



普通高等学校
电类规划教材
电子信息与通信工程

多媒体通信 技术与应用

◎刘勇 石方文 孙学康 编著



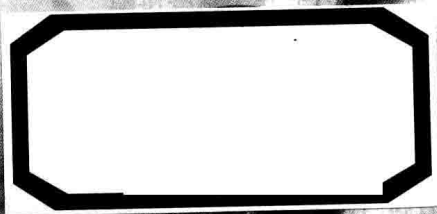
中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS



普通高等学校
电类规划教材
电子信息与通信工程



多媒体通信 技术与应用

©刘勇 石方文 孙学康 编著

人民邮电出版社
北京

图书在版编目 (C I P) 数据

多媒体通信技术与应用 / 刘勇, 石方文, 孙学康编
著. — 北京: 人民邮电出版社, 2017. 8
普通高等学校电类规划教材. 电子信息与通信工程
ISBN 978-7-115-45611-3

I. ①多… II. ①刘… ②石… ③孙… III. ①多媒体
通信—通信技术—高等学校—教材 IV. ①TN919.85

中国版本图书馆CIP数据核字(2017)第190565号

内 容 提 要

本书全面介绍了多媒体通信方面的基本概念、关键技术及多种应用系统。基本内容包括多媒体通信的基本概念及体系结构、数字音频编码、数字图像与视频压缩编码、多媒体通信网络、多媒体流式应用系统与终端、多媒体视频会议应用系统与终端、宽带无线多媒体应用系统与终端。

本书可作为高等院校通信工程、计算机技术等相关专业本科教材或研究生的教学参考书, 也可供从事通信、计算机方面的工程技术人员阅读参考。

-
- ◆ 编 著 刘 勇 石方文 孙学康
责任编辑 李 召
责任印制 陈 犇
 - ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路 11 号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京隆昌伟业印刷有限公司印刷
 - ◆ 开本: 787×1092 1/16
印张: 17.25 2017 年 8 月第 1 版
字数: 418 千字 2017 年 8 月北京第 1 次印刷
-

定价: 49.80 元

读者服务热线: (010)81055256 印装质量热线: (010)81055316

反盗版热线: (010)81055315

广告经营许可证: 京东工商广登字 20170147 号

多媒体通信技术是一门综合的、跨学科的交叉技术，它综合了计算机技术、通信技术以及多种信息科学领域的技术成果。伴随着这些技术领域的发展，多媒体通信技术取得明显进步，展现出广阔的应用前景。本书在介绍音频、视频编解码技术基础上，结合多媒体通信网络，重点介绍了各种多媒体通信应用系统与终端。

本书在内容取材和编写上具有如下特点。

(1) 强化应用内容。本书用3章内容详细阐述了多媒体通信应用系统，包括系统框架结构、关键技术、应用终端及相关实现，具体应用内容如下。

- ① 多媒体流式应用系统与终端。
- ② 多媒体视频会议应用系统与终端。
- ③ 宽带无线多媒体应用系统与终端。

(2) 突出先进技术。本书包括富媒体分发技术、移动网络视频监控等多项多媒体新技术及实用的先进技术。

(3) 循序渐进。多媒体部分内容理论性较强，如音频压缩编码方法、视频压缩编码方法等，因此本书加入了音频技术基础、图像技术基础等内容，由浅入深，便于学生理解、掌握，同时适合自学。

为了便于学习，每一章还提供了内容摘要、小结和习题。

本书第1、4、7章由刘勇编写，第2、6章由石方文编写，第3、5章由孙学康编写。在本书的编写过程中，编者得到北京邮电大学张勖、张碧玲、于翠波、兰丽娜和范春梅老师的热心指导，在此表示衷心的感谢，同时还要感谢周日康、王思远、郝馨、生晓婷、辛雨菡、王晓勤等对本书编写所提供的帮助。

