打印输出和放大时,图形的质量较高。

(3) 图像

图像是指由输入设备捕捉的实际场景画面,或以数字化形式存储的任意画面。静止的图像可用矩阵点阵图来描述,矩阵的每个点称为像素(pixel),整幅图像就是由一些排成行列的像素点组成的,故图像也称为位图。

位图中的位用来定义图中每个像素点的颜色和亮度。对于灰度图常用 4 位(16 种灰度等级)或 8 位(256 种灰度等级)表示该点的亮度。若是彩色图像,R(红)、G(绿)、B(蓝)三基色每色量化 8 位,则称彩色深度为 24 位,可以组合成 2^{24} 种色彩等级,即真彩色。若只是黑白图像,每个像素点只用 1 位表示,则称为二值图。

位图图像适合于表现比较细致,层次和色彩比较丰富,包含大量细节的图像,如自然景观、人物等。由像素矩阵组成的图像可用画位图的软件(如 Photoshop)获得,也可用彩色扫描仪扫描照片或图片来获得,还可用摄像机、数字照相机拍摄或帧捕捉设备获得数字化帧画面。

图形与图像在多媒体中是两个不同的概念,其主要区别如下:

1) 构造原理不同:图像的基本元素是图元,如线、点、面等元素;图像的基本元素是像素,一幅位图图像可考虑为由一个个像素点组成的矩阵。

2) 数据记录方式不同:图形存储的是画图的函数;图像存储的则是像素的位置信息和颜色信息,以及灰度信息。

3) 处理操作不同:矢量图形由运算关系支配,因此,可以分别控制、处理图中的各个部分,如在屏幕上移动、旋转、放大、缩小、扭曲而不失真。图像像素点之间无内在联系,所以在放大与缩小时,部分像素点会丢失或被重复添加而导致图像的失真。

4) 处理显示速度不同:图形的显示过程是根据图元顺序进行的,它使用专门软件将描述图形的指令转换成屏幕上的形状和颜色,其产生需要计算时间。图像是将对象以一定的分辨率解像以后将每个点的信息以数字化方式呈现,可直接快速在屏幕上显示。

5) 数据量不同:图像数据量大,不便于保存和传送,因此,要采用数据压缩算法。图像数据量则相对较小。

6) 表现力不同:图形描述轮廓不很复杂,色彩不很丰富的对象,如几何图形、工程图纸、CAD、3D 造型软件等。图像能表现含有大量细节(如明暗变化、场景复杂、轮廓色彩丰富等)的对象,如照片、绘图等。通过图像软件可进行复杂图像的处理以得到更清晰的图像或产生特殊效果。

随着计算机技术的进步,图形与图像之间的界限已越来越小,这主要是由于计算机处理能力的提高。

(4)声音

声音也叫音频,是指在 20 Hz~20 kHz 频率范围的连续变化的声波信号。对声音可进行录制、存储、播放与合成。音频是数字化音频文件,包括波形声音、语音和音乐。

1)波形声音。波形声音即指数字化的声音,它表示了声音的瞬时幅度,存储的数据文件占有较大的空间。其存储文件格式主要有 WAV 文件、VOC 文件等。

2)语音。人的声音不仅是一种波形,而且还有内在的语言、语音学的内涵,可以利用特殊的方法进行抽取,通常把它也作为一种媒体。

3)音乐。音乐是符号化了的声音,这种符号就是乐曲。MIDI(Musical Instrument Digital Interface)是数字音乐的国际标准。MIDI 文件是其存储的文件格式。

MIDI 信息实际上是一段音乐的描述,当 MIDI 信息通过一个音乐或声音合成器进行播放时,该合成器对一系列的 MIDI 信息进行解释,产生出相应的一段音乐或声音。MIDI 文件紧凑,所占用空间小。通常,MIDI 文件是数字化声音文件的 1/1000~1/200。

(5)视频

视频是由单独的画面序列组成,这些画面以每秒超过 24 帧的速率连续地投射在屏幕上,使观察者产生平滑连续的视觉效果。计算机中的视频信息是数字的,可以通过视频卡将模拟视频信号转变成数字视频信号,进行压缩,存储到计算机中。播放视频时,通过硬件设备和软件将压缩的视频文件进行解压。视频标准主要有 NTSC 制和 PAL 制两种。NTSC 标准为 30 fps,每帧 525 行。PAL 标准为 25 fps,每帧 625 行。常用视频文件格式有 AVI、MPG、MOV 等。

