

数字媒体技术导论复习题

第一章：

媒体的分类：

1. (1)按照人的**感觉**，媒体可分为视觉媒体、听觉媒体等；(2)按照信息的**表现形式**，媒体可分为语言媒体、文字媒体、音乐媒体、图形媒体、动画媒体和视频媒体等；(3)按照信息的**种类**，媒体可分为新闻媒体、科技信息媒体、生活媒体等。(4)国际电信联盟(International Telecommunication, ITU)从**技术**的角度对媒体分如下种类：**感觉媒体(Perception)**，是指**能够直接作用于人的感觉器官，使人产生直接感觉(视、听、嗅、味、触 觉)**的媒体，如语言、音乐、各种图像、图形、动画、文本等。**表示媒体(Presentation)**，是指**为了传送感觉媒体而人为研究出来的媒体**，借助这一媒体可以更加有效地**存储**感觉媒体，或者是将感觉媒体从一个地方传送到远处另外一个地方的媒体，如语言编码、电报码、条形码、语言编码，静止和活动图象编码以及文本编码等。**显示媒体(Display)**，是**显示感觉媒体**的设备。显示媒体又分为两类，一类是输入显示媒体，如话筒、摄像机、光笔以及键盘等，另一种为输出显示媒体，如扬声器、显示器以及打印机等，指用于通信中，使电信号和感觉媒体间产生转换用的媒体。**存储媒体(Storage)**，用于**存储表示媒体**，也即**存放感觉媒体数字化后的代码的媒体**称为存储媒体。例如磁盘、光盘、磁带、纸张等。简而言之，是指用于存放某种媒体的载体。**传输媒体(Transmission)**，传输媒体是指传输信号的物理载体，例如同轴电缆、光纤、双绞线以及电磁波等都是传输媒体。

1.1.2 媒体特性

多样性、集成性、交互性、信息接收/ 使用方便

1.2.1 数字媒体概念

数字媒体是**数字化的内容作品以现代网络为主要传播载体，通过完善的服务体系，分发终端和用户进行消费的全过程。**（我国的数字媒体概念）

或：数字媒体是指**最终以二进制数的形式记录、处理、传播、获取的信息媒体。**

1.2.2 数字媒体特性

(1) 数字化 (2) 交互性 (3) 趣味性 (4) 集成性 (5) 技术与艺术的融合

1.2.3 数字媒体传播模式

数字媒体是利用数字电视技术、网络技术，通过互联网、宽带局域网、无线通信网和卫星等渠道，以电视、电脑和手机为终端，向用户提供视频、音频、语音数据服务、连线游戏、远程教育等集成信息和娱乐服务的一种传播形式。

数字媒体价值产业链包括哪几个过程？

(1) 内容创建 (2) 内容管理 (3) 内容发行 (4) 应用开发 (5) 运营接入 (6) 价值链集成
(7) 媒体应用

1.3.2 数字媒体技术**发展趋势**

将**内容制作技术**以及平台、**音视频内容搜索技术**、**数字版权保护技术**、**数字媒体人机交互与终端技术**、**数字媒体资源管理平台与服务**、**数字媒体产品交易平台**等六个方向为发展重点。

数字媒体技术六个**重点发展方向**

内容制作技术以及平台 音视频内容搜索技术 数字版权保护技术 数字媒体人机交互与终端技术
数字媒体资源管理平台与服务 数字媒体产品交易平台与服务

1.3.3 数字媒体技术研究领域

主要技术范畴包括：

- (1) 数字媒体表示与操作，包括数字声音及处理、数字图像及处理、数字视频及处理、数字动画技术等。
- (2) 数字媒体压缩，包括通用压缩编码、专门压缩编码（声音、图像、视频）技术等。
- (3) 数字媒体存储与管理，包括光盘存储（CD 技术、DVD 技术等）、媒体数据管理、数字媒体版权保护等。
- (4) 数字媒体传输，包括流媒体技术、P2P 技术等。

数字媒体技术主要的研究方向

•数字媒体的技术范畴规划了数字媒体技术的研究领域，其主要的研究方向包括：

(1) 数字声音处理 (2) 数字图像处理 (3) 数字视频处理 (4) 数字动画设计 (5) 数字游戏设计 (6) 数字媒体压缩 (7) 数字媒体存储 (8) 数字媒体管理与保护 (9) 数字媒体传输技术

1.3.4 数字媒体应用领域

数字媒体有着广泛的应用和开发领域，包括教育培训、电子商务、信息发布、游戏娱乐、电子出版、创意设计等。

第二章 数字音频技术基础

本章学习目标：

(1) 知道音频的三个特性及其相关概念。(2) 说出几种声音记录设备。(3) 了解模拟音频处理技术涉及到的设备及各自主要功能(4) 掌握音频数字化的过程(5) 能够列举几种常见数字音频格式，并进行简单的比较。(6) 了解几款常见的数字音频编辑软件，并知道其基本性能。(7) 熟练掌握一款数字音频编辑软件的操作方法。

2.1 音频技术及特性

2.1.1 音频的概念及特性

在物理学上声音的三个基本特性：频率、振幅和波形，对应到人耳的主观感觉就是音调、响度和音色。

- 频率：发声物体在振动时，单位时间内的振动的次数，单位为赫兹（HZ）。
- 振幅是指发声物体在振动时偏离中心位置的幅度，代表发声物体振动时动势能的大小。
- 音色是指声音的纯度，它由声波的波形形状所决定。

•2.1.1 音频的概念及特性

声音的分类 按照人耳可听到的频率范围，声音可分为超声、次声和正常声。人耳可感受声音频率的范围介于 20~20000 赫兹之间。声音高于 20000 赫兹为超声波，低于 20 赫兹为次声波。

2.1.3 模拟音频处理设备

(1) 话筒（Microphone 麦克风）(2) 音箱（speaker，扬声器）(3) 模拟调音台

2.2 音频数字化

2.2.1 数字音频

数字音频是指用一连串二进制数据来保存的声音信号。这种声音信号在存储和电路传输及处理过程中，不再是连续的信号，而是离散的。

2.2.2 音频的数字化过程

数字化的音频信号两种途径：

第一种途径就是将现场声源的模拟信号或已存储的模拟声音信号通过某种方法转换成数字音频；

第二种途径就是在数字化设备中创作出数字音频，比如电子作曲。

音频数字化一般经过三个阶段“采样——量化——编码”。

2.2.2 音频的数字化过程

音频数字化过程的具体步骤包括：

第一步，将麦克风转化过来的模拟电信号以某一频率进行离散化的样本采集，这个过程就叫采样；

第二步，将采集到的样本电压或电流值进行等级量化处理，这个过程就是量化；

第三步将等级值变换成对应的二进制表示值（0 和 1），并进行存储，这个过程就是“编码”。

通过这三个环节，连续的模拟音频信号即可转换成离散的数字信号——二进制的 0 和 1。

2.2.2 音频的数字化过程

数字化过程两个指标：

一是量化深度，也可称之为量化分辨率，是指单位电压值和电流值之间的可分等级数；

二是采样频率，即采样点之间的时间间隔。

两者与音质还原的关系是：采样频率越高，量化深度越大，声音质量越好。

在数字音频的衡量指标中，采样频率的单位是 Hz，量化深度一般用比特（Bit）来度量。

根据 Nyquist（奈奎斯特）定理：要想不产生低频失真，则采样频率至少是录制的最高频率的两倍

2.3 数字音频质量及格式

2.3.1 音频数据率及质量

•数字音频文件存储在计算机中要占据一定的空间，然而不同的采样频率、量化深度和录制时间生成的音频文件大小也不同。

•例如：用 44.1kHz、16bit 来进行立体声（即两个声道）采样（标准的 CD 音质），录制（或采集）3 分钟的音频，那么在该未经压缩的声音数据文件的大小为：

一秒钟内采样 44.1×1000 次，每次的数据量是 $16 \times 2 = 32\text{bit}$ （因为立体声是两个声道），那么 3 分钟的总共数据量是 $44100 \times 32 \times 60 \times 3 = 254016000$ （bit），换算成计算机中的常用单位(Byte)，总共数据量是 $254016000 / 8 / 1024 / 1024 = 30.28\text{MByte}$ 。

