

随着视频处理技术的发展,出现了越来越多的有关视频的应用,如数字电视、视频会议、视频点播和网上电影等。

本章首先介绍视频的基本概念,然后重点介绍模拟视频和数字视频,对视频数字化涉及的一些基本技术进行简单说明;最后以 Premiere 视频制作软件为平台介绍视频的相关处理技术。

5.1 视频的基本概念

视频实际是一系列静态图像序列,如图 5-1 所示。单位时间内连续地播放若干幅静态图像就能得到动态图像组成的视频。每一幅独立的图像称为一帧。帧是构成视频的基本图像单元。连续播放的速率一般为 20fps(帧/秒)或 30fps(帧/秒),人的眼睛就不会感觉失真。

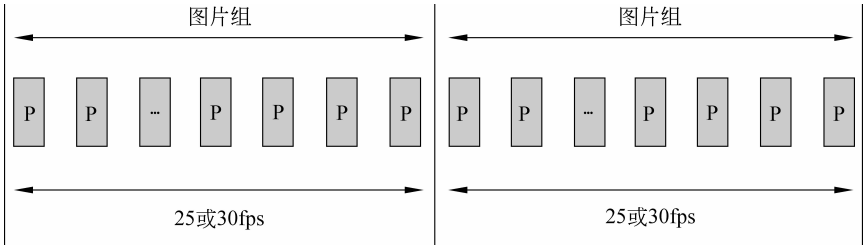


图 5-1 视频示例

人眼对每秒播放一定帧数的多幅静态图像序列之所以不会感觉失真或“停顿”,是由于人眼的视觉暂留效应。视觉暂留(Persistence of Vision)是光对人眼的视网膜所产生的视觉,在光停止作用后,仍保留一段时间的现象。视觉暂留,又称为视觉残留。例如,人眼直视太阳数秒后,人眼将残留一个强光源的影像。视觉暂留效应产生的根本原因是人的视觉神经的反应时间是二十四分之一秒,具有一定的时间差。

维基网站给出了视频暂留的定义,如下:

Persistence of vision refers to the optical illusion whereby multiple discrete images blend into a single image in the human mind and believed to be the explanation for motion perception in cinema and animated films.

动画也是一种视频。动画是物体随时间变化的行为或动作,是利用计算机图形技术进行绘制而获得的,而视频则是模拟信号源(如电视和电影等)经过数字化处理后的图像和同步伴音的混合体。因此,动画与视频虽然都是一帧一帧图像的组合,但其来源或制作方式不同。

视频信息分为模拟视频和数字视频两类。数字视频是以数字化的形式表示和记录的连续图像的信息,可以直接在计算机等多种数字设备上存储、传递和播放。而模拟视频的信息在时间和幅度上都是连续的图像信号,不能直接由数字计算机进行处理。为了在计算机等各种数字视频信息处理设备中处理模拟视频,需要进行数字化处理,即对模拟视频信息进行二进制编码。

5.1.1 模拟视频

模拟视频指的是每一帧图像都是实时获取的自然景物真实图像序列。日常生活中的电视、电影都是模拟视频。模拟视频具有成本低和还原性好等优点,视频画面往往会给人一种身临其境的感觉。但是,模拟视频的最大缺点是不论记录的图像信号有多好,经过长时间的存放之后,信号和画面的质量将大大降低,或者经过多次复制之后,画面失真非常明显。

1. 模拟视频信号的分类

根据视频信号中对图像的亮度信号和色差信号的编码方式,模拟视频信号可以分为如下几类。

(1) 分量视频信号。分量视频信号(Component Video Signal)指的是每个分量可以作为独立的电视信号,如图 5-2 所示。每个分量既可以是 RGB 颜色空间的红色信号 R、绿色信号 G 和蓝色信号 B,也可以是 YUV 颜色空间中的亮度信号 Y、色差信号 U 和色差信号 V,或者是 YIQ 颜色空间中的亮度信号 Y、色差信号 I 和色差信号 Q。分量视频信号对每个分量都独立考虑,使用分量视频信号是表示颜色的最精确的信号。同时,分量视频信号需要较宽的带宽和同步信号。计算机输出的 VGA(Video Graphics Array)视频信号就是一种常见的分量视频信号,它具有分辨率高、显示速率快、颜色丰富等优点,在彩色显示器领域得到了广泛的应用。