(6)动画

动画是采用计算机动画设计软件创作,由若干幅图像进行连续播放而产生的具有运动感觉的连续画面。动画的连续播放既指时间上的连续,也指图像内容上的连续,即播放的相邻两幅图像之间内容相差不大。动画压缩和快速播放也是动画技术要解决的重要问题,其处理方法有多种。计算机设计动画方法有以下两种:

1)造型动画:是对每一个运动的物体分别进行设计,赋予每个对象一些特征,如大小、形状、颜色等,然后用这些对象构成完整的帧画面。造型动画每帧由图形、声音、文字、调色板等造型元素组成,控制动画中每一帧的图元表演和行为的是由制作表组成的脚本。

2)帧动画:是由一幅幅位图组成的连续的画面,就像电影胶片或视频画面一样,要分别设计每个屏幕显示的画面。

计算机制作动画时,只要做好关键帧画面,其余的中间画面都可以由计算机内插来完成,节省了人力物力,同时也提高了工作效率。在各种媒体的创作系统中,创作动画的软硬件环境都是较高的,它不仅需要高速的 CPU 和大容量的内存,而且制作动画的软件工具也比较复杂、庞大。

1.1.2 多媒体的主要特征

多媒体是融合两种或者两种以上媒体元素的信息交流和传播媒体,具体有如下 4 个主要的特点:

(1)信息量巨大

信息量巨大表现在信息的存储量以及传输量上。例如,640×480 像素、256 色彩色照片的存储量为 0.3 MB;CD 双声道的声音每秒的存储量为 1.4 MB;广播质量的数字视频码率约为 216 Mbps;高清晰电视数字视频码率在 1.2 Gbps 以上。

(2)数据类型具有多样性与复合性

多媒体数据包括文本、图形、图像、声音和动画等,而且还具有不同的格式、色彩、质量等。媒体信息具有多样化和多维化,通常不局限于单一媒体元素,而是多种媒体元素的有机组合,从而能够更好地丰富和表现信息。

(3)数据类型间的区别较大

不同媒体间的存储量差别较大;不同媒体间的内容与格式不一,相应的内容管理、处理方法和解释方法也不同。

(4)数据处理复杂

为了能够有效地对多媒体信息进行存储和在网络中进行传输,必须对多媒体信息进行有效处理。数据压缩和解压缩技术、语音识别、多媒体信息检索、虚拟现实等处理技术都是多媒体研究中的重要课题。

1.1.3 多媒体领域涉及的关键技术

多媒体技术是一种基于计算机科学的综合技术,其发展和应用需要一系列关键技术的支持。以下几方面技术是多媒体技术研究与应用领域所涉及的主要关键技术,也是多媒体领域研究的热点。

(1)多媒体数据压缩技术

由于数字化的图像、音频、视频等多媒体数据量非常大,为了能够有效地存储和传输多媒体数据,必须对其进行压缩处理。因此,多媒体数据压缩成为多媒体领域的一项关键技术。目前,几乎所有的多媒体技术的应用都是以数据压缩为基础的。例如,VCD 使用的 MPEG-1 压缩标准,DVD 和数字电视中使用的 MPE G-21 压缩标准,可视电话和视频会议中使用的 H. 261、H. 263、G. 721 和 G. 729 等压缩标准,网络电视和 IPTV 中使用的 MPEG-4 压缩标准等。编码理论研究已有 40 多年的历史,从 PCM 编码理论开始,已经产生了各种各样的针对不同应用和媒体的压缩算法、压缩技术和实现这些算法的大规模集成电路或软件,并日趋成熟。但此方向的研究远未停止,新的多媒体应用不断出现,现有算法有时很难满足新应用的需求,所以必须探索新的压缩算法和技术。例如,随着无线传感器网络应用的普及,必须研究针对无线传感器网络的高压缩率、低复杂度的多媒体压缩技术。

(2)多媒体数据库技术

由于多媒体信息是非结构型的,致使传统的关系数据库已不适用于多媒体的信息管理,需要从多个方面研究多媒体数据库,例如,研究多媒体数据模型、多媒体数据管理及存取方法和多媒体数据库用户界面等。目前,市场上已经出现了多媒体数据库管理系统,但其研究还很不成熟,与实际多媒体数据复杂的管理和应用需求仍有较大的距离。因而,功能强大的多媒体数据库仍是多媒体领域研究的方向之一。

