



思想引领



视频讲解

视频系统的整个流程环节中,信号的采集扮演着至关重要的角色,它直接决定了最终图像质量所能达到的上限。自然光信号通过光学系统的精确引导,被图像传感器接收并转换为电信号。由于光学系统的特性、传感器的性能以及视频系统在传输和存储方面的需求,电信号的处理显得尤为关键。本章将重点探讨彩色图像在视频系统中的采集过程,详细讲解像素分解、模拟电视扫描、图像传感器的工作原理,以及三基色信号的生成、彩色校正和视频系统的 γ 特性等关键环节。

3.1 视频传像基本原理

视频传像的过程大致可分为图像信号的采集、传输和重现三部分。视频传像流程,如图3-1所示。



图 3-1 视频传像流程

首先,图像的采集本质上是一个“光电转换”过程,由摄像机负责捕捉光信号并将其转换成电信号,接着这些电信号经过处理后被输出为视频信号。在视频传像过程中,图像的采集是整个系统的首要步骤,也是质量源头,直接关系到重现画面的质量。

第二部分,即传输部分,可以概括为“电-电传输”。在这一环节中,视频信号通过有线电视网络(如光缆、同轴电缆)、无线广播(如地面广播、卫星广播)或互联网等媒介,被传送至接收端进行接收解码和图像重现。发送端的传输过程包括信号处理、信源编码、复用、信道编码与调制,以及最终的发射(或输出),这些环节在图3-1中被综合表述为“信号处理”和“调制发射”。而接收端的信号处理则是发送端操作的逆过程,它涵盖了信号的接收、解调、信道解码、解复用以及信源解码,这些环节在图3-1中被简化为“信号接收、解调与还原”。

最后的重现部分,本质上是“电光转换”过程,负责将处理后的电信号转换为光信号,从而在显示设备上重现出图像。显然,显示技术对图像的重现有重要的影响,这部分内容将在

第5章进行介绍。

3.1.1 视频的传送

在视频传像系统的前端,摄像机负责捕获景物反射或透射出的光线,并将这些光线转换为适合传输的电信号。由于景物占据三维空间,并且其反射(发出)的光线随时间连续变化,因此,摄像机捕获的视频可以用一个依赖于景物的三维坐标 (x, y, z) 以及时间 t 的函数进行表示。根据三基色原理,任意光色均可通过三个基色分量来表征,如红、绿和蓝三基色。由此,摄像机捕获的视频信号可以用三组函数表示,分别是红分量信号 $f_R(x, y, z, t)$ 、绿分量信号 $f_G(x, y, z, t)$ 、蓝分量信号 $f_B(x, y, z, t)$ 。对于常用的显示设备以及摄像机中的感光器件(即图像传感器)来说,它们重现或转换的都是无深度信息的二维视频图像,因此用 $f_R(x, y, t)$ 、 $f_G(x, y, t)$ 、 $f_B(x, y, t)$ 表示,不包含深度维度 z 。

上述三组函数所表示的视频信号,由于自变量连续取值,因此蕴含的信息量是无限的。然而,视频系统的传输带宽有限,因此需要对视频信号进行分解,“化整为零”,以便将完整的视频信号分解成较小的、有限的单元进行传输。这种做法不仅有助于提升系统的可实现性,还能确保实时性。

1. 视频信号的分解

视频信号可以从时间与空间两个维度进行分解。

从时间维度来看,得益于人眼的视觉惰性特性,连续的视频可以被分解为一幅幅的静止图像,一幅完整的图像被称为一帧。只要图像的换幅频率满足融合频率和临界闪烁频率要求,就能确保视觉上的连续性和无闪烁感。摄像机正是按照一定的时间间隔,逐帧输出这些图像。这样,每帧图像信号可以用 $f_R(x, y)$ 、 $f_G(x, y)$ 、 $f_B(x, y)$ 三组函数来表示。

从空间维度来看,受限于人眼分辨细节的能力,显示器并不需要重现原始图像的所有细节。平面图像可以被精细地分解成许多的小点,即像素(Pixel),像素是组成图像的最小单元。一幅图像所分解出的像素数量和大小会影响该图像的分辨率和清晰度,只要像素的数量足够多、尺寸足够小,人眼感知到的图像就会足够清晰,图像分解成不同像素数量的效果,如图3-2所示。



