



多媒体电脑

选配与使用

杨鸿林 许志坚 主编



机械工业出版社

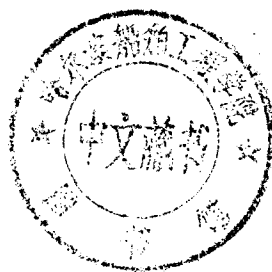


466047

70377
817

多媒体电脑选配与使用

杨鸿林 许志坚 主编



机械工业出版社

JS136/1-4

多媒体技术为什么会如此流行?面对市场上众多的多媒体电脑与多媒体组件,应如何鉴别与选择?接下来,您又该如何最大限度地发挥多媒体电脑的功能呢?对于诸如此类的问题,本书都将为您一一回答。

本书共分十二章,分别介绍了多媒体技术的起源、应用和现状;多媒体电脑常用部件的选配及其安装方法,如声音卡、视频卡、电视卡等;Windows和 Windows 95 中的多媒体部件及其用法;多媒体软件开发的步骤和方法,其中重点介绍了 ToolBook、Visual Basic 等常用多媒体开发软件和工具;本书的最后详细介绍了 Windows 的 MCI(媒体控制接口)的使用方法和命令。

本书内容详尽、语言流畅、图文并茂,适合广大计算机用户及各类多媒体爱好者阅读。

图书在版编目(CIP)数据

多媒体电脑选配与使用/杨鸿林,许志坚主编. —北京
:机械工业出版社,1997.4
ISBN 7-111-05616-7

I. 多… I. ①杨… ②许… II. 多媒体-微型计算机
基本知识 N. TP391

中国版本图书馆 CIP 数据核字(97)第 03025 号

出版人:马九荣(北京市百万庄南街1号 邮政编码 100037)
责任编辑:赵少平
封面设计:张洪波
三河永和印刷有限公司印刷·新华书店北京发行所发行
1997年4月第1版·1997年4月第1次印刷
787mm×1092mm 1/16·13.5印张·330千字
0 001—5 000册
定价:24.00元

前 言

多媒体是最近几年来电脑与信息界里一个新的领域,对于接触电脑的人来说,只要提起多媒体,无不对其如雷贯耳。但是,要全面了解多媒体电脑的方方面面决不是一件容易的事。例如,IDE、SCSI、Sony 接口光驱有何区别?要安装多媒体部件应了解哪些知识?衡量声音卡的性能指标有哪些?解压卡、电视卡、捕捉卡之间有什么区别,它们分别用来干什么?波形音频文件和 MIDI 音频文件有何区别?为什么动画文件有如此之多(如 FLC 文件、MOV 文件)?播放动画的软件有哪些?有哪些软件可以制作动画?Windows 的 AVI 文件是什么?如果用户想开发多媒体软件应该怎么办?有哪些工具可以利用?Windows 和 Windows 95 从哪些方面对多媒体给予了强有力的支持?等等。

本书对诸如此类的问题均给出了明确的答案。本书第 1 章讲述了多媒体电脑的标准、应用领域和发展前景。第 2 章介绍了一些常用多媒体技术,它对用户了解产品的性能、指标是大有裨益的。在接下来的第 3 至 7 章分别介绍了如何安装多媒体部件。第 8、9 两章详细讲述了 Windows 和 Windows 95 中多媒体部件的使用方法。第 10 章介绍了一些播放和制作多媒体数据的软件。第 11、12 章主要是为有兴趣从事或想了解多媒体软件开发的人员编写的,它首先介绍多媒体软件的两种开发方法,然后介绍目前比较流行的多媒体创作软件中 ToolBook 和非常适于开发多媒体软件的 Visual Basic 语言,并对 Windows MCI 进行了较为详尽的讲述。

本书作者曾经长期从事多媒体软件的开发工作,并曾阅读了目前已出版的大量多媒体书籍。这些书籍大致分两类,一类非常专业,如讲述如何用 C++、VB 开发动画等,比较适合专业人士阅读。一类非常简单,尽管有的书籍洋洋洒洒几十万言,但真正有用的内容却实在太少。因此,本书希望告诉读者一些真正有用的内容。全书紧紧抓住多媒体部件装配、使用和多媒体软件开发这样一条主线,由简单到复杂,由低级到高级,使读者对多媒体电脑的方方面面有一个完整的了解,进而为进一步的工作打下良好的基础。

本书由杨鸿林和许志坚主编,参与本书编写工作的还有王一润、刘柏、何光明、张芝先、魏耘、刘浩洋、曹家红、郭玲文、王强、陈晓华、樊爱京和赵陇等。

尽管作者在编写此书时已尽了最大努力,但由于时间仓促,加之作者水平有限,因此,不当之处在所难免,尚希读者批评指正。

编 者

1997 年 2 月

第1章 多媒体综述

媒体一词最初源于英文“media”，它在英文中为中介物或媒介物之意，即传播信息的载体。说到此，大家自然不难想到报刊、书籍、电影、电视、广播等，事实也的确如此，所有这些媒体均以各自不同的形式进行信息传播，报刊、书籍传播的是文字和静态画面（相片），电影、电视传播的是动态画面，广播传送的则是声音。显然，对于广播而言，它只能传送声音，而对于电视，它不但可以传送文字，而且可以传送声音和图像，即前者只能以单媒体方式（声音）传播信息，而后者则可以文、图、声、像等多种媒体形式进行信息传播。至此，我们已不难看出，所谓多媒体即是以多种媒体形式——文字、图形、声音、动画、图像来传播信息。