•**比特率**，即每秒钟音频的二进制数据量。

—上述例题的比特率是 176.4kB/s。

2.3.1 音频数据率及质量

衡量一个数字音频的音质好坏的话，通常可以参考以下指标：

采样频率：即采样点之间的时间间隔，采样间隔时间越短，音质越好。

量化深度(量化分辨率)：是指单位电压值和电流值之间的可分等级数，可分等级越多，音质越好。**音频流**

码率：数字化后，单位时间内音频数据的比特容量，流码率越大音质越好。以

上三个方面的指标中，前面两个指标是绝对指标，而音频流码率是一个**相对指标**，可以间接用来考察音频的质量。

2.3.2 声音文件格式

数字音频的常见格式有以下几种：

(1) **WAV** 格式:WAV 格式支持许多压缩算法，支持多种音频位数、采样频率和声道，WAV 格式对存储空间需求太大不便于交流和传播。(2) **MIDI** 格式:MIDI 是 Musical Instrument Digital Interface 的缩写，又称作**乐器数字接口**，是数字音乐/电子合成乐器的统一国际标准。它定义了计算机音乐程序、数字合成器及其它电子设备交换音乐信号的方式，规定了不同厂家的电子乐器与计算机连接的电缆和硬件及设备间数据传输的协议，可以模拟多种乐器的声音。(3) **CDA** 格式，其取样频率为 44.1kHz，16 位量化位数。CD 存储采用了**音轨**的形式，记录的是**波形流**，是一种近似**无损**的格式。(4)MP3 格式:其全称是 MPEG-1 Audio Layer 3。MP3 能够以**高音质、低采样率**对数字音频文件进行压缩。(5) MP3 Pro 格式:MP3 Pro 可以在基本不改变文件大小的情况下改善原先的 MP3 音乐音质。它能够在用较低的比特率压缩音频文件的情况下，最大程度地保持压缩前的音质。(6) **WMA** 格式:WMA 格式是以减少数据流量但保持音质的方法来达到更高的压缩率目的，其压缩率一般可以达到 1:18。此外，WMA 还可以通过 DRM (Digital Rights Management) 方案加入防止拷贝，或者加入限制播放时间和播放次数，甚至是播放机器的限制，可有力地**防止盗版**。(7)MP4 格式:MP4 在文件中采用了保护版权的编码技术，只有特定的用户才可以播放，有效地保证了音乐版权的合法性。另外 MP4 的压缩比达到了 1:15，体积较 MP3 更小，但音质却没有下降。不过因为只有特定的用户才能播放这种文件，因此其流传与 MP3 相比差距甚远。(8) SACD 格式:SACD (SA=SuperAudio) 是由 Sony 公司正式发布的。它的采样率为 CD 格式的 64 倍，即 2.8224MHz。SACD 重放频率带宽达 100kHz，为 CD 格式的 5 倍，24 位量化位数，远远超过 CD，声音的细节表现更为丰富、清晰。(9) Quicktime 格式:QuickTime 是苹果公司于 1991 年推出的一种数字流媒体，它面向视频编辑、Web 网站创建和媒体技术平台，QuickTime 支持几乎所有主流的个人计算平台，可以通过互联网提供实时的数字化信息流、工作流与文件回放功能。(10) VQF 格式:VQF 格式是由 YAMAHA 和 NTT 共同开发的一种音频压缩技术，它的压缩率能够达到 1:18，因此相同情况下压缩后 VQF 的文件体积比 MP3 小 30%~50%，更便于网上传播，同时音质极佳，接近 CD 音质(16 位 44.1kHz 立体声)。但 VQF 未公开技术标准，至今未能流行开来。(11) DVD Audio 格式:DVD Audio 是新一

代的数字音频格式，与 DVD Video 尺寸以及容量相同，为音乐格式的 DVD 光碟，取样频率为“48kHz/96kHz/192kHz”和“44.1kHz/88.2kHz/176.4kHz”可选择，量化位数可以为 16、20 或 24 比特，它们之间可自由地进行组合（12）MD 格式：采用 44.1kHz 采样的立体声音乐，使用了 ATRAC 算法（自适应声学转换编码）压缩音源。（13）RA 格式：RealAudio 是由 Real Networks 公司推出的一种文件格式，最大的特点就是可以实时传输音频信息，尤其是在网速较慢的情况下，仍然可以较为流畅地传送数据。（14）Liquid Audio 格式：Liquid Audio 是一家提供付费音乐下载的网站。它通过在音乐中采用自己独有的音频编码格式来提供对音乐的版权保护。Liquid Audio 的音频格式就是所谓的 LQT。（15）Audible 格式：Audible 拥有四种不同的格式：Audible1、2、3、4。格式 1、2 和 3 采用不同级别的语音压缩，而格式 4 采用更低的采样率和 MP3 相同的解码方式，所得到语音更清楚，而且可以更有效地从网上进行下载。（16）VOC 文件格式：在 DOS 程序和游戏中常会遇到这种文件，它是随声霸卡一起产生的数字声音文件，与 WAV 文件的结构相似，可以通过一些工具软件方便地互相转换。（17）AU 格式：AU 文件是 UNIX 操作系统下的数字声音文件，由于早期 Internet 上的 Web 服务器主要是基于 UNIX 的，所以这种文件成为 WWW 上唯一使用的标准声音文件。（18）AIFF 格式：AIFF（.AIF）是苹果公司开发的声音文件格式，被 Macintosh 平台和应用程序所支持。（19）Amiga 格式：Amiga 声音（.SVX）：Commodore 所开发的声音文件格式，被 Amiga 平台和应用程序所支持，不支持压缩。（20）MAC 格式：Apple 计算机公司所开发的声音文件格式，被 Macintosh 平台和多种 Macintosh 应用程序所支持，支持某些压缩。（21）S48 格式：S48（stereo、48kHz）采用 MPEG-1 layer 1、MPEG-1 layer 2 声音压缩格式，由于其易于编辑、剪切，所以在广播电台应用较广。（22）AAC 格式：AAC 实际上是高级音频编码的缩写。AAC 是 MPEG-2 规范的一部分。AAC 的音频算法在压缩能力上远远超过了以前的一些压缩算法。它还同时支持多达 48 个音轨、15 个低频音轨、更多种采样率和比特率、多种语言的兼容能力、更高的解码效率。

2.4 数字音频的编辑技术

2.4.1 数字音频的编辑方式

对于数字音频的技术操作具体来说可以归纳为以下六个方面的内容。

（1）**数字录音**：该技术操作是指通过数字方式将自然界中的声源或者存储在其它介质的模拟声音通过“采样——量化——编码”的方式将其变成计算机中或其它数字音频设备中能够识别的数字声音。（2）**数字音乐创作**：该技术操作是指通过相关的数字音频创作工具（如：计算机和 Midi 键盘、Midi 吉它等）直接生成创作数字音频，通常是数字音乐。（3）**声音剪辑**：该操作旨在对数字音频素材进行裁剪或者复制。（4）**合成声音**：也可称之为**混音**，声音合成是指根据需要，把多个声音素材叠加在一起，生成混合效果。（5）**增加特效**：增加特效是指对原始的数字音频素材进行听觉效果的优化调整，以使其符合需要。（6）**文件操作**：对数字音频的文件操作是指对整个音频文件进行的操作，而非改变其音色、音效。

2.4.2 数字音频设备

数字音频处理设备可以分为两类：一类是专用数字音频设备，另一类是非专为处理音频而设计的多媒体计算机。

（1）数字调音台：调音台的作用有两个：其一是将每一路进行优化和调节；其二，对多路声音进行混合输出。（2）数字录音机，采用数字记录方式来存储音频信号。一般可用硬盘记录方式或者光盘记录方式。（3）数字音频工作站：数字音频工作站是一台能够完成从录音、编辑、混合、压缩，一直到最后刻出母盘的全部音频节目制作过程的设备。它最大的特点就是集成度高，免去录音连线的烦恼，且便于携带。

在进行数字音频处理时，除了用到上述几种专用的硬件设备外，还会用到一些其它配套设备，如麦克风、音箱等等。

在处理数字音频时，其关键的硬件技术内核包括：

1）模数转换器：模数转换器（Analog to Digital Converter，ADC）是一个硬件芯片，一般用在带录音功能的音频处理设备之中，其作用就是将模拟的音频电压（流）信号转成数字脉冲电压（PCM）信号。任何 ADC 都包括上面提到的三个基本功能：采样、量化和编码，用来完成从模拟的音频信号向数字音频信号的采集过程。2）数模转换器：数模转换器（Digital to Analog Converter，DAC）也是一个硬件芯片，一般

用在数字音频的重放设备中，用来将数字音频信号还原成模拟的音频信号。3) 数字信号处理器: 数字信号处理器 (Digital Signal Processor, DSP) 是一个专门用来处理数字声音的微型处理器，类似于计算机中的 CPU，可以用来模拟和产生声场，并对声音效果进行控制。

数字音频编辑软件可分为两种：一种是**音源软件**（音序器软件），主要是针对数字音乐创作而言。它是一种可以用来产生和模拟各种乐器或发声物的应用软件。**音源软件中最核心的是音序器**，其主要作用是把音乐元素或事件进行系列或序列编程。另外一种软件是**编辑软件**，可以完成对声音的录音、剪辑、混音合成、特效处理。

(1) 声道 声道 (Sound Channel) 是指声音在录制或播放时**在不同空间位置采集或回放**的相互独立的音频信号，所以声道数也就是声音录制时的音源数量或回放时相应的扬声器数量。