(3) 多媒体存储技术

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,而传统的计算机存储设备如软盘和磁带等根本无法满足这种大信息量的存储要求,需要探索新的存储介质和存储技术解决多媒体信息的存储问题,目前,光盘存储技术和大容量硬盘技术发展非常迅速。同时,为了避免磁盘损坏而造成的数据丢失,需要研究相应的磁盘管理技术,如磁盘阵列(Disk Array)就是在这种情况下诞生的一种数据存储技术。可见,研究探索存取速度快、存储容量大、价格低廉且使用方便的存储设备和存储技术也是多媒体发展和应用的关键技术之一。

(4) 多媒体信息检索技术

随着计算机技术、多媒体技术、网络技术的普及和发展,如何对海量的多媒体信息进行有序化组织、整理,使得人们能够从浩如烟海的信息海洋中快速、准确地寻找到自己所需的有用信息成为当今网络时代最具挑战性的重要课题之一。

从根本上讲,多媒体信息检索就是要解决如何将网上信息进行有效存储、组织、检索以提供用户使用。多媒体信息检索研究与其他计算机研究方向一样,学科交融和交叉在多媒体检索研究也表现得十分明显,它涉及人工智能、心理学、脑科学、计算机视觉、信号处理、统计学、模式识别、数据库、计算机网络、视频通信和人机交互等诸多方面的理论。并且多媒体信息检索的研究目的就是帮助人们更快捷、更方便、更准确地找到需要的多媒体信息,所以多媒体信息检索本质上是理论与实践紧密结合的一项研究。

(5) 多媒体信息安全技术

随着多媒体技术的发展,高质量数字录制设备数量剧增,多媒体盗版问题直接影响着多媒体技术的应用。多媒体的内容拥有者正在急切地寻找能有效保护他们权利的技术和方案。多媒体信息安全技术已经成为多媒体领域另一个研究的热点技术。

多媒体信息安全技术包括:数字水印技术、信息隐藏与信息伪装技术、数字指纹技术、多媒体版权保护技术、多媒体信息认证技术和数字权限管理技术与系统等多个研究方向和研究内容。多媒体信息安全技术的研究也得到了国内外众多研究学者重视,针对多媒体信息安全问题提出了多种解决方案。

(6) 多媒体专用芯片技术

多媒体专用芯片仰仗于大规模集成电路(VLSI)技术,它是多媒体硬件系统体系结构的关键技术。因为要实现音频、视频信号的快速压缩、解压缩和播放处理,需大量的快速计算。而实现图像的特殊效果、图像生成、绘制等处理以及音频信号的处理等,也都需要较快的运算处理速度,因此,只有采用专用芯片,才能取得满意效果。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。多媒体专用芯片技术是多媒体应用产业的一个重要环节,例如,有了价格低廉的 MPEG-2 标准解码芯片就可以有力推动 DVD、数字电视等产业的迅速发展。

(7) 多媒体网络与通信技术

多媒体通信是一个综合性技术,涉及多媒体、计算机和通信等多个领域,长期以来一直是多媒体应用的重要方向。随着通信网络技术的发展,出现了多种网络类型,如宽带互联网、数字电视网、3G、无线局域网、无线城域网、无线个域网和无线自组织网络等。在不同的网络系

统中需采用不同的带宽分配方式、不同的压缩技术和不同的多媒体传输技术,多媒体网络和通信技术的应用研究是多媒体技术领域的热点研究方向之一。

(8) 虚拟现实技术

虚拟现实(Virtual Reality)通过综合应用计算机图像、模拟与仿真、传感器、显示系统等技术和设备,以模拟仿真的方式,给用户提供一个真实反映操纵对象变化与相互作用的三维图像环境所构成的虚拟世界,并通过特殊设备(如头盔和数据手套)提供给用户一个与该虚拟世界相互作用的三维交互式用户界面。利用多媒体系统生成的逼真的视觉、听觉、触觉及嗅觉的模拟真实环境,受众可以用人的自然技能对这一虚拟的现实进行交互体验,犹如在真实现实中的体验一样。

虚拟现实技术的四个重要特征如下:

1)多感知性:指除了一般计算机具有的视觉感知外,还有听觉感知、触觉感知、运动感知,甚至可包括味觉和嗅觉感知,只是由于传感技术的限制,目前尚不能提供味觉和嗅觉感知功能。

