

第3章 多媒体技术基础



1. 了解多媒体基本概念。
2. 掌握多媒体计算机系统结构。
3. 了解多媒体压缩技术。
4. 了解多媒体信息数字化。

自 20 世纪 80 年代以来,随着电子技术和大规模集成电路技术的发展,计算机技术、广播电视技术和通信网络技术这三大领域相互渗透融合、相互促进,从而形成了一门新的技术即多媒体技术。多媒体技术能使计算机具有综合处理声音、文字、图形、图像和视频等信息的能力,因此成为当前最受信息领域关注的热点之一。

3.1 多媒体技术基础

3.1.1 多媒体概述

1. 基本概念

(1) 媒体。在日常生活中,被称为媒体的东西有许多,如蜜蜂是传播花粉的媒体、苍蝇是传播病菌的媒体。但准确地说,这些所谓的“媒体”是传播媒体,并非我们所说的多媒体中的“媒体”,因为这些传播媒体传播的都是某种物质实体,而文字、声音、图像、图形这些都不是物质实体,它们只是客观事物某种属性的表面特征,是一种信息表示方式。在计算机和通信领域所说的媒体(Media)就是人与人之间实现信息交流的中介,简单地说,就是信息的载体,也称为媒介,并不是一般的媒介和媒质。

(2) 多媒体。多媒体的英文单词是 Multimedia,它由 media 和 multi 两部分组成。一般理解为多种媒体的综合,也就是文本、声音、图像、图形、动画及视频等多种媒体的直接结合和综合使用。通过这种方式,我们就可以在听到优美动听的音乐的同时,看到精致唯美的图片,欣赏美轮美奂的影音动画。

(3) 多媒体技术。多媒体技术不是各种信息媒体的简单复合,它是一种把文本、图形、图像、动画和声音等多种信息类型综合在一起,并通过计算机进行综合处理和控制在支持完成一系列交互式操作的信息技术。多媒体代表数字控制和数字媒体的汇合,多媒体技术的发展拓宽了计算机的使用领域,其广泛应用于工业生产管理、学校教育、公共信息咨询、商业广告、军事指挥与训练、建筑规划设计,甚至家庭生活与娱乐等领域。

由于多媒体系统需要将不同的媒体数据表示成统一的结构码流,然后对其进行变换、重



组和分析处理,以进行进一步的存储、传送、输出和交互控制。多媒体技术的应用突破了传统的单靠键盘和显示器以字符的形式交流信息的模式。多媒体技术发展到今天,跟许多技术的进步,如数据压缩技术、大规模集成电路(VLSI)制造技术、大容量的光盘存储器(CD-ROM)、实时多任务操作系统等紧密相关。因为这些技术取得了突破性的进展,多媒体技术才得以迅速的发展,而成为像今天这样具有强大的处理声音、文字、图像等媒体信息的能力的高科技技术。所以,多媒体技术可以说是包含了当今计算机领域最新的硬件和软件技术,它将不同性质的设备和信息媒体集成一个整体,并以计算机为中心综合地处理各种信息。现在所说的多媒体,通常并不单单指多媒体信息本身,主要是指处理和应用它的一套软、硬件技术。因此,通常所说的“多媒体”只是多媒体技术的同义词。

多媒体技术有以下几个主要特点。

- 集成性:能够对信息进行多通道统一获取、存储、组织与合成。
- 控制性:多媒体技术是以计算机为中心,综合处理和控制在多媒体信息,并按人的要求以多种媒体形式表现出来,同时作用于人的多种感官。
- 交互性:交互性是多媒体应用有别于传统信息交流媒体的主要特点之一。传统信息交流媒体只能单向地、被动地传播信息,而多媒体技术则可以实现人对信息的主动选择和控制。
- 非线性:多媒体技术的非线性特点将改变人们传统循序性的读写模式。以往人们读写方式大都采用章、节、页的框架,循序渐进地获取知识,而多媒体技术将借助超文本链接(Hyper Text Link)的方法,把内容以一种更灵活、更具变化的方式呈现给读者。
- 实时性:当用户给出操作命令时,相应的多媒体信息都能够得到实时控制。
- 信息使用的方便性:用户可以按照自己的需要、兴趣、任务要求、偏爱和认知特点来使用信息,任取图、文、声等信息表现形式。
- 信息结构的动态性:“多媒体是一部永远读不完的书”,用户可以按照自己的目的和认知特征重新组织信息,增加、删除或修改节点,重新建立链。

2. 多媒体技术的应用与发展

(1) 多媒体技术的发展。多媒体技术的发展是非常迅速的,让我们来看看这中间几个具有代表性的时刻:

1984年,美国苹果(Apple)公司开创了用计算机进行图像处理的先河,在世界上首次使用位图(Bitmap)概念对图像进行了描述,从而实现了图像进行简单的处理、存储以及传送等。苹果公司对图像进行处理的计算机是该公司自行研制和开发的苹果计算机(Apple),其操作系统名为Macintosh,所以也有人把苹果计算机直接叫做Macintosh计算机。在当时,Macintosh操作系统首次采用了先进的图形用户界面,体现了全新的Window(窗口)概念和Icon(图标)程序设计理念,并且建立了新型的图形化人机界面标准(比微软要早!)

1985年,美国Commodore公司将世界上首台多媒体计算机系统展现在世人面前,该计算机系统被命名为Amiga。并在随后的Comdex'89展示会上,展示了该公司研制的多媒体计算机系统Amiga的完整系列。

同年,计算机硬件技术有了较大的突破,为解决大容量存储的问题,激光只读存储器(CD-ROM)问世,为多媒体数据的存储和处理提供了理想的条件,并对计算机多媒体技术的发展起到了决定性的推动作用。在这一时期,CDDA(Compact Disk Digital Audio)技术也已



经趋于成熟，使计算机具备了处理和播放高质量数字音响的能力。这样，在计算机的应用领域中又多了一种媒体形式，即音乐处理。

1986年3月，荷兰 Philips（飞利浦）公司和日本 SONY（索尼）公司共同制定了交互式激光盘系统标准（Compact Disc-Interactive, CD-I），使多媒体信息的存储规范化和标准化。CD-I 标准允许在一片直径为 5in 的激光盘上存储 650MB 的数字信息量。1987年3月，RCA 公司制定了 DVI（Digital Video Interactive）技术标准，该技术标准在交互式视频技术方面进行了规范化和标准化，使计算机能够利用激光盘以 DVI 标准存储静止图像和活动图像，并能存储声音等多种信息模式。DVI 标准的问世，使计算机处理多媒体信息具备了统一的技术标准。同年，美国 Apple（苹果）公司开发了 HyperCard（超级卡），该卡安装在苹果计算机中，使该型计算机具备了快速、稳定的处理多媒体信息的能力。