(2) 音轨 音轨就是在音频处理软件中看到的一条一条的平行“轨道”。每条音轨分别定义了该条音轨的属性，如音轨的音色，音色库，通道数，输入/输出端口，音量等。

在音序器软件中一条音轨对应于音乐的一个声部或者对于一种乐器，它把 MIDI 或者音频数据记录在特定的时间位置。对于音频工作站软件而言，每一音轨对于一个原始音频素材文件或者前后对应多个音频文件。所有的音频处理软件都可以允许多音轨操作，也就是在某一段时间内，可以同时让多个音频素材同时播放，产生混音效果。

(3) 时序 所谓时序，其实也就是时间的顺序。

2.5 数字音频技术应用

下面就简单的介绍数字音频技术在几个方面的具体应用。

(1) 数字广播 (2) 音乐制作 (3) 影视游戏配乐 (4) 个人家庭娱乐

本章学习目标:

(1) 了解颜色模型概念、颜色表示方法; 理解各模型与 RGB 模型之间的变换方法。(2) 掌握图像的基本属性。(3) 了解位图和矢量图文件的获取和编辑方法。(4) 能较熟练操作数码相机、扫描仪和绘图板等。(5) 初步掌握 Photoshop 软件的核心概念以及基本操作。

3.1 图像颜色的模型

图像颜色的模型，即颜色的表示模型，通常简称为颜色模型，被用来描述人们能感知的和处理的颜色。

在颜色模型中，所有被定义的颜色形成了坐标系的彩色空间。每一种颜色表示颜色坐标系中的一个点，可以使用数值来衡量。

常见的颜色模型包括 RGB (红色、绿色、蓝色); CMYK (青色、洋红、黄色、黑色); HSB (色相、饱和度、亮度); YUV; CIE L*a*b 等。

一般来说，显示时采用 RGB 颜色模型，印刷用 CMYK 颜色模型，彩色全电视信号数字化采用 YUV 颜色模型。

为了便于彩色处理和识别，视觉系统又常采用 HSB 颜色模型。

3.1.1 视觉系统对颜色的感知

• 眼睛看到的自然景观或图像，除了本身的特征外，还与一个重要的因素：颜色。

• 在同一种光线条件下，之所以会看到不同景物具有各种不同的颜色，这是因为**物体的表面具有吸收或反射不同光线的能力**。

• 光不同，眼睛就会看到不同的色彩。色彩的发生，是光对人的视觉和大脑发生作用的结果，是一种视知觉。由此看来，需要经过“光——眼——神经”的过程才能见到色彩。

• 当人的眼睛受到 **380~780nm** 范围内可见光谱的刺激以后，除了有亮度的反应外，同时产生色彩的感觉。一般情况下光进入视觉通过以下三种形式:

(1) 光源 光源发出的色光直接进入视觉，像霓虹灯、日光灯、蜡烛等的光线都可以直接进入视觉。

(2) 透射光 光源穿过透明或半透明物体后再进入视觉的光线，称为透射光。透射光的亮度和颜色取决于入射光穿过被透射物体之后所达到的光透射率及波长特征。

(3) 反射光 **反射光是光进入眼睛的最普遍的形式**。在有光线照射的情况下，眼睛能看到的任何物体都是该物体的反射光进入视觉所致。

3.1.2 RGB 颜色模型

RGB 颜色模型是颜色最基本的表示模型，也是计算机系统彩色显示器采用的颜色模型。其中，R，G，B 分别代表红 (Red)、绿 (Green)、蓝 (Blue) 三色。RGB 颜色模型也称为加色模型，各种颜色由不同比例红、绿、蓝 3 种基本色的叠加而成。当三基色按不同强度相加时得到的颜色称为相加色。 $r[R]$ 、 $g[G]$ 、 $b[B]$ 三个分量各占一个字节 (8 位)，这样共可表示 $2^{24}=16\ 777\ 216$ 种颜色。

3.1.3 CMYK 颜色模型

CMYK 模型以打印在纸上的油墨的光线吸收特性为基础。

当白光照射到半透明油墨上时，色谱中的一部分被吸收，而另一部分被反射回眼睛。哪些光波反射到眼睛中，决定了人们能感知的颜色。

CMYK 模型中也定义了颜料的三种基本颜色——青色 (Cyan)、品红 (Magenta) 和黄色 (Yellow)。在理论上说，任何一种颜色都可以用这三种基本颜料按一定比例混合得到。

由于所有打印油墨都包含一些杂质，因此这三种油墨实际生成土灰色，必须与黑色 (K) 油墨合成才能生成真正的黑色。与 RGB 模型相对，CMYK 模型被称为减色模型。

理论上，在相减混色中，等量黄色 (Y) 和品红 (M) 相减而青色 (C) 为 0 时，得到红色 (R)；等量青色 (C) 和品红 (M) 相减而黄色 (Y) 为 0 时，得到蓝色 (B)；等量黄色 (Y) 和青色 (C) 相减而品红 (M) 为 0 时，得到绿色 (G)。100% 的三种基本颜料合成将吸收所有颜色而生成黑色。

HSB 模型

建立在人类对颜色的感觉基础之上。H 表示色调 (也称色相)、S 表示饱和度、B 表示亮度。

色调反映颜色的种类，是人们眼看一种或多种波长的光时产生的彩色感觉。

饱和度是指颜色的深浅程度或纯度，即各种颜色混入白色的程度。对同一色调的光，饱和度越高则颜色越鲜艳或者说越纯。色调和饱和度通常统称为色度。

亮度是颜色的相对明暗程度。

3.1.5 YUV 与 YIQ 颜色模型

彩色全电视信号采用 YUV 和 YIQ 模型表示彩色电视的图像。

不同的电视制式采用的颜色模型不同。我国和一些西欧国家采用 PAL 电视制式 (在下一章中有讲解)，在 PAL 彩色电视制式中使用 YUV 模型，其中的 Y 表示亮度，UV 用来表示色差，U、V 是构成彩色的两个分量；在美国、加拿大等国采用的 NTSC 彩色电视制式中使用 YIQ 模型，其中的 Y 表示亮度，I、Q 是两个彩色分量。

采用 YUV 颜色模型的有两个优点：

一个优点是解决了彩色电视与黑白电视的兼容问题。这样使黑白电视能够接收彩色电视信号。

另一个优点是可以利用人眼的特性来降低数字彩色图像所需要的存储容量。一幅大小为 640×480 像素的彩色图像，用 8:2:2 YUV 格式 (即 Y 分量用 8 位表示，而对每四个相邻像素 (2×2) 的 U、V 值分别用相同的一个值表示) 来表示，所需要的存储容量为 $640 \times 480 \times (8+2+2)/8=460\ 800$ 字节。若采用 RGB 8:8:8 格式表示，所需要的存储容量为 $640 \times 480 \times (8+8+8)/8=921\ 600$ 字节。

我国的 PAL/D 制式中，亮度 Y 的带宽为 6MHz，色差 U、V 的带宽为 1.3MHz。

3.1.6 CIE L*a*b 颜色模型

L*a*b 颜色模型设计目的是为了得到不依赖于具体设备的颜色标准，从而在实际使用中不论使用何种设备 (如显示器、打印机、计算机或扫描仪) 均能制作和输出完全一致的颜色。

L*a*b 颜色由亮度或光亮度分量 (L) 和两个色度分量组成：a 分量保存从绿色到红色所对应的色彩信息；b 分量保存从蓝色到黄色所对应的色彩信息，单个 a 或 b 无意义，只有 a、b 结合才有意义。

3.2.4 YCrCb 与 RGB 彩色空间变换

YCrCb 与 YUV 的定义基本上是相同的，但应用有所不同。YUV 适用于 PAL 和 SECAM 彩色电视制式的模拟视频图像的表示，而 YCrCb 则适用于数字电视以及计算机用数字视频图像的表示。数字域中的彩色空间变换与模拟域的彩色空间变换不同，YCrCb 与 RGB 空间的转换关系如下：

$$Y=0.299R+0.578G+0.114B \quad Cr=(0.500R-0.4187G-0.0813B)+128$$

$$Cb=(-0.1687R-0.3313G+0.500B)+128$$

3.3 图像的基本属性及种类

图像的属性包含分辨率、颜色深度、文件大小、真/伪彩色、图像的种类等。

3.3.1 分辨率

与图像相关的分辨率包括显示分辨率、图像分辨率、打印分辨率和扫描分辨率等。

(1) **显示分辨率** 显示分辨率是指显示屏上水平和垂直方向上的最大像素点的个数。例如，显示分辨率为 640×480 表示显示屏垂直方向显示 480 个像素，水平方向显示 640 个像素，整个显示屏共含有 307,200 个显像点。显示设备的分辨率越高，屏幕能够显示的像素越多，因此能够显示的图像就越大越精细。分辨率不仅与显示尺寸有关，还要受显像管点距、视频带宽等因素的影响。

(2) **图像分辨率** 图像分辨率是指一幅图像在水平和垂直方向上的最大像素点的个数。若图像像素点距固定，则图像分辨率越大则图像越大；若图像大小一样，图像的分辨率越大，则组成该图的图像像素数目越多，图像看起来就越细致逼真。

