

第一部分 基础篇

第一章 多媒体基础知识

多媒体自 90 年代中期以来已成为计算机领域最热门的话题之一，计算机的多媒体能力也已成为重要的衡量标准，许多使用计算机的人也都希望自己能编制出美观的多媒体软件，那么究竟什么是多媒体呢？如何利用这些多媒体知识编写出较好的软件呢？本章将在进入 AuthorWare 之前，对多媒体的基础知识做一简略介绍。

一、多媒体概述

1. 多媒体定义

“媒体”一词在计算机中有两种含义：

- ①指存储信息的实体，如软盘、硬盘、光盘、磁带、半导体存储器等；
- ②指存储信息的载体，如数字、文字、符号、声音、图形、图像、动画、视频等。

Multi 是一个英文单词的前缀，表示“多”的意思；Medium 是英文“媒体”的意思，其复数形式为 Media；两词合并——MultiMedia，即多媒体。它一般指的是第二种，即多种信息的载体。但这不能说是定义，只是一个新近出现的英文单词。

那么，是不是只要包含多种媒体信息就可以称为多媒体了呢，回答是否定的。如果把上述载体孤立的使用，那就和电话号码、看书读报、收听广播、甚至看电视、电影一样，这都不能称为多媒体。

一般认为多媒体可以如下定义：

以计算机为核心，交互地综合处理文本、图形、图象、动画、音频、视频等多种媒体信息，并使这些信息建立逻辑连接，以协同表示出更丰富更复杂的信息。

2. 多媒体的特征

多媒体有许多特征，归纳起来主要有三个，即：多样性，交互性，集成性。

在多媒体软件开发, 或者叫多媒体创作中, 这三个主要特征都是要予以考虑的。

(1)多样性

早期的计算机只能处理数字、文字(英文和单一字体的汉字)及经过特殊处理的图形, 而多媒体计算机则除了数字、文字以外还可以处理下列信息:

各种字体——各种不同风格的字体(如行楷、行书、魏碑等等)以及艺术字体等;

图形——计算机通过算法生成的画面或画家笔下的作品;

图像——自然空间照片, 即人物、景物等实际场景的静止画面;

(图形和图像可统称为图片, 因有时二者之间的分界是模糊的。比如通过某种特殊效果处理的图像, 就很难区分它们是图形还是图像。)

动画——指由一系列静止画面组成, 按一定顺序快速播放, 产生出活动的感觉;

音频——指各种声音信号;

视频——指由摄影机、摄象机等拍摄的反映真是生活场景的活动画面。

以上这些也是构成多媒体的基本要素, 或称为素材。

人的大脑处理外来信息主要靠五大感官, 即视、听、触、嗅、味, 前三者大约占了全部信息量的 95%。多媒体的多样性已能满足人的视、听感官, 占了大脑处理外来信息的绝大部分。

(2)交互性

电影和电视也包含多种媒体信息, 比如有视频、音乐、对话、字幕、以及动画特技等, 也都可以满足人的视听感官, 但没有交互性, 人只能被动接受。也就是说当您一旦选择了看一部电影的时候, 您就只能随着这部电影情节的发展而发展, 或者干脆不看, 没有第三种选择。

多媒体的交互性指人可以和计算机的多种媒体信息进行交互操作, 从而给用户提供更有效地控制和使用信息的手段, 而不仅仅是被动地接受。也就是说人(用户)可以和计算机交谈, 可以相互沟通, 可以根据需要选择媒体信息, 甚至可以控制电影情节的发展。

(3)集成性

指以计算机为核心, 将不同的媒体设备集成为一个整体, 它包括信息媒体集成和信息处理集成, 即多媒体信息的统一组织和存储以及表现合成等。

一个好的多媒体软件应能给人以舒适的视、听感觉和充分的人机交互性。

3. 多媒体的发展

从研究角度看, 早在有多媒体之前, 人们就已开始研究第五代计算机——人工智能, 但由于种种原因, 智能计算机进展不大, 而主要以满足视听感官和人机交互为典型特征的多媒体功能, 作为一种过渡产品, 却越来越获得人们的青睐。

计算机技术本身的发展为多媒体的发展也提供了条件, 可以从以下几方面看出:

(1)大容量硬盘和光盘的问世

多媒体信息的存储一般都要占较大的容量, 以声音信号为例: 要录制一秒钟的语音信号在计算机中大约要占 11KBytes 的存储容量, 而录制一秒钟的立体声音乐信号要占大约 176KBytes 的存储容量(参考本章3.2); 也就是说一张 1.44MB 的高密软盘只能存放大

约 8 秒钟的立体声音乐。

所以为了满足大量多种媒体信息的存储，必须有大容量的存储媒体。

硬盘的存储容量几乎以每年翻两番的速度增长，1994 年 120MB、1995 年 540MB、1996 年 2GB、1997 年 7GB、1998 年 13GB、……

只读光盘每张 650MB，相当于 450 张高密软盘的容量，一般可存储 14~16 首立体声歌曲或乐曲。

(2) 超大规模集成电路

为了满足集成处理多媒体信息，要求有高速运行的 CPU、主机板、内存及硬盘等。Pentium (奔腾)166 MMX(带有扩展多媒体处理能力的 CPU 芯片)的问世使得计算机的处理能力真正进入了多媒体时代，Pentium II (奔腾 2)更使多媒体电脑如虎添翼。