1.1 多媒体与计算机

我们大家都知道，衡量一台计算机优劣的标准多种多样，但对用户而言，归纳起来却不外乎这么几条，即 CPU 运行速度的快慢、存储器容量的大小（其中包括了高速缓存、主存和外存）以及人机接口是否友好。CPU 的运行速度决定了计算机对你所提出要求的反应快慢。计算机的三种存储器功能则各不相同，高速缓存可以加快主存与辅存之间的数据交换，主存储器容量则决定了计算机一次能从外存调入运行的程序大小，外存包括两部分，即软盘和硬盘，其中软盘的容量是无限的，它仅取决于你身边所拥有的软盘数量，计算机的硬盘容量则是有限的，当硬盘被占满后，用户必须删除一些不常用的内容以腾出空间。人机接口比较复杂，它既涉及到计算机的硬件，也涉及计算机软件。

对于以上三条指标，其中前两条，即计算机的速度和存储器容量完全取决于集成电路技术的进步，仅对微型计算机而言，目前 CPU 的主频最高已达 120MHz，8MB 及 16MB 的存储器芯片，2000MB(2GB)及 4000MB(4GB)的硬盘也已出现，其将来的发展无非是指标上的提高。但对计算机的人机接口而言，情况则比较复杂，因为它涉及的范围比较广，下面我们就这个问题进行探讨。

1.1.1 发展及现状

从根本上讲，计算机和其他所有人类发明的工具一样，不管它怎样发展，其最终目的都是为人服务的。即使它将来具有了一定的智慧，它仍然只是一个工具。所以，人们就必须按照自己的目的（即希望它做什么）将所能提供的数据处理成计算机所能识别的格式告诉给计算机，然后计算机对这些数据加以处理，再以一定的形式输出出来。

紧接下来，大家提出的问题就是，在计算机处理数据的过程中，人是如何对其进行干预的，即人是如何对计算机进行有效控制的呢？我们可以把人对计算机的控制，即人机界面分为三个发展时期：①早期用户通过专门的操作员和机器打交道，用户甚至根本看不到计算机；②出现了分时系统，用户通过终端面对屏幕，但界面仍然不够友善，不易学，不好用；③伴随着微型机发展，人机界面大为改善，特别是出现了图形用户接口（GUI）、下拉弹出式菜单

等等。

1.1.2 缺点

显然,如果你的工作需要经常与计算机打交道,你就会逐渐感到工作太乏味了。因为你所面对的机器既不能听,也不能说,画面过于单调,毫无动感。即使出现了 GUI 之后,也同样存在这种问题。

1.1.3 希望

• 因此,你可能想,如果有一天,你的计算机既能听懂你所说的话,又能说会唱;你既能在上面看电视,又可随时将所喜爱的节目录下来;你还可通过目前现有的各种通信网络(电缆、光纤、卫星电话网及各种计算机局域网或广域网)或将要建立的各种信息高速公路,随时抽看远在千里之外或万里之遥的各种图书馆和资料室里的自己喜爱的各国电影;如果你有一个还算不错的歌喉的话,你还可用它来唱卡拉 OK;你还可利用计算机和远方的亲朋好友进行通话,并可看见对方的动作与表情,从而真正实现古人“海内存知己,天涯若比邻”的梦想;如果需要的话,你还可以利用计算机来收发传真。

总之,你希望将来有一天,你的计算机既可作为电视机,又可以是一架不错的音响,同时它还应是一台传真机。如果配上适当的软件,你的计算机还可以成为各种乐器(如钢琴、鼓、小号等),你可随时通过计算机进行演奏,以调节你的情绪和记录你的各种乐思。

1.1.4 可能

紧接着的问题就是,就目前的技术水平而言,有没有实现的可能呢,或者目前尚未实现,但只是时间问题。我们可以告诉读者的是,答案是肯定的。

随着微电子技术和进步,高清晰度电视(HDTV)、高保真音响(Hi Fi)、高性能录像机、高速通信网和智能化计算机技术正迅速融为一体。一方面是声像产品引入微型机控制和处理,使电视机和各种声响设备实现了数字化;另一方面是计算机应用几乎渗透到人类活动的各个领域。发展的结果必然是把计算机软硬件技术、智能化的人机接口技术、高速通信技术、计算机网络技术和数字化的声像技术集成为一个整体。

1.2 多媒体技术发展简史

综上所述,我们可以看出,多媒体的产生既是一种需要,而且还存在这种可能。那么,多媒体技术又是如何一步步发展的呢?