图像在显示设备的显示效果与图像分辨率和显示分辨率相关。当图像分辨率大于显示分辨率，显示屏幕仅会显示图像的一部分。当图像分辨率小于显示分辨率，图像则只占显示屏幕的一部分。

(3) **扫描分辨率与打印分辨率** 扫描分辨率是指用扫描仪扫描图像的扫描精度，通常用每英寸多少点 (Dots Per Inch, DPI) 表示。图像扫描后的效果很大程度上决定于原图像的精度，但使用扫描仪时选择扫描的精度将直接影响扫描后的图像质量。扫描分辨率越大，得到的图像就越大，像素就越多。例如，用 300 DPI 来扫描一幅 $8'' \times 10''$ 的彩色图像，将得到一幅 2400×3000 个像素的图像。扫描图像的精度一般选择 100-150DPI，若要进行 OCR 识别，为提高识别率，则要将扫描精度上调至 300DPI 以上。

打印分辨率是指图像打印时每英寸可识别的点数，也使用 DPI 为衡量单位。打印分辨率越大，在打印纸张大小不变的情况下，打印的图像将越精细。

3.3.2 颜色深度

颜色深度指一幅图像中的最多使用的颜色数，用来度量在图像中有多少颜色信息来显示或打印像素。较大的颜色深度意味着数字图像具有更多的可用颜色和更精确的颜色表示。

颜色编码二进制位数即为图像的颜色深度值。1 位二进制颜色编码的图像颜色深度为 1，最多有 2^1 种颜色，即每个像素只能有两个可能的颜色值：黑色和白色；4 位颜色的图像，它的颜色深度是 4，它可有 2^4 中颜色（或 16 种灰度等级）；8 位颜色的图像，颜色深度就是 8，它含有 $2^8=256$ 种颜色（或 156 种灰度等级）。

3.3.2 颜色深度

• **24 位** 颜色可称之为真彩色，位深度是 24，它能组合成 2 的 2^4 次幂种颜色，即：16,777,216 种颜色（或称千万种颜色），超过了人眼能够分辨的颜色数量。当用 24 位来记录颜色时，实际上是以 $2^8 \times 2^8 \times 2^8 = 2^{24}$ ，即红、绿、蓝 (RGB) 三基色各以 2 的 8 次幂，256 种颜色而存在的，三色组合形成了一千六百多万种颜色。

32 位 颜色的位深度是 32，实际上是 $2^8 \times 2^8 \times 2^8 \times 2^8 = 2^{32}$ ，即青、洋红、黄、黑 (CMYK) 四种颜色各以 2 的 8 次幂，256 种颜色而存在，四色的组合就形成 4,294,967,296 种颜色，或称为超千万种颜色。

• 常用的颜色深度值范围为 1 到 32。事实上，由于目前的计算机或其它显示设备只能显示 RGB 色彩，即 2^{24} 的真彩色，所以大于这个数值的颜色深度是“不真实”的，也不能完全表现出来。

3.3.3 真彩色、伪彩色与直接色

(1) **真彩色(true color)** 真彩色是指图像颜色与显示设备显示的颜色一致，即组成一幅彩色图像的每个像素值的 R, G, B 三个基色分量都直接决定显示设备的基色强度，这样产生的彩色被称为真彩色。

(2) **伪彩色(pseudo color)** 伪彩色图像的含义是，每个像素的颜色不是由每个基色分量的数值直接决定，而是把像素值当作调色板或彩色查找表的表项入口地址，去查找一个显示图像时使用的 R, G, B 强度值，如果图像中的颜色在调色板或彩色查找表中不存在，则调色板会用一个最接近的颜色来匹配。用查找出的 R, G, B 强度值产生的彩色不是图像本身真正的颜色因此称为伪彩色。

(3) **直接色(direct color)** 每个像素值分成 R, G, B 分量，每个分量作为单独的索引值对它做变换。

也就是通过相应的彩色变换表找出基色强度，用变换后得到的 R，G，B 强度值产生的彩色称为直接色。它的特点是对每个基色都进行了变换。这一点伪彩色是有区别的。

3.3.4 图像的大小及种类

图像大小是指图像以数字表示的大小，单位是千字节(K)、兆字节(MB)或千兆字节(GB)。其大小主要受图像像素和颜色深度影响，与图像的像素数成正比。如一幅图像分辨率为 768*576，颜色深度为 24 的图像的大小为： $768*576*3/(1024*1024)=1.26\text{ MB}$ ，其中 768*576 为图像的总像素个数，每个像素用 24 位表示即为 $24(\text{bit})/8=3\text{Byte}$ 。计算中，第一次除以 1024 得到单位为 KB 第二次除以 1024 得到单位为 MB，最后得到该图像的大小为 1.26MB。

按照图像在计算机中显示时不同的生成方式可以将图像分为矢量图（形）和点位图（像）。所谓矢量图是用一系列计算机指令来表示一幅图，如点、线、曲线、圆、矩形等。在显示图时，也往往能看到画图的过程。矢量图有许多优点，由于矢量图可通过公式计算获得，所以矢量图文件体积一般较小。矢量图文件的大小主要取决于图的复杂程度。当然，矢量图最大的优点还在于当它被放大、缩小或旋转等不会失真。矢量图与分辨率无关，可以将它缩放到任意大小和以任意分辨率在输出设备上打印出来，都不会影响清晰度。

它最大的缺点是难以表现色彩层次丰富的逼真图像效果，遇到这种情况往往就要采用点位图表示。

点位图也简称位图。点位图与矢量图不同，它是把一幅图分成许多的像素，每个像素用若干个二进制位来指定该像素的颜色、亮度和属性。因此一幅图由许多描述每个像素的数据组成，而这些数据作为一个文件来存储，这种文件又称为位图文件。位图与分辨率有关，换句话说，它包含固定数量的像素，代表图像数据。另外，点位图文件占据的存储器空间比较大。但是点位图在表现复杂的图像和丰富的色彩方面有明显的优势。

矢量图和点位图之间可以用软件进行转换，由矢量图转换成点位图采用光栅化(rasterizing)技术，这种转换也相对容易；由点位图转换成矢量图用跟踪(tracing)技术，这种技术在理论上说是容易，但在实际中很难实现，对复杂的彩色图像尤其如此。

所谓图像格式即图像文件存放在存储器上的格式，各种文件格式通常是特定的应用程序创建的。这些文件格式大致上可以分为两大类：一类是属于位图图像文件格式，另一类是属于矢量图形的文件格式。

位图图像常用的文件格式有以下几种：

(1) PSD 图像格式

扩展名是 PSD, 全名为 Photoshop Document，它是 Photoshop 的专用文件格式，也是惟一可以存取所有 Photoshop 特有的文件信息以及所有彩色模式的格式。

(2) BMP 图像格式

扩展名是 BMP, 全名为 Bitmap-File。它是 Windows 采用的图像文件存储格式，在 Windows 环境下运行的所有图像处理软件都支持这种格式。该图像格式采用的是无损压缩，因此其优点是图像完全不失真，其缺点是图像文件的尺寸较大。

(3) JPEG 图像格式

扩展名是 JPG, 全名为 Joint Photographic Experts Group。它利用一种失真式的图像压缩方式将图像压缩在很小的存储空间中，其压缩比率通常在 10:1~40:1 之间。这样可以使图像占用较小的空间，所以很适合应用在网页的图像中。JPEG 格式的图像主要压缩的是高频信息，对色彩的信息保留较好，因此也普遍应用于需要连续色调的图像中。

(4) GIF 图像格式

(5) 扩展名是 GIF, 全名是 Graphics Interchange Format。此种格式的图像特点是文件尺寸较小，支持透明背景，特别适合作为网页图像。此外，GIF 文件格式可在一个文件中存放多幅彩色图形、图像。如果在 GIF 文件中存放有多幅图，它们可以像演幻灯片那样显示或者像动画那样演示。

(6) TIFF 图像格式

扩展名是 TIF, 全名是 Tagged Image File Format。它是一种非失真的压缩格式(最高也只能做到 2~3 倍的压缩比)能保持原有图像的颜色及层次，但占用空间却很大。例如一个 200 万像素的图像，差不多要占用

6MB 的存储容量，故 TIFF 常被应用于比较专业的用途，如书籍出版、海报等，极少应用于互联网上。

矢量图形的文件格式主要有以下几种：

(1) CDR 格式 CDR 是 CorelDraw 中的一种矢量图形文件格式。它是所有 CorelDraw 应用程序中均能够使用的一种文件格式。

(2) DWG 格式 DWG 是 AutoCAD 中使用的一种图形文件格式。

(3) DXF 格式 DXF 是 AutoCAD 中的图形文件格式，它以 ASCII 码方式存储图形，在表现图形的大小方面十分精确，可被 CorelDraw、3dMAX 等大型软件调用编辑。

(4) EPS EPS 是用 PostScript 语言描述的一种 ASCII 图形文件格式，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。

