

第一章 多媒体通信的历史与现状

80年代初,美国、日本和欧洲著名的计算机公司开始致力于多媒体技术的研究,把该技术应用于PC机。首先建立了基于局域网(LAN)的多媒体通信系统,如美国Xerox公司的以太电话(Etherphone),可以说是最早的多媒体通信系统。

进入90年代,企业的国际环境变得日益复杂,竞争日趋激烈,现代企业必须依靠高质量的信息及很高的信息获取能力才能立于不败之地。对企业特别是大公司而言,在经营活动中,希望借助更先进的通信系统达到缩短产品投放市场的时间,改善对用户的服务,减少差旅支出等;对个人而言,由于生活环境日趋复杂,生活节奏越来越快,也希望能随时获取信息,并采用计算机支持协同工作(CSCW)方式来提高工作效率。上述的需求和希望,传统的通信系统是无法完全满足的,唯有建设多媒体通信网才是解决的办法。

未来的家庭对多媒体应用需求主要有远程购物、远程医疗、远程教育、游戏等。

国外多媒体通信的研究开发首先是基于窄带综合业务数字网(N-ISDN),如美国的AT&T、Pacific Bell、奇科(Chico)和加利福尼亚州公立大学(CSUC)于1991年5月开始,利用ISDN BRI(基本速率接口)进行了多媒体DL(Distance Learning 远程学习)网络试验,这项试验全面评价了使用相对窄带ISDN支持DL网络是可行的,可传送较满意的实时视频质量;Bellcore的MediaCom多媒体系统使用Bell Atlantic ISDN线路提供多媒体业务,如远地电子白页(Remote Electronic Whitepages)、公用号码簿等,在实验基础上于1993年进入实用化阶段。1994年Bell Atlantic利用MediaCom系统为医疗行业提供网络支持,用ISDN D信道传药方、帐单等,用B信道传医疗图像、数据、多方会诊等;德国电信公司、德国科学研究中心和SNI于1994年10月共同开发公用多媒体业务,称为COBRA。主要面向中小型商业机构,利用ISDN路由器把各地的以太网和FDDI链接到ISDN。支持远程办公、远程医疗。

目前,基于ISDN网的多媒体会议电视和多媒体检索业务基本上达到实用水平。

国外宽带多媒体通信仍然处在研究开发、现场试验之中,少数系统进入了小规模商业应用。

从1994年7月开始日本进行了国家级的试验项目B-ISDN,对点播电视(VOD)、电子游戏传送点播业务、家庭购物、远程教育、远程医疗、电子图书馆、多媒体会议电视、市民画廊等20多项多媒体通信的应用逐项展开试验。

NTT的光缆骨干网(速率2.4Gb/s、10Gb/s)和ATM骨干网(共10个节点)已建成。200多个企业开始使用ATM进行多媒体试验,试验为期一年。

NTT为了使通信与有线电视(CATV)融合,还在进行CATV多媒体试验,主要是实现VOD业务。光纤接入网由NTT提供,CATV公司提供信息节目。1996年提供VOD、远程医疗等业务。

NEC公司已在奈良开通了基于ATM交换机的多媒体点播系统(Multimedia on Demand),称为HYPERMS,提供家用录相机质量的压缩图像编码(MPEG-1)和高质量运动

图像编码(MPEG-2)图像质量。

AT&T 公司为建设多媒体通信网,连续从美国 PacificBell、Bell Atlantic 电话公司获取 10 亿美元的合同。AT&T 在多媒体产品和业务方面走在前面,已推出 Telemedia 多媒体系统,交互型电视等产品系列。Pacific Bell 公司投资了 1 亿美元,把加州 7400 多个公立学校、图书馆和社区学院连至信息高速公路,提供多媒体会议电视。

美国许多公司开始进行小规模 VOD 现场试验,一些公司准备建设网络,提供多媒体业务。

德国电信局于 1995 年夏在巴登—符腾堡开始最大的“多媒体计划”。该计划为期两年,以 CATV 用户(约 4000 个家庭)为对象,提供的业务主要有 VOD、远程教学等,计划投资 8000 万马克。澳大利亚政府将在 4 年间投资 6000 万美元用于发展澳大利亚多媒体。香港电讯公司在多媒体开发方面也处于世界领先地位。香港电讯公司进行了 VOD 试验,已经投资 830 万美元,从 1994 年 9 月开始,有 50 个香港电讯机构进行技术性试验。香港电讯公司选定具有世界水平的 SYBASE 数据库建立 VOD 系统,首先提供的多媒体业务有包括 VOD、家庭购物、家庭银行等业务。另外,马来西亚电信公司也在试验交互式宽带业务;新加坡也将开始用 ATM 交换机进行 VOD 试验,投资 1000 万美元;韩国电信公司也宣布计划普及视频业务,开始的目标用户为 15000 个。