1990年11月，美国 Microsoft（微软）公司和包括荷兰 Philips（飞利浦）公司在内的一些计算机技术公司成立了“多媒体个人计算机市场协会（Multimedia PC Marketing Council）”。该协会的主要任务是对计算机的多媒体技术进行规范化管理和制定相应的标准。该协会制定了多媒体计算机的 MPC 标准。该标准将对计算机增加多媒体功能所需的软、硬件规定了最低标准的规范、量化指标以及多媒体的升级规范等。1991年，多媒体个人计算机市场协会提出 MPC1 标准。从此，全球计算机业界共同遵守该标准所规定的各项内容，促进了 MPC 的标准化和生产销售，使多媒体个人计算机成为一种新的流行趋势。

1993年5月，多媒体个人计算机市场协会公布了 MPC2 标准。该标准根据硬件和软件的迅猛发展状况做了较大的调整和修改，尤其对声音、图像、视频和动画的播放、Photo CD 做了新的规定。此后，多媒体个人计算机市场协会演变成多媒体个人计算机工作组（Multimedia PC Working Group）。

1995年6月，多媒体个人计算机工作组公布了 MPC3 标准。该标准为适合多媒体个人计算机的发展，又提高了软件、硬件的技术指标。更为重要的是，MPC3 标准制定了视频压缩技术的技术指标，使视频播放技术更加成熟和规范化，并且指定了采用全屏幕播放、使用软件进行视频数据解压缩等项技术标准。

同年，由美国 Microsoft（微软）公司开发的功能强大的 Windows 95 操作系统问世，使多媒体计算机的用户界面更容易操作，功能更为强劲。随着视频音频压缩技术日趋成熟，高速的奔腾系列 CPU 开始武装个人计算机，个人计算机市场已经占据主导地位，多媒体技术得到了蓬勃发展。国际互联网络（Internet）的兴起，也促进了多媒体技术的发展，更新更高的 MPC 标准相继问世。

目前，多媒体技术的发展趋势是逐渐把计算机技术、通信技术和大众传播技术融合在一起，建立更广泛意义上的多媒体平台，实现更深层次的技术支持和应用，使之与人类文明水乳交融。

（2）多媒体技术的应用。近年来，多媒体技术得到了迅速的发展，应用领域也不断扩大，这是社会需求与科学技术发展相结合的结果。多媒体技术的发展也带动了其他一些技术的应用，这些技术同样又促进了多媒体技术的发展。多媒体技术为人们提供了多种表达和交流信息的方式，正在逐渐或已经进入政府部门、军队、学校、科研机构、公司企业以至家庭，并将广泛应用于管理、教育、培训、公共服务、广告、文艺、出版等领域。

多媒体技术的发展使一些原来相对独立发展的产业和行业（如计算机、电视、通信、出



版和娱乐等)开始相互渗透和结合,从而产生了一些全新的产业和应用领域。多媒体技术与多媒体系统的应用多种多样,丰富多彩。从科学研究、商业管理、工业生产一直到家庭娱乐,几乎涉及人类社会生产、生活的各个领域,并且正在不断发展和开拓新的应用领域。目前的多媒体系统大多数还是单机使用的,但实际应用已经提出了把多媒体技术与通信、网络相结合的需求,这就是所谓的“分布式”多媒体技术,它结合了计算机的交互性、通信的分布性和电视的真实性,因此将向人们提供全新的信息服务。下面将从几个方面介绍一些多媒体技术的应用领域,并且举出几个具体的应用例子。

①传统教育行业中的应用。目前的教科书出版商已经开始转入基于 CD-ROM 的教科书出版。电子教科书的出现,使人们可以通过计算机辅助,而得到伴随“交互式指导”的形象化教材。打开计算机,学生们面对的是一本本“活”书。

多媒体辅助教学与多媒体教学软件在课堂教学中的应用越来越广泛。多媒体辅助教学,是指利用计算机帮助教师进行教学或用计算机进行教学,是计算机科学、教育学、心理学等多门学科交叉形成的一门综合性的新兴学科。它既是计算机的一个应用领域,又代表一种新的教育技术和教育方式。多媒体教学软件是一种根据教学目标设计的表现特定的教学内容,反映一定教学策略的计算机教学系统。它可以用于存储、传递和处理教学信息,能让学生进行交互操作,并对学生的学习作出评价的教学媒体。

多媒体技术也可以用于职工培训甚至军事训练,比如通过图、文、声并茂的形式对工人和销售人员培训,可以使他们更快地掌握作业技巧和种种作业规范。一些大型工厂甚至出版通用于全厂各个岗位工人和行政管理人员的多媒体统一行动守则,这样做将使工厂的管理更有章法,也不会因为人为因素使某些作业达不到标准规范的作业要求。

②普通大众生活中的应用。多媒体技术同样也广泛作用于普通大众的生活中,电子出版物的出现,正悄然改变人们的阅读习惯。而电子出版物正是多媒体技术在出版行业中应用的产物。多媒体的应用同样也使图书馆的工作发生了巨大的变化。多媒体技术的应用使得图书馆典藏将由传统的以纸质书刊为主逐渐向以光盘或其他高密度介质发行的电子图书等多媒体馆藏过渡;信息检索系统不仅具有多媒体形式的信息,而且以非线性的结构组织信息,并为读者提供了友好的使用界面;为传统的图书馆信息服务开辟了新的天地,典藏媒体的多样化使信息服务多样化成为可能。

城市中各类无人信息咨询站悄然出现,也是多媒体技术应用的典型。这些咨询站的出现,为人们出行、购物、信息查询等提供了方便。

在商业运营中,广告和销售服务是成功的重要条件。形象、生动的多媒体技术在这方面可以大有作为,主要体现在商品展示、多媒体产品操作手册、销售演示等方面。

在医学教育领域中,交互式多媒体系统可以帮助医生和其他医务人员改进诊断和治疗技能。多媒体技术可以提供非常逼真的病人症状模拟,供医生练习处理和治理病人的疾患,而无需麻烦真正的病人。处在现代医疗卫生中心的医生可以通过多媒体通信网为边远地区的病人提供医疗服务。

多媒体技术将改变我们未来的家庭生活。信息技术领域的多媒体技术和信息高速公路等将丰富现在的家庭生活,集电视、电话、录像、计算机等功能于一体的多媒体技术已日趋成熟,用多媒体计算机便可以收看电视、录像,可打电话,发传真。多媒体技术在家庭中的应用将使人们在家上班成为现实。



随着多媒体通信和视频图像传输数字化技术的发展，及个人计算机和网络的结合，使多媒体会议系统成为多媒体技术最重要的一个应用领域，其效果和使用方便程度比传统的电话会议优越得多。