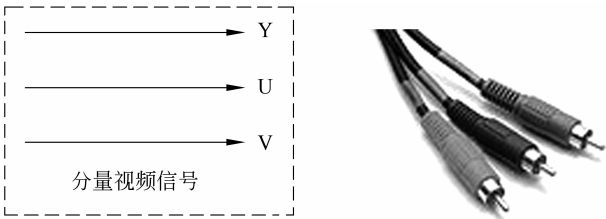


图 5-2 分量视频信号

(2) 复合视频信号。复合视频信号(Composite Video Signal)是将色差信号在亮度信号上进行编码,作为单个信号与亮度信号拥有相同的带宽,如图 5-3 所示。这种类型的视

频信号一般不含伴音信号,传输时对伴音信号单独设置通道。

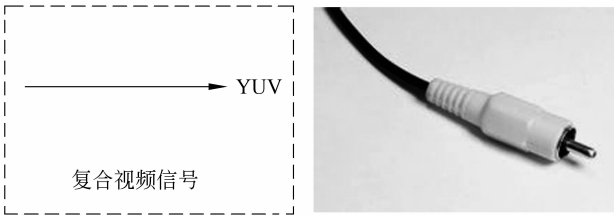


图 5-3 复合视频信号

(3) 分离视频信号。分离视频信号 (Separated Video Signal) 是一种介于分量视频信号和复合视频信号之间的一种方案,将亮度信号和色差信号分成两路独立的模拟信号,一路用于传输亮度信号,一路用于传输色差信号,如图 5-4 所示。这种信号由于将亮度信号和色差信号分开传送,可以减少相互的干扰。

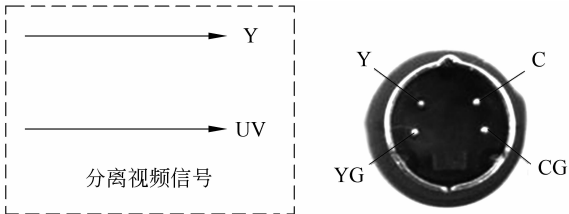


图 5-4 分离视频信号

(4) 射频信号。为了实现模拟视频信号的远距离传输,必须将亮度信号、色差信号、伴音信号和同步信号调制成一个单一的高频信号,这种高频信号就是射频信号,如图 5-5 所示。当电视接收设备接收射频信号时,首先将其还原成电视信号,然后再分别还原视频信号和伴音信号。



图 5-5 射频信号

2. 电视模拟视频

传统的电视视频也是一种模拟视频。电视图像的播放是基于扫描技术实现的。扫描分为水平扫描和垂直扫描。水平扫描指的是扫描点从画面左侧匀速移动到右侧,垂直扫描指的是水平扫描线按均匀间隔在垂直方向上移动。垂直扫描又称为场扫描。电视扫描是水平扫描和垂直扫描的综合。水平扫描所能分辨的点数称为水平分辨率。垂直扫描的

行数称为垂直分辨率。显然,相同尺寸的扫描范围内,水平分辨率越大,分辨的点数越多,点的尺寸越小;同样,垂直分辨率越大,扫描线越多,显示的图像越清晰。

一般电视标准规定:每秒播放 25 帧或 30 帧图像。当电视每秒播放 25 帧时,人眼仍然会感到闪烁,容易疲劳。其解决方法是分两次来播放一幅图像,25 个完整帧/秒的播放就变为 50 个半帧/秒的播放,人眼感到舒服多了。而且,由于一幅图像的行数高达 625,且是隔行扫描,即先扫描 1、3、5、7…行,然后再扫描 2、4、6、8…行,人眼感受到的仍然是连续活动的图像。隔行扫描的示意,如图 5-6 所示。



图 5-6 隔行扫描

每秒传送的图像帧数称为帧频。每秒扫描的场数称为场频。因此,当采用逐行扫描时,帧频等于场频;当采用隔行扫描时,场频是帧频的两倍。

对于逐行扫描

$$f_{\text{Field}} = f_{\text{Frame}}$$

对于隔行扫描

$$f_{\text{Field}} = 2f_{\text{Frame}}$$

这里, f_{Field} 是场频, f_{Frame} 是帧频。

因此,对于 NTSC (National Television System Committee) 电视制式来说,其规定的帧频为 30Hz,则如果是隔行扫描,其场频为 60Hz;对于 PAL (Phase Alternation Line) 和 SECAM (SEquential Couleure Avec Memoire) 电视制式来说,其规定的帧频为 25Hz,如果是隔行扫描,其场频为 50Hz。