20 年前,人们曾把几张幻灯片配上同步的声音,称为多媒体系统。而今天,随着微电子、音像、计算机和通信技术的发展,给多媒体技术赋予了新的内容,多媒体系统也发生了质的变化。

1984 年,美国 Apple 公司在 Macintosh 上为了改善人机之间的接口(界面),大胆地引入位映射(bitmap)的概念来对图进行处理,并使用了窗口(Windows)和图符(Icon)作为用户接口。Apple 公司试图将个人计算机当作启发人们创造性的高级工具来设计,希望个人机不仅仅是计算和办公的工具,也是处理信息、传送信息的工具,处理对象不只是数据和文字,还应

有图形和图像,使非专业人员在上机后,也能很快地操作使用计算机。Apple 公司的设计师们最早用 GUI(Graphics User Interface,图形用户接口)取代了 CUI(Character User Interface,字符用户接口),用鼠标器和菜单取代了键盘操作。1987 年又引入了“超级卡”(Hypercard),使 Macintosh 机成为易使用、易学习,且能处理多媒体信息的机器,一直受到计算机用户的赞誉。为了改善人机之间界面,Microsoft 公司于 1985 年开始了窗口(Windows)的研制,历经 Windows 286、Windows 386,直到 90 年代推出 Windows 3.0,它为图形和多媒体技术率先进入市场做出了贡献。

1986 年,荷兰 Philips 公司和日本 Sony 公司联合研制并推出了交互式紧凑光盘系统 CD-I(Compact Disc Interactive),同时它们还公布了 CD-ROM 文件格式,得到了同行的承认,并成为 ISO 国际标准。该系统把高质量的声音、文字、图形、图像都进行了数字化,并像程序一样放入 650MB 的只读光盘上,用户可以到电视机上显示。后来 CD-I 随着 Motorola 微处理器的发展也不断改进,并广泛用于教育、培训和娱乐。

1991 年,Intel 和 IBM 公司推出了多媒体卡 Actionmedia I,它被作为微通道和 ISA 总线的选件。该卡由采集板和用户板等组成,其视频处理器被升级为 i750B(包括 82750PB 和 83750DB),DVI(Digital Video Interface,数字视频接口)音频信号由 AD 公司的 AD SP2150 处理。同时,这两家公司也开发了在 Windows 3.0 和 OS/2 1.3 下运行的多媒体音频和视频内核 AVK(Audio Video Kernel),AVK 提供了低层编程接口 Beta DVMCI(Digital Video Media Control Interface),后来又扩展到了 Windows 3.1 和 OS/2 2.0 上。世界上已有几百家公司为其开发软件,在美国已广泛应用。

1.3 多媒体电脑标准

多媒体是一项综合性技术,其中包括计算机、声像、通信等。在多媒体技术发展初期,很多产业就非常重视标准化问题。标准化工作的前期是研究、实验、测试,再经过竞争、筛选和优化,它是众多研究单位长期研究开发再和生产厂家结合的结果。多媒体电脑标准是工业界和科技界合作的智慧结晶,它是在最广泛的信息基础上制定的标准,所代表的技术是先进的。标准的出现推动相关工业生产的大幅度增长,产品的成本和价格大幅度降低,并大大改善了多媒体产品之间的兼容性。符合标准的产品具有通用性,其结果,又可导致应用的迅速增长。

早在 1990 年 10 月,在 Microsoft 公司多媒体开发工作者会议上就提出了多媒体 PC 机(MPC)技术规范 1.0。其要求的最低规格如表 1-1 所示。从表中可以看到,多媒体 PC 机实际上是指满足或超出这种规格的特定的一类 PC 机。后来,MPC 理事会重新精炼了多媒体 PC 机的定义,去掉了 80186 处理器,认为其 CPU 最低要用 20MHz 的 386SX。

MPC 规格考虑到了迅速发展多媒体技术,并为今后功能的完善留有充分的余地。例如,美国 IBM 公司于 1991 年 10 月推出了有代表性的多媒体 PC 机系统——IBM PS/2 Ultimedia Model M57 SLC,它的规格与 Microsoft 公司的 MPC 规格略有不同。

1993 年多媒体微机市场委员会发布了多媒体微机的性能标准 2 规格,在建立新的多媒体微机的性能标准 2 规格的同时,也建立了新的多媒体性能标准。新规格是与原有的 MPC 规格相兼容的超级版本。表 1-1 列举了 MPC 标准 1 和标准 2 之间的一些主要区别。多媒体

微机市场委员会还宣布,将给 MPC 标准 1 和标准 2 商标颁发证书。为此,申请使用该证书的硬件要遵照有关规格进行测试。

表 1-1 多媒体微机标准——MPC 标准 1 和标准 2

最低要求	标准 1	标准 2
RAM	2MB	4MB
处理器	16MHz 386SX	25MHz 486SX
CD-ROM 驱动器	持续传送速率 150KB/s, 平均查询时间 1s	持续传送速度 300KB/s, 平均最快查询时间 400ms, CD-ROM XA 能进行多种对话
声 频	8 位数字声频, 8 个音符合成器, MIDI 播放	16 位数字声频, 8 个音符合成器, MIDI 播放
视频显示	640×480, 16 色	640×480, 65536 色
端 子	MIDI I/O 控制杆	MIDI I/O 控制杆

表 1-2 多媒体微机标准——建议值

最低要求	标准 1	标准 2
RAM		8MB
CD-ROM	64KB 板上缓冲区	64KB 板上缓冲区
声 频	—	CD-ROM XA 声频功能支持 IMA 适配器采用的 ADPCM 算法
视 频	640×480, 256 色	在 40%CPU 频带宽的情况下, 每秒传输 1.2MB 像素