2)临场感:指用户感到存在于模拟环境中的真实程度。

3)交互性:指用户对模拟环境中物体的可操作程度和从环境中得到反馈的自然程度,其中也包括实时性。

4)自主性:指虚拟环境中物体依据物理规律进行动作的程度。

虚拟现实技术的实现需要相应的硬件和软件的支持,目前虚拟现实技术还不成熟,与人类现实世界中的行动还有一定的差距,还不能灵活、清晰地表达人类的活动与思维,因此,这方面还有大量的工作要做。

1.2 网络多媒体技术的发展

1.2.1 网络多媒体技术的发展历程

多媒体技术出现于 20 世纪 80 年代中期,由于数字化技术在计算机领域的广泛而卓有成效的应用,使得电视、录像以及通信技术也都开始由模拟方式转向数字化;另一方面,计算机应用开始深入到人们生活、工作的各个领域,这也要求其人机接口不断改善,即由字符方式、文本处理向图形方式、声音和图像处理发展。为此,把电视技术和计算机技术这两项对人类生活产生深刻影响的技术成果结合起来,并相互取长补短,实现信息交流的人为主动控制,以及信息交流形式的多样化,从而促使人们以一种全新的方式应用计算机。

1985 年,美国 Commodore 公司推出世界上第一台多媒体计算机 Amiga 系统。Amiga 机采用 Motorola M68000 微处理器作为 CPU,为了提高多媒体处理能力,Amiga 系统中采用了图形、音响和视频处理的三个专用芯片,同时还提供了一个专用的操作系统,能够处理多任务,并具有下拉菜单和多窗口等功能。

1984 年,美国 Apple 公司在研制 Macintosh 计算机时,为了增加图形处理功能、改善人机交互界面,使用了位图(bitmap)的概念对图形进行处理,并使用了窗口(window)和图标(icon)作为用户接口。这一系列改进所带来的图形用户界面(GUI)深受用户的欢迎,加上引入鼠标

(mouse)作为交互设备,大大方便了用户的操作。在这个基础上,1987年8月,Apple公司又引入了“超级卡”(Hypercard),它使Macintosh成为用户可以方便使用,并且能处理多种媒体信息的计算机。

1985年,Microsoft公司推出了Windows操作系统,它是一个多用户的图形操作环境。Windows使用鼠标驱动的图形菜单,是一个具有多媒体功能、用户界面友好的多层窗口操作系统。

1986年3月,Philips和Sony联合推出了交互式数字光盘系统(Compact Disc Interactive,CD-I),使得光盘成为交互式视频的存储介质。该系统把各种多媒体信息以数字化的形式存放在容量为650MB的只读光盘上,用户可以通过读取光盘内容播放多媒体信息。CD-I系统有两种工作方式:一种是与电视机、录像机和音响设备连接在一起,在系统的控制下,把来自光盘的音频、视频或图像数据传递给这些设备;另一种方式是作为多媒体控制权连接到其他计算机、工作站或小型计算机上。

1987年3月,位于新泽西州普林斯顿的美国无线电公司RCA推出了交互式数字视频系统(Digital Video Interactive,DVI),它以计算机技术为基础,用标准光盘存储和检索静止图像、动态图像、声音和其他数据。1989年3月,Intel公司宣布把DVI技术(包括DVI芯片)开发成一种可以普及的商品。

交互式光盘系统(CD-I)和交互式数字视频(DVI)技术都属于交互式视频领域,但是,CD-I是由视频专业公司按照在音像产品中引入微机芯片控制的设计思想开发出来的,设计目的是用来播放记录在光盘上的按照CD-I压缩编码方式编码的视频信号。而DVI则是由计算机专业公司按照在PC机中采用音视频板卡,软件采用基于Windows的音频/视频内核(AVK)的思路设计的,这就把彩色电视技术与计算机技术融合在一起。两者从不同的角度,按照不同的设计思想,最终实现了一个共同的目标——电视与计算机的有机结合。CD-I和DVI都是交互式视频领域中以光盘(CD-ROM)为存储介质的阶段性成果,其技术分别在后来的VCD和非线性编辑系统中有所体现。

在这段时期,“多媒体”(Multimedia)这一专业术语开始在社会上流传开来,并且取代了已经沿用多年的“交互式视频”。1985年10月,IEEE计算机杂志首次出版了完备的“多媒体通信”专集,是文献中可以找到的最早的出处。1987年成立了交互声像工业协会,1991年,该组织更名为交互多媒体协会(Interactive Multimedia Association,IMA)。

