



**DUOMEITI TONGXIN**

# 多媒体通信

沈瑞琴 晏蓉 主编



**中国铁道出版社**  
CHINA RAILWAY PUBLISHING HOUSE

高等职业教育教材

# 多媒体通信

沈瑞琴 晏蓉 编

中国铁道出版社

2008年·北京

## 内 容 简 介

本书介绍了多媒体音视频信号的压缩编码原理,多媒体通信的协议,包括 IP 电话、视频会议、视频点播、视频监控等几种重要的多媒体通信系统的构成和工作原理。

本书适用于高职计算机网络和通信专业的学生,也可供相关工程技术人员参考。

### 图书在版编目(CIP)数据

多媒体通信/沈瑞琴,晏蓉编. —北京:中国铁道出版社,2007.12

ISBN 978-7-113-08434-9

I. 多… II. ①沈…②晏… III. 多媒体-计算机通信  
IV. TN919.85

中国版本图书馆 CIP 数据核字(2007)第 194616 号

书 名:多媒体通信

作 者:沈瑞琴 晏 蓉 编

---

责任编辑:张永国 电话:010-51873135 电子信箱:zhyg704@sina.com

封面设计:薛小卉

责任校对:孙 玫

责任印制:金洪泽

---

出版发行:中国铁道出版社(北京市宣武区右安门西街 8 号,100054)

印 刷:北京海淀五色花印刷厂

版 次:2008 年 3 月第 1 版 2008 年 3 月第 1 次印刷

开 本:787 mm×1092 mm 1/16 印张:14 字数:351 千

印 数:1~3000 册

书 号:ISBN 978-7-113-08434-9/G·270

定 价:26.00 元

---

### 版权所有 侵权必究

凡购买铁道版的图书,如有缺页、倒页、脱页者,请与本社读者服务部调换。

电 话:市电(010)51873170 路电(021)73170(发行部)

打击盗版举报电话:市电(010)63549504 路电(021)73187



## 前言

电话、视频电话、视频监控等系统,其实完全可以用传统的通信领域的成熟技术实现,而且传输质量优良,但这种系统的缺点是灵活性差、代价高。由于互联网飞速发展,网络基础设施迅速普及,计算机终端处理能力日益强大,通信网与互联网相互借鉴、相互融合,使得多媒体通信系统在互联网上的实现,不但成为可能,而且有自己的优势,成为了互联网上重要的应用之一。多媒体信号使网络世界变得如此有声有色,绚丽缤纷。

第一章“概述”介绍了多媒体通信的概念、系统和关键技术。

音视频信号数据相对于文本数据,“体积”要大得多,传输同样的信息,文本的单位可以是 b/s,音频的单位就是 kb/s,而视频的单位就得是 Mb/s。所以以音视频信息为基本内容的多媒体数据信号,一定要压缩以后才能在带宽有限的网络上传输。介绍各种音视频信号的压缩方法,就是第二章“多媒体信号处理”的内容。

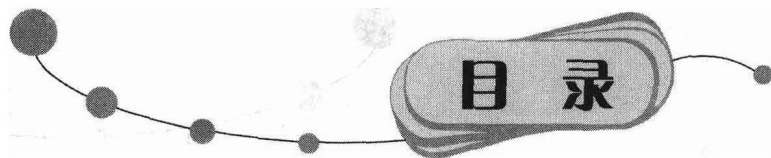
通信网络中设备之间的动作需要沟通和协调,这就是通信信令的作用。互联网中的多媒体通信协议与通信信令作用类似,它们完成诸如呼叫对方、建立连接、拆除连接等作用。目前的多媒体通信协议主要有 H. 323 和 SIP,在第三章“多媒体通信协议”中进行了详细地介绍。

多媒体通信系统中比较典型的、需要信令的有 IP 电话、视频会议系统,它们立足于广域网。视频点播、视频监控系统则相对简单,系统主要运行在局域网或城域网上。在第四章“多媒体通信应用”中介绍了这些系统的组成和工作原理。

本书由沈瑞琴、晏蓉合编。第一、二章由沈瑞琴编写;第三、四章由晏蓉编写。多媒体通信涉及通信、计算机、网络各个领域,知识面宽,其中的许多技术还在不断变化。由于编者水平有限,书中可能会有不妥或谬误,恳请读者指正。

编 者

2007 年 8 月于南京



# 目 录

<b>第一章 多媒体通信概述</b> .....	1
第一节 多媒体基本概念.....	1
第二节 多媒体通信系统.....	2
第三节 多媒体通信的关键技术.....	5
第四节 多媒体通信业务分类.....	9
练习和思考题.....	11
<b>第二章 多媒体信号处理</b> .....	12
第一节 多媒体数据压缩原理.....	12
第二节 音频压缩编码标准及技术.....	32
第三节 视频压缩编码标准及技术.....	50
第四节 流媒体技术.....	94
练习和思考题.....	116
<b>第三章 多媒体通信协议</b> .....	118
第一节 H. 323 协议简介.....	118
第二节 SIP 协议简介.....	131
练习和思考题.....	140
<b>第四章 多媒体通信应用</b> .....	141
第一节 VoIP.....	141
第二节 视频会议系统.....	158
第三节 视频点播系统.....	173
第四节 视频监控系統.....	187
练习和思考题.....	218
<b>参考文献</b> .....	219