③虚拟现实技术（VR）。VR 是 Virtual Reality 的缩写，中文的意思就是虚拟现实，概念是在 20 世纪 80 年代初提出来的，其具体是指借助计算机及最新传感器技术创造的一种崭新的人机交互手段。1992 年美国国家科学基金资助的交互式系统项目工作组的报告中对 VR 提出了较系统的论述，并确定和建议了未来虚拟现实环境领域的研究方向。可以认为，虚拟现实技术综合了计算机图形技术、计算机仿真技术、传感器技术、显示技术等多种科学技术，它在多维信息空间上创建一个虚拟信息环境，能使用户具有身临其境的沉浸感，具有与环境完善的交互作用能力，并有助于启发构思。所以说，沉浸、交互、构想是 VR 环境系统的三个基本特性。虚拟技术的核心是建模与仿真。当前，VR 已不仅仅被关注于计算机图像领域，它已涉及更广的领域，如电视会议、网络技术和分布计算技术，并向分布式虚拟现实发展。

近年来，虚拟现实技术的应用已大步走进工业、建筑设计、教育培训、文化娱乐等方面，它正在逐渐改变着我们的生活。在科技开发上，虚拟现实可缩短开发周期，减少费用。例如克莱斯勒公司 1998 年初便利用虚拟现实技术，在设计某两种新型车上取得突破，首次使设计的新车直接从计算机屏幕投入生产线，也就是说完全省略了中间的试生产。更重要的是，虚拟现实技术已经和理论分析、科学实验一起，成为人类探索客观世界规律的三大手段。用它来设计新材料，可以预先了解成分改变对材料性能的影响。在材料还没有制造出来之前便知道用这种材料制造出来的零件在不同受力情况下是如何被损坏的。在商业上，虚拟现实常被用于推销，例如建筑工程投标时，把设计的方案用虚拟现实技术表现出来，便可把业主带入未来的建筑物里参观，如门的高度、窗户朝向、采光多少、屋内装饰等，都可以感同身受。它同样可用于旅游景点以及功能众多、用途多样的商品推销。因为用虚拟现实技术展现这类商品的魅力，比单用文字或图片宣传更加有吸引力。在医疗上，虚拟现实技术的应用大致上有两类。一是虚拟人体，也就是数字化人体，有了这样的人体模型，医生更容易了解人体的构造和功能。另一是虚拟手术系统，可用于指导手术的进行。在军事上，利用虚拟现实技术模拟战争过程已成为省时、省力、节约经费的研究战争、培训指挥员的方法。随着虚拟现实技术的发展，模拟实验环境，在不进行核试验的情况下，也能不断改进核武器；在教育上，主要有三方面应用。一是虚拟校园，模拟立体的校园实景。虽然目前大多数虚拟校园仅仅实现了校园场景的浏览功能，但虚拟现实技术提供的立体的浏览方式，全新的媒体表现形式都具有非常鲜明的特点。二是虚拟教学，利用简易型虚拟现实技术表现某些系统（自然的、物理的、社会的）的结构和动态，为学生提供一种可供体验和观测的环境。填补了一些课堂中无法实现的实验环境的缺陷。三是虚拟培训，虚拟现实技术的沉浸性和交互性，使学生能够在虚拟学习环境中扮演一个角色，全身心地投入学习，这非常有利于学生的技能训练，如西南交通大学开发的 TDS-JD 机车驾驶模拟装置可模拟列车起动、运行、调速及停车全过程，可向司机反馈列车运行过程中的重要信息，使学生无需到实地就可接受技能训练。

④流媒体技术。流媒体（Streaming Media）技术是当前十分流行的多媒体技术，源于美国，又称流式媒体，是一种新的媒体传送方式。所谓流媒体技术就是把连续的影像和声音信息经过压缩处理后放上网站服务器，让用户一边下载一边观看、收听，而不要等整个压缩文件下载完全到自己的计算机上才可以观看的网络传输技术。该技术先在使用者端的计算机上



创建一个缓冲区,在播放前预先下一段数据作为缓冲,在网路实际连线速度小于播放所耗的速度时,播放程序就会取用一小段缓冲区内的数据,这样可以避免播放的中断,也使得播放品质得以保证。

在实际生活中,流媒体技术广泛应用于娱乐、电子商务、远程培训、视频会议、远程教育、远程医疗、电视台重大节目直播等互联网信息服务的方方面面,它的应用给网络信息交流带来革命性的变化,突破了网站媒体一直以来以文字、图片方式传递信息的局限性。在教育行业中,流媒体技术帮助远程教育消除了距离的阻碍,将需要传递的各种信息,如视频、音频、文本、图片等从教师端传递到远程的学生端,使身在不同地域的学生都能享受到同等的教育。除去实时教学以外,使用流媒体中的 VOD(视频点播)技术辅助教学,达到因材施教、交互式的教学目的,给传统的教学模式带来了很大的冲击;在广电行业中,流媒体直播技术使实时广播电视节目网络化得以实现。随着宽带网的不断普及和流媒体技术的不断改进,互联网直播各种体育赛事、重大事件、电视节目等已毫无压力。在航空探测中,从数据的及时反馈、图像的按时传回到太空行走过程的电视直播等,越来越多的项目中需要依赖流媒体技术。最近,运用 VIEWGOOD WebTable 系统的“远望 5 号”测量船,在海面上监测“嫦娥一号”卫星的运行轨道,给“嫦娥一号”下达指令,指示“嫦娥”不断变轨,使其按照固定轨道顺利运行。WebTable 系统支持视频和桌面多路实时同步直播,独创 KeyBuffer 技术,使广域网时延最短只有 0.1s,满足庞大关注人群的收视需求。WebTable 系统清晰度高,实时性强,并发能力与抗干扰能力更是首屈一指,成功助“嫦娥一号”顺利进行探月工程,为流媒体技术在航空探测中的应用树立了里程碑。

3. 媒体的类型

按照国际电信联盟的定义,媒体有以下 5 种:感觉媒体、表示媒体、显示媒体、存储媒体和传输媒体。感觉媒体指的是用户接触信息的感觉形式,如视觉、听觉和触觉等。表示媒体则指的是信息的表示和表现形式,如图形、声音和视频等。显示媒体是表现和获取信息的物理设备,如显示器、打印机、扬声器、键盘和摄像机等。存储媒体是存储数据的物理设备,如磁盘、光盘、硬盘等。传输媒体是传输数据的物理设备,如电缆、光缆、电磁波等。一般说来,如不特别强调,我们所说的媒体指的就是表示媒体,因为作为多媒体系统来说,处理的主要还是各种各样媒体的表示和表现,其他的媒体类型也都要在多媒体系统中研究,但方法比较单一。

主要的表示媒体有以下几种:

(1) 视觉类媒体。

①位图图像:将所观察到的图像按行、列进行数字化,将图像的每一点都数字化为一个值,所有的这些值就组成了位图图像。位图图像是所有视觉表示方法的基础。

②图形:图形是图像的抽象,它反映了图像上的关键特征,例如点、线、面等。图形的表示不直接描述图像的每一点,而是描述产生这些点的过程和方法,即用矢量来表示。

③符号:符号中也包括文字和文本。由于符号是人们创造出来表示某种含义的,所以它与使用者的知识水平有关,是比图形更高一级的抽象。必须具有特定的知识,才能解释特定的符号,才能解释特定的文本(如语言),符号的表示是用特定值来表示的。

