



北京市高等教育精品教材立项项目

# 多媒体 通信技术

DUOMEITI  
TONGXIN JISHU

孙学康 石方文 刘勇 编著



北京邮电大学出版社  
[www.buptpress.com](http://www.buptpress.com)

北京市高等教育精品教材立项项目

# 多媒体通信技术

DUOMEITI TONGXIN JISHU

孙学康 石方文 刘勇 编著

北京邮电大学出版社

·北京·

## 内 容 简 介

本书内容包括多媒体通信的概念、数字音频技术、图像技术基础、电视技术基础、视音频编码标准、视频信息压缩与处理、多媒体数据信息的获取、存储与显示技术、多媒体数据同步技术、多媒体数据库及其检索技术、多媒体通信网络技术、数据的分布式处理,最后介绍了各和实用的多媒体应用系统。

本书可用作高等院校通信工程、计算机通信专业本科生的专业教材或研究生的教学参考书,也可供从事通信、计算机方面工作的工程技术人员参考。

### 图书在版编目(CIP)数据

多媒体通信技术/孙学康,石方文,刘勇编著. 北京:北京邮电大学出版社,2006

ISBN 7-5635-1203-9

I. 多... II. ①孙...②石...③刘... 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(2006)第010303号

---

书 名:多媒体通信技术

编 著:孙学康 石方文 刘 勇

责任编辑:王琴秋

出版发行:北京邮电大学出版社

社 址:北京市海淀区西上城路10号(100876)

北方营销中心:电话 010-62282185 传真 010-62283578

南方营销中心:电话 010-62282902 传真 010-62282735

E-mail: publish@bupt.edu.cn

经 销:各地新华书店

印 刷:北京源海印刷有限责任公司

开 本:787mm×960mm 1/16

印 张:22

字 数:507千字

印 数:1 3000册

版 次:2006年5月第1版 2006年5月第1次印刷

---

ISBN 7-5635-1203-9/TN·437

定价:32.00元

• 如有印装质量问题,请与北京邮电大学出版社营销中心联系 •

# 编者的话

多媒体通信技术是一门综合的、跨学科交叉技术,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,正在不断发展和完善。相信多媒体通信的广泛应用将极大提高人们的工作效率,减轻社会的交通运输负担,对人们传统的教育和娱乐方式产生革命性的影响。

本教材分析透彻,结构合理,语言流畅,通俗易懂,但并未因此降低对本科学生的要求。全书在取材和编写上具有以下特点:

(1) 内容全面,包括数字音频技术、图像技术基础、电视技术基础、视音频编码标准、视频信息压缩与处理、多媒体数据信息的获取、存储与显示技术、多媒体数据同步技术、多媒体数据库及其检索技术、多媒体通信网络技术、数据的分布式处理和多媒体应用系统。

(2) 有些内容有一定难度,如有关图像信号的数字化、图像信号和语音信号的压缩编码、图像处理技术等方面的知识。因此,本教材加入图像与语音技术基础方面的内容,以帮助学员掌握相关知识。为便于自学,教材中凡涉及到难点、重点和新知识点的部分都增加了相关的基础知识,并在每一章编排了内容摘要、小结和习题等内容。

(3) 内容先进。教材中包括多媒体新技术以及实用先进技术,如流媒体技术、IP组播技术、多媒体业务在软交换网络中的实现、网络广播技术和IPTV技术等。

本书可用作通信专业本科生教材或研究生的教学参考书,也可供从事通信、计算机方面工作的工程技术人员参考。

本书第3、4章由孙学康编写,第2、5、8章由石方文编写,第1、6、7章由刘勇编写。

在本书的编写过程中,得到北京邮电大学李文海教授的热心指导,在此表示衷心的感谢。同时还要感谢为本书编写和出版给予帮助的各位同事。

由于时间紧迫,学识有限,书中的不足之处在所难免,请读者不吝指正。

编者  
2006年5月

# 目 录

## 第 1 章 多媒体通信技术概述

|                       |    |
|-----------------------|----|
| 1.1 多媒体通信的基本概念 .....  | 1  |
| 1.2 多媒体通信中的关键技术 ..... | 3  |
| 1.3 多媒体通信的应用 .....    | 9  |
| 小结 .....              | 10 |
| 习题 .....              | 11 |

