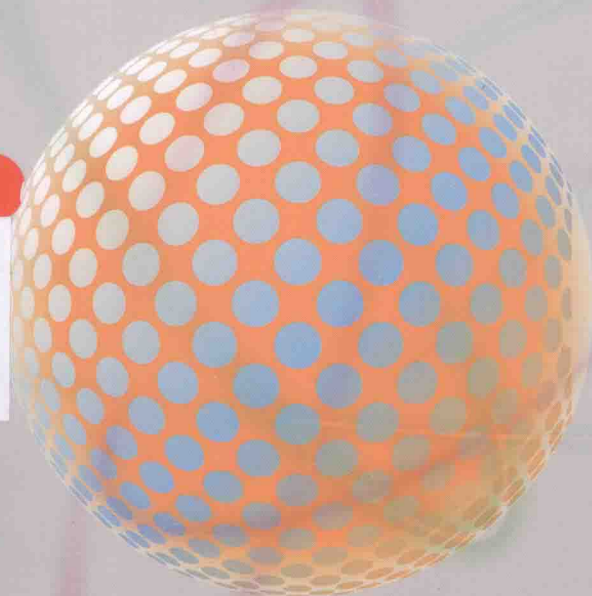


CUOMEITI TONGXIN JISHU
JI YINGYONG YANJIU

多媒体通信技术 (及应用研究)

主 编 马少斌 梁 晔
副主编 黄寿孟 孙玉轩 李 慧
马莲姑 杨婷婷 李筱锋



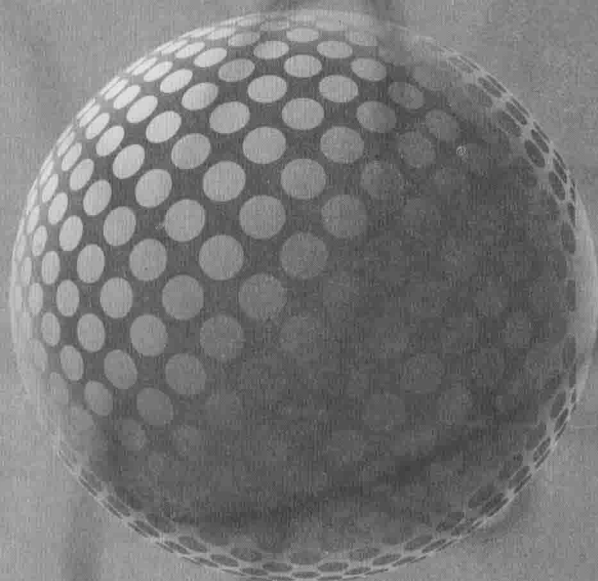
中国水利水电出版社
www.waterpub.com.cn

CUOMEITI TONGXIN JISHU
JI YINGYONG YANJIU

多媒体通信技术 (及)应用研究

主 编 马少斌 梁 峰
副主编 黄寿孟 孙玉钊 李 慧
马建姑 杨婷婷 李筱锋

藏书



中国水利水电出版社
www.waterpub.com.cn

内 容 提 要

本书对多媒体通信技术的基本概念、技术及应用做了全面的介绍。全书共 13 章,在介绍多媒体通信技术相关概念的基础上,重点对多媒体通信中的信息处理技术、通信网络、同步技术、通信终端以及流媒体技术做了比较系统的阐述,最后对一些典型的多媒体通信应用系统做了分析和探讨。本书注重基础理论和基本技术的讲述,同时也对相关标准和前沿技术进行了研究。书中内容丰富、新颖,叙述深入浅出,注重理论与实际应用的结合,更易于读者理解和掌握。

本书可作为高等学校通信工程、计算机通信等相关专业本科生的教材或研究生的教学参考书,也可供从事多媒体通信技术研究和开发的工程技术人员参考使用。

图书在版编目(CIP)数据

多媒体通信技术及应用研究/马少斌,梁晔主编.

—北京:中国水利水电出版社,2014.6

ISBN 978-7-5170-1983-1

I. ①多… II. ①马…②梁… III. ①多媒体通信—
通信技术 IV. ①TN919.85

中国版本图书馆 CIP 数据核字(2014)第 096107 号

策划编辑:杨庆川 责任编辑:杨元泓 封面设计:崔 蕾

书 名	多媒体通信技术及应用研究
作 者	主 编 马少斌 梁 晔 副主编 黄寿孟 孙玉轩 李 慧 马莲姑 杨婷婷 李筱锋
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 电话:(010)68367658(发行部)、82562819(万水)
经 售	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 27.5 印张 704 千字
版 次	2014 年 8 月第 1 版 2014 年 8 月第 1 次印刷
印 数	0001—3000 册
定 价	89.00 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

多媒体通信技术是一门综合的、跨学科的交叉技术,它是计算机技术、通信技术以及广播电视技术长期相互融合、渗透的产物。多媒体通信技术的蓬勃发展开始于 20 世纪 90 年代,即使在今天,仍然在不断发展和完善。实践证明,多媒体通信技术的广泛应用极大地提高了人们的工作效率,减轻了社会的交通负担,并且已经对人们传统的教育和娱乐方式产生了革命性的影响。同时,人们仍然在不断开发它的新的、更多的应用领域。可以预见,在未来的应用中,多媒体通信技术必将影响我们生活的方方面面,使人类的生活更丰富多彩。