(3) 数据压缩技术

为什么多媒体信息一定要进行数据压缩呢，主要从两方面来看：

① 存储问题

以视频为例，如果不压缩，一帧 $640 \times 480 \times 256$ 色(普通 VCD 的分辨率)的画面大约需要 307KBytes (参考本章四.2)，若以每秒钟播放 30 帧计算，一秒钟的视频信号大约需要 9.2MB 的存储空间(这里的计算还未包括音频信号)。即使是 650MB 的光盘，也仅够存放播放 70 秒钟的视频信号！一部电影一般至少一个半小时，您可以简单算一下，要多少张光盘来存储，至少 77 张！也就是说，看一个电影要换 77 次光盘，就算是目前最高档的 VCD 机三碟连放，也要换 26 次光盘，那个电影看着还有意思吗。

而经过压缩之后再进行存储，一张普通光盘就可以存放大约 50 分钟的视频(包括伴音)信号。

② 传输问题

使用 MODEM 通过电话线路传输，常用的有 28.8KBPS(Bit Per Second)、33.6KBPS 和 56KBPS(目前最快的)，若不压缩传送视频信号每秒大约 0.02 帧，也就是说传送 1 秒钟的视频信号大约需要 20 多分钟。

用 CD-ROM 来看 VCD 光盘：单速 CD-ROM 传输率为每秒 150KB，倍速为每秒 300KB，四速为每秒 600KB，…，即使是 24 倍速 CD-ROM，其持续传输率为每秒 3600KB，如果存储的视频信号是未经压缩的，那么最快每秒也只能传送十几帧！远不够每秒 30 帧的播放要求。

综上所述，将占大容量的媒体信息压缩得尽可能小，可以大量节省存储空间和传输时间，尤其在国际互联网中使用这些多媒体素材时，传输时间尤为重要。

笔者认为，多媒体软件不一定非得追求占容量多么大，而是在完成相同功能、体现相同内容的情况下占容量越小越好。

4. 新技术

电视计算机——Teleputer(Television(电视)和 Computer(计算机)的组合词)

计算机电视——Comvision(Computer 和 Television 的组合词)

以上两项不管怎样称呼，总之是把计算机与电视结合起来，以增加交互性和可视性。

网络多媒体，内容极其丰富，可以在网上实时传送图象和声音(如 ATM 网等)。

二、多媒体的研究对象

1. 听觉(音频信号)

主要包括:

语言——鸟有鸟语，兽有兽语，只不过我们听不懂；这里仅指人类语言。

声响——自然界的各种声音。它可以是风声、雨声、鸟叫、虫鸣，以及汽车声、碰撞声、破碎声、摩擦声等等。总之包括自然声响和人类生产生活各方面制造出的声响。

音乐——音乐家的杰作、演奏者的功劳。是通过各种器乐的音调、音阶、休止(此处无声胜有声)等的组合来表达某种思想和意境，是声响的一种抽象。当然歌唱家通过自己运用歌喉的表演也是一种音乐。

语言、声响、音乐之间有时没有明显的界限，比如在钢琴上乱砸一通只能产生无规则声音而不能说是音乐，再比如公共场合大声喊叫也只能是另人厌恶的噪音而非语言。

2. 视觉

包括两个方面

(1)静止的

图象——具体的、自然界的真实画面；比如照相机拍摄下来的照片。

图形——对图象的一种抽象，是画出来的，包括：

二维图形——平面的图画，

三维图形——具有真实感、立体感的一种造型；

文字——包括各种语言文字和符号；是对图形的一种再抽象。特别是中国的方块字，本身就是从图形进化而来，每个字都起源于一个特定的图形、表达某个特定的意义，已形成公认的固定规则。

(2)动态的

影像——由摄影机或摄象机等拍摄下来的实际活动景象。如电影、电视片等就是拍摄后再经过剪辑、加工、配音、电脑特技等制作出来的。

动画——影像的一种抽象，是创作出来的。如动画片。

动态视觉信号在多媒体中我们可以把它们统成为“视频”信号，

3. 其它

触觉能够感受到:

温度——皮肤接触可以感知物体的冷热。

质地——手指触摸一个物体表面可以感觉到光滑、粗糙等，并能感觉出是什么材料。

压力——通过压力可以感知重力。例如当您拿起一只杯子的时候，通过指尖的触觉可以感受到对杯子施加的压力，从而知道杯中有没有水，有多少水。

运动——人能感知 6 个自由度的运动。例如手握一只麻雀，您就能通过手掌的触觉感受到麻雀往各个方向的运动和振动。

传感——将感受到的各种感觉通过神经传到大脑。

例如压电橡胶就可将压力变成电信号(再通过模/数转换变为数字信号送入计算机处理),这一技术可以使机械手带有一定触觉。

另外还有嗅觉,味觉。也许有人会说还有第六感官——直觉。

这些都不在本书讨论范围。

多媒体主要研究前两项,即听觉和视觉信息的采集、存储、加工、播放等。

虚拟现实(Virtual Reality)技术是在多媒体基础上再增加触觉乃至嗅觉和味觉,主要基于仿真技术,是当代多媒体最前沿的课题。

三、数字音频

指数字化了的各种音频信号。所谓数字化就是把各种音频信号转换成计算机能够处理的数据,播放时再把这些数据转换回人耳能够听到的声音。

1. 概述

(1)种类

目前使用最多的主要有两大类:

波形音频——自然界的任何声音,包括语言、声响、音乐等,都是以波的形式传播,所以原则上都可以录制成波形音频文件。波形音频文件以.WAV 为扩展名。

MIDI——Musical Instruments Digital Interface 的缩略语,即乐器数字接口,也称为电子合成模拟音乐或迷笛音乐。MIDI 音频文件以.MID 为扩展名。

另外还有压缩格式的波形音频文件,如扩展名.MP3 等等,由于文件尺寸小而又容纳信息多,因而 Internet 网络上用的较多。

本书主要讨论前两种。

(2)特点

各类音频信号都有共同的特点,它们是:

①音频三要素

音调(声波的振动频率,即音高,人耳能听到的频率范围是 20Hz~20kHz);

音强(声波的振幅,即您所听到的声音的大小);

音色(在基音频率上混入泛音,20kHz 以上的泛音人耳虽不能听到但能感受到)。

②连续性

不能断续;试想您要听一个人说话总是断断续续的会是什么滋味。

③方向感

人能感知声音的来源(方向);立体声音响是通过不同声道的时差和强度使人感受到声音的方向。

一段好的音乐应该是音域很宽且泛音丰富,所以通常用能够表现的音频范围来衡量一个音频设备的质量。

几种音频设备的质量见表 1-1-1。

计算机虽比不上发烧音响的特殊效果，但已接近宽带音响。

表 1-1-1 音频设备及其频率范围

设 备	频 率 范 围
电 话	200Hz~3.4kHz
调幅广播	50Hz~7kHz
调频广播	20Hz~15kHz
宽带音响	20Hz~20kHz
发烧音响	16Hz~40kHz

2. 波形音频

(1)采样与量化

波形音频的特点是非线性连续谱，不可能以某种函数来表达，必须通过采样与量化才能把它数字化，如图 1-1-1 所示。

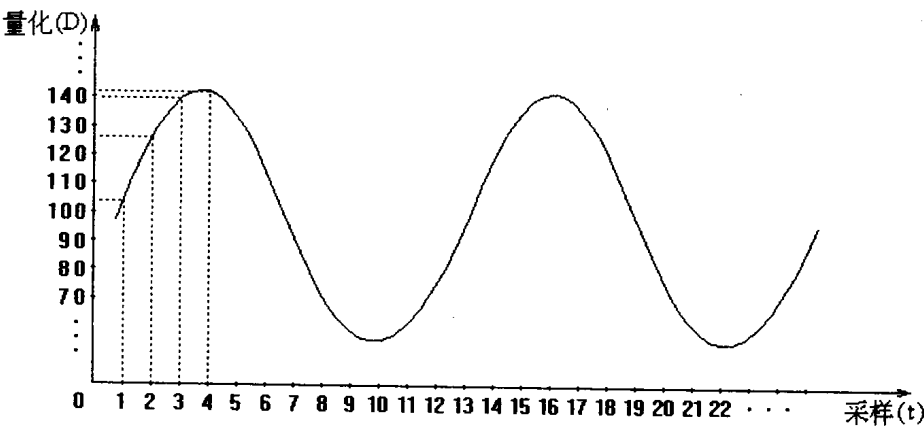


图 1-1-1 采样与量化示意图

图 1-1-1 的横坐标为采样，即每隔一定时间间隔获取一次音频振幅信号。

纵坐标为量化，即把采样所得的振幅模拟量数字化，也就是将采样点的振幅信号转换成数字信息。对于相同振幅的音频信号来说，纵坐标的标注越密集，转换精度就越高。那么什么因素决定纵坐标的标注呢，那就是量化等级。

例如图 1-1-1 中所示在第一个时间点采样所得振幅信号的数值为 104，第二个时间点采样所得振幅信号的数值为 126 等。显然，单位时间内采样的次数越多，两次采样之间的时间间隔就越短，越能反映快速变化的音频特征；量化等级越高，采样所得数字就越接近实际值。

(2)技术指标

如何衡量波形音频数字化的质量呢，这主要用到两个指标：采样频率和量化等级。

采样频率：

每秒钟的采样次数, 单位为 Hz。每秒钟采样一次即为 1Hz, 每秒钟采样 1000 次即为 1000Hz 或记为 1kHz。采样频率越高, 单位时间内的采样点就越密集, 采集到的数据量所能表现的声音信号就越接近连续谱。

目前常用的有三种采样频率

11.025kHz (通常用于录制语音, 因为语言的音域窄且质量要求不高)

22.05kHz

44.1kHz (通常用于录制音乐)

量化等级:

即模拟量转换为数字量时的转换精度。

由于计算机中使用的都是二进制数, 所以量化等级只能用二进制的整数次幂来表示, 也就是用二进制的位数来表示。早期使用的量化等级只有 8 位二进制数, 8 位二进制所能表达的十进制为 $2^8 = 256$, 这个数值的意义是假如把可能采集到的音频最高振幅作为 100% 的话, 它下面只能分为 256 等份, 即纵坐标的标注最大值为 256, 各个采样点采集到的振幅就在这 256 等份中找到最接近的数值来表示。

目前使用较多的是 16 位量化等级, 它能表达的十进制数为 $2^{16} = 65,536$ 。

目前最高的量化等级是 32 位的, 它能表达的十进制数为 $2^{32} = 4,294,967,296$ 。

显然, 量化等级越高, 振幅转换成的数值就越准确, 越接近于真实波形, 失真越小。

声道数:

声道数也是很重要的一个指标, 一般有单声道、双声道和多声道。要想达到立体声效果, 至少要双声道。目前电脑中用的多为双声道。

(3) 数据量计算

波形音频的数据量可以用下列公式计算

采样频率 $\times$ 量化位数 $\times$ 声道数 $\div 8$ (注: 每字节 8 位) = 数据量(字节/秒)

如 44.1kHz 的采样频率、16 位量化等级、双声道的波形音频, 其数据量为:

$$44100 \times 16 \times 2 \div 8 \div 1024 = 172.27 \text{ KBytes/s (千字节/s)}$$

即每秒钟的上述波形音频信号要占 172.2KB 的存储空间。

(4) 优缺点

优点: 任何声波都可记录, 使用灵活、处理方便。

缺点: 数据量太大, 占存储空间太多。

目前多用于语音及重复性较高的声音。语音通常只用 11.025kHz 采样频率、8 位量化等级、单声道, 就能够较好地采集与还原, 数据量只需:

$$11025 \text{ Hz} \times 8 \times 1 \div 8 \div 1024 = 10.77 \text{ KBytes/s (千字节/s)}$$

3. MIDI 音频

(1) MIDI 标准

MIDI 起源于电子乐器。

1982 年一些电子乐器厂家共同达成有关协议。

1988 年正式提交 MIDI 协会, 成为数字音乐的一个国际标准。

它是利用电子合成来模拟各种乐器音效的方法，它并不对声音进行采样和量化，记录的也不是声音本身，它只是像乐谱一样将每个音符记录为一个类似乐谱符号的数字，并且规定了各种音调的混合发音，最后通过合成器将这些数字组合成为音乐播出。

MIDI 用到计算机中能满足较长时间音乐的需要，但它毕竟是电子合成模拟，其声音效果远不如真实乐器的效果。

扩展 MIDI 标准除了上述规定外还规定了两个方面：

- ①电子乐器与计算机相连的硬件标准；
 - ②电子乐器与电子乐器之间、电子乐器与计算机之间的通信协议；
- 使得不同厂家生产的电子乐器之间及与计算机之间能够互相发送和接收音乐数据。

(2)音序器(Sequencer)

音序器是可以进行记录、存储、编辑和播放 MIDI 音乐的计算机软件设备。

好的音序器软件处理电子音乐就象使用文字处理软件处理文字一样，可以自己作曲、演奏和打印乐谱。

(3)优缺点

优点：MIDI 文件数据量比波形音频文件小得多。例如同样半小时的音乐，.MID 文件只需 200KB 左右，而.WAV 文件则要 300MB 以上；

缺点：目前只能记录电子乐器音乐，回放质量受声卡合成芯片的影响较大。

鉴于波形音频和 MIDI 音频的优缺点，目前在开发的多媒体软件中，波形音频通常用来播放解说语音，MIDI 音频通常用来播放背景音乐。

4. 声卡

能够采集并播放声音的计算机外部设备，通常以插卡的形式插在主机板的扩展槽中，也有直接做在主机板上的。

1987 年诞生世界上第一块声卡，Adlib 卡(也称音效卡或魔音卡)

1989 年新加坡 Creative(创通)公司生产了 Sound Blaster 卡(也称声霸卡)

现在的声卡一般都与这两种卡相兼容。

(1)外观及接口

插卡式声卡的外观及接口示意图如图 1-1-2 所示。

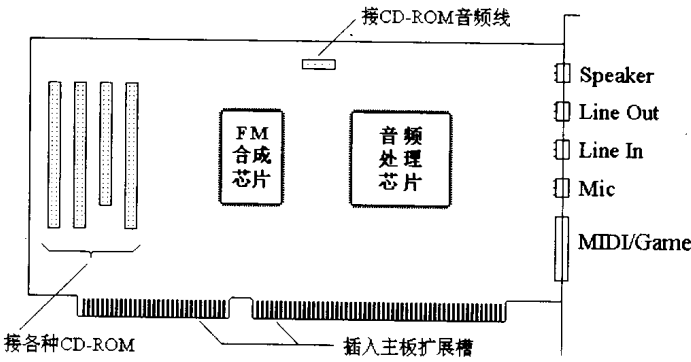


图 1-1-2 声卡外观及接口示意图

图 1-1-2 中右则的各个插口会在主机箱背面显露出来，它们的作用如下：

Speaker——外接耳机或音箱，有几瓦的输出功率，只能接小音箱或有源音箱；

Line Out——输出声音信号，一般可接至音响功率放大器的 Line In 端；

Line In——声音信号输入，一般可接至录音机等声音设备的 Line Out 端；

Mic——麦克风输入；

MIDI/Game——可接电子琴或游戏控制器等。

(2) 声卡的技术指标

① 采样频率与量化等级

目前的声卡都能达到 44.1kHz、双声道、每声道 16 位量化等级。

但模/数转换、前置放大、数字电路、合成芯片等差别较大。

② 合成器

目前多采用 FM 合成法，常用芯片有 OPL-3 或 YMF262 等，可合成 11 种单声。

波表合成技术是最新的 MIDI 合成技术，它是将真实的乐器声波录制下来存放在合成表内，因此用它播放 MIDI 曲调时特别逼真。目前有硬件波表和软件波表两种，当然是硬件波表的效果更好一些。