# 第一章 多媒体通信概述

随着科学技术的迅速发展和社会需求的日益增长,人们已不满足于单一媒体提供的传统的单一服务,如电话、电视、传真等,而是需要诸如数据、文本、图形、图像、音频和视频等多种媒体信息以超越时空限制的集中方式作为一个整体呈现在人们的眼前。与此同时,广域网、局域网等通信网络技术和电话、广播、电视等各种视听技术也取得了较大的技术进展,于是,在通信网络技术、多媒体计算机技术和视听技术三者结合的基础上,产生了多媒体通信。

## 第一节 多媒体基本概念

多媒体通信是一门新技术,由于计算机界和通信界都在研究这门新技术,两个学科对有的概念定义不是很统一。

### 一、媒体和多媒体

若从通信的角度出发,所谓“媒体”是指信息传递和存取的最基本的技术和手段。简单地说,媒体即指信息的载体。

国际电信联盟(ITU-T)将媒体分为5类:

1. 感觉媒体(Perception Medium):由人类的感覺器官直接感知的一类媒体。这类媒体有声音、图形、动画、运动图像和文本等。
2. 表示媒体(Representation Medium):为了能更有效地加工、处理和传输感觉媒体而人为构造出来的一种媒体,或者说是对信息进行数据处理的各种编码。如图像编码(JPEG、H. 261、MPEG等)、文本编码(ASC II、GB2312)和声音编码(G. 723. 1、G. 729)等。
3. 显示媒体(Presentation Medium):进行信息输入和输出的设备。这类媒体有显示屏、打印机、扬声器等输出设备,以及键盘、鼠标、扫描器、触摸屏等输入设备。
4. 存储媒体(Storage Medium):进行信息存储的设备。这类设备有硬盘、光盘、软盘、磁带、ROM、RAM等。
5. 传输媒体(Transmission Medium):用于承载信息,将信息进行传输的传输介质。这类媒体有同轴电缆、双绞线、光纤和无线电链路等。

根据国际电信联盟电信标准部(ITU-T)的定义,多媒体通信中的媒体特指表示媒体,也就是多媒体通信系统中要有存储、传输、处理、显现多种表示媒体信息(即多种编码的信息)的功能。

ITU-T对多媒体(Multimedia)服务的定义,是特指能处理多种表示媒体的服务。多媒体系统和多种媒体系统是不同的,多媒体系统中的媒体相互之间是有关联的,是以时空同步的方式存在的;而多种媒体系统中,媒体与媒体之间可以是毫无关系的。两者间的重要区别在于媒体间的同步性。

## 二、超文本和超媒体

### 1. 超文本

超文本是由节点和链构成的信息网络。

把文本按内部固有的独立性和相关性划分成不同的基本信息块,称为节点。节点是表达信息的单位,是围绕一个特殊主题组织起来的数据集合。节点的内容可以是文本、图形、动画、音频、视频等,也可以是一般的计算机程序。

链用于固定节点间的信息联系,它以某种形式将一个节点与其他节点连接起来。就像文章中出现的“注”,由“注”可以找到一段与之相关的文字或文章,这种由“注”而链接到一段文字或文章的链称为超级链。

超文本是一个非线性的网络结构,具有多种媒体信息、网络结构形式及交互特性。

超文本是一种典型的数据库技术,用户可以对超文本的信息网络进行浏览、查询和注释等操作。

超文本是一种接口模型,它采用“控制按钮”的方式组织接口,“按钮”就是连接节点的链,用户通过按钮访问信息。

### 2. 超媒体

超文本主要以文字的形式表示信息,建立的链接关系主要是文句之间的链接关系。超媒体是超文本的扩展,除了使用文本外,还使用图形、图像、声音、动画或影视片断等多种媒体来表示信息。

## 第二节 多媒体通信系统

### 一、多媒体通信

多媒体通信是指多媒体信息的存储、处理、交换和传输。多媒体通信是多媒体信息处理技术和组网技术的融合。多媒体通信技术的发展打破了传统的单一媒体通信方式和单一电信业务的通信系统格局,反映了通信向高一层次发展的趋势。多媒体通信技术是一种综合技术,涉及多媒体技术、计算机技术、通信技术等多个领域。

### 二、多媒体通信系统

多媒体通信系统是指完成多媒体通信业务的系统,包括多媒体通信终端、通信网络及设备、多媒体应用等设备。多媒体通信系统具有分布式协同多媒体环境,能够通过网络完成多媒体信息的处理和传输,并支持交互式以及广播和多播方式。

#### 1. 多媒体通信网络及设备

多媒体通信网络是多媒体信息传输的载体。多媒体通信网络设备包括网络的交换设备和传输设备。多媒体通信对信息的传输和交换都提出了更高的要求,网络带宽、交换方式及通信协议都直接影响多媒体通信的质量。

多媒体通信网络要求对业务的比特率、传输延迟、延迟抖动和误码率等提供保障,同时能够提供多播和缓冲等功能。多媒体通信网络应具有:能够同时支持音频、视频和数据传输;交换节点的高吞吐量;足够的有效带宽;良好的传输性能,具备呼叫连接控制、拥塞控制、服务质量控制和网络管理功能。各类媒体信息对网络传输能力的要求如表 1-1 所示。