图3-2 图像分解成不同像素数量的效果

光电变换后的信号需要被传输到接收端,传输方式分为同时传输和顺序传输两种方式。同时传输指一幅图像的所有像素信息被同时转换为电信号并行发送,因此传输通道数量与图像总像素数一样庞大,难以实现。而顺序传输则按顺序逐个传送图像的像素单元,只需一个通道,虽然传输速率较高,但由于其在成本和技术上都更为可行,这种传输方式在实际应用中被普遍采用。

2. 视频图像的扫描

最早应用的视频传输系统是模拟电视,采用扫描的方式完成顺序传输。模拟电视系统顺序传输的扫描,如图 3-3 所示,从左至右、从上至下逐行逐帧地将二维视频图像分解为像素电信号顺序传输出去,这一过程称为摄像端扫描。接收端将收到的信号按照相同的规律,逐一扫描到显示器上并恢复出图像,这就是显像端扫描。摄像端与显像端的扫描需要严格同步,才能保证视频图像的完美重现。这里“同步”指扫描的水平和垂直起始位置一样,即“同相”;扫描的水平和垂直速度一样,即“同频”。收发不同步的视频图像重现效果如图 3-4 所示,分别给出了水平和垂直扫描同频不同相的重现效果。如果水平和垂直扫描既不同频也不同相,则画面会出现扭曲,严重时难以辨明图像内容。

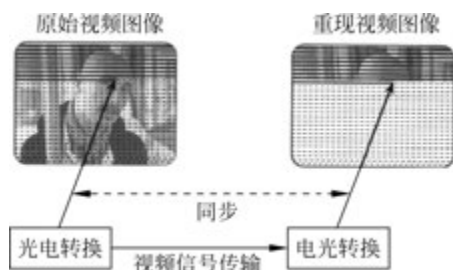


图 3-3 模拟电视系统顺序传输的扫描

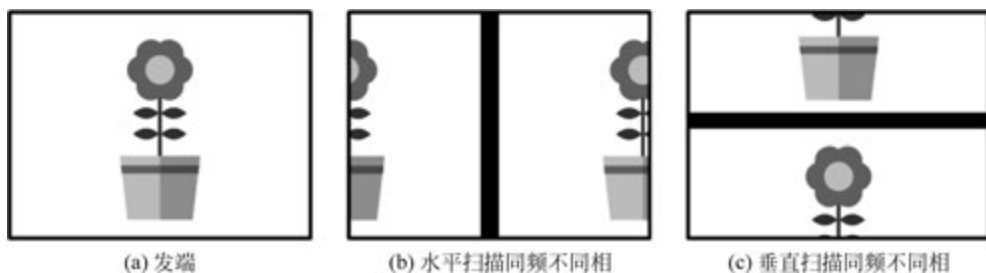


图 3-4 收发不同步的视频图像重现效果

3.1.2 视频的逐行扫描

模拟电视系统,对现代数字视频系统具有深远影响。早期模拟电视系统,摄像机和显示器都是 CRT 结构,核心组件均包括电子枪与偏转系统,两者配合实现发射电子束、加速电子束,并保证电子束在真空管体内匀速直线地运动,并最终在像面上从左到右,从上到下地一行行、一幅幅地扫描运动。不同的是,摄像机发射的是固定大小的电子束流,根据成像面上像素单元的亮暗导致的不同阻抗,完成的是从光像到电信号的拾取。而显像端 CRT 发射的是受视频信号调制电子束流,实现的是从电信号到光图像的重现。CRT 结构的摄像管和显示器如图 3-5 所示。模拟电视系统通常有两种扫描方式:逐行扫描和隔行扫描。

1. 逐行扫描原理

逐行扫描就是图像扫描顺序是从上到下、一行接一行地逐行匀速进行。在逐行扫描中,电子束从左到右,再从右至左回扫,才能完成一整行的扫描。这一个来回的水平扫描时长称为行周期,用 T_H 表示。其中,将从左至右的扫描过程(较慢),称为行扫描正程(简称行正程),所需时长用 T_{Ht} 表示;从右至左的回扫过程(较快),则称为行扫描逆程(简称行逆程),

