



普通高等教育“十一五”国家级规划教材
电子信息科学与工程类专业

多媒体通信技术基础

(第二版)

● 蔡安妮 等编著



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

TN919/24=2

2008



普通高等教育“十一五”国家级规划教材

电子信息科学与工程类专业

多媒体通信技术基础

(第二版)

蔡安妮 等编著

電子工業出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

本书是普通高等教育“十一五”国家级规划教材。作者紧密跟踪国际上多媒体技术发展的动向和研究成果,在分析本领域内大量具有代表性的文献及书籍的基础上,结合多年的科研和教学经验,综合提炼出本书的大纲和内容。全书比较全面地介绍了这一新领域内的主要理论与技术,包括:多媒体技术的特征、视觉特性与彩色电视信号、数据压缩的基本技术、视频数据的压缩编码、音频数据的压缩编码、多媒体同步、多媒体传输网络、多媒体通信终端与系统、视频在分组网上的传输、视频在异构环境下的传输等10章。全书在理论上力求严谨、叙述上尽量深入浅出。

本书选材兼顾到研究生及本科生教学两个方面的需要,同时可以作为从事通信、信息及相关行业的科研和工程技术人员的参考。

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。

图书在版编目(CIP)数据

多媒体通信技术基础/蔡安妮等编著. —2版. —北京:电子工业出版社,2008.7

普通高等教育“十一五”国家级规划教材

ISBN 978-7-121-06967-3

I. 多… II. 蔡… III. 多媒体—计算机通信—通信技术—高等学校—教材 IV. TN919.85

中国版本图书馆CIP数据核字(2008)第091491号

策划编辑:陈晓莉

责任编辑:陈晓莉 特约编辑:李双庆

印 刷: 北京市李史山胶印厂
装 订:

出版发行:电子工业出版社

北京市海淀区万寿路173信箱 邮编100036

开 本:787×1092 1/16 印张:30.5 字数:780千字

印 次:2008年7月第1次印刷

印 数:3000册 定价:45.00元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系。
联系及邮购电话:(010) 88254888。

质量投诉请发邮件至 zltts@phei.com.cn,盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线:(010) 88258888。

再版前言

本书在 2000 年出版的“多媒体通信技术基础”第一版的基础上，根据近年来技术的发展和曾经使用过该书的教师和学生的意见，经过较大幅度的增删和修改而成。本书与第一版一样，作者是按照下述方向努力的：

首先阅读这个新技术领域中绝大多数有代表性的论文、书籍，并根据自己的科研经验，经过消化、提炼，分门别类，使其系统化；然后用自己的语言，尽量通俗、严谨、简繁适当地写出来。

本书选材时尽量兼顾了本科生教学和研究生教学两方面的需要，以及在通信、信息及相关领域中从事研究开发的工程技术人员掌握多媒体技术的需要。在高等院校使用本教材时，可根据本校、本系的特点，选择书中的不同章节、或不同章节中的部分内容分别作为本科生、研究生讲课使用。为了加深理解和拓展各章中的内容，每章之后都附有一定数量的习题和参考文献。

本书第 5 章由苏菲编写；第 3.8 节、4.2、4.3.3 节和 10.2、10.3 节由庄伯金编写；第 1、2 章由孙景鳌第一版编写。

本书取材于众多的文献，作者在此对这些推动多媒体技术发展的人们表示敬意和感谢；同时，对给本书第一版提出意见和建议的老师和学生表示感谢。最后，作者仅以本书缅怀北京邮电大学多媒体通信与图像识别实验室创建人、本书第一版作者之一孙景鳌教授，由于他在实验室对多媒体方向科研的倡导和坚持，才有今天这本书。