由于时间紧迫，编者学识有限，书中难免存在不足之处，请读者不吝指正。

编 者

2017年8月

目 录

第1章 概述	1	3.1.1 视觉特性	45
1.1 多媒体通信的基本概念	1	3.1.2 彩色电视信号的形成	49
1.2 多媒体通信中的关键技术	4	3.1.3 彩色电视信号的数字化	51
1.3 多媒体通信系统与相关业务	7	3.1.4 图像质量的评价	54
1.3.1 多媒体通信的体系结构及业务类型	7	3.2 图像的统计特性	55
1.3.2 多媒体通信系统及应用	9	3.2.1 离散信源的信息熵	55
1.4 多媒体通信的发展趋势	11	3.2.2 无记忆信源的概率分布与熵的关系	56
小结	12	3.2.3 信源的相关性与序列熵的关系	57
习题	12	3.3 视频压缩编码	58
第2章 数字音频编码	13	3.3.1 数据压缩的性能指标	58
2.1 音频技术基础	13	3.3.2 无失真图像压缩编码方法——熵编码	59
2.1.1 人耳听觉特性	13	3.3.3 限失真图像压缩编码方法	64
2.1.2 音频编码的分类	17	3.3.4 预测编码	65
2.1.3 语音质量评价	21	3.3.5 变换编码	71
2.2 常用压缩编码方法	23	3.3.6 小波变换编码	75
2.2.1 差值脉冲编码调制 DPCM 和自适应差值脉冲编码调制 ADPCM	23	3.4 图像与视频压缩编码标准	78
2.2.2 线性预测编码 LPC	24	3.4.1 静止图像压缩编码标准 JPEG/JPEG2000	78
2.2.3 矢量量化 VQ 编码	25	3.4.2 H.261 和 H.263	78
2.2.4 子带编码	27	3.4.3 MPEG 系列	82
2.2.5 感知编码	28	3.4.4 H.264/AVC	89
2.3 音频压缩编码标准	30	3.4.5 H.265/HEVC	91
2.3.1 波形编码标准	30	小结	94
2.3.2 混和编码标准	32	习题	97
2.3.3 MPEG 音频编码标准	34	第4章 多媒体通信网络	98
2.3.4 MPEG-2 AAC 编码标准	38	4.1 多媒体通信对传输网络的要求	98
2.3.5 MPEG-4 音频标准	40	4.1.1 性能指标	98
2.4 多媒体音频信息文件格式	40	4.1.2 多媒体通信的服务质量	102
小结	43	4.2 网络对多媒体通信的支持简介	105
习题	44	4.2.1 网络类别	105
第3章 数字图像与视频压缩编码	45	4.2.2 互联网组网技术	107
3.1 图像技术基础	45		

4.2.3 无线网络组网技术	120	5.7.2 基于 P2P 的 IPTV CDN 系统	182
4.2.4 核心网络技术——NGN	132	5.7.3 基于 P2P + CDN 的流媒体 应用	185
4.3 传输层协议	139	5.7.4 直播、点播、三重播放业务 应用	188
4.3.1 传统 Internet 传输层协议	139	小结	191
4.3.2 RTP 协议	140	习题	192
4.3.3 RSVP 协议	145	第 6 章 多媒体视频会议应用系统与 终端	193
4.4 多媒体同步	147	6.1 会议系统的应用类型	193
4.4.1 多媒体同步概述	147	6.2 多媒体视频会议系统与终端	194
4.4.2 多媒体数据	148	6.2.1 多媒体视频会议系统的组成 方式	195
4.4.3 多媒体时域特征表示	150	6.2.2 多点会议控制方式	197
4.4.4 多媒体同步的四层参考模型	153	6.3 视听通信业务标准	198
4.4.5 同步多媒体集成语言 SMIL	154	6.3.1 视频会议标准体系	199
小结	155	6.3.2 会话初始协议 SIP	208
习题	156	6.4 视频点播 VOD 系统与设备	217
第 5 章 多媒体流式应用系统与终端	157	6.4.1 VOD 基本概念	217
5.1 流媒体	157	6.4.2 VOD 系统结构	219
5.1.1 流媒体技术特点	157	6.4.3 VOD 系统涉及的关键技术	221
5.1.2 流媒体的传输与控制协议	158	6.4.4 视频服务器	223
5.1.3 流媒体系统结构	159	6.4.5 基于有线电视网的 VOD 系统	225
5.2 IPTV 的基本概念及特点	161	6.4.6 基于电信城域网的 VOD 系统	226
5.3 IPTV 业务及对网络的 QoS 要求	161	小结	229
5.3.1 IPTV 业务	161	习题	229
5.3.2 IPTV 对网络的 QoS 要求	162	第 7 章 宽带无线多媒体应用系统与 终端	230
5.4 IPTV 系统结构	163	7.1 概述	230
5.4.1 IPTV 系统架构	163	7.1.1 3G、WLAN 系统互联结构	230
5.4.2 工作流程	164	7.1.2 富媒体分发技术	235
5.5 IPTV 关键技术	165	7.2 移动智能终端	238
5.5.1 内容分发网络 CDN	165	7.2.1 移动终端硬件平台	238
5.5.2 P2P	168	7.2.2 移动终端软件平台	240
5.5.3 节目内容的发布与呈现	171	7.2.3 终端中间件	241
5.5.4 数字版权管理	173	7.3 移动流媒体业务的实现	242
5.6 IPTV 多媒体应用平台与终端	175		
5.6.1 多媒体应用平台	175		
5.6.2 多媒体计算机终端	177		
5.6.3 IPTV 机顶盒	178		
5.7 IPTV 多媒体应用系统	180		
5.7.1 系统整体框架	180		

7.3.1 移动流媒体业务概述	242	7.4.2 移动网络视频监控系统关键技术	262
7.3.2 数字版权管理	247	小结	264
7.3.3 移动流媒体业务流程	255	习题	265
7.4 移动网络视频监控	260	参考文献	267
7.4.1 移动网络视频监控系统概述	260		

