



多媒体技术

数字视频处理技术 I

内容提纲

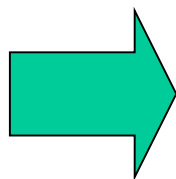
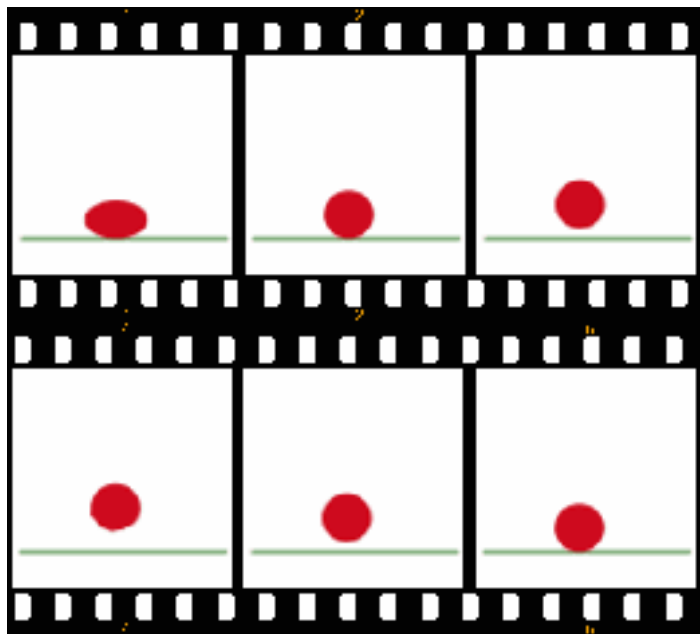
- 动画和视频基础
- 电影与电视
- 电视图像的数字化
- 数字视频的获取
- 数字视频编辑技术
- 视频显示设备

内容提纲

- 动画和视频基础
- 电影与电视
- 电视图像的数字化
- 数字视频的获取
- 数字视频编辑技术
- 视频显示设备

动画的原理

- 当人们看到一件物体时，即使它马上消失了，它在人的视觉中还会停留大约十分之一秒的时间。当投影机以每秒二十四格的速度投射在银幕上，或录像机以每秒三十格的扫描方式在电视荧光屏上呈现影像时，它会把每格不同的画面连接起来，从而在我们脑中产生物体在“运动”的印象。



视频是什么？

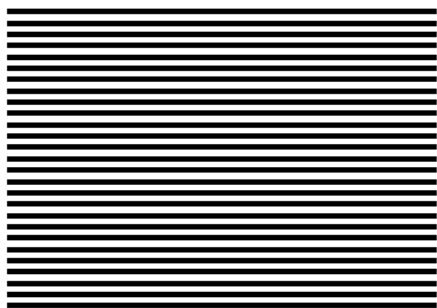
- 视频是图像的集合
- 视频是时间上连续的一系列图像的集合
- 数字视频是时间上连续的一系列数字图像的集合
- 动画和视频信息是连续渐变的静态图像或图形序列，沿时间轴顺次更换显示，从而构成运动视觉的媒体。当序列中每帧图像是由**人工或计算机产生的**图像时，我们常称为**动画**；当序列中每帧图像是通过**实时摄取**自然景象或活动对象时，我们常称为影像视频，或简称为**视频**。

视频的分辨率

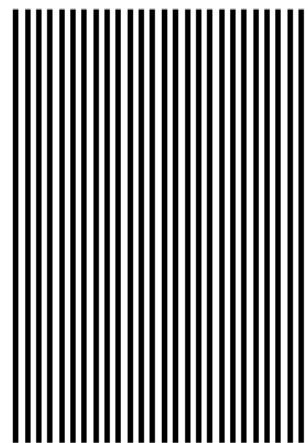
- 视频是由连续快速出现的静止图像组成。在计算机领域，经常使用图像像素的多少来表示视频的分辨率，即水平像素数 \times 垂直像素数。
- VGA：640 \times 480；SVGA：800 \times 600

视频的清晰度

- **线（电视行、TVLine）**：电视的画面清晰度是以水平清晰度作为单位。通俗地说，我们可以把电视上的画面以水平方向分割成很多很多扫描线，分得越细，这些画面就越清楚，而水平线数的扫描线数量也就越多。清晰度的单位是“电视行（TVLine）”也称线。



水平扫描线



垂直扫描线

视频的清晰度

- 电视的画面清晰度是以**水平清晰度（垂直扫描线的数目）**作为单位，电视屏幕由电子枪射出的电子，经由磁场偏向后打在屏幕上而发光，因此每一个图框都由电子枪的扫描线画出来。
- 一个图框的扫描线数是525条或625条。美国国家电视系统委员会（National Television System Committee, **NTSC**）制订的电视画面播映标准，明定每幅画面必需具备525条扫描线，及每秒钟必需播映30幅画面，而欧洲国家（Phase Alteration Line, **PAL**）则有每个图框625条扫描线，且每秒25个图框的电视系统。

常见动画文件格式

- **GIF文件** (*.gif) 最初GIF只是用来存储单幅静止图像，后来发展成为可同时存储若干幅静止图像并进而形成连续的动画。
- **Flic文件** (*.FLI/*.FLC) 是3D Studio中采用的彩色动画文件格式，Flic文件采用行程编码(RLE)算法和Delta算法进行无损的数据压缩。
- **FLASH文件** (*.fla) 是用Macromedia公司的Flash编辑而成的动画文件。在因特网上已经占据了不可替代的霸主地位。

常见视频文件格式

- **AVI文件**（*.avi）即音频视频交错(Audio Video Interleaved)，Windows、OS/2直接支持。
- **QuickTime文件** (*.mov/*.qt) 是Apple计算机公司开发的一种音频、视频文件格式，Apple Mac OS、Windows等主流电脑平台支持。
- **MPEG文件** (.mpeg/.mpg/.dat) 格式是运动图像压缩算法的国际标准，用有损压缩方法减少运动图像中的冗余信息，已被几乎所有的计算机平台共同支持。
- **RealVideo文件** (.rm/.rmvb) 是RealNetworks公司开发的流式视频文件格式，主要用来在低速率的广域网上实时传输活动视频影像。

常见视频编辑软件

- Adobe公司的Premiere
- ULEAD公司的Media Studio Pro
- 会声会影Ulead Video Studio （家庭使用）
- Pinnacle公司的Studio
- Sony公司的Vegas ...



视频编辑的内容

- **格式转换**
- **针对每一帧图像的编辑**
- **删除部分帧图像**
- **将几段视频连接在一起**
- **将一段视频分割为几段**
- **添加艺术化字幕**

内容提纲

- 动画和视频基础
- 电影与电视
- 电视图像的数字化
- 数字视频的获取
- 数字视频编辑技术
- 视频显示设备

电影与电视

•电影原理及历史

以一种连续贴图的方式快速播放，再加上人类“视觉暂留”的因素，产生动画。



视觉暂留

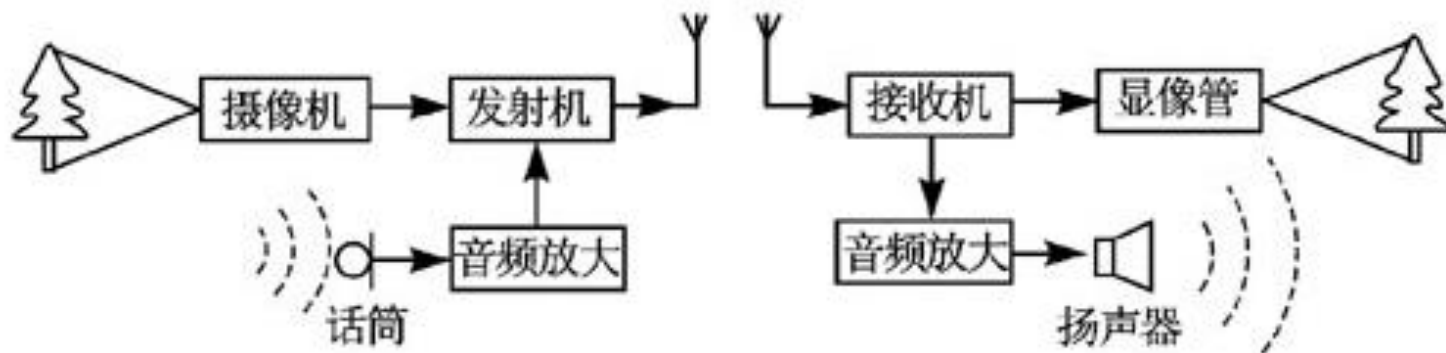
人眼在观察景物时，光信号传人大脑神经，需经过一段短暂的时间，光的作用结束后，视觉形象并不立即消失，这种残留的视觉称“后像”，视觉的这一现象则被称为“视觉暂留”，是“眼睛”和“大脑”联合起来欺骗自己产生的幻觉。