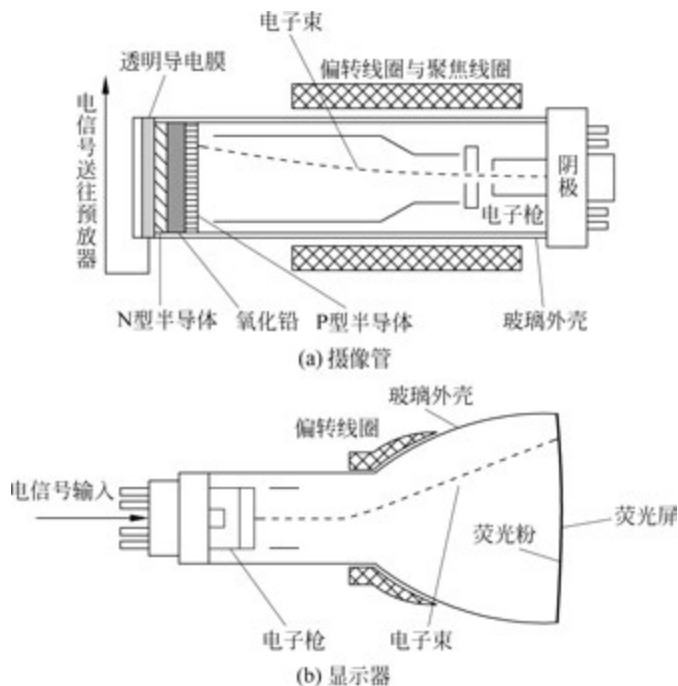


图 3-5 CRT 结构的摄像管和显示器

所需时长用 T_{Hr} 表示。显然, T_H 、 T_{Ht} 、 T_{Hr} 满足式(3-1)。

$$T_H = T_{Ht} + T_{Hr} \quad (3-1)$$

行周期的倒数,称为行频,用 f_H 表示,指一秒内扫描的行数,如式(3-2)所示。

$$f_H = \frac{1}{T_H} \quad (3-2)$$

习惯上,将行逆程与行周期的比值,称为行逆程系数,用 α 表示,如式(3-3)所示。

$$\alpha = \frac{T_{Hr}}{T_H} \quad (3-3)$$

图像在进行水平扫描的同时,还在进行垂直方向的扫描,因此实际上水平扫描的轨迹是略微向右下方倾斜的。

同理,帧周期 T_F 表示在垂直方向上,从上往下扫描一幅图像后(较多行),再从下往上回扫(较少行)的时间。其中,将从上往下扫描的整个过程,称为帧扫描正程(简称帧正程),其总时长用 T_{Ft} 表示,从下到上的扫描过程称为帧扫描逆程(简称帧逆程),时长用 T_{Fr} 表示。对应的,将帧逆程与帧周期的比值,称为帧逆程系数,用 β_F (通常 $\beta_F \approx 0.08$) 表示。同时,帧周期的倒数称为帧频,用 f_F 表示,意味着一秒内扫描的视频帧数。显然, T_F 、 T_{Ft} 、 T_{Fr} 、 β_F 、 f_F 满足如式(3-4)~式(3-6)所示关系。

$$T_F = T_{Ft} + T_{Fr} \quad (3-4)$$

$$\beta_F = \frac{T_{Fr}}{T_F} \quad (3-5)$$

$$f_F = \frac{1}{T_F} \quad (3-6)$$

在模拟视频系统中,采用的是匀速、单向的直线扫描方式。这里的“单向”具体指的是水平方向上自左向右以及垂直方向上自上向下进行扫描以传送图像信息;相反,水平自右向左和垂直自下向上的扫描则不进行图像信息的传送。电子束在屏幕光栅中的扫描线,如图 3-6 所示,屏幕显示的图像被分解为 5 行,这 5 行扫描下来共需 5 倍的行周期扫描时间,恰好就等于帧正程时间 T_{Ft} 。习惯上,将帧正程内扫描的行数称为有效行数(表示图像有效内容的行),用 Z' 表示;同时,将一个帧周期内从上到下和从下到上扫描的总行数,称为标称行数,用 Z 表示。图 3-6 所示的有效行数 $Z'=5$,标称行数 $Z=6$ 。