④视频:视频又称为动态图像,是一组图像按照时间的有序连续表现。视频的表示与图像序列、时间有关。

⑤动画:动画也是动态图像的一种。与视频不同的是,动画是计算机制作出来的图像或



图形，而不像视频采用直接采集的真实图像。动画包括二维动画、三维动画、真实感三维动画等多种形式。

⑥其他：其他类型的视觉媒体形式，如用符号表示的数值、用图形表示的某种数据曲线、数据库的关系数据等。

（2）听觉类媒体。

①波形声音：就是自然界中所有的声音，是声音数字化的基础。

②语音：语音也可以表示为波形声音，但波形声音表示不出语言、语音学的内涵。语音是对讲话声音的一次抽象。

③音乐：音乐与语音相比更规范一些，是符号化了的声音。但音乐不能对所有的声音进行符号化。乐谱是符号化声音的符号组，表示比单个符号更复杂的声音信息内容。

（3）触觉类媒体。

①指点：包括间接指点和直接指点。通过指点可以确定对象的位置、大小、方向和方位，执行特定的过程和相应的操作。

②位置跟踪：为了与系统交互，系统必须了解参与者的身体动作，包括头、眼睛、手、四肢等部位的位置与运动方向，系统将这些位置与运动的数据转变为特定的模式，对相应的动作进行表示。

③力反馈与运动反馈：这与位置跟踪正好相反，是由系统向参与者反馈运动及力的信息，如触觉刺激、反作用力（如推门时的门重感觉）、运动感觉（如摇晃、振动）及温度等环境信息。这些媒体信息的表现必须借助于一定的电子、机械的伺服机构才能实现。

4. 常见的多媒体数据文件格式

多媒体数据、程序均以文件的形式存储在计算机中。常见的多媒体数据文件格式有：

文档文件：DOC、XLS、PPT、WPS、TXT 等。

音频文件：WAV、MP3、MP4、RA、RM、M3U、MPA、VOC、MID 等。

图形图像文件：BMP、GIF、JPG、TIF、PSD、DRW 等。

影像文件：MPG、MOV、AVI、RM 等。

动画文件：GIF、SWF、FLI、DL 等。

3.1.2 多媒体计算机系统

多媒体系统却不只是单一的一门技术，而是多种信息技术的集成，是把多种技术综合应用到一个计算机系统中，实现信息输入、信息处理、信息输出等多种功能。一个完整的多媒体系统由多媒体硬件和多媒体软件两部分构成的。

多媒体硬件系统主要包括计算机硬件、声音/视频处理器、多种媒体输入/输出设备及信号转换装置、通信传输设备及接口装置等。其中，最重要的是根据多媒体技术标准而研制生成的多媒体信息处理芯片和板卡、光盘驱动器等。多媒体软件系统主要包括多媒体操作系统、媒体处理系统工具、用户应用软件等。

1. 多媒体计算机

多媒体计算机一般指多媒体个人计算机（Multimedia Personal Computer, MPC），是一种能够对文本、声音、图像、视频等多种媒体信息进行综合处理，并建立逻辑关系，使之成为一个交互式系统的计算机。1985 年出现了第一台多媒体计算机，其主要功能是指可以把音频视



频、图形图像和计算机交互式控制结合起来,进行综合的处理。

2. 多媒体计算机的硬件设备

CD-ROM、声卡和视频卡是配置多媒体计算机的关键设备。

(1) CD-ROM 驱动器与光盘。CD-ROM 驱动器又称光驱,如图 3-1、图 3-2 所示,是多媒体计算机的基本配置之一。激光头是光驱的心脏,也是最精密的部分。它主要负责数据的读取工作,因此在清理光驱内部的时候要格外小心。CD-ROM 驱动器读取速度是指光存储产品在读取 CD-ROM 光盘时,所能达到最大光驱倍速。因为是针对 CD-ROM 光盘,因此该速度是以 CD-ROM 倍速来标称,不是采用 DVD-ROM 的倍速标称。目前 CD-ROM 所能达到的最大 CD 读取速度是 56 倍速;DVD-ROM 读取 CD-ROM 速度方面要略低一点,达到 52 倍速的产品还比较少,大部分为 48 倍速;COMBO 产品基本都达到了 52 倍速。对于 50 倍速的 CD-ROM 驱动器理论上的数据传输率应为: $150 \times 50 = 7500 \text{KB/s}$ 。

由于软盘的容量小,光盘凭借大容量得以广泛使用。我们听的 CD 是一种光盘,看的 VCD 也是一种光盘。CD 光盘的最大容量大约是 650MB, DVD 盘片单面容量 4.7GB (双面容量 8.5GB), 蓝光的容量则比较大,其中 HD DVD 单面单层容量 15GB、双层容量 30GB; BD 单面单层容量 25GB、双面容量 50GB。



图 3-1 光驱

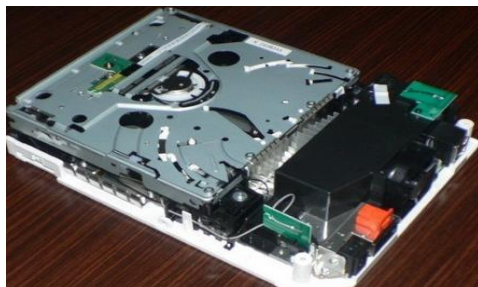


图 3-2 光驱内部结构

(2) 声卡。声卡 (Sound Card) 也叫音频卡 (港台地区称之为声效卡), 如图 3-3 所示, 声卡是多媒体技术中最基本的组成部分, 是实现声波/数字信号相互转换的一种硬件设备。声卡的基本功能是把来自话筒、磁带、光盘的原始声音信号加以转换, 输出到耳机、扬声器、扩音机、录音机等声响设备, 或通过音乐设备数字接口 (MIDI) 使乐器发出美妙的声音。



图 3-3 声卡

声卡是计算机进行声音处理的适配器。它有三个基本功能: 一是音乐合成发音功能; 二是混音器 (Mixer) 功能和数字声音效果处理器 (DSP) 功能; 三是模拟声音信号的输入和输



出功能。声卡处理的声音信息在计算机中以文件的形式存储。声卡工作应有相应的软件支持，包括驱动程序、混频程序（mixer）和 CD 播放程序等。

声卡可以把来自话筒、收录音机、激光唱机等设备的语音、音乐等变成数字信号交给计算机处理，并以文件形式存盘，还可以把数字信号还原成为真实的声音输出。声卡尾部的接口从机箱后侧伸出，上面有连接麦克风、音箱、游戏杆和 MIDI 设备的接口。

（3）视频采集卡。如图 3-4 所示，视频采集卡是将模拟摄像机、录像机、LD 视盘机、电视机输出的视频信号等输出的视频数据或者视频音频的混合数据输入计算机，并转换成计算机可辨别的数字数据，存储在计算机中，成为可编辑处理的视频数据文件。

视频采集卡按照用途可分为广播级视频采集卡、专业级视频采集卡和民用级视频采集卡，它们档次的高低主要取决于采集图像的质量。广播级视频采集卡的特点是采集的图像分辨率高，视频信噪比高，缺点是视频文件所需硬盘空间大，每分钟数据量至少要消耗 200MB，一般连接 BetaCam 摄/录像机，所以它多用于录制电视台所制作的节目。专业级视频采集卡的档次比广播级的性能稍微低一些，分辨率两者是相同的，但压缩比稍微大一些，其最小的压缩比一般在 6:1 以内，输入/输出接口为 AV 复合端子与 S 端子，此类产品适用于广告公司和多媒体公司制作节目及多媒体软件应用。民用级视频采集卡的动态分辨率一般较低，绝大多数不具有视频输出功能。

（4）数码相机。数码相机是一种利用电子传感器把光学影像转换成电子数据的照相机，如图 3-5 所示。与普通照相机在胶卷上靠溴化银的化学变化来记录图像的原理不同，数字相机的传感器是一种光感应式的电荷耦合器件或互补金属氧化物半导体（CMOS）。在图像传输到计算机以前，通常会使用闪存先存储在数码存储设备中，软磁盘与可重复擦写光盘（CD-RW）已很少用于数字相机设备。



图 3-4 视频采集卡



图 3-5 数码相机

3.1.3 多媒体数据压缩技术

从前面对多媒体信息特点的介绍中可以知道，多媒体信息有一个比较显著而且对通信网络有很大影响的特点——数据量很大。如果要在通信网络中支持动态图像信息的存储、传送和处理，将对通信网络各方面的性能提出很高的要求，在不同的应用环境，不同级别的要求下相差也很大。以中等质量，长 1 小时的视像信号显示要求来看，在独立的计算机和通信设备中传输速率要求达到 14.3Mb/s，网络传输速率要达到 114Mb/s，存储容量达到 51.5GB。这个指标要求，以目前发展的通信网络技术和通信技术来看，虽然还不是完全不能达到，但已经需要付出相当昂贵的代价和占用网络中许多宝贵的资源。更为严重的是，上面刚刚所举的例子，仅仅是一个视像信息数据流的要求，而在实际的通信网络工作环境中，完全可能需要支持许多用户的多媒体应用需求，因此在网络的视像服务器中或网络的通信线路中，有可能同时有多个视像



信息数据流在活动, 这将数倍到数百倍地增加对通信网内部的传输速率、存储容量和通信带宽的要求, 如果不采取对视像信号进行压缩的措施, 现有的通信技术和网络技术几乎难以实现。

1. 数据冗余类型

信息数据之所以能进行压缩是因为用来记录和传送信息的载体——数据存在很大的冗余量。在多媒体数据中, 数据冗余的类型主要有下面几种。

① 空间冗余: 空间冗余是静态图像中存在的最主要的一种数据冗余。同一景物表面上采样点的颜色之间往往存在着空间连贯性, 但是基于离散像素采样来表示物体颜色的方式通常没有利用这种连贯性。例如, 图像中有一片连续的区域, 其像素为相同的颜色, 空间冗余产生。

② 时间冗余: 时间冗余是指序列图像中经常包含的冗余。一组连续的画面中往往包含时间和空间上的相关性。而基于离散时间的数据采样来表示运动图像的方式没有利用这样连贯性。例如, 电影画面中人物聊天时, 背景通常不变, 而聊天的人物往往也只是部分肢体位置的改变或动作形态的变化。

③ 结构冗余: 结构冗余是在某些场景中, 存在着明显的图像分布模式, 这种分布模式称作结构。图像中重复出现或相近的纹理结构, 结构可以通过特定的过程来生成。例如, 方格状的地板、蜂窝、砖墙、草席等图结构上存在冗余。已知分布模式, 可以通过某一过程生成图像。

④ 视觉冗余: 视觉冗余是人类的视觉系统对图像场的敏感性是非均匀和非线性的。对亮度变化敏感, 而对色度的变化相对不敏感; 在高亮度区, 人眼对亮度变化敏感度下降; 对物体边缘敏感, 内部区域相对不敏感; 对整体结构敏感, 而对内部细节相对不敏感。可以根据这些视觉特性对图像信息进行取舍。

2. 多媒体数据压缩技术

多媒体信息的特点之一就是数据量非常庞大。以我国使用的 PAL 制视频为例, 每秒播放 25 帧的画面, 每帧影像像数为 720×576 , 如果按 24 位真彩色无压缩方式存放的话, 一帧影像需要 $(720 \times 576 \times 24) \div 8 \approx 1.25\text{MB}$ 的数据量。也就是说, 一张 650MB 的 CD-ROM 只能存储大约 20 秒的无声影视数据。这样庞大的数据量与当前硬件技术所能提供的计算机存储资源和网络带宽之间有很大差距, 成为阻碍人们有效获得和利用信息的一个瓶颈问题。因此, 多媒体数据的存储和传输都要求对数据进行压缩。一般情况是将原始数据压缩后存放在存储设备上或是以压缩的形式来传输, 仅当多媒体数据需要使用时才将其解压缩以还原数据。所以, 所有的压缩系统都需要两个算法: 一个是用于压缩源文件中数据的编码算法, 一个是用于在目的端将数据解压缩以还原数据的解码算法。

目前, 被国际社会广泛认可和应用的通用压缩编码标准大致有 4 种: H.261、JPEG、MPEG 和 DVI。

- H.261: 由 CCITT (国际电报电话咨询委员会) 通过的用于音频视频服务的视频编码解码器 (也称 Px64 标准), 它使用两种类型的压缩: 一帧中的有损压缩 (基于 DCT) 和用于帧间压缩的无损编码, 并在此基础上使编码器采用带有运动估计的 DCT 和 DPCM (差分脉冲编码调制) 的混合方式。这种标准与 JPEG 及 MPEG 标准间有明显的相似性, 但关键区别是它是为动态使用设计的, 并提供完全包含的组织和高水平的交互控制。



- **JPEG**：全称是联合图像专家组（Joint Photograph Experts Group），是一种基于 DCT 的静止图像压缩和解压缩算法，它由 ISO（国际标准化组织）和 CCITT（国际电报电话咨询委员会）共同制定，并在 1992 年后被广泛采纳后成为国际标准。它是把冗长的图像信号和其他类型的静止图像去掉，甚至可以减小到原图像的百分之一（压缩比 100:1）。但是在这个级别上，图像的质量并不好；压缩比为 20:1 时，能看到图像稍微有点变化；当压缩比大于 20:1 时，一般来说图像质量开始变差。
- **MPEG**：是 Moving Pictures Experts Group（动态图像专家组）的英文缩写，实际上是指一组由 ITU 和 ISO 制定发布的视频、音频、数据的压缩标准。它采用的是一种减少图像冗余信息的压缩算法，它提供的压缩比可以高达 200:1，同时图像和音响的质量也非常高。现在通常有三个版本，即 MPEG-1、MPEG-2、MPEG-4 以适用于不同带宽和数字影像质量的要求。它的三个最显著优点就是兼容性好、压缩比高（最高可达 200:1）、数据失真小。
- **DVI**：其视频图像的压缩算法的性能与 MPEG-1 相当，即图像质量可达到 VHS 的水平，压缩后的图像数据传输率约为 1.5Mb/s。为了扩大 DVI 技术的应用，Intel 公司最近又推出了 DVI 算法的软件解码算法，称为 Indeo 技术，它可将数字视频文件压缩至原文件的 1/5 到 1/10。