作 者

2008 年 3 月

目 录

第1章 概论——多媒体技术的特征	1
1.1 概述	1
1.2 多媒体的概念与含义	1
1.3 多媒体产生的技术背景	4
1.3.1 图像压缩编码技术的成熟	4
1.3.2 大规模集成电路技术的发展	5
1.3.3 大容量数字存储技术的发展	5
1.4 多媒体系统的基本类型及相关业务	6
1.4.1 独立商亭式系统	6
1.4.2 多媒体信息检索与查询	7
1.4.3 多媒体会议与协同工作	9
1.4.4 多媒体即时通信	13
1.4.5 点播电视 (VOD)	14
1.5 三网融合及相应的业务	18
1.5.1 网络的融合	18
1.5.2 多重服务与业务融合	20
习题一	20
参考文献	21
第2章 视觉特性和彩色电视信号	22
2.1 人的视觉特性	22
2.1.1 图像对比度与视觉的对比度灵敏度特性	22
2.1.2 空间频率与视觉的空间频率响应	23
2.1.3 视觉的时间域响应	25
2.1.4 彩色的计量和彩色视觉	26
2.2 彩色电视信号	28
2.2.1 扫描——空间频率到时间频率的转换	28
2.2.2 隔行扫描与逐行扫描	28
2.2.3 电视信号的带宽	29
2.2.4 彩色空间的处理	30
2.2.5 全彩色电视信号	31
2.3 彩色电视信号的数字化	33
2.3.1 分量电视信号的数字化	33
2.3.2 复合电视信号的数字化	35
习题二	35

参考文献	36
第3章 数据压缩的基本技术	38
3.1 概述	38
3.2 数据压缩的理论依据	38
3.2.1 离散信源的信息熵	39
3.2.2 信源的概率分布与熵的关系	39
3.2.3 信源的相关性与序列熵的关系	40
3.3 信息率—失真理论	41
3.3.1 通信系统的一般模型	42
3.3.2 信息率—失真函数	43
3.3.3 限失真信源编码定理	44
3.4 取样频率的转换	46
3.4.1 下取样	46
3.4.2 上取样	48
3.4.3 分数比率转换	50
3.5 预测编码	50
3.5.1 差分脉冲编码调制 (DPCM)	51
3.5.2 序列图像中运动矢量的估值	53
3.5.3 具有运动补偿的帧间预测	58
3.6 正交变换编码	61
3.6.1 最佳线性正交变换	61
3.6.2 离散余弦变换	63
3.7 子带编码	66
3.7.1 子带编码的工作原理	66
3.7.2 正交镜像滤波器组	69
3.7.3 时域混叠消除	71
3.8 小波变换编码	74
3.8.1 多尺度分析	74
3.8.2 二进小波变换	76
3.8.3 变换系数的排序和编码	77
3.9 量化	82
3.9.1 均匀量化器	82
3.9.2 最小均方误差量化器	83
3.9.3 最小熵量化器	84
3.9.4 自适应量化	86
3.9.5 DPCM 预测误差的量化	86
3.9.6 DCT 系数的量化	87
3.8.7 子带信号的量化	89
3.10 熵编码	89
3.10.1 熵编码的基本概念	89

3.10.2	霍夫曼编码	91
3.10.3	算术编码	93
3.11	压缩编码算法性能的评价	97
习题三		98
参考文献		101
第4章	视频数据的压缩编码	102
4.1	基于帧的视频编码	102
4.1.1	典型的编码器与解码器	102
4.1.2	视频序列的编码	107
4.1.3	帧内预测编码	108
4.1.4	帧间预测编码的优化	110
4.1.5	基于率-失真优化的编码模式选择	117
4.1.6	低计算复杂度的变换编码与量化	119
4.2	基于对象的视频编码	122
4.2.1	编码器与解码器结构	122
4.2.2	任意形状 VOP 的预测编码	124
4.2.3	任意形状 VOP 的纹理编码	125
4.2.4	形状编码	126
4.2.5	视频对象序列的编码	127
4.3	恒定速率编码器的速率控制	128
4.3.1	高码率应用的速率控制 (TM 5)	128
4.3.2	速率控制中的率-失真模型	130
4.3.3	低延时应用的速率控制 (TMN8)	133
4.3.4	基于率-失真优化的速率控制 (JVT-G012)	135
4.3.5	视频缓存证实器	138
4.4	变速率视频编码	139
4.5	图像和视频压缩编码的国际标准	141
4.5.1	静止图像压缩编码标准 JPEG/JPEG2000	141
4.5.2	视听会议压缩编码标准 H. 261	142
4.5.3	数字声像存储压缩编码标准 MPEG-1	143
4.5.4	通用视频及伴音压缩编码标准 MPEG-2 (H. 262)	144
4.5.5	低比特率视听会议压缩编码标准 H. 263	146
4.5.6	通用音视频对象压缩编码标准 MPEG-4	147
4.5.7	新一代视频压缩编码标准 H. 264/AVC	148
4.5.8	视频压缩编码的国家标准 AVS	150
习题四		150
参考文献		152
第5章	音频数据的压缩编码	153
5.1	概述	153
5.2	人的听觉特性	153