表 1-1 各类媒体信息对网络传输能力的要求

多媒体信息	最大时延(s)	最大时延抖动(s)	平均吞吐量(Mb/s)	可接受的误比特率	可接受的误分组率
音频	0.25	10	0.064	$<10^{-1}$	$<10^{-1}$
视频	0.25	10	100	$<10^{-2}$	$<10^{-3}$
压缩视频	0.25	1	2~20	$<10^{-6}$	$<10^{-9}$
数据文件	1.00	—	2~100	0	0
实时数据	0.001~1.00	—	$<10$	0	0
图形、静止图像	1.00	—	2~10	$<10^{-4}$	$<10^{-9}$

## 2. 多媒体应用设备

多媒体应用设备主要指多点控制单元(MCU)、流媒体服务器、视频服务器、应用共享服务器等。服务器是一种对视/音频数据进行压缩、存储、处理及传输的专用计算机设备。多点控制单元(MCU)是视频会议系统中多点间视频会议信号之间的切换和控制的设备。

## 3. 多媒体通信终端

多媒体通信终端就是用户实现多媒体通信业务的终端设备。只有利用相应的终端设备,用户才能与某种服务或服务的提供者之间进行交互操作。目前的多媒体终端有安装了数据解码和媒体播放功能的计算机系统、安装了机顶盒的电视机系统、H. 320 终端、H. 323 终端、SIP 终端等。

## 三、多媒体通信系统的主要特征

多媒体通信系统的主要特征体现在如下几个方面。

### 1. 集成性

多媒体通信系统中的集成性是指能对如下 4 类信息进行存储、传输、处理、显现的能力。

#### (1) 内容数据(Contentdata)信息

在多媒体通信系统中,信息是以某一种结构的形式存在的,典型的结构有两种,一种是客体结构,其中可处理的最小单元为客体(Object);另一种是文件结构,其中可处理的最小单元为文件。

在这些结构化的信息中,信息由结构框架和结构内容两部分组成。可以形象地将结构化信息看作是装有东西的一个容器,结构框架为容器本身,结构内容为容器中装有的东西。内容部分是真正要传送的实质所在,一般称内容部分的信息为“内容数据信息”。内容数据信息是用单一媒体的标准编码来表示的信息。它包括文本、二维和三维图形、静止图像(连续色调)、二值图像、声音(语音、音乐、噪声)和活动图像(动画片、有声或无声的运动图像)。

#### (2) 多媒体和超媒体信息

多媒体信息和超媒体信息与单媒体信息不一样,它们是结构化的信息,由结构框架和内容数据两部分组成。多媒体与超媒体信息的最小表达形式有两类:一类称为客体;另一类称为文件。

#### (3) 脚本(Script)信息

脚本信息是一组特定的用语意关系联系起来的结构化的多媒体和超媒体信息。它需要提供表示这一组多媒体信息的运作过程和与外部处理模块间的关系。脚本信息提供了应用层级的同步机制,因而十分重要。



#### (4)特定的应用信息

上面所述的信息是3类低层信息,可以由标准来定义和表示。“特定的应用信息”是高层信息,它是与应用密切相关的,它将随应用场合的不同而有很大的不同。它不像前3类信息那样有一般性的表示方法,它的表示方法是基于上述3类的基础之上的。

一个常用的典型例子是目录信息,基于目录信息可以检索到所需的多媒体或超媒体信息,因而目录信息将是按照信息类型的不同(如文档、客体、文件、文本、数据分组等)来分类的,并用内在的关系互相联系起来。这样,用户就可以在检索所需信息前,先利用目录信息来检索所需信息的位置。目录信息就是典型的特定的应用信息。

#### 2. 交互性

交互性是指在通信系统中人与系统之间的相互控制能力。在多媒体通信系统中,交互性有两个方面的内容。其一是人机接口,也就是人在使用系统的终端时,用户终端向用户提供的操作界面;其二是用户终端与系统之间的应用层通信协议。

人机接口(Man-machine Interface)是系统向用户提供的操作界面。目前最好的能用于多媒体通信系统的人机接口为基于视窗(Windows)的人机接口界面。视窗人机接口是一种基于图符的接口方式,它可以提供菜单、按钮、选择框、列表项、对话框、输入域、敏感区等多种复杂的人机接口,以满足多媒体通信系统复杂的交互操作的需要。

用户终端与系统之间的应用层通信协议。在多媒体通信系统中要存储、传输、处理、显示多种表示媒体,而这些表示媒体之间又存在着复杂的同步关系,不同的表示媒体可能以串行的形式传送给用户,也可以以并发的形式传送给用户,以便让用户终端能按照同步关系来复现出多媒体信息。很显然,在多媒体通信系统中,单信道的通信协议就不够用了,需要能支持多信道同时工作的多信道通信协议。在多信道通信协议中,除了要建立一条主信道来支持系统的核心交互工作之外,还要建立起若干条辅助信道来提供并发信息的传送,从而实现完善的多媒体通信的交互过程。

多媒体通信终端的用户对通信的全过程有完备的交互控制能力,这是多媒体通信系统的一个主要特征,也是区别多媒体通信系统还是非多媒体通信系统的一个主要准则。例如,视频点播(VOD, Video on Demand),用户不仅可以观听其显现的图像、声音和文字。还可以对其全过程进行有效的控制。