从世界范围看,电话网(PSTN)仍是主要的通信网,并将长期存在。在 PSTN 上提供多媒体业务有相当的市场需求。据 AT&T 分析,在 PSTN 上的多媒体业务将有数十亿美元的工业潜力。因此,在 PSTN 网如何实现多媒体通信业务是国际上普遍关注的问题。其中,ITU-T(国际电联)在本研究期(1993—1996)开展了低比特率 PSTN 的多媒体终端系列标准的制定,如 ITU-T 制定了 H. 324 建议、V. 34 建议 Modem 速率扩展到 33.6kb/s。MPEG-4 标准已经完成,这是一个用于存储播放场合的低码速运动图像编码标准,另一个低码速运动图像编码标准 H. 263 建议也已经完成,它可以用于会议场合,再加上低码速语音编码取得很大的进展,这样 PSTN 上的多媒体业务已能实用。目前,各国公司急于开发基于 PSTN 的多媒体产品,并扩大各个领域。应用实例有:音频图像电话会议、协同计算、共享白板、远程演讲/远程学习、Internet 电话、技术支持/客户服务、交互式游戏和交易、居家办公、远程购物、电子商品目录。

当前,国外商用多媒体应用市场热点是在多媒体会议电视和多媒体检索业务方面。

目前基于 IP 寻址的多媒体通信异军突起,发展势头极为迅猛。

Internet 目前已经家喻户晓了,Internet 的 Web 技术也已广为使用。随着 Web 技术的发展,特别是 JAVA, JAVA Script 和 plugin 技术的不断引入,Web 系统的能力越来越强大,它已经不只是提供简单的文本、图片等的信息检索服务,它还可以提供动画、声音乃至实时运动图像的服务。目前的 Internet 是不能提供多媒体服务的,其主要原因是 Internet 无法确保服务质量。在网络闲的时候可以提供一些较高档次和一些实时的业务,当网络拥挤时甚至连普通的信息检索都不能提供得很好。虽然说,Internet 不能提供多媒体通信服务,但是 Internet 的核心技术——IP 寻址技术及 TCP/IP 协议是可以用于多媒体通信的。由于它有种种优势,因此基于 IP 寻址的多媒体通信迅速地发展起来了。

目前世界各国、各大厂商都在研究基于 IP 寻址的多媒体通信。国际标准化组织其中包括 ISO(国际标准化组织)和 ITU-T(国际电联)也都积极进行基于 IP 寻址的多媒体通信的

研究制定相应的标准。

基于 IP 寻址的多媒体通信网与 Internet 的最大区别在于前者能从网络的角度保证所提供业务的服务质量和前者的主干网应该是宽带网,以便能向用户提供令人满意的多媒体信息服务。从网络角度来讲,目前进行的主要工作是怎样将 IP 寻址技术与 ATM 相结合的问题,这不仅仅涉及标准化的问题,也涉及具体的产品实现的问题。众所周知基于 IP 寻址的通信是无连接的通信,ATM 则是面向连接的通信。一个面向连接的网要去承载无连接的通信服务,是有很多问题的。目前从标准化角度提出了 LANE, classical IP over 和 ATM MPOA, 还正在研究 MPLS。从产品实现来说已经有 LANE IPOA, IP Switch IP Navigator, IP Switch 和 MPOA 等。新技术还在不断出现以试图解决 IP 网与现有宽带网的关系。

基于 IP 寻址的多媒体通信业务也在发展之中,有一大批原来认为必须在 B-ISDN 中开放的业务,业已证明这类业务可以很好地在 IP 网上开放。最有代表性的业务之一是点播电视(VOD-Video on demand),这个业务被认为是宽带业务中一种具有代表性的业务,在各国的 B-ISDN 实验网中几乎无例外地应用了这一业务。这项业务目前在 IP 网上已能很好地开放,国外已有这样的设备,国内也已研制完成基于 IP 的一整套 VOD 设备,并已投入商业性运行,其使用方便程度和服务质量都很好。

基于 IP 网的另一类业务是多媒体会议业务,由于多媒体会议业务要求的实时性是最高的,特别是在会议中的对话,在所有电信业务中对话的实时性要求最高。因而以前会议或会议业务通常在电路交换型网络中实现。近年来在基于 IP 网的多媒体会议业务有了突破性的进展,H. 323 标准的制定,以及与 H. 323 相关连的,特别是由于 IP 网的特殊性要求图像编码和语音编码尺度可变而提出了 H. 263+ 和 G. 729Extension 等的完成,使得在 IP 网上的多媒体会议业务可以很好地开展了。