在以信息技术为主要标志的高新技术产业中,多媒体技术开辟了当今世界计算机和通信产业的新领域。多媒体通信技术将原来彼此独立的三大技术领域——计算机、广播电视和通信领域融合起来,进而衍生出多种多媒体通信应用系统,影响着人们生活的方方面面。

本章首先介绍了多媒体技术的概念,并对多媒体通信系统的概念及主要特征、多媒体通信中所涉及到的关键技术、多媒体通信的应用做了详细描述,最后就多媒体通信技术的发展趋势进行了分析。

1.1 多媒体通信的基本概念

在了解什么是多媒体通信之前,应首先了解什么是多媒体,什么是媒体。下面首先介绍媒体的概念。

1. 媒体

“媒体”是信息表示、传输和存储的形式载体。“媒体”的英文是 **Medium**,复数是 **Media**。在通信和计算机领域,媒体有两种含义:一是指传递信息的载体,中文常译为媒介,这一类媒体包括文本、音频、图形、图像和视频等;二是指存储信息的实体,中文常译为媒质,这一类媒体包括磁盘、光盘和半导体存储器等。

根据国际电信联盟电信标准局 ITU-T 建议,媒体可划分为 5 大类。

(1) 感觉媒体 (Perception Medium)

感觉媒体是指人类通过其感觉器官,如听觉、视觉、嗅觉、味觉和触觉器官等直接产生感觉(感知信息内容)的一类媒体,这类媒体包括:声音、文字、图像和视频等。

(2) 表示媒体 (Representation Medium)

表示媒体是指一类用于数据交换的编码类媒体,这类媒体包括:图像编码、文本编码和声音编码等。其目的是为了能有效地加工、处理、存储和传输感觉媒体。

(3) 显示媒体 (Presentation Medium)

显示媒体是指进行信息输入和输出的媒体。输入媒体包括:键盘、鼠标、摄像头、话筒、扫描仪和触摸屏等,输出媒体包括:显示屏、打印机和扬声器等。

(4) 存储媒体 (Storage Medium)

存储媒体是指进行信息存储的媒体。这类媒体包括：硬盘、光盘、软盘、磁带、ROM 和 RAM 等。

(5) 传输媒体 (Transmission Medium)

传输媒体是指承载信息，将信息进行传输的媒体。这类媒体包括：双绞线、同轴电缆、光缆和无线电链路等。

根据 ITU-T 定义，多媒体通信中的“媒体”特指表示媒体，即多媒体通信系统应具有处理、显示、存储和传输多种表示媒体（编码信息）的功能。

2. 多媒体

1987 年，第二届国际 CD-ROM 年会上展出了世界上第一台多媒体计算机，首次将彩色电视技术和计算机技术融合在一起，该技术后定名为数字视频交互 (Digital Video Interactive, DVI)，这便是多媒体 (Multimedia) 技术的雏形。多媒体技术一经出现即在世界范围内引起巨大反响，成为人们关注的热点之一。同年，国际上成立了交互声像工业协会，后于 1991 年更名为交互多媒体协会 (Interactive Multimedia Association, IMA)，其时已有 15 个国家的 200 多个公司加入其中。

多媒体计算机区别于普通计算机的一个主要技术特征是在多媒体计算机中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，能在实时的条件下存储、传送活动图像，并显示活动图像，即实现了对视频信号的实时压缩和实时解压缩。多媒体计算机的第二个技术特征是能够使在时间上有相关性的多种媒体保持同步。在电视系统中，伴音信号与图像信号组合成一个信号进行传送，因此它们始终保持着同步关系，但在多媒体计算机中，伴音信号与视频图像信号可以作为两个信号分别存储，因此只有保证它们在读取、处理和显示过程中的正确时间关系才能使它们保持同步。

在多媒体技术发展的过程中，人们一直试图通过一个准确的定义来描述多媒体技术，但是由于多媒体技术是一种融合技术，其中计算机、彩色电视和通信技术具有复杂性和多样性的特点，由此融合起来而产生的多媒体技术，覆盖面更宽，技术更复杂，很难一言以蔽之，结果是人们从各自的角度出发，根据各自的研究方向给出了不同的多媒体技术的定义。

“多媒体数据是由内容上相互关联的文本、图像、声音、动画和活动图像等媒体的数据所形成的复合数据。”

“所谓多媒体是相对单媒体而形成的概念，是指把多种不同的媒体，如文字、声音、图形、图像和视频等综合集成在一起而产生的一种存储、传播和表现信息的全新载体。”

“所谓多媒体技术就是计算机交互式综合处理多媒体信息——文本、图形、图像和声音，使多种信息建立逻辑连接，集成为一个系统并具有交互性。”

简而言之，多媒体技术就是计算机综合处理声、文、图信息的技术，具有集成性、实时性和交互性的特点。

通常将数字化的活动图像信息存储在数据库中，但当数据库与用户多媒体计算机分开时，用户就需要通过通信网络调用远处数据库中的图像信号和伴音信号，这样多媒体技术便延伸至通信领域，多媒体通信技术应运而生。