#### 3. 同步性

多媒体通信系统中的同步性是多媒体通信系统中最主要的特征之一。同步性指的是在多媒体通信终端上显现的图像、声音和文字是以同步方式工作的。例如:用户要检索一个重要的历史事件的片断,该事件的运动图像(或静止图像)存放在图像数据库中,其文字叙述和语言说明则是放在其他数据库中。多媒体通信终端通过不同传输途径将所需要的信息从不同的数据库中提取出来,并将这些声音、图像、文字同步起来,构成一个整体的信息呈现在用户面前,使声音、图像、文字实现同步,并将同步的信息送给用户。

对多媒体通信系统来说,以上3个特征必须是并存的,是缺一不可的。缺少其中之一就不能称其为多媒体通信系统。

多媒体的集成性、交互性、同步性对标准化依赖性很强,如果没有一个国际性的统一标准,多媒体通信的发展和广泛应用是不可能的。提出建议和制定多媒体标准的组织很多,主要为ITU-T(国际电信联盟电信标准部)、ISO(国际标准化组织)和IEC(国际电子技术委员会)等机构,共同研究多媒体的标准化的问题。

### 第三节 多媒体通信的关键技术

在多媒体通信发展过程中涉及到许多问题,其中影响最为显著,同时也是多媒体通信的难点所在的关键技术有多媒体信息处理技术、多媒体通信的同步技术、多媒体通信的网络技术、多媒体通信的终端技术和多媒体数据库技术等。

#### 一、多媒体信息处理技术

多媒体信息处理技术即压缩编码技术。多媒体信息的最大特点是数据量大,其中视频信息的数据量最大。目前,在多媒体信息压缩技术中最为关键的就是音、视频压缩编码技术。一个普通彩色电视节目的速率为  $768 \times 575 \times 25 \times 2 \times 8 = 170 \text{ Mb/s}$ ,一部 2 h 普通彩色电视节目的总数据量  $768 \times 575 \times 25 \times 2 \times 8 \times 3600 \times 2 \div 8 = 160 \text{ GB}$ ,对于高清晰度电视 HDTV 的速率高达  $1000 \text{ Mb/s}$ ,其数据量更加庞大(表 1-2)。因此如果没有压缩编码技术就无法实现多媒体通信。

经过多年的努力,视频压缩技术逐渐成熟,出现了 ITU-T 制定的 H. 261、H. 263、H. 264 和 ISO 制定的 MPEG-1、MPEG-2、MPEG-4、MPEG-7 等一系列视频压缩的国际标准。HDTV 经过压缩后的速率只需  $20 \text{ Mb/s}$  甚至更低,语音信号的压缩技术也得到很大的发展,一路  $64 \text{ Kb/s}$  的语音信息经过压缩后的速率可降低到  $8 \text{ Kb/s}$ ,  $5 \sim 6 \text{ Kb/s}$  甚至  $2.4 \text{ Kb/s}$ 。

表 1-2 各种信息媒体传输的数据量和存储容量

媒 体	传输的数据量要求(Kb/s)	对存储器的存储容量要求(KB/min)
文本	2.4~9.6	18~72
静止图像	32~64	240~480
语音	9.6~128	72~960
高保真声音	64~1500	48~11500
压缩视频	1.5~6000	11500~45000
高质量压缩视频	6000~24000	45000~180000

#### 二、多媒体通信的同步技术

在多媒体通信系统中,可以在如下 3 个层面上实现同步。

##### 1. 链路层级同步

链路层级同步是通过信息流帧结构的特殊设计来实现的。信息流的帧结构又按照不同的应用场合分为两类。第一类是用于会话型(会议型)点与点之间的实时通信,为满足会话的要求,应尽量减少延迟,因而采用比特交织的帧结构;第二类是用于存储读出系统,其应用场合有点播电视、运动图像检索、数字录像机、VOD 等,这种应用场合可以允许一个较大的固有延时而不会造成信息质量的下降,因而采用块交织的帧结构。

##### 2. 表示层级同步(复合客体和超级链)

表示层级同步是通过在客体(或文件)复合过程中引入同步机制和超文本组合过程中引入同步机制来实现的。在多媒体通信系统中,客体(或文件)是可以处理的最小信息单元。

一段文字可以是一个客体,一段语音或音乐可以是一个客体,同样,一幅画面或一段运动

图像片断也可以是一个客体,这种没有复杂结构的客体称为简单客体。另一类客体在多媒体通信系统中也可以当作一个信息单元来处理,但它是有结构的,是由若干客体按某种规律组合而成的,这类客体称为复合客体。在将不同表示媒体的客体复合成一个复合客体的过程中引入同步机制,构成多媒体复合客体;或者用超级链在将不同表示媒体的客体链接过程中引入同步机制,构成超媒体。上述两个过程均在表示层级完成,故称为表示层级同步。表示层级同步有5种类型:绝对时间同步、相对时间同步、链式同步、循环同步及条件同步。