电影与电视

- 电视工作原理

发送端：摄像机（摄像管）把景物转变成电信号（光-电转换）

接收端：由显示设备（显像管）重现原景物（电-光转换）



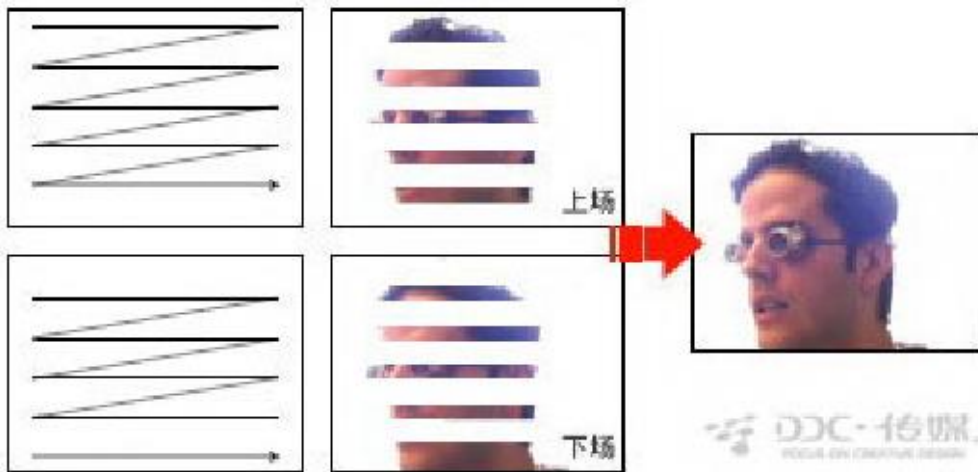
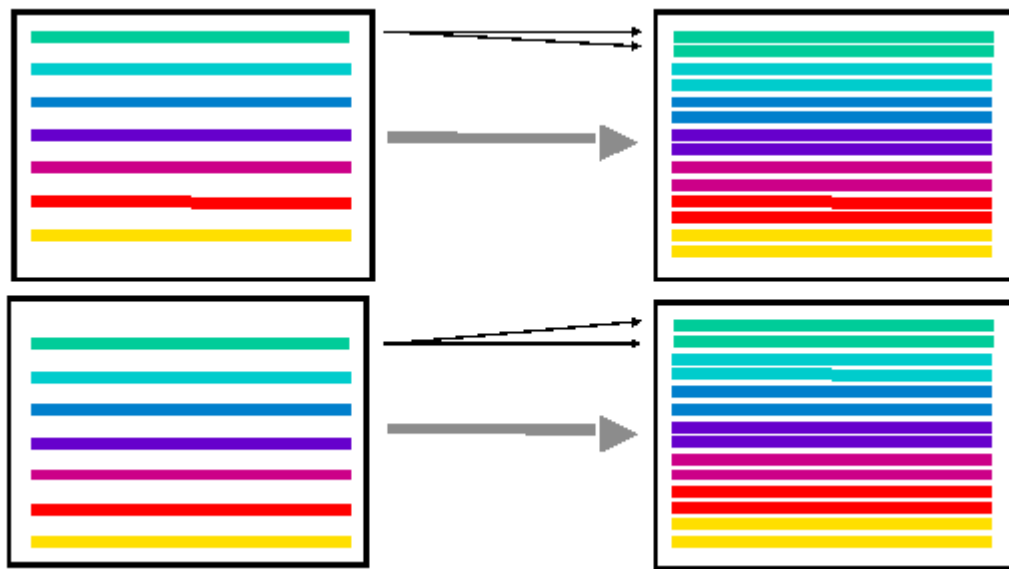
电影与电视

- 电视工作原理

扫描的机制

隔行扫描

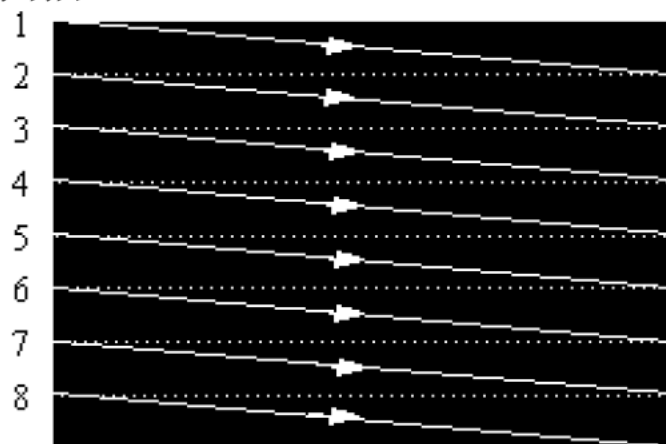
逐行扫描



电影与电视

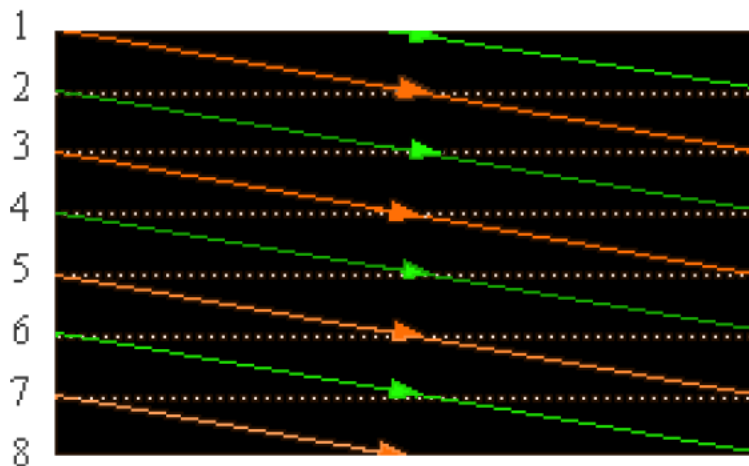
• 电视工作原理

每一帧图像由电子束顺序地一行接着一行连续扫描而成，这种扫描方式称为逐行扫描。把每一帧图像通过两场扫描完成则是隔行扫描，两场扫描中，第一场(奇数场)只扫描奇数行，依次扫描**1、3、5...**行，而第二场(偶数场)只扫描偶数行，依次扫描**2、4、6...**行。隔行扫描技术在传送信号带宽不够的情况下起了很大作用。在电视的标准显示模式中，**i**表示隔行扫描，**p**表示逐行扫描。



逐行扫描

(non-interlaced scanning)



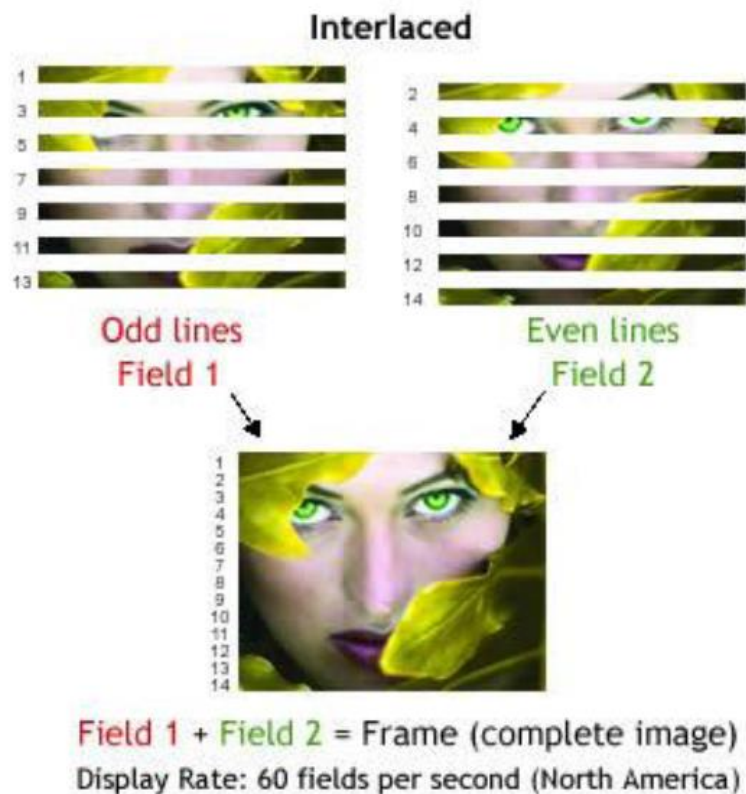
隔行扫描

(interlaced scanning)