3. 模拟电视制式

电视信号的标准又称制式,可以简单地理解为用来实现电视图像或声音信号所采用的一种技术规则或标准。具体说来,电视制式主要规定如下。

- (1) 电视图像信号的传输规则或标准;
- (2) 电视图像信号的显示规则或标准。

一般来说,电视制式包括模拟电视制式和数字电视制式。对于模拟电视来说,又分为黑白电视制式和彩色电视制式。黑白电视制式规定的主要内容包括图像和伴音的调制方式、图像信号的极性、图像和伴音的载频差、频带宽度、频道间隔以及扫描行数。目前世界各国所采用的黑白电视制式包括 A、B、C、D、E、F、G、H、I、K、K1、L、M 和 N 制式,共计 13 种(其中 A、C、E 已不采用),中国采用的是 D、K 制式。

彩色电视制式主要包括 NTSC、PAL、SECAM 三种,如表 5-1 所示。中国大部分地区使用 PAL 制式,日本、韩国及东南亚地区与美国等欧美国家使用 NTSC 制式,俄罗斯则使用

SECAM 制式。

NTSC 制式是 1952 年由美国国家电视标准委员会制定的彩色电视标准。采用正交平衡调幅的技术。美国、加拿大等大部分西半球国家以及中国台湾、日本、韩国和菲律宾等均采用 NTSC 制式。

PAL 制式是 1962 年由西德制定的彩色电视标准。采用逐行倒相正交平衡调幅的技术,克服了 NTSC 制式相位敏感造成色彩失真的缺点。西德、英国等一些西欧国家、新加坡、中国、澳大利亚和新西兰等国家和地区采用 PAL 制式。

SECAM 法文的缩写,意为顺序传送彩色信号和存储恢复彩色信号制,是由法国 1956 年提出、1966 年制定的一种新的彩色电视制式。该制式克服了 NTSC 制式相位失真的缺点。法国、东欧和中东一带采用 SECAM 制式。

表 5-1 各种不同电视制式的技术指标

电 视 制 式	NTSC	PAL	SECAM
帧频/Hz	30	25	25
垂直分辨率	525	625	625
亮度宽度/MHz	4.2	4.43	4.25
色度带宽/MHz	1.3 (I) 0.6 (Q)	1.3(U) 1.3(V)	>1.0(U) >1.0(V)
声音载波/MHz	4.5	6.5	6.5

从表 5-1 可以看出,PAL 电视制式和 SECAM 电视制式比较接近,每秒传输图像的帧数都是 25,垂直分辨率都是 625,都是采用 YUV 颜色空间。对于 NTSC、PAL 和 SECAM 三种模拟电视制式来说,亮度信号使用带宽都远远大于色度信号使用带宽,这也证明了使用 YUV 颜色空间或 YIQ 颜色空间可以方便对数据进行压缩。

5.1.2 数字视频

数字视频指的是基于二进制编码且能够被数字计算机处理的视频信号。与模拟视频相比,数字视频具有以下特点。

(1) 数字视频可以不失真地进行多次复制,而模拟视频每复制一次,则会产生一次误差积累,造成信号失真。日常生活的经验也证明了这点,一个使用数字硬盘存储的电影在被拷贝到另一个硬盘上重新进行播放不会产生任何的失真,而一个胶片电影经过多次翻录后则极有可能造成图像画面的局部失真。

(2) 模拟视频长时间存放后视频质量会降低,而数字视频可以进行长时间存储。例如,一个数字硬盘存储的电影经过几年甚至几十年后重新播放不会有任何失真,而一个胶片电影存放几十年后可能会产生画面的模糊。

(3) 可以对数字视频进行非线性编辑,方便进行特技效果处理等。非线性编辑指的是借助计算机来进行数字化制作,对素材的调用可以瞬间实现,不必在磁带上进行顺序寻找,突破单一的时间顺序编辑限制,具有简便、随机的特点。非线性编辑是相对于传统的

以时间顺序进行的线性编辑而言的。数字视频由于是以二进制编码形式进行存储,且是存放在随机访问的硬盘上,可以使用视频处理工具如 Premiere 等进行非线性编辑,也可方便地进行特效处理。