在多媒体会议业务上,可以产生多种应用系统,最具有代表性的应用系统有三种。

(1) 综合业务多媒体会议系统

这是一项应用面相当广泛的应用。多媒体会议业务目前在国内的普及率是很低的,除了建设了一些 room base 的会议系统外,目前事实上还没有其它可以实用的会议系统,之所以形成这个局面,设备的价格昂贵是一个十分关键的问题,其中包括接入网的问题。基于 IP 的会议系统,由于大量智能化布线大楼的存在及其迅速普及,以及桌面系统价格的大幅下降,在很大程度上克服了上述难点,因而近期它将会有一个很大的发展。

(2) 远程医疗系统

远程医疗是一项综合度很高的多媒体通信业务,多媒体会议是它的一项核心内容,此外还包括多媒体信息检索、多媒体信息分配和多媒体信息采集业务等。远程医疗的发展已经经过了几个阶段,从 PSTN、PSPDN、DDN 到 ISDN 和 IP 网,经过了业务从低级到高级的发展过程,目前国内专业人士普遍认为 IP 网上的远程医疗将是方向,也正在进行新一代远程医疗系统的开发工作。

(3) 远程教学系统

远程教学与远程医疗有很大的相似性,它也是一项综合度很高的多媒体通信业务。由于校园 LAN 的建立,包括一些中小学都相继建立起智能布线系统,使得基于 IP 网的通信要求得到了很好的解决,目前远程教学 90% 以上是 IP 业务。

国内在多媒体通信方面发展很快,特别是智能化布线在国内的普及,它在很大程度上解

决了多媒体通信原来最难的解决的问题——用户接入问题,使得多媒体业务的发展的普及成为可能。

目前正在着手组建的国内最大的多媒体通信网,除了骨干网外,部分省的骨干网也将采用宽带网技术,这是一个基于 IP 寻址技术的多媒体网,在这网上将要开放十大类多媒体业务。从技术上来讲这是一个非常先进的网,许多技术在世界上都属首次,因而这个网的建成,将标志我国的多媒体通信技术进入了世界先进行列。

第二章 多媒体通信概述

多媒体通信是一门新技术,目前计算机界和通信界都在研究这门新技术,由于着眼点不同和使用研究方法不同,因而从论述上也不会相同,要对一门技术进行深入研究,概念和定义必须首先明确,本章从通信的角度出发来探讨多媒体通信中的一些概念和定义。

2.1 概念和定义

2.1.1 媒体

根据国际电联(ITU-T)的定义,媒体共有五类,它们是:

(1) 感觉媒体(Preception Medium)

感觉媒体指的是由人类的感官直接能感知的一类媒体。这类媒体有:声音、乐音、动画、运动图像、图形和噪音等。

(2) 表示媒体(Representation Medium)

表示媒体指的是用于数据交换的编码。这类媒体有:图像编码(JPEG、JBIG、H. 261、MPEG 等),文本编码(ASCII 码、GB2312、T. 101)和声音编码等。

(3) 显示媒体(Presentation Medium)

显示媒体指的是进行信息输入和输出的媒体。这类媒体有:显示屏、打印机、喇叭等输出媒体和键盘、鼠标器、扫描器、触摸屏等输入媒体。

(4) 存储媒体(Storage Medium)

存储媒体指的是进行信息存储的媒体,这类媒体有:硬盘、软盘、光盘、磁带、ROM、RAM 等。

(5) 传输媒体(Transmission Medium)

传输媒体指的是用于承载信息,将信息进行传输的媒体。这类媒体有:同轴电缆、双绞线、光纤和无线电链路等。

那么什么是多媒体通信?多媒体通信中的媒体究竟是指什么?根据国际电联的定义,多媒体通信中的媒体特指表示媒体(Representation Medium),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。

2.1.2 多媒体和超媒体

多媒体通信技术是一项新技术,这里对最常用的两个名词术语作一解释。

(1) 多媒体(Multimedia)

多媒体本身不是一个名词,而是一个形容词,它只能用作定语。因而单独说多媒体是没

有意义的,只有将它与名词相联系如:多媒体终端(Multimedia Terminal)和多媒体系统(Multimedia System)才是正确的说法。ITU-T 对多媒体服务的定义,是特指能处理多种表示媒体(Representation Media)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连的,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。

(2) 超媒体(Hyper media)