多媒体通信主要研究多媒体数据的表示、存储、恢复和传输。多媒体数据是由在内容上相互关联的文本、图像、图形、音频、视频和动画等多种媒体数据构成的一种复合信息实体。其中,有着严格时间关系的音频、视频等类型的数据称为连续媒体数据,其他类型的数据称为离散媒体数据。一般来说,多媒体数据至少包含两种媒体数据,其中一种必须为连续媒体数据。

在本书编写过程中,编者注重难易结合,对涉及的难点、重点和新知识的部分增加了相关的基础知识的叙述。如图像信号和语音信号压缩编码、图像处理技术有一定的难度,本书加入了图像与语音技术基础方面的内容,帮助读者掌握相关知识。读者可根据不同要求进行取舍。

本书共分 13 章,全书主要内容包括绪论、多媒体数据压缩编码技术、多媒体数据压缩编码标准、多媒体信息处理技术、多媒体数据库与检索技术、分布式多媒体系统、多媒体通信网络技术、多媒体通信用户接入技术、多媒体通信同步技术、多媒体通信终端技术、超媒体与流媒体技术、多媒体数字水印技术、多媒体通信应用系统研究等。

本书是编者根据近年来从事多媒体通信技术教学和实践的体会,并参考了国内外相关文献,在原有讲义的基础上编写而成。全书力求对基础技术做到系统深入的介绍,对新技术做到文献材料翔实可靠,对具体应用做到具体分析。

由于受理论水平、实践经验及资料所限,虽然多次修改,书中疏漏与缺点一定存在。热忱欢迎同行和广大读者朋友批评指正。

编 者

2014 年 3 月

目 录

第1章 绪论	1
1.1 基本概念	1
1.2 多媒体通信的体系结构	2
1.3 多媒体通信的特征	3
1.4 多媒体系统的基本类型和相关业务	4
1.5 多媒体通信的关键技术	12
第2章 多媒体数据压缩编码技术	15
2.1 概述	15
2.2 预测编码	21
2.3 变换编码	24
2.4 统计编码	28
2.5 其他编码	33
第3章 多媒体数据压缩编码标准	35
3.1 静止图像压缩编码标准	35
3.2 视频压缩编码标准	42
3.3 音频压缩编码标准	61
第4章 多媒体信息处理技术研究	68
4.1 图像处理技术	68
4.2 音频处理技术	78
4.3 视频处理技术	95
4.4 动画处理技术	105
第5章 多媒体数据库与检索技术	124
5.1 多媒体数据库概述	124
5.2 多媒体数据库的数据模型	127
5.3 多媒体数据库管理系统	131
5.4 多媒体数据库的体系结构	134
5.5 基于内容的检索技术	138

第 6 章 分布式多媒体系统	158
6.1 分布式多媒体系统概述	158
6.2 分布式多媒体数据库系统	164
6.3 计算机支持的协同工作系统	167
第 7 章 多媒体通信网络技术	171
7.1 多媒体通信对传输网络的要求	171
7.2 网络类别	175
7.3 现有网络对多媒体通信的支持情况	178
7.4 多媒体通信协议与标准	212
第 8 章 多媒体通信用户接入技术	222
8.1 接入网概述	222
8.2 铜线接入网技术	229
8.3 光纤接入网技术	238
8.4 HFC 接入网技术	242
8.5 宽带无线接入网技术	244
第 9 章 多媒体通信同步技术	251
9.1 概述	251
9.2 多媒体同步参考模型	264
9.3 分布式多媒体系统中的同步	266
9.4 多媒体同步控制机制	270
第 10 章 多媒体通信终端技术	280
10.1 多媒体通信终端概述	280
10.2 多媒体通信终端的标准	282
10.3 基于 N-ISDN 网的多媒体通信终端	282
10.4 基于 IP 网络的多媒体通信终端	285
10.5 其他多媒体通信终端	292
10.6 基于计算机的多媒体通信终端	301
第 11 章 超媒体与流媒体技术	305
11.1 超文本与超媒体概述	305
11.2 超文本与超媒体系统	308
11.3 Web 超媒体系统	316
11.4 流媒体概述	323
11.5 流媒体传输协议	330

11.6 典型的流媒体系统.....	339
第 12 章 多媒体数字水印技术	353
12.1 概述.....	353
12.2 文本数字水印技术.....	365
12.3 图像数字水印技术.....	369
12.4 音频数字水印技术.....	375
12.5 视频数字水印技术.....	379
第 13 章 多媒体通信应用系统研究	384
13.1 概述.....	384
13.2 多媒体会议系统.....	384
13.3 视频点播(VOD)系统	392
13.4 视频监控系统.....	398
13.5 远程教育系统.....	403
13.6 IP 电话系统	408
13.7 网络电视(IPTV)系统	414
13.8 远程医疗系统.....	418
参考文献.....	431

第1章 绪论

多媒体通信是计算机、通信和多媒体技术相结合的产物,目前它已经成为通信的主要方式之一。现在的社会已进入信息时代,各种信息以极快的速度出现,人们对信息的需求日趋增加,这个增加不仅表现为数量的剧增,同时还表现在信息种类的不断增加上。一方面,这个巨大的社会需求(或者说是市场需求)就是多媒体通信技术发展的内在动力;另一方面,电子技术、计算机技术、电视技术及半导体集成技术的飞速发展多媒体通信技术的发展提供了切实的外部保证。由于这两个方面的因素,多媒体通信技术在短短的时间里得到了迅速的发展。

1.1 基本概念

1. 媒体

媒体(Medium)是信息表示和传输的载体。Medium 源于拉丁文,本身具有中介、中间的含义。在日常生活中被称为“媒体”的东西很多,如报纸广播是传播新闻的媒体,蜜蜂是传播花粉的媒体。准确地说,这些所谓的“媒体”指的是传播媒体。

计算机领域中的媒体包含有两种含义:一是,用以存储信息的实体,如磁盘、磁带、光盘和半导体存储器;二是,信息的载体,如数字、文字、声音、图形、图像和视频等。这里可以将就是计算机领域中的媒体分为感觉媒体、表示媒体、显示媒体、存储媒体和传输媒体 5 种形式。

(1) 感觉媒体(Perception Medium)

能直接作用于人的感官,使人直接产生感觉的一类媒体。感觉媒体包括:

- 视觉媒体,文字、景象。
- 听觉媒体,语言、音乐、自然界的各种声音。
- 触觉媒体,力、运动、温度。
- 味觉媒体,滋味。
- 嗅觉媒体,气味。

(2) 表示媒体(Representation Medium)

为了加工、处理和传输感觉媒体而人为地研究、构造出的一种媒体,它能够将感觉媒体从一个地方向另一个地方传输,以便加工和处理。

表示媒体有各种编码方式,如语音编码、文本编码、静止图像编码和运动图像编码等。根据属性的不同,表示媒体可进行如下分类:

• 按照时间属性划分,可以分为离散媒体和连续媒体。离散媒体是指不随时间变化而变化的媒体,如图形、静态图像、文本等。连续媒体则是指随时间变化而变化的媒体,如声音、视频、动画等。

• 按照空间属性划分,可以分为一维媒体、二维媒体和三维媒体。如单声道的音乐信号被称为一维媒体。二维媒体则指立体声、文本、图形等。三维图形、全景图像和空间立体声则被称为三维媒体。

• 按照生成属性划分,可以分为自然媒体和合成媒体。自然媒体是指采用数字化方法从自然界获取的媒体,如图像、视频等。合成媒体则是指通过计算机创建的媒体,如合成语音、图形、动画等。

(3) 显示媒体(Presentation Medium)

一种对感觉媒体的抽象描述形成的媒体。通过表示媒体,人类的感觉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。

(4) 存储媒体(Storage Medium)

一种用于存储表示媒体的物理设备,主要指与计算机相关的外部存储设备。

(5) 传输媒体(Transmission medium)

指的是将媒体从一个地方传输到另一个地方的物理载体。传输媒体是通信的信息载体,如双绞线、同轴电缆、光纤等。

2. 多媒体

多媒体(Multimedia)可以理解是直接作用于人感官的文字、图形、图像、动画、声音和视频等各种媒体的统称。多媒体包括许多东西,是文字、图形、图像、动画、声音和视频等各种媒体的组合。需要注意的是,用户也包括在多媒体内。对于多媒体而言,用户不仅仅是一个被动的观众,还是可以控制,可以交互作用,可以让它按用户的需要去做。

计算机领域内所指的多媒体一般是融合了两种以上媒体的人—机交互式信息交流和传播媒体。多媒体是信息交流和传播媒体,从这个意义上说,多媒体与电视、报纸、杂志等媒体具有相同的功能。

3. 多媒体通信

多媒体通信中的“多媒体”一词指的是在内容上相互关联的文本、图形、图像、音频和视频等媒体数据构成的一种复合信息实体。在多媒体通信过程中所传输和交换的是一个既有声音,又有图像,也可能还有文字、符号等多种信息类型的综合体,而且这些不同的媒体信息是相互联系、相互协调的。多媒体通信技术具有良好的人机界面,并可以在时间轴上和空间域内进行随意加工处理,给人们提供了综合的信息服务多媒体通信是多媒体技术和通信技术结合的产物,它将计算机的交互性、通信的分布性和广播、电视的真实性融为一体。

由于多媒体技术的介入使原来泾渭分明的各通信领域逐渐变得互相介入、互相融合,传统的电话将发展成为可见对方活动影像的可视电话;传统的单向广播型电视通信发展成双向选择型系统,即交互式影视节目自选型;在有线电视通信网络上传输计算机信息,在计算机通信网络上传电视信号。由于采用了多媒体技术,使多媒体计算机变成了录音电话机、可视电话、图文传真机、立体声音响设备、电视机和录像机等综合设备,特别是 Internet 的广泛应用,使我们真正进入了信息时代。

1.2 多媒体通信的体系结构

图 1-1 为国际电联 ITU-TL 211 建议为 B-ISDN 提出的一种适用于多媒体通信的体系结构模式。

一般应用	特殊应用
多媒体通信平台	
网络服务平台	
传输网络	

图 1-1 多媒体通信的体系结构

多媒体通信体系结构模式主要包括下列 5 个方面的内容。

1. 传输网络

它是体系结构的最底层,包括 LAN(局域网)、WAN(广域网)、MAN(城域网)、ISDN、B-ISDN(ATM)、FDDI(光纤分布数据接口)等高速数据网络。该层为多媒体通信的实现提供了最基本的物理环境。在选用多媒体通信网络时应视具体应用环境或系统开发目标而定,可选择该层中的某一种网络,也可组合使用不同的网络。

2. 网络服务平台

该层主要提供各类网络服务,使用户能直接使用这些服务内容,而无须知道底层传输网络是怎样提供这些服务的,即网络服务平台的创建使传输网络对用户来说是透明的。

3. 多媒体通信平台

该层主要以不同媒体(正文、图形、图像、语音等)的信息结构为基础,提供其通信支援(如多媒体文本信息处理),并支持各类多媒体应用。

4. 一般应用

该应用层指人们常见的一些多媒体应用,如多媒体文本检索、宽带单向传输、联合编辑以及各种形式的远程协同工作等。

5. 特殊应用

该应用层所支持的应用是指业务性较强的某些多媒体应用,如电子邮购、远程培训、远程维护、远程医疗等。

1.3 多媒体通信的特征

多媒体通信技术是多媒体技术、计算机技术、通信技术和网络技术等相互结合和发展的产物。在物理结构上,由若干个多媒体通信终端、多媒体服务器经过通信网络连接在一起构成的系统,就是多媒体通信系统。在计算机领域,人们也将该系统称为分布式多媒体系统。多媒体通信系统必须同时兼有多媒体的集成性、计算机的交互性、通信的同步性 3 个主要特征。

1. 集成性

多媒体通信系统能够处理、存储和传输多种表示媒体,并能捕获并显示多种感觉媒体,因此多媒体通信系统集成了多种编译码器和多种感觉媒体的显示方式,能与多种传输媒体接口,并且

能与多种存储媒体进行通信。

2. 交互性

多媒体通信终端的用户在与系统通信的全过程中具有完备的交互控制能力,这是多媒体通信系统的一个重要特征,也是区别多媒体通信系统与非多媒体通信系统的一个主要准则。例如,在数字电视广播系统中,数字电视机能够处理与传输多种表示媒体,也能够显示多种感觉媒体,但用户只能通过切换频道来选择节目,不能对播放的全过程进行有效的选择控制,不能做到想看就看、想暂停就暂停,因此数字电视广播系统不是多媒体通信系统。而在视频点播(VOD)中,用户可以根据需要收看节目,可以对播放的全过程进行控制,所以视频点播属于多媒体通信系统。

3. 同步性

同步性是指在多媒体通信终端上所显示的文字、声音和图像是以在时空上的同步方式工作的。同步性决定了一个系统是多媒体系统还是多种媒体系统,二者的含义完全不同,多种媒体是各种媒体的总称,例如图像、文本和声音等,它们中的任何一种都不是多媒体,只有将它们融合为一体,使它们具有时空上的同步关系,这才是多媒体。同步性也是在多媒体通信系统中最难解决的技术问题之一。

1.4 多媒体系统的基本类型和相关业务

多媒体计算机是多媒体技术的最直接、最简单的表现形式。因其本身具有存储、运算、处理和显示的能力,具有独立的功能,如动画显示、播放 VCD 节目等,因此,多媒体计算机一出现便立即在家庭、教育和娱乐方面得到广泛的应用。但是,多媒体技术真正的意义在于与网络的结合,在于通过网络(局域网和广域网)为用户以多媒体的方式提供信息服务。

基本的多媒体系统除了以多媒体计算机为基础的独立(Stand-Alone)商亭式系统之外,通过网络提供业务的系统可以分为两大类:一类是人与人之间交互的系统,如多媒体会议与协同工作、多媒体即时通信等;另一类是人机交互的系统,如多媒信息检索与查询、点播电视等,本节中将分别对这些系统进行介绍。

1.4.1 独立商亭式系统

在多媒体技术问世不久,便出现了用于导购、导游和教学等方面的商亭(Kiosk)系统。例如,商亭放置在大商场的人口处,将商场的主要产品拍成电视录像,并配之以介绍产品性能、价格、位于商场的什么位置等信息的伴音。将这些录像数字化、压缩编码以后存放到商亭内。顾客一进商场的大门,就可以从计算机屏幕上挑选自己中意的项目;屏幕上将按照要求,显示出你所感兴趣的诸如电视机、电冰箱、家具等商品的图像、价钱,以及售货员向你介绍商品性能的配音等。

凡是以一台多媒体计算机为核心的应用系统,我们都称为独立商亭式系统。在这类系统中,除了各种媒体的采集、表示、压缩存储和解压缩播放之外,如何组织素材,并运用多媒体手段将信息有效地、具有感染力(或艺术性)和方便地提供给用户是制作应用软件时应考虑的重要问题。这里涉及的不仅有技术、有艺术,甚至还有社会、心理学等方面的问题。多媒体制作软件(如 Au-

thoring Tool、Authorware 等),或者原有操作系统的多媒体扩展(如 Video for Windows),是为制作应用软件而提供开发环境的软件。它不仅向应用程序的开发者提供多媒体输入/输出设备的接口,更重要的是,还提供建立媒体数据之间的空间布局和播放时间顺序等关系的手段。因此,开发优秀的创作软件本身远比开发应用软件困难。

在这类系统中,操作系统的实时性是值得重视的另一个问题。在嵌入式系统或工业控制机中常常涉及实时操作系统,在那里强调的是对事件中断的实时响应。而在多媒体系统中,由于视频和音频数据需要在一定时间约束条件下(如每秒 25 幅图像)连续不断地送到输出设备上供用户聆听和观看,因此这里操作系统的实时性强调的是,处理这些有时间要求的连续媒体流的能力。

提供更友好的人机接口是商亭式系统技术发展的一个方向。除了使用键盘和鼠标,人们还试图通过触摸、声音、手势,甚至表情等多种模态的接口对系统进行控制,从而构成更人性化的多媒体交互环境。

1.4.2 多媒体信息检索与查询

通过因特网进行信息查询已是当前十分普及的应用。多媒体信息检索与查询 MIS(Multi-media Information Service)系统在根据关键字等对文本资料的查询之外,也同时具有活动图像和声音的查询能力。从通信方式而言,MIS 是点对点(信息中心对一个用户),或一点对多点(信息中心对多个用户)的双向非对称系统。从用户到信息源只传送查询命令,要求的传输带宽较小,而从信息源传送到用户的信息则是大量的、宽带的。

MIS 所涉及的两个重要技术问题是:首先,如何向用户提供丰富的信息和如何让用户快速、有效地查询与浏览这些信息;其次,如何合理、有效地组织多媒体数据的存储和检索。

为了对第一个问题有所认识,首先让我们回顾一下人人都熟悉的读书过程。对于阅读一本小说来说,人们通常是从头至尾逐页阅读的,或者说是按顺序阅读的。但在有些情况下,特别是在技术或社会科学领域,在阅读某本书的过程中,经常需要从另一本书或论文查找某个论点,或者说,在几本书之间需要交叉参考的情况常常发生。图 1-2 表示出用电子的方法来实现交叉参考的情况,这实际上已经是大家在因特网的查询中十分熟悉的过程:用鼠标点击黑框所标的地方,就会显示出箭头所示的有关参考信息,看完该信息后可以回到原来的页面,或者再进入其他页面……。箭头指向的页面(信息单元)可能与原来的页面在同一个文件中,也可能在其他文件

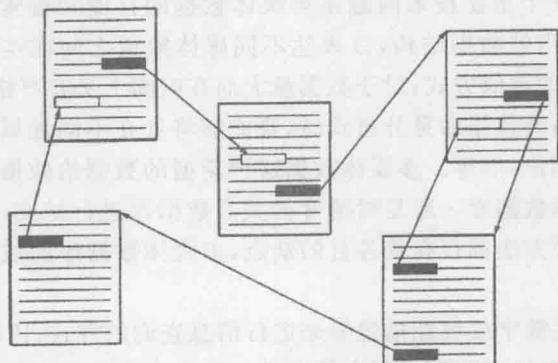


图 1-2 超文本文件结构

里。这种信息的非顺序(或称为非线性)的组织结构称为超文本(Hypertext),超文本中信息单元之间的链接称为超链(Hyper Link)。当上述信息组织方式不仅用于文本,还包括其他媒体数据、特别是音频和视频数据时则称为超媒体(Hypermedia)。超文本和超媒体这两个词在很多文献中也常常被混用。

超媒体为用户提供了一种在文件内部和文件之间迅速查找和浏览多媒体信息的方法,但是人们希望在更大的范围内迅速、有效地获取信息,这就不能不提到近年来推动因特网突飞猛进发展的 WWW 技术。WWW(WorldWide Web)最初是 1989 年在日内瓦 CERN 启动的一个研究项目的名称,由于它的巨大成功,现在 WWW 已经意味着在超媒体原理下发展起来的一系列概念和通信协议。Web 这个词也代表了世界范围内由因特网相互连接起来的众多的信息服务器所构成的巨大的数字化的信息空间,也有的学者将之称为超空间(Hyperspace)。

WWW 的基本思想和它所解决的问题主要体现在如下几个方面:

①在超空间中没有一个统一的管理者。任何人都可以创建超文本文件、将其与其他文件链接,并放入超空间中去。标准的超文本文件采用 HTML(Hyper Text Markup Language)格式。

②定义了一种在超空间中寻找所需要的文件的机制,称为统一资源定位器 URL(Uni-versal Resource Locator)。通过 URL 可以知道每个文件处于哪一台机器,叫什么名字,以及以何种机制可以将该文件传输到需要链接它的地方去。

③具有一个统一的、简单的用户界面,无论查询到的信息来自本机,还是来自远方的服务器,用户从界面上看起来都是一样的。实现 WWW 用户端功能的软件称为浏览器(Browser)。通过浏览器不仅能够调取 HTML 格式的文件,还可以调取以任何形式存储在已有的数据库、或信息库中的信息(虽然此时不具备超链接功能)。

以上 3 个问题的解决,使得世界上使用不同硬件和软件的分离的信息系统,通过因特网构成了一个庞大的统一的信息系统,从而为用户打开了通往一个大得难以想象的信息库的大门。这正是 WWW 取得巨大成功的原因。为了使用户不至于面对浩瀚的信息而不知所措,人们又进一步设计了帮助用户过滤掉无用信息、尽快找到所需要的信息的专门软件,这就是所谓的搜索引擎。

随着声音和活动图像等实时信息的逐步增加,因特网正在演变成世界范围内最大的 MIS 系统。以上所介绍的如何向用户有效地提供和查找信息的技术也广泛地应用在其他 MIS 系统中。由于关于这些技术的书籍已经很多,本书将不准备进一步讨论这方面的内容。

MIS 系统涉及的第 2 个重要技术问题是多媒体数据的存储和检索。与存储传统的数据不同,多媒体数据需要有适当的数据结构,以表达不同媒体数据之间在空间上与时间上的相互关系;对不同媒体要有合理的存储方式;对于数据量大而在时间上又有严格要求的音频和视频数据流,要有实时的提取算法;当数据库是分布式时,要能够将处在不同地域的服务器所提供的信息协调起来同步地提供给用户,等等。多媒体数据这种新型的数据给数据库的设计带来了一系列的新问题。目前的多媒体数据库一般是对通常的关系数据库进行扩充,或者采用面向对象的数据库来实现。但是这两种方法都存在着各自的缺点,多媒体数据库的成熟仍需要相当长的一段时间。

此外,传统的、利用关键字或属性描述等来进行信息查询的方式,比较适用于文字信息,用来对声音、图像等多媒体信息的查询则有不方便之处。基于内容的检索是伴随着视频和音频查询而发展起来的技术。利用这种技术,给出(或从查找对象中自动提取出)所要求的特征,例如图像

中物体的形状、颜色等,就能找出具有同样、或类似特征的物体的图像来。更高级的查询方式则是给出“概念”或“事件”,如国旗、山脉、骑自行车的人等,找出具有同样概念或事件的图像或视频来。这种方式也称为基于语义的检索。基于内容和基于语义的检索涉及图像和视频的分析与理解、语义提取、模式识别与人工智能等,是当前多媒体领域中的一个重要研究方向。由于本书侧重于多媒体通信,因此将不准备讨论这方面的内容。

最后需要指出,信息检索与查询业务的发展引发了网上交易的商机。既然查询到了某种合意的商品,为什么不可以订货、交钱,然后等着商品送货上门呢?这就是当前令人瞩目的电子商务。虽然电子商务所涉及的主要技术,如身份认证、安全保障、网上货币交易等,并不属于多媒体技术,但是电子商务的发展无疑将是推动 MIS 系统和业务发展的强大动力。

1.4.3 多媒体会议与协同工作

可视电话和会议电视是早在多媒体出现之前就已经存在的人与人之间进行通信的手段。计算机支持的协同工作 CSCW(Computer Supported Cooperative Work)也是早在 20 世纪 80 年代初在计算机领域内提出的概念。它是指用来支持多个用户共同参与一件工作(如共同编辑文件、修改设计图等)的计算机系统及其相关的技术,但合作者之间不能“见面”与交谈。多媒体的出现为这两种交流形式提供了结合的基础,合作者既能看得见、听得到,又能一起处理事务,使他们真正象聚集在同一个房间里面对面地交流与工作。这种通信系统和业务称为多媒体协同工作 MMC(Multimedia Collaboration)。近几年出现的多媒体远程医疗诊断系统、多媒体远程教育系统等都是 MMC 的典型应用。

多媒体一词出现以后,电视会议系统在概念上变得有些混乱,在这方面简要地叙述一下它的发展过程,对于理清概念是有帮助的。

1. 直接利用广播电视的电视会议

直接利用广播电视系统达到会议的目的的想法出现于 1956 年。其方式是将讲话人的形像和声音由电视台直接广播出去,听众则通过电视机收看。这种电视广播讲话的会议形式一直沿用至今。这是一种“一言堂”的会议方式:一个人(如国家领导人)讲,其余的人都是观众。

在 20 世纪 80 年代,这种直接利用广播电视“开会”的方式又增加了新的内容。例如,电视台邀请某些政治家或学者对最近发生的重大事件发表看法,进行辩论。辩论时各参与者处在各自所在城市的电视演播室内,电视信号通过双向的电视信道与主持会议的电视台相互传送,如图 1-3 所示。图中 A、B、C、D 是 4 个位于不同城市或国家的 4 个电视台的演播室,主持人(主席)在

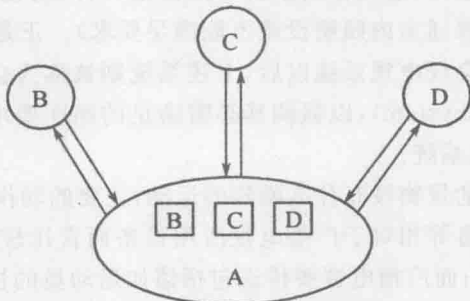


图 1-3 广播电视式的会议

A,参加辩论的3个人分别坐在B、C、D内,B、C、D的电视信号经卫星信道送到A,分别在B、C、D 3个电视显示屏上显示出来。A 演播室的信号向电视观众播放的同时,也送给B、C和D,使辩论者也能看到、听到其他3方面的情况。位于A的导演还可以将几路信号进行组合向广大观众广播。

2. 可视电话

1964年AT&T公司提出可视电话的通信方式,通话的人相互能看得见,不过这只能是两个人之间的会议系统。在汉语中,两个人谈话算不上会议,但是按英语的含意,两个人相见、交谈就是会议或会见(Meeting)。可视电话从一出现就遇到了普通百姓家“用不起”的问题,此问题伴随着它经历了从模拟信号至数字信号的发展。

“用不起”的问题很容易理解。以模拟信号为例,1路模拟电话信号所占的带宽是4 kHz,模拟可视电话信号约为1 MHz带宽,是前者的250倍,因此,通话的费用大体上也应该是这个比例。以图像压缩编码技术现在所达到的水平,可视电话的图像信号用64 kb/s(或更低一些)传送时,其图像质量已经能够为用户所接受,这是一个巨大的进步。但是,数字的可视电话机(终端)必须具备压缩编码器(将图像信号压缩后传送出去)和解码器(将收到的编码信号还原为可以显示的模拟图像信号),因此,可视电话终端的造价要比普通电话机高得多。这就是为什么可视电话作为一个独立的业务在相当长的时间内发展不起来的主要原因。近年来,随着集成电路技术的发展和终端集成的业务增多,可视电话只是终端(如手机)能够实现的多种业务中的一种,这个问题才得到了缓解。

虽然上述两种系统广义地讲可以叫做会议电视系统,但是它们毕竟不是以通常意义上的会议为主要目的的。在人们传统概念中的会议电视系统,是下述专用的会议系统。

3. 会议室会议电视系统

这是传统的会议电视系统,是专门为会议的目的而设计的。其原理与广播电视系统类似,由电视摄像机对着主会场、主席等拍摄,通过电缆、光缆、微波或卫星信道送到分会场收看。如果系统再复杂一些,要求主会场也能看到、听到分会场发言的情况,传输信道则是双向的,分会场也要有摄像机,将分会场的信号送到主会场。主会场(或者通信网的某个节点上)有信号切换设备,用来选取某一分会场的信号,并将该信号送至其他分会场;或者将几个分会场的信号“综合起来,以分画面的形式送给各个会场。在有的系统中,主会场还可以对分会场摄像机的摄取方向等进行控制。图1-4是会议室会议电视系统会场的示意图。

这类系统的一个重要特点是,需要像电视台的演播室一样,对被拍摄的景物(人、黑板、会场的全景等)给以专门的照明(普通室内照明设施不能满足要求)。正是由于这个原因,在近年来出现不需要有特殊照明要求的会议电视系统以后,上述系统则被称为会议室会议电视系统(Meeting Room Video Conference System),以强调其必需满足的照明要求。它既包括早期的模拟信号系统,也包括目前的数字化系统。

由于会议电视系统拍摄的景物没有什么剧烈的运动,主要的动作是讲话人嘴唇、眼睛或头部的运动,所以摄像机、信道设备等相对于广播电视所用设备而言比较简单。对于模拟系统,传送1 MHz的视频信号就可以了;而广播电视要传送包括诸如运动员的快速动作在内的高速运动的图像,因而需要6 MHz带宽。同样,对于数字系统,由于会议电视图像的相对静止或者说运动缓慢,因此在同样的图像分辨率下,会议电视的数据率可以被压缩更大的倍数。另外,还是由于同

样的原因,为了保证动作的连续,电视图像每秒钟需要传送 25 帧,而在会议电视中每秒传送 10~15 帧即可以被接受。而且,会议电视的情况与人们看电视不同,看电视是一种艺术欣赏,人们对图像的分辨率要求高,而会议则是在较长时间看一个固定的面貌,与会者在心理上对图像分辨率的要求大为降低。数据率为 384 kb/s 的系统所给出的图像质量已经可以令人足够满意了。

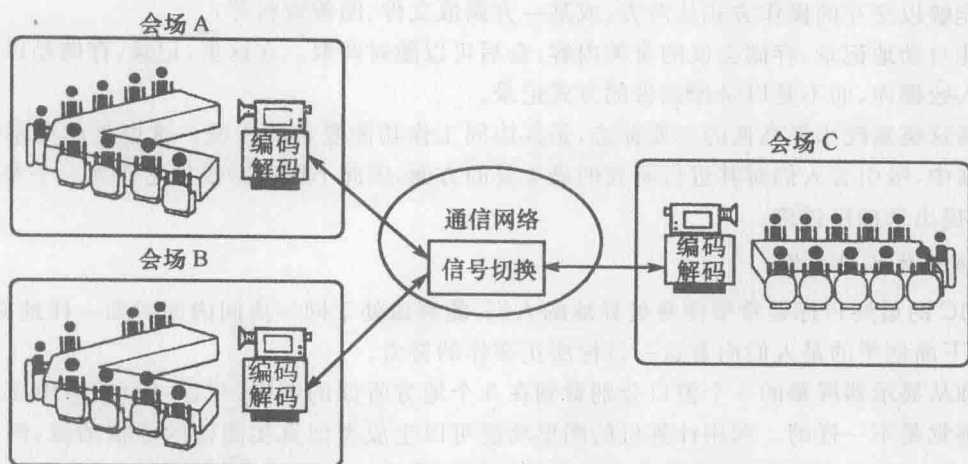


图 1-4 会议室会议电视系统

4. 桌面会议电视系统

用计算机取代传统的会议电视系统中的解码设备和电视机或者电视显示器,这是桌面会议电视系统的基本特征。

开会时,演讲的人往往需要写字、画图、作曲,以表述其演讲的内容。在传统的会议电视中,这些操作通常要在真实的黑板上进行。此时,电视摄像机不仅要拍摄讲话者上半身的像、他的来回的走动,还要能够对整个黑板进行拍照,这要求摄像机有较大的视野和较高的灵敏度,因此其照明条件必须达到演播室的标准。