## 第 2 章 音频技术基础

|                                     |    |
|-------------------------------------|----|
| 2.1 声学基础知识 .....                    | 12 |
| 2.1.1 音频信号特性 .....                  | 12 |
| 2.1.2 人耳听觉特性 .....                  | 13 |
| 2.2 音频信息编码分类 .....                  | 16 |
| 2.2.1 波形编码 .....                    | 16 |
| 2.2.2 参数编码 .....                    | 17 |
| 2.2.3 混和编码 .....                    | 18 |
| 2.3 常用压缩编码方法 .....                  | 20 |
| 2.3.1 差值脉冲编码调制自适应差值<br>脉冲编码调制 ..... | 20 |
| 2.3.2 线性预测编码 .....                  | 21 |
| 2.3.3 矢量量化编码 .....                  | 22 |
| 2.3.4 子带编码 .....                    | 24 |
| 2.3.5 感知编码 .....                    | 25 |
| 2.4 音频信息压缩编码标准 .....                | 27 |
| 2.4.1 波形编码标准 .....                  | 27 |
| 2.4.2 混和编码标准 .....                  | 29 |
| 2.4.3 MPEG 音频编码标准 .....             | 32 |
| 2.4.4 MPEG-2 AAC 编码标准 .....         | 36 |
| 2.4.5 MPEG-4 音频标准 .....             | 38 |
| 2.5 多媒体音频信号文件格式 .....               | 39 |
| 小结 .....                            | 41 |
| 习题 .....                            | 42 |

## 第 3 章 图像技术基础

|                            |    |
|----------------------------|----|
| 3.1 视觉特性 .....             | 43 |
| 3.2 图像质量的评价 .....          | 49 |
| 3.3 图像信号数字化 .....          | 51 |
| 3.3.1 图像信号的表述 .....        | 51 |
| 3.3.2 取样和二维取样定理 .....      | 52 |
| 3.3.3 量化与编码 .....          | 54 |
| 3.3.4 取样、量化对图像质量的影响 .....  | 56 |
| 3.4 电视技术基础 .....           | 57 |
| 3.4.1 电视系统的组成 .....        | 57 |
| 3.4.2 彩色电视信号的形成与传送原理 ..... | 58 |
| 3.4.3 视频信息的数字化 .....       | 63 |
| 3.4.4 数字电视系统的结构及特点 .....   | 66 |
| 小结 .....                   | 73 |
| 习题 .....                   | 76 |

## 第 4 章 视频信息压缩与处理

|                           |    |
|---------------------------|----|
| 4.1 图像统计特性 .....          | 77 |
| 4.1.1 图像的信息量 .....        | 77 |
| 4.1.2 离散信源 .....          | 78 |
| 4.1.3 图像的信息熵 .....        | 78 |
| 4.2 信息压缩方法及其分类 .....      | 81 |
| 4.2.1 图像信息中存在的冗余类型 .....  | 81 |
| 4.2.2 图像编码的基本过程 .....     | 83 |
| 4.2.3 压缩编码方法及其分类 .....    | 83 |
| 4.2.3.1 数据压缩技术的性能指标 ..... | 85 |
| 4.3 无失真图像压缩编码方法 .....     | 86 |
| 4.3.1 哈夫曼编码 .....         | 86 |
| 4.3.2 游程编码 .....          | 88 |
| 4.3.3 算术编码 .....          | 89 |
| 4.4 限失真图像压缩编码方法 .....     | 91 |
| 4.4.1 率失真函数 .....         | 92 |