5.2.1 响度级和响度	154
5.2.2 听觉灵敏度	154
5.2.3 听觉掩蔽	155
5.2.4 临界带宽	157
5.3 音频信号的数字化	157
5.4 音频自适应差分脉冲编码调制	158
5.5 音频子带编码	159
5.6 线性预测编码	161
5.7 码激励线性预测编码 (CELP)	162
5.7.1 感知加权滤波器	162
5.7.2 合成分析法	163
5.7.3 CELP 编解码原理	164
5.7.4 G.729 编解码器	165
5.8 感知编码	167
5.9 MPEG-1 音频编码	168
5.9.1 概述	168
5.9.2 MPEG-1 心理声学模型	169
5.9.3 编码层次	171
5.10 MPEG-2 音频编码	174
5.10.1 MPEG-2 BC	175
5.10.2 MPEG-2 AAC	176
5.11 杜比 (AC-3) 编码	178
5.12 音频压缩编码的国际标准	180
习题五	183
参考文献	183
第 6 章 多媒体同步	184
6.1 多媒体数据	184
6.1.1 连续媒体数据与静态媒体数据	184
6.1.2 多媒体数据内部的约束关系	185
6.1.3 多媒体数据的构成	187
6.2 多媒体数据时域特征的表示	188
6.2.1 时域场景和时域定义方案	188
6.2.2 时域参考框架	190
6.2.3 描述时域特征的时间模型	190
6.2.4 同步容限	196
6.3 多媒体同步的 4 层参考模型	197
6.4 分布式多媒体系统中的同步	200
6.4.1 分布式多媒体系统的结构	200
6.4.2 同步规范的传送	200
6.4.3 影响多媒体同步的因素	201

6.4.4	多级同步机制	203
6.5	连续媒体内部的同步	203
6.5.1	基于播放时限的同步方法	203
6.5.2	基于缓存数据量的同步方法	205
6.6	媒体流之间的同步	207
6.6.1	基于全局时钟的时间戳方法	207
6.6.2	基于反馈的流间同步方法	208
6.7	接收与发送时钟的同步	210
6.7.1	基于接收缓存器的方法	210
6.7.2	基于时间戳的锁相方法	211
6.7.3	基于网络时间协议的方法	212
6.8	同步算法小结	213
	习题六	214
	参考文献	215
第7章	多媒体传输网络	217
7.1	概述	217
7.2	多媒体信息传输对网络的要求	217
7.2.1	性能指标	217
7.2.2	网络功能	221
7.2.3	服务质量的保障	222
7.3	网络类别	223
7.3.1	电路交换网络和分组交换网络	223
7.3.2	面向连接方式和无连接方式	225
7.3.3	资源预留、资源分配和资源独享	225
7.4	电路交换广域网对多媒体信息传输的支持	226
7.4.1	电路交换网络	226
7.4.2	多点控制单元	227
7.5	分组交换广域网对多媒体信息传输的支持	228
7.5.1	帧中继	228
7.5.2	SMDS	230
7.6	ATM网对多媒体信息传输的支持	231
7.6.1	ATM原理	231
7.6.2	ATM协议结构	233
7.6.3	ATM服务类型和ATM适配层	234
7.6.4	ATM性能	235
7.6.5	ATM网应用	235
7.7	以太网对多媒体信息传输的支持	237
7.7.1	传统的共享介质局域网	237
7.7.2	以太网帧交换	240
7.7.3	吉比特以太网	241