在多媒体通信系统中,另一个经常出现的词汇是超媒体。要讲清超媒体首先要了解什么是超级链接(hyperlink),超级链接一般人都把它看成一个十分神秘而新颖的概念,其实不然,它是一个十分古老的概念。在出版物中经常会出现“注”,由“注”你可以找到一段与之相关的文字或一篇与之相关的文章,这种由“注”而链接到一段文字或文章的链即称为超级链,同样道理,超级链也可以将若干不同媒体链接起来,其集合便称为超媒体。“注”这种概念十分古老,因而超级链也是十分古老的概念,只不过超媒体这种媒体组织方式不适用于文章的组织,文章一般采用线性顺序组织方式,这就是为什么几千年来超媒体技术没有获得发展的原因。由于近年来计算机技术的发展,一种特别适宜于超媒体信息的排序和检索方式的发展,才使超媒体技术获得了广泛应用。

2.2 多媒体通信的三大特征

具有如下三个特征的通信系统称为多媒体通信系统。

1. 集成性

多媒体通信系统中的集成性指的是能对下述四类信息进行存储、传输、处理、显现的能力。

(1) 内容数据(Content Data)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(File)。

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。这内容部分是真正要传送的实质所在,我们称这内容部分的信息为“内容数据信息”。内容数据信息是用单一媒体的编码标准来表示的信息。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

多媒体和超媒体客体可用他们在实时交换环境下的适当标准来进行表示和编码,它用到单媒体的一些表示标准。

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,是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(Object),另一类称为文件(File),围绕这两种表达形式,产生两类国际标准。

多媒体和超媒体信息的标准必须具有下列特点:

• 在表示客体(或文件)之间可以有不同的时间同步算法(绝对时间关系同步、相对时间关系同步、链接同步、循环同步和条件同步)。

- 具有表示客体(或文件)间空间复合的能力和机制。
- 可以用超级链(Hyperlink)去引用外部的表示信息。
- 可以定义用户的不同输入请求。
- 定义客体(或文件)间的链接,例如:事件(Event)和反应(Action)的链接。
- 描述与客体相联系的项目(Project)信息,详细说明它是如何在用户面前显现的。
- 提供一种可引用内容或将这些内容包含在 MH 客体之中的机制。

(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(MHI)。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。

脚本信息至少需具备如下一些特点:

- 能控制对结构的操作;
- 宣布全局控制事件;
- 复杂的定时操作;
- MHI 客体的表示;
- 外部处理机的调用;
- 库函数的调用;
- 定义校核点及从校核点的恢复功能。

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

上面所述的信息是三类低层信息,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。它不像前三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三类的基础之上的。

一个重要的例子是:用户工作站执行了 Scriptware 后,产生的返回踪迹(trace)是随着不同应用场合(debugging, execution, learning, charging, statistics)而不同的。

另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、数据包等)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到所需信息的位置。目录信息就是典型的特定的应用信息。

2. 交互性

交互性指的是在通信系统中人与系统之间的相互控制能力。很显然,交互性是多媒体通信系统必备的一个特性,但并非其独有的特性,目前许多通信系统也有着程度不等的交互性。如信息检索系统,它一般都提供菜单和征询单两种用户与系统的交互界面,用户可以通过点菜单或填写征询单,将用户的要求告诉系统,系统根据用户的要求,将满足条件的信息送给用户。用户与系统通过这一简单的交互过程完成了通信过程。

在多媒体通信系统中,交互过程将不再是这么简单。诚然,菜单和征询单这一类简单的交互过程在多媒体通信系统中仍将使用,以提供简洁而明了的交互操作,但是光有简单的菜

单和征询单的交互过程是不能满足多媒体通信系统的全部需要的,它将需要更为复杂的交互操作过程。

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、输入域、对话框、敏感区、敏感字段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。

当然,应该指出,在多媒体通信系统中基于视窗的人机接口界面与PC机的视窗区别是很大的。从用户操作的观点来看,二者给用户的操作感觉是十分相像的,基本上是PC机的视窗提供的人机操作方式,多媒体通信系统中基于视窗的人机接口界面基本上都支持;从系统角度来考虑,二者的差别就十分大了。PC视窗的全部操作是本机操作,而多媒体通信系统基于视窗的人机接口则完全不是本机操作,而是本地终端与远地数据库(或其它主机)的交互操作,它的每一个动作,如出一张菜单,给出一个列表项等则全部受到远地主机的控制,因而将是一个十分复杂的通信过程。

除了人机接口之外,多媒体通信系统中交互性的另一个方面是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将要存储、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的特征,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:数字彩色电视机可以对多种表示媒体(图像编码、声音编码)进行处理,也能进行多种感觉媒体(声、文、图)的显现,但用户除了能进行频道切换来选择节目外,不能对它的全过程进行有效的选择控制,因此,彩色电视系统不是多媒体系统。点播电视VOD(Video on demand)就不一样了,它可以对其全过程进行有效的控制,想看就看,想停就停,因此,VOD系统是多媒体通信系统。