3. 多媒体通信

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。从本质上讲,多媒体通信就是把多媒体信息数字化以后,在通信网络上通过数据传输技术进行传输,然后再实现多媒体信息远程应用的过程。在这个过程中,需要维持多媒体信息的多样性和数据的海量性、信息的同步性等特点。

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。一般来说,一个多媒体通信系统应该具备以下三个特征。

(1) 集成性

集成性包括多种媒体的集成和多种业务的集成。

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,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。

此外,在一台多媒体终端上既可以处理各种媒体中数据量最大的电视信号,也可以处理其他数据量较小的多项业务,如电子邮件、信息查询等,从而实现多项业务的融合。

(2) 交互性

交互性包括人一机交互和人与人的交互。

在人一机交互中,系统向用户提供操作界面,用户通过一系列指令使系统实现相应功能,这种人与系统之间“对话”式的操作就是交互操作。在多媒体通信系统中,是对多媒体信息以交互的方式进行操作,例如,对活动图像进行“暂停”“回放”等录像机式的操作。

此外,一些多媒体系统,例如视频会议、可视电话、协同工作等,还支持人与人之间的交互。与传统的支持人与人交互的电话通信系统相比,这些系统使用了多种媒体,不仅包括语音,还包括图像、视频等,并且融合了除通话外的多项业务,如修改文件、共同观看等。

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,在传统的模拟电视系统中,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,因此普通电视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,对播放的全过程可以控制,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,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。同步性是多媒体通信系统中最主要的特征之一。

1.2 多媒体通信中的关键技术

多媒体通信技术是一门跨学科的交叉技术，它涉及到的关键技术有多种，本节我们分别对这些技术做简单介绍，其详细内容我们将在后续章节中阐述。

1. 多媒体数据压缩技术

多媒体信息数字化后的数据量非常巨大，尤其是视频信号的数据量更大，例如一路以分量编码的数字电视信号，数据率可达 216Mbit/s，那么存储 1 小时这样的电视节目需要近 80GB 的存储空间，而欲实现远距离传送的话，则需要占用 108~216MHz 的信道带宽。显然对于现有的传输信道和存储媒体来说，其成本十分昂贵。因此，为节省存储空间，充分利用有限的信道容量传输更多的多媒体信息，需对多媒体数据进行压缩。多媒体数据的压缩包括对视频数据和音频数据的压缩，二者采用的基本压缩技术相同，只是视频信号的数据量比音频数据量大得多，压缩难度更大，所以通常以视频信号为例来讨论多媒体数据的压缩技术。

从图像压缩编码的发展过程看，可以分为两个阶段，即第一代、第二代图像压缩编码方法。第一代图像压缩编码方法以仙农信息论为理论基础，不关心图像的具体内容，主要考虑图像信源的统计特性。这一方法是通过在空间和时间上对图像取样得到的一组像素值来表示图像视频序列（声音则是通过在时间上对波形进行取样得到的一系列样值进行表示），进行压缩时采用一般信号分析的方法消除其中的冗余数据。这种基于像素（或基于波形）的压缩方法即为第一代图像压缩编码方法。第一代图像压缩编码方法于 20 世纪 80 年代初已趋于成熟。“第二代图像压缩编码方法”这一术语出现于 20 世纪 80 年代中期，其编码方法主要用于获得极低码率的压缩图像数据，为此第二代压缩编码方法从研究人类视觉特性出发，通过人眼识别图像所依据的关键特征来构造图像模型。目前第二代技术尚未发展到成熟阶段。

有关图像压缩编码的国际标准主要有：JPEG/ JPEG2000、H.261、H.263、H.264/AVC、H.265/HEVC、MPEG-1、MPEG-2/H.262、MPEG-4、AVS 和 HEVC 等。JPEG 标准是由 ISO 和 ITU-T 组织的联合摄影专家组（Joint Picture Expert Group）于 1991 年提出的用于压缩单帧彩色图像的静止图像压缩编码标准，其后在 2000 年年底，联合摄影专家组又制定了具有更高编码效率的静止图像压缩标准 JPEG2000；H.261 是由 ITU-T 为在窄带综合业务数字网（N-ISDN）上开展速率为 $p \times 64\text{ kbit/s}$ 的双向声像业务（例如可视电话、视频会议）而制定的全彩色实时视频图像压缩标准，其中 $p = 1 \sim 30$ ，因此 H.261 也称为 $p \times 64$ 标准；H.263 是由 ITU-T 制定的低比特率的视频图像编码标准，主要用于 64kbit/s 及以下速率的应用，如可视电话和视频会议；H.264/AVC 是 ISO 活动图像专家组（MPEG）和 ITU-T 的视频编码专家组 VCEG 组成的联合视频组 JVT（Joint Video Team）于 2003 年制定的一个视频压缩编码标准，该标准不仅压缩比高，还具有良好的网络适应能力，能够在恶劣的网络传输条件下提供较高的抗误码性能；MPEG 标准是由 ISO 活动图像专家组（MPEG）制定的一系列运动图像压缩标准，MPEG-1 是为速率为 1~1.5Mbit/s 的数字声像信息的存储而制定的，该标准通常用于提供录像质量（VHS）视频节目的光盘存储系统；MPEG-2/H.262 是由 ISO MPEG 和 ITU-T 于 1994 年共同制定发布的运动图像压缩标准，初衷是提供一个广播电视质量（CCIR 601 格式）的视频信号，后来该标准的适用范围不断扩大，成为能够对图像信号进行不同分辨率和