7.7.4	以太网 QoS 保障	244
7.7.5	以太网应用	244
7.8	IP 网对多媒体信息传输的支持	246
7.8.1	传统的 IP 网	246
7.8.2	IP 多播	249
7.8.3	新一代 IP 协议 (IP v6)	251
7.8.4	IP 核心网交换 (路由) 与传输	256
7.9	IP QoS 保障机制	261
7.9.1	综合服务模型与资源预留协议	261
7.9.2	区别服务	263
7.9.3	多协议标记交换	265
7.9.4	通用多协议标记交换	266
7.9.5	流量工程和基于约束的寻径	267
7.9.6	MPLS 与 Diffserv 的结合	269
7.10	无线局域网对多媒体信息传输的支持	270
7.10.1	无线传输的特点	270
7.10.2	无线局域网的构成	271
7.10.3	基本的介质接入控制方式	272
7.10.4	先进的介质接入控制方式	274
7.11	蜂窝移动通信网对多媒体信息传输的支持	277
7.11.1	蜂窝的概念	277
7.11.2	多址接入	278
7.11.3	蜂窝移动通信系统的发展	279
7.11.4	无线网络中多媒体传输的特殊问题	283
7.12	宽带用户接入网	285
7.12.1	数字用户线路接入	285
7.12.2	光缆接入	289
7.12.3	光缆—同轴电缆混合接入	291
7.12.4	宽带无线接入	296
7.13	下一代网络	300
7.13.1	网络融合模型	300
7.13.2	IP 多媒体子系统	301
7.13.3	下一代网络框架	302
	习题七	303
	参考文献	304
第 8 章	多媒体通信终端与系统	306
8.1	传输层协议	306
8.1.1	应用层分帧和集成层次处理	306
8.1.2	传统的因特网传输层协议	307
8.1.3	实时传输层协议 RTP 及 RTCP	309

8.1.4	RTP/RTCP 的扩展	312
8.1.5	UDP-Lite	314
8.1.6	流控制传输协议 SCTP	316
8.2	多媒体数据的复接与分接	320
8.2.1	同步时分复用	320
8.2.2	分组复用	321
8.2.3	多媒体数据的封装	326
8.3	视听通信系统与终端	329
8.3.1	视听通信终端的一般框架	329
8.3.2	同步电路交换网视听业务标准 (H. 320)	329
8.3.3	ATM 网视听业务标准 (H. 310 和 H. 321)	331
8.3.4	分组交换网视听业务标准 (H. 322 和 H. 323)	332
8.3.5	公用电话网视听业务标准 (H. 324)	335
8.3.6	移动电话网视听业务标准 (H. 324M 和 3G-324M)	337
8.3.7	H 系列系统间的互通	339
8.3.8	视听系统的通信控制协议	341
8.4	会话发起协议 SIP 和会话描述协议 SDP	344
8.4.1	会话发起协议 (SIP)	344
8.4.2	SIP 消息格式和会话描述协议 (SDP)	347
8.4.3	SIP 与 H. 323 信令的比较	348
8.5	协同计算与组通信	349
8.5.1	协同计算的概念	349
8.5.2	组通信	350
8.5.3	应用共享控制	351
8.6	多媒体流式应用系统与终端	353
8.6.1	流媒体的概念	353
8.6.2	多媒体流式应用系统的结构	353
8.6.3	实时流协议 RTSP	356
8.6.4	流媒体文件	357
8.6.5	流式应用系统的关键技术	360
8.6.6	流服务器	362
8.6.7	内容分发网络	369
8.6.8	对等网络实时媒体流的分发	373
8.6.9	3GPP 分组交换流服务 (PSS)	377
8.6.10	IPTV	378
	习题八	383
	参考文献	384
第 9 章	视频数据的分组传输	385
9.1	概述	385
9.2	服务质量的保障	385