另外,数字多媒体技术应用的的关键问题是对图像进行压缩编码和解码。国际标准化组织 (ISO) 和国际电报电话咨询委员会 (CCITT) 两家联合成立了专家组 JPEG (Joint Photographic Experts Group), 在过去几年里一直致力于建立适用于彩色和单色、多灰度连续色调、静态图像的数字图像压缩国际标准。经过方案评选、实验测试,于 1990 年确定了大框架,1991 年作为委员会提案提交投票,表决通过了 ISO/IEC 10916G 标准,即“多灰度静止图像的数字压缩编码”。

为了实现全屏幕运动图像压缩编码及解码,还成立了专家小组 MPEG (Moving Picture Expert Group), 1990 年开始工作。MPEG 在定义协商基础上提交的 MPEG-1 为 ISO CD11172 号标准,于 1992 年被通过。它包括三个部分:MPEG 视频、MPEG 音频和 MPEG 系统。MPEG-1 的平均压缩比为 50:1。

关于电视电话/电视会议 $P \times 64\text{Kb/s}$ 标准 (CCITT H. 216), 早在 1984 年就曾提出。经过研究和改进,1988 年的建议为: $P \times 64\text{Kb/s}$ 中的 P 是一个可变参数,取值为 1~30, $P=1$ 或 2 时,支持四分之一中间格式每秒帧数较低的视频电话;当 $P \geq 6$ 时,可支持通用格式每秒帧数较高的电视会议。

在这一阶段,除了上述的标准化工作外,为了开拓多媒体应用,另一个关键是降低多媒体产品的成本。多媒体产品最大的市场是个人和家庭。开拓家庭市场的先决条件是价格低,

只有低价格的多媒体产品,才能广泛为家庭用户所接受,这在中国显得更为突出。目前来说,已开发的希望用于消费者市场的产品价格与消费者可能接受的价格之间存在很大差距。因此,在我国,多媒体技术的发展目前还处于初期应用和标准化阶段。估计还要持续几年,才能逐步进入多媒体广泛发展与应用阶段。

1.4 多媒体技术的应用

多媒体既然有这样或那样的好处,那么,就用户而言,多媒体能给我们的工作和生活带来哪些具体变化呢?我们相信读者在了解了多媒体的历史、现状、特点之后,心中已有一定想法。我们下面再为读者做一详细的整理。

1.4.1 多媒体将大大改变目前的教育面貌

由于多媒体的引入,将大大改善学习效果,这点已为许多教育方面的专家或学者所肯定。如多媒体系统已使原来电脑辅助教学(CAI)单调乏味的教学画面变得如电视画面一样生动活泼。多媒体使用的领域不只限于学校的学习环境,其他如员工训练、器材使用方法的训练、向顾客介绍使用产品的解说等软件亦均可加入多媒体元素。另外,对具有危险性和花费过高或费时过长的模拟也是多媒体最拿手的工作,多媒体可以模拟制造出其他任何产品无法达到的真实性及亲切感,使学习者获得活生生的教师与真实的学习环境。

诚如 PC Magazine(1992)所言:“多媒体维持电脑‘交互性’的特色,具有接受用户指挥的反应能力,可以根据用户的要求执行不同的工作。而多媒体最吸引人的地方就在于用户和媒体之间具有交互性,用户才是真正的主角,这一点是传统的电子媒体所望尘莫及的。”

1.4.2 多媒体的出现使得销售、咨询、展示和信息管理更形象

多媒体系统可以用来当做说服他人的有力工具。尽管其他媒体也可以做到其中的某些项目,但是多媒体可以集合一种以上的媒体,使说明更生动,更得心应手。多媒体软件可以告诉人们如何使用产品及产品有什么样的功能与好处。以往使用电子试算表、条状图或饼图等均为生硬难懂的数字数据,如今改用多媒体来显示可以让人们感觉更舒适,更容易接受那些重要的观念与意见。生理学家认为人类的每一项感觉器官均为接受消息、改变观念行为的最佳项,多媒体便是利用多种媒体交互或同时多重刺激来传达消息的,因此,其效果也不是单一媒体可以相比拟的。

例如,房地产公司的业务员使用多媒体技术,不用把客户带到现场就可以通过计算机屏幕引导客户“身临其境”地沿着指定的路线,看到整幢楼房的外部 and 内部结构及装修,而且可以动态地删去当天已售出的房子,并给出不同的价格;旅游导引系统可以让游客在出发前,就能对不同地方的景点产生身临其境的感觉。另外,如果在产品的展示上采用触摸屏技术,客户便可能根据自己的爱好,仅需在屏幕上一指,即可沿着自己的思路去选择各种产品。

1.4.3 集书籍、电视、电影、录音带等多种媒体于一身,从而满足人们的更高需求

书籍、电视、电影等均为人类文化生活中不可缺少的资料,但是没有一种如多媒体一样可以如此完整的集合文化元素。许多人认为录像可以比书本或照片来得真实,也有人喜欢听

了再听某个伟人的演说录音,但是不管是书籍也好,电视、电影、录音带也好,这些都只抓住了人类文化生活当中的某一部分或某些层面而已,都不够完整。多媒体除了可以集合以上所提任何一种媒体外,同时还可以加入一些其他媒体所没有的动态数据处理功能。

