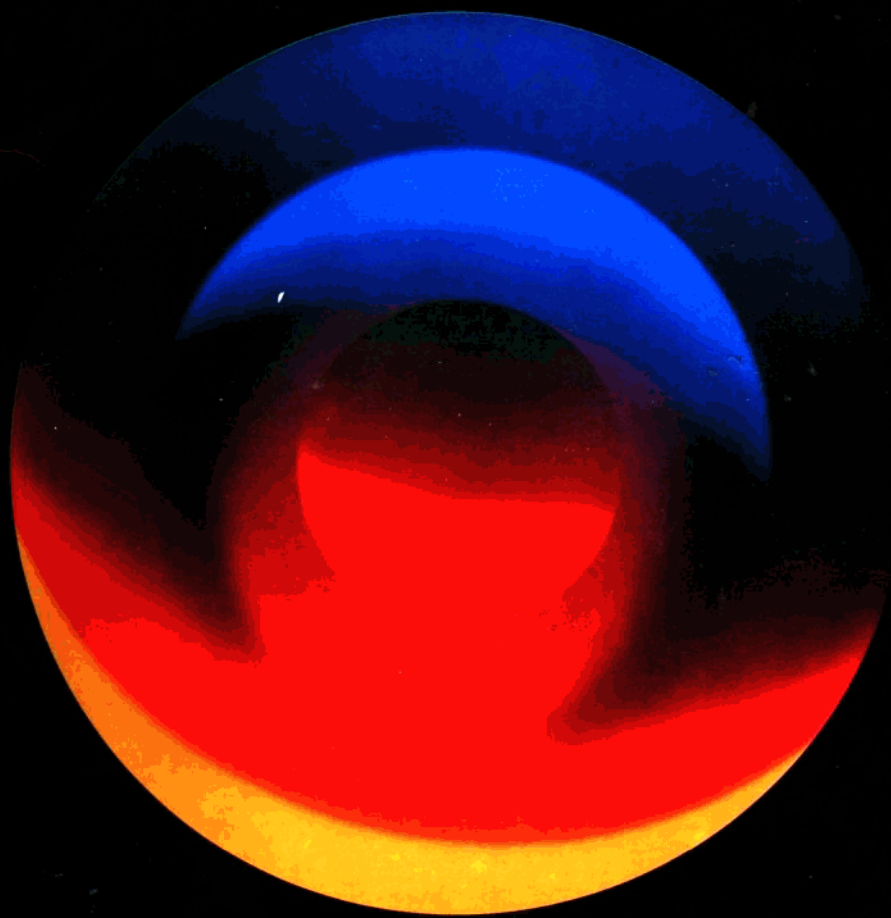


# 多媒体通信系统



电信新技术培训系列教材

DIANXIN XINJISHU PEIXUN  
XILIE JIAOCAI

人民邮电出版社

● 王行刚  
编著

电信新技术培训系列教材

# 多媒体通信系统

王行刚 编著

人民邮电出版社

## 内 容 提 要

本书是为在职人员编写的新技术培训教材。

本书概要介绍多媒体通信系统,全书共五章。第一章为系统概论,主要从基本概念、基本技术、图像压缩、音频编码、网络支持几方面论述;第二章介绍了多媒体通信的终端设备;第三章简述了通信网,主要从通信网的发展、高速局域网、帧中继、ATM 等方面介绍;第四章为应用系统,介绍了基于 Internet 的多媒体通信系统、多媒体与超媒体信息检索系统、会议电视系统、交互式电视系统及多媒体可视电话系统。第五章为国际标准,重点介绍 JPEG 标准、H. 261 建议、H. 26X/MPEG-2 标准及多媒体和超媒体对象的表示和编码标准。

本书避免繁琐的数学推理,条理清楚,文字简练,通俗易懂。

本书可供从事通信和电子技术的人员参考。

电信新技术培训系列教材

### 多媒体通信系统

王行刚 编著

责任编辑 马月梅

\*

人民邮电出版社出版发行

北京崇文区夕照寺街14号

北京成金印刷厂印刷

新华书店总店科技发行所经销

\*

开本:787×1092 1/16 1996年11月 第一版

印张:7.5 1997年3月 第二次印刷

字数:192千字 印数:10001-18000册

ISBN7-115-06369-9/TN·1154

定价:10.00元

## 前 言

当前,电信新业务、新技术迅速发展,广大干部和职工急需提高业务、技术和管理水平,以适应通信大发展的需要。1992年11月以来,已由人民邮电出版社陆续出版了《移动通信》、《电信网》、《程控交换》、《数字通信》、《光纤通信》、《数字微波》等6种“电信新技术培训系列教材”。