9.2.1	端到端的 QoS 保障	385
9.2.2	QoS 规范	386
9.2.3	QoS 预制机制	388
9.2.4	QoS 控制机制	390
9.2.5	QoS 管理机制	393
9.2.6	有关 QoS 保障机制的研究	395
9.3	应用层拥塞控制	395
9.3.1	TCP 拥塞控制	395
9.3.2	TCP 流量模型和 TCP 友好性	396
9.3.3	基于速率的拥塞控制	398
9.3.4	拥塞检测	399
9.3.5	媒体流的速率调整	402
9.4	传输层差错控制	402
9.4.1	自动重传技术	402
9.4.2	循环冗余码	403
9.4.3	前向纠错码	405
9.4.4	前向删除复原码	410
9.4.5	数据交错	416
9.4.6	前向差错控制与 ARQ 的结合	417
9.5	编码层差错控制	417
9.5.1	误码和包丢失对已压缩视频信号的影响	418
9.5.2	抗误码编码	419
9.5.3	错误掩盖	427
9.5.4	编解码器交互差错控制	431
9.6	差错控制小结	434
	习题九	435
	参考文献	436
第 10 章	视频在异构环境中的传输	438
10.1	概述	438
10.2	可伸缩性编码	439
10.2.1	可伸缩性编码的概念	439
10.2.2	空间域可伸缩性编码	440
10.2.3	时间域可伸缩性编码	442
10.2.4	质量可伸缩性编码	444
10.2.5	可伸缩性编码的联合应用	445
10.2.6	基于小波变换的可伸缩性视频编码	446
10.3	视频转码	447
10.3.1	视频转码框架	447
10.3.2	运动向量的再利用	450
10.3.3	DCT 域的运动补偿	453

10.3.4	转码的速率控制	454
10.4	多描述视频编码	456
10.4.1	多描述编码的概念	456
10.4.2	多描述的生成	458
10.4.3	视频多描述编码	462
10.5	视频信源模型	463
10.5.1	概述	463
10.5.2	视频信源的统计特性	464
10.5.3	自回归过程模型	465
10.5.4	马尔柯夫过程模型	468
	习题十	470
	参考文献	470

第1章 概论——多媒体技术的特征

1.1 概述

在技术发展史上,计算机、通信和广播电视一直是三个互相独立的技术领域,各自有着互不相同的技术特征和服务范围。但是,近几十年来,随着数字技术的发展,这三个原本各自独立的领域相互渗透、相互融合,形成了一门崭新的技术——多媒体。多媒体技术的最初体现是配之以声卡、视卡的多媒体计算机。它一出现立即在世界范围内,在家庭教育和娱乐方面得以广泛的应用,并且由此而激发了小型激光视盘(VCD和DVD)的迅速发展,促进了数字电视和高清晰度电视(HDTV)的迅速发展。多媒体技术的应用与发展,又反过来进一步加速了这三个领域的融合,使多媒体通信成为通信技术今后发展的主要方向之一。

有许多技术,从它们开始出现时就给人以清楚明了的概念。例如电话技术,从最初用两根电线把两部简陋的电话机连接起来实现远距离通话时起,它就被称为电话技术;后来经历了人工交换、步进制交换、程控交换,以至于发展到数字式移动电话,仍然是电话技术。多媒体技术所遇到的情况则有所不同。使人们不容易清楚地建立起“什么是多媒体”的概念的因素很多。首先,通信、计算机与彩色电视本来都是技术面宽而复杂的技术,由它们融合在一起而产生的多媒体技术,其技术覆盖面自然就更宽,技术的交叉更为复杂。这就使得多媒体不能像其他诸如电话、电影、电视、汽车、马车等事物那样一目了然。另外,为了经济上或商业上的利益,某些商家把本来不属于多媒体的技术说成是多媒体技术,人为地造成了概念上的混乱。此外,新闻报道中某些不准确的用词也产生了概念上的误导。由于上述种种原因,造成了这样的局面:“如果你向10个不同的人请教多媒体一词的含义,你至少会得到10种不同的答案”^[1]。

鉴于上述情况,我们力图用这一章的篇幅,使读者对多媒体和多媒体技术建立起一个比较完整、比较全面的概念。在本章讨论中所涉及的具体技术问题,将分别在后面的有关章节中深入讨论。