3.4.1 位图的获取设备与技术

位图的获取通常用扫描仪，以及摄像机、录相机、激光视盘与视频信号数字化卡一类设备。

获取位图图像的三种常用方法：

(1) 通过数字转换设备采集，如：扫描仪或视频采集卡 (2) 通过数字化设备摄入，如：数码相机、数字摄像机 (3) 从数字图库中收集，如：光盘、网络、硬盘。

3.4.1 位图的获取设备与技术

3.4.1.1 扫描仪

扫描仪 (Scanner) 是一种高精度的光电一体化的高科技产品，它是将各种形式的图像信息输入计算机的重要工具，是继键盘和鼠标之后的第三代计算机输入设备。

3.4.1 位图的获取设备与技术

(1) 扫描仪种类及原理

目前市场上流行的扫描仪有：手持式扫描仪、滚筒式扫描仪和平台式扫描仪。

平台式扫描仪也称平板式扫描仪、台式扫描仪。这类扫描仪光学分辨率在 300dpi-8000dpi 之间，色彩位数从 24 位到 48 位，扫描幅面一般为 A4 或者 A3

小滚筒式扫描仪是手持式扫描仪和平台式扫描仪的中间产品。小滚筒式的设计是将扫描仪的镜头固定，而移动要扫描的物件通过镜头来扫描，运作时就像打印机那样，要扫描的物件必须穿过机器再送出，因此被扫描的物体不可以太厚。

其它扫描仪还包括大幅面扫描用的大幅面扫描仪、笔式扫描仪、条码扫描仪、底片扫描仪、实物扫描仪，还有主要用于专业印刷排版领域的滚筒式扫描仪等。

扫描仪一般由光源、光学透镜、扫描模组、模拟—数字转换电路加塑料外壳构成。

扫描仪的性能参数包括光电器件、分辨率等。

(3) 扫描仪的分辨率

扫描仪的分辨率分为光学分辨率和最大分辨率。其中，光学分辨率是指扫描仪物理器件所具有的真实分辨率，而最大分辨率相当于插值分辨率，并不代表扫描仪的真实分辨率。

光学分辨率是扫描仪的光学部件在每平方英寸面积内所能捕捉到的实际的光点数，它的数值是由光电元件所能捕捉的像素点除以扫描仪水平最大可扫尺寸得到的数值。

一般来说，扫描仪的纵向分辨率是横向分辨率的两倍，有时甚至是四倍。判断扫描仪光学分辨率时，应以最小的一个为准。

•最大分辨率其实就是插值分辨率，是通过在相邻像素间求出颜色或灰度的平均值，从而通过计算机对图像进行分析，对空白部分进行数学填充（这一过程也叫插值处理），得以增加像素数的办法而提高的分辨率。在最大分辨率方式下可以增加图像的像素值，但不能增添更多的图像细节。

•色彩分辨率，它是表示扫描仪分辨彩色或灰度细腻程度的指标，它的单位是 bit (位)。色彩分辨率的含义是用多少个位来表示扫描得到的一个像素。例如:1bit 只能表示黑白像素，它们分别代表黑与白。8bit 可以表示 256 个灰度级 (2⁸=256)，它们代表从黑到白的不同灰度等级。

3.4.1.2 数码相机

（1）数码相机的工作原理

数码相机的工作原理如下：当按下快门时，镜头将光线会聚到感光器件 CCD（电荷耦合器件）上，它代替了普通相机中胶卷的位置，功能是把光信号转变为电信号。感光器件得到了对应于拍摄景物的电子图像，A/D（模/数转换器）器件获得数字信号。微处理器（MPU）对数字信号进行压缩并转化为特定的图像格式，例如 JPEG 格式。最后，图像文件被存储在内置存储器中。通过数码相机，我们可以通过 LCD（液晶显示器）查看拍摄到的照片。

（2）数码相机的工作部件

数码相机是由镜头、CCD、A/D（模/数转换器）、MPU（微处理器）、内置存储器、LCD（液晶显示器）、PC 卡（可移动存储器）和接口（计算机接口、电视机接口）等部分组成，通常它们都安装在数码相机的内部。

3.4.2 矢量图的获取设备与技术

矢量图形是通过电脑的绘图软件创作并在电脑上绘制出来。

（1）什么是绘图板

•绘图板，又叫数位板，它是一种专门针对电脑绘图而设计的输入设备，通常由一块板子和一支笔组成，主要面向美工、设计师或者绘图工作者、美术爱好者等用户。

绘图板最大的特色就是具备压力感测功能（与写字板的区别）

（2）绘图板的功能

绘图板可以模拟各种各样的画家的画笔，如模拟最常见的毛笔，用力重时，毛笔能画很粗的线条，很轻时它可以画出很细很淡的线条。它也可以模拟喷枪，而且还能根据笔倾斜的角度，喷出扇形等等的效果。除了模拟传统的各种画笔效果外，它还可以利用电脑的优势，实现使用传统工具无法实现的效果，例如根据压力大小进行图案的贴图绘画，使用者只需要轻轻几笔就能很容易绘出一片开满大小形状各异的鲜花的芳草地。

绘图板作为一种硬件输入工具，结合 Painter、Photoshop 等绘图软件，可以创作出各种风格的作品：油画、水彩画、素描...

绘图板除了可用于绘画之外，还可以用于照片处理、3D 制作、游戏制作和影视制作等。

（3）绘图板与手写板的区别

手写板的作用和键盘类似，局限于输入文字或者绘画，也具有一些鼠标的功能。手写板一般是使用手写笔，或者手指在特定的区域内书写文字，通过各种方法将笔或者手指走过的轨迹记录下来，然后识别为文字。与手写板不同，绘图板是一种专门针对电脑绘图而设计的输入设备，主要面向美工、设计师或者绘图工作者、美术爱好者等用户。手写板则主要是为了解决文字输入而设计的。

在原理上，大部分绘图板是靠电磁笔的感应识别的，可以和桌面的分辨率进行绝对对应的。

手写笔上一般没有感应装置，甚至可以用别的东西例如牙签、手指等代替手写笔在上面进行书写。

绘图板的主要性能指标包括压感级数、有效尺寸、解析度、读取速度和接口类型等。

3.5.1 图像处理软件简介

图像处理指对已有的数字图像进行再编辑和处理。图像处理的软件包很多，目前常用的有：Photoshop、Firework、Acdsee、Windows 的画图软件等。

在 Windows 环境下，这些图像应用软件在其功能上都具有一定的共性，包括：

（1）支持多种图像数据格式，具有图像编辑、变形变换、优化处理等功能。（2）可选定某个区域进行裁剪、复制、粘贴、水平或竖直翻转、镜像、旋转、变形、透视等操作。（3）具有不同的效果处理功能，包括可调亮度、色度、去噪音、模糊、锐化、边界等，还包括其他一些特技。（4）具有一定的绘图功能。

3.5.2 图像处理基本概念

Photoshop 的主要核心概念包括图层、通道、滤镜、蒙板、路径、选择区、文字与绘图。

3.5.2.1 图层

Photoshop 将一个图像按不同的图层（layer）来记录和编辑。一个图层就是一个相对独立的图像单元，每层可以独立选择和处理，而决不会影响到其他的图层信息。各层之间可按不同的透明度和前后顺序叠在一

起。

3.5.2.2 通道

在 Photoshop 中有两种类型的通道——颜色通道和 Alpha 通道。**颜色通道是用来保存图像颜色信息的**。当用户打开或创建一个新的图像文件时，程序将自动创建颜色信息通道。

Alpha 通道是由用户建立的用于**保存选区**的通道，Alpha 通道中将选区作为 8 位的灰度图像来保存，其中白色部分表示完全选中的区域，黑色部分表示没有选中的区域，而灰色部分表示被不同程度选中的区域。在 Photoshop 中打开一幅图像时会自动产生默认的色彩通道。色彩通道的功能是存储图像中的色彩元素。图像的默认通道数取决于该图像的色彩模式，如 CMYK 色彩模式的图像有四个通道，分别存储图像中的 C、M、Y、K 色彩信息。可以把通道想象成彩色套印时的分色板，每块板对应一种颜色。黑白、灰度、半色调和调配色图只有一个色彩通道；RGB、Lab 图有三个色彩通道，另有一个复合色彩通道用于图像的编辑。一般利用 alpha 通道来保存和编辑蒙板，以便于高级图像编辑，创造出不同的图像效果。

图像的通道越多，文件容量越大。

3.5.2.3 滤镜

滤镜是图像处理软件所特有的，它是众多图像处理软件为进行图像特殊效果处理制作而设计的系统处理接口。

3.5.2.4 蒙板

蒙板（Mask）是浮在图层之上的一块挡板，它本身不包含图像数据，只是**对图层的部分数据起遮挡作用**。当对图层进行操作处理时，被遮挡的数据将不会受影响。

在 Photoshop 中，有三种方式来创建和保存蒙板：图层蒙板、快速蒙板和 alpha 通道方式。

3.5.2.5 路径

路径（Paths）：路径提供一个**精确定义选择区**的方式。路径是一种不包含点阵的矢量对象，因此独立于图像数据之外，也不会被打印输出。

路径是 Photoshop 中的重要工具，其主要用于选取和剪裁复杂的形体轮廓，也可以用来绘制光滑线条。一旦建立了一个路径，可把它保存到路径控制板中，也可以转换成选择区域；还可以用前景色绘制路径曲线或填充路径包围的区域。此外，选择区也可以转换成路径。