3. 同步性

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它数据库中。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。

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,可以这样说信息

的同步与否,决定了系统是多媒体系统还是多种媒体系统。另外,多媒体通信系统中的同步性也是在多媒体通信系统中最为困难的技术问题之一。一般来说,多媒体通信系统是一个资源受限的系统,所谓的资源受限指的是两种情况:通信速率受限和终端内存受限。如果这两个方面没有限制,同步本来不会有很大的技术难点,譬如说信道通信速率不受限,那么只要发端完全安排好信息媒体间的关系,在接收端就完全忠实的复现出来,信息同步将不成问题,当然在信道的通信速率受限的情况下,接收端的信息间同步就要困难得多;另外,如果接收端存储器的存储量是无限的,将所有信息全部接收下来,然后在终端内同步播出,这种场合下同步问题也好解决,但实际上这个条件经常是无法满足的,因而使同步问题就变得很困难了。

多媒体通信系统中,同步可以在三个层面上实现,这三个层面是:①链路层级同步,它通过信息帧结构的合理设计来实现;②表示层级同步,它通过在客体(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;③应用层级同步,它通过脚本的设计来实现。

(1) 链路层级同步

在实时多媒体通信系统中,特别是引入运动图像的多媒体通信系统中,链路层级同步是十分重要的。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的,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。第二种是用于存储读出系统的,其应用场合有点播电视、运动图像检索、数字录相机、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。

· 比特交织的帧结构

用于会议电视的,由 ITU-T H. 221 建议定义的帧结构是一种典型的比特交织的帧结构,以单一 64kb/s 信道的帧结构,如图 2-1 所示。其中每 80 个 8 比特组成一个帧,80 个第 8 比特组成公务子信道,每帧中公务子信道(SC)的比特 9~16 称 BAS,BAS 的值确定了图像、语音数据在信道中的占有率,从而确定了图像、语音和数据是以比特交织的方式存储在去信息的数据流中。

1	2	3	4	5	6	7	8	8比特组编号
子	子	子	子	子	子	子	FAS	1
信	信	信	信	信	信	信	BAS	9
道	道	道	道	道	道	道	(ECS)	17
1	2	3	4	5	6	7		24
								⋮
								80

图 2-1 单一 64kb/s 信道的帧结构

· 块交织的帧结构

用于数字录像的存储读出系统的国际标准 MPEG-1 的帧结构,就是一种块交织的帧结构。由于录像在回放时引入一个固有延时是不会影响其质量的,只要这个固有延时不要太大是不会被人感知的。块交织的帧结构的优点是简单,其缺点是将会引入较大的固有延时,这种特点正好能满足存储读出系统的要求。

(2) 表示层级同步(复合客体和超级链)

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,一段运动图像片段也可以是一个客体,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引入同步机制,构成多媒体复合客体;或者用超级链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,故称表示层级同步。表示层级同步有五种类型。

• 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,如图 2-2 所示。

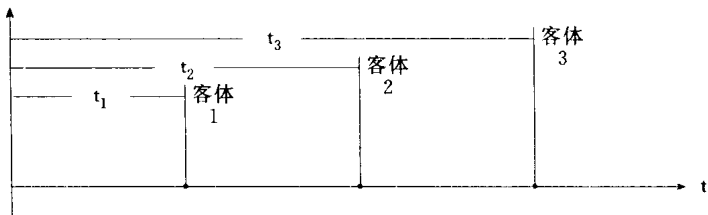


图 2-2 绝对时间同步

• 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之中,确定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,如图 2-3 所示。

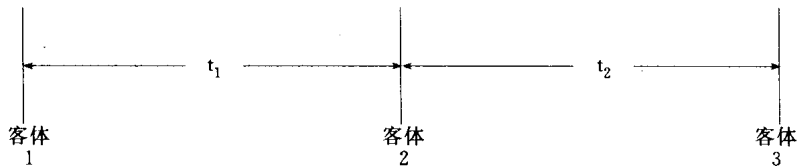


图 2-3 相对时间同步

• 链式同步

在复合客体的复合过程中或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。

• 循环同步

在复合客体的复合过程中或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重复出现两次或多次,这种同步称循环同步。

• 条件同步

在复合客体的复合过程之中或超媒体由超级链的链接过程中,只有当某一个条件满足后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(脚本)

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(MHEG 标准)或文件(Hytime 标准中),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

通信还是不够的,还必须引入应用层级同步——脚本技术。什么是脚本(Script)?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,它将来实现完善的多媒体通信。以电影为例,在一部电影里有许多演员,每个演员有多少不等的台词片段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是一个一个客体(包括复合客体 and 简单客体),要将一个一个客体来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。