不同输出比特率的编码的通用标准；事实上 ISO 活动图像专家组最初制定的一系列标准中有 MPEG-3，主要用于提供 HDTV 质量的视频信号，但由于后来 MPEG-2 的适用范围逐渐扩大以致能够支持 MPEG-3 的所有功能，于是 MPEG-3 被取消；MPEG-4 是由 ISO MPEG 制定的，初衷是用于甚低码率（低于 64Kbit/s）应用的一个通用标准，计划采用第二代压缩编码方法，但由于第二代算法还不够成熟，MPEG-4 就转而支持那些已有标准不能覆盖的那些应用，如交互式多媒体服务等；AVS（Audio Video Standard）是由我国制定的一个视频编码国家标准，具有自主知识产权，该标准提出了一系列优化技术，能够以较低的编码复杂度实现与国际标准相当的技术性能；HEVC 又被称为 H.265，是由 ISO MPEG 和 ITU-T VCEG 组成的联合视频编码组 JCT-VC（Joint Collaborative Team on Video Coding）制定的新的视频压缩国际标准，该标准旨在处理更高分辨率和更大尺寸的图像。

音频信号的压缩与图像压缩相比，其不同之处在于图像信号是二维信号，而音频信号是一维信号，数据压缩难度相对较低。在多媒体技术中涉及的声音压缩编码的国际标准主要有 ITU-T 制定的 G 系列标准，如 G711 和 G721、G729 等；MPEG 组织制定的 MPEG-1 和 MPEG-2 音频标准；用于数字电视广播和 HDTV 系统的 AC-3 标准；基于特定应用的地区性编码标准，如移动蜂窝网络中的 AMR 语音编码等。

2. 多媒体数据库及检索技术

多媒体数据库用于存储多媒体数据。传统的关系数据库仅适合存储结构化的数字、文字和数值信息等，但不适合存储非结构化的多媒体数据，其局限性主要体现在：多媒体数据内部有各种复杂的时域、空域以及基于内容的约束关系；需要实时提取音视频数据流，这些数据流通常在时间上有严格要求且数据量很大；若多媒体数据采用分布式存储，则数据库还需将不同数据源的信息进行同步后，再提供给用户。因此多媒体数据需要采用适当的数据结构进行存储，如将关系数据库进行扩充或采用面向对象的数据库来实现。目前多媒体数据库技术仍不成熟，需进一步发展。

此外，对多媒体数据进行检索与查询的相关技术也得到发展，如基于内容的检索和基于语义的检索。其中基于内容的检索是通过多媒体数据中的某些特征检索出具有同样或类似特征内容的多媒体数据。如给出图像中物体的颜色、形状特征，可以检索出具有相同或类似颜色、形状特征的物体的图像来。基于语义的检索则是更高级的检索方式，通过给出“概念”或“事件”等语义，找出具有相同“概念”或“事件”的多媒体数据。例如，给出“拿手机的人”“河流”等概念或事件，可以检索出具有相同或相似语义的图像或视频。基于内容的检索和基于语义的检索代表当前多媒体领域的重点研究方向。

3. 多媒体通信网络技术

网络应用的需求是推动网络技术发展的主要动力。随着视频会议、视频点播、多媒体即时通信、多媒体信息检索与查询等多媒体网络应用的开展，能够满足多媒体应用需要的通信网络必须具有高带宽、可提供服务质量的保证和实现媒体同步等特点。

首先，网络必须有足够高的带宽以满足多媒体通信中的海量数据，并确保用户与网络之间交互的实时性。就单个媒体而言，对网络带宽要求最高的媒体是实时传输的活动图像。其次，网络应提供服务质量的保证，目的是满足多媒体通信的实时性和可靠性的要求。为了使

用户拥有实时的感觉,网络对语音和图像的单程传输延时应在 100~500ms 之间,一般应小于 250ms。而在交互式多媒体应用中,系统对用户指令的响应一般应小于 1~2s。最后,网络必须满足媒体同步的要求,包括媒体间同步和媒体内同步。由于多媒体信息具有时空上的约束关系,例如图像及其伴音的同步,因此要求多媒体通信网络应能正确地反映媒体之间的这种约束关系。

传统网络,无论是通信网、计算机网还是电视广播网,虽然都可以传递多媒体信息,但都不是理想的解决方案。有线电视网络是单向的,不支持多媒体的交互;计算机通信网不提供可靠的服务质量保证;通信网络的技术复杂性高,开销巨大。为了适应多媒体业务的发展需要,有必要进行网络“融合”以提供理想的多媒体业务。