这套书出版后,我局曾组织了三期电信处长、电信局长、总工程师等同志参加的学习班,收到了较好的效果。广大学员反映这套书具有简明、实用和便于自学等特点,但品种还不够全,还不能满足需要,特别是新业务、新技术的短期培训教材尚不配套,有必要进一步增新补缺。为此,我局根据广大电信职工和管理干部的要求,结合企业实际工作的需要,又组织编写了《分组交换》、《电信新业务》、《卫星通信》、《图像通信》等一批教材,并将陆续出版。

由于时间仓促,经验不足,书中难免有缺点和不足之处,希望各地在使用过程中,及时把意见反馈给我们,以便今后修订。

邮电部电信总局

1993年6月

# 目 录

<b>第一章 系统概论</b> .....	(1)
<b>第一节 基本概念</b> .....	(1)
一、什么是媒体 .....	(1)
二、什么是多媒体 .....	(1)
三、多媒体通信 .....	(3)
<b>第二节 基本技术</b> .....	(3)
一、多媒体通信应用 .....	(4)
二、网络 .....	(4)
三、多媒体通信终端 .....	(6)
<b>第三节 图像压缩</b> .....	(7)
一、基本方法 .....	(7)
二、处理步骤.....	(11)
<b>第四节 音频编码</b> .....	(13)
一、数字音频.....	(13)
二、基本方法.....	(14)
三、编码标准.....	(16)
<b>第五节 网络支持</b> .....	(17)
一、宽带网络.....	(17)
二、复用技术.....	(20)
三、网络性能.....	(22)
四、终端入网.....	(24)
<b>第二章 终端设备</b> .....	(26)
<b>第一节 计算机的发展</b> .....	(26)
一、计算机处理的“媒体”.....	(26)
二、计算机的通信能力.....	(26)
三、计算机的求解能力.....	(27)
<b>第二节 多媒体计算机的发展</b> .....	(28)
一、模式识别.....	(28)
二、CD-ROM .....	(29)
三、多媒体芯片.....	(31)
<b>第三节 多媒体通信终端的发展</b> .....	(32)
一、从简单终端到复合终端.....	(32)
二、基于计算机的终端.....	(33)
三、基于多媒体计算机的终端.....	(34)
<b>第四节 多媒体个人计算机</b> .....	(35)

一、MPC 标准 .....	(35)
二、多媒体升级套件 .....	(37)
三、多媒体 PC 产品技术 .....	(38)
<b>第五节 多媒体通信终端平台</b> .....	(41)
一、硬件平台 .....	(41)
二、软件平台 .....	(43)
三、信息服务 .....	(44)
<b>第三章 通信网</b> .....	(46)
<b>第一节 现代通信网的发展</b> .....	(46)
一、综合化与宽带化 .....	(46)
二、智能化 .....	(47)
三、个人化 .....	(48)
四、多媒体化 .....	(49)
<b>第二节 高速局域网</b> .....	(49)
一、FDDI/CDDI .....	(49)
二、100Mb/s 以太网 .....	(51)
三、交换以太网 .....	(52)
四、ATM-LAN .....	(53)
<b>第三节 帧中继技术</b> .....	(55)
一、FR 技术的兴起 .....	(55)
二、基本原理 .....	(56)
三、技术特点 .....	(57)
四、国际标准 .....	(57)
五、FR 的发展 .....	(60)
<b>第四节 ATM 网</b> .....	(60)
一、基本原理 .....	(60)
二、ATM 特点 .....	(62)
三、网络配置 .....	(63)
四、接入控制 .....	(65)
<b>第四章 应用系统</b> .....	(66)
<b>第一节 基于 Internet 的多媒体通信系统</b> .....	(67)
一、Internet 的发展 .....	(67)
二、多媒体电子邮件系统 .....	(68)
三、多媒体信息发现系统 .....	(70)
四、视频和音频通信系统 .....	(71)
<b>第二节 多媒体与超媒体信息检索服务系统</b> .....	(72)
一、系统组成 .....	(73)
二、信息类型 .....	(74)
三、信息交换 .....	(75)

四、支撑系统.....	(76)
<b>第三节 会议电视系统</b> .....	(77)
一、会议电视业务的网络支持.....	(78)
二、会议电视设备.....	(79)
三、会议电视网举例.....	(80)
<b>第四节 交互式电视系统</b> .....	(81)
一、交互式电视业务类型.....	(81)
二、技术途径.....	(83)
三、特殊设备.....	(87)
<b>第五节 多媒体可视电话系统</b> .....	(87)
一、可视电话的发展.....	(87)
二、标准.....	(89)
三、终端.....	(91)
四、应用.....	(92)
<b>第五章 国际标准</b> .....	(93)
<b>第一节 JPEG 标准</b> .....	(93)
一、概述.....	(93)
二、算法.....	(94)
<b>第二节 H.261 建议</b> .....	(95)
一、概述.....	(95)
二、图像格式和图像数据结构.....	(95)
三、视频编解码器.....	(98)
四、DPCM 编码 .....	(100)
<b>第三节 H.26X/MPEG-2 标准</b> .....	(101)
一、概述 .....	(101)
二、视频比特流的句法和语义 .....	(102)
三、编码算法 .....	(105)
四、轮廓和水平 .....	(105)
五、应用 .....	(106)
<b>第四节 多媒体和超媒体对象的表示和编码标准</b> .....	(106)