自20世纪90年代以来,多媒体技术逐渐成熟,多媒体技术从以研究开发为重心转移到以应用为重心。由于多媒体技术是一种综合性技术,它的实用化涉及计算机、电子、通信、影视等多个行业技术协作,其产品的应用涉及各个用户层次,因此,提出了对多媒体相关技术标准化的要求。

多媒体相关标准涉及多个技术领域,包括多媒体计算机标准、静止图像编码标准、视频编码标准、音频编码标准和多媒体通信标准等。最早出现的多媒体标准是多媒体个人计算机标准,1990年10月,在微软公司会同多家厂商召开的多媒体开发工作者会议上提出了MPC 1.0标准。1993年,由IBM、Intel等数十家软硬件公司组成的多媒体个人计算机市场协会(The Multimedia PC Marketing Council,MPMC)发布了多媒体个人机的性能标准MPC 2.0。1995年6月,MPMC又宣布了新的多媒体个人机技术规范MPC 3.0。在多媒体个人计算机标准制

定的同时,多媒体编解码技术标准工作也迅速开展起来,多媒体编解码技术的标准主要由国际电信联盟(International Telecommunications Union,ITU)和国际标准化组织(International Organization for Standardization,ISO)两个协会制定。ITU 制定的压缩编码标准主要有静止图像编解码标准 JPEG 和 JPEG 2000,视频编码标准 H. 261、H. 263 和 H. 264,音频编码标准 G. 721、G. 727、G. 728 和 G. 729 等。ISO 制定的标准主要有 MPEG-1、MPEG-2 和 MPEG-4 等。

1.2.2 网络多媒体技术的发展趋势

目前,多媒体技术正向网络化、智能化、标准化、多领域融合和虚拟现实等几个方向发展。

(1) 网络化

随着宽带网络的快速发展,网络传输速度和质量快速提高,各种基于网络的多媒体系统,例如,可视电话系统、点播系统、电子商务、远程教学和医疗等得到迅速发展。

多媒体通信网络环境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、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得大量信息,反过来又以最有效的方式为社会创造更大的利益。

(2) 智能化

未来的多媒体系统会具有越来越高的智能性,可以与人类进行自然的交互,系统本身不仅能主动感知用户的交互意图,还可以根据用户的需求做出相应的反应。目前正在研究的图像理解、语音识别、全文检索、基于内容处理的多媒体系统是使多媒体系统智能化的主要手段。

(3) 标准化

多媒体标准仍是研究的重点。各类标准的研究将有利于产品规范化,应用更方便。因为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很高的要求,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) 多领域融合

多媒体技术正在向各个技术领域渗透,如自动控制系统、人机交互系统、人工智能系统、仿真系统等,几乎所有的具有人机界面的应用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,正在为人们所接受,并称为研究课题之一。

(5) 虚拟现实

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、智能接口等技术相结合,建立高层次虚拟现实系统。

多媒体技术总的发展趋势是具有更好、更自然的交互性,更大范围的信息存取服务,为未来人类生活创造出一个在功能、空间、时间及人与人交互方面更完善的崭新世界。

1.3 网络多媒体处理技术的应用领域

多媒体技术的发展使计算机的信息处理在规范化和标准化的基础上更加多样化和人性化,特别是多媒体技术与网络通信技术的结合,使得远距离多媒体应用成为可能,也加速了多媒体技术在经济、科技、教育、医疗、文化、传媒、娱乐等各个领域的广泛应用。多媒体技术已成为信息社会的主导技术之一。

1.3.1 教育、培训领域中的应用

多媒体技术对教育产生的影响比其他领域的影响要深远得多。利用多媒体计算机的文本、图形、视频、音频及其交互式的功能,可以编制出计算机辅助教学软件,即课件。课件具有生动形象、人机交流、即时反馈等特点,能根据学生的水平采取不同的教学方案,根据反馈信息为学生提供及时的教学指导,创造出生动逼真的教学环境,改善学习效果。而且教师可以根据情况随时修改程序,不断补充新的教学内容。由于有人机对话功能,师生的关系也发生了变化,改变了以教师为中心的教学方式,学生在学习中担当更为主动的角色;学生可以参与控制以调整自己的学习进度,通过自己的思考进行学习,从而取得良好的学习效果。多媒体技术不仅改变传统的教学方式,也将使教材发生巨大的变化。将来的教材不仅有文字和静态图像,还将具有动态图像和语音等多种形式。