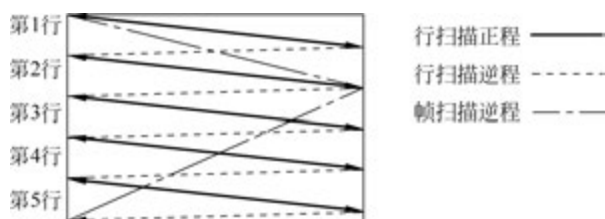


图 3-6 电子束在屏幕光栅中的扫描线

显然,帧正程时间 T_{Ft} 等于帧周期 T_F 减掉帧逆程时间 T_{Fr} ,因为行扫描具有固定的行周期,因此有效行数 Z' 的计算表达式推导如式(3-7)所示。

$$\begin{aligned} T_{Ft} &= T_F - T_{Fr} = T_F(1 - \beta_F) \\ Z' \cdot T_H &= Z \cdot T_H(1 - \beta_F) \\ Z' &= Z(1 - \beta_F) \end{aligned} \quad (3-7)$$

此外,行频等于帧频乘以标称行数,如式(3-8)所示。

$$f_H = Z \cdot f_F \quad (3-8)$$

为了实现逐行扫描中电子束匀速、单向的直线运动,早期模拟电视系统采用 CRT 结构的摄像管或显像管,会向其水平偏转线圈和垂直偏转线圈输入相应的水平、垂直锯齿波电流。这些电流会造成水平偏转线圈和垂直偏转线圈产生适当的磁场,促使电子束流进行精确的水平与垂直偏转,确保实现匀速且单向的直线扫描。水平扫描电流与垂直扫描电流的波形示意图,如图 3-7 所示,该图相关参数与图 3-6 相对应。当水平偏转线圈通入行扫描频率的锯齿波电流时,产生的磁场使得电子束沿水平方向进行来回偏转,这种锯齿波电流被称为水平扫描电流。同样地,当垂直偏转线圈通入帧扫描频率的锯齿波电流时,电子束将沿垂直方向进行来回偏转,该锯齿波电流则被称为垂直扫描电流。

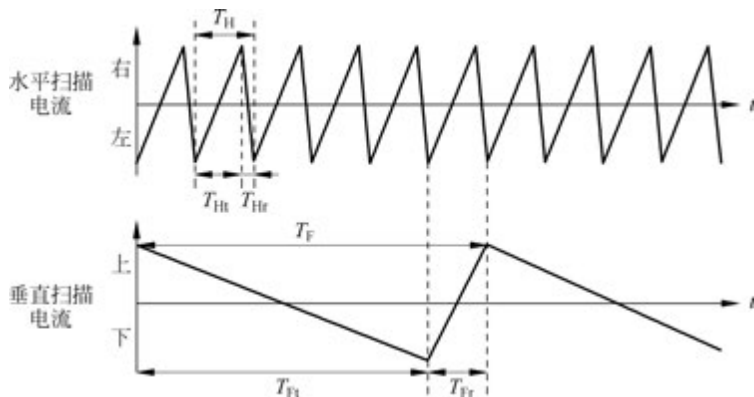


图 3-7 水平扫描电流与垂直扫描电流的波形示意图

扫描电流与电子枪发射的电子束流之间存在着严格的同步关系。以图 3-7 为例,在显像管中,当水平扫描电流值由最小逐渐增大至最大时,会产生一个均匀变化的磁场,这个磁场引导电子枪发射的电子束流从屏幕的最左端扫描到最右端,并在屏幕表面留下一条随图像信号变化的直线轨迹。这条轨迹是电子束流激打荧光屏所产生的辉光,此时正处于行正程阶段。相反地,当水平扫描电流从最大值逐渐减小到最小值时,电子束流则从屏幕的最右端扫描回最左端。不过,在这个过程中,电子枪会停止发射电子束流,因为此时正处于行逆程阶段,这样做是为了避免点亮屏幕从而干扰正程画面的正常显示。

类似的,如图 3-7 所示,当垂直扫描电流从最大值逐渐减小至最小值时,同样会产生一个均匀变化的磁场,促使电子束流从上至下扫描。因此,在帧正程阶段,电子束流在水平扫描的同时,还在垂直方向上从上往下逐行扫描,使得光像一行一行地在屏幕上依次显示出来。由于显示器的余晖效应和人眼的视觉暂留特性,这些逐行显示的光轨迹能够在人眼中拼凑成一个完整的图像。而在帧逆程期间,电子束流从下往上快速一行行扫描,但此时电子枪会全程停止发射电子束流,避免干扰帧正程图像的正常显示。