# 第一章 系统概论

## 第一节 基本概念

### 一、什么是媒体

媒介(medium),指一个居间的手段,尤指思想表达或交流的手段。在信息技术领域,媒体(medium)指信息的载体。由于信息被人们感觉、加以表示、使之显现、实现存储或传输的载体各有不同,因此媒体可区分为:

(1)感觉媒体(perception medium),指人们所感觉的信息的自然种类,如声音、图形、图像和文本等。

(2)表示媒体(representation medium),指被交换的数据的类型,它们定义了信息的特性,如用它的编码形式加以描述。相应地有语音编码、图像编码和文本编码等表示媒体。

(3)显现媒体(presentation medium),指为人们再现信息的物理手段的类型(输出设备),或者指获取信息的物理手段的类型(输入设备)。如显示器、扬声器、打印机等输出类显现媒体,以及键盘、鼠标器、扫描器等输入类显现媒体。

(4)存储媒体(storage medium),指存储数据的物理手段的类型,如磁盘、光盘、磁带等存储媒体。

(5)传输媒体(transmission medium),指传输数据的物理手段的类型,如同轴电缆、光缆、双绞线、无线电链路等传输媒体。

(6)交换媒体(interchange medium),指在系统之间交换数据的手段类型,它们可以是一存储媒体,或者是一传输媒体,或者是两者的某种组合。

### 二、什么是多媒体

“多媒体(multimedia)”作为一个形容词来理解时,它必须附加到一个名词上,才能确定其准确的含义。“多媒体”通常附加到技术、业务、计算机、通信、应用等名词上,形成一系列新的专用词汇。它们引发出以下一些新的概念。

#### 1. 群体性

从信息处理的角度看,同一台计算机早就可以分别处理数据、文字、图形以及声音等多种感觉媒体,为什么过去不把计算机称为多媒体计算机呢?

从信息传输与交换的角度看,同一个公用电话交换网(PSTN)早就可以分别提供电话、电报和数据(利用 modem)业务,为什么过去不把 PSTN 称为多媒体通信网呢?

第一个重要的差异在于,多媒体概念强调多种媒体间存在着相互联系这一“群体”性,即无论是传输、交换或处理,其对象是具有“群体”特征的信息;而不是一个个“分立”的信息类



型,尽管它们在时间、空间上可能并存在同一系统之中。

多媒体间相互联系的“群体”性,主要体现在多种媒体间存在着“同步性”,即在时间上多种媒体间相互有紧密联系(如画面与声音同步),在空间上多个媒体间也有紧密联系(如图像画面与文字注释在空间分布上具有协调性)。

从这种群体性来鉴别,一般的计算机虽然按“多道程序”运作时,不同“道”的程序可能分别在处理数据、文字或声音,但通常这些数据、文字和声音不是作为一个有机的群体对象来处理的。多媒体计算机的处理对象则强调将不止一类的感觉媒体作为一个整体来加以表示、存储、处理和显现。最典型的“群体”处理对象是以活动图像为基体的,例如最常见的是活动图像配以声音的群体——视像/视频(video)信息。实际上,计算机从处理数据、文字到直接处理视像,是计算机堪称多媒体计算机的最显著标志。自然,多媒体计算机更一般的含义是指其加工对象是群体信息,而信息群体中只要大于一类信息,并不一定要包含活动图像/视频信息。

总之,“群体”是区别于“个体”的概念,多个互不无联系的“个体”的集合并非“群体”。“多媒体”引入的是一个“群体”概念,内含所涉及的多个媒体之间是具有同步性和协调性的整体。

## 2. 交互性

如果仅从具有群体性特征来判断,广播式电视系统可谓是“多媒体”通信系统了,因为影像和声音是作为一个整体加以采集、传输和显现的。但是,多媒体概念不仅仅涉及群体性一个特征,而是具有一“群”特征,这些特征之间也是不可分割的,否则就不是一个完整的多媒体概念。多媒体概念的第二个重要特征是“交互性”。

计算机技术的一大特点是它可以提供用户与系统之间的交互控制的能力。这种交互式控制的能力,通常是信息技术领域中一些新系统的重要特征。交互式控制能力可以从控制的对象、控制的程度和交互控制的速度等性能来刻画。

从“交互性”的角度看,用户对广播式电视机的控制程度很低,只能通过切换频道来选择节目,并不能通过控制“信源”(电视台)来选择所需要的信息,即“信宿”与“信源”之间不存在交互作用。所以从“群体性”和“交互性”两种属性并存的角度看,广播式电视系统尚不属于多媒体通信系统的范畴。