多媒体通信系统中,同步可以在三个层面上实现。但这并不是说,一个多媒体通信系统必须同时具有这三个层面的同步,但它必须至少用到了其中一种同步方式。当然,同步方式用得越多,系统的性能就越完善。

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。

2.3 多媒体通信业务

多媒体通信业务种类繁多,可以这样说,今后的通信新业务都将会是多媒体业务,特别是宽带通信业务则将全部是多媒体业务,为了能对多媒体通信业务有一个全面的了解,对多媒体通信作一个分类是必要的。

根据国际电联(ITU-T)的定义,多媒体通信业务共分为6种。

会议业务(Conference Services)

——多点间,人与人之间通信。

谈话业务(Conversation Services)

——点与点,人与人之间通信。

分配业务(Distribution Services)

——点对多点,机器对人的通信。

检索业务(Retrieval Services)

——点对点,人与机器间的通信。

采集业务(Collection Services)

——多点对一点,机器与机器或机器与人间通信。

消息业务(Message Services)

——点对点,人→机→人,存储转发。

在这6类多媒体通信业务中,进一步归类的话,可以归为以下4类。

(1) 人与人之间交谈型多媒体通信业务

会议业务和谈话业务在多媒体通信的体系中,实际上是一类业务,它们都是人与人之间的通信业务,采用的标准基本上是一样的,比如说它们的显示格式都是用CIF(Common Image Format)和QCIF;在ISDN环境中采用的标准都是H.320,在PSTN环境下都采用H.324等,都是用于人与人之间的交谈。它们的不同之处是会议业务是多点之间的人与人

之间的交谈,谈话业务则是两个人之间的交谈,另外的不同之处是图像质量和声音质量,一般来说会议业务的要求更高一些。

(2) 人与机之间交互型信息检索多媒体通信业务

多媒体分配业务和多媒体检索型业务在多媒体通信的体系中可以归为同一类,即人与机之间的检索型多媒体通信业务。这类业务完成的是人与主机之间的交互操作,操作有两个方面的内容,其一是用户通过人机接口向主机发送检索请求,主机接收到用户的要求后,将满足用户要求的信息传送给用户,以完成用户和主机的交互过程,实现检索的目的;其二是用户通过人机接口与主机交互信息,通过交互人与机之间完成某种交易工作,如远程购物、电子银行。多媒体分配业务和多媒体检索业务的主要区别是:多媒体检索业务是一个人对一台机器之间(点对点)的交互业务;而多媒体分配型业务则是一台主机对多个人(一点对多点)的人机交互业务,这种业务的一个最典型的例子是,当在开一个多媒体会议时,参加会议者有几位,当讨论到一份文件时,一个会议参加者去检索,主机根据用户的要求将需满足条件的信息发送给全体会议参加者,这就是多媒体分配业务。

(3) 多媒体采集业务

多媒体采集业务是一种机与机或人与机之间多点向一点汇集信息的业务。信息采集业务目前用得十分普遍,如城市交通的信息采集系统,智能大楼的全楼监控系统,工业监控系统等,当然目前的系统并非多媒体系统,随着信息业的发展,信息处理技术的发展,多媒体采集业务将是信息采集和监控系统今后的发展趋势。

(4) 多媒体消息业务

多媒体消息业务是一种存储转发型多媒体通信业务。用户 A(人)将要传送给用户 B 的多媒体信息发送给系统中的主机,在主机中存放起来。用户 B 得知有传送给他的信息后,从主机中提取由用户 A 发送来存入在主机中的信息,并在用户 B 的终端中显现出来,这就是多媒体消息业务的典型通信流程。多媒体电子信箱是典型的多媒体消息业务。

在实用生活中,这些业务并不都是以孤立的形式存在的,而是以相互交织的形式存在的。多媒体业务的应用实例如表 2-1 所示。

表 2-1 多媒体业务的应用实例

应 用	具 体 应 用	业 务 类 型
多媒体合作应用	电子出版	多媒体会议业务
	协同设计	多媒体会议业务
	联合软件开发	多媒体会议业务
多媒体会议和群体工作	管理会议	多媒体会议业务
	技术会议	多媒体会议业务
	商业协商	多媒体会议业务
学习和教育	讲课	多媒体会议业务
	教育信息检索	多媒体交互型信息检索业务
	远程培训	多媒体交互型信息检索业务
远程专家协商	远程故障诊断和维护	多媒体会议业务
	远程故障检索	多媒体会议业务
	专家系统	多媒体交互型信息检索业务