3.5.2.6 选区

选择区的作用是在当前层上定义一个编辑区域，使得对该层图像的编辑操作仅对选择区域内的数据有效，而区域外的数据将不会受到影响。

3.5.2.7 文字与绘图

（1）色彩工具（2）绘画工具（3）画线工具：用直线工具可画出不同粗细的直线或带箭头的直线。（4）文字工具：利用文字工具，可在图像中加入各种效果的文字，工具分文字和文字蒙板两类。文字和文字蒙板的区别在于：文字工具产生的文字会直接填入前景色，并生成一个新的图层；而文字蒙板产生的文字实际是文字选择区，还需进一步处理才能形成图像数据。

3.5.3.1 图像处理的一般流程

一般的图像处理流程包括以下七个步骤：

（1）确定图像主题及构图（2）确定成品图的尺寸大小以及画面基调（3）获取基本的数字图像素材（4）对素材进行处理（5）图片上叠加文字说明或绘制的图形（6）整体效果调整（7）图像的输出

4.1 电影的放映原理

人们之所以能够看到电影屏幕上的活动影像，是利用了人眼的视觉暂留特性。科学实验证明，人眼在某个视像消失后，仍可使该物像在视网膜上滞留 0.1—0.4 秒左右。而在电影放映的过程中，电影胶片以**每秒 24 格画面匀速切换**，这就相当于每一格画面给人眼的刺激是 1/24 秒（相当于 0.04），由于人的眼睛有视觉暂留的特性，一个画面的印象还没有消失，下一个稍微有一点差别的画面又出现在银幕上，连续不断的印

象衔接起来，就组成了活动电影。

4.1.2 电视工作原理

电视是根据人眼视觉特性以一定的信号形式实时传送活动景物（或图像）的技术。在发送端，用电视摄像机把景物（或图像）转变成相应的电信号，电信号通过一定的途径传输到接收端，再由显示设备显示出原景物（或图像）。

4.1.3 电视制式简介

•目前世界上现行的彩色电视制式有三种：NTSC 制、PAL 制和 SECAM 制。

上面提到的 NTSC 制、PAL 制和 SECAM 制都是兼容制制式。这里说的“兼容”有两层意思：一是指黑白电视机能接收彩色电视广播，显示的是黑白图像，另一层意思是彩色电视机能接收黑白电视广播，显示的也是黑白图像，这叫逆兼容性。为了既能实现兼容性而又要彩色特性，彩色电视系统应满足下列几方面的要求：

- ①必须采用与黑白电视相同的一些基本参数，如扫描方式、扫描行频、场频、帧频、同步信号、图像载频、伴音载频等等。
- ②需要将摄像机输出的三基色信号转换成一个亮度信号，以及代表色度的两个色差信号，并将它们组合成一个彩色全电视信号进行传送。在接收端，彩色电视机将彩色全电视信号重新转换成三个基色信号，在显像管上重现发送端的彩色图像。

各制式的区别主要就是规定的扫描频率、周期等特性的不同

（1）PAL 制式的主要特性

625 行/帧，25 帧/秒

一帧图像的总行数为 625，分两场扫描。场扫描频率是 50Hz，周期为 20ms；帧频是 25Hz，是场频的一半。

（2）NTSC 制式的主要特性

525 行/帧，30 帧/秒

一帧图像的总行数为 525 行，分两场扫描。场扫描频率是 60 Hz，周期为 16.67 ms；帧频是 30 Hz。

（3）SECAM 制式的主要特性

4.1.4 视频信号类型

彩色视频信号包括三种：复合视频信号、分量视频信号和 S 端子视频信号。

4.2.1 数字化方法

电视图像数字化常用的方法有两种：

- （1）先从复合彩色电视图像中分离出彩色分量，然后数字化。
- （2）首先用一个高速 A/D 转换器对彩色全电视信号进行数字化，然后在数字域中进行分离，以获得所希望的 YCbCr，YUV，YIQ 或 RGB 分量数据。

4.2.3 数字视频属性

（1）数字视频的概念

数字视频是将传统模拟视频(包括电视及电影)片段捕获转换成计算机能处理的数字信号。

4.2.3 数字视频与模拟视频相比有何优点？

首先，数字视频是由一系列二进位数字组成的编码信号，它比模拟信号更精确，而且不容易受到干扰；其次，视频信号数字化后，对数字视频的加工处理只涉及到反映数字视频数据在计算机硬盘中的排列，即访问地址表。播放、剪辑数字视频只是控制着计算机硬盘的磁头读出是 1 还是 0，与信号本身并不接触，不涉及到实际的信号本身。这就意味着不管对数字信号做多少次处理和控制，画面质量几乎是不会下降的，可以多次复制而不失真；

第三、可以运用多种的编辑工具（如编辑软件）对数字视频进行编辑加工，对数字视频的处理方式也是多种多样，可以制作许多特技效果。将视频融入计算机化的制作环境，改变了以往视频处理的方式，也便于视频处理的个人化、家庭化；

最后，数字信号可以被压缩，使更多的信息能够在带宽一定的频道内传输，大大增加了节目资源。并且还

可以突破单向式的数字信号传输，实现交互式的信号传输。

(2) 数字视频的属性

数字视频的参数有：视频分辨率、图像深度、帧率、视频文件格式。

视频分辨率：是指视频信号本身的分辨率，只与视频信号的带宽有关。

图像深度：与静态图像一样，视频的图像深度决定其可以显示的颜色数。某些编码（压缩算法）使用固定的图像深度，在这种情况下该参数不可调整。选择较小的图像深度可以减小文件的容量，但同时也降低了图像的质量。

帧率：每秒的帧数(fps)，用来表示视频文件每秒钟能够显示的帧数。高的帧率可以得到更流畅、更逼真的画面。

压缩质量：选择了一种压缩算法后还可以调整压缩质量，这个参数常用百分比来表示，100%表示最佳效果压缩。同一种压缩算法下，压缩质量越低，文件容量越小，丢失信息也越多。

4.2.4 数字视频文件格式

目前，视频文件格式可以分为适合本地播放的本地影像视频和适合在网络中播放的网络流媒体影像视频两大类。

(1) 本地影像视频

AVI 格式，这种视频格式的优点是图像质量好，可以跨多个平台使用，其缺点是体积过于庞大。压缩标准不统一是其主要问题。

DV-AVI 格式，它可以通过电脑的 IEEE 1394 端口传输视频数据到电脑，也可以将电脑中编辑好的视频数据回录到数码摄像机中。这种视频格式的文件扩展名一般是.avi，所以也叫 DV-AVI 格式。

MPEG 格式，MPEG 文件格式是运动图像压缩算法的国际标准，它采用了有损压缩方法减少运动图像中的冗余信息，从而达到压缩的目的(其最大压缩比可达到 200:1)。

DivX 格式，是由 MPEG-4 衍生出的另一种视频编码(压缩)标准，也叫 DVDrip 格式，它采用了 DivX 压缩技术对 DVD 盘片的视频图像进行高质量压缩，同时用 MP3 或 AC3 对音频进行压缩，然后再将视频与音频合成并加上相应的外挂字幕文件而形成的视频格式。其画质直逼 DVD 并且体积只有 DVD 的数分之一。MOV 格式，美国 Apple 公司开发的一种视频格式，具有较高的压缩比率和较完美的视频清晰度等特点，但是其最大的特点还是跨平台性，即不仅能支持 MacOS，同样也能支持 Windows 系列。

(2) 网络影像视频

ASF 格式，是微软为了和 Real Player 竞争而推出的一种视频格式。用户可以直接使用 Windows 自带的 Windows Media Player 对其进行播放。由于它使用了 MPEG-4 的压缩算法，所以压缩率和图像的质量都很不错。

WMV 格式，也是微软推出的一种采用独立编码方式并且可以直接在网上实时观看视频节目的文件压缩格式。WMV 格式的主要优点包括：本地或网络回放、可扩充的媒体类型、部件下载、可伸缩的媒体类型、流的优先级化、多语言支持、环境独立性、丰富的流间关系以及扩展性等。

RM 格式，这种格式的另一个特点是用户使用 Real Player 播放器可以在不下载音频/视频内容的条件下实现在线播放。另外，RM 作为目前主流网络视频格式，可以通过其 Real Server 服务器将其它格式的视频转换成 RM 视频并由 Real Server 服务器负责对外发布和播放。

RMVB 格式，是一种由 RM 视频格式升级延伸出的新视频格式。RMVB 视频格式打破了原先 RM 格式那种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。这样在保证静止画面质量的前提下，大幅地提高了运动图像的画面质量，从而图像质量和文件大小之间就达到了微妙的平衡。