③ 兼容性

若仅用于 Windows 下，只需该卡提供驱动程序即可，Windows 本身可以兼容；

若用在 DOS 下，则必须要求兼容以下几个标准：

Adlib

Sound Blaster 或 Sound Blaster Pro

MicroSoft Windows Sound System 2.0

MPC Level 2 (参考本章六.2)

④ MIDI 标准

MIDI——乐器数字接口，MIDI 协会标准；

General MIDI——规定了 MIDI 文件中乐器声音的排列顺序；

MPU-401——Roland 公司指定的 MIDI 标准，也为大多数乐器制造商采用。

有声卡的电脑才能播放声音，这是不言而喻的。但播放声音的设备往往是独占的，比如当正在播放一个波形音频文件时，其它的波形音频就不能同时播放，此时若想既有音乐又有语音的话，就必须事先将音乐和语音合成在同一段波形音频文件中，但这样不利于分别控制语音和音乐。幸好波形音频和 MIDI 音频是分别两个设备，也就是说一段波形音频和一段 MIDI 音频可以在同一个声卡上同时播放。

四、图形与图像(静态视觉)

1. 人的视觉特点

(1) 可见光

我们知道可见光波段的波长范围很窄，只在 380~780nm(毫微米)之间。从长波到短波分别为：红、橙、黄、绿、靛、蓝、紫。

人眼对颜色变化反应较弱，一般感觉红紫最暗、橙蓝次之、黄绿最亮。

(2)视觉范围

亮度范围：

人眼对亮度变化较敏感。在亮度适中时，上下限亮度之比约为 1000:1。也就是说，当人眼适应了某种亮度环境时，最亮的和最暗的比不能超过 1000:1，过亮的护眼，过暗的看不清楚。

视力集中范围：

就是当您注视一点时能够看清楚的范围，一般只有 $2\sim 3^\circ$ 。在 $2\sim 3^\circ$ 范围之外的其余地方是逐渐模糊的。例如当您拿着一本书距眼睛一尺远的地方阅读时，如果眼珠不动，不可能念完整个一行；看着一个字的时候，顶多前后各一个字能看清楚。

视力集中范围以外的都可以叫做余光，在余光范围内，人眼对活动的东西还是比较敏感的。在视力集中范围内则比较侧重静止的图象。

视野范围：

左右约 180° ，对称。就是说当您平视前方并注视某一点时，左右两边余光可以看到的范围大约为 180° ，此范围因人而异，不同的人略有不同。

上下约 110° ，不对称，上边的视角小于下边的视角。

(3)注视点

所谓注视点是指容易引起人的目光注意的地方。下面列出五种，您不妨试试。如果您还能找到更多的注视点，欢迎和作者联系。

高反差交界处，即颜色反差特别强烈的地方，如图 1-1-3(a)；

拐角处，如图 1-1-3(b)；

闭合图形内侧，如图 1-1-3(c)；

变化区域，如图 1-1-3(d)；

不规则的地方，如图 1-1-3(e)。

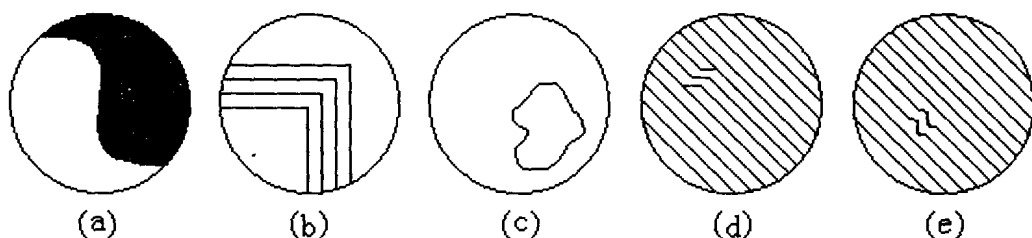


图 1-1-3 注视点测试

(4)人的视觉优点

直观，或可以说“一目了然”，因为人眼就象照相机镜头一样，不需要逐点、逐行去扫描，当您一睁开眼睛，视力所及范围内的景物一下子“尽收眼底”；

确定，也可以说“一看便知”，比如见到一个熟识的朋友，不管他是否戴了一顶帽子或是换了发型，无须仔细分析他的身高、长相，看一眼就能知道他是谁；

高效，笔者把这暂称为“一睹记住”，比如您看到一件什么东西，猛的看一眼后立即

闭上眼睛，在您眼前还会较长时间(比看一眼的时间要长得多)浮现出这样东西的形状。

了解人的视觉特点就可以在设计用户界面时加以注意，每屏显示不要有过多的重点，需要引起用户注意的地方最好不要太分散，重点的地方要能抓住人的视线；颜色搭配应尽量以暗色作背底，亮色作前景等等。

2. 图片特性(包含图形与图象)