|                                    |     |                                         |     |
|------------------------------------|-----|-----------------------------------------|-----|
| 4.1.2 预测编码和变换编码 .....              | 93  | 小结 .....                                | 194 |
| 4.1.3 矢量量化编码 .....                 | 109 | 习题 .....                                | 195 |
| 4.1.4 子带编码 .....                   | 111 |                                         |     |
| 4.5 几种新型的图像压缩编码技术 .....            | 116 | <b>第6章 多媒体通信网络技术</b>                    |     |
| 4.5.1 小波变换编码 .....                 | 116 | 6.1 多媒体通信对传输网络的要求 .....                 | 196 |
| 4.5.2 分形编码 .....                   | 119 | 6.1.1 概述 .....                          | 196 |
| 4.5.3 模型编码 .....                   | 121 | 6.1.2 多媒体传输网络的性能指标 .....                | 197 |
| 4.6 数字图像处理 .....                   | 122 | 6.2 网络对多媒体通信的支持简介 .....                 | 200 |
| 4.6.1 数字图像处理系统模型及其<br>主要处理方法 ..... | 122 | 6.2.1 多媒体通信网络的现状 .....                  | 200 |
| 4.6.2 图像增强 .....                   | 123 | 6.2.2 IP网 .....                         | 201 |
| 4.6.3 图像复原 .....                   | 125 | 6.2.3 IP城域网 .....                       | 210 |
| 4.7 图像压缩技术标准 .....                 | 127 | 6.2.4 ATM广域网 .....                      | 216 |
| 4.7.1 图像种类与编码标准 .....              | 127 | 6.2.5 以软交换为核心的下一代网络 .....               | 219 |
| 4.7.2 静止图像压缩编码标准<br>JPEG .....     | 128 | 6.3 多媒体通信协议 .....                       | 232 |
| 4.7.3 H.261与H.263 .....            | 131 | 6.3.1 IPv6 .....                        | 233 |
| 4.7.4 MPEG系列 .....                 | 136 | 6.3.2 组播路由选择协议与因特网<br>组管理协议(IGMP) ..... | 241 |
| 4.7.5 H.264 .....                  | 151 | 6.3.3 实时传输协议(RTP) .....                 | 249 |
| 小结 .....                           | 153 | 6.3.4 资源预留协议(RSVP) .....                | 253 |
| 习题 .....                           | 158 | 6.3.5 实时流协议(RTSP) .....                 | 256 |
|                                    |     | 小结 .....                                | 257 |
|                                    |     | 习题 .....                                | 258 |
| <b>第5章 多媒体通信系统中的关键技术</b>           |     | <b>第7章 多媒体数据的分布式处理</b>                  |     |
| 5.1 多媒体信息输入输出及存储技术 .....           | 159 | 7.1 分布式多媒体系统 .....                      | 259 |
| 5.1.1 视频信息输入输出技术 .....             | 159 | 7.1.1 分布式多媒体系统的概念 .....                 | 259 |
| 5.1.2 音频信息的输入输出技术 .....            | 164 | 7.1.2 分布式多媒体系统的结构 .....                 | 260 |
| 5.1.3 语音识别技术和语音合成技术 .....          | 166 | 7.1.3 分布式多媒体系统的服务模型 .....               | 261 |
| 5.1.4 其他输入输出技术 .....               | 169 | 7.1.4 分布式多媒体系统的特征 .....                 | 261 |
| 5.1.5 多媒体存储技术 .....                | 170 | 7.2 DMS的操作及管理 .....                     | 262 |
| 5.2 超媒体技术 .....                    | 175 | 7.2.1 分布式多媒体系统的QoS管理 .....              | 263 |
| 5.2.1 超文本技术 .....                  | 176 | 7.2.2 分布式多媒体系统的操作系统 .....               | 267 |
| 5.2.2 超媒体的组成 .....                 | 178 | 7.2.3 分布式多媒体系统的终端设备 .....               | 273 |
| 5.2.3 超媒体系统结构 .....                | 180 | 7.2.4 分布式多媒体系统的媒体服务器 .....              | 281 |
| 5.2.4 分布式超媒体系统WWW .....            | 182 | 7.3 流媒体技术 .....                         | 284 |
| 5.3 多媒体同步技术 .....                  | 183 | 7.3.1 流媒体概述 .....                       | 284 |
| 5.3.1 多媒体同步概述 .....                | 183 | 7.3.2 流媒体开发平台简介 .....                   | 286 |
| 5.3.2 多媒体数据 .....                  | 184 | 7.3.3 流媒体的调度技术 .....                    | 290 |
| 5.3.3 多媒体时域特征表示 .....              | 187 | 小结 .....                                | 294 |
| 5.3.4 多媒体同步的四层参考模型 .....           | 191 | 习题 .....                                | 295 |
| 5.3.5 同步多媒体集成语言SMIL .....          | 193 |                                         |     |

|                                |     |
|--------------------------------|-----|
| <b>第 8 章 多媒体通信应用系统</b>         |     |
| 8.1 概 述 .....                  | 296 |
| 8.2 多媒体视频会议系统 .....            | 297 |
| 8.2.1 多媒体视频会议系统发展概述 .....      | 297 |
| 8.2.2 多媒体视频会议系统的<br>组成方式 ..... | 298 |
| 8.2.3 多点会议控制方式 .....           | 301 |
| 8.2.4 多媒体会议国际标准 .....          | 302 |
| 8.2.5 互联网环境下的多媒体<br>会议系统 ..... | 311 |
| 8.2.6 我国的视频会议网 .....           | 313 |
| 8.3 视频点播(VOD)系统 .....          | 317 |
| 8.3.1 概述 .....                 | 317 |
| 8.3.2 VOD 系统结构 .....           | 319 |
| 8.3.3 VOD 系统涉及的关键技术 .....      | 320 |
| 8.3.4 视频服务器 .....              | 322 |
| 8.3.5 基于有线电视网的 VOD 系统 .....    | 325 |
| 8.3.6 基于电信城域网的 VOD 系统 .....    | 326 |
| 8.4 网络电视(IPTV) .....           | 329 |
| 8.4.1 概述 .....                 | 329 |
| 8.4.2 IPTV 业务需求说明 .....        | 330 |
| 8.4.3 IPTV 的基本结构 .....         | 331 |
| 8.4.4 IPTV 的相关技术 .....         | 332 |
| 8.4.5 利用电信城域网实现 IPTV .....     | 333 |
| 8.5 多媒体电视监控系统与报警系统 .....       | 335 |
| 8.5.1 概述 .....                 | 335 |
| 8.5.2 远程多媒体监控系统的<br>工作原理 ..... | 337 |
| 8.5.3 远程多媒体监控系统结构 .....        | 338 |
| 8.5.4 多媒体远程监控系统具体应用 .....      | 340 |
| 小结 .....                       | 341 |
| 习题 .....                       | 312 |
| 参考文献 .....                     | 313 |