续表

应 用	具 体 应 用	业 务 类 型
远程医疗	诊 断	多媒体会议业务
	专家会诊	多媒体会议业务
	文件检索	多媒体交互型信息检索业务
	X—射线图像检索	多媒体交互型信息检索业务
电子交易	远程购物	多媒体交互型信息检索业务
	家庭银行	多媒体交互型信息检索业务
娱 乐	卡拉 OK	多媒体交互型信息检索业务
	游 戏	多媒体交互型信息检索业务
	点播电视(VOD)	多媒体交互型信息检索业务
多媒体信息应用	旅游信息检索及预约	多媒体交互型信息检索业务
	房地产交易	多媒体交互型信息检索业务
多媒体检索应用	远程文件	多媒体交互型信息检索业务
	远程图书馆	多媒体交互型信息检索业务
多媒体交互应用	远程投票	多媒体交互型信息检索业务
多媒体监视	远距离证实(诉状)	多媒体交互型信息检索业务
	安全和警卫	多媒体会话业务
	远距离小型会议	多媒体会话业务
多媒体邮件应用	多媒体电子信箱	多媒体消息业务
多媒体分配型应用	NVOD	多媒体分配业务

思 考 题

- 1. 什么是媒体？媒体共有几类？
- 2. 什么是多媒体？
- 3. 什么是超级链？什么是超媒体？
- 4. 请叙述多媒体通信的三大特征。
- 5. 请解释一下，多媒体系统与多种媒体的区别。
- 6. 同步是怎样实现的？
- 7. 多媒体通信有多少种业务？请指出各种业务的特点。

第三章 多媒体通信中的信息类别

多媒体系统中的媒体特指的是表示媒体,也就是指信息的编码。根据国际电联的文件,在多媒体通信系统中可以处理的信息分为:内容数据(Content Data)信息、多媒体和超媒体信息、脚本(Script)信息及特定的应用信息 4 大类。

3.1 内容数据(Content Data)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(File)。

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。这内容部分是真正要传送的实质所在,我们称这内容部分的信息为“内容数据信息”。内容数据信息是用单一媒体的编码标准来表示的信息。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)。

1. 文本

文本含有三个方面的内容:

- 符号
- 符号的字型和字体
- 在数据传送和操作管理中的符号编码

西文/文本显示的标准格式为 24 行×80 列,对单字节编码的文本来说,满屏存储量约为 2kB,中文文本显示格式为 24 行×40 列,中文为双字节编码,满屏存储量也约为 2kB。

ISO646、ISO10646、T. 101、ASCII 和 GB2312 等是一些著名的国际标准和工业标准。

2. 图形

图形编码一般有如下几种方法。

(1) 镶嵌图形法

这是一种最简单,但又极高效的图形编码,用它组出来的图形很像室内装潢中用的马赛克拼图一样。马赛克块越多,图形质量越好。

(2) 动态再定义图形

动态再定义字符是一种很特殊的构图技术,它是一种点阵组图形,可以组成质量相当不错的图形,只是编码效率不太高,它适用于一种图形多次重复的场合。

(3) 几何图形法

几何图形法是用点、直线、三角形、矩形、多边形、圆弧等几何元素来表示图形。以直线为

例,直线只要有起点、终点坐标、线宽和颜色就可以充分表示一条直线了。对于椭圆,只要有左上角、右下角、线宽、颜色、内填色等参数就可以表达了。因而这是一种很高效的编码方法。它的局限性在于不是所有图形都可以用几何图形元素来表示的。

(4) 增量

增量其本质上是以折线来代替曲线,当变化区间变得很小时,折线中的直线段即称为增量。

国际上对增量作过深入的研究,提出 6 种增量编码法,其中两种是最好的,这二种是:区域增量编码和轨迹增量编码。

著名的图形编码的国际标准和工业标准有:T. 101、T. 150、ISO8632、ISO7942 等。

3. 静止图像与二值图像

大多数个人计算机显示器的显示格式为 480 行,每行 640 个像素,因而每屏约有 300k 个像素,对于二值图像来说一个像素用 1 比特来表示,因而每屏为 300kbit。对彩色图像(Y CrCb)如采用 4:1:1,则需 $12 \times 300\text{kbit} = 360\text{kB}$,如 RGB 均为 8bit,则需 $24 \times 300\text{kbit} = 720\text{kB}$ 。还有一种彩色技术,即调色板技术,每个像素采用 8bit 编码(256 色),则全屏存储量为 300kB。

静止图像是与时间无关的相片图像,是颜色、色饱和度、强度连续变化的二维相片。对于一种典型的静止图像,每幅图像的编码比特数为 $640 \times 480 \times 24\text{bit} \approx 900\text{kB}$ 。