2. 逐行扫描参数

在设计逐行扫描的电视系统时,关键参数(如帧频、行频等)的确定至关重要。这些参数需确保收发端能够保持同步,并且在满足传输带宽要求的同时,还要兼顾人眼的视觉特性。这样做的根本目的,是力求让人眼所感知到的重现图像与原始图像尽可能一致。

(1) 帧频的选择。在设计逐行扫描电视系统时,帧频的选择对系统有重要的影响,需综合考虑多个因素。首先,为保证重现图像具有连续感,帧频需高于人眼的融合频率 20Hz,以避免观众感知到画面内容的顿挫和跳跃,使运动画面更加流畅。其次,帧频应满足人眼临界闪烁频率特性,即 45.8Hz 以上,以消除图像闪烁感。同时,还需考虑交流电源频率的干扰,选择与当地电源频率(如中国及欧洲为 50Hz,美国、加拿大等为 60Hz)一致的帧频,避免画面干扰现象。最后,帧频的选择还需兼顾传输带宽资源的限制,尽管提高帧频能提升画面还原度和视觉体验,但也会增加传输难度和设备性能要求。因此,帧频的选择要综合考虑上述四个因素。

传统胶片有声电影的帧频为 24Hz,已满足人眼融合频率要求,但通过采用 2 开口遮光器使画面显示频率倍频至 48Hz,以达到无闪烁效果。对于电视信号的显示,传统显示器刷新频率与视频信号播放帧频一致,而现代智能电视机则可以采用提高刷新频率的技术提供更好的显示效果。对于运动速度极快的竞技游戏,系统需提供更高的帧频,使相邻两帧之间的时间更短、主体相对位移更小,从而带来更加流畅的视觉体验。只有当相邻两帧图像上主体移动的距离相对于人眼的张角不超过 7.5 分时,观众才能感受到连续平滑的画面。因此,游戏视频中,主体移动速度较高,系统必须提供更高的帧频,才能使相邻两帧之间的时间更短、主体相对位移更小,从而让主体运动看起来更流畅。

(2) 行频的选择。因为行频等于帧频乘以标称行数,因此在帧频确定之后,行频的选择就取决于每帧的标称行数。

确定每帧标称行数需遵循以下原则:首先, Z 应为正整数,以使各帧图像的光栅能精确重合;其次,电视系统需具备适当的分解力(电视系统传送图像细节的能力,用图像水平或垂直方向上能分辨的像素数表示),以满足人眼的分辨力要求;最后,形成的图像信号带宽不应过高。

假设屏幕宽高为 $W \times H$,屏幕从上至下交替显示黑白条纹,每行黑条或白条的高度均

为一个像素高度 d 。人眼的分辨角与扫描行数的关系示意图如图 3-8 所示,假设像素高度 d 恰好满足人眼分辨力的要求。设最佳观看距离 L 为画面高度 H 的 6 倍,人眼的分辨角 θ 为 $1'$,则根据式(1-8)可计算出,画面分成 573 行,满足上述要求。由于这里计算得出的 573 行是有效行数,若帧逆程系数 β_F 为 0.08,则根据式(3-7)可求得标称行数约为 623 行。

在实际应用中,由于 CRT 显示器使用真空管分频电路从行频推导出场频,标称行数最好是小整数因子的乘积,这样电路设计更为简单。因此,标称行数不能随意设定,标清 525 线电视系统,标称行数 525 可分解为 $7 \times 5^2 \times 3$; 而我国标清电视采用 625 行的标称行数, $625 = 5^4$, 均满足小整数因子乘积的要求。

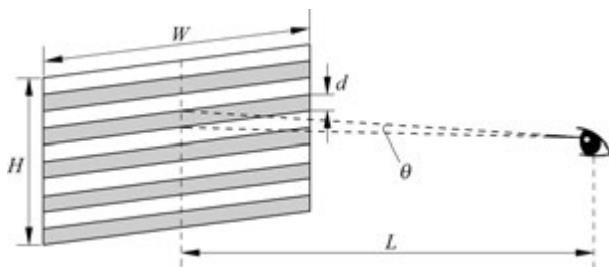


图 3-8 人眼的分辨角与扫描行数的关系示意图

3. 基带视频信号最高频率

由于模拟视频传输系统的带宽是有限的,视频信号的带宽会直接影响到视频传输系统能够传输的视频节目频道的数量。因此,有必要确定视频基带信号的最高频率,因其基本上代表了视频信号的带宽(视频信号最低频率接近 0)。