# 第 1 章 多媒体通信技术概述

21 世纪的人类社会是信息化的社会。信息正影响着人们学习、工作和生活的各个方面。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,广泛影响着人类的的生活和工作。多媒体技术的真正潜力在于与通信网络的融合,它是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——融合起来而形成的一门崭新技术。经过近十几年的迅速发展,出现了多种多媒体通信应用系统。

本章首先介绍多媒体技术的概念,并对多媒体通信系统的概念及主要特征、多媒体通信中涉及到的关键技术和多媒体通信的应用作了详细描述,最后分析了多媒体通信技术的发展趋势。

## 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,应首先了解什么是媒体及什么是多媒体。

### 1. 媒体

媒体是指信息传递和存储的最基本的技术和手段,即信息的载体。媒体的英文是 Medium,复数是 media。根据原 CCITT(国际电报电话咨询委员会)的定义,媒体可划分为五大类:

#### (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、气味和冷、热等。

#### (2) 表示媒体(Representation Medium)

表示媒体是指用于数据交换的编码表示,这类媒体包括图像编码、文本编码、声音编码等。其目的是有效地加工、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、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、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。



在多媒体技术中的“多媒体”，通常是指感觉媒体的组合，即声音、文字、图像、数据等各种媒体的组合。

## 2. 多媒体 (Multimedia)

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DVI(Digital Video Interactive)，这便是多媒体技术的雏形。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，后更名为交互多媒体协会(IMA, Interactive Multimedia Association)，其宗旨在于推动多媒体产业的发展并减少现存对多媒体技术应用的障碍，现已有280多个公司加入。

在多媒体技术发展的这十几年间，人们一直试图通过一个准确的定义来描述多媒体技术。但由于多媒体技术是一种融合技术，其中的计算机、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，其覆盖面更宽，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。目前公认比较准确的概念是由Lippincott和Robinson于1990年2月在《Byte》杂志上发表的两篇文章中给出的“多媒体技术”的定义：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。

多媒体技术最简单的表现形式就是多媒体计算机。多媒体计算机相对于普通计算机的一个根本不同点在于，在多媒体计算机中增加了对活动图像(包括伴音在内)的处理、存储和显示的能力，其硬件配以声卡、视频采集卡等。其主要特征体现在它能够有效地对电视图像数据进行实时的压缩和解压缩，并能够使在时间上有相关性的多种媒体保持同步。

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。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。

一般来说，同时具有以下三个特征的通信系统方可称为多媒体通信系统：

### (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。

## 1.2 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,它涉及的关键技术有多种,本节分别对这些技术作简单介绍,其详细内容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,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216 Mbit/s,存储1小时这样的电视节目需要近80 GB的存储空间,而要实现远距离传送,则需要占用108~216 MHz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。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,须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,可以分为三个阶段,即第一代、第二代、第三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基础,考虑图像信源的统计特性,采用预测编码、变换编码、矢量量化编码、子带编码、小波变换编码、神经网络编码等方法。第一代图像压缩编码方法于20世纪80年代初已趋于成熟,目前利用第一代技术对视频图像的压缩可以得到8~48 kbit/s的最低码率。第二代图像压缩编码方法充分考虑了人眼的视觉特性,从人类的主观特性出发,采用基于方向滤波的图像编码方法和基于图像轮廓一纹理的编码方法,利用第二代图像压缩编码方法可以获得极低码率的图像数据。目前,第二代技术尚未发展到成熟的阶段。第三代图像压缩编码方法考虑到了图像传递的景物特征,采用分形编码方法和基于模型的编码方法,其中基于模型的压缩编码方法是日前最活跃的研究领域,代表着新一代的压缩编码方向。

有关图像压缩编码的国际标准主要有JPEG、H. 261、H. 263、MPEG-1、MPEG-2/H. 262、

MPEG-4 等。JPEG 标准是由 ISO 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。H. 261 是由 ITU-T 第 15 研究组为在窄带综合业务数字网 (N-ISDN) 上开展速率为  $p * 64 \text{ kbit/s}$  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。H. 263 是由 ITU-T 制定的低比特率视频图像编码标准, 用于提供在  $30 \text{ kbit/s}$  左右速率下的可接受质量的视频信号。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  $1 \sim 1.5 \text{ Mbit/s}$  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 活动图像专家组和 ITU-T 第 15 研究组于 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的视频信号, 后来该标准的适用范围不断扩大, 成为能够在不同传输速率下对图像信号进行编码的通用标准。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 即“甚低速率 (低于  $64 \text{ kbit/s}$ ) 视听编码标准”, 共有 4 个版本, 目前第一、二版已制定完成 (1998 年第一版, 1999 年第二版), 第三、四版正在制定中。MPEG-4 的显著特点是采用了面向对象、基于内容的压缩编码, 引入了视听对象 AVO 的概念, 支持固定和可变速率视频编码, 使得多媒体系统的交互性、互操作性和灵活性大大增加。

MPEG 7 和 MPEG-21 是 MPEG 系列标准中的新成员, MPEG-7 的正式名称为“多媒体内容描述接口”, 用于描述多媒体内容数据。此前的 MPEG 标准着眼于对视听内容的编码, MPEG-7 表示的则是与内容有关的信息, 即是对多媒体内容本身的描述, 其目标是提供一组丰富的标准化工具来描述多媒体内容。MPEG 7 能够对多媒体资料进行快速并有效的检索、识别、处理和过滤, 使用户能够直接使用非文本的多媒体内容并与之进行交互。例如, 通过在键盘上演奏音符来查询音乐、用图像的颜色和纹理来获得具有相似特征的图像、用“尖叫的刹车声”查找汽车追逐的场景等等。不仅如此, MPEG-7 还能广泛地用于其他与多媒体信息内容管理相关的领域。MPEG-21 即“多媒体框架”, 这一概念是在 1999 年 10 月的 MPEG 会议上提出的, 它是一种用于协调处于不同地点、与不同类型的网络相联接的各种设备间进行多媒体信息交互的综合性标准。它是基于框架的开放标准, 能够为不同的网络用户提供透明的和不断增强的多媒体资源。

有关音频信号的压缩编码技术基本上与图像压缩编码技术相同, 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。相比较而言, 其数据压缩难度较低。在多媒体技术中涉及的声音压缩编码的国际标准主要有 G. 711、G. 721、G. 722、G. 728、G. 729、G. 723. 1 等, 在这里不再赘述, 详细内容请参阅第 2 章。

## 2. 多媒体数据的获取与显示技术

多媒体数据的获取与显示主要体现在用户和多媒体系统之间进行交互的多媒体终端上。由于多媒体业务存在多种形式, 而且不同的业务要求使用不同的多媒体终端, 因此多媒体终端对多媒体数据采用的获取和显示技术也不相同。通常使用的多媒体终端主要有多媒体计算机和针对某种应用的专用终端设备, 如各种机顶盒、可视电话等。下面以多媒体计算机为例进行

说明。

多媒体数据的获取与显示主要指人机交互的界面,它介于用户和计算机系统之间。用户通过人机界面向系统提供命令、数据等输入信息,系统则将输出信息通过人机界面呈现给用户。多媒体技术出现之前的人机界面普遍采用字符界面,输入工具是键盘,而现在的多媒体计算机采用鼠标、跟踪球、电子笔、触摸屏、摄像头和视音频采集卡等作为输入设备,呈现给用户的是图形化的界面;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。

在多媒体信息输入技术中,为克服键盘输入文字的不便而发展起来的技术有联机手写识别技术、脱机手写文稿与脱机印刷文稿识别技术等,这些技术的核心是文字识别技术,其中汉字识别是文字识别领域最为困难的问题,它涉及模式识别、图像处理、数字信号处理、自然语言理解、人工智能、模糊数学、信息论、计算机、中文信息处理等多个学科。从识别的角度来看,手写识别难于印刷体识别,而脱机手写识别又难于联机手写识别。对汉字而言,联机手写识别是一种用户利用手写板实时地把汉字输入计算机的方法,计算机通过手写板将笔尖的移动轨迹用坐标点序列表示,把汉字转化为--维的线条(笔画)串来处理。目前我国的联机手写识别技术已趋于成熟,对于一般书写的正确识别率可达90%以上,某些产品的识别率甚至已经接近理论识别率的极限。手写输入产品除了在电脑上直接应用外,同时还向PDA(Personal Digital Assistant,个人数字助理)、手机、机顶盒等信息产品领域发展。汉字识别领域中最后一个十分困难的问题是脱机手写汉字识别,它处理的是二维汉字点阵图像,目前仍然处于实验室研究阶段。

语音识别技术也是多媒体信息输入技术中的一种重要技术,其目的是能够让机器听懂人类的语言。语音识别技术与文字识别技术一样,都涉及到多门学科。语音识别的难点在于语音信号的多变性、动态性、瞬时性和连续性。真正意义上的语音识别研究始于上个世纪50年代,它是随着计算机的出现发展起来的。而直到1996年,语音识别技术才首次得到商用。我国在上个世纪80年代中期之后,由于国家“863计划”的积极推动和国内各科研机构的艰苦努力,语音识别技术经历了从特定人到非特定人、从小词量到大词汇量、从孤立字到连续语音的发展历程。今天,语音识别技术已经取得了巨大进展,加之语音识别技术与语音合成技术的结合使人们能够甩掉键盘,直接通过语音来操作手机,语音技术的应用已经成为一个具有竞争性的新兴高技术产业。

总之,随着技术的进步,多媒体数据的获取和显示技术越来越多样化,例如图形识别、图像理解等先进技术的运用,使人们能够越来越方便地使用多媒体终端。

### 3. 多媒体数据库及其检索技术

随着多媒体数据在Internet、CAD(Computer Aided Design,计算机辅助设计)系统和各种企事业信息系统中被越来越多地使用,用户不仅要存取常规的数字、文本数据,还包括声音、图形、图像等多媒体数据。传统的常规关系型数据库管理系统可以管理多媒体数据。但从20世纪70年代开始,人们将目光集中在基于图像内容的查询上,即通过人工输入图像的各种属性建立图像的元数据库来支持查询,由此开展图像数据库的研究。但是随着多媒体技术的发展,

由于图像和其他多媒体数据越来越多,对数据库容量要求也越来越大,此时以传统的数据库管理系统管理多媒体数据的方法逐渐暴露出了它的局限性,集中表现在以下三个方面:第一,多媒体数据所包含的信息量非常大,用人工注释难以准确描述;第二,多媒体数据随时变化,因而难以统计及预测;第三,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,传统的数据库系统未曾涉及到这些方面。由此开发出新的多媒体数据库系统,同时,基于内容的多媒体信息检索研究方案也应运而生。

基于内容的方法是从新的角度来管理多媒体信息,多媒体数据的内容包括概念级内容、感知特性(例如图像的颜色、纹理,声音的音色、音质等)、逻辑关系(例如音视频对象的时空关系等)、信号特征(通过信号处理方法,例如小波变换获得的媒体特征)、特定领域的特征(与应用相关,例如人的面部特征、指纹特征)等。20世纪90年代初,国际上就开始了对于内容的多媒体信息检索方面的研究,从基本的颜色、纹理的检索,到综合利用多媒体数据的多种特征进行检索。目前基于内容的图像检索系统有:IBM Almaden研究中心开发的第一个用于商用的检索系统——QBIC(Query By Image Content),它提供了对静止图像及视频信息基于内容的检索手段;麻省理工学院媒体实验室研发的 Photobook 系统,它能够在存储图像时按人脸、形状或纹理特性自动分类,提供基于内容的检索;卡内基·梅隆大学的 Informedia 数字视频图书馆系统也属于基于内容的视频检索系统。在基于内容的音频信息检索方面,美国的 Muscle fish 公司推出了较为完整的原型系统,对音频的检索和分类有较高的准确率。同时,MPEG-7 标准(即多媒体内容描述接口)正在制定中,其目标是建立一种描述多媒体信息的标准,并将该描述与所描述的内容联系起来,实现快速有效的检索。

目前,基于内容的多媒体检索在国内外尚处于研究、探索阶段,诸如算法处理速度慢、漏检误检率高、检索效果无评价标准等都是未来需要研究的问题。毫无疑问,随着多媒体内容的增多和存储技术的提高,对基于内容的多媒体检索的需求将更加迫切。

#### 4. 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、远程教育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、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。通常,网络应提供的数据传输速率要在 100 Mbit/s(理论上最多 50 个 MPEG-1 视频流)以上才能充分满足各类多媒体通信应用的需要。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 100~500 ms 之间,对静止图像的延时应少于 1 s,对共享数据进行传输时应没有误码。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反映媒体之间的这种约束关系。

现有的几种通信网络(电信网络、计算机通信网络和有线电视网络)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;电信网络的技术复杂性高,开销巨大。为了适应多媒体网络

的发展需要,三网合一将成为必然趋势。多媒体通信网络应充分利用现有的网络基础设施,不应再专为某种特定应用而构建新的网络,避免为增加新的应用又引入不同的网络。

目前的多媒体网络应用主要涉及到网络传输、服务质量、服务模式等方面的技术。因为传统的基于分组交换的IP网络只提供尽力而为型的服务,要在IP网上得到QoS的保证,需要额外的协议或标准,目前有综合服务模型(IntServ)、RSVP、区分服务模型(DiffServ)、多协议标签交换(MPLS)等。这正是当前网络研究的热点,技术还未完全成熟。而IETF新公布的IPv6协议则在一定程度上满足了QoS的服务质量要求。对于电路交换网,则通过已有的ITU-T H.32X和T.120等相关的系列标准来获得有保证的服务质量。而ATM网从协议本身就考虑了对多媒体应用的支持,可以为多媒体应用预留资源。

多媒体应用还需要组播服务,其发展是多媒体应用的趋势。目前,应用在传统IP分组网上的组播技术正在得到进一步的研究和推广,相关的协议有因特网组管理协议(IGMP;Internet Group Management Protocol),组播路由协议DVMRP、MOSPF、PIM和CBT等。ATM网的组播技术也是研究的一个重要内容,其信令控制已经可以采用组播技术,但是数据分发的组播技术还没有成熟。

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网络所涉及的内容十分广泛,几乎涵盖了所有新一代的网络技术,形成了基于统一协议的由业务驱动的分组网络。它采用开放式体系结构来实现分布式的通信和管理。电信网络向NGN过渡将成为必然趋势,这是众多标准化组织研究的重点,也是各运营商和设备厂商讨论的热点。

## 5. 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、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(DVI)、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,可以被视为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—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IDE(Integrated Drive Electronics,电子集成驱动器)和SCSI(Small Computer System Interface,小型计算机系统接口)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9~66 Mbit/s的传输速率,SCSI接口的硬盘传输速率已达160 Mbit/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(RAID; Redundant Array of

Inexpensive Disks), 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(SAN: Storage Area Network)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,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。

### 6. 多媒体数据的分布式处理技术

随着多媒体应用在 Internet 上的广泛开展,其应用环境由原来的单机系统转变为地理上和功能上分散的系统,需要由网络将它们互联起来共同完成对数据的一系列处理过程,从而构成了分布式多媒体系统。分布式多媒体系统涉及了计算机领域和通信领域的多种技术,包括数据压缩技术、通信网络技术、多媒体同步技术等,并需考虑如何实现分布式多媒体系统的 QoS 保证,在分布式环境下的操作系统如何处理多媒体数据,媒体服务器如何存储、捕获并发布多媒体信息等问题,与这些问题相关的技术复杂而多样,目前仍存在大量技术问题亟待解决。

适用于分布式多媒体系统的业务多种多样,不同业务所用的多媒体终端也各不相同。目前常用的多媒体终端有多媒体计算机终端,以及针对某种特定应用的专用设备,如机顶盒、可视电话终端设备等。对于像机顶盒、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,并且为了保证多媒体数据处理的实时性,磁盘采用快速调度算法——Scan-EDF 算法。除此之外,硬件系统通常还包含光盘驱动器、音频卡、视频卡、扫描卡等各种媒体输入/输出设备,其发展趋势是逐步将这些功能卡和接口卡集成到芯片或主板上。目前,Intel、AMD、IBM 等公司生产的一系列 CPU 都具备多媒体信息的处理能力。

终端的软件平台则包括系统软件以及各种应用软件,其中操作系统是软件的核心。多媒体操作系统有 Apple 公司的 QuickTime、微软公司的 Windows XP 等。微软公司于 2004 年推出的 Windows XP Media Center Edition 2004 操作系统是专为多媒体个人电脑设计的,旨在让更多的电脑走出办公室,进入家庭的客厅。装有该操作系统的电脑可以通过遥控器来控制计算机的多媒体功能,包括打印和编辑数码相片、听广播,还可以从内容服务供应商那里下载数码音乐和电影。