数字视频的一个最大规模的应用是数字电视。与模拟电视制式类似,数字电视也必须具有相应的标准。1997 年,数字视频广播(Digital Video Broadcasting,DVB)联盟发表了数据电视技术规范,包括卫星电视传输标准 DVB-S(Satellite)、有线电视传输标准 DVB-C(Cable)和地面数字电视传输标准 DVB-T(Terrestrial)。其中,DVB-S 规定了卫星数字广播调制标准,使原来传送一套 PAL 制式节目的频道可以传播四套数字电视节目,大大提高了卫星传输的效率。DVB-C 规定了在有线电视传播数字电视的调制标准,使原来传送一套 PAL 制式节目的频道可以传播四至六套数字电视节目。DVB-S 和 DVB-C 这两个全球化的卫星和有线传输方式标准,目前已作为世界统一标准被大多数国家所接受。而对于地面数字电视传输标准,则有三个,分别为欧盟的 DVB-T 标准、美国的先进电视制式委员会(Advanced Television System Committee,ATSC)标准和日本的综合业务数字广播(Integrated Services Digital Broadcasting,ISDB)标准。数字电视标准如表 5-2 所示。

表 5-2 数字电视标准

电视标准	DVB-S	DVB-C	DVB-T	ATSC	ISDB
视频编码	MPEG-2	MPEG-2	MPEG-2	MPEG-2	MPEG-2
音频编码	MPEG-2	MPEG-2	MPEG-2	AC-3	MPEG-2
复用方式	MPEG-2	MPEG-2	MPEG-2	MPEG-2	MPEG-2
带宽	—	—	8M	6M	27M

当前,世界各国都根据本国的具体情况,慎重地选择地面数字电视传输标准。从世界范围看,除了美国外,还有加拿大、阿根廷、韩国等国家采用美国的 ATSC 标准。而欧洲所有国家和澳大利亚、新加坡、印度等国则选用了欧洲联盟的 DVB-T 标准。

目前,我国已经公布了自己的地面数字电视传输标准(Digital Television Terrestrial Multimedia Broadcasting,DTMB),实现了固定电视和公共交通移动电视的数字电视信号传送,并成为第四个地面数字电视传输国际标准。

5.2 视频数字化

5.2.1 视频数字化的方法

如果需要在数字计算机中处理模拟视频,必须要将模拟视频转换为数字视频,这就是视频的数字化。一般来说,视频数字化有两种方法。

(1) 复合数字化(Recombination Digitalization)

复合数字化指的是首先使用一个高速的模/数 (Analogy/Digital) 转换器对全彩色电视信号进行数字化,然后再分离亮度和色差信号,最后再转换成 RGB 分量。

(2) 分量数字化(Component Digitalization)

分量数字化指的是首先将复合视频信号中的亮度和色差信号分离,得到 YUV 或 YIQ 分量,然后使用三个模/数转换器对三个分量分别进行数字化,最后再转换成 RGB 分量。

不管是复合数字化还是分量数字化,数字化的过程都是要将模拟信号经过采样、量化和编码的过程转为数字信号,如图 5-7 所示。

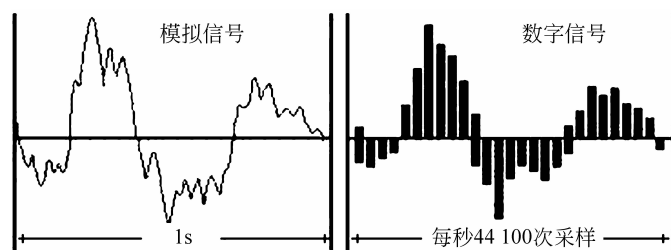


图 5-7 模拟信号和数字信号

5.2.2 图像子采样

模拟视频转换为数字视频的过程,同样也涉及采样的问题。对彩色电视图像进行采样时,可以采用两种采样方法:(1)使用相同的采样频率对图像的亮度信号和色差信号进行采样;(2)使用不同的采样频率对图像的亮度信号和色差信号进行采样。

对于第二种采样方法,对色差信号使用的采样频率比对亮度信号使用的采样频率低,这就是图像子采样。图像子采样是根据人的视觉系统所具有的特性:人眼对色差信号的敏感程度比对亮度信号的敏感程度低。基于这个特性可以对组成图像像素颜色的亮度信号和色差信号使用不同的采样率,即对亮度信号的采样率高于对色差信号的采样率。