1.2 多媒体的概念与含义

1984年美国RCA公司在普林斯顿的David Sarnoff实验室,组织了包括计算机、广播电视和信号处理三个方面的40余名专家,综合了前人已经取得的科研成果,经过4年的研究,于1987年3月在国际第二届CD-ROM年会上展出了世界上第一台多媒体计算机。这项技术后来定名为DVI(Digital Video Interactive)。这便是**多媒体**(Multimedia,由Multiple Media两个词组合而成)技术的雏型。

需要指出,在多媒体计算机展出的前一年,即1986年,由Philips和Sony公司合作研究成功了CD-I(Compact Disc Interactive)技术。这项技术在原理上与DVI很相似,它构成一个独立的光盘播放器,播放器内包含有ROM和解码器。ROM用来存储实时操作系统(软件),解

码器用来将从 CD-ROM 上读取的压缩的数字活动图像信号解码,并还原成模拟电视信号。这个播放器可以直接与电视机相连接显示出电视图像。DVI 与 CD-I 之间的实质性差别在于,前者的编、解码器是置于微机中由微机的操作系统来控制完成计算的,这就把彩色电视技术与计算机技术融合在一起;而后的设计目的,只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电视信号(等同于后来的 VCD 播放器)。这便是在 DVI 技术出现之后,人们就立即失去了对 CD-I 的兴趣的原因。

多媒体技术一出现,在世界范围内立即引起巨大的反响,因为它清楚地展现出信息处理与传输(即通信)技术的革命性的发展方向。国际上在同一年内立即成立了交互声像工业协会,当该组织 1991 年更名为交互多媒体协会 IMA(Interactive Multimedia Association)时,已经有 15 个国家的 200 多个公司加入了。

多媒体计算机与普通计算机有什么不同,它的出现何以如此引人瞩目?早期的计算机只能进行数学运算,后来又具有了文字处理能力,经过若干年的发展之后又增加了图形与动画的功能。而多媒体计算机则是增加了对包括伴音在内的活动图像(即动作连续的电视图像)的处理、存储和显示的能力。这在技术上是一个质变性的飞跃。

为了比较深入地理解这一飞跃,让我们来作一个简单的计算。我国的电视图像在一帧内有 625 行,去掉在扫描逆程(不显示图像的部分)中占去的约 50 行,出现在一幅画面上有效的扫描行数是 576 行左右。电视画面的宽与高尺寸之比是 4:3,要保证图像在水平方向上的单位距离内可分辨的像素数与垂直方向上相等,那么在图像水平方向的像素数应为 $576 \times 4/3 = 768$ 。将这样一幅单色的电视图像数字化,其取样点数应不少于 768×576 。

根据三基色原理,一幅彩色图像是由红、绿、蓝 3 幅单色图像组成的,每秒钟要传送 25 幅彩色图像才能保证电视图像的连续性。如果每一个像素采用 8 比特量化,1 秒钟时间内需要传送的数据量则为

$$768 \times 576 \times 3 \times 8 \times 25 \approx 265 \quad (\text{Mb})$$

8 个比特(bit)是一个字节(Byte),265 Mb 的数据也就是 33 兆字节(MB)。1 部 2 小时电视电影的总数据量为

$$33 \times 60 \times 60 \times 2 = 238 \quad (\text{GB})$$

在 20 世纪 80 年代,计算机硬盘的容量是 40 MB(注意通常用大写 B 表示字节,小写 b 表示比特),要通过微机存储这样一部电影的数据需要 5940 个 40 MB 的硬盘,这显然是没有任何实际意义的。换句话说,普通微机不具有存储和处理未经压缩的连续活动图像的能力。

多媒体计算机的出现,标志着人们已经在有效地利用图像压缩编码的研究成果方面,在对于电视图像数据的实时压缩(使得能在实时的条件下存储、传送活动图像)、实时解压缩(以显示活动图像)等技术问题方面取得了突破性的(具有划时代意义的)进展。