信息检索系统、实时业务处理系统和过程控制系统等计算机应用系统,通常具有高度的交互控制能力。从要具备“交互性”特征的角度来看,多媒体技术是离不开计算机技术的。

## 3. 多值性

多媒体概念引入的另一新概念是——“多值性”(multivalued)。在一个信息技术领域的“系统”中,这种多值属性反映在,在一次信息采集、传输、处理和显现过程中,应涉及到多种表示媒体、传输媒体、存储媒体或显现媒体,或者多个信源和信宿的交互作用。这种多值性将创造出系列新业务、新系统、新技术和新应用。

例如,在“多媒体”通信业务的多值性特征中包含一次单一的通信会话时可以使用多条连接,可以增/减通信资源和用户等。

综合以上群体性、交互性、多值性的多媒体内涵,将其与技术、业务、通信、计算机等本意

结合,就可界定一系列新词汇(多媒体通信、多媒体计算机、多媒体技术)的意义。通常,这都会引发出不同于传统意义的新范畴。它们是通信技术和计算机技术的新领域,特别是两者结合(computer & communication, C&C)的新领域,包括一系列新的技术、产品和应用,以及相关的新产业和新市场。实际上,多媒体技术、产品和市场深刻地反映出各项信息技术的大融通,而多媒体技术的广泛应用将推动世界进入一个新的多媒体世界。

### 三、多媒体通信

多媒体通信是一种新的通信业务,它既不同于基于点对点式通信的电话业务或数据业务等传统电信业务,也不同于基于广播式通信的无线电视业务和有线电视(CATV)业务。多媒体业务是基于群体通信(group communication)的尚待在发展中逐步界定其内涵的新业务。

多媒体业务(multimedia service)的特征是,在一次单一的通信会话(communication session)期间,可以有多个参与方,多条连接,而且通信资源和用户数可以增减。

多媒体业务类型可分交互型业务和分配型业务两大类。

#### 1. 交互型业务

交互型业务,包括会话型(如会议电视,文本传送,可视电话等)、电子信函型(如声、图、文电子信箱,文本传递等)和检索型(如可视图文,文件检索,视频检索等)业务。

(1)会话型业务用于在两点之间或点与多点之间同时传递话音、活动画面和视频扫描静止图像与文件,也包括高速数据的传输。该业务是即时发生的,双方或几方预定的,也可以是永久型的。用户信息的流量可以是双向对称的或不对称的。

(2)电子信函型业务包括传递活动图像和伴音的电子信箱业务,以及传递混合文件的电子信箱业务。它具有存储—转发和消息处理的功能。该业务可以是点对点或点对多点进行的,可以是双向对称或单向的。

(3)检索型业务包括宽带可视图文、高分辨率图像检索、文件/数据检索等,应用于远程教学、远程诊断、远程购物及娱乐等。该业务是即时进行的,可以是点对点或点对多点方式。

#### 2. 分配型业务

按用户能否进行单独演示控制分为两种:

(1)用户不能控制的分配型业务,包括常规电视分配型业务、文件分配型业务、高速不受限数字信息的分配型业务,可用于电视节目的分配和电子报纸等。该业务是广播型的,用户不能控制广播信息的起始时间和顺序。

(2)用户能够进行单独演示控制的分配型业务,即点播型业务,包括全频道广播视频通信,可用于远程教学、新闻检索和节目点播等。该业务是点播型的,用户可以控制节目的起停和顺序等。

## 第二节 基本技术

多媒体通信技术,可分为三部分:多媒体通信终端技术;支持多媒体业务的通信网络技

术;多媒体通信应用系统技术。驱动多媒体通信技术发展的关键因素是应用,实际上是由应用系统的需要推动多媒体通信终端技术的发展以及通信网络作相应地适应性变革和创新。因此,首先从当前多媒体通信应用最富吸引力和潜力的方面来分析。

## 一、多媒体通信应用

在两大类多媒体业务——交互型业务和分配型业务中,最富“多媒体”特征的是新型的交互业务(interactive service)。

这些新型交互型业务有些什么特点呢?

(1)业务的多样性。例如,在商贸业务方面有商贸电视业务(transaction TV service)、电子订票、电子图书、电子报刊、家庭银行(home banking)、家庭购物(home shopping)等。在公众信息服务方面,提供交通信息、气象信息、旅游信息、经贸信息、社区信息以及政府信息等。在教育、医疗方面有远程教学(distance learning)、远程诊断、远程软件等。在娱乐方面有点播电视(video-on-demand, VOD)或交互电视业务(ITV service)、电视游戏(TV game)等。

(2)一对多的服务方式。新型交互型业务除了可实现一对一(点对点)交互外,还可实现一对多(一对多点)的业务,即一个业务提供者可以同时为多个顾客(消费者)服务,提供相同或不同的业务。

