**编解码**

　　编解码器（codec）指的是一个能够对一个信号或者一个数据流进行变换的设备或者 程序。这里指的变换既包括将信号或者数据流进行编码（通常是为了传输、存储或者加密）或者提取得到一个编码流的操作，也包括为了观察或者处理从这个编码流 中恢复适合观察或操作的形式的操作。编解码器经常用在视频会议和流媒体等应用中。

**容器**

　　很多多媒体数据流需要同时包含音频数据和视频数据，这时通常会加入一些用于音频和视 频数据同步的元数据，例如字幕。这三种数据流可能会被不同的程序，进程或者硬件处理，但是当它们传输或者存储的时候，这三种数据通常是被封装在一起的。通 常这种封装是通过视频文件格式来实现的，例如常见的\*.mpg, \*.avi, \*.mov, \*.mp4, \*.rm, \*.ogg or \*.tta. 这些格式中有些只能使用某些编解码器，而更多可以以容器的方式使用各种编解码器。

　　FourCC全称Four-Character Codes，是由4个字符（4 bytes）组成，是一种独立标示视频数据流格式的四字节，在wav、avi档案之中会有一段FourCC来描述这个AVI档案，是利用何种codec来 编码的。因此wav、avi大量存在等于“IDP3”的FourCC。

　　视频是现在电脑中多媒体系统中的重要一环。为了适应储存视频的需要，人们设定了不同的视频文件格式来把视频和音频放在一个文件中，以方便同时回放。视频档实际上都是一个容器里面 包裹着不同的轨道，使用的容器的格式关系到视频档的可扩展性。

**参数介绍**

**采样率**

　　采样率（也称为采样速度或者采样频率）定义了每秒从连续信号中提取并组成离散信号的采样个数，它用赫兹（Hz）来表示。采样频率的倒数叫作采样周期或采样时间，它是采样之间的时间间隔。注意不要将采样率与比特率（bit rate，亦称“位速率”）相混淆。

　　采样定理表明采样频率必须大于被采样信号带宽的两倍，另外一种等同的说法是奈奎斯特 频率必须大于被采样信号的带宽。如果信号的带宽是 100Hz，那么为了避免混叠现象采样频率必须大于 200Hz。换句话说就是采样频率必须至少是信号中最大频率分量频率的两倍，否则就不能从信号采样中恢复原始信号。

　　对于语音采样：

* 8,000 Hz - 电话所用采样率, 对于人的说话已经足够
* 11,025 Hz
* 22,050 Hz - 无线电广播所用采样率
* 32,000 Hz - miniDV 数码视频 camcorder、DAT (LP mode)所用采样率
* 44,100 Hz - 音频 CD, 也常用于 MPEG-1 音频（VCD, SVCD, MP3）所用采样率
* 47,250 Hz - Nippon Columbia (Denon)开发的世界上第一个商用 PCM 录音机所用采样率
* 48,000 Hz - miniDV、数字电视、DVD、DAT、电影和专业音频所用的数字声音所用采样率
* 50,000 Hz - 二十世纪七十年代后期出现的 3M 和 Soundstream 开发的第一款商用数字录音机所用采样率
* 50,400 Hz - 三菱 X-80 数字录音机所用所用采样率
* 96,000 或者 192,000 Hz - DVD-Audio、一些 LPCM DVD 音轨、Blu-ray Disc（蓝光盘）音轨、和 HD-DVD （高清晰度 DVD）音轨所用所用采样率
* 2.8224 MHz - SACD、 索尼 和 飞利浦 联合开发的称为 Direct Stream Digital 的 1 位 sigma-delta modulation 过程所用采样率。

　　在模拟视频中，采样率定义为帧频和场频，而不是概念上的像素时钟。图像采样频率是传感器积分周期的循环速度。由于积分周期远远小于重复所需时间，采样频率可能与采样时间的倒数不同。

* 50 Hz - PAL  视频
* 60 / 1.001 Hz - NTSC 视频

　　当模拟视频转换为数字视频的时候，出现另外一种不同的采样过程，这次是使用像素频率。一些常见的像素采样率有：

* 13.5 MHz - CCIR 601、D1 video

**分辨率**

　　分辨率，泛指量测或显示系统对细节的分辨能力。此概念可以用时间、空间等领域的量 测。日常用语中之分辨率多用于图像的清晰度。分辨率越高代表图像品质越好，越能表现出更多的细节。但相对的，因为纪录的信息越多，文件也就会越大。目前个 人电脑里的图像，可以使用图像处理软件，调整图像的大小、编修照片等。例如 photoshop，或是photoimpact等软件。