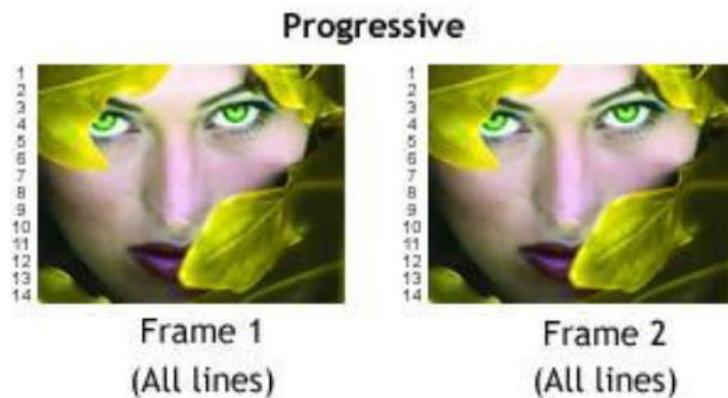
行频、场频、帧频

电影与电视

• 电视工作原理



逐行扫描和隔行扫描的显示效果主要区别在稳定性上面，隔行扫描的行间闪烁比较明显，逐行扫描克服了隔行扫描的缺点，画面平滑自然无闪烁。



Display Rate: 60 frames per second (North America)

电影与电视

- **电视工作原理**

- **I=Interlace**, 隔行扫描: 就是一个画面分成两次送出, 先送奇数线的画面后, 再送偶数线, 显示时再将它们合成完整画面.
- **P=Progressive**, 逐行扫描: 就是每次都送完整的画面, 不需要奇偶画面重新组合.
- 隔行扫描显示的屏幕上观看时, 交错画面要把它重组回来, 这个动作叫做“去交错”.

电影与电视

- 电视工作原理

帧速率和场

先有场后有帧

PAL制式：分别隔一场抽掉奇数行和偶数行



被抽掉奇数场的偶数场



被抽掉偶数场的奇数场

电影与电视

- 电视扫描术语

- 场频/场帧率（field rate）：符号为 f_f ，每秒钟扫描的场数。场频是根据人的视觉特性和电网频率（50Hz或60Hz）确定的，目的是使屏幕上显示的图像看起来不会让人感到闪烁。
- 帧频/帧率（frame rate）：符号为 f_F ，每秒钟扫描的帧数，并用fps做为单位。
- 行频/水平行速率（horizontal line rate）：符号为 f_H ，每秒钟扫描的行数。例如，NTSC制精确的帧率为29.97Hz，525行每帧，因此行频为 $29.97 \times 525 = 15734$ 行/秒

电影与电视

• 电视扫描格式

- D1为480i格式，和NTSC模拟电视清晰度相同，525条垂直扫描线，483条可见垂直扫描线，4:3 或16:9，隔行/60Hz，行频为15.75KHz。
- D2为480P格式，和逐行扫描DVD规格相同，525条垂直扫描线，480条可见垂直扫描线，4:3 或 16:9，分辨率为 640×480 ，逐行/60Hz，行频为31.5KHz。
- D3为1080i格式，是标准数字电视显示模式，1125条垂直扫描线，1080条可见垂直扫描线，16:9，分辨率为 1920×1080 ，隔行/60Hz，行频为33.75KHz。

电影与电视

- **电视扫描格式**

- D4为 720p格式，是标准数字电视显示模式，750条垂直扫描线，720条可见垂直扫描线，16: 9，分辨率为 1280×720 ，逐行/60Hz，行频为45KHz。
- D5为1080p格式，是标准数字电视显示模式，1125条垂直扫描线，1080条可见垂直扫描线，16: 9，分辨率为 1920×1080 逐行扫描，专业格式。

电影与电视

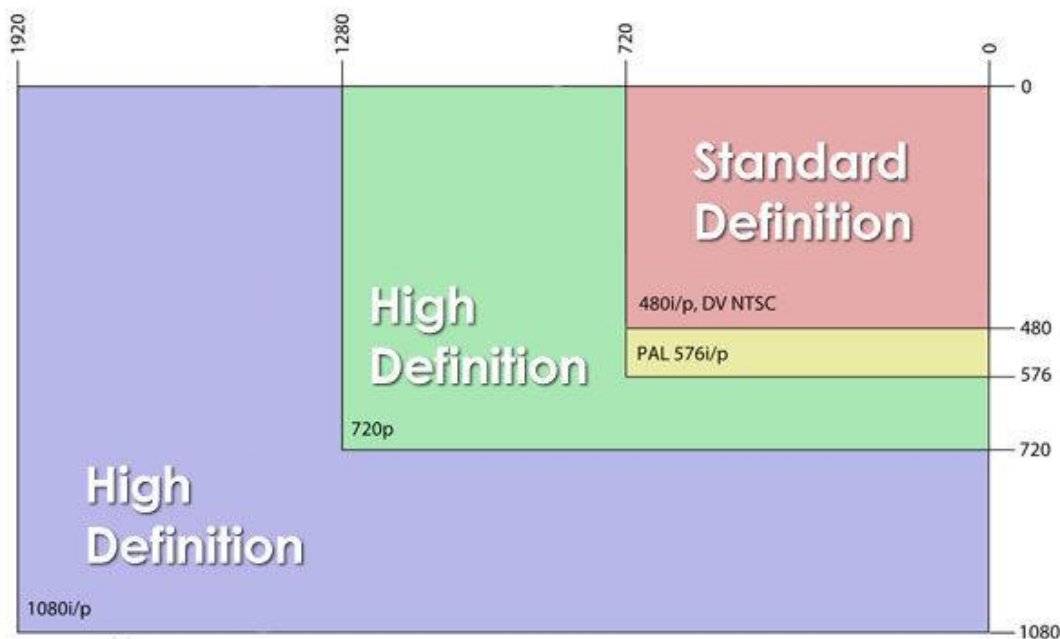
1、电视制式

美/加/日/韩/台——29.97帧/秒——NTSC(525线-480可视)——720×480

澳/中/欧/南美——25帧/秒——PAL(625线-576可视)——720×576

法国/部分非洲——25帧/秒——SECAM(625线-576可视)——720×576

2、标清与高清



电影与电视

- **PAL制式**

- 625行(扫描线)/帧， 25帧/秒(40 ms/帧)
- 每场中25行为场回扫， 575行是有效行
- 宽高比4:3 （垂直视线角 15° 、水平视线角 20° ）
- **隔行扫描**， 2场/帧， 312.5行/场。奇数扫描行构成的场叫**奇数场**，偶数扫描行构成的场叫**偶数场**，奇数和偶数场交错组成一个**电视帧**
- 颜色模型：YUV

电影与电视

- **NSTC制式**

- 525行(扫描线)/帧, 30帧/秒
- 宽高比4:3
- **隔行扫描**, 一帧分成2场(field), 262.5线/场
- 在每场的开始部分保留20扫描线作为**控制信息**, 故只有485条线的可视数据
- 每行63.5微秒, 水平回扫时间10微秒(包含5微秒的水平同步脉冲), 所以显示时间是53.5微秒
- 颜色模型: YIQ

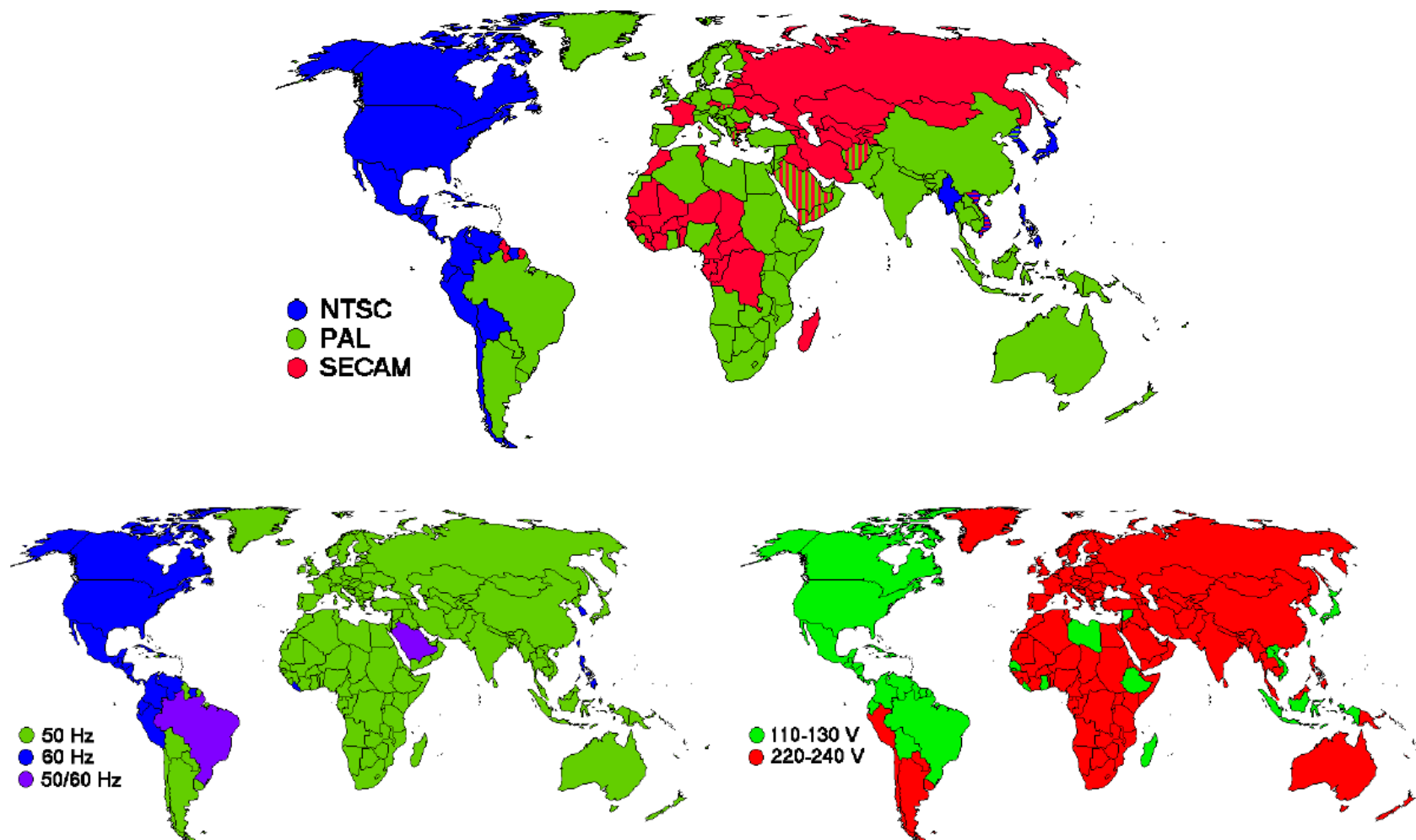
彩色全电视信号

- 黑白电视中只需传送一个亮度信号，而在彩色电视机中，则需要在满足与黑白电视兼容、而且不增加为黑白电视所规定的信道带宽（如 6 MHz）的条件下，同时传送亮度信号和2个色差信号
- 如何找到一个可取的方案去实现这一要求，是彩色电视制式所解决的问题。不同的色差信号传送方案就形成了不同的彩色电视制式

彩色全电视信号

- 当前世界上主要的彩色电视制式有 3 种：
 - NTSC：正交平衡调幅制
 - 美国、加拿大、日本、韩国、菲律宾等地区
 - PAL：逐行倒相正交平衡调幅制
 - 德国、英国、中国、朝鲜等
 - SECAM：顺序传送彩色与存储制
 - 法国、前苏联及东欧国家

彩色全电视信号



Source: <http://www.stjarnhimlen.se/tv/tv.html>

彩色全电视信号

- U和V（用副载波 w_{sc} ）经平衡正交调制，加到亮度Y上，形成彩色全电视信号
- 对于NTSC制，U信号调制在副载波的零相位上，而V信号调制在90°相位上，如下：

$$Cvbs = Y + U \sin w_{sc} t + V \cos w_{sc} t$$

色度信号

彩色全电视信号

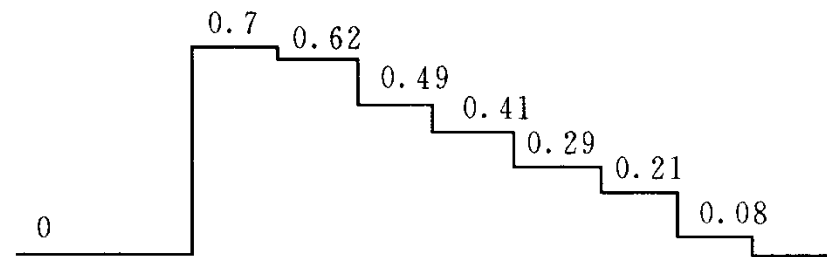
- NTSC制式系统对信道微分相位敏感，导致图像色调失真。
- PAL制式是为克服这一缺点而提出的，具体做法是将V分量调制相位逐行倒相180°。PAL制式色度信号表达式如下：

$$Ch = U \sin w_{sc}t + S(t)V \cos w_{sc}t$$

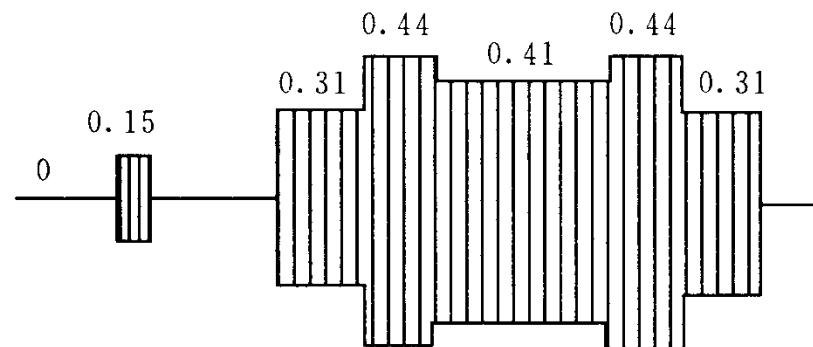
- S(t)称为PAL开关函数，它是双极性矩形脉冲，其重复周期为行周期的两倍，幅度为+1和-1

- 彩色全电视信号各部分（以一行彩条波形为例）
- 将亮度信号、色度信号和相关的同步信号进行混合，构成彩色全电视信号，或称复合电视信号

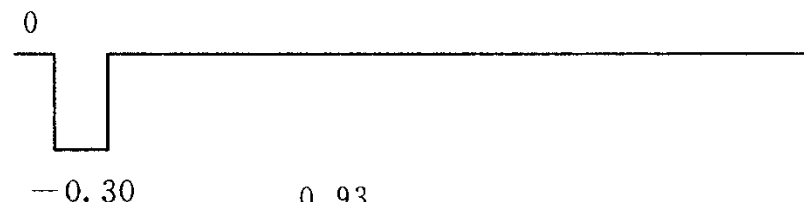
亮度信号 Y



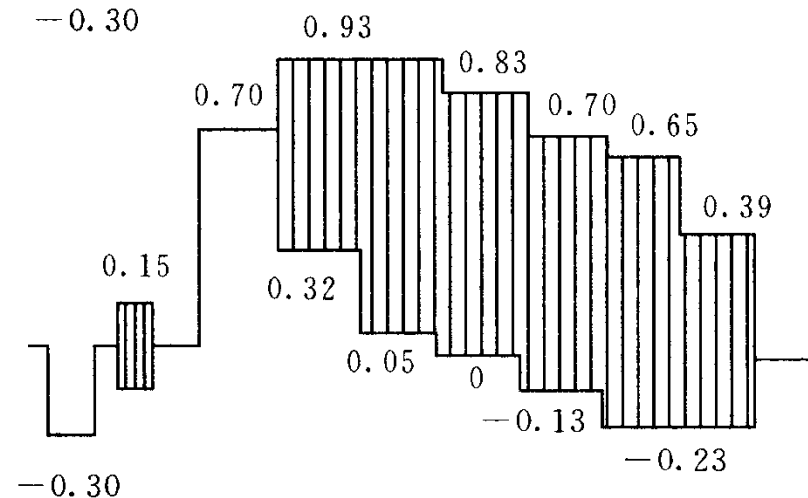
色度信号 C



复合同步信号 S



彩色全电视信号



视频信号类型

- 复合电视信号 (CVBS)

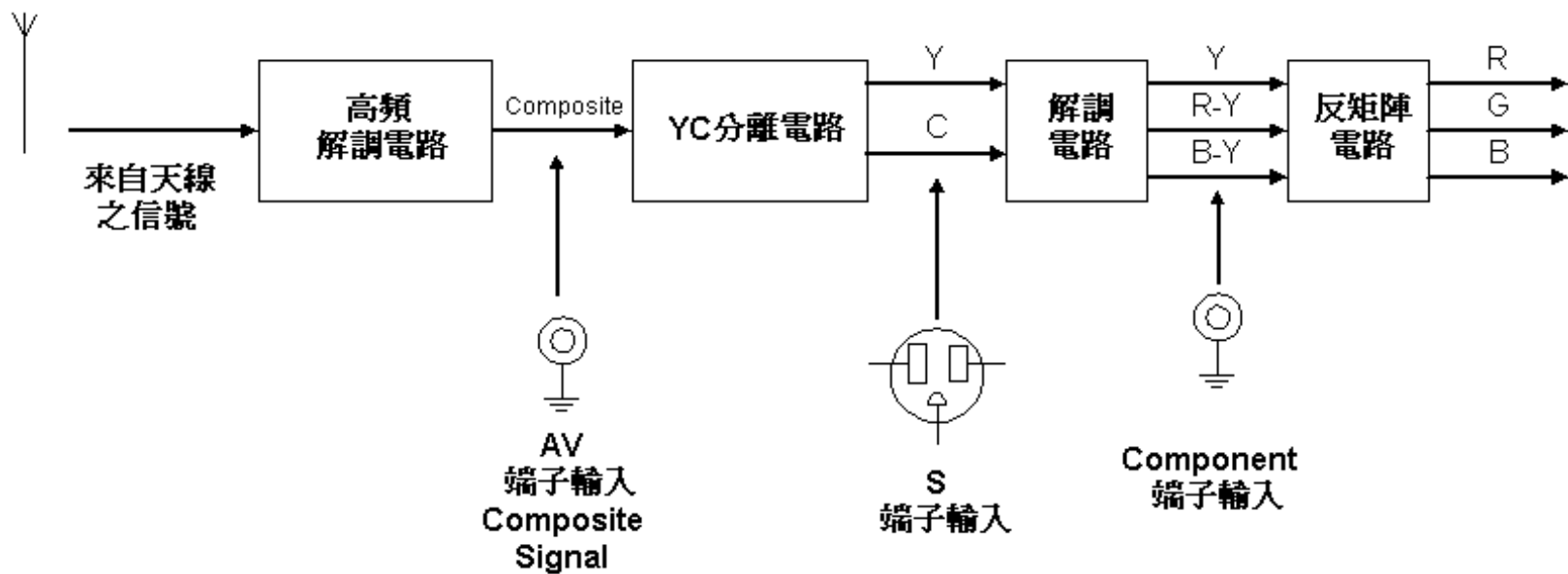
- 包含亮度信号、色差信号和所有定时同步信号的单一电视信号，或称**全电视信号**

- 分离电视信号 (S-Video)

- 是亮度和色差分离的一种电视信号(Y/C)

- 分量电视信号

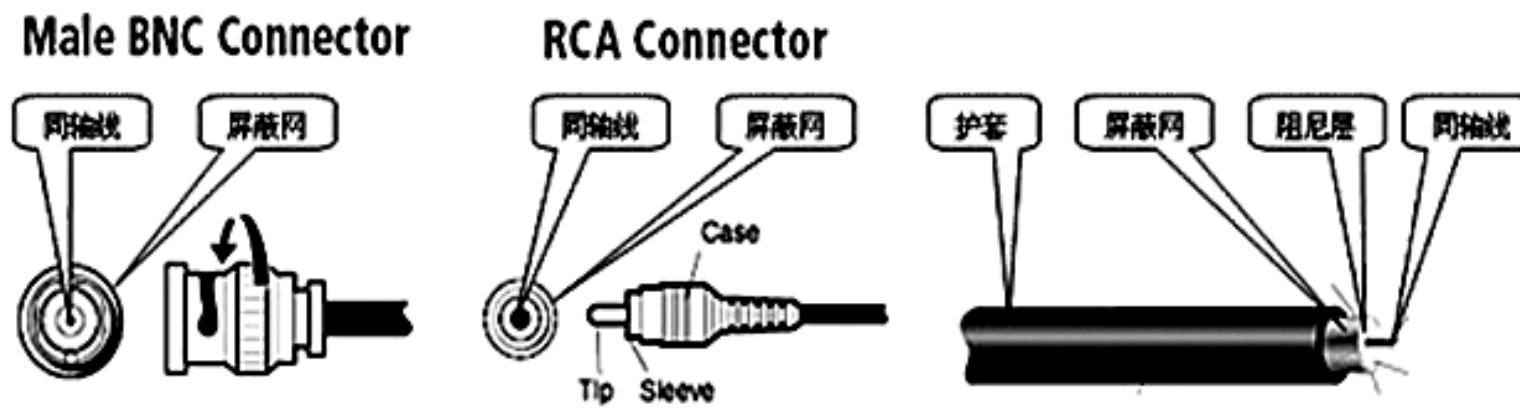
- 是指每个基色分量作为独立的电视信号，如：RGB或YUV



视频信号类型

• 复合视频信号

- 复合视频信号是将亮度信号和色度信号采用频谱间置方法复合在一起。这种方法易导致亮色串扰、清晰度降低等问题。
- 复合电视信号接口的基本特性：
 - ①传输介质：单根带屏蔽的同轴电缆
 - ②传输阻抗：75Ω
 - ③常用接头：BNC接头、莲花(RCA)接头
 - ④接线标准：插针是同轴信号线，外壳公共地是屏蔽网线



复合视频信号常用接线头和标准

视频信号类型

- 分量视频信号

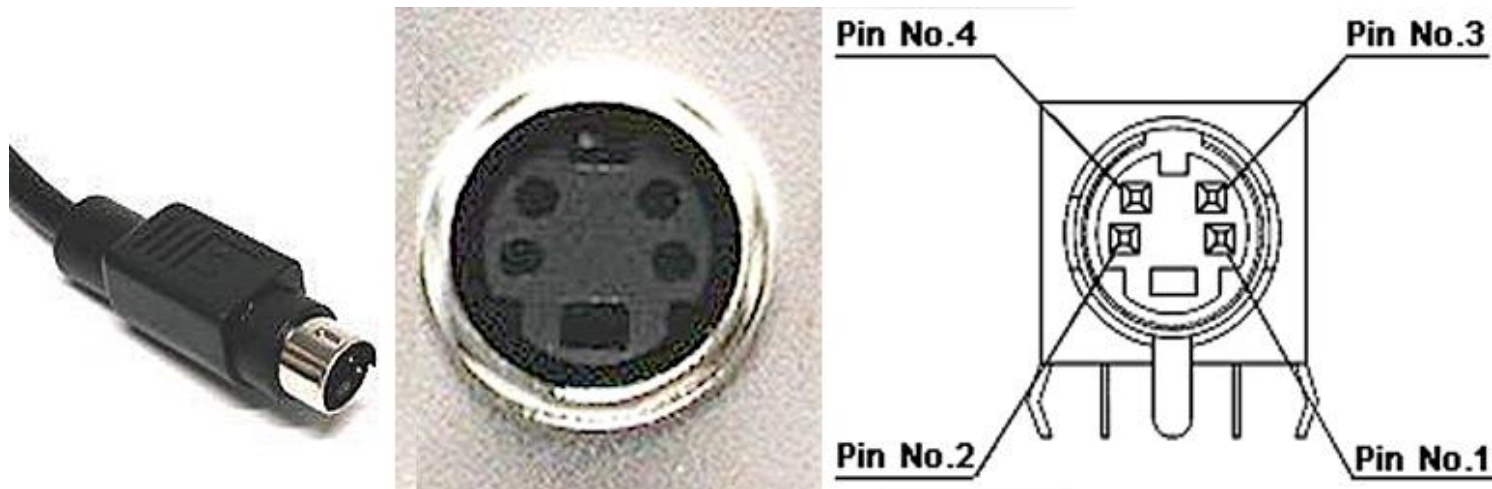
- 分量视频信号(Component Video Signal)是指每个基色分量作为独立的视频信号。每个基色既可以分别用R、G和B表示，也可以用亮度-色差表示，如YIQ、YUV。
- 使用分量视频信号是表示颜色的最好方法，但需要比较宽的带宽和同步信号。



视频信号类型

- **S-Video信号（S端子视频信号）**

- S端子视频信号俗称S端子信号，它同时传送两路信号：亮度信号Y和色度信号C。由于将亮度和色度分离，所以图象质量优于复合视频信号，色度对亮度的串扰现象也消失。
- 色度信号仍须解调，所以其图象质量的提高是有限的。
- S端子用四芯插头，一些计算机显卡或是非线性编辑卡也有用七芯插头，其外形与S端子一样，只是又包含了复合视频信号。



S-Video四芯插头（座）

内容提纲

- 动画和视频基础
- 电影和电视
- 电视图像的数字化
- 数字视频的获取
- 数字视频编辑技术
- 视频显示设备

电视信号的数字化

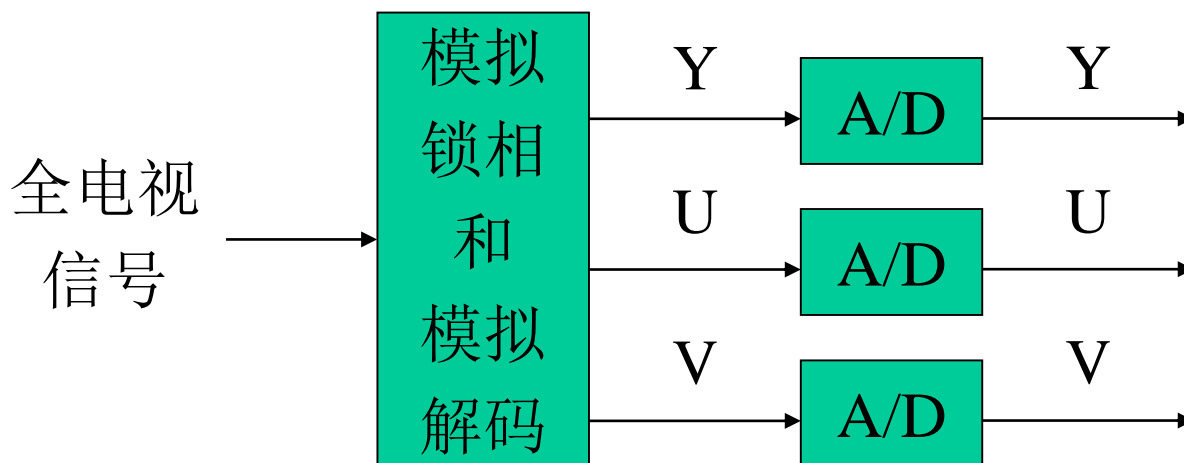
- 数字电视图像有很多优点
 - 例如，可直接进行随机存储使电视图像的检索变得很方便
 - 复制数字电视图像和在网络上传输数字电视图像都不会造成质量下降
 - 很容易进行非线性电视编辑等等

电视信号的数字化

- 常用的方法有两种：
 - 分量电视信号的数字化
 - 先从复合彩色电视图像中分离出彩色分量，然后数字化
 - 复合电视信号的数字化
 - 对色度信号和亮度信号共频带所形成的复合电视信号直接进行数字化

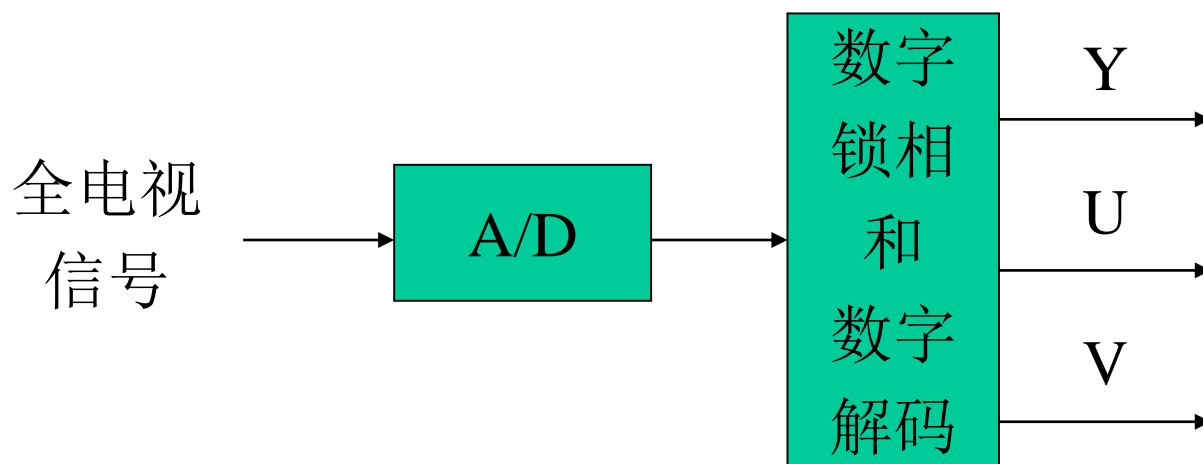
电视信号的数字化

- **分量电视信号的数字化**：首先把模拟的全彩色电视信号（如来自录象带、激光视盘、摄像机等）分离成YCbCr, YUV, YIQ或RGB彩色空间中的分量信号，然后用三个A/D转换器分别对它们数字化



电视信号的数字化

- **复合电视信号的数字化**：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



电视信号的数字化

- 早在20世纪80年代初，国际无线电咨询委员会 CCIR (International Radio Consultative Committee) 就制定了彩色电视图像数字化标准，称为**CCIR 601标准**，现改为**ITU-R BT.601标准**
- 该标准规定了彩色电视图像转换成数字图像时使用的采样频率，RGB和YCbCr两个彩色空间之间的转换关系等

电视信号的数字化

- 为了保证信号的同步，采样频率必须是电视信号行频的倍数。CCIR为NTSC、PAL和SECAM制式制定的共同的电视图像采样标准：

$$f_s = 13.5\text{MHz}$$

- 这个采样频率正好是PAL、SECAM制行频的864倍，NTSC制行频的858倍，可以保证采样时采样时钟与行同步信号同步
- 对于4:2:2的采样格式，亮度信号用 f_s 频率采样，两个色差信号分别用 $f_s/2 = 6.75\text{MHz}$ 的频率采样

电视信号的数字化

表7-04 彩色电视数数字化参数摘要

采样格式	信号形式	采样频率	样本数/扫描行		数字信号取值
		(MHz)	NTSC	PAL	范围(A/D)
	Y	13.5	<u>858</u> (720)	<u>864</u> (720)	220级 (16 ~ 235)
<u>4:2:2</u>	Cr	6.75	429 (360)	432 (360)	225级 (16 ~ 240)
	Cb	6.75	429 (360)	432 (360)	(128 ± 112)
	Y	13.5	858 (720)	864 (720)	220级 (16 ~ 235)
<u>4:4:4</u>	Cr	13.5	858 (720)	864 (720)	225级 (16 ~ 240)
	Cb	13.5	858 (720)	864 (720)	(128 ± 112)

电视信号的数字化

- CCIR为NTSC制、PAL制和SECAM制规定了共同的电视图像采样频率

- 对PAL制、SECAM制：

$$f_s = 625 \times 25 \times N = 15625 \times N = 13.5 \text{ MHz}$$

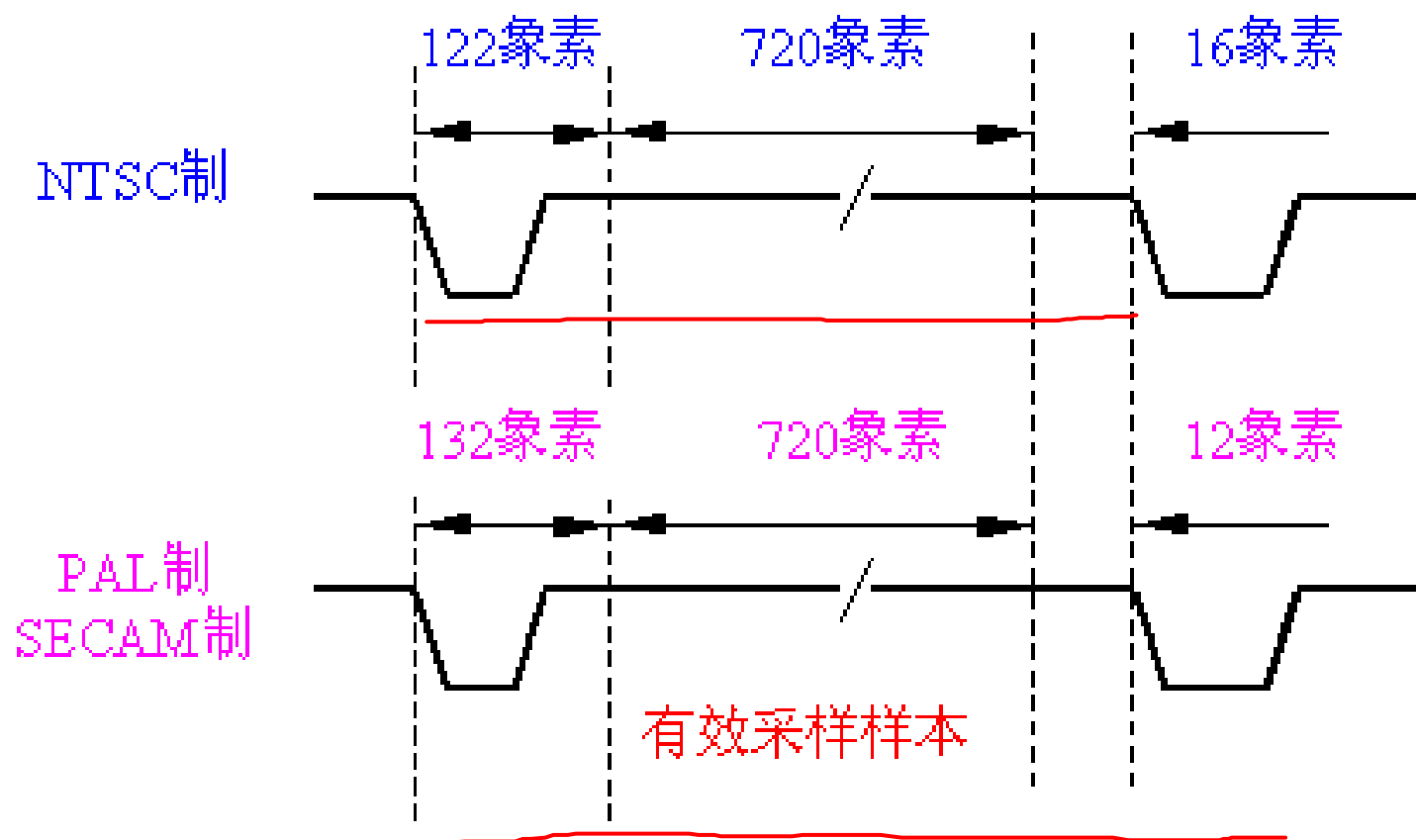
N=864, N为每一扫描行上的采样数目

- 对NTSC制：

$$f_s = 525 \times 29.97 \times N = 15734 \times N = 13.5 \text{ MHz}$$

N=858, N为每一扫描行上的采样数目

电视信号的数字化



电视信号的数字化

- 人眼对彩色细节的分辨能力远比对亮度细节的分辨能力低，可利用人的视觉特性降低数据量
 - ITU-R BT.601推荐使用4:2:2的彩色电视图像采样格式。Y用13.5 MHz的采样频率，Cr、Cb用6.75 MHz的采样频率
 - 根据不同的应用，分量信号还可以采用其它不同的采样格式，4:4:4、4:2:2、4:2:0等

电视信号的数字化

- 4种采样格式

- 4:4:4 在每条扫描线上每4个连续的采样点取4个亮度Y样本、4个红色差Cr样本和4个蓝色差Cb样本，相当于每个像素用3个样本表示；
- 4:2:2 在每条扫描线上每4个连续的采样点取4个亮度Y样本、2个红色差Cr样本和2个蓝色差Cb样本，相当于每个像素用2个样本表示；
- 4:1:1 在每条扫描线上每4个连续的采样点取4个亮度Y样本、1个红色差Cr样本和1个蓝色差Cb样本，相当于每个像素用1.5个样本表示；

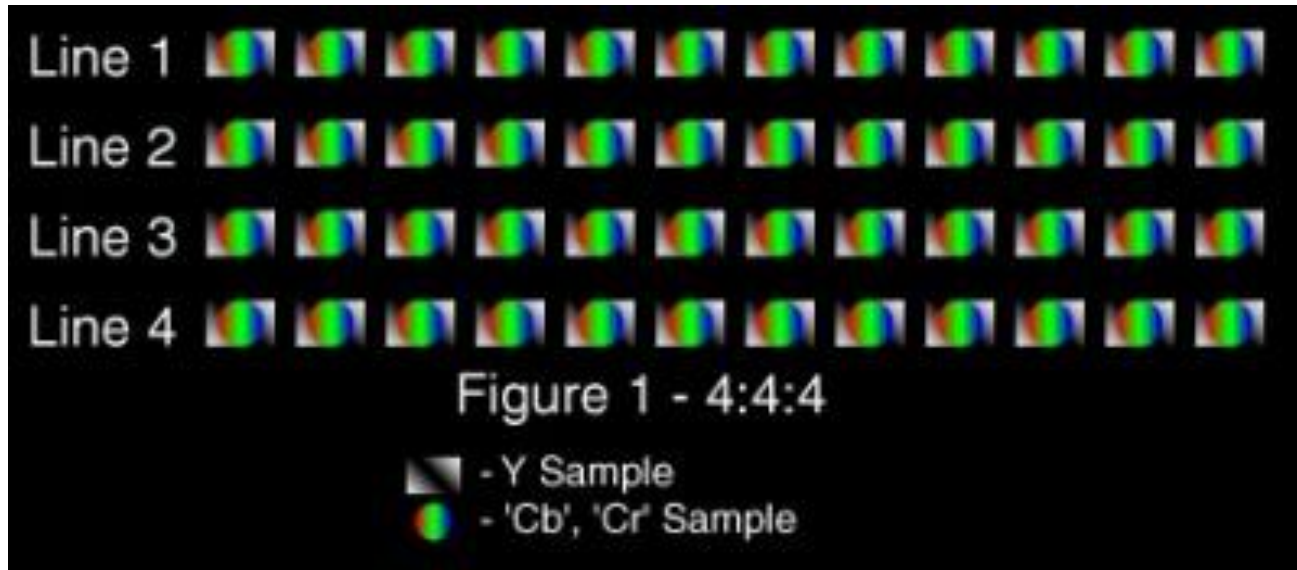
电视信号的数字化

- 4种采样格式

- 4:2:0 在**水平和垂直方向**上每2个连续的采样点取2个亮度Y样本、1个红色差Cr样本和1个蓝色差Cb样本，相当于每个像素用1.5个样本表示；

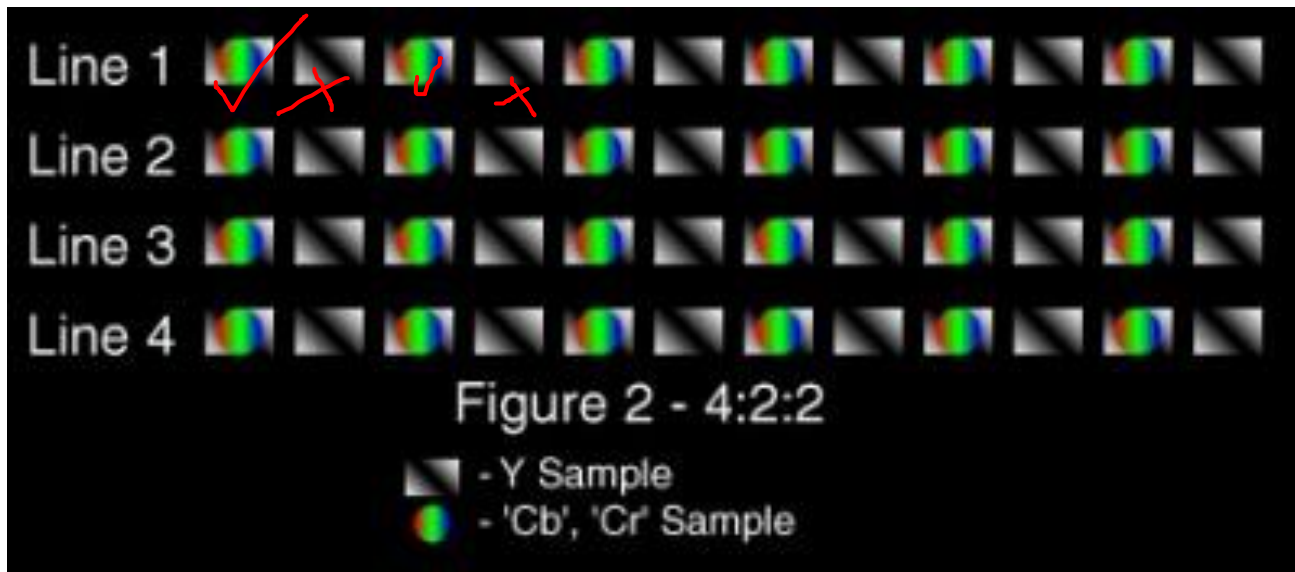
YCbCr 4:4:4

- 色度信号分辨率最高的格式是4:4:4
 - 每4点Y采样，就有相对应的4点Cb和4点Cr
 - 在这种格式中，色度信号的分辨率和亮度信号的分辨率是相同的
 - YYY YCbCrCbCrCbCrCbCr



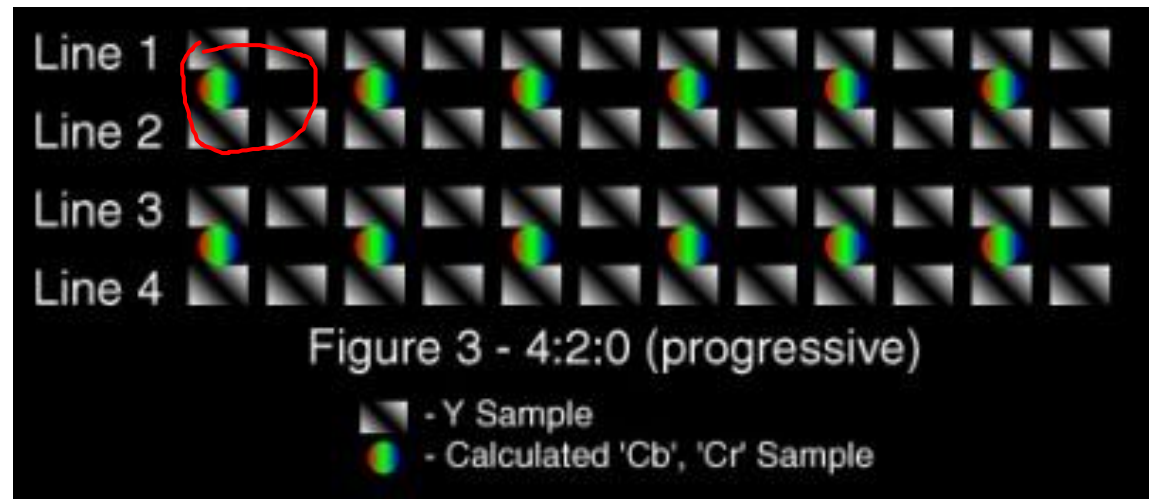
YCbCr 4:2:2

- 每4点Y采样，就有2点Cb和2点Cr
 - 每个像素都有与之对应的亮度采样，同时一半的色度采样被丢弃，色度采样信号每隔一个采样点才有一个
 - 当整张画面显示的时候，缺少的色度信息会由两侧的颜色通过内插补点的方式运算得到
 - YYY YCbCrCbCr



YCbCr 4:2:0

- 色度采样在每条横向扫描线上只有亮度采样的一半，扫描线的条数上，也只有亮度的一半
 - 无论是横向还是纵向，色度信号的分辨率都只有亮度信号的一半。如果整张画面的尺寸是720*480，那么亮度信号是720*480，色度信号只有360*240。
 - “缺失”的色度采样不单单要由左右相邻的采样通过内插补点计算补充，整行的色度采样也要通过它上下两行的色度采样通过内插补点运算获得。
 - YYYYCbCr



数字视频文件格式

- 目前，视频文件格式可以分为两大类
 - 适合本地播放的本地影像视频
 - 适合在网络中播放的网络流媒体影像视频
- 尽管后者在播放的稳定性和播放画面质量上可能没有前者优秀，但网络流媒体影像视频的广泛传播性使之正被广泛应用于视频点播、网络演示、远程教育、网络视频广告等等互联网信息服务领域

数字视频文件格式

- 数字视频文件格式

- MPEG (Moving Pictures Experts Group)
 - MPEG-1、 MPEG-2、 MPEG-4
- MOV
- ASF (Advanced Streaming Format) , MPEG-4算法
- WMV (Windows Media Video)
- RM (Real Media)
- RMVB

本地影像视频

- **AVI格式(Audio Video Interleaved)**，即音频视频交错格式
 - 所谓“音频视频交错”，就是可以将视频和音频交织在一起进行同步播放
 - 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**
 - 是由索尼、松下、JVC等多家厂商联合提出的一种家用数字视频格式
 - 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

本地影像视频

- **MPEG格式**，英文全称为Moving Picture Expert Group，即运动图像专家组格式
 - 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)。
 - 目前MPEG视频格式常见的压缩标准是：MPEG-1、MPEG-2、和MPEG-4。

本地影像视频

- **DivX格式**，是由MPEG-4衍生出的另一种视频编码标准，也即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，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
 - 默认的播放器是苹果的QuickTime Player。具有较高的压缩比率和较完美的视频清晰度等特点
 - 最大特点是跨平台性，即不仅能支持MacOS，同样也能支持Windows系列

网络影像视频

- **ASF格式** (Advanced Streaming Format)
 - 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。使用了MPEG-4的压缩算法
- **WMV格式** (Windows Media Video)
 - 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
 - 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

网络影像视频

- **RM格式**，Real Networks公司所制定的音频视频压缩规范称为Real Media
 - 用户可以使用RealPlayer或RealOne Player对符合RealMedia技术规范的网络音频/视频资源进行实况转播，并且RealMedia可以根据不同的网络传输速率制定出不同的压缩比率，从而实现在低速率的网络上进行影像数据实时传送和播放。
 - 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

网络影像视频

- **RMVB格式**，是一种由RM视频格式升级延伸出的新视频格式
 - RMVB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：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给快速运动的画面场景
 - 在保证了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。