(3)以视频交互业务为基础。此乃这些新型交互型业务的最大特点。通过对视频、音频信号的数字编码和压缩与解压缩,可以比模拟技术实现更多的业务、更高的质量。

(4)跨网络的可操作性。交互型业务系统由业务提供者(provider)、网络经营者(network operator)和用户(subscriber 或 consumer)组成。业务提供者制作和供应业务,必须具有巨大容量的业务资料库,具有快速响应用户请求并实时分配所需服务的能力。

每个载运业务的网络必须有足够的带宽,方便信息存取、抗扰和安全能力。而不同经营者的网络之间还必须是可互操作(inter-operability)的,以便实现跨网络的交互业务传送。

目前,新型交互型业务在很多国家和地区正在形成研究、开发和试验的热潮。将来,新型交互型业务能否被广大消费者接受主要取决于以下一些因素。

(1)费用。消费者使用新型交互型业务的费用包括须增加的设备(如多媒体终端)的费用和日常使用的费用(月付费)。

(2)业务内容适合要求。据估计,新型交互型业务使用的优先顺序是:交互电视,交互教育节目,日常新闻频道,家庭银行,体育节目等。

(3)使用操作方便。广大消费者期望交互型终端设备除具备必要的功能、可接受的价格外,还必须使用方便、可靠,越简单越好。

以上这些业务和应用需求,是通过一系列的有关技术及其产品和产业来实现的。这些技术涉及相当广泛的领域,如数字技术、计算机技术、通信技术、微电子技术、显示技术等。

## 二、网络

由于多媒体业务、多媒体通信尚处兴起阶段,目前还难于对这些术语给出一个完整的科学的定义式的描述。在本章第一节对多媒体业务的描述,只是多种见解之一。此外,还有一些有代表性的描述,例如:

- 多媒体业务是至少包括两种不同信息类型的业务;

·多媒体通信是指以机器可理解形式存在的诸如文字、语音、图形、静止图像、声音和电视等多媒体信息的表示、存储、查询和传播；

·将多媒体通信理解为一整套通信技术，能够把巨量的存储与访问文本、静止与活动图形、照片、电视以及声音联系在一起。

从整个通信技术的发展来看，当前的通信体制主要是建立在对通信的低层目标支持上，其主要目的是在远端真实地再现输入信号。而信号本身可以是语音波形、组合电视信号、传真信号或计算机语言符号等。它们的性质差别很大，因此要用不同的业务终端来接受和再现，用不同的方式来传递，用不同的网络来支持由不同信号所组成的业务。

人类通信的本来目的是交流信息。因此，通信的高层目标应当是支持人类对各种信息的自然交流。这要把把所有信息转化为可以进行统一操作的对象，即数字媒体。理论上早已证明了数字化的可能性，而计算机技术和通信技术等的的发展已使得在数字形式下对信息进行操作变得经济、可靠和有效。

可见，一方面通信向高层发展，趋向多样化和复杂化；另一方面，对高层的支持又趋向于通信低层的单一化和简单化。数字化是走向单一化和简单化的第一个层面，在此之上要进而解决通信网络功能的单一化和简单化。

发展综合业务数字网(ISDN)时首先建立了使用户—网络接口(user-network interface, UNI)单一化和简单化的概念，试图用统一UNI来规范用户对网络的接入。但要支持各种信息类型在网络中传递、相互之间的转换及其控制，还必须对网络的传递模式进行统一。

发展宽带综合业务数字网(B-ISDN)时，进一步试图解决传递模式统一问题，相应地提出了“异步转移模式”(asynchronous transfer mode, ATM)。ATM采用固定分组长度的信元(cell)来承载信息，在网上进行复用、分用、交换和传输，并且仅需要最基本的网络处理功能。B-ISDN网络功能的多元化体现在ATM适配层，目前CCITT只定义了5种不同的适配层类型。从支持业务多样化的趋向看，ATM适配层仍比较简单，还不够多样化。

未来的电信网不再只是支持几种特定业务，而应当全面支持包括交互型和分配型业务的各种可能的业务。但是，对未来可能的各种业务、特别是多媒体业务进行规范会遇到本质上的困难，因为多媒体通信所要支持的人类自然通信本身是非常复杂、多样的。因此，更为实际可行的出路是，转而对提供业务的软硬件平台进行规范。

多媒体通信的实现需要在不同层面上进行通信组织工作，而这些通信组织工作是和交互性连在一起的。可以说，交互性存在于通信组织的各个层面以及通信网的各类接口中。

表1-1列出了多媒体通信的组织和交互的一些组织层面、功能举例和交互层面。

表 1-1

通信组织层面	功能举例	交互层面
显现方式的组织	显现媒体选择、屏幕划分风格等	显现界面的交互
媒体间相互关系的组织	媒体同步、控制和转换	连接级别的交互
演示材料的组织	数据、文件结构等	交流内容的交互
演示进程的组织	超媒体、交互式多媒体程序等引导控制方式	交流进程的交互
交流环境的组织	协同工作所需各种功能	通信各方的交互
通信环境的创造	虚拟现实等	通信者和通信环境的交互