4.3.1 数字视频的获取方式

数字视频的获取渠道有很多种，其主要的获取途径包括：从现成的数字视频库中截取、利用计算机软件制作视频、用数字摄像机直接摄录和视频数字化。

另外，还可以通过一些重要的设备和技术来实现从模拟视频到数字视频信号的转化，这个过程就称为视频数字化。在实际工作中，电视机、激光视盘、摄像机等都可提供丰富多彩的模拟视频信号，通过视频采集设备获取数字视频。

4.3.2 数字视频获取设备及特性

4.3.2.1 摄像机

(1) 摄像机的组成和功能

摄像机主要由以下几个部分组成：镜头系统、主机、寻像器和附件。

(2) 摄像机的分类

摄像机根据制作节目图像质量的要求可将摄像机分为三档次：广播级、业务级、家用级。

(3) 录像机的分类

•按用途分主要有以下三种：广播级录像机、专业用录像机、家用录像机。

4.4.1 视频编辑基本概念

(1) 什么是视频编辑

视频编辑包括了两个层面的操作含义：其一是传统意义上简单的画面拼接；其二是当前在影视界技术含量高的后期节目包装——影视特效制作。

就技术形式而言，视频可以分为两种形式：线性编辑和非线性编辑。

5.1.1.2 动画的分类

动画可以从不同角度进行分类。

(1) 传统动画和计算机动画 (2) 平面动画和三维动画

从空间的视觉效果上看，又可分为平面动画和三维动画。平面动画又可称为二维动画。这种动画无论画面的立体感有多强，终究只是在二维空间上模拟三维空间效果，同一画面内只有物体的位置移动和形状改变，没有视角的变化。而三维动画中不但有物体本身位置和动作的改变，还可以连续的展现视角的变化。

5.1.2 动画的制作过程

传统动画的制作是一个复杂而繁琐的过程，无论是手绘动画还是模型动画，其基本规律和思路是一致的。简单的说，其关键步骤包含六个：

(1) 由编导确定动画剧本及分镜头脚本；(2) 美术动画设计人员设计出动画人物形象；(3) 美术动画设计人员绘制、编排出分镜头画面脚本；(4) 动画绘制人员进行绘制；(5) 摄影师根据摄影表和绘制的画面进行拍摄；(6) 剪辑配音。

5.1.3 计算机动画及优势

计算机动画又称为数字动画，是指在制作过程中用计算机来辅助或者替代传统制作颜料、画笔和制模工具的一种动画制作方法及其最终成果。

可以从两个方面去理解这一含义：其一，广义上的理解，是指在制作动画时采用数字技术（计算机技术）而得到的动画；其二就是狭义上的理解，是指在制作、存储、传输、重现等过程全部运用数字技术。

按照计算机及其软件在动画制作中的作用而言，计算机动画可分为：计算机辅助动画和计算机创作动画两种。

—计算机辅助动画属二维动画，其主要用途是辅助动画师制作手绘动画，简化手绘动画的工具和手段；

—计算机创作动画则完全用计算机来替代传统动画制作工具而得到的动画，一般也把它称之为“无纸动画”。如网络中常见的Flash动画，一般都是完全用计算机来绘制、作图、上色并使其运动的；又如计算机三维造型动画，则是用计算机建模来替代粘土和钢架的建模。

按照计算机动画制作当中动画运动的控制方式分类。按照这种分类可分为：实时（Real-Time）动画和逐帧动画（Frame-by-Frame）两种。

—逐帧动画也称为帧动画或关键帧动画，在表现画面中某一运动时，将该物体运动的过程在计算机中按照

画面播放的先后顺序逐一的画出来，也即通过一帧一帧显示动画的图像序列而实现运动的效果。

—**实时动画**是用**算法**来实现物体的运动，它并不是将运动物体的动作按照时间点逐一的画出来，而是只记录最开始的状态和最终的状态，中介的运动过程通过计算机自动产生。实时动画也称为算法动画

5.1.3.2 计算机动画的优势

对于制作工艺而言，计算机的使用，大大简化了工作程序，方便快捷，提高了效率。

5.2.1.1 二维动画概述

数字二维动画与传统二维动画到底有哪些相似与不同？

(1) 数字二维动画与传统二维动画的相似之处

—**平面**上的运动。二维动画是在平面上表现运动事物的运动和发展，尽管它有动态的变化，但其单个镜头画面的视点是单一的，不能改变的。即便是数字二维动画亦是如此。

—共同的技术基础——“**分层**”技术。对于传统手绘动画而言，动画师将运动的物体和静止的背景分别绘制在不同的透明胶片上，然后叠加在一起拍摄，其目的是：减少了绘制的帧数，实现透明、景深和折射等不同的效果。而在计算机二维动画技术中也是基于这一概念的，在计算机软件中的各层上绘制，然后直接电脑合成。

共同的**创意来源**。不管是手绘二维动画还是计算机二维动画，最终成果的来源都是来自于创作者——人的创意，而非计算机或者是其它智能化仪器。无论最终是以动画影片的形式出现还是最终以网页和数字游戏的形式出现，人的创意是无法替代的，是首位的，这也是动画制作的基点。

制作的**基本流程**相似。在制作流程的大概阶段上，计算机二维动画与传统手绘动画是相似的，都要经历“总体规划、设计制作、具体创作和拍摄制作”这四个阶段，仅仅在具体步骤和手段上有所差异。

(2) 数字二维动画与传统二维动画的相异之处

—二者的**实现工具和手段**有差异。传统手绘动画在制作工具上依靠的是画笔、颜料、曲尺、明胶片、摄影机和银盐胶片来完成动画的加工和记录过程。而数字二维动画在进行动画制作过程中更多的是依靠计算机，用计算机中鼠标和虚拟的颜料来替代实际的画笔和颜料，用软件中的图层来替代明胶片，用计算机及其硬盘来替代摄影机和银盐胶片。

—二者的创作步骤稍有出入。由于制作工具的不同带来具体操作步骤和方法上的更改。

—二者在最终成果形式以及应用领域上有所不同。传统二维动画一般最终成果形式就是影视作品，只能在电影或者电视屏幕上播放展现的一种作品方式。而现在数字二维动画，其成果形式可以是影视作品，也可以是网络动画、游戏动画或者计算机演示动画。不仅仅局限于剧情和过程的单向展现，更多的应用于互动娱乐之上。

5.2.2.2 二维动画软件中的技术概念

(1) 图层 (2) 时间轴 (3) 图库 (4) 场景

简述用 Flash 制作小球弹跳动画的基本步骤。

5.3.1 三维动画制作的基本流程：造型→赋予材质→添加灯光和摄像机→设置动画→渲染输出

5.5.1 数字动画的应用领域

(1) **电影业** 在电影业方面的应用包括动画影片的制作和数字特效，也就是人们口头上常说的“电脑特效”。(2) **电视片头和电视广告** (3) **科学计算和工业设计** (4) **模拟、教育和娱乐** (5) **虚拟现实和 3D Web**

8.1.1 多媒体数据中的信息冗余有哪些？

空间冗余 结构冗余 时间冗余 视觉冗余 知识冗余 信息熵冗余

8.2 数据压缩分类

按信息压缩前后比较是否有损失进行划分

按信息压缩前后比较是否有损失，可以划分有损压缩和无损压缩。

无损压缩指使用压缩后的数据进行重构，重构后的数据与原来的数据完全相同。常用的无损压缩算法有**霍夫曼(Huffman) 算法**和**LZW 算法**。

有损压缩是指使用压缩后的数据进行重构，重构后的数据与原来的数据有所不同，但不影响人对原始资料表达的信息造成误解。

按数据压缩编码的原理和方法进行划分

—**统计编码**，主要针对**无记忆信源**，根据信息码字**出现概率的分布特征**而进行压缩编码，寻找概率与码字长度间的最优匹配，对于出现频率大的符号用较少的位数表示，而对于出现频率小的符号用较多的位数来表示。（霍夫曼编码、算术编码、香农编码、词典编码）

—**预测编码**是利用空间中**相邻数据的相关性**来进行压缩数据的，编码器记录与传输的不是样本的真实值，而是它与预测值的差。（PCM、DM、DPCM）

—**变换编码**是将图像**时域**信号转换为**频域**信号进行处理。（DCT）

—**分析—合成编码**是指通过对源数据的分析，将其分解成一系列更适合于表示的“基元”或从中提取若干更为本质意义的参数，编码仅对这些基本单元或特征参数进行。

按照媒体的类型进行压缩划分

图像压缩标准 声音压缩标准 运动图象压缩标准

8.2.2 霍夫曼编码

霍夫曼编码（Huffman）是运用**信息熵原理**的一种无损编码方法，这种编码方法根据源数据各信号发生的概率进行编码。

在**源数据中出现概率大的信号，分配的码字越短；出现概率越小的信号，其码字越长**，从而达到用尽可能少的码表示源数据。

霍夫曼编码的算法

初始化，根据符号概率的大小顺序对符号进行排序。

把概率最小的两个符号组成一个新符号（节点），即新符号的概率等于这两个符号概率之和。

重复第2步，直到形成一个符号为止（树），其概率和等于1。

分配码字。码字分配从最后一步开始**反向**进行，即从最后两个概率开始逐渐向前进行编码，对于每次相加的两个概率，给概率大的赋“0”，概率小的赋“1”（也可以全部相反，如果两个概率相等，则从中任选一个赋“0”，另一个赋“1”）。