#### (1)绝对时间同步

在复合客体的复合过程中或者超媒体由超级链的链接过程中,确定各客体间的时间同步关系是以初始客体的时间为基准的。

#### (2)相对时间同步

在复合客体的复合过程中或者超媒体由超级链的链接过程中,确定各客体间的时间同步关系是以相对于前一客体的时间为基准的。

#### (3)链式同步

在复合客体的复合过程中或者超媒体由超级链的链接过程中,许多客体构成一个链,客体一个接着一个出现,这就是链式同步。

#### (4)循环同步

在复合客体的复合过程中或者超媒体由超级链的链接过程中,一个客体或某子客体按预定要求出现两次或多次,这种同步称为循环同步。

#### (5)条件同步

在复合客体的复合过程中或者超媒体由超级链的链接过程中,只有当某个条件满足才能引发一个客体的出现,这种同步方式称为条件同步。

### 3. 应用层级同步

多媒体通信系统中,最高一级的同步是应用层级的同步。应用层级同步采用的技术为脚本同步技术。在多媒体通信系统中,信息存储、处理的最基本单元是客体或文件,当然,这里的客体包含简单客体和复合客体(简单文件或复合文件)。一般来说,用了复合客体后,简单的多媒体通信已能进行,但要实现复杂的功能齐全的多媒体通信还远远不够,必须引入应用层级同步即脚本同步。脚本是一种特殊的文本,它用语意关系将多媒体或超媒体的运作过程和外部处理模块联系起来构成脚本信息,从而实现完善的多媒体通信。当然,这样讲过于抽象,举一个例子来说就清楚了。以电影为例,在一部电影里有许多演员,每个演员有许多不同的台词片断和各种各样的场景镜头,如果这些东西零零星星地放着,显然什么也不是,要把它们变成一部电影,还要有一个电影脚本将它们有序地联系起来。

在多媒体通信系统中,情况就非常类似,电影中的演员、台词片断和场景镜头,在多媒体通信系统中就像一个个客体(包括复合客体和简单客体),要将一个个客体组成完整的有声有色的多媒体信息,就要用脚本将它们联系起来。在多媒体通信系统中,脚本同步是最高一层的同步,也是十分重要的一级同步。

在远程教学系统中,教师不仅讲课,还可以提问;学生不仅听课,还可以向教师提问。这种远程教学的交互性,在一般的广播电视教学中是不可能的。在这个过程中,相互之间交流的信息内容也是十分丰富的,有语音、图像、文字、还可能有CAI(计算机辅助教学)的应用软件,这体现了多媒体通信的集成性。在教学过程中,教和学双方通过对方的图像、声音以及表述的文字等信息同步的表现才能达到和谐的交流,这体现了多媒体通信的同步性。可以想象,这种远程

教学的效果是令人满意的,因为它生动、直观、交互性强,因此能充分调动学生的学习积极性。

多媒体通信系统中,同步可以在3个层面上实现。这并不是说,一个多媒体通信系统必须同时具有这3个层面的同步,但它必须至少用到其中一种同步方式。当然,同步方式用得越多,系统的性能就越完善。

### 三、多媒体通信的网络技术

多媒体通信网络就是能够处理和传送文本、图形、图像、视频和音频等交互式多媒体信息的网络。它是多媒体计算机技术、通信网络技术和视听技术相互融合、相互渗透的产物。

多媒体通信网络不同于传统的数据通信网络,两者的区别如表1-3所示。

表 1-3 多媒体通信网络与传统的数据通信网络比较

比较项	多媒体通信网络	传统的数据通信网络
传送的数据类型	多媒体信息	传统的、简单的数据类型
数据率	高	低
信息传送方式	面向数据流的高度突发性	突发性
可靠性	可能有一些损失	无损失
延时性要求	低延时	不要求
通信模式	点对点 and 点对多点	点对点
同步性	同步性高	不要求

多媒体通信网络技术包括宽带网络技术以及接入网技术。传统的网络(X.25、DDN、帧中继、PSTN)都是单业务的网络,不适合直接多媒体传输。在多媒体通信系统中,网络上传输的是多种媒体综合而成的一种复杂的数据流,它不但要求网络对信息具有高速传输能力,还要求网络具有对各种信息的高效综合能力。在目前看来,以ATM技术为核心的B-ISDN是多媒体通信的理想网络。这是因为它具有以下特点:以固定长度(53 B)的信元进行高速交换,网络时延小;能处理突发性信息,可以动态分配带宽;收发端通过虚电路VC进行连接,并保证提供网络和终端连接建立时所商定的服务质量QOS,如最小带宽、最大时延、延时变化和服务类型等;通过建立多条虚电路可以实现多点连接等。

但是从网络的发展来看,在IP网络上实现多媒体通信是世界各国的主要目标。这是因为IP网络采用TCP/IP协议(统一上层通信协议,支持任何现有网络)而应用日益广泛。然而IP网络是面向非连接、非实时、无QOS的网络,其带宽不易控制、时延不能保障、QOS不能保证等特点又不利于多媒体通信业务的发展,因此必须解决这些相关问题。到目前为止已经采取了增加协议、IP宽带化、增加QOS等措施,例如为保证实时业务的运行,IP网络采用了实时传送协议(RTP)和实时传送控制协议(RTCP);为给实时业务或者其他业务提供足够的带宽,IP网络采用了资源预留协议(RSVP);快速提高IP骨干网的速率等。目前IP网络上的业务正在从非实时业务向实时业务演变,从窄带业务向宽带业务发展,宽带IP网络正向着满足多媒体通信需求的方向发展。事实上,IP网络已经被列为下一代网络(NGN)的核心网。

接入网是目前通信网中的一个瓶颈,虽然全光网、无源光网络(PON)、光纤入户(FTTH)被认为是理想的接入网,但光终端设备价格偏高、无源光网络的稳定性和实用性等问题还没有完善地解决。而现阶段大量的窄带双绞铜线因为价格低廉而得到广泛的应用。在此基础上改造的宽带接入网络如高速数字用户线(HDSL)、不对称数字用户线(ADSL)、甚高速数字用户

线(VDSL)等技术已获得较大的成功。另外还可以充分利用 CATV 网络的带宽资源,使其适应多媒体通信业务的传输,如混合光纤同轴(HFC)、交换型数字视频系统(SDV)、交互型数字电视系统(IDV)等。

#### 四、多媒体通信的终端技术

多媒体通信终端是能集成多种媒体信息,能对多媒体信息实现同步,并具有交互功能的通信终端。它必须完成信息的采集、处理、同步、显示等多种功能。

##### 1. 基于 PC 的多媒体终端

在基于 PC 的多媒体终端上可实现如视频点播(VOD)、高清晰电视(HDTV)、远程教学、远程医疗等宽带多媒体业务。

目前,多媒体终端设备主要可分为两大类:

一类是多媒体 PC 机结合媒体播放软件而成的软终端,它以其功能强大、实现简单而成为目前支持宽带多媒体业务的主要终端设备。这种终端的缺点是可移动性差。另一类是由专业硬件芯片制成的个人多媒体终端。是一种新型的支持宽带流媒体业务的终端,具有可移动性、易操作性和目的专一性。这种个人多媒体终端是实现任何人(Whoever)在任何时间(Whenever)、任何地点(Wherever)与世界上任何人(Whoever)进行任何方式(Whatever)通信(称为个人通信)的最佳设备。当然个人多媒体终端应该包括以下几个方面功能:

(1)它是一部功能强大的电话终端,能够支持目前普通 POTS 电话终端或移动电话终端的基本业务和大部分的补充业务,例如振铃、播放回铃音、忙音、DTMF 等信号音。

(2)为了支持宽带流媒体,它必须具有编解码功能,特别是解码功能,才能对所接收的流媒体进行解码,从而实时播放。由于要考虑到个人多媒体终端的性能一般都没有 PC 机强大,因而为了保证解码的速度,解码功能一部分可以由专业的 DSP(数字信号处理)芯片完成,例如可在个人多媒体终端中集成可以对 MPEG4、MP3、REAL 等的解码芯片。除了这些专业解码芯片外,个人多媒体终端为了支持基本的 VOIP 语音通话,还必须具备支持语音编码(如 G. 711, G. 729, G. 723. 1 算法)的编码与解码硬件 DSP 芯片或软件包。

(3)为了能够缓存所要播放的媒体流,个人多媒体终端必须具备较大的存储空间,一般来说,至少具有 16 MB 的内存空间。

(4)为了能够运行个人多媒体终端内的各种协议栈和软件包,终端必须具备功能比较强大的专业嵌入式 CPU,并且为了有效统一地管理各种软硬件资源,最好在个人多媒体终端内配备嵌入式操作系统。

(5)个人多媒体终端还需要具备相关的视、音频播放装置,例如彩色液晶屏等。

(6)为了能够支持宽带传输,个人多媒体终端还需要配备高速以太网接口,如 100 MB 或 1 000 MB 的网口。

(7)为了支持流媒体的接收,个人多媒体终端还必须具备基本的 TCP/IP 协议栈,传输流媒体的实时传输协议,如 RTP、RTCP、RTSP 等。终端为了能够与应用服务器进行协商,从而建立会话,还必须支持 SIP、H. 323 或 MGCP 信令控制协议。除此之外,智能电话还需要支持 HTTP、RADIUS 等相关协议。

##### 2. 信息家电(网络家电)

网络冲浪电视机能使用户边看电视节目,边在互联网上巡游,并可以把与电视节目有关的网址下载,还可通过 E-mail 传送视频、音频及静态图像,也可连接打印机,将所需网络的资料

打印出来；

网络可视电话与家用电脑的网络连接装有因特网软件声卡、麦克风扬声器和视像屏幕，不仅可以在网上面对面聊天，还能通过因特网拨号到世界各地直通长途电话，收费便宜；

网络信息冰箱则以游戏机上网为先导，冷冻室门是 38 cm 液晶显示屏，可接收有线电视或网络服务的信息，并可通过主页传递家庭电子邮件或本地区各类信息；

网络全球收音机通过通信卫星接收全球 1 000 多家广播电台的信号，调频清晰，音质尤佳，并可以进入因特网接收各种综合服务的数据信息……

网络家电的国际市场可谓色彩斑斓，日新月异。

## 五、多媒体数据库技术

数据库是指某实体相关的一个可控制的数据集合，而数据库管理系统(DBMS)则是由相关数据和一组访问数据库的软件组合而成的，它负责数据库的定义、生成、存储、存取、管理、查询和数据库中信息的表现(Presentation)等。传统的 DBMS 处理的数据类型主要是字符和数字。传统的数据库管理系统在处理结构化数据、文字和数值信息等方面是很成功的。但是，随着技术的发展，各种非结构化数据(如图形、图像和声音等)的大量出现，传统的数据库信息系统就难以胜任了，因此需要研究和建立能处理非结构化数据的新型数据库——多媒体数据库。多媒体数据库管理系统(MMDBMS)不但要对传统数据库管理系统的功能加以改进，还要增加一些新的功能。

多媒体数据库的基本技术主要包括：多媒体数据的建模、数据的压缩/还原技术、存储管理和存取方法、用户界面技术和分布式技术等等。

为了适应技术的发展和应用的变化，MMDBMS 应该具有开放的体系结构和一定的伸缩性，同时 MMDBMS 还需要满足如下要求：具备传统数据库管理系统的功能；具备超大容量存储管理能力；有利于多媒体信息的查询和检索；便于媒体的集成和编辑；具备多媒体的接口和交互功能；能够提供统一的性能管理机制以保证其服务性能等等。

## 第四节 多媒体通信业务分类

### 一、交谈型业务(点-点、多点间)

#### 1. 视频会议

视频会议的目的就是利用多媒体通信网络和多媒体终端，使身处异地的与会者就同一议题参与讨论。与会者不仅可以听到发言者的声音，而且还能看到发言者的图像及背景，同时还可以交流有关该议题的数据、文字、图表等信息。现在，大多数的视频会议业务是通过电信网络或计算机网络召开的，它的会议控制信息交流是由专门的多点控制单元(MCU, Multi-Control Unit)来担任的，而网络中的每个节点都具有发起、召集一个会议的功能。会议期间，每个节点都可显示多个窗口：视频窗口显示远程节点的活动彩色视频图像；共享窗口提供所有与会节点的共享显示区，使与会人员可共同浏览、修改和编辑文件，共享鼠标，共同绘制各种图表，对共享区内容的修改等操作结果可被大家同时看到；个人窗口是用户私有区域，系统应允许用户在参加会议的同时，在私有区域做一些私人的工作，如准备发言稿、做笔记或进行注释等；控制窗口用来控制会议的召开、会议的响应、节点的发言请求、数据发送等会议事务，它还提供操作菜单，用于对答窗口的调整、信号质量的调节等。

## 2. 网络教育

现代信息网络的快速发展催生了以交互式多媒体宽带网络为教育信息传输平台的网络教育。所谓网络教育,就是利用多媒体通信网络来实现时空分离的由教师和学生共同来完成的教育、教学活动,以有效促进学生学习,从而达到预期学习目标的一种新型教育方式。网络教育除具备教与学相互分离的特征外,更体现了个性化教育。

网络教育几乎融入了 20 世纪 80 年代以来通信信息领域所有的最新技术,其建立在计算机技术、网络技术、多媒体技术、双向电子通信技术及教育传播理论和现代学习理论紧密结合的基础上,以多媒体计算机网络承载和传播包括图、文、声、像在內的覆盖整个教学系统的各种要素信息,形成交互式的教学模式。网络教育可以实现实时“面对面”远程授课、视频课件点播、同步课业辅导、远程交流讨论、交互式答疑等丰富的教学功能,突破了课堂教学和课本教学信息单一化的局限,能充分开发、组合和利用各种教育信息资源,将多学科、多层次的丰富信息通过多种途径传播,有利于全方位培养人才,有利于学习者全面发展。

## 3. 远程诊疗

多媒体通信网络的建立与网络视频技术的发展,为远程诊疗开辟了一个广阔的应用领域。身处现代医疗卫生中心的医生可以通过多媒体通信网传输的视频信息为远在千里的病人提供诊断服务。通过多媒体终端,医生不仅可以“面对”病人进行观察和询问,同时还可以通过远端的医疗传感器或仪表对病人进行多项病理检查,检查的结果可立即传送到中心,为医生诊断提供依据。如果需要,通过远程医疗网络还可组织各地的医疗专家为患者进行会诊,讨论医疗方案。

## 二、检索型业务(人-机之间)

### 1. 网上直播

网上直播是网络视频应用的一个热点,它使观众摆脱了对电视的依赖。中国中央电视台每年的“春节联欢晚会”除了电视直播外,还同步在网上直播,世界各地的观众只要登录中央电视台的网站,就可以直接在多媒体计算机上收看。越来越多的网站也采用网上直播的形式同步报道国内外的重大事件,大大加快了视频信息传递的时效性。目前,网上直播大多采用流媒体技术。流媒体技术可以在低宽带环境下提供高质量的音、视频;特别是流媒体技术中的智能流技术,可以根据网络带宽的动态变化而自动地调整播放质量,避免在播放过程中的信息流中断。

### 2. 网上视频点播

网上视频点播(WVOD, Web-based Video On Demand)就是用户在视频网站上选择自己喜爱的节目表单,并将选择的节目表单信息提交到视频服务器;视频服务器根据用户的选择将存放在节目库中的视频检索出来,合成一个个视频数据流,并通过宽带网络传输到用户端;用户利用安装在多媒体计算机上的网络视频播放器(如 Windows Media Player、RealOne Player)进行播放。网上视频点播是一种融合了多媒体计算机技术、网络通信技术和视听技术的新技术。

网上视频点播利用网络和视频技术的优势,彻底改变了用户被动收看节目的状态,实现了节目的按需收看和任意播放,为用户提供实时、交互和按需点播服务。网上视频点播系统主要由视频服务器、传输与交换网络、客户端多媒体计算机组成。目前,很多新闻传媒机构都提供各种视频新闻节目,许多开展远程教育的院校也将教师课堂授课实况制成网络视频存放在视

频服务器中,允许学习者按需点播学习。

在网上视频点播的视频服务中心必须解决的一个突出问题是:众多的用户可能在同一时刻要求传送同一影片,并分别要求进行快进、快退等录像机式的操作,而该影片在数据库中只存有一份拷贝,一旦视频服务中心为某个用户开始提供服务以后,则向该用户传送的数据流就不允许有任何中断。这就需要研制出一种比较理想的视频服务器,它应具有良好的管理功能和相应的硬件结构。

### 三、采集型业务(多点向一点、多点向多点汇聚信息)

采集型业务(Collection Services)是一种机器与机器或人与机器之间多点向一点汇聚信息的业务。

#### 1. 监控系统(多点向一点汇聚)

监控系统是一种防范能力较强的综合系统,是安全防范系统的重要组成部分。监控系统以其直观、方便、信息内容丰富而广泛应用于各种场合,如城市交通监控系统、智能楼宇监控系统、工业自动化监控系统等。

#### 2. 数字式医院(多点向多点汇聚)

数字式医院综合先进的医疗设备、通信网络和整合的医疗信息系统,对临床工作的流程带来根本性的变化,从而提高病人护理质量,减少医疗错误,节省时间和费用。其基本特点是“三化、四通、四性”。三化即无纸化、无胶片化和无线化;四通即医疗、药品、经济和人力资源四类信息的通畅;四性即网络模式下医院信息具有实时性、准确性、共享性和全面性的特点。当然数字式医院还可实现远程医疗、资料查询等其他类型的业务。

### 四、消息型业务(存储转发)

消息型业务主要有多媒体邮件系统。

普通的电子邮件(E-mail)只有传送文件的功能,而多媒体邮件系统除可传送文件外,还具有传送图像、视频、音频的功能。多媒体邮件系统对于跨越时区或不同语种的地区特别重要。当发送者估计接收者不在场,或在信息可以被延迟发送而不要求交互通信的情况下,发送者可以在建立、编辑某个邮件以后,通过多媒体邮件系统将它传送到对方的邮箱中;而接收者可以在任何方便的时候打开邮箱,取出邮件。

与上述3种系统相比,多媒体邮件对通信的要求是最低的。因为它是一种非实时的存储转发系统。

## 练习和思考题

1. 什么是媒体? 媒体共分几类?
2. 什么是多媒体? 什么是超媒体?
3. 多媒体通信的主要特征是什么?
4. 多媒体通信涉及哪些关键技术?
5. 多媒体通信业务主要分哪几类?
6. 结合自己的体会,说说在现实生活中有哪些多媒体应用。



## 第二章 多媒体信号处理

### 第一节 多媒体数据压缩原理

通常,把在时间和幅度上都是连续的信号称为模拟信号。时间和幅度都用离散的数字表示的信号就称为数字信号。由于数字信号具有易于处理、存储、远距离传送、没有积累失真、可高品质地被还原等优点,随着信息技术的发展,通信技术已由模拟时代全面转向数字时代。

模拟信号数字化至少要经过采样、量化、编码 3 个过程。为了接收端能够正确地还原,采样时必须满足采样定理,即采样频率  $f_s \geq 2f_m$  ( $f_m$  是模拟信号的最高频率)。

多媒体的原始信号如声音信号、运动图像信号,它们都是模拟信号。将它们数字化转换后数据量非常庞大,为了使多媒体通信正常有效的工作,必须对其进行适当的处理,这就是多媒体数据的压缩编码。

#### 一、数据压缩的基本概念

##### 1. 多媒体数据压缩的目的

多媒体数据包括视频数据、音频数据、文本数据等。视频数据包括图像数据和图形数据;而图像数据又包括静止图像数据和运动图像数据,图形数据包括单幅图形和动画。音频数据是指语音数据、音乐数据、声音数据。文本数据是对文字和字符进行编码而得到的数据。

多媒体数据尤其是视频数据、音频数据在没有压缩前的数据量是相当大的,表 2-1 列出了几种视频信号的原始数据率和压缩后的数据率,将这样的数据信号直接去存储或传输,不仅要求海量的存储器、大宽带的传输信道,而且在经济上造成了高成本。因此多媒体信息的压缩技术是多媒体通信领域中的关键技术之一。

表 2-1 视频信号压缩前后的数据率

应用种类	比特数/像素	像素数/行	行数/帧	帧数/s	亮色比	b/s(压缩前)	b/s(压缩后)
HDTV	8	1 920	1 080	30	4:1:1	1.18 Gb/s	20~25 Mb/s
普通电视 CCIR601	8	720	480	30	4:1:1	167 Mb/s	4~8 Mb/s
会议电视 CIF	8	352	288	30	4:1:1	36.5 Mb/s	1.5~2 Mb/s
桌上电视 QCF	8	176	144	30	4:1:1	9.1 Mb/s	128 Kb/s
电视电话	8	128	112	30	4:1:1	5.2 Mb/s	56 Kb/s

CD-ROM 的容量大约是 650 MB,双层双面 DVD 的容量大约有 17 GB,这样,对于未压缩的电视信号,CD-ROM 仅可存储 23.5 s,DVD 光盘仅可存储大约 15 min 节目。

多媒体数据压缩的目的,简单地讲就是为了最有效地利用有限资源(包括存储器资源、信道资源、计算资源等)。或者说,压缩就是用尽可能少的比特数来表示源信号并能将其还原。因此,压缩的任务就是保持信源信号在一个可以接受的状况的前提下把需要的比特数减到最少程度,以此来减少存储、处理和传输的成本。