以软交换为核心的 NGN 网络为多媒体通信开辟了更广阔的天地。NGN 网络采用开放的分层体系架构来实现语音、数据和多媒体业务。在这种分层体系架构下,与业务有关的功能独立于与传输有关的技术,各功能部件之间采用标准的协议进行互通,能够兼容 PSTN 网、IP 网、移动通信网等多种网络技术,提供丰富的用户接入手段,支持标准的业务开发接口,并采用统一的分组网络进行传送。虽然 NGN 网络目前仍面临很多问题,但网络融合将成为必然趋势。

4. 多媒体信息存储技术

由于多媒体信息的信息量巨大,因而在多媒体信息传输时,为保证其传输质量必须对其实时性提出较高的要求,同时还需保持媒体间的同步关系。这些特点对多媒体系统的存储设备提出了很高的要求。既要保证存储设备的存储容量足够大,还要保证存储设备的速度要足够快,带宽要足够宽。满足上述要求的存储设备有多种,包括硬盘、光盘和磁带等等。

磁带是以磁记录方式来存储数据的,它适用于需要大容量的数据存储,但对数据读取速度要求不是很高的某些应用,主要用于对重要数据的备份。光盘则是以光学介质来存储信息,光盘的种类有很多,例如,CD-ROM、CD-R、CD-WR、DVD 和 DVD-RAM 等。而以光盘为主存储介质的光盘库存储系统不仅具有巨量的存储特性,还能够实现数据的网络共享。光盘和光盘库的存储容量大,成本低,尤其是光盘更换方便,可以被看作是一种无限容量的存储设备,但是对这种设备中的数据读取有时不能立即得到响应,有时还需人工干预。硬盘及磁盘阵列则具有更快速的数据读取速度。硬盘是电脑中最重要的一种数据存储设备和数据交换媒介,按照接口类型不同,可以分为电子集成驱动器(Integrated Drive Electronics, IDE)和小型计算机系统接口(Small Computer System Interface, SCSI)两种。其传输速率的快慢直接影响计算机系统的运行速度。目前新型增强 IDE 接口的硬盘具有 9~66Mbit/s 的传输速率,SCSI 接口的硬盘传输速率已达 160Mbit/s。虽然硬盘的存取速度已经得到了很大提高,但仍然满足不了处理器的要求。为了解决这个问题,人们采取了多种措施,其中一种就是由美国加州大学伯克利分校的 D.A.Patterson 教授于 1988 年提出的廉价冗余磁盘阵列(Redundant Array of Inexpensive Disks, RAID)。RAID 将普通 SCSI 硬盘组成一个磁盘阵列,采用并行读写操作来提高存储系统的存取速度,并且通过镜像、奇偶校验等措施提高系统的可靠性。由于硬盘及磁盘阵列的读取速度快,因此能够为实时媒体流提供即时读写能力,并支持大量用户同时访问,但是一般价格较高,容量不太大,适用于存储经常被访问的内容。为了进一步提高数据的读取速度,同时获得大容量的存储,存储区域网络(Storage Area Network, SAN)

技术应运而生。SAN 是一种新型网络，由磁盘阵列连接光纤通道组成，以数据存储为中心，采用可伸缩的网络拓扑结构，利用光纤通道有效地传送数据，将数据存储管理集中在相对独立的存储区域网内。SAN 极大扩展了服务器和存储设备之间的距离，拥有几乎无限的存储容量以及高速的存储能力，真正实现了高速共享存储的目标，满足了多媒体应用的需求。

5. 多媒体终端技术

多媒体通信终端是多媒体通信系统的重要组成部分，它面向用户，为用户提供与系统的交互功能，并将多种媒体数据进行集成，通过同步机制将多媒体数据提供给用户。多媒体通信终端应实现信息采集、处理、显示和数据同步等基本功能，其中涉及到的关键技术包括信源编码技术（如图像压缩编码技术）、信道编码技术（如频带传输技术、纠错技术）和信号处理与识别技术（如语音识别技术、人脸识别技术）等。

适用于多媒体通信系统的业务多种多样，不同业务所使用的多媒体终端也各不相同，如多媒体计算机终端、多媒体移动终端以及针对某种特定应用的专用设备，例如机顶盒、可视电话终端设备等。像机顶盒、可视电话终端这样的专用设备，其硬件平台多采用专用集成电路来完成对信息的高速处理，而对于多媒体计算机终端，其硬件系统则是较高配置的计算机主机硬件，输入设备采用鼠标、跟踪球、电子笔、触摸屏、摄像头和视音频采集卡等，呈现给用户图形化的界面，输出手段也非常丰富，可采用声音、图形图像以及活动视频作为信息的显示形式。随着移动通信系统的发展，利用移动终端设备开展多媒体业务势在必行。移动通信终端为用户提供极大的灵活性，使用户能够在任何时间、任何地点进行通信。移动终端在为用户提供通话的同时，还可以看成是一台小型计算机，对多种多媒体应用进行处理，这就要求移动终端具有极大的智能性。移动终端的缺点是利用电池作为电源，其功率和寿命会受到限制，并且由于移动终端体型较小，限制了其存储容量以及数据的处理能力。尽管如此，随着市场需求的不断增加，必将推动多媒体移动终端技术的迅速发展。