1.4.4 多媒体营造了一种有参与感的环境,从而为娱乐和游戏带来了巨大的变革

多媒体系统可以创造一种环境,用户身在其间可以充分发挥其各个感官的功能,如文字、图画、动画可以产生视频;声音、音乐、碰撞等可以产生音频;而允许用户操作交互则更发挥了人类眼、耳、手等器官同时学习接受的功能。现今有许多认知心理学家正在研究这种多项显示(Multiple Representation)是否可以增进人类了解问题的能力。因为人的认知形态(Cognitive Style)包括视觉、听觉与触觉(Reissmen, 1965)三种不同的模型(Model),多媒体即可适应不同认知形态的个别差异。使用过多媒体产品的人皆认为多项显示至少比单调显示更容易集中注意力和扩大视野。虽然目前仍无学术研究上充分的证据支持多媒体的功能,但使用过多媒体系统的人皆称“喜欢它”,这已经是一个非常好的开始。

此外,多媒体最引人入胜者,莫过于它能引发人类的想象力。动画可以千变万化,随心所欲地去创作,图像可以使原来肉眼看不清楚的东西放大和旋转。这种媒体与电脑的结合可以使人类的感官与想象力互相配合,产生前所未有的思想空间与创造资源。

目前,一种新的娱乐形式——双向(交互对话式)电影已经出现。观众可以从预先安排的几种情节发展中选择某一种让故事进行下去。而更能使用户感兴趣的“双向电视”也不再是一个遥远的梦想了。日本将于1997年发射BS-4新一代广播卫星,这样本世纪内就能实现这种引人入胜的双向电视了。另外,旅游娱乐界还希望能利用虚拟现实(Virtual Reality)使观众对双向电影有亲临现场之感。

1.5 流行多媒体产品

目前市场上的多媒体产品主要有CD-ROM、声音卡、图像卡、多媒体升级套件、编辑工具、多媒体计算机/工作站、多媒体软件和各种CD-ROM节目,与之配套的还有扫描仪、数字化仪、摄像机等。

1.5.1 CD-ROM 光盘驱动器

CD-ROM驱动器是多媒体电脑的核心部件,其作用同软盘驱动器差不多,其接法也类似软盘驱动器,但它是用来读取光盘的。目前的光盘一般都能识别各种格式。也就是说,你既可用它来播放CD唱片和CD视盘,也可用它来读取各种CD软件。当然,这都需要适当的软件来支持,同时还需要一定的硬件来配合,如声音卡和图像卡等。

1.5.2 声音卡

如果没有声音卡,就无法充分利用多媒体产品。声音卡的主要功能有:播放CD唱盘、进行声音编辑(录制、播放、修改)、文本到声音的转换以及作为FM电子琴等。

1.5.3 图像卡

图像卡主要用于捕捉、数字化、冻结、存储、输出、放大、缩小和调整来自激光视盘机、录像机或摄像机的图像,同时它还可进行一些音频的相关处理。

1.5.4 多媒体升级套件

多媒体升级套件主要由 CD-ROM 驱动器、声音卡和若干 CD-ROM 光盘组成。声音卡通常带有 CD-ROM 接口。

多媒体套件的主要优点是能确保声音卡、CD-ROM 和 CD-ROM 驱动器之间的一致性,这样用户就不必另外再找驱动软件、电缆或其他零部件。

1.5.5 多媒体相关设备

1. 音箱

普通计算机上都有用于报警的喇叭,但是由于它功率太小且频率响应非常有限,所以将它用作仅需发出“嘟嘟”响声的报警器还可以,要用它作为多媒体的扩音器就万万不能了。

用于多媒体的音箱必须为有源音箱,其功率可选 80W、120W 或 200W 均可,这要根据用户对音质的要求而定。

2. 麦克风

用户要想把声音加到自己的文件中,就需要一个麦克风。如果不是用于创作专业的高保真音乐,低价麦克风就能作得很好。

3. 多媒体相关设备

其他多媒体设备还有扫描仪、录像机、摄像机、录音机、激光唱盘机、激光视盘机等。

1.5.6 多媒体 CD-ROM 节目

所谓多媒体 CD-ROM 节目即为存放在 CD-ROM 盘上的应用软件产品。它大致可分为教育、参考、医学、旅游与地理、商业、地图、娱乐等几类,典型产品如《中国邮林》、《美国鸟类》、《康普顿百科全书》、《FINDIT 韦伯斯特词典》、《牛津医学教程》、《家庭医生》、《东方旅游》、《美国商业电话号码簿》以及各种游戏等。

1.6 多媒体技术发展前景

当前国内外多媒体技术的应用主要是基于 CD-ROM 的单机系统。随着多媒体在通信中的发展,特别是高速网络(100Mb/s)成本的下降,使多媒体的普及应用成为可行。只有多媒体通信和网络技术的广泛应用,才能极大地提高人们的工作效率,减少社会交通运输负担。目前美国、欧洲、日本等发达国家已认识到这个问题的重要性,而且把建造高速、宽带、能传送多媒体的网络列入高科技发展计划。