下面以 YUV 颜色空间为例,说明几种常见的图像子采样。

1. 4: 4: 4 YUV 格式的图像子采样

这种图像采样格式不是真正的子采样格式,它是指在每条扫描线上每 4 个连续的采样点取 4 个亮度 Y 样本、4 个红色色差 U 样本和 4 个蓝色色差 V 样本,共有 12 个样本表示 4 个像素,每个像素使用 3 个样本表示,如图 5-8 所示。扫描线编号隔行顺序递增,说明此为隔行扫描。压缩比为 $3: 3 = 1.0$ 。实际上,4: 4: 4 YUV 格式的图像采样不是一种图像子采样。

采样前的 4 个像素: $[Y_1 \ U_1 \ V_1], [Y_2 \ U_2 \ V_2], [Y_3 \ U_3 \ V_3], [Y_4 \ U_4 \ V_4]$;

4: 4: 4 采样后的码流为: $Y_1 \ U_1 \ V_1 \ Y_2 \ U_2 \ V_2 \ Y_3 \ U_3 \ V_3 \ Y_4 \ U_4 \ V_4$;

采样后的 4 个像素: $[Y_1 \ U_1 \ V_1], [Y_2 \ U_2 \ V_2], [Y_3 \ U_3 \ V_3], [Y_4 \ U_4 \ V_4]$ 。

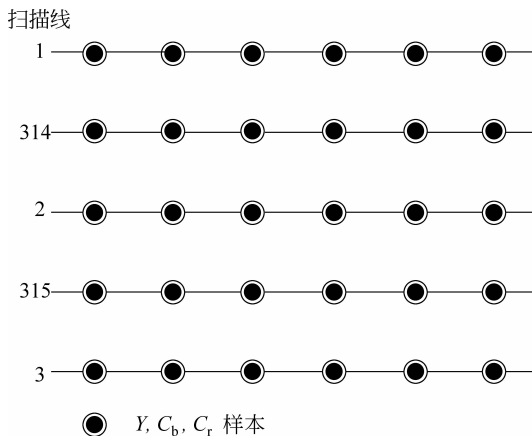


图 5-8 4: 4: 4 采样

2. 4: 2: 2 YUV 格式的图像子采样

这种子采样格式是指在每条扫描线上每 4 个连续的采样点取 4 个亮度 Y 样本、2 个红色色差 U 样本和 2 个蓝色色差 V 样本, 共有 8 个样本表示 4 个像素, 平均每个像素使用 2 个样本表示, 如图 5-9 所示。压缩比为 $3: 2 = 1.5$ 。

采样前的 4 个像素: $[Y_1 U_1 V_1], [Y_2 U_2 V_2], [Y_3 U_3 V_3], [Y_4 U_4 V_4]$;

4: 2: 2 采样后的码流为: $Y_1 U_1 Y_2 V_2 Y_3 U_3 Y_4 V_4$;

采样后的 4 个像素: $[Y_1 U_1 V_2], [Y_2 U_1 V_2], [Y_3 U_3 V_4], [Y_4 U_3 V_4]$ 。

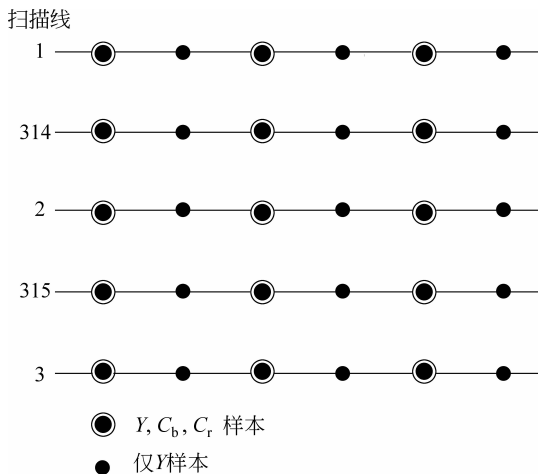


图 5-9 4: 2: 2 采样

3. 4: 1: 1 YUV 格式的图像子采样

这种子采样格式是指在每条扫描线上每 4 个连续的采样点取 4 个亮度 Y 样本、1 个红色色差 U 样本和 1 个蓝色色差 V 样本, 共有 6 个样本表示 4 个像素, 平均每个像素使用