1.3 多媒体通信系统及相关业务

1.3.1 多媒体通信的体系结构及业务类型

1. 多媒体通信的体系结构模式

在多媒体通信中，国际电联 ITU-T 在其 I.211 建议中为 B-ISDN 提出了一种适用于多媒体通信的体系结构模式，如图 1-1 所示。

由图 1-1 可知，该体系结构模式自上而下主要包括 5 个方面内容。

(1) 传输网络

传输网络位于体系结构最底层，主要为实现多媒体通信提供最基本的物理环境。传输网络包括局域网（LAN）、城域网（MAN）、广域网（WAN）、ISDN、B-ISDN 和 FDDI 等高速数据网络。在进行多媒体通信

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模式

时，可选择其中的某一种网络，也可将不同网络组合起来使用，这取决于多媒体通信系统的具体应用环境或系统开发目标。

(2) 网络服务平台

网络服务平台的作用是为用户提供各类网络服务，其目的是使传输网络对用户来说是透明的，即用户无需知道底层传输网络如何提供网络服务，只需直接使用这些网络服务内容即可。

(3) 多媒体通信平台

多媒体通信平台向上支持各类多媒体应用，向下与网络服务平台相连。该层主要基于不同媒体的信息结构，为其提供通信支援，如进行多媒体文本信息处理等。

(4) 一般应用

在多媒体通信中，一般应用是指人们常见的一些多媒体应用，如多媒体信息检索、远程协同工作等。

(5) 特殊应用

在多媒体通信中，特殊应用是指如电子邮购、远程医疗等业务性较强的一些多媒体应用。

2. 多媒体通信的业务类型

随着多媒体技术的发展和用户需求的提高，不论是多媒体通信的一般应用还是特殊应用，其业务种类还在不断增加。为此国际电联 ITU-T 在其制定的 F.700 系列标准中，对现有的或即将开展的音像和多媒体业务制定了一个标准架构，如图 1-2 所示。

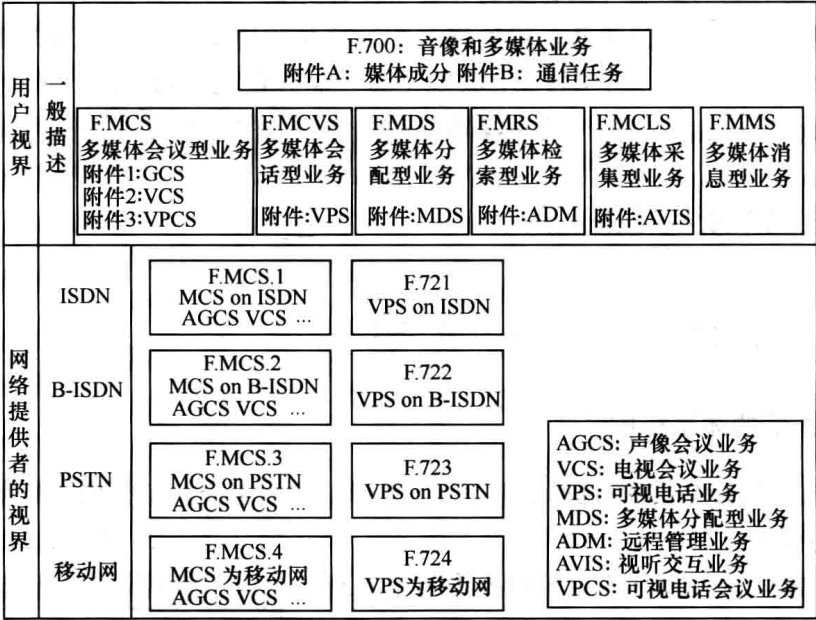


图 1-2 音像和多媒体业务的标准架构

图 1-2 中上部分从用户视界的角度给出了音像和多媒体业务的一般描述，下部分是从网络提供者的视界给出了适用于不同音像和多媒体业务的各种网络的一般描述。

根据 ITU-T 的定义，从用户视界的角度，多媒体通信业务可分为 6 种类型。

(1) 多媒体会议型业务 (Conference Services)

该类业务包括视听会议、声像会议等业务类型,能够进行多点通信,实现人与人之间的双向信息交换。

(2) 多媒体会话型业务 (Conversation Services)

该类业务包括多媒体可视电话、数据交换等业务类型,通过点到点通信,实现人与人之间的双向信息交换。

(3) 多媒体分配型业务 (Distribution Services)

该类业务包括广播式视听会议等业务类型,通过点对多点通信,实现机器与人之间的单向信息交换。

(4) 多媒体检索型业务 (Retrieval Services)

该类业务包括多媒体数据库、多媒体图书馆等业务类型,通过点对点通信,实现机器与人之间的单向信息交换。

(5) 多媒体采集型业务 (Collection Services)