多媒体在实时通信中的同步策略、方法、技术等最终发展趋势是实现分布式多媒体计算机协同系统(Computer Supported Collaborative Work)。一个分布式多媒体计算机系统主要特征是:具有分散性,以区别于当前的集中式单机多媒体系统;具有可操作性,即图、文、声、像一体化的多媒体设备;具有交互性,是指“双向电视”和“双向电影”,以区别普通电视和电影;系统透明性,即对分布式多媒体计算机系统用户而言,整个系统中的资源看上去就像使用一台多媒体个人机一样,但实际上是一个采用网络互连的分布式多媒体工作环境。最后应指出:分布式多媒体技术是一门前沿的研究课题,尚有很多问题需要解决,有若干理论需要突破,而理论上的突破将把多媒体技术应用推向一个划时代的新阶段。

第2章 多媒体基础技术

为便于读者更好地学习后面的内容,我们在本章为读者介绍一些多媒体产品使用及开发需要了解的基本知识。

首先,我们应确认这么一点,信息技术的核心是信息共享问题,如制片厂制作的电影、电视台拍摄的电视片,其目的决不是供他们自己使用,而是供大家观赏的。因此,要解决信息共享问题就必须解决信息表示、信息传送、信息存储和信息处理等问题。

对于多媒体而言,也同样是这个问题。在早期多媒体技术和产品开发时,人们出于经济上的原因,曾仿效家电技术使用录音带录音和录像带录像。虽然利用这种方式也能把各种信息资源汇集在一起,但是由于模拟方式固有的缺点,限制了多媒体技术的进一步发展。这些问题主要表现在:模拟信号在传输过程中易衰减、丢失且抗干扰性较差;模拟信号不适合电脑进行加工、处理,而这点尤其重要。

因此,在多媒体技术中采用数字化技术是必然的。那么,采用这种方法是否就一点问题都没有了?如有又是哪些问题,该如何解决呢?这正是本章要告诉读者的。

2.1 问题的提出

多媒体技术最终要处理的无非是图、文、声、像四类信息,即解决这四类信息的表示、传送、存储和处理问题。下面,我们首先就这三类信息的特点进行叙述。

2.1.1 文字与图形信息

凡是对于电脑有一些了解的用户都知道,在电脑中,信息存储的最小单位是位,即0或1,而一个英文字符在电脑中占用一个字节(即8位),一个汉字则占用一个字(即两个字节)。显然,这对于今天动辄4MB、8MB或16MB的内存空间,以及数百兆字节、甚至几千兆字节的硬盘来讲,要存储文字信息,可谓轻而易举。自然,由于文字信息占用空间较少,因而,对它的处理、传输也不会存在什么问题。对于图形信息而言,由于它们通常比较规则,比如圆、矩形、椭圆等。所以图形信息经过处理后其数据量也很小。

2.1.2 图像信息

我们以一幅中等分辨率的彩色图像为例,假定它的幅面为 640×480 点,每点有 2^{24} 灰度级,那么它应占用 $640 \times 480 \times 24$ 位即921 600字节,也就是说,该幅图像要占用近1MB存储空间。进一步讲,如果它是一幅动态图像的话,则至少要每秒30帧(只有这样,人眼才能产生一种连续感,而不会产生跳跃),而这30帧图像就需要近30兆字节的存储空间。即使目前容量相当大的4000MB硬盘,也不过存储2min动态图像。显然,要存储、传输及处理这么大的数据量,这对电脑及通信而言都是非常难的。因此,要真正把多媒体技术投入应用首先要解决的问题就是图像的压缩编码及解码。如果没有数据压缩技术的进步,多媒体电脑就很难

进入实际应用阶段。

2.1.3 音频信息

对于音频信息而言,也存在同样的问题。以目前的激光盘 CD-DA 声音数据为例,采用 PCM 采样,采样速率为 44.1kHz,16 位两通道立体声,则 600MB 的光盘也只能存放 1h 的数据。

因此,如果要处理上述视频和音频数据,则必须进行数据压缩编码,且其压缩比至少要达到 200:1,否则就无法实现声音和图像的动态播放。也就是说,要使多媒体电脑能综合处理声、文、图、像等信息,就必须解决如下四个问题:

- (1) 视频信息和音频信息的获取问题。
- (2) 视频信息和音频信息的压缩和解压缩问题。
- (3) 视频信息和音频信息的实时处理和特技。
- (4) 视频信息的显示和音频信息的立体声输出。

这是建立多媒体电脑硬件支撑平台必须具备的功能,尤其是视频和音频信息的输入和输出、压缩和解压缩功能,一定要用硬件实现。

2.2 多媒体电脑产生声音的三种主要方法

下面我们就首先来谈谈在多媒体电脑中声音是如何处理的。大家都知道,声音信号是携带信息的极其重要的媒体,是多媒体技术和多媒体产品开发中的一个重要内容。声音的种类繁多,如人的话音、乐器声、动物发出的声音、机器产生的声音以及自然界的雷声、风声、雨声、闪电声等。多媒体电脑产生声音的方法主要有三种:采样及重放、CD 唱片重放、通过 MIDI 驱动内置或外置的合成器。