二值图像是一种特殊的静止图像,对每一个像素只有二种状态(1 或 0),因而无色调和强度变化。一幅典型的二值图像比特数为 $640 \times 480 \times 1\text{bit} = 307.2\text{kB}$ 。

由于一种多灰度的静止图像可以看成是多平面二值图像的组合,因而在编码技术的研究中没有绝对的界限。一般来说,游程编码和 JBIG 编码适用于二值图像,以离散余弦变换为基础的 JPEG 算法适用于静止图像,但也不绝对如此,已有实验证明 JBIG 算法也可用于彩色静止图像,当量化比特数小于 6bit 时,JBIG 算法还是最优的。

JBIG 算法和 JPEG 算法已经是国际化的并受到广泛的应用。目前还有不少新的算法正在进行研究,如小波变换,分形编码和模型编码等。

4. 声音

声音是指人们在听觉范围内的语言、音乐、噪声等音频信息。对于电话的话音(0~3.4 kHz)抽样速率为 8kHz,每一个抽样进行 8bit 量化(用 μ 率或 A 率),这样不压缩的语音码率为 64kb/s。对立体声要用 44.1kHz 抽样,每个抽样用 16bit 量化,这样未压缩数码率为 705.6kb/s。

语音编码技术标准绝大多数是(ITU-T)(CCITT)提出来的。G. 711 描述了 8kHz 抽样,8bit 量化的语音编码(PCM),其码率为 64kb/s;G. 722 描述高质量语音(0~7kHz)的编码,采用子带编码和自适应脉码调制(ADPCM)技术,其码率为 48、56、64kb/s;G. 728 是采用码激励线性预测技术,将电话语音质量的话音压缩至 16kb/s;G. 723.1 是最新建议,它将电话语音压缩至 5.3kb/s 和 6.3kb/s;在 MPEG 中也有语音编码,通常被称为 MPEG Audio,它的编码率为 8~32kb/s。MPEG Audio 算法也用于数字式盒式录音机和数字式语言广播,它采用子带编码压缩技术。

5. 运动图像

运动图像指的是人们所能见到的活动图像(如电视、电影、录像等)。由于要使图像活动起来需要每秒钟内有足够多的帧数(如电视每秒 25 帧),这样运动图像未压缩的数据量就很大(对每秒 25 帧的广播电视,未压缩码率为 22Mb/s)。由于运动图像相关性很强,这包括图像的帧内相关性和两幅连续图像的帧间相关性,该特性可以充分用于编码压缩。这导致了为数众多的运动图像压缩标准。

运动图像压缩标准可分为两大类:

(1) 点对点实时通信

用于点对点实时通信的运动图像压缩编码,其最大的要求要实时性好,应用场合为可视电话和会议电视。目前主要有两种标准:H. 261 用于 $P \times 64\text{kb/s}$ 的场合,H. 263 用于 20kb/s 以下的场合。

(2) 存储读出系统

MPEG 是用于存储读出系统(如信息检索系统,数字录相机,VCD 和数字电视等),而不是用于实时点对点通信,其原因是 MPEG 算法将引入较大的固有时延。

MPEG-1 是家用录相机质量的压缩图像编码,其最高码速为 1.5Mb/s,这个速率光盘(CD-ROM)的读出速度是相适配的(150kb/s),目前流行的 VCD 用的就是 MPEG-1。

MPEG-2 是为高质量运动图像编码(如:HDTV、SDTV 等),最高编码速率为 40Mb/s,广播电视质量的 MPEG-2 其码率为 6~8Mb/s。

MPEG-4 已经完成标准化的工作,其码速为 20kb/s 以下。

3.2 多媒体和超媒体信息

多媒体和超媒体客体可用他们在实时交换环境下的适当标准来进行表示和编码,它用到单媒体的一些表示标准。

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,是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(object),另一类称为文件(file),围绕这两种表达形式,产生两类国际标准。

多媒体和超媒体信息的标准必须具有下列特点:

(1) 在表示客体(或文件)之间可以有不同的时间同步算法(绝对时间关系同步、相对时间关系同步、链接同步、循环同步和条件同步)。

(2) 具有表示客体(或文件)间空间复合的能力和机制。

(3) 可以用超级链(Hyperlink)去引用外部的表示信息。

(4) 可以定义用户的不同输入请求。

(5) 定义客体(或文件)间的链接,例如:事件(Event)和反应(Action)的链接。

(6) 描述与客体相联系的项目(Project)信息,详细说明它是如何在用户面前显现的。

(7) 提供一种可引用内容或将这些内容包含在 MH 客体之中的机制。

近年来,开始对多媒体和超媒体客体(Object)的表示和编码进行了标准化。最具代表性的有下述几种: