

学校的理想装备

电子图书·学校专集

校园网上的最佳资源

中小学信息科学知识

多媒体与多媒体技术



第一章 多媒体总论

一、信息与媒体

简单地说，信息是“有关事物运动状态和方式的知识，它用于消除不确定性”。比如：招生的广告、股票的价格表、一幅香港的地图、一段录像内容、一个呼机的响声都是信息，它们都向人们传递着关于特定事物的知识。一般讲，信息是通过编码形成的可加以识别处理的数据。信息已成为当今世界的一项重要资源，与物质和能源一样受到人们极大的重视。信息产业被称作“现代社会的先驱产业”，在发达国家将其视为第一产业，信息科学技术革命已成为现代科学技术革命的主流。

信息不是物质，它必须依附于一定的载体，表示、传输、存储信息的载体也称作媒体。信息的媒体可分为表示媒体、存储媒体和传输媒体。

表示媒体是用于表征信息内容的媒体，例如图书中的文字；存储媒体是记载信息的媒体，例如图书的纸张；传输媒体是传递信息时的媒体，例如看书时需要光线，光线就是图书信息的传输媒体。

二、多媒体技术

多媒体技术是随着计算机的发展而发展起来的。早期的计算机以科学计算为主。随着计算机的发展，人机的交流大大增加，人们希望机器能说话，能看东西，与人有更好的交流，为人们提供更多的信息。多媒体技术实际上就是音频技术、视频图像技术及通讯传输技术在计算机上的实现。通过多媒体计算机可以处理声音、视频图像（静态、动态）并在与人交互的方式进行工作，为人类提供更大的方便，帮助人们记忆巨大的文字信息、图像信息、声音信息并能快速地提取为人类服务。

三、多媒体计算机

多媒体计算机又称为 MPC，对于多媒体计算机，人们要求它是能处理多种感觉的媒体，一般包括文字、图形、图像、声音、动画、活动影像，前三种称为静态媒体，后三种称为动态媒体，而在所说的多媒体计算机中至少包括一种是动态媒体，并且各种媒体协同动作以达到同一目的。比如：一个最简单的儿童英语教学的软件，在屏幕上有一个老虎，当鼠标点击老虎后，会出现 tiger 这个英文单词和一个老虎的动画镜头，同时从音箱里发出老虎的叫声。一般来讲多媒体只有在计算机中才能实现，因为只有计算机技术可以将文字、图形、图像、声音、动画、活动影像综合处理，并且有人机的交互性，而录像机、电视机不可以称为多媒体。MPC (Multimedia Personal Computer) 即多媒体个人计算机，1990 年 11 月，由 Microsoft 公司联合、AT&T、NEC 等全球十多家计算机厂商提出的一个新概念，并制定了相应的标准，MPC 是未来高性能多媒体应用的最佳机种，可用来作为电子图书、地图等的工作平台，未来的 windows 操作系统将具备 TV/PC 多任务功能，使 MPC 既可作 PC 使用，又可作 TV 的 Control Box，甚至可自行通过 PC 附加多媒体信息到电视节目上。

四、多媒体技术发展简史

在多媒体技术的发展过程中，1984 年 Apple 公司推出的 Macintosh 机被

认为是多媒体计算机的开始。Macintosh 机使用 Bitmap (位图) 和 Icon (图标) 作为用户界面,在此基础上,Apple 公司又于 1987 年使用了 HyperCard (超级卡),使 Macintosh 机成为当时极为流行的多媒体计算机,直到今日 Macintosh 机在多媒体计算机界仍占很重要的地位。

美国 Commodore 公司 1985 年推出 Amiga 计算机也是多媒体计算机的早期产品。

1986 年 Philip (菲利普) 和 Sony (索尼) 两家公司联合推出了交互式光盘系统——CD-I (Compact Disc Interactive) 系统,将文字、图像、声音存贮于 650 兆的只读光盘介质上,用户能以交互的方式同时播放光盘中的文字、图像、声音等内容。

1987 年 3 月,美国 RCA 公司展示了交互式数字影像系统——DVI (Digital Video Interactive) 系统。这是一个使用 PC 技术、光盘技术来存储和检索活动景象、静止图像、声音和其它数据的系统。之后,Intel 公司买下了 DVI 技术,Intel 公司于 1989 年 3 月宣布将把 DVI 开发成普及的大众化商品,并将 DVI 芯片用于 IBM PS/2 机主板上。在 1991 年的美国计算机展览会上,DVI 系统获“Comdex91”最佳奖。

1990 年 11 月,由 Philips 等 14 家厂商组成的多媒体市场协会成立,制定了 MPC 这个计算机标志及技术规范。MPC 标准的第一级为 MPC1,随后又制定了 MPC2 及以后的一系列标准。这标志着多媒体技术的发展必然产生了相应的国际标准。

1991 年,第六届国际多媒体技术和 CD—ROM 大会标志着多媒体技术进入新的发展阶段,大会宣布的 CD—ROMXA 标准是对原有 CD—ROM 标准在音频方面的扩充。

1993 年发布的 2 级 MPC 标准包括了全活动的视频图像,使用 16 位的音频信号采样。JPEG 成为 ISO/IEC 的 10918 号标准。

1994 年 11 月,MPEG—1 成为国际标准。

五、多媒体技术的应用与发展

多媒体技术主要有以下几个方面的应用:

(1) 教育。多媒体技术特别适应于教学,多媒体在教学中的应用,丰富了人们学习时的感觉,特别是声音和视频动画等给人的印象极深。许多的软硬件厂商都看好这一领域,如《CSC 电脑家庭教师》、《Studio Classroom on CD—ROM 空中英语》以及多种计算机的交互式教程:如《中文版 NT4.0 教程》等。

(2) 家庭娱乐。家用电脑进入家庭还是近二三年的事,但多媒体电脑却已成为购买的对象。一台配置齐全的多媒体电脑,在家庭中相当于一台 CD 唱机、一台 VCD 影碟机、配上电视卡还可以当作电视机、配上视频捕捉卡和 CCD 镜头相当于一台数字照相机、加装正版游戏软件就是一台高档次的 3D 级游戏机,更不用说多媒体电脑还可以加工人的声音、图片、制作电子 MIDI 音乐等等。

(3) 商业展示、广告业及公用设施。多媒体在商业及传媒业中的应用也极其广泛。许多厂商利用多媒体技术展示自己的产品;许多部门将多媒体引入原有系统,使其更加生动丰富并有吸引力;在公用设施方面我们也会见到商场的指示图变成了触摸屏,破烂不堪的邮局的邮政编码本变成了及其方便

的触摸屏，指点江山便会得到邮政编码……

(4) 电子出版。由于光盘的超大储存容量，莎士比亚全集、伊索寓言、安徒生童话、世界博物馆、世界美术馆、贝多芬交响乐等等，从纸上出版品到音像出版品，都已经以多媒体方式呈现在读者眼前。除了一般图书外，字典、百科、期刊、报纸也都有了光盘版。它那声色俱全的形式很快吸引住出版人的眼光，他们纷纷和计算机界结合，试图在这新的出版领域中抢得先机。

(5) 专业影视制作与音频制作。我们见到的《侏罗纪公园》、《灭绝者》、《玩具总动员》等影视作品中饱含制作人员的多媒体技术，影片中的许多精彩镜头都是利用多媒体技术制作出来的，这些镜头不可能用通常的摄制方法获得，是人类想象利用多媒体技术的完美再现。在音响方面，许多的发烧友利用 32 位或 64 位的声卡创造着迷人的 MIDI 音色组合，体验古典音乐、现代流行音乐的听觉感受。

今后多媒体的发展主要在以下技术领域：

- 视频的硬件与软件技术；
- 音频的硬件与软件技术；
- 多媒体演示系统的制作工具；
- 支持多媒体的数据库及其管理系统；
- 支持多媒体的操作系统；
- 数据压缩与还原技术；
- 支持多媒体的外部设备（CD—ROM 与触摸屏）；
- 支持多媒体 A/V 功能的芯片；
- 多媒体信息的网络传输；
- 多媒体计算机的体系结构；
- 动画、广告与可视化技术。

第二章 多媒体信息的计算机表示

一、文字

文字是计算机中的一种最常用的表达方法，多媒体信息也少不了使用文字。西文文字通常用 ASCII 码表示。ASCII 是美国信息交换标准代码 (American Standard Code for Information Exchange) 的缩写，是由 7 位二进制编码组成的字符集，包括大小写字母、标点、数字、数字符号及控制字符等共 128 个字符。

中文文字目前使用的是 1980 年制定的国家标准 GB2312—80，分二级字库共 6763 个汉字和 850 个符号。采用双字节编码，即两个 ASCII 码大小表示一个汉字。港台地区与海外使用的另一种汉字编码标准是 BG5，它的字符集更大一些，但仍不能满足对古今所有汉字的编码。目前，国际计算机界已开始采用了共同的标准，即 ISO10646，该标准有 2 万多汉字，也可处理日本、韩国等许多国家的文字。

二、音频

音频 (Audio) 就是声音的信息表示，通常指在 15 ~ 2000Hz 的频率范围的声音信号。声音进入计算机的前提是对各种声音进行数字化。这个数字化过程有对声音采样和对声音加以量化两个步骤。采样就是把模拟声音信号变换成数字信号，采样的频率 (一定时间内采样的数量) 越高，声音“回放”的质量就会越好，但要求的记录空间也越大。量化就是把采样所得的值用二进制表示，一般可用 8 位、16 位、32 位、64 位二进制数表示 (也叫分辨率)，位越多声音的质量越好，但存储空间也成倍增加。下表表示采样的速率、分辨率与存储空间的关系。

采样频率	分辨率 (位)	立体声或单声道	1 分钟 (Mbyte)
44.1kHz	32	立体声	21
44.1kHz	32	单声道	10.5
44.1kHz	8	立体声	5.2
44.1kHz	8	单声道	2.6
22.05kHz	32	立体声	10.4
22.05kHz	8	单声道	1.3

因此，要获得优秀的音质必然会占用大量的存储空间，所以声音信号需要压缩存储，这将在后面的章节中讨论。

还有一种声音是电子合成的声音，它分为音乐合成和语音合成。音乐合成一般采用 MIDI 标准，语音的合成没有统一标准。

MIDI (Musical Instrument Digital Interface，乐器数字接口) 标准是 80 年代初制定的，它描述一段音乐的音符、音调、使用什么乐器等，并通过音乐或声音合成器 (Synthesizer) 解释播放，产生音乐或语音。MIDI 与数字化声音各有优势，数字化声音比较自然，但会占用大量的存储空间，同样播放时间长度的 MIDI 音乐要比数字化音乐的存储空间小 200 ~ 1000 倍。MIDI 可以比较方便地修改、处理细节，比较适合于音乐创作。

三、图像

图像是多媒体信息的重要组成部分，它再现了人们的视觉。一般静态图像分为两种：位图(Bitmap)和矢量图(Vector—Drawn)。位图适用于逼真、精细的照片式图像，而矢量图则适用于直线、方框、圆圈、多边形以及其它可用角度、坐标和距离来进行数量化表示的图形。上述两种图像按不同文件格式存储，并且可以相互转换。

位图：位图一般是用矩阵来表示，单色的图像可用一维矩阵，即1位的位图表示。16色的图像用二维矩阵，即4位的位图表示，256色的图像用8位的位图表示，32768色的图像用16位表示，24位则可以表示1670多万种颜色，即我们常在计算机中说的真彩色。

位图一般可用Windows中的画笔一类的程序画出，也可以用程序从屏幕上直接抓取，甚至用扫描仪或数字化设备获取照片、电视、电影图像。比如我们经常在大商场见到的“电脑画像”、“电脑婚纱摄影”等均属后者。

抓取来的图像，可在很多程序中进行编辑，替换颜色、插入文字等。位图编辑软件主要有Windows下的画笔，Micintosh机的HyperCard，Aldus公司的Photoshop，或HiJaak Pro等。

矢量图：是一种描述性的图像，它存储的是量化后的数字。矢量图的元素有直线、矩形、椭圆形、多边形、正文等。矢量图多用于计算机辅助设计(CAD)、三维动画等一些电脑创作。

对于一个特定的元素(如一种颜色的正方形)，矢量图比位图在存储空间上要省很多，但对于复杂的图形，矢量图要用许许多多的元素来描述，并且要将他们还还原到屏幕上，就会占用很多的系统资源和时间，位图的刷新速度则相对要快一些。矢量图与位图之间可以进行转换。

不论用哪种表示方法，图像都要按一定的格式存储在软盘或硬盘上。

Windows平台下最通用的图像格式有以下几种：

文件扩展名	图像格式
BMP	BitMap
DIB	Microsoft Windows DIB
PAL	Microsoft Palette
DRW	Micrografx Designer/Ddraw
GIF	CompuServe GIF
JPG	JPEG
PCX	PC Paintbrush
PIC	Lotus1-2-3Graphics
TGA	Truevision TGA (Targa)
TIF	Tiff
WMF	Windows Metafile

四、动画

人在观察过物体时，物体的影像会在视网膜上保留一段时间。当人们看到一系列每次改变很小交替速度很快的图像，就会有物体是连续运动的感觉，动画就是利用这种原理制作出来的。

计算机动画研究始于60年代初，1963年BELL实验室制作了第一部计算

机动画片。当时主要致力于研制二维动画。70年代初期开始研究三维计算机动画，直至80年代中后期才进入实用阶段。随着具有实时处理能力的超级图形工作站的出现，以及三维造型技术、真实感图形生成技术的迅速发展，开发出了一些实用化、商品化的三维动画系统。到90年代初，计算机动画技术成功地应用于电影特技，取得了轰动性的效果。如电影《终结者》、《侏罗纪公园》等影片，均采用计算机动画制作了大量现实生活中没有或用传统技术无法获得的镜头，与真人表演天衣无缝，给人耳目一新的感觉，前者获奥斯卡最佳电影特技奖，后者成为90年代世界最卖座片，标房收入高达数亿美元，由此可见计算机动画技术的重要作用。

五、视频 (Video)

视频是多媒体技术的重要组成部分。全世界广播视频标准有 NTSC、PAL、SECAM 三种。

NTSC 制式：由美国国家电视系统委员会 (National Television Systems Committee) 于 1953 年制定的一种兼容的彩色电视制式，美国、日本等国家广为使用。这一制式又称为正交平衡调幅制。每一帧视频由 525 行水平扫描线构成，分辨率为 525 行水平扫描线。传输速率为每秒 30 帧，每帧由两次扫描完成，每一次扫描画出一个场 1/60 秒，两个场构成一帧。由两个场构成一帧的过程称为隔行扫描，可以防止荧光屏的闪烁。

PAL 制式：是原西德 1962 年制定的一种兼容的彩色电视制式。PAL (Phase Alternate Line) 是指“相位逐行交变”。我国和大部分西欧国家使用这种制式。它是把彩色加到黑白电视中去的一种集成方法。水平扫描 625 行、每秒 25 帧、隔行扫描、每场需要 1/50 秒。

SECAM 制式：法国及俄罗斯等东欧国家使用的彩色电视制式。SECAM (Sequential Color and Memory) 也是 625 行，50Hz，但其基本技术及广播的方法与 NTSC 和 PAL 有很大区别。

除以上广为使用的广播视频标准以外，目前许多国家和厂商正致力于开发新的视频技术。

1. 高清晰度电视 HDTV (High definition Television)

国际学术界也使用 ATV、HRV 等其它简称。国际无线电咨询委员会 CCIR 将高清晰度电视定义为：当观看距离为屏幕高度的三倍时，系统显示效果等于或接近于一名正常视力者在观看原视景物或演示时的临场印象的一种新型电视系统。通俗地说，HDTV 就是画面高度清晰、音响高度逼真的电视。与传统电视相比，主要有四大特点：

(1) 高清晰度画面。电视画面的清晰度主要取决于单位时间内所接收到的横向扫描线的条数。普通彩电的扫描线是每秒 525/625 条，而 HDTV 的横向扫描线则高达 1125/1250 条，因此，HDTV 能提供相当于普通彩电 4 倍~5 倍的高清晰度画质。

(2) 新的电视显示手段。传统的彩色电视的阴极管显示 CRT 技术带有明显的局限性，无法用来制作 76.2 厘米以上的超大屏幕彩电。

(3) 近于人类自然视域的画面结构。HDTV 的画面纵横比由现在普通彩电的 3:4 (近似于正方形) 变为 9:16 (长方形)，这种长方形屏幕与现代电影院的电影屏幕非常相近，避免了现代宽银幕影片转为电视片播出时的画面损失。

(4) HDTV 采用数字高频广播 (DAB) 技术作为电视音响传播手段, 突破了原有电视传音的模拟声道限制, 使收听、收看效果大大提高。

2. 视频会议系统 (Video Conferencing System)

视频会议系统是多媒体网络的重要应用, 不同地点的人员, 可以通过显示器或电视屏幕来传达文件、进行讨论、协调工作、共享信息等。人们无需关心地理位置上的差异, 只需把自己的方案、档案资料准备好, 就可以随时交与“与会”各方, “面对面”地讨论问题。

第三章 多媒体技术

多媒体技术涉及面很广泛，基本技术包括有信息的光存储技术、图像输入和输出技术、图像压缩技术、音频技术、视频技术、网络传输技术等。

一、多媒体信息的光存储技术

光盘是一种大容量的存储介质，它成功地解决了计算机存储容量问题。和磁盘相比，光盘的主要优点是：

- (1) 存储信息量大，每片可保存 680MB 信息。
- (2) 记录介质磨损小，受环境污染的影响小，读写信息时光头不接触光盘表面，信息保存时间可达 50 年之久。
- (3) 存取速度快。一般光盘采用随机存取方式，平均寻道时间低于 0.1 秒。

(4) 单位成本低。光盘单位存储量的制造成本要大大低于磁盘。

光盘主要只读型光盘 CD—ROM、一次写入型光盘 WORM 和可抹型光盘三类。CD—ROM 是多媒体系统中的主要支持产品。有关光盘的内容将在第五章内予以详细介绍。

二、图像输入和输出技术

1. 多媒体系统的图像来源

多媒体应用中所需要的图像主要有以下几种来源：

- (1) 直接购买的数字化图像。它可以存储在 CD—ROM 光盘、磁盘或磁带上。
- (2) 通过计算机和专门软件所创作的图像。可利用 Paintairbrush paintbrush 等软件生成一些图形、屏幕上的图符、动画等。
- (3) 用彩色扫描仪将照片或艺术作品扫描后得到数字的图像。可以将照片、艺术作品变换为全彩色的位图图像。
- (4) 利用电视摄像机捕获的图像。摄像机与插在计算机内的数字化板相连接。将来自摄像机的模拟信号转换成数字数据，从而获得所需的图像。

2. 图像的输入设备

计算机输入图像的设备有扫描仪、电视摄像机、光盘和磁带放像机。

(1) 扫描仪。扫描仪在图像处理领域有广泛的应用。扫描仪的主要技术性能指标有分辨率和灰度等级。无论是正片还是负片扫描仪，其分辨率均用每英寸点数 DPI 表示。高档扫描仪分辨率可调，调节范围较宽，最低为 30DPI，最高达 600DPI。分辨率越高，识别最小细节的能力越强，产生的图像越清晰。单色扫描仪的灰度等级是指识别和反映像素点明暗程度的能力。灰度等级越高，产生的图像越逼真。彩色扫描仪扫描图像时，要对像素点分色，把一个像素点分解为红(R)、绿(G)和蓝(B)三基色的组合。对第一个基色的深浅程度也用灰度等级表示，这就是采色精度。扫描仪还有其它技术性能指标，如扫描区域的尺寸、速度等。

(2) 电视摄像机、录放像等视频设备。电视摄像机、录放像机等视频设备所产生的模拟图像在垂直方向有固定的分辨率，即每帧 625 行扫描线。数字化图像的精度由视频信号捕捉卡决定，但是每帧画面的水平和垂直方向 625 线。电视画面的高宽比为 3 4，为了保证画面的水平和垂直方向有相同

的分辨率，因而水平方向的精度=4/3 × 垂直精度 4/3 × 625=833 列。

3. 图像的采集和存储

输入的图像可分为静态图像和动态图像。静态图像的输入只须考虑分辨率、彩色精度和数字化后的数据量。对于动态图像，则还要着重考虑图像的数字化速度和输入计算机中的速度。

数字化的图像产生大量的数据，尤其是动态图像，数字化后数据量大得惊人，这给计算机存贮图像带来了困难。尽管光存贮技术已使计算机外存容量激增，但与图像的数据量相比，存贮容量仍相差甚远。例如，以 512 × 512 的分辨率和 24bit 精度量化动态电视画面，每分钟产生的数据为：512 × 512 × 24 × 25 × 60/8=2300M 即使光盘容量为 1200M，一片光盘也只能存贮半分钟的电视节目。降低图像数据量的最简单办法就是牺牲图像的分辨率和彩色精度。在多媒体系统中，图像的分辨率和彩色精度要根据应用需求确定。减少图像数据量的另一个有效措施就是数据压缩。图像数据有很大的冗余度，采用压缩编码技术可以大大降低图像的数据量。

4. 图像的处理和应用

图像输入到计算机后，应对它进行各种处理，例如对图像进行滤波、增强、特征抽取等信号处理手段以及旋转、放缩、剪裁、着色、与文本声音混合等编辑手段，以满足各种应用需求，因此，配备有一系列的图像的编辑软件与处理软件。

5. 图像的输出设备

(1) 打印机。图像的打印输出由打印驱动程序完成，打印机的技术指标与扫描仪相同，主要有分辨率和彩色精度，分辨率也用 DPI 表示，一般为 200 ~ 300DPI。彩色精度最高可达 16777216 种彩色。

(2) 监视器。大多数监视器能接收 RGB 基色信号，因而可将数字图像信号转换为 RGB 模拟信号。有些监视器也接收复合视频信号或 S—视频信号，此时就需要将数字化的 RGB 信号转换为模拟的复合视频信号或 S—视频信号。

图像信号经显示卡转变为 VGA 或其它显示标准的信号，就可以在计算机屏幕上显示。显示卡有 EGA、VGA、SV-GA 和 XGA 等多种标准，它们提供了不同的显示分辨率和彩色精度。CGA 的分辨率为 640 × 200，EGA 为 640 × 350，其它为 640 × 400，640 × 480，800 × 600，1024 × 768，1280 × 1024 不等，彩色精度有 16 色，256 色，32 768 色至 1677 万色。

三、音频技术

多媒体系统产生声音的方法主要有数字音响、CD 唱片重放、通过 MIDI 驱动内置或外置的合成器三种。

1. 数字音响

数字音响是计算机技术与音响相结合的产物。通过 A/D 转换对声音采样。重放时，作 D/A 转换，把数字再变回波形。数字化采样要占用很大的存储空间，采样频率越高、比特 (bit) 数越多，声音质量越好。一般来说，采样频率要高于被采样波形最高频率一倍以上，可获得满意的音质。由于人耳的听觉上限大约在 20kHz，因此，采样频率达到 CD 唱片的采样频率为 44.1kHz，要求 40kHz 以上即可能达到最佳听觉效果。比特数 (位长) 也是衡量采样质量的一个参数，它是指每次采样用几个 0 或 1。8 比特的采样能描绘 $2^8=256$

种变化，16 比特的采样则能描绘 $2^{16}=65536$ 种变化。其声音质量比前者要高得多。多媒体系统声音采样的比特数有 8 和 16 两种。采样时可以选择单声道或立体声。采样数据存入光盘或硬盘。

音频信号的数字化方法一般有下列三种：脉冲编码调制 (PCM)；差分脉冲编码调制 (DPCM)；自适应差分编码调制 (ADPCM)。这三种方法的差别在于采样频率和采样精度相同时，数字化音频数据量有所不同，前者方法比后者要多。

2. CD 唱片重放

多媒体系统的 CD-ROM 驱动器可以直接放送 CD 唱片。

3. 通过 MIDI 驱动内置或外置的合成器

MIDI 是乐器数字接口 (Musical Instrument Digital Interface) 的缩写，它是一个通过电缆将电子音乐设备联接起来的协议。这个协议现已成为设计人员共同遵守的一个标准。它规定了电子乐器与计算机之间进行连接的电缆与硬件方面的标准，以及电子乐器之间、电子乐器与计算机之间传送数据的通信协议。以下介绍与 MIDI 有关的几个主要概念：

(1) MIDI 设备。它包含了处理 MIDI 信息的微机及有关硬件接口。一台 MIDI 设备可以有 1~3 个端口，它们分别称之为 MIDI IN, MIDI OUT, MIDI THRU。MIDI IN 接收来自其它 MIDI 设备的 MIDI 信息，MIDI OUT 发送本设备生成的 MIDI 信息，MIDI THRU 将从 MIDI IN 端口传来的信息转发到相连的另一台 MIDI 设备上。

(2) MIDI 合成器。合成器是一种用数字信号处理器或其它类型的芯片产生音乐和声音的电子设备。它产生波形，通过声音发生器送往扬声器。

(3) MIDI 软件音序器。是用于记录、编辑和播放 MIDI 文件的软件。音序器可以帮助专业音乐工作者和音乐爱好者通过 MIDI 文件进行作曲，也可以帮助计算机作曲，用于乐曲修改及播放。计算机作曲软件一般用高级语言编程，按照特别的作曲和算法，设置各种音乐参数（如音高、节、拍、音量、接合、音色等）写出乐曲程序。这样产生的乐曲程序还不能直接在合成器上演奏。必须经过编译，把乐曲程序文件变换成 MIDI 文件，才可以通过音序器作进一步的加工，或在合成器上播放出乐曲。

(4) MIDI 文件。MIDI 信息用一种标准的格式记录，并作为文件而存储，这个文件称为 MIDI 文件。MIDI 文件实际上就是数字形式的乐谱，它由一系列的乐符组成，还包括每个乐符的键，通道号、音高、音长、音量、键落下的速度，以及乐器的配置等。

四、视频技术

下表说明计算机监视器显示的图像与电视机显示图像的区别：

	接收的图像 信号形成	屏幕显示	产生彩色 方式	扫描方式
计算机图像	数字信号单位为 BIT	VGA 监视器分为横向 640 个像素点，纵向 480 个像素点，光强一 个点一个点地变化， 它经过 640 个点	256 种彩色	通过扫描
电视机图像	模拟信号，连续变化。通 过频率和强度的变化来 承载信号，改变信号的波 形则表示不同的彩色和 亮度	512 线。光强有强有弱 之分，从无到最亮，信 号连续经过屏幕	可显示任意高 度的彩色	隔行扫描

从上表可知，监视器显示图像与电视机显示图像有明显的差别，通过视频技术可以使 PC 显示技术与 TV 显示技术完美地结合到一起，彼此通用，发挥更大的功效。

视频技术包括有视频数字化和视频编码技术两个方面。

视频数字化就是指将模拟视频信号经模数转换和彩色空间变换转换为计算机可处理的数字信号，使计算机可以显示和处理视频信号。视频编码技术则是将数字化的视频信号经过编码成为电视信号，从而可以录制到录像带中或在电视上播放。

五、网络传输技术

计算机网络是计算机技术和通信技术二者相结合的产物。计算机网络的突出特点是综合利用了当代所有重要信息技术的研究成果，通过信息的收集、识别、存储、交换、传输和处理技术，把分散在广泛区域中的许多信息处理系统有机地连接起来，组合成一个规模更大、功能更强、可靠度更高的信息综合处理系统。

压缩技术的发展，为多媒体信息网络传输提供了基本条件。图像压缩技术使电话网传输图像成为可能，若与视频技术结合还能实现半动态传输小窗口的视频图像。例如在 64000 波特率的电话网上实现每秒 1 帧的小窗口视频图像的传输。

目前，技术水平已发展到可在综合业务数字网 ISDN 上实现可视电话和电视会议系统。

ISDN 的建立为活动图像和电视电话开辟了广阔的前景。由于彩色活动图像要求 25 帧/秒 ~ 30 帧/秒，因此对压缩比要求很高，运算速度要求更快。图像压缩编解码器是电视电话的关键器件。随着多媒体技术的进步，可视电话将成为多媒体通信终端而获得巨大的发展。

真正解决多媒体的传输，还需要更高速的网络支持。高速光缆局域网 FDDI 和双绞线网络已达到 100M 的速度，覆盖距离可达 100 千米，新的 1000M 的网络也正在逐步趋于成熟。网络传输技术是多媒体发展及应用的基础。

六、触摸屏技术

触摸屏是指一种能对触摸产生反应的屏幕。当人的手指或其他物体触到屏幕不同位置时，计算机能接收到触摸信号并按照软件要求进行相应的处理，这种技术就是触摸屏技术。触摸屏设备由软件和硬件组成。一般的触摸屏是在手指触到显示器屏幕时，将手指的屏幕位置坐标传给计算机。这种输入方法具有直观、方便、快速的特点。触摸屏主要有以下四种类型：

1. 电阻膜式触摸屏

将一个两层导电且高透明度的物质薄膜涂层涂在玻璃或塑料表面上，再按到屏上，或直接涂到屏上。当手指触到屏幕时，在接触点处产生电接触，使电阻发生改变，在屏幕的 X、Y 方向上分别测出电阻的变化量则可确定触摸位置，属早期产品。

2. 电容式触摸屏

将一个接近透明的金属性涂层覆盖在一个玻璃涂层上，当用户接触这个涂层时，电容会改变，使连接的震荡器频率发生变化，测量频率变化的大小即可确定触摸位置。这也是早期的产品。

3. 红外线触摸屏

通常在屏幕的一边用红外器件发射红外光，在另一边设置接收装置来检查光线的遮挡情况。这有两种方式：一种是利用互相垂直排列的两列红外发光器件在屏幕上方与屏幕平行的平面内组成一个网格，而在相对应的另两边用光电器件接收红外光，检查红外光的遮挡情况。当手指指在屏幕上时，部分光束被接收，光电器件因接收到光线减少而发生电平改变。另一种是用扇开的光束从屏幕两角照射屏幕，在与屏幕平行的平面内形成一个光平面，人手触摸时，通过测量投射在屏幕其余两边的阴影覆盖范围来确定手指的位置。

4. 表面声波触摸屏

表面声波 (Surface Acoustic Wave 缩写为 SAW) 又称为表面弹性波。SAW 触摸屏由压电传感器、反射器、触摸屏器件组成，它们可以固定在一块用环氧化物做在平的或弯曲的玻璃表面上的小型压力楔块上，也可直接固定在一台显示器的玻璃表面上，传感器和反射器一起工作，当发射的声波穿过玻璃表面时，一只手触到 SAW 触摸屏，则被触之处使声波发生衰减，这一信号的衰减被接收到并被转换成 XY 轴的坐标传给计算机。

七、多媒体信息的压缩技术

多媒体信息的数据量是非常庞大的，不仅在网络上进行实时传输时需要的传输速率大大超过了当前网络的数据传输率，而且对网络服务器的存储容量及多媒体信息的同步也都提出了很高的要求。为了解决这些矛盾，在不断改善多媒体网络环境的同时，更重要的是必须对多媒体数据进行压缩。一幅 500×500 的 24bit 真彩色图像约需 6MB 的存储量，如用传输速率为 64kbps 的 ISDN 信道进行传送，需要 94 秒才能完成，但用 JPEG 进行压缩后，存储量可降低近 20 倍，用同样的信道进行传输只需 5 秒。由此可见数据压缩的重要性。

多媒体数据压缩技术研究的主要问题包括：数据压缩比、压缩/解压缩速度、简洁的算法。

根据对压缩数据进行解压缩后的数据是否与压缩前的原始数据完全一致作为标准可把数据压缩方法划分为无失真压缩(可逆压缩)和有失真压缩(不

可逆压缩)两类。

国际通信组织 CCITT 对公用电话和公用网制定了一系列的音频压缩标准,包括 G.711、G.721、G.722、G.728,主要是采用脉冲编码调制(PCM)或自适应差分脉冲编码调制(ADPCM)对采样频率为 8~16kHz 的音频进行压缩。

数字化视频中包含有音频信号,所以对视频的数据压缩方法一直很受重视。主要的视频压缩标准有 JBIG、JPEG、MPEG、H.261 等。

(1) JBIG (Joint Bilevel Image Group) 是国际标准化组织 ISO 制定的二值图像压缩标准,是一种采用累进操作方式的无失真压缩方法,其压缩比可达 10^3 。该方法也可用于灰度值或彩色图像。

(2) JPEG (Joint Photograph Expert Group) 是 ISO 和 CCITT 联合制定的对连续色调、多级灰度、彩色或单色静态图像进行压缩的标准,采用顺序(Sequential)和累进两种工作方式。其压缩比为 $10^1 \sim 100^1$,当压缩比小于 40^1 时,基本上可以认为是无失真的。

(3) MPEG (Moving Picture Expert Group) 是 ISO 制定的用于数字存储媒体运动图像及其伴音的压缩编码,包括三个部分即:MPEG 视频(MPEG—Video),它研究视频信号的压缩算法、MPEG 音频(MPEG—Audio),它研究音频信号的压缩算法和 MPEG 系统(MPEG—System),它研究音/视频信号的同步和复用问题。MPEG 图像采用 CIF 图像格式,视频动态速率为 30 帧/秒。由于视频和音频需要同步,所以 MPEG 压缩算法要兼顾对视频和音频。目前已制定出 MPEG—1、MPEG—2 和 MPEG—4 标准。MPEG—1 最后产生一个电视质量的、视频和音频压缩形式的、位速率约为 1.5Mbps 的 MPEG 单一位流;MPEG—2 基于 3Mbps~4Mbps 或 4Mbps 以上速率压缩存储活动图像,质量可达广播级和 HDTV 级;MPEG—4 以 9Mbps~40Mbps 速率压缩存储活动图像,质量为基本级。MPEG 标准的平均压缩比为 50^1 ,是一种失真压缩方法。

(4) H.261 是 CCITT 为可视电话和电视会议制定的压缩标准,也称为 $p \times 64\text{ kbit/s}$ 标准($p=1, 2, \dots, 30$),在低速时($p=1$ 或 2 ,速度为 64 或 128kbit/s)只支持四分之一中间格式 QCIF (Quarter Common Intermediate Format) 图像格式,此时可采用亚帧(Sub—frame)技术,即隔 1 (或 2, 3) 帧处理一帧;当 $p > 6$ 时,支持通用中间格式 CIF,其压缩比可达 48^1 ,压缩比虽较 MPEG 稍低,但实时性很好。

第四章 多媒体系统的图像压缩技术

在多媒体系统中，数字视频技术中的核心是图像压缩技术，这是因为图像的采集、存储、显示、传输等都涉及到大容量存储技术和实时传输等问题。例如，在 VCA 显示中，幅中分辨率为 640×480 ，256 色的图像，若每个像素占 24 位，则需要 7.37Mbits。一个 100M 字节的硬盘只能存放约 100 帧的图像数据，若要达到每秒 25 帧的全动态显示要求，则无压缩的数据传输率一般应为 184kB/s。这样，数据压缩比必须达到数十倍以上，这还不包括音频数据及其它信息。

人们通过多年的对数据压缩进行的研究，用软件、硬件以及软、硬件结合的技术，实现了压缩算法。本节主要介绍数据压缩的基本原理、图像压缩标准等内容。

一、图像压缩原理

可以把黑白数字图像看作一个矩阵，其行和列的交点标识为图像中的一个点，相应的矩阵单元则表示该点的灰色梯度。这种数字陈列的单元称为图像单元或像素。数字转化视频就是由这样的一连串单元组成的。

对于彩色图像来说，每个像素由红 (R)、绿 (G) 和蓝 (B) 三种颜色表示。为了有效地编码，彩色像素可分成灰度 (Y) 和色度 (U 和 V)。因此可以将彩色图像看作是对应于三个成分的三个矩阵。为得到高的压缩比，U 和 V 的处理不同于 Y 外，各个矩阵都可以用与黑白图像相同的方法进行编码。

图像压缩可以分成两种类型：无失真的压缩和有失真的压缩。前者即能够精确地生成原始图像；后者则有某些不太容易引起注意的失真。因为数据的压缩都可以看成是一种变换，所以解压缩（数据恢复）被看作是一种反变换。实现变换的方法称为编码技术。

无失真编码在图像帧序列的压缩中有较大的潜力，各帧图像之间有关相同的部分，原则上可以只存储相邻帧之间的差异部分。

图像压缩技术指标主要有三个：（1）图像压缩比，指压缩前后所需的图像存储量之比；（2）实现压缩的算法难易程度或速度；（3）重现精度，指重现时的图像与原图像相比较有多少失真。第三个指标涉及到人们的主观判断，可以区分两幅图像的优劣，在多数场合下，则还需要有其它的客观指标来区分。

二、彩色视频信号的编码

1. 彩色视频信号

一般来说，可以通过三基色红、绿、蓝（即 RGB）表示图像中的彩色。RGB 构成了一个三维的彩色。此外，还有其它的彩色空间表示，它们可以通过坐标变换而相互变换。不同的电视制式 and 多媒体系统采用不同的彩色空间。例如，PAL 制式采用 Y、U、V 彩色空间，NTSC 制式采用 Y、I、Q 彩色空间。在 YUV 表示中，Y 表示亮度信号，U 和 V 是色差信号。

黑白电视机可以接收它的亮度信号 Y，它是一种与黑白电视兼容的彩色电视信号。由于人的眼睛对亮度信号敏感，对色差信号 U、V 不敏感，因此可以降低表示 U、V 的数据量，一般 Y、U、V 采用 $4 \times 2 \times 2$ 或 $4 \times 1 \times 1$ 格式，可以压缩彩色电视的传输带宽及压缩彩色图像的数据量。

从摄像机、录像机、视盘等视频信号源输出的标准彩色全电视信号，经过同步分离和锁相回路产生一致的时钟信号，再由彩色译码电路、A/D 变换器分出 RGB 或 YUV 信号，并将模拟量转换成数字的信号。数字信号 R、G、B 或 Y、U、V 可以存储在帧存储器里。

2. 电视信号的帧内预测编码

点的亮度值为数字化的 RGB 或 YUV 信号，可以通过编码减少在一个画面内的空域冗余信息，达到数据压缩的目的。

差分脉冲编码调制 (DPCM) 编码方法属于一种预测编码，它通过画面上坐标为 (m, n) 的像素点的三个相邻点 $(m, n-1)$ 、 $(m-1, n-1)$ 、 $(m-1, n)$ 的数值，来预测 (m, n) 点上的数值。若记 (m, n) 像素点的亮度值为 Y_{mn} ，它从邻域像素点的亮度值得到的预测值以 Y_{mn} 表示， Y_{mn} 的预测公式为：

$$Y_{mn} = a^1 Y_{m(n-1)} + a^2 Y_{(m-1)(n-1)} + a^3 Y_{(m-1)n}$$

其中 a^1, a^2, a^3 是选定的参数，它们可以通过使误差

$$e_{mn} = Y_{mn} - Y_{(m-1)(n-1)}$$

达到数字期望为零、均方误差最小，或通过最小二乘法得出。

对误差信号 e_{mn} 进行量化、编码，必然比对原来的图像信号进行量化、编码需要的比特数要小得多。在量化的过程中引起信息损失，使图像失真。但是，如果合理地选择量化步长，使量化误差不超过人眼的可见阈值，图像质量还是可达到主观保真度的要求。

在 DPCM 编码中，如果对量化步长及预测公式的参数能根据图像的局部特征作自适应的调整，就称为自适应差分脉冲码调制 (ADPCM) 编码。

3. 电视信号的帧间编码

电视图像由沿时间轴的帧序列组成，每秒传送 25 帧或 30 帧画面。相邻帧的时间间隔只有 1/25 秒或 1/30 秒，大多数像素点的亮度及色信号帧间变化都不大。可以通过减少帧序列内时域的冗余信息，采用帧间编码技术，达到数据压缩的目的。

根据人眼的视觉特性，对于静态图像的空间分辨率高，但在图像切换时，视觉灵敏度降低。传送帧内画面时，要有足够的空间分辨率，但不必每帧都传送，这样，可增大帧传送的间隔。依靠帧存储器把未传送的帧补充进去，对于画面中的运动部分，可以降低空间分辨率传送。

如果只传送帧间相对应像素点的亮度（或色差）差值超过阈值的像素，而且只传送差值。这样，一帧图像只需传送较少部分的像素。

三、JPEG 静态图像压缩标准

JPEG (Joint Photographic Experts Group) 是联合图片专家小组的缩写。在图像采集、处理与传输的硬件中，为了使不同厂商的产品具有图像压缩共同的依据和兼容性，各公司都重视建立通用的图像压缩标准。1986 年由 CCITT 和 ISO 两个国际组织联合成立了一个联合专家组简称 JPEG，致力于建立适合于彩色和单色多灰度的连续色调静态图像的压缩标准，1991 年形成 ISO/IEC10918 文件，1992 年正式成为一个适用于连续色调、多级灰度、彩色或单色静态图像数据压缩国际标准。

JPEG 标准采用了混合编码方法。它定义了二种基本压缩算法：

(1) 基于空间线性预测技术，即差分脉冲码调制 (DPCM) 的无失真压缩

算法。

(2) 基于 DCT 的有失真压缩算法。并进一步应用游程编码和熵编码。

JPEG 标准的应用目标是：

- (1) 压缩比及图像保真度可在较大的范围内调节，供用户选择；
- (2) 可用于任何连续色调数字图像，不限制图像的内容；
- (3) 在一定能力 CPU 上即可实现，复杂的软件本身要易于操纵。

1. 无失真预测编码压缩算法

JPEG 选择了一种简单的线性预测编码方法，具有重建图像质量好的优点。对于中等复杂程度的彩色图像，无失真预测编码、解码可达到大约 2:1 的压缩比。

2. 基于 DCT 的有失真压缩

基于 DCT 的压缩编码算法包括两种层次的系统：基本系统和增强系统，而且这种算法还定义了二种工作方式：顺序方式和累进方式。基本系统采用顺序工作方式，只采用哈夫曼编码，解码器中只存储两套哈夫曼表。增强系统除包含基本系统外，还可采用自适应二进制算术编码。

DCT 编码的基本过程：首先进行 DCT 正变换，再对 DCT 系统进行量化，并分别对量化后的系数进行差分编码或游程编码，最后再进行熵编码。

(1) 离散余弦变换 (DCT)。JPEG 采用 8×8 大小的子块进行二维离散余弦变换。编码器的输入端以原始图像的 8×8 子块作为输入。除了对原始图像进行顺序的分割外，还将原始图像的采样数据从无符号整数变成有符号整数。

解压缩编码的过程是一个逆过程，DCT 正变换可以看成是一个谐波分析仪，而逆变换看成是一个谐波合成器。 8×8 数据块输入分解至 64 个正交基信号，每个基信号对应于 64 个二维空间频率中的一个。输出的 64 个基信号的幅值，称作 DCT 系数。这 64 个变换系数中包括一个代表直流分量的“DCT 系数”和 63 个代表交流分量的“AC”系数。

DCT 逆变换把 64 个 DCT 系数经逆变换重建 64 点的图像。由于在计算过程中不可避免的精度损失以及系数的量化，这个 64 点的图像不能完全恢复到原始图像。

(2) 量化。对 DCT 系数作量化处理，可以减少幅值，增加零值，达到压缩数据的目的。

JPEG 采用线性均匀量化器，将 64 个 DCT 系数除以量化步长，四舍五入取整。量化步长是量化表的元素。量化表元素与 DCT 系数在 8×8 矩阵中的位置有关。JPEG 给出了供参考用的量化表，也可以由用户来确定量化表。

量化的目的是在使图像质量达到一定的保真度前提下舍弃些次要的信息。不同频率的余弦函数对人类视觉的效果不同，可根据不同频率的视觉阈值来选择量化表中元素的大小。量化的结果一般都是频率低的分量系数大，而频率高的分量系数大多为零。

(3) DC 系数差分编码与 AC 系数游程编码，64 个变换系数中，DC 系数是处于左上角的，它实际上等于 64 个图像采样值的平均值。相邻的 8×8 子块之间的 DC 系数有较强的相关性。JPEG 对于量化后的 DC 系数采用差分编码，二相邻块的 DC 系数的差值为 $DC_i - (DC_{i-1})$ 。

其余 63 个 AC 系数量化后通常出现较多零值，JPEG 算法采用游程编码，并建议在 8×8 矩阵中按照“Z”形的次序进行，可增加零的连续次数。

系数编码后都采用统一的格式表示，包含二个符号的内容：第一符号占一个字节，对于 DC 系数而言，它的高 4 位总为零；对于 AC 系数而言，它表示到下一个非零系数前，所包含的连续为零的系数个数。第一字节的低 4 位表示 DC 差值的幅值编码所需的比特数，或表示 AC 系数中下一个非零幅值编码所需的比特数。第二个符号字节表示 DC 差值的幅值，或下一个非零 AC 系数的幅值。

还需要指出，由于第一字节中只有 4 位表示游程长度，量大值为 15。当游程长度大于 15 时，可以插入一个至多个 (10) 字节，直至剩下的零 AC 系数个数小于 15 为止。

因此，63 个 AC 系数表示为由二个符号对组成的序列，其中也可能插入 10 个字节，快结束字节以全零表示。

(4) 熵编码。对于上面给出的码序列，再进行统计特性的熵编码。这仅对于序列中每个符号对中的第一个字节进行，第二个幅值字节不作编码，仍然直接传送。

JPEG 建议使用二种熵编码方法：哈夫曼编码和自适应二进制算术编码。对于哈夫曼编码，JPEG 提供了针对 DC 系数、AC 系数使用的哈夫曼表（包括对于图像的亮度值与色度值二种情况），在编码和解码时使用。JPEG 解码器能够同时存储最多 4 套不同的熵编码表。

3. 基于 DCT 的增强型系统

增强系统增加了二种累进操作的方式，它们在编码方法和步骤上与顺序方式基本相同，不同的是它们采取多次扫描的方式，实现对每个图像分量的编码。第一次扫描得到粗糙的压缩结果，根据粗糙的压缩数据，重建一幅质量较低的可识别图像。通过多次扫描，作进一步的压缩，直至最后全部所需的压缩。在这个过程中，由于不断增加的信息，使重建图像的质量达到满意的结果。

由于不是一次完成压缩编码，在量化器的输出之后，有一个缓冲区，用以存储量化后的 DCT 的全部系数。对此缓冲区存储的系数进行多次扫描，分批完成压缩编码。

二种累进方式分别为：

(1) 按频段累进。对 64 个 DCT 系数，每次只对其中一些频率的系数进行压缩编码、传送，然后再逐渐增加其它频率的系数，直到全部系数传送完毕。

(2) 按位逼近。对量化后的 DCT 系数按照其数位由高至低分成若干段，然后依次对各段进行压缩编码。首先对量高有效的 n 位进行压缩编码、传送，再依次对下面的段进行。

4. 基于 DCT 的分层操作方式

JPEG 提供了一种分层操作方式。它是对原始图像的空间分辨率的不同层次之间加以转换而编码。空间水平分辨率与垂直分辨率都以 2 为倍数降低，形成一个金字塔形。它的编码过程为：

(1) 把原始图像的空间分辨率降低。

(2) 对降低分辨率的图像采用基于 DCT 的压缩编码（可为顺序方式、累进方式或无失真编码）。

(3) 对低分辨率的图像解码，重建图像，然后，用手动的办法恢复图像的水平与垂直分辨率。

(4) 把分辨率已升高的图像的预测值与原图像的差值采用基于 DCT 的压缩编码。

(5) 重复步骤(3)、(4)，直至达到图像的完全的分辨率。

5. 压缩比和图像质量

JPEG 标准的压缩编码与解码是有失真的，主要原因是变换后系数的量化而引起失真。压缩的效果与图像内容本身有较大的关系，高频成份少的图像比高频成份多的图像可获得更高的压缩比，而图像仍有较好的质量。对于给定的质量因子(分为 1~255 级)，采用相应的编解码参数，达到相应的压缩效果。

JPEG 没有专门规定彩色图像的编码算法。可以对彩色空间中各个分量分别进行压缩编码。当彩色空间各分量彼此独立时，可获得最好的压缩效果。

四、MPEG 动态图像压缩标准

“声音”和“图像”是多媒体计算机技术处理的两种新媒体，都具有数据量庞大的特点，特别是全动态、全屏幕的活动图像，数据量更大。如果将未经压缩的活动图像直接记录在 CD 光盘上的话，一张 650GB 的 CD 光盘只能记录 40 秒钟的信号，一部 90 分钟的影片需要用 1500 张光盘来保存，这是不可想象的事情。同时，要想处理这样大的数据量，就要求硬件能够处理庞大的数据流，对硬件提出很高的要求，这是目前硬件难以做到的。为了解决这一困难，必须对图像数据进行大规模的压缩，减少图像数据的冗余度。目前的 VCD 和 DVD 即是分别采用 MPEG1、MPEG2 标准来实现活动图像及其伴音数据的压缩和还原的。

MPEG 是活动图像专家组 (Moving Picture Experts Group) 的缩写。该专家组成立于 1988 年，现属于 ISO/IEC 信息技术联合委员会第 29 分委员会的第 11 工作组，缩写为 JTC1/SC29.WG11。

MPEG 专家组的工作项目在开始时分 MPEG1 (码率达 1.5MB/s 的编码)、MPEG2 (码率达 10MB/s 的编码)、MPEG3 (码率达 40MB/s 的编码) 共 3 个项目。MPEG3 项目于 1992 年 7 月撤消，因为当时 MPEG2 的工作内容及进展已包括了 MPEG3。随着技术发展的需要，1993 年 7 月正式成立了 MPEG4 项目组，它的目标是甚低码率小于 64kB/s 的视频音频编码。目前，MPEG1、MPEG2 标准已正式公布执行。

1. MPEG1 数据压缩标准

MPEG1 标准制定于 1991 年，是目前所采用的最有效的电脑视频影像、声音数据压缩标准。MPEG1 标准包括视频、音频和系统 (音频和视频同步) 三部分。完整的视频要求视频和音频同步，MPEG1 将具有电视质量的视频、音频数据合为一个数据流，并将数据流的速率降至 1.5MB/s (其中图像数据率占 1.2MB/s，立体声伴音占 0.25MB/s 左右)。

MPEG1 压缩的基本方法为：在单位时间内首先采集并保存第一帧图像的信息，此后在对单帧进行有效压缩的基础上，只存储其余帧图像中相对第一帧图像发生变化的部分，以达到图像数据压缩的目的。它包括时间上的压缩和空间上的压缩两个方面。

(1) 时间上的压缩 (即帧间压缩)

MPEG1 视频压缩算法减少时间冗余量的方法主要是以 16×16 的块作为运动补偿单元的运动补偿技术，即对比视频影像的前、后两张画面，如果两者

完全一样或接近一样，便可去除其中一帧画面。去除的判断阈值视所要求的质量而定。较好的 MPEG 算法可在 2 帧 ~ 3 帧画面间作一次对比运算，而一般较差的算法只能在 12 帧 ~ 30 帧画面间作一次对比运算。对比运算做得越频繁，压缩效率越高，但运算量也越大，因此要求有更高性能的硬件设备。

(2) 空间上的压缩 (即帧内压缩)

在视频影像的同一帧画面中，往往有许多完全重复的部分或几乎相同的部分，这就可以只记录一次其相同或相近的部分，而去除其余部分，以达到节省存储空间的目的。MPEG 标准通常采用变换编码、矢量量化编码以及混合使用变化编码、视图加权标量量化、游程编码等技术来降低空间冗余度。

MPEG1 采用有损图像质量的非对称压缩算法，压缩时间 (大约几十小时) 远远大于解压时间 (1 个多小时)。压缩只有一次，需大量运算，进行图像的比较分析，而解压可有上千次，运算量较少。

2. MPEG2 数据压缩标准

MPEG2 是建立在 MPEG1 的基础上、以提高图像质量为目的的通用国际编码标准，共包括系统、视频、音频、符合性测试、软件、数字存储媒体的指令和控制、非向后兼容音频、10 比特视频、实时接口等 9 个项目。其中第 1 到 3 部分作为 MPEG2 的核心，已在 1994 年 11 月正式公布执行，其它的部分将在 1995 到 1997 年陆续公布。

MPEG2 的压缩方法与 MPEG1 的方法相似，基本算法相同，但增加了场间预测。MPEG2 的传输码率可以调整，支持从可视电话到 HDTV 多种应用，针对不同的用途，MPEG2 标准又分为 4 个等级和 5 个档次。

(1) MPEG2 等级。

低等级 (Low)：这是最低层，是 CCIR601 图像格式的 1/4，GI 则采用这种图像格式；

主要等级 (Main)：图像质量达到广播级，相当于 CCIR6-01 图格式；

高 1440 等级 (High1440)：相当于每行 1440 取样的 HD 图像质量；

高等级 (High)：相当于每行 1920 取样的 HDTV 图像质量。

(2) MPEG2 档次。

简单档次 (Simple)：除了没有内插图像帧以外，与主要档次 (Main) 相同。简单档次的目的是为了节省 RAM；

主要档次 (Main)：没有分级性，但质量最高；

信噪比可分级档次 (SNR)：相对主要档次有所改进，在信噪比方面有分级性；

空间域分级档次 (Spatially)：在图像空间清晰度方面有分级性；

高档次 (High)：支持 4:2:2 格式和全部分级性。

(3) MPEG2 图像档次和等级的组合依据 MPEG2 标准的 4 个等级和 5 个档次，可组成 20 种不同的图像格式和压缩化，有定义可采用的只有 11 种。

第五章 光盘

光盘是采用光学方法进行信息存取的盘形介质。光盘存储是多媒体技术中最重要的一种技术，自其问世以来，受到各界极大的重视，成为当今最重要的一种信息存储方法。

光盘具有存储密度高、容量大、稳定性好、保存时间长等许多优点。激光的相干性很好，用聚光系统可以聚焦成极少的光点，因而光盘上每位信息占的空间仅有1微米左右，光盘的记录密度为 5×10^5 位/毫米²是目前录像磁带存储器的40倍，计算机磁盘的100倍。光盘的容量大，普通的CD—ROM可存储680GB的数据，只需要100张直径30厘米的双面光盘就可以记录藏书量2000万册的美国国会图书馆的全部藏书。而且光盘为非接触式读写，不会发生机械损伤，载体保存时间长达几十年甚至上百年。光盘存取数据时间短，数据传输率高，现代的光盘驱动器的速度仅次于硬盘。

光盘的应用范围相当广泛，从商业广告、文件存储到音乐影视节目的播放、图书资料的发行等等。自70年代以来，光盘产业在与信息记录和处理各个领域得到了世界范围的迅速发展，应用前景令人振奋。

一、光盘的种类

光盘可分为：只读光盘、一次写入/多次读取光盘、可擦写光盘三大类。

1. 只读光盘

全称是只读光学存储器 (Optical Read Only Memory)，其代表性产品又分为以下几个小类：

(1) LD (Laser Video Disk) 直径30厘米的大光盘，记录信息一般为影视信息，又称大影碟；

(2) CDA (Compact Disk—Digital Audio) 音乐光盘，记录信息为数字音频信息；

(3) CD—ROM 一般记录与计算机软件相关的信息；

(4) VCD (Video—CD) 记录用MPEG1标准压缩的动态音像信息，又称为小影碟；

(5) DVD (Digital Video Disk) 记录用MPEG2标准压缩的动态音像信息，容量比VCD大许多。

(6) CD—I (Compact Disc Interactive) 交互式光盘，一般记录静态图像及同步音频信息；

(7) CD—G 记录音频信息及静态图像。

2. 一次写入/多次读取光盘

英文名称是WORM, Write Once Read Multiple。其代表性产品是CD—R (CD Recordable)。

3. 可擦写光盘

英文名称是R/W, Rewritable 代表性产品有：

磁光盘 (MO: Magneto—Optical)

相变盘 (PC: Phase Change)

二、只读光盘 CD—ROM

只读光盘在多媒体技术中占有举足轻重的地位，因为它容量大、稳定性

高、非接触式读写、占空间小易保存、载体保存时间长。但只读光盘的最大特点是厂商制造出产品后，用户不能对其内容进行修改，只可以读取。只读光盘的驱动器叫 CD—ROM 驱动器，已成为多媒体计算机 MPC 必须配置的外设。下面我们先介绍 CD—ROM 的原理、结构及制作过程，再介绍各种格式的光盘，然后探讨 CD—ROM 驱动器。

1. 只读光盘结构、原理及制作

(1) 只读光盘的结构。

光盘是由基片、记录介质、保护层等几部分组成。

基片材料一般为金属（如铝片）、玻璃（如硼硅酸玻璃）或塑料（如聚碳酸酯）。基片介质要厚度均匀、有较高的透光性和足够的光学均匀性，并且还要有足够的强度和抗热变形性，以及不会受腐蚀。

记录介质材料有 Te、TE—c、Te 合金、Te—Se、卤化银膜，非晶半导体膜，非晶硫碲化合物，热塑材料，光磁材料，光色材料等。记录介质要有一定的记录灵敏度、高分辨率、高信噪比，能够实时记录和瞬时读出，并且容易制成长期稳定且均匀完好的记录膜。

记录介质是光盘组成成分中最重要的部分，它直接关系到光盘片的质量，一般地说，记录介质的结构有四种：

单层结构：就是把记录介质直接涂在基片上，这种结构的光盘较难获得充分的对比度；

二层结构：使用电介质膜和金属膜二层结构，提高光盘记录灵敏度，并可用于可擦写光盘；

三层结构：用铝反射膜、透明的电介质层和记录层三层形成记录介质层，三层结构比二层结构记录灵敏度更加提高。

液晶结构：液晶结构可用在可擦写型光盘中，由于液晶在比室温稍高的温度下可以改变其状态，故利用此特点，可进行对光盘信息的改写。

保护层一般使用树脂材料，它直接涂在介质表面上以保护记录介质。保护层需要坚硬牢固，既可防擦伤又可防潮，以达到保护记录介质的作用。

(2) 光盘数据的制作。

准备数据：包括对数据的处理和数据的调制编码，也就是说，将信息按照它们要在光盘上存储的形式划分成以 2048 个字节为一个单元的许多数据块，在每一块数据块上再附加 288 个字节作为系统数据供系统使用，如 ECC 错误检测等。

用户数据追加同步模式、标头、EDC 码、ECC 码后，就可进行数据的调制过程。调制的目的是为了数据均衡，使得在保持分辨率的同时提高字节密度，保证一块数据中出现的错误不会影响其它数据块，同时还要降低 EDC 信号量的频率以符合伺服装置的要求。为了达到以上要求，8 位二进制数据被转换为 14 位调制数据，同时在每 14 位字节之间再追加 3 个字节称为归并位，使得在 0 和 1 代码中，符合下列规定：在“1”之间至少有两个“0”，在“1”之间至多有 11 个“0”。当然这是由处理过程自动计算出的。

母盘制作：先在抛光盘玻璃主盘上涂上光敏抗蚀材料，然后在记录台上用激光对抗蚀涂层进行刻制，沿着螺线状导向槽由盘心逐渐向外，直至完全占据整个记录区。采用所需要的调制码来调节激光束的通、断。对已曝光的有抗蚀涂层的主盘进行蚀刻，并放在蒸发箱内，然后在主盘表面上涂一薄层导电层。把盘放进电镀槽内，在玻璃主盘上生成一层镀镍层。将镍层从玻

璃上分离开，用它制成模版，母盘便作好了。

盘片的制作：把由主盘制成的模版放进成型机，注入熔化的丙烯树脂材料，冷却凝固后就可以复制出与原主盘刻制特征完全一致的盘片。在盘片信息记录面一侧，用蒸发法或溅射法把记录层和光吸收层连续地涂敷、镀制，使其厚度为 30 微米。蒸发镀制记录层之后，再涂上保护层，使之与盘基相粘合，便完成对盘片的制作。

(3) 光盘的读取。

光盘信息记录方式有许多种，因此信息读出的方法也各不相同。比较常见的是凹坑记录信息的读出。这是对于采用凹坑记录信息方式的光盘，照射到光盘信息径迹上的激光波长约 0.8 微米，接近激光波长。当激光落在无信息坑的平面上，光能量会大致全部反射回（80%）；当光点落在小坑上时，只有少量的光被反射（10%），这样使得光敏二极管输出强度变化的比例达到 10^{-1} 以上，就可读出信号。

对于可擦写光盘，激光头输出的光分两种，一种是强光，一种是弱光。强光用于记录信息，弱光用于读取信息。

(4) 光盘格式的标准。

CD—DA 标准：红皮书标准——数字音频（适用于记录或播放音频 CD）。

1981 年为 CD—DA 制定的规格，称为红皮书。这是最先建立的规格。它采用脉冲编码调制 PCM，用 44.1kHz 的采样频率对音频信号进行采样和编码，记录在 CD—DA 盘片上。CD—DA 可以得到超级 Hi Fi 的音响效果。

CD—ROM 标准：黄皮书标准——CD—ROM（适用于 CD—ROM 的读写）。

1985 年为 CD—ROM 制定的规格，称为黄皮书。它对光盘的物理格式和盘地址作了规定。

1986 年又制定了它的逻辑格式，在 1987 年正式作为国际标准 ISO9660。

CD—I 标准：绿皮书标准。

1988 年 Philips 公司和 Sony 公司联合推出 CD—I 系统，并对其软、硬件部分作了详细的规定，称为绿皮书。它是所有规格中最复杂的一种，采用了多种压缩编码技术，完成对各种数据的编码。对于音频信号采用 PCM 和 ADPCM 技术进行压缩编码，可获得四种不同音质的信号。绿皮书中的 ADPCM 技术使 CD—I 光盘的容量较红皮书规格扩大 16 倍。

CD—ROMXA 标准。

这是 1989 年 Philips 和 Microsoft 合作推出的、并在 1991 年国际多媒体和 CD—ROM 大会上宣布为扩展结构体系标准。它减少了数字化音频信号的光盘存储空间，是对原 CD—ROM 规格的扩充。具体是用 ADPCM 作为压缩编码，使用三种不同的采样频率。CD—ROMXA 规格是黄皮书和绿皮书规格的结合。

Photo—CD 标准。

1992 年 Kodak 公司推出，等待计算机厂商的确认。Photo—CD 光盘用于记录数字化照片信息。

CDTV 标准。

1991 年由 Commodore 公司专门设计并规定了它的格式，由于 CDTV 没有采用 CD—I 标准，只能在指定的系统中使用。

DVI 标准。

1990 年由 IBM 和 Intel 公司联合推出，DVI 光盘只能在 DVI 系统环境下使用，但 DVI 技术具有很强的功能，将全屏幕全动态视频图像、高清晰度静

态图像、双声道高保真音频信号及传统的图形和文本集成到一个交互的环境，可实现动态视频/音频数据的实时采集、实时压缩存储/反压缩播放，成为全数字化多媒体技术的代表。

VCD 标准：白皮书标准。

视频光盘 Video CD (适用于播放 MPEG 解码视频)

橙皮书标准。

多对话式 Multi—Session (适用于在可记录 CD 上增加文件)

2. 几种常见的只读光盘

(1) CDA (Compact Disk—Digital Audio) 又称“激光唱盘”。

1981 年，Philips 公司和 Sony 公司制定了关于 CD—DA (Compact Disc—Disc—Digital Audio) 的红皮书标准，至此激光唱盘有了统一的标准。标准规定：CDA 的直径为 12 厘米，厚为 1.2 毫米。一张 CDA 盘可记录 74 分钟的以 44.1kHz 采样、16 位的数字立体声信息。

CDA 盘是早期的 CD 标准，它虽然可以在计算机的 CD—ROM 中进行播放，但由于不符合 ISO9660 标准，所以不可以象普通的 CD—ROM 光盘一样用 DIR 命令来显示其内容。CDA 自 1982 年大量进入市场后，以其高清晰度的音质，成为录音盒带的主要替代品，CD 机成为家庭音响的主要组成部分。

(2) LD (Laser Video Disk) 通称“大影碟”。

LD 是由 Philips 公司于 70 年代初推出。直径 30 厘米的大光盘，记录信息一般为影视信息。它的分辨率为 425 线，采用 MPEG1 标准进行图像压缩。LD 有“主动放演”和“慢放演”两种规格。“主动放演”为每秒 25 转恒角速度，每面可放演 30 分钟带宽 5MHz 的 PAL 制电视节目；“慢放演”为每秒 11 米左右的恒线速度，每面可放演 60 分钟全带宽 PAL 制式的电视节目。

从旋转速度上分，LD 有常线速 (CLV) 型和常角速 (CAV) 型。

CLV 型以恒定的线速旋转。从内圈至外圈，轨道的长度有所不同，因而 LD 在读不同的圈时，盘的旋转速度不一样。在读中心轨道时，每分钟为 1800 转，读最外圈轨道时每分钟为 600 转。采用 CLV 格式，每个轨道可以存储多一帧的图像。它使得每个盘面的播放时间比 CAV 格式多一倍。12 英寸盘每面可播放 60 分钟，8 英寸盘每面可播放 20 分钟。

CAV 型 LD 以恒定的角速旋转。对于 NTSC 制式，CAV 盘的旋转角速度为每分钟 1800 转。对于 PAL 制式 CAV 盘的旋转角速度为每分钟 1500 转。每帧图像存储于一个轨道上，轨道的长度从内圈至外圈逐渐增加。12 英寸的 CAVLD 每面可播放 30 分钟，8 英寸盘面 14 分钟。

LD 视盘一般都具有交互功能，概括地说，对视盘的交互控制有四个不同的级别：

第一级的交互性最低。第一级程序只能控制在某章停下来，或在预先编号的帧停住。视盘的每章都有一个相连的代码编号。这一级别的交互不接受来自微机系统的控制，原则上由给定的逻辑顺序而定。

第二级交互控制的程度较高。当视盘播放时，交互命令被转储到视盘播放机内嵌式微处理器中。当用户进行选择时，程序转移到选定的帧位置，并根据存储的命令播放一个帧序列。由于控制程序是事先存储在视盘内，因此，二级程序的视盘只适用于不需经常改变、使用期限较长的节目。

第三级交互是目前多媒体平台使用的一种典型方式。视盘上不存储程序代码，只存放音频和视频数据。微机通过与视盘播放机之间的通信控制视盘

的操作，可以达到更高层次的交互性。在这一级交互中，用户不必受给定的帧序列的限制，可以通过触摸屏输入进行选择，还可将计算机内存存储的数据与视盘中的数据合并后再显示在屏幕上。目前它是多媒体应用中采用的主要方式。

第四级交互将程序或命令与 LD 数据结合起来存储在单一的光存储介质上。这种新的光盘称为 LD—ROM，它能同时存储高保真的模拟视频、音频信号和数字数据。

(3) CD—ROM 又称为“电脑光盘”。

CD—ROM 一般记录与计算机软件相关的信息，一般用于计算机的外部存储器，可以记录数字数据及音频、图像等信息。现代的 CD—ROM 有两种格式，格式 1 (Model 1) 格式可提供 680MB 的存储容量，格式 2 (Model 2) 格式可存储 780MB 的信息。

电脑光盘的内容极其丰富，大致可分为以下几大类

计算机软件类：包括计算机操作系统，大公司为了发行操作系统软件而制作的光盘，如 Microsoft 公司的 NT Serv-er 4.0，SUN 公司的 Solaris 2.5 等等，都以电脑光盘形式发售。其它的计算机软件也可以通过电脑光盘载体发售，如 Microsoft 公司的语言开发工具 VC++ 4.2、VB5.0 等，再就是电脑厂家发放的随机光盘，一般其中软件的内容是 OEM 厂家提供的随机软件及配套的硬件的驱动程序以及一些随机赠送的游戏软件等。

多媒体教学系统：包括计算机软件的教学系统如：NT4.0 互动式教程；Learn By Yourself 系列，Access，Visual Basic，PowerPoint 等软件的学习指导；

图像专辑：如 3DS 图像库、背景图库、矢量图形库、曲线字型库等，是计算机辅助设计的最佳工具盘；

百科全书类光盘：如：中国少儿百科全书光盘版，京华揽胜，天文学百科等；

语言教学系统：如：空中英语教室，大家学英语等；

家庭娱乐性光盘：如微软的 HOME 系列光盘，从西餐烹调到飞机百科，从图书书架到百科全书，从高尔夫球到旅游地图，极富娱乐性；

图书类：如《计算机世界·微电脑世界光盘版》、《中国计算机用户光盘版》；

纯游戏类光盘：如天龙八部、命令与征服、红色警报等。

(4) VCD (Video—CD) 小影碟。

VCD 最早源于 Philips 和 JVC 公司 1992 年 10 月推出的“卡拉 OK CD”规格，在这个规格的基础上，1993 年 3 月，又联合了 Sony 公司和 National 公司制定了 VCD1.0，同年 10 月，经改进制定了 VCD1.1，1994 年 7 月又完成了 VCD2.0 规格的制定，为 VCD 产品的标准化奠定了基础。

VCD 的主要特点

制造工艺简单 VCD 盘的制造工艺流程与普通的 CD 唱盘相同，现在的 CD 生产厂商无需增加设备投资，即可生产 VCD 盘；VCD 播放机的生产也很简单，只需在普通的 CD 唱机上加 VCD 视频/音频解码电路即可。

适应面广 VCD 盘既可以在 VCD 播放机上播放，也可以在带有 MPEG1 解压卡的多媒体计算机 (MPC)、CD 唱机、LD 视盘机上播放。由于目前 MPC、CD 唱机、LD 视盘机的普及面越来越广，商品化的 MPEG1 解压卡也很多，价格也

很便宜，许多已购买上述设备的用户只需少量投资，即可欣赏 VCD 盘上的节目。

性能价格比高 从性能方面来看，VCD 的图像质量低于 LD，与 VHS 相当，但在实际播放时要优于 VHS；VCD 的音质与 LD 相当，优于 VHS。VCD 盘片略低于录像带的价格，大大低于 LD 盘的价格。VCD 盘体积小，易于保存，反复、多次播放也不会造成损伤。

VCD 播放系统的原理与构成

VCD 播放系统的构成主要有 3 部分：CD 驱动器、MPE- G1 解码器、微处理机。CD 驱动器用于读取 VCD 盘上经压缩的图像和音频数据；MPEG1 解码器用于将压缩的图像和音频数据还原；微处理机用于控制整个播放过程。

VCD 播放系统的基本工作原理为：CD 驱动器首先从 VCD 光盘上读取被压缩的串行数据流信号(由图像数据和音频数据组成)和控制信息，然后经 MPEG 数据流分析器将串行数据流信号分解为图像数据信号和音频数据信号，这些被压缩的图像和音频数据信号再分别经 MPEG 视频译码器和 MPEG 音频译码器进行还原，最后再经过各自的数模转换器(DAC)变成模拟的图像信号和音频信号；这些模拟的图像信号和音频信号经过相应的放大即可在有关的图像设备和语音设备上播放。由于 VCD 采用的是 MPEG1 的有损图像质量的非对称压缩算法，在对图像数据进行压缩时即有数据丢失，因此用 VCD 播放系统播放的图像将达不到原始图像的水平。

(5) DVD (Digital Video Disk)。

DVD 的标准

由于 CD 的容量仅有 650MB，难以满足高质量影视和大数据量多媒体应用的要求。这些年来，由于众多厂商的努力，CD、激光光学、反射材料和盘片制造等技术都取得了巨大进步，数字编码和压缩算法已达到相当高的水平，集成电路和伺服机构也有了相当大的发展，使制造更高容量、更高速度的激光存储介质成为可能。

1995 年 9 月 18 日 Toshiba/Time Warner 和 Sony/Philips 为主的两大阵营制定出一个单一标准 DVD 标准。DVD 包括用于存储音乐的唱片 DVD—Audio、用于存储电影和视盘的 DVD—Video，以及用于存储计算机数据的 DVD—ROM 等分支。其标准主要分为以下三种：DVD 容量标准、DVD 视频标准、DVD 音频标准。

DVD 系统的构成和相关技术

DVD 系统的构成主要有 DVD 驱动器、MPEG2 解码器和 DVD 盘三部分。

DVD 盘是用来存储高清晰度活动图像及其伴音数据的。由于 DVD 对图像质量要求很高，必须达到广播级水平。相对 VCD 来讲，DVD 的数据量将是 VCD 的 4 倍，要想在一张光盘上装下 135 分钟活动图像的数据，要求 DVD 盘的存储量必须是 VCD 盘容量的 7 倍~8 倍。为了增大光盘容量，在 DVD 盘的生产工艺上采用了一种新的技术：即采用波长为 635 毫米~650 毫米的红激光刻盘，由于红激光波长短，轨道的道间距减少，可以大大增加光盘的存储容量。

DVD 驱动器是用来读取 DVD 盘上数据的设备。为了达到读取广播级图像数据的目的，DVD 驱动器的速度必须提高到原来 CD 驱动器速度的 4 倍以上。另外，要读取 DVD 盘上比 VCD 更小的数据孔径，必须采用波长更短的激光器，目前 DVD 驱动器采用的是波长为 635 毫米~650 毫米的红激光。

MPEG2 解码器的作用是将经 MPEG2 标准压缩的数据进行还原。目前世界上

生产 MPEG2 解码器的厂家很多，产品也不少，如 IBM、NEC 和汤姆逊公司等。

DVD 的技术规格：外形采用双面结构，直径为 12 厘米，厚 1.2 毫米，由厚 0.6 毫米的两张盘片粘合而成；DVD 的存储采用 SD 技术，存储量为每面 4.7GB，可存储 135 分钟的经 MPEG2 格式标准压缩的广播级电视图像及其伴音，数据读取则采用 MMC 的单面方法；信号调制方式采用 MMC 的 EFM Plus 方式；误差修正方式采用 SD 的 RS-PC 方式。

DVD 图像参数：DVD 的技术核心是 MPEG2 标准，MPEG2 标准的图像格式共有 11 种组合，DVD 采用的便是其中主要等级的图像格式，使其图像质量达到广播级水平。具体的图像参数为：PAL 制——720 × 576，25 帧/秒；NTSC 制——720 × 480，30 帧/秒。

三、一次写/多次读光盘

1. 一次写/多次读光盘

一次写/多次读光盘是用户自己将所需的信息记录在光盘上，这种光盘提供了可写一次的机制，即一次性写入或一次次续写。当把信息写入光盘后，光盘介质发生物理变化，信息永久记录于此光盘上，不可修改、删除或重写，这种光盘被称为 WORM。

早期的 WORM 为 5.25 英寸大小，价格较高，基本上没有投入大批量使用。现在一种新型的 WORM 型产品 CD-R 投入市场后，盘片价格较低，且不断下降（进口的 CD-R 盘片 20~30 元，国产的 CD-R 盘片 10 元），CD-R 光驱从 4000 元到 10000 元之间，使一次写/多次读光盘真正有了实用的价值。

2. CD-R 盘片

CD-R 盘片的结构与 CD-ROM 基本相同，盘的直径为 120 毫米，孔径 15 毫米，厚 1.2 毫米。两者均为三明治式结构，底层为聚碳酸酯，上层是保护膜。两者的区别仅在于 CD-R 比 CD-ROM 多一记录介质层，并且有正弦波形的预刻槽，用于 CD-R 盘片在驱动器中的转速控制。记录介质层为有机染料，如酞花菁 (phthalocyanine)、花菁 (cyanine) 或金属，反射层通常采用反射率、熔点较高的金属，最常用的是金。CD-R 采用金膜是由于金具有良好的抗腐性和很高的光反射率。此外，CD-R 夹层中的多膜与底层之间还有一层染料层，在写入器大功率光束的激励下，染料层变形，在底层形成与 CD-ROM 那样的“地”与“坑”，以表示二进制数据。

常用的 CD-R 盘片按容量区分有三种：CD-R18、CD-R63 和 CD-R74，分别是对应于 18、63 和 74 分钟或者 175、553 和 650MB 的存储容量。与激光唱片一样，信息只记录在单面。

CD-R 盘片的使用寿命约为一般 CD-ROM 盘片的 3 倍，可达 100 年。

据报道，IBM 的研究人员已在实验室研制成 2 个和 4 个记录面的 CD-R，并宣称，制作 10 层或 20 层盘片在技术上已不存在问题。

3. CD-R 一次写入的原理

CD-R 一次写入的原理是，在对应 CD 导入区（存放文件目录表）的前面增加了 PMA (Program Memory Area) 和 PCA (Program Calibration Area) 两个区。其中 PCA 区用于调整写入激光波长的功率，因为不同的 CD-R 的染料型记录层的激光响应波长不同；PMA 区用于存放追加数据信息的目录，当整盘数据录入完成后，再将 PMA 区中的目录写入导入区，这时 CD-R 就成为记录有信息的只读型光盘。

通常 CD—R 有两种记录方式，一种是用激光在记录层烧蚀凹坑，另一种是用激光加热使染料型记录层变色，变色部分的反射率比附近区域低得多；CD—R 的记录格式和 CD 相同，即 CLV 和 MER；记录的数据格式，则是根据需要，按照相应标准，把空白 CD—R 写成 CD—DA、CD—ROM 或 CD—I 格式。读出都是利用反射激光光强的变化来实现，与 CD 读出原理相同，所以 CD—R 上记录的信息可以在通用 CD 驱动器上读出，即与 CD 读兼容，这也是 CD—R 受市场重视的原因。

4. CD—R 的应用

CD—R 是一种正在兴起的新技术，即使在国外，它的应用也起步不久，许多应用领域还有待于进一步开拓。一般地说，它适应于那些只需要制作少量 CD—ROM，或者由于安全等原因不便将信息外传的 CD—ROM 软件制作部门。具体地说，目前可用于以下几个方面：

(1) 大容量多媒体存储。CD—R 用来存放固定不变或不经常改变的信息是最理想的介质，可广泛应用于各种平台的多媒体文档管理系统。尤其是对于那些拥有敏感数据的单位，诸如国防、安全、银行等部门，CD—R 具有它所特有的优越性。

(2) 信息发布。有些企业、商业宣传部门经常需要制作一些多媒体宣传材料，例如产品发布展览会上演示或信息交换用的资料等，且 CD—R 具有携带、邮寄方便等优点。

(3) CD—ROM 样品。CD—ROM 软件制作人员可先在办公室里制作少量的 CD—ROM 样品，待测试、修正后再送工厂批量生产。这样可以大大提高 CD—ROM 制作的成品率，缩短制作周期。

(4) 硬盘后备。目前，多数硬盘仍然采用磁带作为数据后备。磁带上的数据不但日久容易丢失，而且工作方式是顺序读出，速度很慢。采用 CD—R 就可以避免这些缺陷。新型的 CD—R 系统均可对盘片进行逐次写入。因此，一张 CD—R 盘片在没有写满以前可多次使用。

(5) 照相 CD。将照相底片或幻灯片连同声音记录在 CD—R 上，可通过多媒体计算机欣赏图像，聆听声音。这种将 CD—R 用作电子相册的业务，在国外有些照相馆已经向用户开放。照相 CD 技术还可用作建立医疗图像库。

(6) 图文存储。在有些文档存储系统中，需要将某些原始文件原封不动地记录下来，例如法律文件、信用卡、医疗记录与工程图等。以往，这些领域多采用缩微胶卷来保留原始文件。采用 CD—R 不但可大大提高检索速度，而且可降低成本。

5. CD—R 系统构成

CD—R 写入系统主要包括：计算机、CD—R 写入器、1GB 以上的硬盘、SCSI 接口以及预制作和制作软件。

计算机系统可以是 IBM 兼容机、Macintosh、Sun SPARC 工作站等机种；操作系统可以是 DOS、Windows 或 UNIX 等。为了获得较好的写入效果，最好用 RAM 4MB 以上、CPU 性能相当或超过 486 的机种。

CD—R 写入系统要求配置一个 1GB 以上的高速硬盘。因为在写入器向 CD—R 写入的过程中数据流不能中断，对于倍速的写入器，硬盘的数据传送率最好在 600kB/s 以上。在向 CD—R 写入以前，需要先在硬盘上建立一个 CD—R 映象，因此硬盘容量最好大于被写入数据容量的 2 倍。

写入系统的核心是 CD—R 写入器。这种写入器最早是由 YAMAHA 和 Sony

推出的，最初的售价为 2.5~5 万美元。最近价格已跌到 1200 美元以下，从而打开了 CD—R 广泛应用的大门。

CD—R 写入器的主要技术指标是数据传送率、缓冲容量以及是否支持逐次写入。CD—R 的工作速度用数据传送率来衡量，这是指写入头向盘片送入数据的速度。与 CD—ROM 驱动器一样，有单速、双速、三速和四速四种，传送率分别为 150kB/s、300kB/s、450kB/s、以及 600kB/s，目前最流行的是双速写入器。

缓冲器用来存放来自硬盘的数据，并由此将数据写入 CD—R。它必须有足够大的容量，以保证写入数据的连续性，在写入过程中不能变空。一般写入器的缓冲器容量为 256kB~2MB，最大可扩充到 32MB。

CD—R 写入器的另一个指标是一张盘片上的数据是否可以分若干次写入，即所谓逐次写入。目前，多数写入器已具备这种功能。大部分写入器可兼用作 CD—ROM 读出器，读出速度不一定与写入相同，例如双速写入的机器在读出时可能只是单速。平均无故障时间为 2~5 万小时，国内售价为 3.5~4 万元之间。

CD—R 写入器与计算机之间采用小型机的标准接口 SCSI，目前用得较多的卡是 Adaptec1540/1542，它们能提供较高的数据传送率。1GB 以上的硬盘也可用此卡与写入器连接。但有经验的 CD—R 制作人员建议，为了工作可靠，最好两者分别使用二块独立的 SCSI 卡。

预制作与制作软件主要完成四项工作：在硬盘上生成 CD—R 映象并将所需文件转换成 ISO 9660 格式；对 CD—R 映象进行仿真；将映象写入 CD—R，建立目录表（TOC）；将写入数据与写入前数据进行比较。

目前，常用的预制作与制作软件有 CD—GEN、WinOn- CD、EasyCDpro、CDWRITE、Personal Scribe 等。有些软件是与写入器一起提供的，也有单独出售的。

四、可重写光盘与可重写光盘机

1. 可重写光盘类型

(1) 相变型可擦写光盘。

自 1983 年日本 National 公司展出第一台可重写相变型视频光盘以来，真正拉开了可重写光盘的序幕。此后又于 1989 年开发成功可直接重写的 3.5 英寸相变光盘，其单面容量达 280MB，采用分离式光头，平均存取时间小于 54.5 毫秒。

利用晶态和非晶态两种状态表示数据的 1 和 0，而利用两者之间的相转变，来达到可擦重写的目的。读、写、擦过程如下：将高功率和窄脉宽的激光束聚焦到薄膜上，Insb 基化合物或碲基化合物介质吸热后迅速升温至熔点并在聚冷条件下形成非晶态，由于晶态和非晶态材料的折射率与周围区域有明显反差，用小功率激光束可检出此反射率的差异，这就是读出过程。擦除过程则是用较长脉宽和较低功率的激光束作用于记录点，使该点温度上升到低于材料的熔点而高于非晶态的转变温度，使介质产生重结晶而恢复成多晶状态。这种擦、写、读过程所需的时间与激光束的功率、材料温升速率等有关。近年来大量研究工作都是围绕着降低擦除时间（加快晶化速度），提高晶态和非晶态的反衬度以及材料稳定性等方面进行的。

目前单束光可重写的相变型光盘已实用化。如采用波长为 680 纳米光头，

实现 1.2 微米的窄道宽，密度比原来提高 2 倍。3.5 英寸相变光盘容量为 600MB，与目前 5.25 英寸磁光盘相当。日立公司不久前也报道了其容量达 160MB 的 2.5 英寸相变型光盘驱动器，其装置厚度仅 15 毫米。

相变光盘的传输速率受晶相转变速度的影响，但它可在双面介质上采用单光束直接重写。由于这一特点，相变光盘不仅整体容量大（双面可重写），而且具有向下兼容的特点，这就是所谓的多功能光盘驱动器，这种多功能光盘驱动器，能兼容相变可重写和一次写入光盘盘片的 CD—ROM 信息，所以仍受到用户的青睐。

（2）磁光型可擦写光盘。

可擦重写磁光盘系统于 1988 年正式推出，一问世即呈供不应求之势，发展极为迅速。磁光型光盘的读写方式分磁场调制和激光调制两种。所用介质为稀土—过渡金属（Re—Tm）非晶态合金。磁场调制方式是在恒定功率激光照射下，介质的温度升至居里温度（ T_c ）或补偿温度（ T_{com} ），光头中磁场线圈的调制信号使磁场反向磁化，在写入激光束很快离开聚焦点后产生某一方向的磁化区域。光调制则是在恒定磁场作用下，通过调制激光脉冲信号来改写数据的。信息的读出是利用磁光克尔效应。信息的擦除过程与写入方法相同，把磁化方向又反转过来。相对于相变型光盘而言，磁光型的光头部分增加了磁场线圈，即增加了光头复杂程度。但是由于磁光记录材料的寿命相对较高，状态也较稳定，制备过程容易控制，因此目前情况下，厂家更多地选择了磁光型可重写光盘。

同相变光盘一样，为了保持向下兼容，也开发了多功能的磁光盘驱动器，HP、Sony 和 Maxtor 三家公司均安装这种能处理可重写和一次写入两种介质的多功能驱动器。最近又有消息报道，松下电器产业公司试制成了可对磁光盘记录/再生的 3.5 英寸相变光盘装置。这种光盘装置实际上是磁光盘驱动器，但附加了可对相变光盘记录再生的电路。利用盘盒上的写保护孔，判别介质是磁光盘还是相变光盘，若是相变光盘，则首先将转速从 3000rpm 降至 1800rpm，以便使用同一激光源（35 毫伏）对相变介质进行写/读。据称，若能加大或改变激光器功率则转速可以不变。这种情况，一方面反映了多功能光盘的发展趋势，另一方面说明即使是最早生产相变光盘的松下公司，也已介入磁光这个领域。

相变光盘最显著的特点是：可实现直接重写；信号电平比磁光盘的高，因而信噪比高；不需磁场元件，因而光头结构简单，重量轻，易实现集成化，可提高伺服跟踪精度和数据传输率。但是相变光盘介质的稳定性及重写循环次数还不够理想。磁光型技术成熟，主要也是体现在磁光介质方面，擦写循环次数超过 106 次，工作寿命在 15 年以上。为了实现与相变光盘的竞争，磁光盘的直接重写技术、分离光头技术等，也在进一步研究之中。总之，可重写光盘正朝着大容量小型化方向高速度发展。

2. 可重写光盘机

与可重写光盘相对应，可重写光盘机目前主要有相变型（P—C）和磁光型（M—O）两种形式，相变型光盘具有直接盖写能力，而早期的磁光盘则一般要先擦、后写、再校，即先转一圈抹去原来的数据，再转第二圈写入新数据，第三圈进行校验。但目前很多公司已开发出可直接重写的磁光盘驱动器，采用双光头三光束技术，在转一圈内完成擦除、写入、读出校验三项工作。可重写光盘机因其可直接重写的特殊性，故可作为计算机硬盘的后备存储器

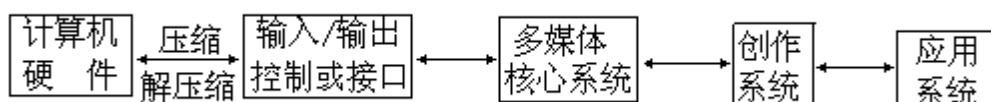
或二次存储装置，甚至还可作为个人机或网络服务器、工作站上具有可换性的一级存储器。随着技术进步，生产批量的增大，其价格会不断下降。另外，值得一提的是多功能光盘驱动器，这种驱动器可以读写 WORM 介质和可重写介质。目前多功能光盘机的类型较多，有的可以同时使用一种 WORM 介质和 MO 介质，也有的可使用相变 WORM 盘和相变可改写盘，National 公司还推出了 3.5 英寸磁光、相变两用盘等等。多功能驱动器把两种驱动器的功能组合在一起，从而降低了制造成本和销售价格，同时用户可通过一台光盘驱动器进行多种应用，而不需要把单独的 WORM 驱动器和可重写驱动器一起挂在系统上。多功能光盘机除可作为二次存储装置使用外，还可在 CAD/CAM/CAI、数据采集、医疗卫生、金融事务处理等方面得到广泛应用。

第六章 多媒体个人计算机

一、多媒体计算机技术的概念及体系结构

具有多媒体功能的计算机被称为多媒体计算机，其中最广泛、最基本的是多媒体个人计算机（Multimedia Personal Computer，简称 MPC）。具备多媒体功能的计算机系统即是多媒体计算机系统。

多媒体计算机系统的结构如下图所示：



(1) 计算机硬件。是多媒体系统的基础，包括计算机及附属声卡、视卡、CDROM 等；多媒体计算机硬件及板级产品多媒体计算机要能综合处理声、文、图信息、必须解决下述四个问题：视频和音频信息的获取问题；视频和音频信息的压缩和解压缩问题；视频和音频信息的实时处理和特技；视频信息的显示和音频信息立体声输出。这是建立多媒体计算机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以及压缩和解压缩功能，一定要用硬件板卡实现。

(2) 多媒体信息的压缩与解压缩。一般多媒体信息需要巨大的存储空间，所以在处理这些信息时要对他们进行压缩，而且要求压缩的速度极快，一般这个层次都是在芯片中实现。而且整个 IC 界对压缩与解压缩规定了许多标准；

(3) 输入/输出控制或接口。包括对多媒体硬件设备的驱动、控制和软件调用规定；

(4) 多媒体核心系统。也就是多媒体操作系统，它是介于操作系统与上层软件之间的系统，负责多媒体信息在操作系统与应用软件之间的处理；

(5) 创作系统。是为开发者提供的开发工具程序包，一般创作系统除编辑多媒体的工具外，还有播放功能。多媒体创作系统的设计目标是缩短多媒体应用软件的制作开发时间，降低对制作人员素质的要求，多媒体编辑工具可分成下述三类：高档编辑工具：适合电影、电视系统专业编辑工具；中档编辑工具：适合教材、娱乐系统的制作编辑；低档编辑工具：适用于商业介绍资料、简报及家庭学习材料的编辑。目前世界上比较流行的，在 Windows 平台上运行的编辑工具有：Macromedia 公司的 Authorware Professional、Aimech 公司的 Icon Author，Asymetrix 公司推出的 Multimedia ToolBook，MacroMedia 公司推出的 Action 等。

(6) 应用系统。就是最终与用户见面栩栩如生的各类软件。多媒体应用系统利用多媒体数据库和多媒体编辑工具，可以方便、迅速地编制出极有效益的多媒体应用系统。如多媒体办公自动化系统；多媒体工程数据库系统；多媒体人事档案管理系统；多媒体地理管理系统；各种电子出版物。

二、MPC 标准

由于多媒体计算机技术是因实际需要而产生发展起来的，最初各厂商都独自生产自己的多媒体产品，现在市场上也存在多种不同标准的多媒体集成一体机或多媒体集成系统，著名的有 MPC 标准、Apple 公司的 Macintosh 系

统(其操作系统是 System 7.0, 数字视频平台为 Quick Time)、Corm—modore 公司的 Amiga 系统(其多媒体描述语言为 AmigaVision)、Philips/Sony 公司开发的 CD—1 系统(其操作系统是 CD—RTOS, Compact Disk Real Time Operating System)、In- tel/IBM 公司合作开发的 DVI 系统(其核心是 AVSS, 即 An- dio/Video Support System) 和 AVK, 即 Audio/VideoKernel) 等。但其中发展最快、得到大部分厂商支持的是 MPC 标准。

世界几家较大的多媒体计算机厂商, 包括 Microsoft, IBM, Philips, NEC 等于 1990 年成立了多媒体计算机市场协会 (Multimedia PC Marketing Council, INC), 以进行多媒体标准的制定和管理, 该组织制定的标准即是 MPC (MultimediaPC, 多媒体个人计算机) 标准。1991 年该组织根据当时的 PC 机发展水平制定了多媒体 PC 机的基本标准即 MPC 标准, 对多媒体 PC 机及相应的多媒体硬件规定了必需的技术规格, 要求所有使用 MPC 标志的多媒体产品都必须符合该标准的要求。随着计算机和多媒体产品性能的不不断提高, 多媒体计算机市场协会 1993 年 5 月对 MPC 标准中的大部分规定进行了更新, 发布了多媒体 PC 机的新标准, 即 MPC2 标准, 原来的 MPC 标准被称为 MPC1 标准。

现在, 计算机和多媒体产品有了进一步的发展, 更名为多媒体 PC 机工作组 (The Multimedia PC Working Group) 的多媒体计算机市场协会 1995 年 6 月份公布了最新的多媒体 PC 机标准, 即 MPC3 标准。

MPC1 标准规定了 PC 机多媒体的基本要求, 提供了多媒体 PC 机的基本框架; MPC2 更新了 MPC1, 进一步扩展了多媒体 PC 机的结构, 使多媒体计算机技术逐步标准化; MPC3 并不是用以替换 MPC2 的, 它只是对多媒体 PC 机的表现能力有了更高的要求, 为多媒体技术的广阔应用打开了大门。

三、MPC 的发展趋势

未来的 MPC 除了在中多媒体功能上不断加强外, 还会朝着如下三个方面发展。

1. 一体化, 即向多媒体一体机方向发展

所谓多媒体一体机 (Moniputer), 简单地说, 就是显示器加上 PC 再加上多媒体, 当然并不是 MPC 加个显示器套在一个外壳里就可称为一体机。除了外形的一体化以外, 一体机的设计与普通台式机的差别在于, Moniputer 是显示器和电脑主机一体成型的“模块化”电脑, 不仅符合人体工程学的原理, 而且结构简洁, 安装容易, 即插即用。

早期的多媒体电脑, 大多是在电脑整机基础上加装多媒体硬件, 使之具备多媒体功能。由于功能的不断加强和提升, 不同厂商不同规格的多媒体部件堆在一部机器上, 不仅破坏了 PC 原本的简洁美观, 也造成了使用、维护和功能上的麻烦与损失。因此, 当一些知名 PC 厂商开始推出自己的 MPC 品牌时, 一体成型也成了一种趋势。

最早的是 1993 年由 PC 界的龙头老大 (当时还不是) Compaq 公司推出的 Presario。目前, 我们在市场上可以见到的一体化的 MPC 有: IBMAptiva2168X/Y/240、CompaqPre- sario5522、IPC 万智能及 AcerAspire 等。

从市场的角度看, 多媒体电脑的很大市场在家庭, 至少厂商是期望以此打破电脑进入家庭的瓶颈, 而一体化的 MPC 正好可以从多方面满足家庭的需

用：简洁美观、占用空间小、易装易用，如同家里的彩电或音响一样。我们注意到，Apple 电脑多年来在美国及全球家用电脑市场上经久不衰，Apple 电脑从 Apple 到 Pertorma，十多年来均保持着一样的简洁美观，两者之间恐怕有某种必然的联系吧。可以想见，由 MPC 的市场特点，将会有更多的一体机问世，并成为 MPC 的主流。

2. 网络与通信功能成为标准配置

虽然无论是 MPC1、MPC2 还是 MPC3，都没有将网络与通信方面的要求列入，但是眼下可见的 MPC，很多都具有网络与通信的功能，并且在宣传时都竭力突出这方面的能力。Fax/Modem 及网络通信软件已经成了 MPC 不可缺少的基本配置。

这种现象与电脑发展的大趋势是一致的。我们处在一个网络时代，今天我们已能感到网络无所不在，明天我们将会感到，我们无时不在网络之中，由于网络的普及，对于各类信息来源（本地的或远程的）、多种信息类型（数字的或模拟的）及多种信息载体（语音的、数据的、图形的、图像的、视频的等）的信息的获取、存储、转发、处理，就成为电脑的一个很重要的甚至是核心的任务。

由于多媒体计算机大多用于家庭，如何引起消费者的购买热情是一个关键。一年多以前，多媒体热是由于 VCD 影碟的大量流行，如今影碟的热劲已经退去，新的热点在哪儿，厂商显然看好网络，特别是因特网的兴起。尽管国内因特网的用户还很少，而且其发展也受到诸多方面的影响，但因特网连入千家万户似乎是不可逆转的潮流。在这情况下，多媒体电脑自然都要标榜有网络与通信功能，以争取用户的青睐。由于因特网给人们带来信息获取方式的全面改变，深深地改变着人们的生活方式和模式，因此，应该说把因特网作为卖点，比看几张影碟对人们更有持久的吸引力。方正卓越多媒体电脑提出，使其卓越电脑成为信息处理中心，显然比单纯的娱乐中心更具远见。

随着网络技术的飞速发展和网络建设的快速推进，未来家用电脑的主流是简单便宜的网络电脑（NC），还是功能更强大的个人电脑（PC），目前是仁者见仁、智者见智，看不出个眉目。但依笔者之见，NC 也许更合适的市场在学校、银行或公司，此外功能单一的网络电脑也不符合中国人的消费习惯。结合了网络功能的多媒体电脑应该是未来中国家庭购买电脑时的首选品种。

当上因特网、有自己的 Email 地址，不再是身份与地位的象征，当网上购物、远程医疗、视频点播等成为人们生活中的一部分时，网络与通信功能对于 MPC 来说，显然是不可缺少的。

3. 家电化的趋势

多媒体电脑的家电化趋势也很明显，因为家庭是它的主要市场。无论从销售渠道还是从电脑功能的设置以及外观、使用等方面，多媒体电脑还会继续呈现家电化的趋势。但是，这里有一个问题，电脑毕竟还不同于一般的家用电器，过份的家电化倾向可能达不到预期的目的——被家庭接受。某个品牌的多媒体电脑几乎无所不包的家电功能，市场上反响并不热烈。也许纯粹家电化的应该是多媒体电视，而不是多媒体电脑。

第七章 多媒体软件

多媒体软件是多媒体技术的核心。多媒体软件必须运行于多媒体系统之中，才能发挥其多媒体功效。多媒体软件综合了利用计算机处理各种媒体的新技术，如数据压缩、数据采样等等，能灵活地高度运用多媒体数据，使各种媒体硬件和谐协调地工作，使多媒体系统形象逼真地传播和处理信息。多媒体软件的主要功能是让用户有效地组织和运转多媒体数据。

多媒体软件大致可分成四类：

- (1) 支持多媒体的操作系统；
- (2) 多媒体数据准备软件；
- (3) 多媒体编辑软件；
- (4) 多媒体应用软件。

以下分别予以介绍。

一、多媒体操作系统

操作系统是计算机的核心，它控制计算机的硬件和其它软件的协调运行，管理计算机的资源。因此，它在众多的软件中占有特殊重要的地位，它是最基本的系统软件。所有其它系统软件都是建立在操作系统的基础上的。

操作系统具有两大功能：首先是通过资源管理提高计算机系统的效率，即通过 CPU 管理、存储管理、设备管理和文件管理，对各种资源进行合理的调度与分配改善资源的共享和利用状况最大限度地发挥计算机的效率。其次，改善人—机界面向用户提供友好的工作环境。操作系统是用户与计算机之间的接口。窗口系统是图形用户界面的主体和基础。窗口系统是控制位映象、色彩、字体、光标、图形资源及输入设备。

MS—Window3.0 版上已开发了许多多媒体应用软件，它和 Macintosh Toolbox 成为多媒体应用的主要支持软件。Windows3.1 提供了多媒体支持和目标连接嵌入等功能，是 PC 机上开发多媒体的最佳环境。

为多媒体而设计的操作系统（OS），要求易于扩充，数据存取与格式无关，面向对象的（object—Oriented）结构，能同步数据流，用户界面直观。例如，Microsoft 公司用 WME（win—dows Multimedia Extensions，即 Windows 的多媒体扩充版），这是在 OS 的层次上支持和增设的多媒体功能。

在 Windows 下，应用程序先同 MMSYSTEM 对话，再由后者推动低层多媒体设备驱动程序某个 MCI 驱动程序。

Windows 的 MCI（Media Control Interface 的缩写，媒体控制接口）可直接控制硬件。对 Windows MCI 的命令包括：MCI—PLAY，MCI—RECORD 及 MCI—STATUS。Windows 多媒体设备驱动程序执行为各类多媒体设备规定的标准功能，例如打开和关闭设备，发送和接收缓冲的连续波数据或单个的 MIDI 信息，设备的音量等。

二、多媒体数据准备软件

多媒体数据准备软件主要包括数字化声音的录制、编辑软件；MIDI 文件的录制、编辑软件；全动视频片段的录制；图像的获取。

1. 数字化声音的录制、软件

Creative Wavestudio——该软件支持多文档接口（MDI）和下拉菜单及

鼠标操作，它能从多种声源录音并将它们混合控制。编辑命令有：剪裁、贴进、回声、淡化、放大和反向等。在硬件许可条件下，录制声音进可选择立体声、单道声。

Creative Soundo*LE——该软件录下的声音文件能采用 OLE 技术插入到任何 Windows 文件中。可以在多个输入源中选择进行录音，在硬件许可的条件下可选择声音压缩方法，可选择立体声或单声道，8 位或是 16 位采样强度，11 或 22 或 44kHz 采样频率。获取的数据直接写入硬盘，录音时间（长度）不受内存大小的限制。

2. 录制、编辑 MIDI 文件的软件

Creative Apprentice for Windows——这是一个对用户友好的图形 MIDI 音序器。用户在不同的形式下观看多个控制图标并编辑 MIDI 音序：钢琴滚筒、五线谱或事件表。用鼠标可以描述速度变化、调整力度或细调弯音。在重放音乐时改变音色分配或改变速度。此外，还能控制数字声音卡，使之与 MIDI 音序同步。

3. 从视频源中获得图像的软件

Creative Video Kit——该软件是随 Creative Video Blaster 发行的静态图像定格获取软件，其主要功能特点：

- (1) 帧捕获
- (2) 活动/静态视频缩放
- (3) 输入声音的软件调节

4. 录制、编辑全动视频片段的软件

Microsoft Video for Windows——这是一套根据 Windows 的全动视频的获取及编辑，并能把获取的 AVI 格式的全动视频片段通过 Windows OLE 技术嵌入到任何的 Windows 实用软件。

三、多媒体编辑软件

多媒体编辑软件又称为多媒体创作工具，它的主要作用是支持应用开发者从事创作多媒体应用软件。

1. 功能

一套实用的多媒体编辑软件，应具备以下功能：

(1) 编程环境。提供编排各种媒体数据的环境，能对媒体元素进行基本的信息控制操作，包括循环分支、变量等价及计算机管理等。此外，还具有一定的串处理、定时、动态文件输入/输出等功能。

(2) 媒体元素间动态触发。所谓动态触发是指用一个静态媒体元素（如文字图表、图标甚至屏幕上定义的某一区域）去激活一个动作或跳转到一个相关的数据单元。在跳转时用户应能设置空间标记，以便能返回起跳点。

多媒体应用经常要用到原有的各种媒体的数据或引入新的媒体，这就要求多媒体编辑软件具有输入和处理各种媒体数据的能力。

(3) 动画。能通过程序控制来移动媒体元素（位图、文字等），能制作和播放动画。制作或播放动画时，应能通过程序调节物体的清晰度、速度及运动方向。此外，还应具有图形、路径编辑，各种动画过渡特技（如淡入淡出、渐隐渐现、划入划出、透视分层等）等能力。

(4) 应用程序间的动态连接。能够把外面的应用控制程序与用户自己创作的软件连接，能由一个多媒体应用程序激发另一个应用程序，为其加载数

据文件，然后返回第一个应用程序。更高的要求是能进行程序间通信的热连接（如动态数据交换），或另一对象的连接嵌入。

（5）制作片段的模块化和面向对象化。多媒体编辑软件应能让用编成的独立片段模块化，甚至目标化，使其能“封装”和“继承”，使用户能在需要时独立取用。

（6）具有良好的开拓性。多媒体编辑软件能兼顾尽可能多的标准，具有尽可能大的兼容性和扩充性。此外，性能价格比较高。

（7）设计合理，容易使用。应随附有详细的文档教材，这些材料应描述编程方法，媒体、输入过程，应用示例及完整的功能检索。

由上述可见，多媒体编辑软件的基本思想是将程序的“底层”操作模块化。例如，用户只需输入图像文件名称和显示坐标即可通过“显示模块”在屏幕上显示图像出来。如果在显示图像时，需要有伴音，则只需再引入一个音频播放的模块。这样，用户在上机时尽可能不借助印刷文档而掌握软件的基本使用方法。这不仅提高了多媒体软件创作的效率，而且为大批非计算机工作者进入这个领域提供了方便。

多媒体编辑软件应设计成操作简便、编辑时易于修改、菜单布局合理。

2. 常用的多媒体编辑软件

此处重点讨论在 PC MS—window 版本上运行的多媒体编辑软件。

（1）Authorware Professional。该软件以 Windows 作为操作环境，系统采用 Objectauthoring 框架，以 ICON 方式来编辑程序，能编排各种类型的多媒体数据，能任意地中断流程以供编辑或执行。

该软件由英国 Authorware 公司推出，用于交互或教育培训。它的创作环境是由多个图符组成了一种图符界面：Display, Animator, Erase, Wait, Decision, Interaction, Calculation, Map, Movie, Sound, Video。这些图符小，只有黑白色，但功能十分强大。通过选择这些图符可构成应用程序结构的流程图。每个图符都有多种功能，例如，Calculation 图符可用于完成教学计算，管理系统变量，还可用于调用包括用户定义过程在内的特殊函数，或跳到其它文件或应用程序上。此外，Calculation 图符还有记录功能，可记录用户的正确答案次数及回答次数，从而判断用户的学习状况。

Authorware Profession 容易学会，无需编程，便于操作，它是开发教育方面多媒体应用项目的一种较好的编辑软件。

（2）IconAuthor——这是运行在 Windows 环境下的一套包括开发全特征多媒体应用软件所需全部工具的多媒体编辑软件。它主要用于教育培训、信息咨询领域。它提供图形流程图模式，50 个设计精美的图标位于屏幕左边，表示所有功能。

（3）Multimedia Toolbook。这是一套用来编辑全特征多媒体应用程序的多媒体编辑软件。它是一个面向对象的开发环境，它并不是专为多媒体应用开发而设计的，但它提供的 openscript 语言容易为初学者掌握，因而制作多媒体软件比较方便。它可提供预先编好程序的图形对象，可以用 Windows 的剪贴功能将图形放在自己的程序之中，加以控制 CD—LD 等设备；它可提供多媒体屏幕原形。

Toolbook 在多媒体 Windows 支持下连接各种多媒体设备，包括影像迭加设备、视频磁盘、录像机等，它可支持声音及音乐等媒体，通过 MCI 命令串在播放后可触发事件。

Toolbook 可显示外办存储的位图文件，并嵌入到多媒体应用中。

四、多媒体应用软件

这是多媒体开发者利用媒体编辑工具组织编排大量的多媒体数据而成为多媒体的最终产品。

多媒体应用软件的 6 个阶段制作流程如下：

(1) 成立研制组确定主题目标，组织人员分工。如美工设计、剧本编写（包括音频和视频剧本）程序设计人员（进行多媒体应用编程）、媒体制作人员（准备录音室、电视制作室等，获取声音、图像的文件，准备对音像媒体元素进行必要的加工）。

(2) 编写剧本。按照主题要求，收集、整理、组织有关的资料，拟定大纲，编写成剧本，加旁白说明。

(3) 数据准备。这个阶段是文字、声音、图像、图形等生成阶段。

设计出分镜头脚本，并生成相应的文本文件。

用创作、转录、效果等方式产生音乐。

旁白说明的录制及速度控制。

把收集的图片输入计算机。

采集编辑全动视频片断。

(4) 利用多媒体编辑软件制作。数据准备完成后，便可按照剧本在编辑软件指导下，组织编排数据，形成产品，完成多媒体应用软件的开发。

(5) 请计算机专家进行程序加工，完善编辑软件。

(6) 产品评审。组织专家评议审定。通过软件的演示和报告，审查是否达到预期目的，进行必要的调整和修正。

目前，我国有一些企业已经开发了几种的多媒体应用软件，例如，通用多媒体图像/图形编辑处理软件系统等。立足国内跟踪国际发展，根据市场需要，开发各种媒体应用软件，以实现中国特色的多媒体技术产业化。

第八章 多媒体信息检索服务

一、多媒体信息检索服务及其构成

多媒体与超媒体信息检索系统是一种人机交互系统，也是多媒体通信系统中一个十分重要的分支。由于多媒体通信系统是一个全新的课题，因而出现了一批新的术语。为便于正确了解这些术语的含义，现将其中最主要的术语解释如下：

· 超媒体 (Hypermedia) ——通过外部树形的链 (Hyperlink) 将多种类型的媒体连成一个集合，这个集合称超媒体。

· 脚本 (Script) ——描述一个显示或输入对象中各对象间关系 (包括相关对象间的时空关系、同步) 的技术规范。

· 脚本软件 (Scriptware) ——实现脚本的软件。

1. 多媒体信息检索服务的系统结构

(1) 信息提供者。信息提供者是多媒体信息检索服务的重要组成部分。它是多媒体和超媒体检索系统中的全部信息的产生者。是它们将多媒体信息产生出来，并以一定规则传送出去。

信息提供者的三个最基本的要求是：有编辑多媒体与超媒体信息的能力；可向数据库上载它产生的多媒体和超媒体信息，并有能力修改数据库；可向下加载检索途径 (Trace)。

(2) 信息库。多媒体信息检索服务系统中，全部信息是存放在信息库中的。由于多媒体与超媒体信息的特殊性，其库结构也将是特殊的，特别是需要从硬盘中大量读取信息的场合。因而，需要特殊算法来满足需要。

信息库的三个基本要求是：可存储多媒体与超媒体信息；能通过交互检索已存储的 M&H 信息；能接收信息提供者的上载数据。

(3) 信息系统管理者。多媒体信息检索服务系统是个很复杂的系统，它不是一个简单的点对点通信，因而在系统中将有一个信息管理者存在。

信息管理者的最基本要求是：管理多媒体和超媒体信息；管理多媒体和超媒体信息目录；管理系统用户；向上装载已处理的检索过程；向下装载未处理的检索过程；计费；接入控制；保护；系统监测。

(4) 信息系统用户。多媒体和超媒体信息检索系统的用户是系统信息的实际消费者，它受信息管理者管理并接入系统，通过多种途径来检索所需信息。

信息消费者的最基本要求是：查阅多媒体和超媒体信息目录，查看系统的有效信息；根据有效的信息目录来检索多媒体与超媒体信息；下载和修改多媒体与超媒体信息；多媒体与超媒体信息的远程查询；上载未处理的检索过程。

2. 多媒体信息检索服务的网络拓扑结构

多媒体信息检索服务可有三种典型的网络拓扑结构。

(1) 点对点。这是一种常用的拓扑结构，用户终端与数据库是点对点连接。引伸开去，在这种结构中也包含多个用户对一个数据库主机的量形连接。由于在这种网络拓扑结构中，一个用户只能查一个数据库，因而这种拓扑结构适用于专网信息系统。

(2) 一点对多点。在这种拓扑结构中，一个用户可接续多个数据库主机，因而也可查找多个数据库的信息资源，这是一种适用于公网信息系统的网

络拓扑结构。在这种拓扑结构中，有一点要强调指出的，这里对不同数据库的路由选择是由终端来实现的（或由终端用户来实现），因而系统中数据库的增减需及时通知全部用户，否则将产生系统对多点的网络拓扑结构差错。

（3）一点对多点（通过接入点设备）。在这种网络拓扑结构中，一个用户可接续多个数据库主机，因而也可查找多个数据库的信息资源，这也是一种适用于公网信息系统的网络拓扑结构。与上一种不同之处是，它增加了一个接入点设备（AP），因而从性能上就大大优于前一种方式了。由于在网络中增加了接入点设备，它将负责对系统中数据库的管理，这样对不同数据库的路由选择已不是用户终端（或用户）的事，而是由接入点设备来自动完成。系统中数据库的增减只需在接入点处登记，而无须通知所有用户；另外，接入点设备还实现对用户终端的控制和管理。因而与上一个网络拓扑结构相比，这一种结构更为实用。

3. 多媒体信息检索服务系统处理的信息

（1）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、二值图像、声音和活动图像。

（2）多媒体和超媒体信息。多媒体和超媒体对象可用它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。多媒体和超媒体表示信息的标准，至少必须拥有下列特点：在表示对象之间可有不同的时间同步算法；具有表示对象间的空间复合的能力和机制；可以用 Hyperlink 去引用外部的表示信息；可以定义用户的不同输入请求；定义对象间的链接，例如事件（Event）和作用（Action）的链接；描述与对象相联系的主题信息，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 对象之中的机制。

（3）脚本信息。脚本信息将选用一些合用的标准来编码和表示。它将用单媒体信息的编码标准和 MH 信息的编码标准。脚本信息是一组特定的用语义关系联系起来的结构化的 MHI（多媒体和超媒体信息）。它需要提供这一组多媒体信息的运作过程及与外部处理模块间的关系。

脚本信息至少需具备如下特点：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 对象的表示；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。

（4）特定的应用信息。上面所述的信息是三类低层的信息，可以由标准来定义和表示。第四类信息是高层信息，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三类的基础之上的。

一个重要的例子是，最终用户工作站执行了 Scriptware 后，产生的返回踪迹（Trace）是随着不同的应用场合而不同的。另一个常用的典型例子是目录（号码簿）信息，通过目录信息可检索到所需的多媒体或超媒体信息，因而目录信息是按照信息类型的不同（如：文档、对象、文件、文本、数据包等）来分类的，并用内在关系互相联系起来。这样，用户就可以在检索所需信息前，先利用目录信息来找到所需信息位置。目录信息就是典型的特定应用信息。

二、多媒体信息检索服务标准及信息交换

1. 多媒体信息检索服务的国际标准和工业标准

国际电联 CCITT 和国际标准化组织 ISO 在多媒体信息检索服务领域已经进行了大量的标准化工作，现在已有一批建议可供使用。另外，在多媒体的开发中，一些工业发达国家为本身设备的开发也制定了一些工业标准。鉴于这些标准已经广泛使用，因而很可能在今后的国际化中被采用。所以，在多媒体和超媒体信息检索服务系统的研究中，国际标准和工业标准都是值得重视的。下面将围绕多媒体信息检索服务的国际标准和工业标准进行讨论。

(1) 单媒体内容数据信息

文本

文本含有三方面内容：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。已有的国际标准有：ISO646；CCITT T.51、T.52、T.61、T.101。目前还有一个新的国际标准正在制定之中，这就是 ISO10646，它基本上是一个 16bit 编码结构。该国际标准将把世界上所有的编码表都纳入统一的编码结构之中。

在已有文本的工业标准中，最著名的是 ASCII 码，这是美国的国家标准，另外还有 CDA 和 RTF 等。

图形

图形编码一般有如下几种方法：镶嵌图形法；动态再定义图形；几何图形法；增量。关于图形编码的国际标准有：CCITT T.101 建议（可视图文的表示层句法，定义了全部图形编码方法）；CCITT T.105 建议（关于增量编码的建议）；ISODP8632（CGM）；ISO7942（GKS）；ISO9637（CGI）。工业标准有：WMF；DXF；PICT；PostScript。

二值图像与静止图像

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。

其国际标准有：CCITT T.4、T.6；CCITT T.81；CCITT T.82；ISODIS11544；ISODIS10918。

工业标准有：GIF；PCX；DIB；TIFF；IFF；JPG。

声音

语音编码技术国际标准绝大部分都是 CCITT 提出的。Rec.G.711 描述了 8kHz 抽样、8bit 量化的语音编码；Rec.G.721 描述了低速率、低质量的语音编码；G.722 则是高码率、高质量的语音编码。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，它采用的是子带编码压缩技术，因而能获得低的码速和高的语音质量。另外还有 T.101 的 AnueX.ODA 将声音引入了文件体系结构中去。在工业界大多数采用 PCM 和 ADPCM 技术。

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）。在运动图像方面的国际标准最主要的是：CCITT T.261；ISO11172（MPEG）。它们均是基于 DCT 变换的压缩编码技术。H.261 是用于点对点实时通信的场合，在可视电话和会议电视场合都采用 H.261。其数据速率为 P*64kbit/s。MPEG 是用于存储读出系统（如信息检索系统、录像机、数字电视等），而不是用于实时点对点通信，其原因是 MPEG 算法将引入固有时延。MPEG1 是家用录像机质量级的压缩运动图像编码，其最高码速为 1.5Mbit/s，这个

速度与光盘 (CDROM) 的读出速度是相适配的 (150KB/s)。MPEG2 是高质量的运动图像编码 (如 HDTV), 最高编码速率为 40Mbit/s。MPEG 还正在致力于研究低速率的运动图像编码 (希望能达 10kbit/s)。除上述两种主要的国际标准外, 在其他场合还可用 JPEG。虽然说 JPEG 是专门为静止图像编码而研究的, 但也可用于运动图像编码。

工业标准有: CD1; DVI, DVI 使用了两种编码格式: RTV、PLV; Macromind 和 Animator 是动画的工业标准。

(2) 多媒体和超媒体信息

近年来, 开始对多媒体和超媒体对象的表示和编码进行标准化。对多媒体和超媒体对象的表示和编码涉及到以下内容:

- 对单媒体数据或多媒体对象的引用和嵌入。
- 时间同步和空间同步。
- 项目信息 (Projector): 项目信息指的是与内容有关的显示属性信息。譬如对声音内容的信息来说, 它的项目信息指的是音量、声道等属性信息。
- 演奏方式 (Rendition) 信息: 演奏方式信息指的是一类与媒体直接相关的信息。譬如对时基 (Timebase) 媒体来说, 可以有“速率设置”的 Rendition 信息; 对可见 (Visible) 媒体来说, 可以有“尺寸选择和设置”的 Rendition 信息。
- 将单媒体 (Content) 信息聚焦和复合起来, 构成多媒体对象的方法。
- 多媒体对象间的超级链 (对象间用链接关系链起来)。
- 输入信息 (客户与系统的交互) 等。

目前, 多媒体和超媒体对象的表示和编码已有多种, 最具代表性的是 MHEG; HyTime; Hyper ODA; MPEG 等。

MHEG

MHEG 是 ISO/IECJTC1/SC29/WG12 和 CCITTSG8/

Q11 组成的联合工作组的研究成果。该标准是用于实时交换的多媒体和超媒体信息对象的表示法。

HyTime

HyTime 是 ISO/IECJTC1/SC18/WG8 的工作成果。Hy-Time 主要研究的有: 多媒体同步的表示, 超媒体在文档内或文档间的链接。HyTime 适用于综合的开放型多媒体和超媒体信息系统, 以及在开放环境下的文档交换和操作管理。

HyTime 标准的最近版本是 1992 年版, 标准号是 DIS- ISO10744。HyTime 与 MHEG 在很多方面是一致的, 但它们的使用方法和应用环境是不同的。涉及到文档的处理和交换可用 HyTime, 涉及到对象的处理与交换则是用 MHEG。Hy-Time 从标准的设计角度出发, 试图用于枢纽间的信息交换, 而不是专门用于实时信息交换。

HyperODA 与 ODA

HyperODA 与 ODA 是 ISO/IECJTC1/SC18/WG3 和 CCITTSG8/Q27 的成果。HyperODA 与 ODA 主要是用于文档的交换与操作管理。HyperODA 是一种围绕的体系结构 (外部元素必须根据 ODA 标准化), 仅限于 ODA 环境下的文档交换。ODA, 即开放文件体系结构, 是用于交换文件的描述方法。ODA 不支持多媒体信息的处理与交换, 但可以在 ODA 的内容 (Content) 中引用其他方式编码的多媒体信息。Hyper-ODA 是基于 ODA 的, 但能处理和表示多媒体和超媒体信

息。

其他标准

包括 MPEG 和 MPEG1 (ISO11172)。MPEG1 定义了 MPEG 中交织的数据码流 (图像、语音和文本) 的帧结构, 这是一种表示时间同步特性的信息表示法。当然, 这种表示法本质上是线性的, 只适用于表示比较简单的信息, 主要用于表示运动图像及与其相关联的声音数据。在 MPEG 的帧结构中, 还保留了若干比特供文本交互用。

工业标准

· RIFF: 是包在其他格式文件外的一种包封结构, 在 RIFF 的标记上, 标有分量的类别和尺寸。这些被 RIFF 包起来的分量, 则是用独立有效的常用的文件格式来表示。

· MCI: 提供一种高层的命令接口 (API), 用于多媒体系统的源文件中, 它使用的编码格式为 RIFF。

· Quicktime: 用来描述并行的媒体数据流和它们之间的同步。Quicktime 是将多媒体与格式化的 Movie 环境相联系的方法。还有专门用于 PC 的 Hypertext、Hypermedia 等格式, 但都是专用的, 无交换能力。

(3) 脚本层

这由 ISO/IECJTC1/SC18CCITTSG1/Q17 和 ISO/IECJTC1/SC29CCITTSG8/Q9 研究, 至今尚无国际标准。

工业标准有: HyperCard、Toolbook 和 Scriptx。它们都使用了能够表示超级结构 (Hyperstructure) 的通用编程语言作为通用的 Script 语言。

2. 信息交换

从显示的观点, 信息交换 (Information Interchange) 可分为两大类: 一类为实时信息交换, 另一类为非实时信息交换。

(1) 实时信息交换

实时信息交换指的是边传输边进行显示的一种信息交换方式。实时信息交换的国际标准有: CCITT.170 建议; CCITT.175 建议; CCITT.176 建议; CCITT.41X 建议; CCITT.43X 建议。

在 T.170 建议中, 描述 AVI 中的某些业务需要, 并定义了“ ”、“ ”、“ ”、“A”和“B”五种接口。

T.175 建议描述“ ”接口 (从信息消费者到主机), 它包括文件传送、选择、浏览、远程执行等。

T.176 建议描述“B”接口, 这是一个实时信息交换的协议, 它覆盖了 MHI 的 RTI。

协议“ A ”为应用层协议, 目前已有 ContentData 的 RTI 的“ A ”协议。这些协议有: T.102、T.103、T.105、T.43x 和 T.41x。

一个重要的工业标准是 X.11.X.11 是本地 (局域) 网络客户/服务器 (Client/Server) 协议, 用于图形和窗口信息。

(2) 非实时信息交换

非实时信息交换指的是, 终端必须等到一个单元的信息 (文件或数据库记录) 完全传送完毕后, 才能显示的一种信息交换方式。非实时信息交换的标准有: FTAM; RDA; DFR; DTAM。FTAM 从本质上来说, 只能操作和传送完整的文件, 然而, 对其他类文件, 它也可操作和传送其中的一个对象, 并且还能对文件的存储执行某些管理功能。RDA 提供从数据库中将其中一项 (如,

一条记录) 传送出去。这个标准分为多个部分: 首先, 它定义了通用的服务和协议; 其次, 分别为不同的数据库体系结构和数据库操作管理作了专门的定义。目前, 只有一个专门定义已完成, 这就是 SQL 数据库语言。当然, 今后还将发展另外的数据库语言和组织。特别是存储对象结构的文本数据库将会用 SGML 来描述。

DFR 很像 FTAM 和 RDA, 但只能用于文档操作。文档由其属性来描述, 属性的用法与 ODA 相近。DTAM 是用于存取和操作管理 ODA 文档的。

三、多媒体信息检索服务硬件技术

1. 终端设备

用于多媒体信息检索服务的终端, 既可以用通用的计算机, 也可以用专用的终端设备。

(1) 总线。

在多媒体设备中, 除中央处理器 (CPU) 外, 还需辅助以特殊处理器, 这些处理器通过总线和附加的控制器将内存储器 and 外部设备连接起来。由于总线速度受限而制约了设备的性能, 目前解决的办法是采用局部总线。

大多数个人计算机和 workstation 都提供扩展槽, 扩展卡可插入这些槽中与系统总线相连接。由于不同的计算机平台有不同的总线系统, 因而一般来说, 扩展卡不具有通用性。

(2) 本地存储。

本地存储指的是计算机或 workstation 的本机存储, 一般有四类: 工作存储器 (RAM)、外部永久性存储器 (硬盘)、备份存储器 (磁带)、可移动式存储器 (软盘、CDROM、可拆卸式硬盘)。

对存储器的要求为: 存取时间/寻道时间; 读写时间; 存储容量; 可移动性; 可写性: 存取/编码和文件结构; 模拟存储。

(3) 多媒体的显现。

多媒体的显示输出与用户的感官要求密不可分。因此, 在研究多媒体的显现硬件设备时, 首先需研究一下, 用户感觉对不同媒体的显现要求和限制。

首先讨论视觉的限制, 根据实验可知, 人的视觉分辨率为 $1/120$ 度, 这相应于在一个人的正常视距内, 屏幕上分辨力 380dpi (点数/英寸)。从目前或将来可及的技术来看, 视频显示器只能达到上述分辨率的 $1/10$ 。人眼对亮度的辨别率为 240 ~ 300 级灰度, 而且人眼对不同颜色的分辨率是不相同的。目前采用的 24bit 量化 (红、绿、蓝各 8bit, 分别有 256 种灰度) 是与人眼适配的。在大多数场合下, 色彩比分辨率更重要。

再来讨论人耳的听觉。正常人的耳朵可以分辨出 80db 强度的变化 (16bit 可以表示 96db 的动态范围)。人耳的感知频率为 20Hz ~ 20kHz。根据纳奎斯特抽样定律, 抽样频率必须大于 2 倍的最高频率, 因此对音质要求很高的场合 (能覆盖人耳感知频率的全程), 抽样频率为 44.1kHz, 对每个抽样进行 16bit 量化。对于立体声来说还分左、右 2 个声道, 这样未压缩的数字化立体声的码率为 $44.1 \times 4KB/s$ 。在研究了图像和语言的显现问题后, 下面来讨论现有多媒体硬件设备。

显示器: 现有的显示器为: 640×480 , 768×512 , 800×600 , 1024×768 , 1024×1024 , 24bit 量化 (RGB 各 8bit)。它们在显示卡上各自带的存储器容量为 900、1152、1406、2304 和 3720KB。为了减少存储容量, 通常用调色板

法 (262114 种选择, 共有 1680 百万种颜色) 来显现 256 种颜色。也有 RGB 各 5bit 量化, 可直接表示 32768 种颜色。

声音: 为了避免混叠, 抽样频率应为所需最高频率的 2 倍。语音编码器采用的压缩编码算法通常有: PCM、DPCM、ADPCM 和线性预测编码 (LPC) 等。

目前市场上已有的声音卡: Multi Sound Board; ProAudio Spectrum; Mpc Audio One; Sound Blaster Pro。

静止图像的硬件有: JPEG Board; i750; Kodak PhotoCD。

视频卡有: C Cude; CD1; DVI; JPEG Decoders。

(4) 目前与多媒体接口的软件平台有 Microsoft Windows、Macintosh System、Xwindow System。

2. 网络环境

(1) FDDI

FDDI 是以令牌协议为基础的高速局域网, 其速度为 100Mbit/s。

(2) MAN (MetropolitanAreaNetwork)

MAN 网是用于填充局域网和公用广域网之间的空白, 这类网络的 IEEE 标准是 DQDB。

(3) ISDN

ISDN 是一种广域网, 它提供两种接入类型: 基本接入速率为 2B+D; 一次群接入速率为 30B+D。另外还有: H0 信道 (384kbit/s); H11 信道 (1536kbit/s); H12 信道 (1920kbit/s)。

(4) BISDN

BISDN 即宽带 ISDN, 它采用的是 ATM 规程, 速率可达 155Mbit/s。

(5) 其他

其他网还有 PSTN、DBN (DigitalBroadcastingNetwork) 等。

3. 主机

多媒体信息检索服务对主机的要求很高, 主要有如下三方面: 向多用户传送大量的实时数据; 主机有很大的存储容量; 有相应强的计算能力。

这里值得一提的是主机的硬盘速度, 因为在多媒体信息检索服务中硬盘的存取速度往往是系统的瓶颈, 它需要与通信网的通信速率相匹配。在目前, 如工作在一次群 NISDN 环境下 (30B+D1920kbit/s), 与目前硬盘的存取速率 (8Mbit/s) 是基本适配的, 不至于成为瓶颈。当通信环境进入 BISDN, 通信速率将达 155Mbit/s, 那么硬盘的存取速率在 200Mbit/s 以上才能适配。有幸的是, 据预测, 硬盘在 10 年内可望达到这个存取速率。

四、多媒体可视图文系统

多媒体可视图文系统是一种远程多媒体信息检索系统, 它通过公用分组交换网、局域网和数字数据网将分布在不同地域的多媒体信息源组织起来, 通过电话网、综合业务数字网和用户专线向广大用户提供多媒体信息服务。它是一种公用的开放型的信息服务系统。多媒体可视图文系统是现在业已进行商业运行的可视图文业务网的升级系统, 它对现有系统是向下兼容的。它是 90 年代初国际上刚完成标准化的最新一代的多媒体远程信息检索系统。

多媒体可视图文系统由用户终端设备、多媒体可视图文接入点设备、多媒体信息数据库和业务管理中心四大部分组成, 为了与其他网相连构成广域

的高速信息环境，从而获得更多的信息，网内还有三类设备：专用集线器（HUB）与 ATM 相连、路由器与其他局域网相连、网关与因特网相连。

用户终端是用户与系统的接口，用户是通过用户终端接入到系统中去，并通过用户终端获取多媒体信息。多媒体可视图文系统中用户终端有三种：一种是电话网上的用户终端，用户终端通过电话网与系统相连，其通信速率为 14.4kB/s；第二种是综合业务数字网上的用户终端，它通过电话线与综合业务数字网的交换机相连，其通信速率为 128kB/s；第三种是专线接入终端，它通过专线直接和系统相连，其通信速率是 64kB/s。尽管这三种用户终端的接入方式不同，接入速率也不同，其得到的服务则是相同的。它的人机界面为 Windows 界面，因而可以提供复杂的人机交互方式来获得真彩色的静止照片、无失真的二值图片、手绘真迹图形和文字、各种图表、曲线和文字，还有声音和合成音乐。终端所得到的声、文、图信息是以同步方式经多通道向用户提供的。

多媒体可视图文接入点设备是系统中的核心设备，它实现对用户终端的管理、对数据库的路由寻找、计费、统计等功能。它还要与业务管理中心进行通信，接受业务管理中心的各项管理命令，并向业务管理中心报告状态信息。多媒体可视图文接入点设备提供五种网络接口，其中三种网络接口是面向用户的，这三种网络接口是：电话网通信接口（通信速率为 14.4kB/s）、综合业务数字网通信接口（通信速率为 $D=16\text{kB/s}$ ， $2B=144\text{kB/s}$ ）和专线用户接口（通信速率为 64kB/s）。另外两种网络接口是面向数据库的，这两种网络接口是：分组网通信接口（通信速率为 64kB/s），它主要是用于和文本数据库、脚本数据库以及业已运行的所有可视图文数据库相连，以获取信息；局域网通信接口（通信速率为 10MB/s），它主要是用于和图像数据库、声音数据库相连，还可以通过专用集线器与 ATM 交换系统相连，或通过网关与因特网相接以提供宽带广域通路。

多媒体可视图文数据库有四类，一类为文本数据库，这类数据库与现在已经运行着的可视图文数据库没有本质上的不同，不同之处在于它提供超级文本链接，它主要是提供文字和简单图形信息内容；第二类数据库为脚本数据库。这是一类很特殊的数据库，它主要是为了能实现声、文、图同步而专门设计的数据库；第三类数据库是图像数据库和复合对象数据库，它存放经过压缩的彩色图像、不失真二值图像等信息。这是面向对象的数据库，它还存放复合的多媒体对象信息；第四类数据库是声音和音乐数据库，它存放经过压缩的声音和音乐，它也是面向对象的数据库。多媒体可视图文系统中，声、文、图是分别存放在不同的数据库中，在终端程序的调度下，经由不同的信息通道，由数据库将数据信息传送给终端，这些信息在终端同步起来，向用户提供多媒体信息服务。

业务管理中心是实行全网业务管理的，多媒体可视图文系统是一种业务统一管理、计费统一管理的集中式管理信息系统。业务统一管理指的是数据库的登记、删除、系统中数据库目录的编制是统一管理的。信息系统的业务统一管理给用户的使用带来极大方便，用户对在网的信息资源，可从多媒体接入点设备提供的数据库目录中一览无遗，网中接入点设备的数据库目录正是业务管理中心编制并下载下来的。业务管理中心除了实现全网业务管理外，还对全网的接入点设备进行监控，并向全网采集统计信息，从而确保全网的服务质量。

多媒体可视图文系统中的网关是专门为因特网设置的，多媒体可视图文终端用户通过网关进入因特网获取信息，并可和因特网用户交换信息。

多媒体可视图文系统是严格遵循国际标准来实现的。在通信系统中，国际标准化是很重要的。多媒体可视图文系统是现有可视图文系统从通信业务层次上向高技术发展的延续。因而在现有可视图文系统中应用的全部国际标准，在多媒体可视图文系统中将全部采用。此外，在多媒体可视图文系统中还采用了如下一些国际标准。

1. 增强型可视图文人机接口 (VEMMI)

可视图文系统中，用户终端与系统的界面是简单的问答式交互界面，用户终端可采用基于 DOS 的接口方式。在多媒体可视图文系统中，人与系统的接口大大复杂化了，必须使用 Windows 的人机接口方式，国际电联提出的 VEMMI 协议就是这样的一种人机接口。要指出的是，Windows 是 PC 机本机的人机接口，而 VEMMI 则是一种具有 Windows 接口功能、终端与远地数据库之间的人机接口协议，这是一个很复杂的协议，多媒体可视图文系统中实现了这一协议。

2. 多媒体可视图文系统的应用层协议

多媒体可视图文系统中，用户终端与系统中信息流的交互再不会是简单的单一信息流，而是多种信息流的工作状态。根据脚本中对信息流的安排，或者复合对象中各种不同信息流间的关系，多种信息流的开发工作情况是要考虑的。多媒体可视图文的应用层协议正是针对此来设计的。它除了有一条主信道外，还可有若干与主信道同时工作的辅助信道，它们可以用于同时传送文本、语音、音乐和图像信息。多媒体可视图文系统中采用了这一建议并实现了它。

3. 多媒体可视图文系统中的表示层句法

多媒体可视图文系统是一种资源受限的通信系统，即通信速率受限和终端存储量受限。为了要在资源受限的通信系统中有很好的效率和很好的服务质量，系统中大量采用先进、高效的编码技术，如文本、图形编码继续沿用可视图文表示层句法；曲线采用区域增量编码技术，这是 ISO 和 ITUT 两大国际标准化组织从最优的六种增量编码方案中优化出来的，它对手写真迹文本有最佳的压缩效率；二值图像采用 JBIG 算法，这是基于自适应算术编码算法基础之上，是目前最佳二值图像压缩算法；彩色静止图像采用 JPEG 算法，这是基于离散余弦变换 (DCT) 算法，理论已经证明，DCT 是仅次于卡-洛变换的准最佳算法，加之它有快速算法，因此是目前国际标准化了的一种很好的算法；声音采用两种压缩编码算法，其符合 ITUTT.101 建议 AnnexE，码率为 8kb/s 和 16kb/s 两种语音；音乐采用 MIDI 编码，这是一种合成音乐。由于通信信道速率受限，没有引入运动图像的压缩编码算法。

4. 多媒体可视图文系统中的同步

同步是多媒体通信系统中的一个极为重要的特征。在多媒体通信系统中，同步可以在三个层面上实现，这三个层面分别为：帧级、表示层级和应用层级。多媒体可视图文系统中不引入运动图像，因而没有帧级层面的同步。它在应用层级进行了同步，即应用脚本技术来实现的；它在表示层级进行了同步，这一同步是在构成复合对象时引入时间同步和空间同步机构来实现的。

5. 多媒体可视图文系统中的超文本技术多媒体可视图文系统中，大

量采用超级文本技术

超级文本是一种非线性的文本组织方式，特别适合计算机中文本的组织。多媒体可视图文系统中利用敏感区（图像）和敏感字段（文本）作为超级链，以此来链接内容相关的文本从而构成超级文本。由于大量采用了 1994 年以后的国际标准，因而多媒体可视图文系统在技术上是先进的，也确保了系统良好的继承性和可发展性。多媒体可视图文系统中还大量采用面向对象技术。用户接口、用户终端、接入点设备和数据库都采用面向对象技术。用户与系统的接入协议（VEMMI）是面向对象的接口协议；用户终端的体系结构是面向“脚本”、面向对象的；数据库管理系统也是面向对象的。

多媒体可视图文系统的应用领域极为宽广。从大类来分，可用于如下三个方面：

（1）多媒体信息检索

这是多媒体可视图文系统的最主要功能，用户通过终端提供的多种人机交互方法，向数据库索取各种数据资料。在用户和数据库间进行交互式人机通信过程中，用户只读取数据库中的数据，不改变数据库中的任何内容。这类业务很多，如电子图书馆、电子新闻、电子报纸、电子杂志、法律、文化艺术、体育消息、旅游资料、百科全书等。

（2）多媒体交易型业务

这是多媒体可视图文系统的重要功能，用户通过终端提供的多种人机交互方法，不仅向数据库索取各种数据资料，还要修改数据库的内容。在用户和数据库主机间进行的人机通信过程中，用户对数据库的操作是一个“既读又写”的过程。这类业务的典型应用有：电子银行、电子购物、证券期货交易、房地产交易、机票预定和客房预约等。

（3）多媒体消息业务

多媒体消息处理业务是一种存储转发型业务，是一种人—机—人的通信业务。当用户 A 要将多媒体信息传送给用户 B 时，他先通过系统将信息存在某一个确定的数据库主机中，用户 B 通过系统从该数据库主机中获取由 A 放入的信息，从而实现多媒体信息由用户 A 向用户 B 的传送过程，数据库主机起到存储转发的功能。这类业务的典型应用有：多媒体电子信箱和多媒体文件传送。

多媒体可视图文是新一代的信息检索系统，它采用了大量新技术和全面采用了最新国际标准，它将是今后多媒体远程信息检索系统的主流系统之一。

第九章 多媒体通讯

一、多媒体通讯的类型

顾名思义，所谓多媒体通讯，就是把声音、图像等多种信息载体传到远方，或者说通讯的双方不但可以听到声音、传送文件，甚至还可以看到对方。

从通讯的眼光来看，这种多媒体通讯是继电报、电话、传真以后的新一代的第四代通讯手段。

当今信息社会的发展，对通讯提出了更高的要求，要不但能听到对方的声音，还要能看到对方的面貌。这就是可视电话的诞生。而多媒体通讯，则又进一步增加了新的功能。它和可视电话有什么不同呢？它的最大不同是多媒体通讯把多媒体电脑和通讯结合起来，不只是可以同时传图像和声音，更重要的是它具有一切电脑所具有的功能，而且把双方的电脑结合了起来，所以这种多媒体通讯称为“个人间会议”（Personal Conferencing），也简称为 PC。

1. 电话会议

在进行通话时，双方不但可以采用手机，而且还可以采用微音器和扬声器，这样就可以不只供一人使用，在房间内的其他人也可以参与讨论。在附加一些设备以后，还可以进行一点对多点的通话，各点之间也可以互相通话，所有这些都未必经过电话局就可以实现。

由于语音信号是数字化的，而且是经过压缩的，所以很容易进行数字加密，因而全部电话会议都可以在保密的情况下进行。这一特点，使其特别适用于军用或其它要求保密的场合。

2. 文件会议（Document Conference）

多媒体通讯可以在双方的显示屏上显示同一个文件，这个文件可以先是从甲传给乙，或是再从乙方传给甲，总之是实时双向的。而且双方可以就同一文件展开讨论，所以，也是交互式的。而且由于它是和计算机联结在一起的，所以它可以利用计算机的一切优点，例如文件的储存、检索、修改、打印等一切功能。所有这一切，都不知比普通传真机要方便多少。最重要的是，因为双方可以实时地就某一文件进行讨论、修改，最后还可以签字画押，一切复杂的需要面对面的讨论过程，都可以在短短的十几分钟的文件会议中解决，免除了从甲地到乙地的长途跋涉，提高了效率。

3. 电视会议（Video Conference）

这种个人间电视会议不同于正式的“电视会议”。后者适用于大型正式会议，需要有多台摄影机等或是遥控摄像机、大屏幕显示连续的活动画面等……。这种个人间的电视会议，主要用于看到对方。然而双方在交谈时可看到对方是很重要的，因为这可以增加讨论时的亲切感，看到对方的面部表情也是一种很重要的交流方式。

正式的电视会议系统需要数字线路，例如美国的 T—1 线路、X.25 数字线路，或是 ISDN 线路，所有这些并不是到处都有的，而且租金昂贵，不能和普通的电话线路相比。

个人间电视会议则可利用电话线，它通常有一种“快摄”功能，所谓“快摄”就是你可以先把摄影机对准你所要传送的物体或景象，只要按一下鼠标上的键，就可以把一幅高分辨率（640×480）的彩色图像在十几秒的时间里传送到对方。这在很多场合是十分有用的。例如，在商品贸易中，可以利用

这一功能来展示不同的样品给对方看。

二、在多媒体通讯中的关键技术

由于多媒体通信是一项综合性技术，涉及多媒体、计算机及通信等许多领域，因此通信技术十分繁杂，牵扯到的技术十分广泛。如：系统平台技术，包括数/模转换、音/视频信号的压缩及解压缩、音/视频信号的抓取和播放、音/视频信号的混合及同步、数字信号处理、VGA 和 TV 相互转换技术等；多媒体技术，包括多媒体操作系统、多媒体数据库管理系统、多媒体文件存储等；多媒体网络技术，包括网络管理技术、高速网络协议、开放式文件结构、不同网络间的传输技术、ISDN 通信技术等等。这里仅谈谈其中几个关键技术。

1. 音/视频信号的压缩技术

由于要传输多种媒体信息，多媒体通信的数据量极大而且应用中又需要实时处理，因此不可避免地要对所传输的多媒体数据进行压缩。多媒体数据的压缩/解压缩技术可以说是多媒体信息处理的瓶颈。

(1) 视频压缩技术。原始信源数据存在大量的冗余，在多媒体通信技术中，图像和视频信号的压缩十分重要。压缩的方法有多种，可以分为两种类型：有损压缩和无损压缩。无损压缩利用数据的统计特性来进行数据压缩，可以完全恢复原始图像而不引起任何失真。典型的编码有 Huffman 编码、行程编码、算术编码和 Lempel—Lev 编码。无损压缩的压缩率一般为 $2^{-1} \sim 5^{-1}$ 。为进一步提高压缩率，可以使用有损压缩法，这种方法不能完全恢复原始数据，而是利用人的视觉特性使解压缩后的图像看起来与原始图像一样。主要方法有预测编码、变换编码、模型编码、基于重要性的编码以及混合编码方法等。压缩比随着编码方法的不同差别较大。

图像、视频的压缩目前主要有三种国际标准可以遵循：ISO JPEG、ISO MPEG 和 CCITT H.261。JPEG 用于静止图像的压缩，适用于黑白及彩色照片、传真和印刷图片；H.261 标准的制定是为可视电话和电视会议服务，因此，标准中建议的视频编码算法有实时处理的能力，延时控制在最小。MPEG 用于动态图像的压缩。其中第一个阶段的目标 (MPEG—I) 是以约 1.5Mb/s 的速率传输电视质量的视频信号，亮度信号的分辨率为 360×240 ，色度信号的分辨率为 180×120 ，每秒 30 帧。MPEG 第二阶段的目标 (MPEG—II) 是对每秒 30 帧的 $720 \times 572 \times$ 分辨率的视频信号进行压缩。在 MPEG—II 的扩展模式下可以对 1440×1152 的视频信号进行压缩编码，因此可以作为对高清晰度电视的压缩编码方法。

(2) 音频压缩技术。音频信号是多媒体信息的重要组成部分。音频信号可以分成三种：电话质量的语音、调幅广播质量的音频信号和高保真立体声信号，它们的频率范围分别是： $200\text{Hz} \sim 3.4\text{kHz}$ 、 $50\text{Hz} \sim 7\text{kHz}$ 、 $20\text{Hz} \sim 15\text{kHz}$ 。

音频信号的压缩方法也有多种。无损压缩法包括不引入任何数据失真的各种熵编码。有损压缩法又可分为波形编码、参数编码和同时利用这两种技术的混合编码方法。波形编码利用抽样和量化过程来表示音频信号的波形，使编码后的音频信号与原始信号的波形尽可能匹配。它主要根据人耳的听觉特性进行量化，以达到压缩数据的目的。波形编码的特点是在较高码率的条件下可以获得高质量的音频信号，适合于高保真语音和音乐信号。参数编码的压缩率很大，但计算量大，保真度不高，适合于语音信号的编码。混合编

码介于波形编码和参数编码之间，集中了这两种方法的优点。

三种常用音频信号分别有不同的压缩标准。电话质量的语音信号压缩编码的一些主要质量评判使用 MOS 标准，它是利用多人打分的平均值衡量语音质量的一种主观评估方法，满分为 5 分。

调幅广播质量的音频标准有 CCITT 制定的 G.722 标准。G.722 采用子带编码方法，能够在每秒 8kB 的存储量下给出相当好的音乐信号，很适合于需要存储大量高质量音频信号的多媒体系统使用；目前国际上比较成熟的高保真立体声音频压缩标准为“MPEG 音频”，MPEG 音频压缩技术的传输速率为每声道 32kB/s ~ 448kB/s。

2. 高速网络

对于多媒体通信而言，高速网络是基石，否则多媒体信息的特性将很难实现。现在市场上已有 FDDI，传输速率在 100Mbps，是高速 LAN 的一种；蓬勃发展的 100Mbps 的 Ethernet 也是高速网络的新产品；宽带综合业务数字网 (B-ISDN) 被许多人预言为未来高速网络的主干。

FDDI 可以说是高速局域网的先驱，它是用于数据传输的通信网络，可提供实时传输，但它没有考虑所传送的媒体的数据特性，不能为不同的媒体数据选择不同的协议，也不能动态分配带宽；100Mbps 的 Ethernet 既能处理多个 LAN 用户的同时需求，也可以应付高性能 PC 及其复杂应用程序带来的特性高峰。当突发传输成为关键因素时，选择快速以太网是工作组的适当方案；ISDN (综合业务数字网) 的出现大大改善了网络条件。它可以提供标准接口，可以支持多种通信业务，但服务内容有限，带宽也不很高，对媒体特性也仅从带宽上考虑，很难满足复杂多媒体应用的要求。B-ISDN 克服了 ISDN 的局限性，在用户网络接口上至少能提供 H1 (135Mbps) 以上的接口速率。采用异步传输模式 (ATM) 进行高速分组传送，极大地提高了网络的通信处理能力。

3. 多媒体通信网的交换技术

交换是网络的核心。网络交换技术是保持多媒体通信网络在扩展后的性能及可管理性的一个新兴方案。ATM 交换技术被 CCITT 定为 B-ISDN 的基础。ATM 是在光纤大容量传输媒体的环境中分组交换技术的新发展，它所传送的单元长度固定。ATM 交换速度比路由器快。采用 ATM 技术的通信网可支持语音、数据以及图像传送、提供专用高带宽以及可预测性能，可直接支持多媒体计算机。

4. 用户接入网技术

用户接入网在多媒体通信中起着十分重要的作用，接入网的发展也是当前国际上的一个主要问题，即常说的“竞争在最后一英里”。在多媒体通信中，有三种用户接入网可供选择：

(1) 非对称数字用户环 ADSL (Asymmetric Digital Subscriber Line)

ADSL 是在一根双绞铜线上传输高速数字信息。它的传输能力在上行、下行两个方向上是不对称的。上行比特率为 16kB/s，下行比特率为 1.5 ~ 8MB/s。这是由于一般情况下，查询语句很短而被查内容很多，在网上跑的信息量是非对称的。ADSL 的基本原理是采用了多载波技术，最大特点是利用现有电话网络开展宽带视频服务。

虽然现在“光纤到户”“光纤到路边”讲得很多，其实这不是在任何情况下都是一种最佳的选择。在最靠近用户的最后一段通信线路内铜缆技术仍大有可为。

(2) 光纤同轴电缆混合网 HFC (Hybrid Fiber Coaxial)

HFC 由光纤节点经由同轴电缆送至用户,有足够宽的带宽。这是一种通带 (Passband) 调制传送方式,在有线电视网中易于与有线电视并存兼容。在发达国家,有线电视 (CATV) 已相当普及,约有 60% ~ 90% 的家庭的有线电视接到了 CATV 网上。我国情况很特殊,我国有 2 亿多台电视机,远远超过电话机,而且连接到 CATV 网上的家庭数目增长非常快。对用户来说,利用有线电视网接入更大范围的通信网是一个很好的手段。发展 HFC 的思路就是让小范围的同轴电缆网 (CATV 网) 和大范围的光纤网结合起来。

利用 HFC 技术已经实现将 PC 接入因特网,它很有可能发展成为宽带通信网的主体,并与多媒体计算机结合实现多媒体通信。

(3) 光纤用户环路 FITL (Fiber In The Loop)

光纤用户环路以其频带宽、可传高速数据甚至高清晰度电视并能适用于 B-ISDN 和信息高速公路的优点引起广泛重视。但是 FITL 中的光纤到家庭 (Fiber To The Home) 成本昂贵,限制了它的大规模应用。在光纤用户环中,有前景的是光纤到大楼 (Fiber To The Building) 和光纤到路边 (Fiber To The Curb)。

5. 多媒体通信协议及标准化

通信对于多媒体信息数据的标准化的要求比独立的多媒体计算机更加严格,因此应采用国际标准。近几年来,多媒体通信协议的标准化工作是国际通信界讨论与研究的最热门话题之一。目前关于多媒体通信协议种类很多,但尚无成熟的标准。国际电信联盟标准化部门 (ITU-T)、国际标准化组织 (ISO) 和国际电工委员会 (IEC) 正积极通过协作来制定通用标准。有人预测目前因特网逐步采用的 RTP/RTCP (实时传输协议/实时控制协议) 有可能发展为通用标准。

第十章 多媒体网络

一、多媒体网络的概念及环境组成

1. 多媒体网络的概念及特性

利用现有或专门的网络来传输多种媒体信息就构成多媒体网络，多媒体网络系统本质上也是一种计算机网络，可以是局域网，也可以是广域网。它与普通局域网的主要区别在于能够对多媒体数据进行获取、存储、处理、传输等操作。多媒体网络应具备以下特性：

(1) 实时性。即数据传输不能有延迟。由于网上传输的多媒体信息主要涉及人的听觉、视觉等，具有很强的时间相关性和连续性。因此，必须做到能够及时地获取、传输、显示各种相关的信息。例如用户利用可视电话进行交谈，要求能够实时地听到对方的声音和看到对方的影像，不能有太长的延迟，否则通话将无法进行。

(2) 同步性。多媒体网络用户应能同步地显示图、文、声、像等信息，把完整的信息显示给用户。

(3) 高速性。能够达到多媒体的声音、视频等巨量数据的传输要求，它是实时性的基础。

(4) 综合性。多媒体网络节点应能够同时处理多种媒体信息，并同时显示多种媒体信息以实现多媒体数据处理。

2. 多媒体网络的环境组成

网络多媒体环境的组成包括：应用程序、影像内容、存储设备、视像服务器、具有解压缩及语音能力的个人电脑等。另外，高速网络也是其中一部分。

视像质量的好坏取决于两个主要因素：一个是每秒钟传送的帧页数目 (Frame/SEC)；另一个则是视像内容的大小。在每秒钟传送的帧页数目较少的环境，可将其用于视像会议中，尤其用于远程通讯连接。人们的眼睛可将每秒30个视像帧页看成足够连续的影像，但每秒提供的30个视像帧页并不代表传输带宽。

编码方法将决定30个视像帧页的内容及其组合方式。目前常用的编码/解码方法包括：DVI、Indeo、QuickTime、H261及Motion—JPEG，而ISO11172 (MPEG标准)为最常使用的标准。

MPEG是采用不同位元传输的非对称式影音编码/解码的方法，且MPEG提供量化及压缩标准。现在有多种规格解释，其中以C—CUBE为首要。标准MPEG传输速度与传输质量的关系为：1.152MB/s、3.45MB/s为影片/广播，4.608MB/s为实况转播，8.064MB/s为现场、14MB/s为HDTV。

在视窗环境中，MPEG影像虽然可以全屏播放，但并非真正支持全屏幕视像播放。因为视像经过解压缩后，展开并运用内插方式填补画面缺空，所以在对角线上会产生锯齿状。

MPEG—1推出至今已两年多，但其需要相当高的CPU运算能力。位于加州的C—Cube公司已开发出低价格的MPEG压缩芯片，以提高相关MPEG产品的开发。例如Sigma公司的Reel—Magic个人解压缩接口卡；Opti Vision及OptiBase的影音压缩卡等。但是，有关MPEG编辑工具以及视像数据库的应用仍然十分缺乏。值得高兴的是，目前已有相当多的厂商投入生产MPEG产品的行列。

在多媒体应用中，语音的输出是相当重要的一环。MIDI 是串行输出端口的硬件规格及音乐数据文件建立的标准，其定义电子合成设备、键盘以及其它 MIDI 设备间的通讯，使得语音与视像可方便地集成起来。

未经压缩的全屏幕视像需要大约 27MB/S 的带宽，普通光盘机具有 150kB/S 的传输速率，若需从光盘机中得到电视质量，则需经过 180 倍的压缩，而 MPEG 可达此压缩要求（可达 200 倍压缩）。经 MPEG 压缩编码后的数据仅需 1.2MB/s 的网络带宽，相当于每分钟 9MB 的存储容量。一张光盘可以存储 74 分钟的 MPEG 全屏幕压缩影音。在传统以太网络的每一网段上，理论上最多可支持六个多媒体使用者，但由于 CSMA/CD 协议限制，因此并不可行。在 100VG—AnyLAN 上，由于 IEEE—802.12 标准的优先存取协议结构，100VG 可支持多达 80 个使用者。

多媒体网络所有的用户均需要有影像压缩及语音输出的能力，且网络必须是主从式结构。

二、多媒体网络的硬件与软件

1. 硬件方面

多媒体个人电脑及视像服务器需 486/66 以上 CPU、16M 以上的内存以及数百 MB 的硬盘（容量大小取决于存储的影像内容。对于 MPEG 压缩，每分钟需 9M）。多媒体个人电脑需配备 MPEG 解压缩的语音视像卡。以提供 MPEG 语音视像的解压缩，减少 CPU 的负担。另外，还需要 1024 × 768 高解析度的显示器。

2. 软件方面

多媒体个人电脑需要最新版本的 DOS 及 Windows。服务器还需要网络系统，例如 Novell 的 Netware 3.12 或视像服务器软件（如 Novell NVS、Starlight StarWare、IBM MultiMedia），以实行多人共享同一文件，存储设备共享，以及多人视像服务下的视像控制等。

三、网络多媒体的要素

多媒体网络需考虑的四大因素：网络带宽、延迟性及数据传输优先顺序、网络兼容及成本。现分述如下：

1. 网络带宽

以 MPEG—1 为例，即使是采用最佳的解码及编码压缩技术，对于每个用户而言仍需要占用 1.2MB/s 的传输带宽，甚至高达 6MB/s。采用路由器、网桥或交换器来分隔多部视频服务器的网络，可借此提升网络传输带宽。低成本的 100MB/s 网络产品将是最佳选择。如 100VG—AnyLAN 单一网段中就可提供高达 80 个用户同时使用传统以太网理论上提供 10MB/s 的传输带宽，但因使用 CSMA/CD 网络数据存取方法，故使得实际网络带宽只有 6MB/s 到 7MB/s。而 100VG 采用最有效率的需求优先权存取方式，提供 97MB/s 的有效带宽，因此比传统以太网高出 16 倍。

2. 延迟性及数据传输优先顺序

质量优良的全屏幕视像需每秒 20 ~ 30 个帧页的传送。

这与通常的数据传送大不相同。例如，一个用户正在为 500MB 的网络服务器做文件备份，他可能不会在意耗费 45 秒或 55 秒的等待时间。但是，如果他同时也正在看病患者的心电图或公司的教育培训影片时，数据传输优先

顺序的传送时间却是十分重要的。假设在应用层遗漏 30 个帧页中的一帧页，他或许感觉不出异常，但若是需要等待一个帧页，并且通过网络缓慢传送，则他将会感觉质量的差异。通过 100VG 控制多媒体的数据库存取的优先顺序，可达到各项软件的应用需求。

3. 网络兼容性

由于 100VG 使用现存的以太网和令牌环网数据格式，所以无需修改现有的网络操作系统、应用程序及网络结构等。对于不同网络的数据格式的转换，如 ANSI—FDDI 及 IEEE—802.3 10MB/s 间的转换，可通过网桥来达到速率匹配。

4. 成本

任何新技术都有不同的软、硬件成本。与现有的以太网相比，100VG 的网卡和集线器的硬件成本仍算是廉价的（ISA 网卡为 399 美元；100VG 集线器为每端口 250 美元）。由于 100VG—AnyLAN 是根据 10BaseT 以太网架设的，因此人员培训成本、规划设计及安装均相同。

专家们认为，多媒体网络从现在到将来都有其重要的商业机会。随着网络视像服务器技术的出现（如 StarLight、Nov-ell/Fluent），将使多媒体信息在局域网能顺畅地流动，而快速网络技术将在其中扮演重要角色。

四、多媒体数据传输

多媒体数据传输的物理基础是多媒体网络，同时要涉及到多媒体的数据压缩技术。

多媒体网络具有上述高速、同步、实时等特性，传输音频、视频等巨量数据的需要对多媒体网络提出了远高于一般网络的特殊要求，首先是要有很高的带宽，由于多媒体数据的量很大，从而要求具有很大的存储空间，传输带宽也要求很高，因此就要对数据进行压缩。但高倍率的压缩不仅影响媒体本身的质量同时也是以损失原始数据信息量为代价。在使用静态、慢速或小画面传输来限制数据量时，也会影响通信质量。因此，必须加大带宽，使得通信网能适应多媒体数据量的增长。

其次，是对实时性的要求。满足了前面的高带宽的要求之后，如果通信协议不合适，也会影响多媒体数据的实时性。一般来说，电路交换方式延时短，但占用专门信道；而分组交换方式则延迟偏长，且不适于数据量变化较大的应用使用。不同的媒体对通信网、通信协议及高层软件提出了不同的要求。例如，对语音传输，最大可接受的延迟为 0.25 秒，使用分组交换方式传输的语音最大时滞（分组与分组之间的延迟）应小于 10 毫秒，否则就会感到说话声不连续；但其对传输速率的要求相对低一些，可以采用 64kbps 的信道。对图像而言，延迟大一些不会产生多大影响，错一个像素（例如红色变成绿色）也影响不会太大，但丢失一个分组（包），在图像中会出现一个黑块，这是不能容忍的。对数值传输而言，不允许出现任何错误，但时滞对数值和图像不会产生任何影响。鉴于各种媒体之间特性的不一致，一般采用“服务质量”（Quality of Service）来描述，传输时也往往用“服务质量”来决定传输策略，例如对语音可采用延迟断、延迟变化小的策略，对数据传输则可采用可靠、保序的传输策略等。“服务质量”是说明多媒体性能目标的元组，通过该元组的性能说明，可以对通信系统性能进行指定。“服务质量”基本参数包括传输速率、最大延迟、最大时滞、可用性、可靠性、传输位率、

利用率等。传输位率是其中最重要的参数，但不同的系统强调的参数往往不同，而且“服务质量”参数的设置一般采用分层方式，不同层的参数表现形式不同。

多媒体信息的数据量是非常大的，不仅仅在网上实时传输多媒体信息时需要的传输速率大大超过了当前网络的数据传输速率，而且对网络终端系统的存储容量及多媒体信息的同步也提出了过高的要求。

为了解决这些矛盾，必须采用数据压缩技术。例如一幅 500×500 的 24bit 真彩色图像约需 6MB 的存储容量，如用传输速率为 64kbps 的 ISDN 信道进行传送，需要 94 秒才能完成，但用 JPEG 压缩之后，存储量可降低近 20 倍，用同样的信道进行传输只需 5 秒。可见数据压缩的重要性。

数据压缩是以一定的质量损失为容限，按照某种方法从给定的信息源中推出已简化的数据表述。它通过减少信号空间的量，使信号能安排到给定的消息集或数据样本集中。数据之所以能够压缩是基于原始信源的数据存在着很大的冗余度。衡量一种数据压缩技术的好坏有三个重要的指标。一是压缩比要大，即压缩前后所需要的信息存储量之比要大；二是实现压缩的算法要简单，压缩、解压缩速度要快，尽可能地作到实时压缩及解压缩；三是恢复效果要好，要尽可能地恢复原始数据。随着通信技术和计算机技术的发展，数据压缩技术也日趋成熟，适合各种应用场合的压缩方法不断产生。目前常用的压缩编码方法可以分为两大类：一类是冗余压缩法，也称无损压缩法、熵编码；另一类是熵压缩法，也称有损压缩法。有损压缩法压缩了熵，会减少信息量。

多媒体压缩大体可分为音频和视频两部分。CCITT 在 1984 年公布了基于自适应差分脉冲编码调制 (ADPCM) 的音频压缩标准 G721，其先后又公布了 G722、G728 等采样频率为 8 或 16kHz 的音频数据压缩标准。但比较成熟的高保真立体声音频压缩标准却是“MPEG 音频”。

视频信号的压缩相对于音频信号而言更为重要，因数字化视频中含有音频信号。主要的视频压缩标准有 JPEG、MPEG、JBIG 等。

五、多媒体网络与通信

现在运行中的网络主要有电话网、有线电视网、局域网、ISDN 等，设计中的理想网络有宽带综合业务数字网 B-ISDN 等。

电话交换网是以传输模拟信号为主的，带宽很窄，一般不用于多媒体通信。

有线电视网则只能单向传输、交互性低，也不适合作为多媒体网络使用；高速局域网和 ISDN 是目前多媒体通信的基础，B-ISDN 是未来多媒体通信的主要发展方向。

1. 高速局域网

对局域网的多媒体扩展主要是针对高速局域网进行的。高速局域网是指传输速率大于 100Mbps 的局域网，典型的有光纤分布数据接口 FDDI (Fiber Distributed Data Interface)、100VG-AnyLAN 和 ATM-LAN 等。FDDI 是 ISO 标准的高速局域网，其数据传输率高达 100Mbps，采用基于令牌环的网络拓扑结构，提供了分布式应用和图像传输所必须的带宽，并可作为更大网络环境的基础。FDDI 的网络模型是基于 OSIRM 的一种扩展，主要是对数据链路层和物理层的扩充。

FDDI 的主要技术特性有以下几点：

(1) 高可靠性。FDDI 采用双层反向环形结构传输数据，可以保证网络中某一节点失效时，数据仍可在网络中其余节点之间传输。

(2) 消除时钟偏移。FDDI 采用分布式时钟方案，每一节点都有独立的时钟和弹性时钟缓冲器，以保证在网络的运行过程中不出现时钟偏移。

(3) 优秀的容量分配策略。FDDI 采用标记传递协议，在发送站发出帧后立即发送标记帧，既支持长的通信流，又支持突发通信；既支持同步通信，又支持异步通信。可以保证各节点间既可发同步帧，又可发异步帧，并满足多帧对话的要求。

(4) 能保证信息同步的数据编码。FDDI 采用 4B、5B 编码技术，编码后的五位码中 0 码不超过 3 位，能得到足够的同步信息。

100VG—AnyLan 是 HP 公司和 AT & T 公司对其 100Base—VG 技术的扩展，并被 IEEE 认可为 IEEE802.12 标准。它融合了以太网和令牌网的优点，数据传输率可达 100Mbps，并可选择路由与 FDDI、ATM 骨干网和广域网连接，是一种多媒体网络的优选网络。100VG—AnyLan 的网络模型也是对 OSIRM 数据链路层和物理层的扩充。

100VG—AnyLan 具有如下技术特性：

(1) 优先级控制 网中一般的数据定为普通优先级，而与时间相关性高的多媒体数据定为高优先级，具有优先访问网络的权利。

(2) 通道分配与编码 先进行帧和位的处理，再对处理后的 5 位数据进行 5B、6B 编码，形成 6 位数据，再加入同步信息、错误检测、帧标志码等形成网络传输格式。

异步转移模式 (ATM) 是一种新的复用与交换机制，具有很高的带宽、远距离传输、延迟低、独立带宽及带宽可根据需要进行动态配置的优点，其数据传输率可高达 2Gbps，是理想的多媒体网络模式。其网络模型是一种具面状特征的分层结构，包括物理层 PHY、ATM 层、ATM 适配层 AAL，管理面、控制面和用户面对它们进行操作。ATM 采用分组交换方式，以固定长度为 53 字节的信元，通过硬件进行交换，具有分组交换和线路交换的特点，不会因交换而造成延迟，适合于大容量多媒体数据的传输通信。ATM 的基本特征包括有如下几点：

(1) 简洁的网络结构。ATM 采用基于终结点和开关之间的点对点链路代替了共享存取传输媒体，有效地解决了共享地址空间的网络中用户对网络无限制存取问题。ATM 是由开关形成的网络，每条链路都能独立操作，改善了传输速率与开销问题。

(2) 固定长度的信元。采用 53 字节的定长信元 (5 字节信息头和 48 字节数据体)，信息头中包含数据的类型及发送端、接收端等信息，还可将对时间敏感的媒体数据加上时间码，以保证多媒体数据的无缝传输。

(3) 面向连接的技术。ATM 用面向连接的技术代替了把数据报文发送到网络上，再由路由器传送的方法，用嵌入式的层次化寻址代替了以往的复杂数据链路层和网络层地址复合机制，大大提高了整个网络的效率。

(4) 可扩充性强。ATM 为用户提供了规模可缩放的网络，把逻辑子网和物理子网分开，允许网络管理员建立一个跨越物理子网的、在物理上分开的逻辑用户组成的逻辑用户组或虚拟网络。

ATM 网改变了传统局域网的媒体分享性，而使网络具有切换性，大大提高

了网络的扩展性、定位性和节点数，具有很高的数据传输率，可以为每一个网络用户提供 100Mbps 的数据传输率。

2. ISDN 和 B—ISDN

ISDN 现已在美国、欧洲、日本等发达国家投入实际使用，并已受到越来越广泛的重视。采用 ISDN 可以实现利用单一的系统接口来控制一般用户的所有通信需求，可以在数字电话网上将语音、电子邮件、传真、因特网业务和视频服务等综合起来，主要用于传输静态图像和要求较低的动态视频。

ISDN 是一种广域公用数字网，具有 B、D 两种信道，B 信道以 64kbps 的速率传输电话呼叫和数据，D 信道以 16kbps 的速率传输编号、信令和维持信息。ISDN 具有两种基本的工作方式，即基本速率和基群速率。基本速率方式在 144kbps 线路上提供 2 个 B 信道和一个 D 信道（简称 2B+D），速率较低，服务对象是广大的终端用户；基群速率方式在四线 2Mbps 线路上提供 23B+D（美国、日本）或 30B+D（欧洲），速率较高，主要服务对象是团体用户和网络服务器。

B—ISDN 是一种基于光纤的分组交换 ATM 网，具有电路交换延迟小、分组交换速率高（最高可达 2.4Gbps）及速率可变等特点，适合于传输高保真的立体声、普通或高清晰度电视节目等。B—ISDN 现还处于研究阶段，CCITT 1990 年公布了基本 B—ISDN 试验网标准，1992 年公布了商业网标准，现在仅有少数的几个试验网在运行，预计在下个世纪才能正式投入使用。

作为一种全新的网络，B—ISDN 并不是对现有通信网 ISDN 在带宽上的简单扩展，它具有如下的重要技术特征：

（1）具有全新的信息传送方式、交换方式和用户接入方式。

（2）其运行的业务种类繁多，特征差异大，能传送现有的任何业务及将来可能出现的业务，从传输速率为 Mbps 级的控制数据到 100Mbps ~ 150Mbps 级的 HDTV 信号都能在网络上传输。

（3）具有很高的灵活性和网络重组能力，能适应各种新技术和新业务的变化。

（4）在网络内采用统计复接方法，能使网络的资源得到充分有效的利用。

3. 因特网中的多媒体信息传输

因特网是全球最大的网络系统，已连接世界上 90 多个国家和地区的近 5 万个网络、320 万台计算机和 3200 万以上的用户。预计到 2000 年，将有 100 万个网络、1 亿台计算机和 10 亿用户使用因特网。因特网被认为是未来信息高速公路的雏形。

在因特网上传输多媒体信息的主要障碍是数据传输的时间延迟，因为因特网所基于的 TCP/IP（Transmission Control Protocol/Internet Protocol）协议不给特殊的传输分配路径或更多的带宽，这将引起大数据量的多媒体信息传输产生不稳定和不可预测的时间延迟，严重影响多媒体信息传输的实时性。解决该问题的主要方法是 IETF（Internet Engineering Task Force）开发的资源预留协议 RSVP（Resource Reservation Protocol）。

RSVP 运行于 IP 的上层，提供接收方初始化的资源预留设备，以便于应用程序数据流的传输。

RSVP 的另一重要优点是其兼容性好，虽然它是以 IP 为底层协议的，但仍

能工作于任何的物理网络结构中，包括以太网、令牌网、FDDI 等。

因特网进行多媒体信息传输的另一问题是传输速率。因特网用户可以通过 4 种方式连入因特网：电话网、ISDN、局域网、广域网。用户通过因特网传输、获取、共享数据的通道必须经由以上网络的一种或多种，其中速率最低的网络将决定数据传输的速率。如果传输通道的最低级网络是电话网，则只能分别进行语音、传真、邮件和静态图像，以及质量很差的视频，并且同步性很难保证；如最低级网络是 ISDN，则可将语音、电子邮件、传真、视频服务等综合起来，主要用于传输静态图像和要求较低的动态视频，基本上能保证同步；如最低级网络是数据网（局域网、广域网），则能实现语音、邮件、传真、中等质量视频的同步传输。

六、多媒体网络技术的应用

多媒体网络技术的应用领域非常广，主要包括信息检索、远程教育、异地医疗诊治、娱乐、办公室自动化等。

信息检索是最简单的一种多媒体网络应用，这实际上是以文件为单位进行数据的传输和共享，即将多媒体数据以文件的形式存储在多媒体文件服务器中，网络用户可通过网络进行数据共享。这种应用要求的带宽较低。

多媒体网络能够实现真实、交互的远程教学，教师和学生不仅可“面对面”地交谈，而且还能通过网络即时交换和共享教学信息，所以这种远程教学不仅具备面对面教学的交互、灵活、反应快等优点，而且教师和学生能方便地共享信息，教师能简单地将教学内容以文字、图像、视频或文件的形式传输给学生，并能即时地查阅、批改所有学生通过网络所提交的作业，这有可能从根本上改变现有的教育方式。总的来说，用多媒体技术进行远程教学，有以下特点：

（1）模拟现实环境。多媒体同步传输图、文、声、像等信息能够很好地反映现实世界，增强了学习兴趣，提高了学习效果。根据有关专家实验的结果，人类感官对信息的理解程度因刺激的方式而不同，在视觉 83%、听觉 11%、触觉 3%、味觉 2%、嗅觉 1% 的比例下综合刺激才能产生最佳效果。人类对信息记忆的保持能力与信息的提供方式有关，谈话情况是：2 小时后 70%、72 小时后 10%；观看情况：2 小时后 72%、72 小时后 20%；二者并重：2 小时后 85%、72 小时后 65%。多媒体计算机技术可以通过视觉、听觉，甚至触觉、味觉等多种方式对用户的感官进行刺激，其效果是任何一种单一的途径所无法比拟的。

（2）信息共享。这是通信网络的显著特点，凡在网上的用户均可以通过网络即时交换和共享信息。教师将教学内容以文字、图形、视频或文件的形式传给学生，并能够即时地查阅、批改所有学生通过网络所提交的作业。

（3）智能化。计算机辅助教学的真正目的是要作到“因材施教”。在软件设计时采用超文本的思想，并在超媒体的链中融入知识和规则，使其具有推理的能力。可以有针对性地选择学习内容，根据学生的测试结果给出下一步的学习计划，使学生根据自己的实际需要进行学习，明显地提高了学习效率。

通过多媒体网络，医生、专家可以真实地观察、诊断万里之外的病人，并给出医疗处方或指导当地的医生进行手术，实现异地医疗诊治。这将大大减轻病人的痛苦、使病人得到及时的治疗，有时甚至能及时抢救病危的病人，

同时也大大降低了医疗的费用，发挥出城市大医院专家或名医的医疗水准，充分共享人类的智慧。

娱乐是多媒体技术的重要应用领域。多媒体网络实现的交互电视 VOD (Video-on-Demand) 改变了传统电视由电视台单向控制的播放方式，用户可以自己选择播放的节目甚至与电视台共同参与节目的制作和播放；多媒体网络的宽带特征，使其可以传输高清晰度电视 HDTV 信息，网络用户可以享受高质量的电视影像；视频重放可以实现网络电影数据库，及在网络服务器上存储电影的数字视频信息，用户通过网络检索、选择需要的电影，并在自己的多媒体终端上重放电影的视频信息，用户通过网络检索、选择需要的电影，并在自己的多媒体终端上重放电影的数字视频信息；多媒体网络中的两个或多个用户可以参与进行声、视频同步的大型真实多媒体对抗游戏，其感觉比与计算机对抗要有趣得多。

视频会议系统 (Video Conferencing System) 是多媒体网络的重要应用，不同地点的人员，可以通过显示器或电视屏幕来传达文件、进行讨论、协调工作、共享信息等。人们无需关心地理位置上的差异，只需把自己的方案、档案资料准备好，就可以随时交与“与会”各方，“面对面”地讨论问题。基于电话网的视频会议系统其图像质量和同步性较差，有实用价值的是基于 ISDN 或局域网的视频会议系统。

多媒体网络可以提高办公室自动化的质量和效率，除视频会议系统外，文件会议和白板应用是通过多媒体网络实现办公室自动化的主要方式。身处不同地理位置的双方可以共享、共同修改、存储、显示数据，并能存储成文件永久保存。文件会议和白板应用可以集语音、传真、文件、图像、视频于一体，可以直接将传真文字识别成文件保存，实现无纸办公。

七、多媒体网络技术的发展前景

多媒体网络技术是目前广为关注的信息高速公路的基础技术之一，也是现代计算机技术最活跃的分支之一。

按国家信息基础结构或信息高速公路构想的要求，未来的多媒体网络应是高速信息传输网，即在全国乃至全世界范围内，以光纤和高速数字交换设备构成的高速宽带信息网络。在网络中连接有大量专用和公用数据库，各种用户通过多媒体终端享受商务、金融、教育、卫生、文化、娱乐等广泛的服务。多媒体网络将是光纤作为主要的传输介质、基于 B-ISDN 的网络，它具有如下特征：

(1) 双向传输。网络中的信息是双向交流传输的，信息消费者同时也是信息提供者。

(2) 高精度、快速数字化传送。以数字形式传输数据、文字、声音、图像和电视节目等内容，在对数据进行高效压缩后，传输效率将比现有水平高出数千倍。这将使第三产业、广播电视业、娱乐业和超级市场发生根本性变化。

(3) 宽带化。传输速率应是 G 级 (1Gbps=1000Mbps) 以上的，采用 ATM、SDH 技术的 B-ISDN 作为网络基础。

(4) 开放式系统。人人都可通过电话线或电视接口进入网络，人人都有共享信息的权利。

(5) 规模化。多媒体网络应是全球网，能将全世界的各种大小网络连接

起来，异构网互联、互操作性、不同协议的转换和通用（兼容性）、操作透明性等将是要解决的突出问题。

（6）智能化。即进行智能化的网络管理，实时监测、调度、管理网络状态，对系统配置的规划、安全、重组、计价等进行统一的智能管理。

（7）多层结构。未来的多媒体网络应是一个巨大的网络之网络，具有多层结构，每一层都是网络。其中最重要的是服务网络，其重点是增值业务网。

总之，多媒体网络技术将随着信息社会的不断深化而得到广泛的研究和应用，其中技术问题的解决程度将直接关系到未来信息高速公路的成败。