在应用级,交互性反映为协同工作能力。例如,能否方便地召集群体(group)通信;所提供的环境是否满足群体工作的特定要求;是否提供了足够的辅助交流工具;能否做到用户间信息的实时共享;以及用户是否能按自己喜好选择不同的显现方式、承载媒体和演示进程等。

在业务级,交互性反映在保持媒体间的时间同步关系、多线索任务的时序安排、媒体转换点和连接控制桥的互连及兼容方案等。

总之,交互性是通信高层目标的一个重要方面。甚至可以认为,通信的最终目标就是实现“在带宽和语义约束下的极大交互性”。

### 三、多媒体通信终端

用户终端的功能有两方面:面向用户提供一种自然的通信环境,屏蔽掉各种复杂的网络功能;面向网络下达其所需实施的各种功能,屏蔽掉应用环境的复杂性和不确定性。

多媒体终端的模型,可以设想分为三级:人机界面级;业务级;物理级。总的思想是将从应用到入网的整个过程分解成几个级别,各有侧重地实现功能优化。但目前尚未形成国际标准。

CCITT 第 15 组的综合视像业务(IVS)和第 18 组关于宽带业务的建议,都强调统一用户终端平台在方便用户使用、增强互通性和降低成本方面的意义。

新一代的用户终端平台将是基于计算机的、通信技术与计算机技术紧密结合的产物。因此,多媒体终端技术将涉及现代计算机和通信的许多基本技术和最新技术,同时也有其特殊的部分。

按照以计算机技术为基础,可以把多媒体终端技术划分为以下几个方面。

#### (1) 系统体系结构与组织

研究并解决多媒体通信终端与系统的功能结构(体系结构)与物理结构(系统组织),包括研究多媒体业务的概念、分类、模型,多媒体通信的协议、标准或规范,多媒体通信终端的开放式系统结构,多媒体通信的人—机交互方式、协议和接口技术,多媒体通信终端与系统的性能评价等等。

#### (2) 硬件

多媒体通信终端的硬件,包括基本的计算机硬件系统和各种附加设备。典型的附加设备有各种网络接口板/卡,视频和音频数据采集、压缩与解压缩板/卡,各种显现媒体(如摄像机、扫描仪、高分辨率或大屏幕显示器、双声道扬声器或高音质扬声器、彩色喷墨输出设备等),各种人机交互控制手段(如键盘、鼠标器、触摸屏、光笔等)等。

#### (3) 软件

多媒体通信终端与系统需要配置的软件包括多媒体操作系统,多媒体数据库管理系统,多媒体应用开发工具与环境,各种各样的多媒体应用软件(如多媒体电子邮件软件、多媒体会议系统控制软件、点播节目管理软件、远程教学管理软件等)。

#### (4) 数据

多媒体通信系统的应用常常离不开多媒体数据库,包括集中式的数据库和/或分布式数据库。这些多媒体数据库通常具有容量大、响应快和更新快等特点。特大容量的多媒体数据库的建库工作量是巨大的,需要在建库方法与手段方面有新的技巧和工具。

在多媒体数据采集、编辑、节目或脚本创作等方面,也提出了一些新课题,需要充分利用计算机软硬件技术创作相应的工具。

#### (5)特殊部分

多媒体通信终端与系统技术中比较特殊的部分有:视频和音频数据压缩和解压缩技术,群体通信控制技术(包括多种媒体的同步控制技术),自然的人-机接口技术(包括人一机通信协议、交互式控制技术),以及各种各样促进经济和社会信息化的多媒体通信应用技术等。

## 第三节 图像压缩

图像压缩可分静止图像压缩和活动图像压缩两大类。前者侧重图像精度,而对实时性要求不高;后者侧重实时性,可根据传输信道数据码率和接收端图像质量再分档次,例如可分为:

- 电视电话(video telephone),10kb/s-64kb/s
- 会议电视(video conference),64kb/s-2Mb/s
- 常规电视(television, TV),1.5Mb/s-8Mb/s
- 高清晰度电视(high density televisivn,HDTV),20Mb/s-30Mb/s

以下介绍各种图像压缩方法的基本概念。

### 一、基本方法

#### 1. 量化(quantization)

量化是一种直观的数据压缩方法,其过程相当于将输入数据的取值范围加以限制,利用量化查找表(look up table)使一个输出值对应于若干个输入值。考虑人的视觉对不同亮度值域的敏感程度不一样的特点,在一定输出图像质量的前提下,可调节量化查找表达到最佳的压缩比。

根据量化查找表的性质,量化算法可分为两种:线性和非线性算法。根据图像数据的性质,可对不同的图像区域采取不同的量化方法,或称动态量化,这是目前一些实用编码方法中较多采用的方法。