**图像分辨率**：

　　用以描述图像细节分辨能力，同样适用于数字图像、胶卷图像、及其他类型图像。常用' 线每毫米'、 '线每英吋'等来衡量。通常，“分辨率”被表示成每一个方向上的像素数量，比如640x480等。而在某些情况下，它也可以同时表示成“每英吋像素” （pixels per inch，ppi）以及图形的长度和宽度。比如72ppi，和8x6英吋。

**视频分辨率**：

各种电视规格分辨率比较视频的画面大小称为“分辨率”。数位视频以像素为度量单位，而类 比视频以水平扫瞄线数量为度量单位。标清电视频号分辨率为 720/704/640x480i60（NTSC）或768/720x576i50（PAL/SECAM）。新的高清电视（HDTV）分辨率可达 1920x1080p60，即每条水平扫瞄线有1920个像素，每个画面有1080条扫瞄线，以每秒钟60张画面的速度播放。

**画面更新率fps**

　　Frame rate中文常译为“画面更新率”或“帧率”，是指视频格式每秒钟播放的静态画面数量。典型的画面更新率由早期的每秒6或8张（frame per second，简称fps），至现今的每秒120张不等。PAL (欧洲，亚洲，澳洲等地的电视广播格式) 与 SECAM (法国，俄国，部分非洲等地的电视广播格式) 规定其更新率为25fps，而NTSC (美国，加拿大，日本等地的电视广播格式) 则规定其更新率为29.97 fps。电影胶卷则是以稍慢的24fps在拍摄，这使得各国电视广播在播映电影时需要一些复杂的转换手续（参考Telecine转换）。要达成最基本的视 觉暂留效果大约需要10fps的速度。

**压缩方法**:有损压缩和无损压缩

　　在视频压缩中有损（Lossy ）和无损（Lossless）的概念与静态图像中基本类似。无损压缩也即压缩前和解压缩后的数据完全一致。多数的无损压缩都采用RLE行程编码算法。有损 压缩意味着解压缩后的数据与压缩前的数据不一致。在压缩的过程中要丢失一些人眼和人耳所不敏感的图像或音频信息，而且丢失的信息不可恢复。几乎所有高压缩 的算法都采用有损压缩，这样才能达到低数据率的目标。丢失的数据率与压缩比有关，压缩比越小，丢失的数据越多，解压缩后的效果一般越差。此外，某些有损压 缩算法采用多次重复压缩的方式，这样还会引起额外的数据丢失。

* 无损格式，例如WAV，PCM，TTA，FLAC，AU，APE，TAK，WavPack(WV)
* 有损格式，例如MP3，Windows Media Audio（WMA），Ogg Vorbis（OGG），AAC

**帧内压缩和帧间压缩**

　　帧内（Intraframe）压缩也称为空间压缩（Spatial compression）。当压缩一帧图像时，仅考虑本帧的数据而不考虑相邻帧之间的冗余信息，这实际上与静态图像压缩类似。帧内一般采用有损压缩算法， 由于帧内压缩时各个帧之间没有相互关系，所以压缩后的视频数据仍可以以帧为单位进行编辑。帧内压缩一般达不到很高的压缩。

　　采用帧间（Interframe）压缩是基于许多视频或动画的连续前后两帧具有很大 的相关性，或者说前后两帧信息变化很小的特点。也即连续的视频其相邻帧之间具有冗余信息，根据这一特性，压缩相邻帧之间的冗余量就可以进一步提高压缩量， 减小压缩比。帧间压缩也称为时间压缩（Temporalcompression），它通过比较时间轴上不同帧之间的数据进行压缩。帧间压缩一般是无损的。 帧差值（Frame differencing）算法是一种典型的时间压缩法，它通过比较本帧与相邻帧之间的差异，仅记录本帧与其相邻帧的差值，这样可以大大减少数据量。

**对称编码和不对称编码**

　　对称性（symmetric）是压缩编码的一个关键特征。对称意味着压缩和解压缩占 用相同的计算处理能力和时间，对称算法适合于实时压缩和传送视频，如视频会议应用就以采用对称的压缩编码算法为好。而在电子出版和其它多媒体应用中，一般 是把视频预先压缩处理好，尔后再播放，因此可以采用不对称（asymmetric）编码。不对称或非对称意味着压缩时需要花费大量的处理能力和时间，而解 压缩时则能较好地实时回放，也即以不同的速度进行压缩和解压缩。一般地说，压缩一段视频的时间比回放（解压缩）该视频的时间要多得多。例如，压缩一段三分 钟的视频片断可能需要10多分钟的时间，而该片断实时回放时间只有三分钟。