3.2 多媒体数据的数字化

在多媒体技术中，如何获得及处理多媒体素材是十分重要的环节，不同的多媒体素材需要不同的采集方法和不同的软件来处理。多媒体素材主要包括文本、图形图像、声音、音像、动画等。

3.2.1 文本素材及其数字化

目前，文字素材的处理技术已经非常成熟。文字素材处理的常用手段是通过文字处理软件直接输入，当然可以利用语音识别系统通过话筒输入，也可以采用触摸屏技术手写输入，或利用数字扫描技术识别输入。

1. 利用键盘文字输入

文字录入的最常用的方法就是利用输入法通过键盘输入并编辑处理。英文可以直接通过键盘处理，对于汉字则需要采用相应的输入法来输入。

2. 语音技术

发展了几十年之久的语音技术在计算机硬件的发展和普遍应用的驱动下，已经从模式识别和人工智能的一个分支提升为一门综合人类智能各项研究的独立学科。语音技术包括语音识别、说话人的鉴别和确认、语种的鉴别和确认、关键词检测和确认、语音合成、语音编码等，其中最具有挑战性和最富有应用前景的是语音识别技术。语音技术的非常重要的应用之一就是语音文字输入法。目前，市面上有很多语音输入法，在 Office 2003 版本中也支持语音输入。著名的语音输入法主要有：IBM 语音输入法、语音输入法 2008 等。

3. 手写输入

手写字的输入方法与笔绘板基本相同，也是首先把手写字的位置信息转换成二进制的数字



编码信息,并写入显示存储器,实时显示在显示屏上。但随后的处理就不同了,手写字识别技术会把写入的手写字信息在计算机内用一定的人工智能技术,对手写字进行智能化的识别,并产生与之匹配的数字编码的中文或英文字符,同时所产生的字符也送给显示屏显示,最后经过输入者确认后才作为正确的输入结果。这种书写技术可以使我们与计算机进行信息交流,它不仅可以输入文本文字,而且可以输入计算机或通信设备能够“理解”的命令,对它们进行控制 and 操作。这也是手写字技术与笔绘板技术的不同之处。笔绘板技术虽然也可以进行手写输入,但实际上其输入结果并未被计算机识别和理解,仅仅是对它“照葫芦画瓢”地进行存储、传送。手写字识别技术,特别是中文字识别技术经过十几年的研究已经获得很大的发展,虽然在使用中还存在一定的局限,如一定的错误率,但由于它作为一种多媒体与人工智能结合的技术,特别适合我国的国情——汉字的结构比较复杂而且规律性较少,所以发展仍然十分迅速,最近还在家用计算机市场一度出现火爆。值得一提的是,目前手写输入在手机输入法中已成为主流。

3.2.2 声音素材及其数字化

在multimedia的数据领域中,音频数据是不可缺少的部分。multimedia的音频信号的来源主要有商品语音库和录音制作两种。录音制作中,声音的采集和还原播放是由声卡完成的,音频处理软件可以对音频信号进行编辑处理,以获得精彩的声音效果。

1. 数字音频信号的获取

描述声音的模拟信号是一个连续的模拟波形信号,不能直接由计算机处理,故必须将其数字化。声音信号的计算机获取过程就是声音信号的数字化处理过程。经过数字化处理之后的结果就是数字音频,这样音频格式通常保留了声音的波形特点,所有又称波形声音。数字音频是在时间上离散的数据序列,其数字化过程如图3-6所示。

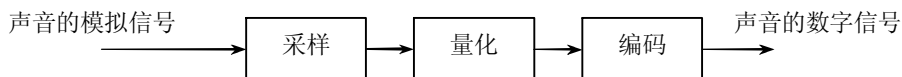


图3-6 声音信号的数字化过程

采样:把模拟音频转成数字音频的过程,就称作采样,所用到的主要设备便是模拟/数字转换器(Analog to Digital Converter,即ADC,与之对应的是数/模转换器,即DAC)。采样的过程实际上是将通常的模拟音频信号的电信号转换成二进制码0和1,这些0和1便构成了数字音频文件。采样的频率越大则音质越有保证。由于采样频率一定要高于录制的最高频率的两倍才不会产生失真,而人类的听力范围是20Hz~20kHz,所以采样频率至少是 $20 \times 2 = 40\text{kHz}$,才能保证不产生低频失真,这也是CD音质采用44.1kHz(稍高于40kHz是为了留有余地)的原因。采样精度决定了记录声音的动态范围,它以位(bit)为单位,比如8位、16位。8位可以把声波分成256级,16位可以把同样的波分成65536级的信号。可以想象,位数越高,声音的保真度越高。

采样精度:样本大小是用每个声音样本的位数bit/s(即bps)表示的,它反映度量声音波形幅度的精度。例如,每个声音样本用16位(2字节)表示,测得的声音样本值是在0~65536的范围里,它的精度就是输入信号的1/65536。样本位数的大小影响到声音的质量,位数越多,



声音的质量越高，而需要的存储空间也越多；位数越少，声音的质量越低，需要的存储空间越少。采样精度的另一种表示方法是信号噪声比，简称为信噪比（Signal-to-Noise Ratio, SNR）。

量化：采样后得到的采样值（振幅值）表示形式音频型号的量化。通常用二进制编码来保存采样值。离散化的数据经量化转变成二进制表示一般要损失一些精度，这主要是因为计算机只能表示有限的数值。例如用 8 位（1 字节）二进制表示十进制整数，只能表示出 -128~127 之间的整数值，也就是 256 个量化级。如果用 16 位二进制数，则具有 64K（65536）个量化级。量化级的大小决定了声音的动态范围，16 位的量化级表示人耳刚刚听得见极细微的声音到难以忍受的巨大噪声这样一个声音范围。量化级对应的二进制位数称为量化位数，有时也直接称采样位数或抽样位数。显然量化位数越多，音质越好，但数据量也越大。

编码：就是对量化级以及二进制数码按一定数据格式表示的过程。音频信号的编码涉及音频信号的压缩技术，也就是说音频信号的编码是使用一定的压缩技术对音频信号按相应的音频格式存放的过程。常见的音频文件格式主要有 MP3、WAV、WMA、RA、RM、RMX 等。

2. 常见的音频处理软件

常见的音频处理软件有很多，典型的如微软的 Windows 自带的“录音机”，其录制的声音文件格式为 WAV 格式，录制好的音频文件可以通过音频文件格式软件转换为其他格式；Cool Edit Pro 是一个非常出色的数字音乐编辑器和 MP3 制作软件；Goldwave 是一款相当不错的数码录音和编辑软件，编辑好的声音文件可以用多种格式存放。