(1)分辨率

通常说到分辨率一般指如下几种：

显示分辨率或叫屏幕分辨率

指在一定尺寸中划分的行数与列数，行列的交点称为像素，相同分辨率不同尺寸的显示器像素大小不同。常见分辨率有：

640×480，800×600，1024×768，1280×1024，……

能够分辨出两个相同颜色像素的距离称为点距，点距越小清晰度越高。常见点距有：

0.39，0.31，0.28，0.26mm，……

打印分辨率

指打印机输出的实际效果，一般以 DPI(Dot Per Inch，即每英寸点数)标定。通常有：

180，300，600，1200，2400，……

图片分辨率

指根据行数与列数决定图片的尺寸。

相同分辨率的图片在不同分辨率的屏幕上显示出的大小是不同的，当然清晰度也不相同。如果原图片清晰度不高的话，将其适当缩小，看起来会更清晰。

同一图片的显示大小和打印大小也不相同。

(2)颜色数与颜色深度

颜色数

指一个像素可以用多少种颜色来表达，如 16 色、256 色等等。

颜色深度

指表达这些颜色所需的二进制位数，如表 1-1-2 所示：

表 1-1-2 颜色数与颜色深度的关系

深 度	表达式	颜 色 数	称 为
1	2 ¹	2	二值图(黑白)
4	2 ⁴	16	16 色
8	2 ⁸	256	256 色
16	2 ¹⁶	65,536	64K 色或 16 位图
24	2 ²⁴	16,777,216	16M 色或 24 位图或真彩色

3. 图片格式

目前有很多种图形文件格式，但主要有三大类，即位图、矢量图和印刷图。较常见的有：

.BMP 或.DIB (BitMaP)设备独立位图，无压缩，存储空间只取决于图片分辨率和颜色深度，计算方法如下：

$$\text{每行点数} \times \text{每屏行数} \times \text{颜色深度} \div 8$$

如：640×480×256 色(颜色深度为 8)的一幅图片容量为：

$$640 \times 480 \times 8 \div 8 \div 1024 = 300 \text{ KBytes(千字节)}$$

BMP 图支持任何运行于 Windows 下的软件。

.PCX 位图，长度编码压缩，支持多种排版、图形艺术、视频捕获软件。

.TIF (Tag Image File Format) 位图，有多种压缩法，支持多种排版、图形艺术软件。

.GIF (Graphics Interchange Format) 位图，LZW 压缩(即 8 位 256 表项压缩)，主要为图形交换与传输而设计的一种数据流格式，许多应用软件都可以通过它进行图形转换，Internet 上使用较多。

.JPG JPEG(Joint Photographic Experts Group—联合摄影师专家组)标准位图，舍弃压缩算法，是目前压缩率最高的位图，由于压缩率高，占容量小，可大量节省传输时间，所以在 Internet 中被大量使用。目前已接近成为事实上的计算机图像工业标准。

以上所有位图都是逐点描述，所以可获得自然图象，由于采用不同的压缩标准从而得到不同格式的图片文件，这些格式大多数软件均可识别，另外许多图形处理软件还有自己的专门图片格式。

.WMF(Microsoft Windows MetaFile) 类似于矢量图的图元文件，由于是由计算产生，而非逐点描述，故占容量很小，支持 Word, Visual Basic, PageMaker, CorelDraw 等。

.PS 或.EPS (PostScript——页描述语言)它虽然只是一种印刷排版语言使用的图形格式，但由于 PostScript 已几乎成为排版系统的事实上的工业标准，所以绝大多数排版软件和几乎所有的照排机都支持这种格式。

4. 色彩空间

描述每个象素点颜色的方法称为色彩空间，也叫做颜色模式，通常有：

RGB 红绿蓝三基色(也称为加型颜色模式)，是一种以黑色为背底的配色方案，多用于屏幕显示；

YBR 亮度、蓝红底色；

HSI 色调、饱和度、强度；

CMYK 青粉黄黑(也称为减型颜色模式)，是一种以白色为底的配色方案，多用于打印。有时在屏幕上看到的图片颜色非常好看，但打印出来却不尽如人意，这通常是由于不同的颜色模式所致。若想得到打印图片与显示的一样美观，就需要将其转换为 CMYK 模式再打印。

5. 图像采集

将图象数字化的过程称为图象采集，下列图象采集设备较为常用：

①扫描仪

扫描仪主要有三种，按其扫描方式可分为：

手持式 小巧、简易、价廉，扫描速度不稳定，效果差；

平板式 价格中等，扫描效果中等：

滚筒式 价格最高，扫描效果最好。

②数码相机

通过镜头及 CCD 光耦合转换,可直接存成 JPG 格式文件,或通过专用软件转换成常见格式的图形文件。多数数码相机都是通过串口将照片送入计算机,也有可直接存入软盘的。目前有许多生产厂家诸多品牌的数码相机,价格从几千元到十几万元不等。其差异主要在分辨率上,现在最低的分辨率也得在 640×480 、24 位颜色深度,最高的可达 1280×1024 、32 位颜色深度。

③摄像机+视频捕获卡

如果您已经有摄像机，那么只要增加一块视频捕获卡，即可以采集图象。一般是取摄像机摄入图象的一帧作为静态画面。

6. 图象处理

图象处理的含义非常丰富，如底色处理、淡化处理、加灯光效果、加一些文字、图片叠加、放大缩小等等，总之，对采集到的图象进行再加工，都可以称之为图象处理。而这种加工往往是必须的，要根据多媒体软件的实际需要选择和加工图片。

图象处理的工具软件很多，目前用得最多且最广泛的是 Adobe 公司的 PhotoShop 软件，该软件可以实现非常丰富的您所能想到的各种加工处理。

五、数字视频(动态视觉)