流媒体技术也是一种分布式多媒体技术,它主要解决了在多媒体数据流传输过程中所占带宽宽,用户下载数据等待时间长的问题。为了提高流媒体系统的效率,提出了流媒体的调度技术、流媒体的拥塞控制技术、代理服务器及缓存技术等。在互联网迅速发展的时代,流媒体技术也日新月异,它的发展必然将给人们的生活带来深远影响。

## 1.3 多媒体通信的应用

多媒体通信系统的应用非常广泛,可以提供 VOD 视频点播、远程教学、远程办公、远程医疗、多媒体电子邮件、可视电话、桌面视频会议、数字图书馆、电子百科全书等多种多样的业务,下面简要介绍几种典型应用。

### 1. 可视电话系统

可视电话系统是较早提出的一种多媒体通信系统,其目的是使电话网能够传送视频信号,使用户在打电话的同时能够看到对方。可视电话与传统电话机相比,除了具有语音处理部分以外,还应包括图像的输入输出部分以及对图像信号的处理部分。可以通过在普通电话机上加装屏幕,利用专用芯片和专用电路来提供可视化功能,也可以通过个人电脑,加以相应的硬件来完成可视电话的功能,或者通过电视机加装机顶盒来提供图像的输入输出及处理功能。

据专家推测,可视电话将在近几年内普及,它不仅能给家庭生活带来方便和乐趣,还是其他多媒体通信应用(例如远程医疗会诊、远程教学、新闻采访、推销采购、远程购物、远程安全监控等)的好帮手。

### 2. 视频会议系统

视频会议又称会议电视或视讯会议,它是一种实时的、点到多点的多媒体通信系统。视频会议系统基于计算机网络使在异地的多个会场召开视频会议,从而减少出差经费开支。在召开会议时,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。

Internet 的迅猛发展使 IP 网络几乎遍及世界的每一个角落。IP 视频会议是目前研究的热点,成为视频会议发展的主流。为了保证音视频数据在 Internet 上的实时传输,下一代 Internet 采用了若干协议,例如 IPv6、RTP/RTCP、RSVP 等。其面向的人群逐渐向个人化方向延伸,最终将发展到家庭,功能也不仅限于单纯的会议功能,而是向远程教学系统、远程监控系统等方面发展。

### 3. 多媒体电子邮件

多媒体电子邮件不同于目前使用的 E-mail。E-mail 只有文字,而多媒体电子邮件除了包含文字之外,还包含其他媒体,例如一段音频或视频。

多媒体电子邮件是一种非实时的存储转发系统,对传输信道要求不高,发送时可以采用低速率,等待信道空闲时传送。

### 4. 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。

### 5. 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。

### 6. 虚拟现实

虚拟现实也称虚拟环境,它是由计算机模拟的三维环境,使介入其中的人产生身临其境的感觉,给人以各种感观刺激,如视觉、听觉、触觉等。虚拟现实技术通过将计算机加上先进的外围设备来模拟生活中的一切,包括过去发生的事、正在发生的事或将要发生的事。它是一种全新的人机交互系统,可以应用于驾车模拟训练、军事演习、航天仿真、教育、娱乐等多种领域,具有巨大发展潜力。

自1993年美国宣布“信息高速公路”计划后,全球范围内掀起了开展信息高速公路建设的热潮,到目前为止,其发展势头更加迅猛。我国的三大网络——电信网、有线电视网和计算机网也遵循这一规律,开发出各种先进技术,最终实现三网合一,而在实现融合的过程中又会出现许多高新技术,从而加速网络的进一步发展。目前许多专家认为,三网合一后的网络所使用的协议是IP协议,而基于ATM的多媒体通信网也备受关注。此外,宽带交换技术也在不断地向前发展。最近,国内外有关研究部门正在研究一种新的交换方式——DTM动态同步传输模式,该方式能够更好地适应未来宽带实时通信传输的需要。总之,宽带化、IP化、无线化是多媒体通信的发展方向。

科学技术的任何进步都可能极大地促进多媒体通信的发展。可见,未来的网络将一定会沿着满足信息社会、知识经济发展的需求这一方向不断前进,给人们的生活、工作和娱乐带来深刻的影响。

## 小 结

本章概括介绍了多媒体通信技术的相关概念及涉及的若干关键技术,主要内容有:

(1) 媒体。媒体是指信息传递和存储的最基本的技术和手段,即信息的载体。媒体可划分为五大类:感觉媒体、表示媒体、显示媒体、存储媒体和传输媒体。

(2) 多媒体。由于多媒体技术是一种融合技术,覆盖面宽,技术复杂,很难一言以蔽之。目前公认比较准确的概念是: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。

(3) 多媒体通信系统。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