2.2.1 音频信息采样频率及位长

数字音响是电脑技术与音响相结合的产物。大家知道,过去音响都是模拟型的,话筒把声音变成模拟的电流,再送到电路进行放大及处理,然后送到扬声器放音或变成相应的磁场记录到磁带上。这种方法虽然比较简单,但存在许多难以克服的缺点,比如在电路中产生的各种失真畸变以至信号丢失,混进各种难以分离或消除的噪音等。因此,就必须对各种模拟信号进行转化,我们称这个转化过程为采样,即每隔一定时间对信号进行采集。

1. 采样频率

我们知道声音本是一种波,如果以极小的间隔把波形变成一系列数字,从技术上讲就是从模拟量到数字量的转换,简称 A/D 转换,也就是对声音进行采样。到重放的时候再倒过来处理,作 D/A 转换,把数字再变回波形。这种处理方法当然比模拟式的要复杂得多,但它能消除模拟音响的大部分缺点,当今的电脑处理模拟量一般都用采样方法,随着电脑技术的发展,采样技术日益完善。

在采样的时候,采样点之间的间隔越小,采出的样就越细腻逼真。间隔越小也就意味着采样频率越高。但是由于电脑的工作速度总是有限的,电脑的存储容量也不可能无限扩充,

因此,采样频率不可能无限地高上去。一般认为,采样频率高于被采样波形最高频率一倍以上就可以获得满意的质量。人耳的听觉上限大约在 20kHz,所以 40kHz 以上的采样频率已足可使人满意了,CD 唱片的采样频率就是 44.1kHz。多媒体电脑常用的采样频率有 11.025kHz、22.05kHz 和 44.1kHz 等几种,用户可以根据情况选用。

2. 采样位长

决定采样质量的还有一个参数,就是采样的比特数,也称位长,指每次采样用几个 0 或 1。如果说采样频率是水平坐标,比特数就是垂直坐标,它决定采样的动态范围。8bit 的采样能描绘 $2^8=256$ 种变化,16bit 的采样就能描绘 $2^{16}=65536$ 种变化。有人用画画的颜色来形容它,用 65536 种颜色画的画更丰富多彩。多媒体电脑声音采样的比特数有 8bit 和 16bit 两种。

采样时还可以选择单声道或立体声。数字采样处理立体声实际就是把左右声道的信号分别记录或重放。

数字化采样要占用很大的存储空间,采样频率越高、比特数越多,则声音质量越好,但占用的空间也更大,使用立体声时占用空间比单声道时占用空间要翻一番。各种情况下每分钟占用存储空间的数量如表 2-1 所示。

表 2-1 音频采样及占用存储空间对照表(一分钟)

单声道或立体声		单声道		立体声	
采样频率	采样位长	16bit	8bit	16bit	8bit
44.1kHz		5.28MB	2.64MB	10.56MB	5.28MB
22.05kHz		2.54MB	1.27MB	5.28MB	2.54MB
11.025kHz		1.27MB	0.635MB	2.54MB	1.27MB

采样数据像其他文件一样存在硬盘或光盘中,用 .WAV 作扩展名。

2.2.2 CD 唱片

常见的 CD 唱片也是用数字采样技术制作的,只是不存成 .WAV 文件,而把 1 和 0 变成微小的长短凹坑直接刻写在盘片上,重放时用激光读出这些数字,再通过 D/A 转换变成模拟信号。多媒体电脑的 CD-ROM 驱动器可以直接放送 CD 唱片。

2.2.3 MIDI 合成

MIDI 产生声音的方法与前面讲的采样有很大的不同。MIDI 文件以 .MID 为扩展名,它并没有记录任何声音,实际上只是发送给合成器的一系列指令。例如命令某个 MIDI 通道的某个音以某种力度开始发声,某个 MIDI 通道的音色改变为第几号音色等。具体出声音的硬件是合成器或音源(Tone Generator),它们自己又有一套给声音编程的参数,一般相当复杂。由于 MIDI 文件不含任何声音成分,所以通常比采样声音文件小得多。编写 MIDI 音乐要用专门的软件,它们一般都具有多种输入方式、丰富的编辑命令和友好的接口。

2.3 音频数据压缩和编码

数字信号可以存到磁盘上用于编辑和重放。这种应用虽然很容易想到,但也只是在最近几年,由于微型计算机有了足够高的处理速度,才使这种应用变成了现实。

音频数据的处理过程可分成两大类:实时处理和非实时处理。实时处理是指处理过程能够用与硬件转换相同的速度去接受和处理样本数据。这就意味着处理过程不能被中断,只能处理连续的数据流。当没有足够的能力处理连续的数据流时,就要用非实时处理。

例如,数字激光唱盘上的立体声音频信号,每个通道每秒有 44 100 个样本要实时处理。如果用由 40 个乘法器构成的数字滤波器,那么,对每个样本就要做 40 次乘法运算和 40 次加法运算。两个通道的乘加运算量就需要 $2(\text{通道}) \times 44\,100(\text{样本/s}) \times (40 \times 2)(\text{乘加运算/样本}) = 7.056 \times 10^6 \text{ 次乘加运算/s}$ 。当今的数字信号处理器 DSP(Digital Signal Processor)已经具有这样高的处理速度来处理这种实时信号。