多媒体计算机区别于普通计算机的第二个技术特征,是解决了同时存储、读取和显示两个在时间上紧密相关的数字信号(即伴音信号和图像信号)时,如何在时间上保持同步的问题。同步的必要性很容易理解,屏幕上说话的人嘴形与声音配不上就是不同步。在电视中,图像信号与伴音信号是组合成一个信号传送的,二者之间总是保持着同步。在多媒体计算机中,数字形式的伴音与视频图像信号是可以作为两个信号分别存储与读取的,要保证同步就需要考虑二者的时间关系。对需要考虑时间关系的信号进行编程,要比存取一般的计算机数据复杂。

与存储一般的计算机数据类似,数字形式的电视图像和伴音也存储在数据库(或文件)中。

视频文件可以与用户的解码器同在一台计算机内(多媒体计算机就是这样),也可以放在远距离之外(如在电视台内,或者在租借电视光盘、磁带的影像店内),二者之间用通信网络连接起来,实现从远距离以外调取活动图像与伴音信号,这便出现了**多媒体通信**的概念。

现在我们来讨论一下**多媒体**一词中“多”字的含义。在多媒体计算机出现之前,计算机虽然经历了从最初的数学运算功能,逐步发展到具备了文字编辑、动画制作以及通过电话线路或计算机网络收、发 E-mail 等许多功能,但在这一发展过程中,计算机仍然还是一台处理一般数据的计算机。而多媒体计算机所增加的处理、存储、随机地读取包括伴音在内的电视图像的功能,不仅是在功能数量上又多了一项,而是有一个质变的过程,因为声音和电视图像在传统上是属于通信技术和电视技术的研究对象的。多媒体计算机已经不再是原有意义上的计算机,它已经把计算机延伸到电视技术和通信技术领域了。这是“多”字的一个方面的含意。

前面的一个简单计算已经可以清楚地说明,传送包括伴音在内的电视信号,是当前各种通信手段中传送数据量最大的一种。既然技术已经发展到了能用计算机处理这种数据量最大的业务,其他数据量小的,如电话、传真、可视电话、远程医疗诊断、远程教育等项技术与业务,则都可以融合到一起了。这就意味着在办公室及家庭中,上述所有的业务(包括有线电视)都将在一台多媒体终端上实现。这是多媒体中“多”字的第二层含意。也有人用**集成**一词来描述多媒体将多种媒体、多项业务集合在一起的这一特点。

有的学者从数据本身的特点来理解多媒体技术与其他技术的区别。多媒体数据是由内容上相互关联的文本、图形、图像、声音、动画、活动图像等媒体的数据所形成的**复合数据**。这一数据合成的过程是在计算机控制下完成的^[2]。在这里首先说明了多媒体所涉及的信号是数字化的,而不是模拟的;其次应当注意“内容上相互关联”和“合成”这两个要点,内容上毫不相干的文字、图形、声音、活动图像等数据的集合并不是多媒体数据。通常对时间敏感的声音、活动图像的数据称为**实时数据**,其他类型的数据则是**非实时数据**。如果说某个数据是多媒体数据,则意味着该复合数据中至少包含有一种实时数据和一种非实时数据^[3]。

在 20 世纪 90 年代初期,国外的许多杂志、书籍经常通过与广播电视相比较的方式,通俗地介绍多媒体技术的含意与特点。看电视的时候,只能是电视台播放什么节目人们就看什么节目,观众不能按自己的意愿实时地让电视台播放个人想看的内容,更不能像在自己家中看录像节目那样进行“快进”、“快退”、“静止画面”、“慢动作”、“倒带”等操作。利用多媒体技术则可以这样做,不管图像节目库距离你有多远,你都可以根据“菜单”(目录),调取你想看的节目,而且可以通过遥控器,像使用录像机一样任意地完成上述种种操作。这个例子所代表的技术与业务在多媒体中称为**点播电视 VOD(Video On Demand)**。

这个例子通俗、形象地概括了多媒体技术所具有的另一个基本特征,即**交互性(Interactive)**。人一机交互是计算机固有的技术特征。通过键盘或鼠标打开一个文件,就会显示你所要的内容(比如正在完成的一篇论文),经过修改、补充之后,再写入到给定的文件名下,或者通过一条打印指令在打印机上将文章打印出来。这种人与计算机之间通过指令“对话”式的操作就是交互操作,这已经是人们几十年来所熟知的常识。而在上面介绍的关于 VOD 的例子中,强调的是以交互的方式对活动图像(包括伴音)进行操作。如果一台机器只能以交互的方式处理文字或静止图像,这仍旧是单纯的计算机技术,不属于多媒体技术。