用计算机取代电视机以后,这些操作可以通过键盘、鼠标在计算机上进行,直接得到代表文字、图形的数据,传送到其他与会者的计算机上显示。这就是最简单的“白板”的概念。这时,讲话的人只需要坐在摄像机前面讲话即可(无需走动)。这不仅降低了照明要求,也降低了对摄像机的视野和灵敏度要求,摄像机大为简化。

在桌面电视会议系统中,计算机的引入导致了系统功能上的变化,便于增加与会各方进行协同工作的模式,以向多媒体协同工作的方向发展。

5. 视听会议系统

这一名称在汉语中没有多少新的含义,因为“电视会议”一词已经标明了“又看”、“又听”的特点。Audiovisual 会议系统一词主要出现在与国际标准有关的文件中,它既包括传统的会议室会议电视系统,也包括桌面会议电视系统。

6. 多媒体会议电视系统

在多媒体出现之后,无论是会议室系统还是桌面系统,都在一定程度上开始融入计算机协同工作的功能,因此要给多媒体会议电视系统下一个严格的定义比较困难,但是下面所叙述的衡量准则是值得借鉴的。

如果说一个系统已经超出了电视会议的范围,进入多媒体会议系统的范畴,该系统应该在比较高的水平上具备了下述功能:

①具有比较复杂的协同工作功能,使得身处异地的人员可以同时使用同一种软件工作,例如用同一文字处理软件修改文件,用同一 CAD 软件修改设计图,等等;

②能够以交互的操作方式从对方、或某一方调取文件、图像资料等;

③能自动地记录、存储会议的有关内容,会后可以随时调取。在这里,记录、存储是以数字的方式存入数据库,而不是以录像磁带的方式记录。

判断这类系统水平高低的主要标志,是其协同工作功能复杂的程度。这也是此类系统在未来的发展中,吸引着人们对其进行研究的最主要的方面,因此下面有必要把它作为一个专门的问题,单独提出来加以讨论。

7. 多媒体协同工作

MMC 的最终目标是希望使身处异地的人们,能够像处于同一房间内面对面一样地交谈、协商工作,下面例举的是人们向着这一目标所正在作的努力。

教师从显示器屏幕的 3 个窗口分别看到在 3 个地方听课的学生,与在一个教室中面对全体学生的感觉是不一样的。利用计算机的图形功能可以生成类似真实图像的虚拟图像,例如具有天花板、窗户、灯具的教室,并将从 3 个地方传送来的学生的现场图像与计算机生成的虚拟教室图像结合在一起,构成一个全体学生在内的完整的教室全貌,将会给人以更真实的感觉。

在现实生活中举行会议时,除了主席和与会人的公开讨论外,某个与会者有时需要和邻座说一些不愿意让别人听到的悄悄话,或暗中递个条子;有时某个与会者会从文件包中拿出一份文件递给另一个人,然后两人根据文件内容小声商量一下,再正式在会议上发言;如此,等等。在多媒体会议中,要实现类似现实生活中的这些简单的行为要涉及许多技术问题。

显示器的屏幕是平面的,无论屏幕上显示的景像是多么地有立体感,人们仍然是身在其外,而不是身在其中。如何将虚拟现实(Virtual Reality)与协同工作结合起来,使人们在虚拟的三维环境之中协同工作是一个值得研究的课题。

人们会面时的第一个动作往往是一边握手,一边说“你好”。如果 MMC 终端可以用语言(不是键盘)输入、并配有机器手,可能使你感受到远方合作者向你握手问好的真实感觉。除了听觉和视觉之外,将其他的感觉,如触觉、嗅觉等结合进协同工作环境;或者将多媒体协同工作与机器人技术结合起来,使合作者能够共同进行除了屏幕上的工作(如编辑文件之类)以外的事情,这些都是研究者在探索的问题。

总之,多媒体协同工作将从各种不同的方面,向着能够使得被空间距离分开的人,在必要的时候可以像已经聚在一起,有面对面地一起工作的条件与自我感觉的方向发展。但要真正达到这一目标,要走的路途还相当遥远。这里包括的不仅是技术问题,还有许多为社会学和心理学家们所感兴趣和值得研究的问题。

从通信的角度来看,MMC 系统是对通信系统要求最高的应用。它要求一点对多点,或者多点对多点的实时的不间断的信息传输。在复杂的协同工作系统中,要实现“开小会”、“说悄悄话”、“传条子”等,还要能够随时建立、撤销某些私有信道。当涉及视、听之外的其他形式的传感器时,通信机制的复杂程度则会更高。