3.2.3 图形与图像素材及其数字化

图形、图像是人认识感知世界的最直观的渠道之一，图形、图像的重要性是毋庸置疑的。图形是指由外部轮廓线条构成的矢量图，即由计算机绘制的直线、圆、矩形、曲线、图表等。用一组指令集合来描述图形的内容，如描述构成该图的各种图元位置维数、形状等。描述对象可任意缩放不会失真。在显示方面，使用专门软件将描述图形的指令转换成在屏幕上显示的形状和颜色。适用于描述轮廓不很复杂，色彩不是很丰富的对象，如几何图形、工程图纸、CAD、3D 造型软件等。图像是用数字任意描述像素点、强度和颜色的位图。描述信息文件存储量较大，所描述对象在缩放过程中会损失细节或产生锯齿。在显示方面，它是将对象以一定的分辨率将每个点的色彩信息以数字化方式呈现，可直接快速在屏幕上显示。分辨率和灰度是影响显示的主要参数。图像适用于表现含有大量细节（如明暗变化、场景复杂、轮廓色彩丰富）的对象。

1. 图形与图像的采集和处理

图形的获取是通过用一组指令集合来描述图形的内容，如描述构成该图的各种图元位置维数、形状。显示时通过专门的软件将描述图形的指令转换成在屏幕上显示的形状和颜色。

图像是由若干点阵（像素）组成的点位图。黑白线条图常用 1 位值来表示，灰度图常用 4 位（16 种灰度等级）或 8 位（256 种灰度等级）来表示点的亮度，通常的彩色图像用 8 位、16 位、24 位、32 位来表示像素点的颜色层次。图像的获得可以通过如数码相机捕捉实际场景画面，用画图软件编辑，用数字扫描设备获取，用抓图软件从显示器的屏幕上抓取等方法获得。图像处理有几个重要的技术指标。

- 分辨率：分为屏幕分辨率和输出分辨率两种，前者用每英寸行数表示，数值越大图形（图像）质量越好；后者衡量输出设备的精度，以每英寸的像素点数表示。



- 色彩数和图形灰度：用位 (bit) 表示，一般写成 2^n 次方， n 代表位数。当图形 (图像) 达到 24 位时，可表现 1677 万种颜色，即真彩。灰度的表示法类似。

2. 常见的图形、图像处理软件

Photoshop：它是由 Adobe 公司开发的图形处理软件之一，主要应用于图像处理、广告设计。最先只是在 Apple 机 (MAC) 上使用，后来也开发出了 for window 的版本。

Freehand：它是 Macromedia 公司 Studio 系列软件中的一员，是一个功能强大的平面矢量图形设计软件，无论是要绘制机械图，还是要绘制建筑蓝图，无论是想制作海报招贴，还是想实现广告创意，Freehand 都是一件强大、实用而又灵活的利器。

Illustrator：它是美国 Adobe 公司推出的专业矢量绘图工具。Adobe Illustrator 是出版、多媒体和在线图像的工业标准矢量插画软件。

CorelDRAW：它是加拿大 Corel 公司的平面设计软件。CorelDRAW Graphics Suite 非凡的设计能力广泛地应用于商标设计、标志制作、模型绘制、插图描画、排版及分色输出等诸多领域。

3. 图形与图像的文件格式

BMP (Bit Map Picture)：PC 上最常用的位图格式，有压缩和不压缩两种形式，该格式可表现从 2 位到 24 位的色彩，分辨率从 480×320 至 1024×768 。该格式在 Windows 环境下相当稳定，在文件大小没有限制的场合中运用极为广泛。

DIB (Device Independent Bitmap)：描述图像的能力基本与 BMP 相同，并且能运行于多种硬件平台，只是文件较大。

PCP (PC paintbrush)：由 Zsoft 公司创建的一种经过压缩且节约磁盘空间的 PC 位图格式，它最高可表现 24 位图形 (图像)。过去有一定市场，但随着 JPEG 的兴起，其地位已逐渐下滑。

DIF (Drawing Interchange Formar)：AutoCAD 中的图形文件，它以 ASCII 方式存储图形，表现图形在尺寸大小方面十分精确，可以被 CorelDRAW、3DS 等大型软件调用编辑。

WMF (Windows Metafile Format)：Microsoft Windows 图元文件，具有文件短小、图案造型化的特点。该类图形比较粗糙，并只能在 Microsoft Office 中调用编辑。

GIF (Graphics Interchange Format)：在各种平台的各种图形处理软件上均可处理的经过压缩的图形格式。缺点是存储色彩最高只能达到 256 种。

JPG (Joint Photographics expert Group)：可以大幅度地压缩图形文件的一种图形格式。对于同一幅画面，JPG 格式存储的文件大小是其他类型图形文件的 $1/10$ 到 $1/20$ ，而且色彩数最高可达到 24 位，所以它被广泛应用于 Internet 上的 homepage 或 Internet 上的图片库。

TIF (Tagged Image file Format)：文件体积庞大，但存储信息量亦巨大，细微层次的信息较多，有利于原稿阶调与色彩的复制。该格式有压缩和非压缩两种形式，最高支持的色彩数可达 16M。

EPS (Encapsulated PostScript)：用 PostScript 语言描述的 ASCII 图形文件，在 PostScript 图形打印机上能打印出高品质的图形 (图像)，最高能表示 32 位图形 (图像)。该格式分为 Photoshop EPS 格式 Adobe Illustrator EPS 和标准 EPS 格式，其中后者又可以分为图形格式和图像格式。

PSD (Photoshop Standard)：Photoshop 中的标准文件格式，专门为 Photoshop 而优化的格式。



CDR (Coreldraw): CorelDRAW 的文件格式。另外, CDX 是所有 CorelDRAW 应用程序均能使用的图形(图像)文件, 是发展成熟的 CDR 文件。

IFF (Image File Format): 用于大型超级图形处理平台, 比如 AMIGA 机, 好莱坞的特技大片多采用该图形格式处理。图形(图像)效果, 包括色彩纹理等逼真, 简直是原景再现。当然, 该格式耗用的内存外存等的计算机资源也十分巨大。

TGA (Tagged Graphic): 是 True Vision 公司为其显示卡开发的图形文件格式, 创建时期较早, 最高色彩数可达 32 位。VDA、PIX、WIN、BPX、ICB 等均属其旁系。

PCD (Photo CD): 由 KODAK 公司开发, 其他软件系统对其只能读取。

MPT (Macintosh Paintbrush) 或 MAC: Macintosh 机所使用的灰度图形(图像)模式, 在 Macintosh Paintbrush 中使用, 其分辨率只能是 720×567 。