值得注意,人一机交互与传统的通信领域中的**双向通信(Two Ways)**是本质上不同的两个概念。双向通信意味着通信可以在两个方向上进行,如图 1-1(a)所示,北京的电视节目经过线路 1 传到上海,上海的电视节目经过线路 2 传到北京,二者使用了对称的信道,但相互独立。

所谓对称是指两条线路的信道带宽或数据率等传输参数与指标是相同的。而图(b)表示由北京的通信终端自动地从位于上海的数据库内调取电视节目时的交互工作。这项业务使用了双向的非对称信道,上行线(由用户向信息中心)传送指令,是低速率的;下行线传送读取的信号,是高速率的。需要强调的是,这项业务中的上、下行信道中传输的信息是相互关联的,体现了人一机交互操作。

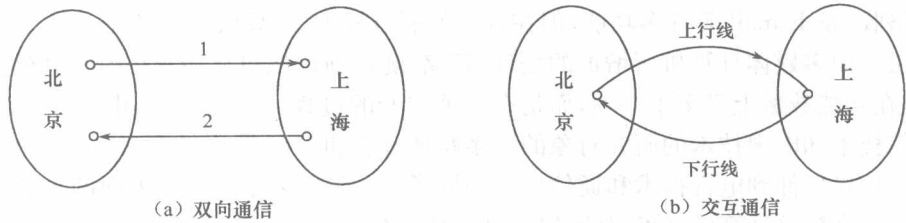


图 1-1 交互工作与双向通信

除了人一机交互之外,一些多媒体系统支持人与人之间的交互,例如,可视电话、多媒体会议和协同工作就是典型的例子。传统的电话也是实现人与人交互的通信系统;支持人与人交互的多媒体系统与电话的最大区别在于,它使用了更多的媒体(除相互听得见外还相互看得见)和允许同时进行更多的业务(如除通话外还可以共同观看、修改文件等)。

通过上面的讨论,我们可以对多媒体的特征作一个简要的总结,即它具有集成性、同步性和交互性。其中集成性包括多种媒体的集成和多种业务的集成;交互性包括人与人的交互和人一机交互。

1.3 多媒体产生的技术背景

一种新技术的产生与发展往往是与其特定的技术背景相联系的,是以其他有关技术的发展作为基础的。实际上,多媒体技术之所以能够在 20 世纪 80 年代末期出现,主要得益于下述几个方面的技术成果。

1.3.1 图像压缩编码技术的成熟

在通信领域中,人人都知道数字通信具有模拟信号通信所无法比拟的优越性。模拟信号在传输过程中产生了失真或者混进去噪声,在接收端难以使其恢复原形。数字信号则不同,因为发出的脉冲信号形状是已知的,如果在传输中产生失真或叠加上噪声,在接收端经过放大、幅度切割等整形处理,失真和噪声则被消除,又恢复成原来的形状。

数字通信的缺点是将模拟信号变为数字信号以后,对信道带宽的要求大幅度增加。以电话为例,一个模拟话路只需要 3.4 kHz 的带宽。变成数字信号时,取样频率取 8 kHz(根据取样定理,取样频率的数值最低不得低于被取样信号的最高频率的 2 倍),每个取样点采用 8 比特量化,一路数字电话的数据率则为 64 kb/s。当用二进制码传输时,每赫兹带宽最高只能传送 2 b/s(采用多进制码传输时,这个数字可以高一些)。可见一路电话从模拟传送改为数字传送,对信道带宽的要求提高了很多。彩色电视所遇到的情况则更为困难。我们在 1.2 节里已粗略地估算了彩色电视信号的数据率。按照国际标准,一路按分量进行编码的彩色电视信号(不包括伴音),编码后的数据率 $R=216 \text{ Mb/s}$,而一路模拟彩色电视信号的带宽只有 6 MHz。