该类业务包括远程监控、投票等业务类型,通过多点到一点通信,实现机器与机器或机器与人之间的单向信息交换。

(6) 多媒体消息型业务 (Message Services)

该类业务包括多媒体文件传送等业务类型,通过点到点通信,实现人-机器-人之间的单向信息交换(存储转发)。

以上多媒体业务,有些特点很相似,因此也可以做进一步的归类,划分为以下4种类型。

(1) 人与人之间进行的多媒体通信业务:会议型业务和会话型业务都属于此类。会议型业务是指在多个地点的人与人之间的通信,而会话型业务则是在两个人之间的通信。另外从通信的质量来看,会议型业务的质量要高些。

(2) 人机之间的多媒体通信业务:多媒体分配业务和多媒体检索业务都属于此类。多媒体检索业务是一个人一台机器的点对点的交互式业务,而多媒体分配型业务是多人一台机器的一点对多点的人机交互业务。

(3) 多媒体采集业务:多媒体采集业务是一种多点对一点的信息汇集业务,一般是在机器和机器之间或人和机器之间进行。

(4) 多媒体消息业务:此类业务属于存储转发型多媒体通信业务。此类多媒体信息的通信不是实时的,需要先将发送的消息进行存储,待接收端需要时再接收相关信息。

在实际应用中,上述业务不仅可以单独存在,也可以进行相关交织处理,在教育科研、军事、医学和娱乐等多个领域为用户提供服务。

1.3.2 多媒体通信系统及应用

一个多媒体通信系统,在其终端与终端之间、终端与多媒体服务器之间均有网络相连。通过网络提供业务的多媒体通信系统可以分为人机交互系统和人与人之间交互的系统,其中人机交互系统包括多媒体信息检索与查询系统、视频点播系统等,人与人之间交互的系统包括多媒体会议与协同工作系统、多媒体即时通信系统等,其中多媒体会议与协同工作系统又包括视频会议系统、多媒体远程教育系统 etc。下面简单介绍几种多媒体通信系统及应用。

1. 多媒体信息检索与查询系统

多媒体信息检索与查询 MIS (Multimedia Information Service) 系统使用户通过输入关键字或者声音、视频能够对相关资料进行查询和浏览。为使用户能够快速检索到相关信息, MIS 系统采用超文本、超媒体技术对信息进行管理。

超文本 (Hypertext) 是一种信息的非线性的组织结构。它采用非顺序的网状结构, 按照人们的“联想关系”组织各信息点。当各信息点不仅有文本信息, 还包括音频、视频等其他媒体数据时, 这种组织结构即为超媒体 (Hypermedia)。超媒体为用户提供了一种在文件内部和文件之间快速有效地检索和查找多媒体信息的方法。

2. 视频点播系统

传统的有线电视系统, 其模式为电视台单向播放节目, 用户被动接收。视频点播 (Video on Demand) 系统则可以为用户提供不受时空限制的交互点播, 使用户能够随时点播自己希望收看的节目。该系统将节目内容存储在视频服务器中, 随时根据用户的点播要求取出相应的节目传送给用户。用户点播终端可以是多媒体计算机, 也可以是电视机配机顶盒。

视频点播系统是一个开放式平台, 可以集成多种多媒体应用, 广泛应用于远程教育、数字图书馆、新闻点播和网上购物等。

3. 视频会议系统

视频会议 (Video Conference) 系统是一种实时的、点到多点的多媒体通信系统。能将音频、视频、图像、文本和数据等集成信息从一个地方通过网络传输到另一地方。

视频会议系统基于计算机网络, 使处于异地的多个会场构成一个会议环境, 从而可召开视频会议。这样不同会场的与会者既可以听到对方的声音, 又能看到对方的形象以及对方展示的文件、实物等, 同时还能看到对方所处的环境, 使与会者具有身临其境的感觉。

4. 远程教育系统

远程教育系统是以现代传媒技术为基础的多媒体应用系统, 学生通过通信网络实时或非实时地接收教师上课的内容, 包括教师的声音、图像以及电子教案。如果是实时的远程教学, 学生还可以随时向教师提出疑问, 教师可以马上回答, 并且根据需要, 教师也可以看到学生的图像和声音, 从而模拟学校的课堂授课方式。对于非实时的教学, 教师可以将自己授课的内容做成课件放到网上, 学生可以在自己希望的任何时间和地点按照自己的学习速度和方式来学习。

5. 多媒体即时通信系统

即时通信系统 (Instant Messaging System, IMS) 可以使用户通过网络传递即时的短消息, 消息格式可以是文本、图片, 甚至视频等。用户既可以采用文本短消息进行交流, 还可以进行文件传输、音频或视频信息的即时沟通。

即时通信系统通常与呈现 (Present) 服务相结合。呈现服务是一种辅助通信手段, 使通信双方能够了解对方的状态, 如在线、隐身和离线等。通过呈现服务, 一方面用户可以使自