下面是目前常见的几种视频编码技术方案；

**1）MPEG-2**

MPEG-2图像压缩的原理是主要利用了图像中的两种特性：空间相关性和时间相关性。另外它综合采用了运动补偿的帧间预测、空间域离散余弦变换、自适应量化和可变长编码的混合编码。同时它开始用半像素精度运动矢量搜索，引入了“帧”和“场”的编码方法，和MPEG-1向下兼容；主要针对存储媒体、数字电视、高清晰等应用领域，分辨率为：低（352\*288）、中（720\*480）、次高（1440\*1080）、高（1920\*1080）。

**2）MPEG-4**

采用基于对象的编码理念，即在编码时将一幅景物分成若干在时间和空间上相互联系的视频音频对象，分别编码后，再经过复用传输到接收端，然后再对不同的对象分别解码，从而组合成所需要的视频和音频。实现了从基于像素的传统编码向基于对象和内容的现代编码的转变。

主要特点是支持交互性、高压缩比及能用存储性；

**3）H.264/AVC**

主要包括2个层次：视频编码层和网络抽象层。前者主要致力于有效的表示视频内容，后者格式化VCL视频表示，提供头部信息，适合多种传输和存储媒体。采用了全新的编码技术：帧内预测，可变大小的图像分块，多预测参考帧，1/4和1/8像素精度的运动估计，残差图像的整数变换编码等。由于其压缩效率高、容错能力好、网络适应能力强等特点，能适用于不同网络中的视频传输。

**4）Windows Media Video**

采用了MPEG4的压缩算法，所以压缩率和图像的质量都很不错，由于Windows Media video是微软的杰作，所以Windows系统对Windows MediaVideo给予了很好的支持。

**5）RealVideo**

采用REAL公司可变编码率的REAL格式的视频编码技术。它能在保持文件较小的条件下获得较高编码率的视频质量。主要定位在视频流应用方面的，不太适合专业场合，在网络和娱乐场合占有不错的份额。

**6）Sorenson Video**

Sorenson Media公司随 QuickTime5发布的编码器，采用的是sorenson公司的压缩算法，质量很不错，已经成为QuickTime 的标准视频编码，网络上大部分电影预告片都采用这种编码。

**7）Intel Indeo video**

由Intel架构实验室开发，常见的视频编码，主流的有4.5和5.10两种，质量比Cinepak和R3.2要好，可以适应不同带宽的网络，但必须有相应的解码插件才能顺利地将下载作品进行播放。

**8）Dvix和xviD**

Dvix主要是将影像部分以MPEG4来压缩，音效部分则以MP3来压缩；采用了空间和时间压缩技术。在档案大小上占了相当大的优势，画质部分的表现也很不错。XviD是目前世界上最常用的视频编码解码器，是第一个真正开放源码的，它是基于Open DivX编写的。Xvid支持多种编码模式，量化方式和范围控制，运动侦测和曲线平衡分配等众多编码技术。

目前常见的几种音频编码技术方案；

**1）MPEG Audio Layer 3**

MP3采用了名为“感官编码技术”的编码算法：编码时先对音频文件进行频谱分析，然后用过滤器滤掉噪音电平，接着通过量化的方式将剩下的每一位打散排列，最后形成具有较高压缩比的MP3文件 。能够在音质丢失很小的情况下把文件压缩到更小。

**2）FLAC**

采用了无损压缩技术，使用了同步代码和CRC校验码，不会破坏任何原有的音频资讯。被编码的音频(PCM)数据没有任何信息损失，解码输出的音频与编码器的输入的每一个字节都是一样的。

**3）AAC**

采用了临时噪声重整、后向自适应线性预测、联合立体声技术和量化哈夫曼编码等最新技术，通过结合其他的功能来提高编码效率,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。

**4）Ogg Vorbis**

主要算法是利用MDCT（改进离散余弦变换）技术，MDCT先将整体图像分成N\*N像素块，然后对N\*N像素块逐一进行DCT变换。是一种灵活开放的音频编码，能够在编码方案已经固定下来后还能对音质进行明显的调节和新算法的改良。可以不断导入新技术逐步完善。

**5）Windows Media Audio**

采用的压缩技术与MP3压缩原理近似，但它并不删减大量的编码 ，音质要强于MP3格式，更远胜于RA格式，是以减少数据流量但保持音质的方法来达到比MP3压缩率更高的目的，WMA的压缩率一般都可以达到1：18左右。

**6）PCM**

通过抽样,对样值幅度量化,编码三个步骤将模拟信号转换为数字编码。可以提供话音、图象传送、远程教学等业务，特别适用于对数据传输速率要求较高，需要更高带宽的用户使用。

**7）RA（RealAudio）**

采用的SureStream（自适应流）技术是RealNetworks公司具有代表性的技术。它通过RealServer将A/V文件以流的方式传输，然后利用SureStream方式，根据客户端不同的拨号速率（不同的带宽），让传输的A/V信息自动适应带宽，并始终以流畅的方式播放。