霍夫曼编码的特点

霍夫曼编码构造出来的编码值**不是唯一的**。

对不同信号源的编码效率不同

由于编码长度可变，因此**译码时间较长**；没有保护功能，可能会导致**错误传播**（码串中一个码译错导致其他代码出错）；编码长度的不统一，也使得硬件实现有难度。

行程编码

行程编码又称行程长度编码（Run Length Encoding, **RLE**），是一种**熵编码**。这种编码方法广泛地应用于各种图像格式的数据压缩处理中。

行程编码的原理是在给定的图像数据中寻找**连续重复**的数值，然后用两个字符取代这些连续值。即将具有相同值的连续串用**其串长和一个代表值**来代替，该连续串就称为行程，串长称为行程长度。

行程编码分类

定长（表示串长的二进制位数）编码 不定长编码

算术编码

算术编码不是将单个信源符号映射成一个码字，而是把整个信源表示为实数线上的0到1之间的一个区间，其长度等于该序列的概率，再在该区间内选择一个代表性的**小数**，转化为二进制作为实际的编码输出

8.2.4 词典编码

词典编码（dictionary encoding）技术属于无损压缩技术，主要是利用数据本身包含许多重复的字符串的

特性。可以用一些简单的代号代替这些字符串，就可以实现压缩，实际上就是利用了信源符号之间的相关性。字符串与代号的对应表就是词典。

8.2.4.1 LZ77 算法

LZ77 算法在某种意义上又可以称为“滑动窗口压缩”，该算法将一个虚拟的、可以跟随压缩进程滑动的窗口作为词典，要压缩的字符串如果在该窗口中出现，则输出其出现位置和长度。

LZ77 算法具体步骤

(1) 把编码位置设置到输入数据流的开始位置。(2) 找窗口中最长的匹配串(3) 以“(Pointer, Length) Characters”的格式输出，其中 Pointer 是指向窗口中匹配串的指针，Length 表示匹配字符的长度，Characters 是前向缓冲存储器中的不匹配的第 1 个符。(4) 如果前向缓冲存储器不是空的，则把编码位置和窗口向前移(Length+1)个字符，然后返回到步骤。

PCM 编码包括如下三个过程：(有损)

一采样，将模拟信号转换为时间离散的样本脉冲序列。量化，将离散时间连续幅度的抽样信号转换为离散时间离散幅度的数字信号。编码，用一定位数的脉冲码组表示量化采样值。

PCM 编码的优点

有很强的抗干扰性

能方便的利用计算机编程，实现各种智能化设计。

8.2.6 增量调制(DM)

增量调制也称 Δ 调制(delta modulation, DM)，它是一种预测编码技术，是 PCM 编码的一种变形。

DM 是对实际的采样信号与预测的采样信号之差的极性进行编码，将极性变成“0”和“1”这两种可能的取值之一。如果实际的采样信号与预测的采样信号之差的极性为“正”，则用“1”表示；相反则用“0”表示，或者相反。(一位系统)输入信号变化快：斜率过载；输入信号变化慢：粒状噪声

8.2.7 差分脉冲编码调制

差分脉冲编码调制(Differential Pulse Code Modulation, DPCM)是利用样本与样本之间存在的信息冗余度来进行编码的一种数据压缩技术。

差值脉冲编码调制是利用信号的相关性找出可以反映信号变化特征的一个差值量进行编码。

DPCM 的基本工作原理

根据过去的样本去估算(estimate)下一个样本信号的幅度大小，这个值称为预测值，然后对实际信号值与预测值之差进行量化编码，从而就减少了表示每个样本信号的位数。

它与脉冲编码调制(PCM)不同处在于，PCM 是直接对采样信号进行量化编码，而 DPCM 是对实际信号值与预测值之差进行量化编码。

声音压缩标准——MP3

MP3 是 MPEG-1 的标准草案中音频编码的 Layer 3。MP3 最大特点是能以较小的比特率、较大压缩比达到近乎完美的 CD 音质，制作简单，交流方便。MP3 压缩编码是一个国际性全开放的编码方案，其编码算法流程大致分为时频映射、心理声学模型、量化编码三大功能模块，这三个功能模块是实现 MP3 编码的关键。

MP4

MP4 是 MPEG-2 AAC(ISO/IEC 13818-7)技术(Advanced Audio Coding)。

MP4 的特点是音质更加完美而压缩比更大。它增加了诸如对立体声的完美再现、比特流效果音扫描、多媒体控制、降噪等 MP3 没有的特性，使得在音频压缩后仍能完美的再现 CD 的音质。

8.3.2 图像压缩标准

JPEG 包含两种基本的压缩算法：无损压缩算法(基于差分脉冲调制)和有损压缩算法(基于离散余弦变

换)。

JPEG 算法压缩编码步骤

(1) 使用正向离散余弦变换 (Forward Discrete Cosine Transform, FDCT) 把信息从空间域变换成频率域的数据, 并利用数据的频率特性进行处理; (2) 使用加权函数对 DCT 系数进行量化 (信息损失), 这个加权函数对于人的视觉系统是最佳的; (3) 使用霍夫曼可变字长熵编码器对量化系数进行编码。

JPEG 压缩编码算法的主要计算步骤

(1) 正向离散余弦变换 (FDCT)。 (2) 量化 (quantization)。 (3) Z 字形编码 (zigzag scan)。 (4) 使用差分脉冲编码调制 (differential pulse code modulation, DPCM) 对直流系数 (DC) 进行编码。 (5) 使用行程长度编码 (Run Length Encoding, RLE) 对交流系数 (AC) 进行编码。 (6) 熵编码 (entropy coding)。

JPEG2000

JPEG2000 弥补了传统 JPEG 压缩技术缺陷, 有以下优点

高压缩率 支持无损压缩和有损压缩 渐进传输 对感兴趣区域压缩

8.3.3.1 MPEG 标准

运动图像专家组 MPEG 是由国际标准化组织 ISO 和国际电工委员会 IEC 联合成立的, 负责开发电视图像数据和声音数据的编码、解码和它们的同步标准。这个专家组开发的标准称为 MPEG 标准。 (视频、音频、系统)

MPEG-1 和 MPEG-2 是 MPEG 组织制定的第一代视音频压缩标准, 为 VCD、DVD 及数字电视和高清晰度电视等产业的飞速发展打下了牢固的基础

MPEG-4 是基于第二代视音频编码技术制定的压缩标准, 以视听媒体对象为基本单元, 实现数字视音频和图形合成应用、交互式多媒体的集成, 目前已经在流式媒体服务等领域开始得到应用

MPEG-1 标准

- 1、采用 JPEG 压缩算法去掉空间冗余
- 2、采用运动补偿算法去掉时间冗余
- 引入了心理声学模型 (根据人耳的听觉特性分配编码比特位数)

MPEG-2 标准

主要特点是:

(1) MPEG-2 解码器兼容 MPEG-1 和 MPEG-2 标准。4) 允许在画面质量、存储容量和带宽之间选择, 在一定范围内改变压缩比。

MPEG-4 标准

MPEG-4 标准于 1998 年 11 月公布, 是各种音频/视频对象的编码, 包括了系统、电视图像、音频、一致性测试和参考软件、传输多媒体集成框架等, 采用基于对象的编码、基于模型的编码等第二代编码技术, 提高编码效率

MPEG-4 的目标是支持多种多媒体的应用, 特别是多媒体信息基于内容的检索和访问, 可以根据不同的应用需求现场配置解码器。其编码系统也是开放的, 可以随时加入新的有效的算法模块。

MPEG-7 标准

MPEG-7 标准于 2001 年公布, 称为多媒体内容描述接口, 包括系统、描述定义语言、电视图像、音频、多媒体描述框架、参考软件以及一致性测试七个部分。

MPEG-7 标准的目的是产生一个描述多媒体内容的标准, 支持对多媒体信息在不同程度层面上的解释和理解, 从而使其可以根据用户的需要进行传递和存取。

MPEG-7 注重的是提供视听信息内容的描述方案, 并不包括针对不同应用的特征提取方法和搜索引擎。

MPEG-21 标准

•MPEG-21 基于两个基本概念: 分布和处理基本单元 (Digital Item, DI) 以及 DI 与用户间的互操作。MPEG-21 可表述为以一种高效、透明和可以互操作的方式支持用户交换、接入、使用甚至操作 DI 的技术。

H. 26X 系列视频标准

H. 26X 系列视频标准是国际电信联盟 ITU 的视频编码专家组（ITU-T）制定的系列图像压缩标准，主要有 H. 261、H. 263、H264 等。这些视频标准主要应用于实时视频通信领域，如会议电视、可视电话等。

8.3.3.3 AVS 标准