数字信号处理除了需要强大的处理能力之外,数字信号本身也占据可观的存储容量。一小时的超级高保真立体声音乐,就要占据 $44100(\text{样本/s} \cdot \text{通道}) \times 2(\text{Byte/样本}) \times 2(\text{通道}) \times 3600(\text{s}) = 6.350 \times 10^8 \text{ Byte}$,即近 600MB 存储空间。所以,必须对数据进行压缩,下面我们就来谈谈这方面的情况。

2.3.1 数据压缩

数据压缩是数字信号处理中最基本的一种处理方法。压缩的目的也很清楚,就是要满足存储容量和传输速度的要求,而付出的代价就是要进行大量的计算。压缩有两种基本的方法,即无损压缩和有损压缩,有损和无损主要是针对原始信号而言的。

无损压缩是用在要求重构的信号与原始信号完全一致的场合。其最典型的例子就是磁盘文件的压缩,这种压缩最基本的要求就是不能丢失任何信息,要绝对忠实于原文件。根据目前的技术水平,这种压缩方法可以把数据压缩到原来的 1/2 到 1/4,即其压缩比为 2:1 或 4:1。

有损压缩是用在重构的信号不一定要完全忠实于原始信号的场合,即它容许有一定程度的失真。如语音和电视图像信号的压缩就可以采用有损压缩。这种压缩比相当高,如静态图像的压缩比可以达到 10:1 到 100:1,甚至更高,这主要取决于用户对重构出的图像质量的要求,以及所采用的压缩算法。典型的有损压缩算法主要有 ADPCM(Adaptive Differential Pulse Code Modulation,自适应差分脉冲编码调制)、CD-XA(Compact Disc Extended Architecture)、子带编码(Subband Coding)、JPEG(Joint Pictures Experts Group)、MPEG(Moving Pictures Expert Group)等,其中 ADPCM、CD-XA 和子带编码主要用于压缩音频信号,JPEG 压缩算法用于静态图像压缩,MPEG 压缩既用于电视图像又用于声音。

2.3.2 几种压缩编码方法

1. 脉冲编码调制(PCM)

脉冲编码调制 PCM(Pulse Code Modulation)是概念上最简单、理论上最完善的编码系

统,也是最早研制成功、使用最广泛的编码系统,但它也是数据量最大的编码系统。

PCM 编码的原理实际上就是把连续模拟信号变成离散的幅度信号,再把离散的幅度信号变成数字的离散信号。典型 PCM 编码如表 2-2 所示。

表 2-2 典型 PCM 编码实例

信号名称		信号频率	采样率	位/样本	数据传输率
电话语音		3.4kHz	8kHz	8	64Kb/s
音乐(超级 HiFi)		20kHz	44.1kHz	16	705.6Kb/s
彩色电视	亮度	6MHz	13.5MHz	8	(13.5+6.75+6.75) ×8b=216Mb/s
	两个色差		6.75MHz	8	

如果采用相等的量化间隔对采样得到的信号进行量化,那么这种量化称为均匀量化,这种量化方法的缺点是占用存储空间大。如果根据样本值的幅度大小去改变量化间隔,那么就可以节约一定数量的存储空间。非均匀量化的实质是,编码时对信号进行压缩,还原时对代码进行扩展,而压缩和扩展的特点可以根据应用要求加以选择。目前美国和日本通常采用的是 μ 律压扩标准,中国和欧洲采用的则为 A 律压扩标准。

2. 增量调制(DM)

DM 编码是 PCM 编码的一种变型。PCM 编码是对每个采样信号的整个幅度进行量化编码,因此它具有对任何随机波形进行编码的能力。DM 编码则是对实际的采样信号与预测的采样信号之差的极性进行编码,将极性变换成两种可能的取值之一。如两个信号之差的极性为正,可用 1 表示,极性为负,则用 0 表示。

这种编码的特点是编码比较简单,但是也出现了两个问题:一是会出现“斜率过载(Slop Overload)”,二是产生“粒状噪声(Granular Noise)”。所谓斜率过载是指增量调制器的输出不能保持跟踪输入信号的快速变化。粒状噪声则是指当输入信号与预测信号的差值为零时,出现随机交替变化的“0”和“1”,这种情况主要出现在输入信号变化较慢的时候。为了解决这个问题,人们又提出了自适应增量调制 ADM(Adaptive Delta Modulation)。

3. 自适应增量调制(ADM)

这种编码方法的核心是,假定增量调节器的输出为 +1 和 -1,每当输出不变时量化阶增大 50%,每当输出值改变时,量化阶减小 50%。利用这种方法可将斜率过载和粒状噪声减少到最低。

4. 自适应差分脉冲增量编码调制(ADPCM)

数字电话中,典型地使用采样率为 8kHz、每个样本位为 8bit、数据传输率为 64Kb/s 的 A 律或 μ 律 PCM 编码,这就是 CCITT 推荐使用的 G.711 标准:话音频率脉冲编码调制(Pulse Code Modulation of Voice Frequences)。随着数字电话和数据通信容量的日益增长,人们希望在不明显降低话音质量的前提下进一步提高传输效率,为解决这个问题,就要对话音信号进一步压缩。方法之一就是采用自适应差分脉冲编码调制(ADPCM)。

ADPCM 是利用样本之间的高度相关性来压缩数据的一种波形编码技术。CCITT 为此