#### 2. 统计编码(statistical coding)

统计编码也称熵编码(entropy coding),是一种无损(loss less)编码(也称信息保持型编码)。

图像信号是一种随机信号。如果每一图像像素出现的概率是非均匀的,并且假设某一像素值的概率已知,则可根据其概率的大小而赋予一定长度的码字表示。出现概率大的像素值用短码字表示,出现概率小的像素值则用长码字表示。这样,平均每像素所需码字的长度将比用等长度码字表示每个像素时短,从而达到压缩的目的。著名的霍夫曼(Huffman)编码乃是熵编码中的最佳编码方法。

### 3. 预测编码(predictiv coding)

预测编码利用了图像信号在空间域的相关性。在一帧自然图像中,由灰度缓慢变化的区域构成的平缓区常常占很大部分。由于这些平缓区中的一个像素与它周围的像素的区别不大,因此可用它周围的像素的值的近似代替。该像素在接收端恢复时,可用近似值+误差值来无失真地重建。

通常,先预测某像素的灰度值,即用其前后左右的若干像素的线性组合来表示近似预测的灰度值;然后对这个像素的真值与预测值之差进行编码,即对其误差(error)进行编码。

预测编码可以获得高质量的解码图像,但压缩比较低,主要用在帧间图像压缩方面。预测编码中最典型的方法是“差分脉冲编码调制”(differential pulse coding modulation, DPCM),它已成为数字电视的基本压缩方法。

### 4. 变换编码(transform coding)

变换编码的目的是消除图像像素之间的相关性。通常是对图像数据进行线性变换,使变换后的系数之间的相关性降低。经过线性变换后原数据的个数并没有减少,但是如果变换后的某些系数为零或接近于零,则可舍去这些系数而只对非零系数进行编码再传输,达到对原数据进行压缩的目的。在接收端则进行逆变换、恢复图像。

通常,将图像分成不重叠的方块(例如 $8 \times 8$ 方块),对每一数据块独立进行变换。分块编码的缺点是,在高压缩比时会出现灰度不均匀的“方块效应”。

变换编码方法中有著名的 KLT(Karhunen-Loeve transform)变换和离散余弦变换(discrete cosine transform, DCT)等。前者是最佳的线性变换压缩编码方案,可以作为评价其它线性变换方法的基准。后者具有快速算法,方法简单,易于实现, DCT 芯片目前已广泛使用。

### 5. 子带编码(sub-band coding, SBC)

图像子带编码是从声音的子带编码移植过来的。它将图像通过若干个带通滤波器,从而将原始图像的频谱分成几个频率段;然后对不同的频率段采用不同的后续编码方法,实现对原图像的压缩。

人类视觉系统对噪声的敏感程度在不同频率段中是不同的。这种自然生理现象可被利用来对不同频率段采用不同失真度的压缩。典型的 SBC 是采用高通、低通两个滤波器,将原图像在频率域分成两个子带;也可以将子带再一分为二,形成多层次、多频带的 SBC 过程。由于 SBC 是对整个图像进行的,不会出现方块效应。

### 6. 小波变换(warelet coding)

多分辨率的图像描述称为图像的小波分解。小波的图像分解方案属于子带分解的特例。

小波变换对图像进行压缩的原理与子带编码方法一样,是将原图像信号分解成不同的频率区域(在对原图像进行多层次分解时,总的数据量与原数据量一样,不增不减);后续的压缩编码方法则根据人的视觉特性及原图像的统计特性,对不同的频率区域采用不同的压缩编码方法,尽可能减小数据量。

小波变换的特点是,在空间域里将信号分解为分辨率不同的层次,在进行多层次分解运算的同时形成了频率域中的多层次分解。在频域中的每个层次上,高频分量与低频分量的分布与原数据中频率分布的方向有关。例如,在算法上如只考虑图像的 X 方向、Y 方向与对角方向,则在频率域中这三个方向上的频率变化将分别体现出来。小波变换方法始于 1989 年。

#### 7. 矢量量化(vector quantization, VQ)

矢量量化利用相邻图像数据之间的高度相关性,将输入的图像数据序列分组。每一组  $m$  个数据被描述成为一个有  $m$  个元素的矢量。将每一矢量与码本中的矢量(称为码字)比较,选择达到最佳匹配的码字的序号进行编码。由于对码字序号的编码长度小于对矢量的编码长度,从而实现数据压缩。

矢量量化的解码方法比较简单:接收方根据收到的序号从与编码端一致的码本中找出该序号对应的码字,实现对原数据的近似重现。可见矢量量化是一种信息有损失的压缩方法,但可以获得较高的压缩比。

为了减小失真,码本的容量自然应增大,使得矢量匹配的搜索运算时间增长。此外,由于要对原数据分块,例如要将二维图像分成若干个  $M \times N$  的矩形块(每一个矩形内的数据排列成为一个  $m = M \times N$  个元素的一维矢量),因此在高压缩比时也会出现方块效应。