3.2.4 视频素材及其数字化

连续的图像变化每秒超过 24 帧(frame)画面以上时, 根据视觉暂留原理, 人眼无法辨别单幅的静态画面, 看上去是平滑连续的视觉效果, 这样连续的画面叫做视频。视频技术最早应用于电视系统, 但是现在已经发展为各种不同的格式以利消费者将视频记录下来。网络技术的发达也促使视频的纪录片段以串流媒体的形式存在于因特网之上并可被计算机接收与播放。

1. 视频采集与处理

所谓视频采集就是将模拟摄像机、录像机、LD 视盘机、电视机输出的视频信号, 通过专用的模拟、数字转换设备, 转换为二进制数字信息的过程。在视频采集工作中, 视频采集卡是主要设备, 它分为专业和家用两个级别。专业级视频采集卡不仅可以进行视频采集, 并且还可以实现硬件级的视频压缩和视频编辑。家用级的视频采集卡只能做到视频采集和初步的硬件级压缩, 而更为“低端”的电视卡, 虽可进行视频的采集, 但它通常都省却了硬件级的视频压缩功能。

2. 常见的视频格式

MPEG/MPG/DAT: MPEG 是 Motion Picture Experts Group 的缩写。这类格式包括了 MPEG-1、MPEG-2 和 MPEG-4 在内的多种视频格式。MPEG-1 相信大家接触得最多的了, 因为目前其正在被广泛地应用于 VCD 的制作和一些视频片段下载的网络应用上面, 大部分的 VCD 都是用 MPEG-1 格式压缩的(刻录软件自动将 MPEG-1 转为 DAT 格式), 使用 MPEG-1 的压缩算法, 可以把一部 120 分钟长的电影压缩到 1.2 GB 左右大小。MPEG-2 则是应用于 DVD 的制作, 同时在一些 HDTV(高清晰电视广播)和一些高要求视频编辑、处理上面也有相当多的应用。使用 MPEG-2 的压缩算法压缩一部 120 分钟长的电影可以压缩到 5~8GB 的大小(MPEG-1 的图像质量与 MPEG-2 是无法比拟的)。

AVI: Audio Video Interleaved (音频视频交错) 的英文缩写。AVI 格式调用方便, 图像质量好, 但缺点就是文件体积过于庞大。

RA/RM/RAM: RM 是 Real Networks 公司所制定的音频/视频压缩规范 Real Media 中的一种, Real Player 能做的就是利用 Internet 资源对这些符合 Real Media 技术规范的音频/视频进行实况转播。在 Real Media 规范中主要包括三类文件: RealAudio、Real Video 和 Real Flash (Real Networks 公司与 Macromedia 公司合作推出的新一代高压缩比动画格式)。Real Video (RA、RAM) 格式一开始定位在视频流应用方面, 也可以说是视频流技术的始创者。



ASF: 是 Advanced Streaming Format (高级串流格式) 的缩写。ASF 是 Microsoft 为了和现在的 Real Player 竞争而发展出来的一种可以直接在网上观看视频节目的文件压缩格式。

WMV: 一种独立于编码方式的在 Internet 上实时传播多媒体的技术标准。WMV 的主要优点在于: 可扩充的媒体类型, 可在本地或网络回放, 可伸缩的媒体类型, 流的优先级化, 多语言支持, 扩展性强等。

RMVB: 这是一种由 RM 视频格式升级延伸出的新视频格式, 它的先进之处在于 RMVB 视频格式打破了原先 RM 格式采用的平均压缩采样的方式, 在保证平均压缩比的基础上合理利用比特率资源, 就是说静止和动作场面少的画面场景采用较低的编码速率, 这样可以留出更多的带宽空间, 而这些带宽会在出现快速运动的画面场景时被利用。

MP4/3GP: 手机常用视频。

3. 常见的视频播放软件

暴风影音: 作为对 Windows Media Player 的补充和完善, 当前暴风影音定位为一种软件的整合和服务而存在, 而非一个特定的软件。它提供对绝大多数影音文件和流的支持, 包括 RealMedia、QuickTime、MPEG-2、MPEG-4(ASP/AVC)、VP3/6/7、Indeo、FLV 等流行视频格式; AC3、DTS、LPCM、AAC、OGG、MPC、APE、FLAC、TTA、WV 等流行音频格式; 3GP、Matroska、MP4、OGM、PMP、XVD 等媒体封装及字幕支持等。

豪杰超级解霸: 超级解霸 9 是集影音播放器、格式转换于一体的多功能播放系统, 包含视频解霸、音频解霸和豪杰 DAC 提取、制作、专辑, 辅助工具, 音视频转换工具等几大部分。

Windows Media Player: 是微软公司出品的一款播放器。通常简称 WMP。Windows Media Player 是一款 Windows 系统自带的播放器, 支持通过插件增强功能, 在 V7 及以后的版本, 支持换肤。可以播放 MP3、WMA、WAV 等音频文件, 由于竞争关系微软默认并不支持 RM 文件, 不过在 V8 以后的版本, 如果安装了解码器, RM 文件可以播放。视频方面, 可以播放 AVI、MPEG-1 文件, 安装 DVD 解码器以后可以播放 MPEG-2、DVD 文件。用户可以自定媒体数据库收藏媒体文件。

RealPlayer: RealPlayer 是网上收听、收看实时音频、视频和 Flash 的最佳工具, 让你享受更丰富的多媒体体验, 即使你的带宽很窄。RealPlayer 是一个在 Internet 上通过流技术实现音频和视频的实时传输的在线收听工具软件, 使用它不必下载音频/视频内容, 只要线路允许, 就能完全实现网络在线播放, 极为方便地在网上查找和收听、收看自己感兴趣的广播、电视节目。支持播放各种在线媒体视频, 包括 Flash、FLV 格式或者 MOV 格式等, 并且在播放过程中能够录制视频。

小结

本章主要介绍多媒体技术, 着重介绍多媒体、多媒体计算机系统、多媒体数据压缩技术、多媒体数据的数字化、多媒体技术的应用及发展以及多媒体信息处理工具软件。

多媒体技术不是各种信息媒体的简单复合, 它是一种把文本、图形、图像、动画和声音等多种信息类型综合在一起, 并通过计算机进行综合处理和控制在, 能支持完成一系列交互式操作的信息技术。

多媒体计算机系统是多种信息技术的集成, 是把多种技术综合应用到一个计算机系统中,



实现信息输入、信息处理、信息输出等多种功能。一个完整的多媒体系统由多媒体硬件系统和多媒体软件系统两部分构成。

多媒体信息的特点之一就是数据量非常庞大。因此，对多媒体数据的存储和传输都要求对数据进行压缩。一般情况是将原始数据压缩后存放在存储设备上或是以压缩的形式来传输，仅当多媒体数据需要使用时才将其解压缩以还原数据。所以，所有的压缩系统都需要两个算法：一个是用于压缩源文件中数据的编码算法，一个是用于在目的端将数据解压缩以还原数据的解码算法。

在多媒体技术中，如何获得及处理多媒体素材是十分重要的环节，不同的多媒体素材需要不同的采集方法和不同的软件来处理。多媒体素材主要包括文本、图形、图像、声音、音像、动画等。