人的视觉暂留一般在 $1/24 \sim 1/10$ 秒，所以只要把多幅相互关联的图像连续播放达每秒 24 帧以上，看起来就是连续动态的。

1. 视频标准(即制式)

目前世界各地使用的视频标准不同，主要有三大制式，见表 1-1-3。计算机的视频显示一般为 NTSC 制。

表 1-1-3

视频制式与标准

制 式	标 准	国 家 和 地 区
NTSC 制	525 线/帧, 30 帧/s, 60Hz	美国*、加拿大、日本等
PAL 制	625 线/帧, 25 帧/s, 50Hz	前西德*、我国、欧洲大部
SECAM 制	625 线/帧, 25 帧/s, 50Hz	法国*、东欧部分、前苏联等

* 号标明为发明国

2. 视频数字化

(1) 捕捉

前述视频标准都是指模拟信号，必须经过数字化才能被计算机处理；

视频信号数字化的过程称为捕捉(Capturing)。有两类捕捉：

①帧捕捉(Frame Grabber)——只将视频一帧存为静态图象；

②视频捕捉(Video Capture)——将连续帧转为数字图像序列。

(2) 视频压缩

压缩的目的在本章1.3中已有说明，这里我们主要来看一下视频压缩的特点和方法。

特点：

①要有尽可能高的压缩率；因为视频信号尺寸非常大，如果没有很高的压缩率，就不足以实现视频的存储和播放要求。

②可以是非实时压缩，实时播放。也就是说先每帧压缩时不必强调压缩速度，播放时必须达到30帧/秒，连续播放。

压缩方法：

视频压缩的方法很多，大体可分为两类：

①有损压缩——压缩时会放弃一些不重要的信息(比如一些大面积颜色相近的地方)，回放时通过算法弥补。有损压缩的压缩率一般较高。典型的有损压缩标准为MPEG(Moving Picture Experts Group——运动图象专家组)制定的一种基于硬件方案的压缩标准，其视频文件的扩展名为.MPG。

②无损压缩——不丢失信息的压缩方法，一般也称为打包。压缩率较低。

VCD一般采用有损压缩图像、无损压缩声音。

解压缩：

压缩后的视、音频信号要想回放必须首先经过解压缩，为了达到连续播放效果，解压缩的速度就必须快。主要有两类解压缩方法：

①硬件解压缩——解压缩卡(MPEG Decompression Card)，可用于80486以上的任何PC机型，一般需插在主机板上的扩展槽内，并有连线与显示卡适当连接。

②软件解压缩——满足MPEG标准的软件解压方案，无须任何附加硬件设备，但最好在Pentium 166 MMX以上或显示及内存等配置较高的Pentium 133以上使用。

软解压主要有Xing、Video Play、金山影霸、超级解霸等软件。

(3) 视频编码

将解压缩之后的数字信号还原为视频信号的过程称为视频编码。

一般都是还原成NTSC制视频信号，多数解压缩卡还可以通过硬件选择还原为PAL制信号。

综上所述，视频信号从进入计算机到返回播放要经过“捕捉”、“压缩”、“存储”、“解压缩”、“编码”五个过程。

(4) 全运动和全屏幕视频

全运动(Full Motion)——必须达到每秒钟30帧以上的播放速度，这样才不会感到跳跃和闪烁。我们称之为全运动播放。

全屏幕(Full Screen)——至少应能达到 Standard VGA 的 $640 \times 480 \times 256$ 色的运动图象,甚至能达到 Super VGA 的 $800 \times 600 \times 256$ 色的运动图象,而不是只能在屏幕上的一个小窗口中播放。

全运动和全屏幕同时达到才是播放视频信号,特别是播放 VCD 的最低要求,前述对软件解压缩的机型配置要求就是以同时满足全运动和全屏幕为标准的。

3. Windows 的视频处理

Windows 目前采用纯软件的视频处理,一般达不到全运动和全屏幕。其主要使用的视频驱动程序有三:

①Video for Windows

MicroSoft 公司自行开发的视频驱动程序,目前最高为 1.1E 版。该驱动程序使用的数字视频文件扩展名为.AVI。

Video for Windows 安装后在 Windows 的媒体播放器(Media Player)中的设备(Device)菜单中会增加“Video for Windows”菜单项,选此菜单项即可打开.AVI 文件播放。

②QuickTime for Windows

这是从 Macintosh(苹果)机上移植过来的 Windows 视频驱动程序。该驱动程序使用的数字视频文件扩展名为.MOV。但它也支持 Windows 的.AVI 等视频格式。

当然 Windows 中也支持安装硬件解压缩卡,但它必须要求解压缩卡的生产厂家提供驱动程序。

③AAPlay

安装 AAPlay 就可以播放由 Animator 或 3D Studio 等动画软件制作的文件扩展名为.FLC 或.FLI 等动画格式文件。

4. MPEG 与 VCD

MPEG(Moving Picture Experts Group—运动图象专家组)制定的一种基于硬件的、带有同步音频的视频实时压缩/回放标准。可达到 $1024 \times 768 \times 128K$ 色、24~30 帧/秒的全运动全屏幕视频图象及 CD 音质的伴音;使用的数字视频文件扩展名为.MPG 或.DAT。

VCD 采用的就是 MPEG 格式。

5. 视频卡

目前多媒体领域对视频处理远不成熟,尚在积极研究中,新成果、新产品层出不穷,就目前情况而言,视频卡主要有以下几种:

①帧捕捉卡(Frame Grabber Card)

只能单帧捕捉成静态图像。

②视频捕捉卡(Video Capture Card)

可将摄像机或录象机输出的视频信号转换成数字信号输入给计算机。

低档的只能非实时捕捉,且不能达到全屏幕;

高档的可以实时捕捉,一般带有视频处理专用芯片,甚至可进行实时压缩。

③视频转换卡(PC-TV Card 或 VGA-TV Card)

可将 VGA 信号转换成 NTSC 制或 PAL 制或 SECAM 制等电视信号输出,也可将这些制式的电视信号转换成 VGA 信号输入。

④视频叠加卡(Video Overlay Card)

可提供视频特技效果,如在视频信号上叠加文字等。

注:本章介绍的有关音频、视频及图像的采集并非本书重点,这里的介绍仅作为了解。本书重点是利用已有的多媒体素材编制相应软件。本套丛书将另外有关于多媒体素材制作的书奉献给有兴趣的读者。

六、多媒体个人计算机

具备多媒体功能的个人计算机就可以称为多媒体 PC 机——MPC(MultiMedia Personal Computer)。另外还有多媒体工作站、多媒体专用机等。本节只介绍 MPC。

1. MPC

①Macintosh 机

1984 年,美国 Apple 公司首次公布了一种新机型 Macintosh,此机能用人工合成语言进行自我介绍,开创了 MPC 的先河。

新一代的 Macintosh 机不需要附加任何硬件就可以播放声音及具备 8 位、16 位甚至 32 位的图形处理能力。操作系统采用 System 7.x,其功能和效率都远远超过 DOS 甚至 Windows,其对多媒体文件的管理十分方便。

但由于 Apple 公司的封闭政策和不开放的软硬件技术,使得为其开发软、硬件配套产品的厂家很少,因此给用户也带来诸多不便。过高的价格也影响了它的普及。

②IBM PC 机

由于 IBM(International Business Machine Corp.——美国国际商用机器公司)技术公开,吸引了诸多厂家为其开发软、硬件配套产品,一步一步地走到今天的 MPC 阶段。它的重要特点是需要什么配什么,灵活,价格低廉,软件极其丰富。目前我国国内使用的微机绝大多数是 IBM 或 IBM 兼容机。

本书所有内容主要针对 IBM 的多媒体微机。

2. MPC 标准

(1)MPC 标准的制定

1990 年 11 月,以 MicroSoft(微软)公司为主的主要多媒体厂商联合成立了多媒体微机市场协会(MPC Marketing Council),首次制定了 MPC 标准,规定了以 Intel CPU 为基础的微机扩展成为多媒体机的最低软硬件规范,后来称之为 MPC Level 1 标准。

1993 年 5 月,进一步规定了 MPC Level 2 标准。

后来由于硬件的迅速发展,几乎所有的微机都可以满足多媒体要求,所以未再规定新的标准,一直沿用 MPC Level 2 标准。

(2)MPC 的硬件配置

MPC 标准的硬件配置最低要求如表 1-1-4 所示。

由表 1-1-4 可以看出，要想满足 MPC Level 2 标准，必须 Intel 80486 及以上微机或相当于 486 及以上的其它厂家 CPU 的兼容机，配有 16 位及以上的声卡，两倍速及以上的光盘驱动器。

(3)MPC 的软件支持

MPC 的多媒体功能主要由 Windows(MicroSoft 公司产品)或 OS/2 Warp(IBM 公司产品)操作系统支持，目前使用最为广泛的是前者。本篇第二章将较详细地介绍 Windows 的多媒体支持。

表 1-1-4		MPC 硬件标准
	MPC Level 1	MPC Level 2
CPU	80386sx	80486sx
RAM	2MB	4MB
HD	30MB	160MB
Display	640×480×16 色	800×600×64K 色
Sound Card	44.1kHz, 8 位	44.1kHz, 16 位
CD-ROM	平均寻道时间 1s 持续传输率 150KB/s 支持 ISO 9660, CDAudio	平均寻道时间 400ms 持续传输率 300KB/s 支持 ISO 9660, CDAudio, CDDA XA, PhotoCD, CDI

七、光盘

光盘作为一种存储介质，对多媒体有着特别重要的意义。它适合于记录大量的、长时间存储的、不需要高速检索的信息，如：百科全书、语言词典、大量照片、美术图片、计算机软件，当然还有大家都熟悉的光盘电影、光盘音乐等。

1. 光盘分类

(1)分类

读写光盘——可以随时写入、随时读出的光盘。

目前有几个厂家都在生产，有多种记录格式，尚无统一标准，互不兼容，也就是说各家生产的光盘只能在自家驱动器中使用。写出的光盘也不能在 CD-ROM 中使用。

刻写光盘(CD-Recorder)——一次写多次读。

光盘刻写机也就是 CD-R 驱动器，其使用的刻写光盘一次写入后不能擦除。刻写好的光盘可在 CD-R 中使用，也可在 CD-ROM 中读出。

只读光盘(CD-ROM—Compact Disk-Read Only Memory)——厂家制作好的光盘，在只读光盘驱动器中只能读不能写。目前绝大多数多媒体电脑上安装的都是 CD-ROM 驱动器，使用只读光盘。