#### 8. 网状编码(mesh based coding, wireframe coding)

将原图像用线状网(如三角形网、矩形网、四顶点网格)分成不同的区域,再将网格内部的图像灰度信息抽象到网格的顶点(网的节点)之上,实现对网格内部数据的压缩。解码时,在网的拓扑确定之后,每一网格内的数据值由接收到的该网格顶点的信息经过插值恢复。

若对不同图像采用固定拓扑的网,则在接收端网的恢复具有唯一性。而建立在非固定拓扑之上的网,需要首先确认网的生成具有唯一性。如果网的拓扑在接收端不能唯一地恢复,则图像重建出现失真。

#### 9. 混合型编码(hybrid coding)

这种方法将两种以上的编码方法接续起来,以求得最大的压缩效果。

例如, SBC-VQ 混合编码是用子带变换(SBC)先将图像分解成两个频带,再对不同频带施以不同失真率的矢量量化编码(VQ)。

再如, DCT-Entropy 编码是将 DCT 输出的码流施行统计编码。当 DCT 产生的数据是以非均匀概率出现时,这种混合方法可以实现进一步的数据压缩。

#### 10. 帧间编码(interframe coding)

帧间编码是针对时间上的序列图像,特别是实时的序列图像。当帧与帧之间时间上的相关性比一帧之内图像(帧内)数据空间上的相关性大得多时,帧间编码可以获得比单纯用帧内压缩方法更高的压缩比。目前,帧间编码多用预测编码得到帧间误差,然后再对该误差编码。

帧间编码的难点是快速运动物体的位移估值,而正确的位移估值可以使两帧图像的相关程度得到补偿。



上述几种编码方法均属于波形编码(wareform coding)类,均基于图像信源是平稳过程的马尔柯夫(Markov)随机场这一假设,它们可达的压缩比受图像熵值的限制,大体是一个数量级。

在信息保持型编码类,图像的熵值是每像素平均码数的下限。在信息有损型编码类,对于一定的失真程度,存在一个平均编码数的下限不可超越。因此,上述几种编码方法不可能获得压缩比上的突破,除非突破波形编码概念的限制。

### 11. 分析-合成法编码(analysis-synthesis coding,ASC)

ASC 不属于波形编码类,它是建立在对图像信息的理解之上,编码过程是对图像的分析过程,解码过程则是对图像信息的合成过程。对有先验知识的单物体(如人的头部)的编码,“模型法”有望达到高压缩比。对于无先验知识的图像,则图像分析是困难的。

基于先验知识进行图像分析的例子是,对于运动的人的头部图像,可以只传入头、五官形状、位置、运动、光线等参数,接收端即可由这些参数及有关人头的模型库重建原图像,而图像重建是计算机绘图的过程。

ASC 一般建立在网状编码的基础上,但它不仅以网格形式利用图像数据局部的相关性;并且根据对图像的理解形成模型或从模型出发理解图像,采用模型方法描述局部与整体之间的结构,从像素点(pictoric)的描述升华到符号(symbolic)的描述,从而达到比网状编码更高的压缩比。

### 12. 分型编码(fractal coding)

这种方法在于发掘自然物体(如云彩、森林)在结构上的自相似性(它是图像整体与局部相关性的表现),通过去除这种相关性来实现图像数据压缩。

图像数据的自相似性可以理解为,它的所有部分都可以用本图像中其它部分通过  $W$  运算近似求得;而  $W$  运算仅仅是比例、旋转、镜像和移位。数学上的描述即是一种迭代运算:

$$X \doteq W(X)$$

对于给定的图像信号  $X$ ,如能找到迭代运算算子  $W$ ,而  $W$  能重建  $X$ ,因此只传输  $W$  便可在接收端恢复  $X$ ,从而达到压缩的目的。

分型编码与波形编码相比,它不但去除了数据之间局部的相关性,而且去除了整体与局部之间的相关性,所以可望达到波形编码方法所不及的压缩比。不过,目前分型编码尚未达到预期的压缩比。这种方法的关键在于迭代算子  $W$  的确认及其运算时间。

### 13. 视频信号打包编码(packet video coding)

实时图像数据可以采取将数据分组打包传输。对于非定长的码字(每一个被编码的长度不一样),在单位时间内其信息量是不均匀的。因此,当信息量大时可在单位时间内多发送数据包,而当信息量小时则少发数据包,以保证图像质量的恒定。这样便可解决固定数码率信道所无法解决的问题;当图像细节增多导致输出缓冲器溢出,而不得不牺牲图像质量的缺陷。

打包编码方法需要研究解决一系列编码策略问题:如何将数据分组打包;如何确定包的优先级;如何对待传输中包的丢失问题,包括如何减小丢失包造成的失真等。