多媒体技术在教育与培训方面的应用可以用以下的“6C”概括。

(1) CAI——计算机辅助教学

CAI(Computer Assisted Instruction)是多媒体技术在教育领域中应用的典型范例,它是新型的教育技术和计算机应用技术相结合的产物,其核心内容是以计算机多媒体技术为教学媒介而进行的教学活动。

(2) CAL——计算机辅助学习

CAL(Computer Assisted Learning)也是多媒体技术应用的一个方面。它着重体现在学习信息的供求关系方面。CAL向受教育者提供有关学习的帮助信息,例如,检索与某个科学领域相关的教学内容,查阅自然科学、社会科学以及其他领域中的信息,征求疑难问题的解决办法,寻求各个学科之间的关系和探讨共同关心的问题等。

(3) CBI——计算机化教学

CBI(Computer Based Instruction)计算机化教学是近年来发展起来的,它代表了多媒体技术应用的最高境界。CBI将使计算机教学手段从“辅助”位置走到前台来,成为主角。CBI必将成为教育方式的主流和方向。

(4) CBL——计算机化学习

CBL(Computer Based Learning)是充分利用多媒体技术提供学习机会和手段的事物。在计算机技术的支持下,受教育者可在计算机上自主学习多学科、多领域的知识。实施CBL的关键,是在全新的教育理念指导下,充分发挥计算机技术的作用,以多媒体的形式展现学习的内容和相关信息。

(5)CAT——计算机辅助训练

CAT(Computer Assisted Training)是一种教学的辅助手段,它通过计算机提供多种训练科目和练习,使受教育者迅速消化所学知识,充分理解和掌握重点与难点。

(6)CMI——计算机管理教学

CMI(Computer Managed Instruction)主要是利用计算机技术解决多方位、多层次教学管理的问题。教学管理的计算机化,可大幅度提高工作效率,使管理更趋科学化和严格化,对提高教学管理水平发挥重要的作用。

1.3.2 电子商务领域中的应用

网络多媒体技术已经渗透到我们生活的方方面面,网络多媒体技术还可以很好地应用于电子商务中,例如,在广告和销售服务工作中,采用多媒体技术可以高质量、实时、交互地接受和发布商业信息,进行商品展示、销售演示,并且把设备的操作和使用说明制作产品操作手册,以提高产品促销的效果,为广大商家及时地赢得商机。此外,各种基于网络多媒体技术的演示查询系统和信息管理系统,如车票销售系统、气象咨询系统、病历库等也在人们的日常生活中扮演着重要的角色,发挥着重要的作用。

1.3.3 电子出版领域中的应用

当CD-ROM、DVD-ROM光盘出现以后,由于其存储量大,能够将文字、图形、图像、声音等信息进行存储和播放,出现了多种电子出版物,如电子杂志、百科全书、地图集、信息咨询、剪报等。电子出版是多媒体技术应用的一个重要方面,我国国家新闻出版总署对电子出版物曾有过如下定义:电子出版物是指以数字代码方式将图、文、声、像等信息存储在磁、光、电介质上,通过计算机或类似设备阅读和使用,并可复制发行的大众传播媒体。

电子出版物中信息的录入、编辑、制作和复制都借助计算机完成,人们在获取信息的过程中需要对信息进行检索、选择,因此电子出版物的使用方式灵活、方便、交互性强。

电子出版物的出版形式主要有电子网络出版和电子书刊两大类。电子网络出版是以数据库和通信网络为基础的一种出版形式,通过计算机向用户提供网络联机、电子报刊、电子邮件以及影视作品等服务,信息的传播速度快、更新快。电子书刊主要以只读光盘、交互式光盘、集成卡等作为载体,容量大、成本低是其突出的特点。

1.3.4 网络通信领域中的应用

多媒体技术与网络通信技术的结合产生了可视电话、视频会议、多媒体电子邮件、信息点播和计算机协同工作(Computer Supported Cooperative Work, CSCW)等应用技术,这些技术的应用在某种程度上已经改变了人们的生活方式和习惯,并将继续对人类的生活、学习和工作产生深刻的影响。

信息点播包括桌面多媒体通信系统和交互电视。通过桌面多媒体信息系统,人们可以远距离点播所需的信息,如电子图书馆、多媒体数据的检索与查询等。点播的信息可以是各种数据类型,包括立体图像和感官信息。用户可以按信息的表现形式和信息的内容进行检索,系统根据用户的需要提供相应的服务。交互式电视和传统电视的不同之处在于用户在电视机前可