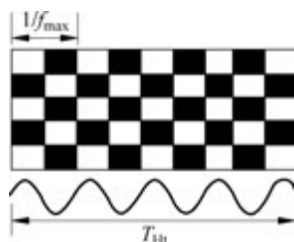


图 3-9 黑白“棋盘格”图像与行正程对应关系

当视频图像的像素呈现为黑白相间的“棋盘格”时,该图像达到细节最丰富的状态。黑白“棋盘格”图像与行正程对应关系如图 3-9 所示,图像只在行正程期间呈现(忽略帧逆程)。显然,此时的视频信号频率达到最大值 $f_{\max}$ 。由于存在一些限频因素,图像信号的频率不可能无穷大,因此图像信号的波形更接近于正弦信号,而非锐利的方波脉冲。视频图像信号对应的频率等于“黑白”两个相邻像素传输时间(一个正弦波周期)的倒数,即单位时间内能够容纳的正弦波数量。假设单个像素的传输时间为 T_1 ,则视频信号的最大频率 $f_{\max}$ 有

$$f_{\max} = \frac{1}{2T_1} \quad (3-9)$$

假设图像像素为方形,水平方向上的最大像素数为 $N_{\max}$,则 $T_1 = N_{\max} / T_{Ht}$,可得

$$f_{\max} = \frac{1}{2} \frac{N_{\max}}{T_{Ht}} \quad (3-10)$$

由于视频传输系统带宽的限制,原始视频信号经传输后会导致高频分量丢失,从而使得视频终端显示的图像的细节缺失,清晰度下降。因此,视频传输系统的带宽可以通过其传输视频图像细节的实际能力来衡量。图像实际呈现的水平方向上黑白相间的条纹数被定义为水平分解力,用 N 表示。用 N 替代式(3-10)中的 $N_{\max}$,即可求得图像所在视频传输系统的实际带宽 Δf ,即

$$\Delta f = \frac{1}{2} \frac{N}{T_{\text{Ht}}} \quad (3-11)$$

假设一视频系统, $T_{\text{H}} = 64\mu\text{s}$, $T_{\text{Hr}} = 8\mu\text{s}$, 宽高比为 16 : 9, 当传送的视频图像为 560 条黑白竖直相间的条纹时, 可计算该视频系统的带宽 Δf 为

$$\Delta f = \frac{1}{2} \frac{N}{T_{\text{Ht}}} = \frac{1}{2} \frac{560}{(64 - 8)\mu\text{s}} = 5\text{MHz}$$

若仍在该系统下, 当接收显示设备的通频带宽为 4MHz 时, 可得此时显示图像的实际水平分解力为

$$N = 2\Delta f T_{\text{Ht}} = 2 \times 4\text{MHz} \times (64 - 8)\mu\text{s} = 448(\text{线})$$

又假设一视频系统标称行数为 625, 逐行扫描, 帧频为 50Hz, 宽高比为 4 : 3, β_{F} 为 0.08, α 为 0.1875。假设图像为方形像素, 水平最大分解力与有效行数之比即为宽高比, 可得

$$N_{\text{max}} = \frac{4}{3} Z' = \frac{4}{3} Z(1 - \beta_{\text{F}}) \quad (3-12)$$

而行正程时间 T_{Ht} 为

$$T_{\text{Ht}} = T_{\text{H}}(1 - \alpha) = \frac{(1 - \alpha)}{f_{\text{H}}} = \frac{(1 - \alpha)}{Z \cdot f_{\text{F}}} \quad (3-13)$$

将式 (3-12) 和式 (3-13) 代入式 (3-10), 可求得该逐行视频系统的最大频率为 14.74MHz。

逐行扫描对于早期的模拟电视系统来说, 所需带宽相对较大。为了节约频谱资源, 降低设备成本和复杂度, 模拟电视系统采用了隔行扫描技术。隔行扫描技术的引入, 有效地平衡了图像质量与传输带宽之间的需求。

3.1.3 视频的隔行扫描

1. 隔行扫描的基本原理

隔行扫描是按照隔行顺序扫描, 将一帧图像分为两场子图像, 先从上至下扫描奇数行形成一场子图像, 称为奇数场(简称奇场); 再从上至下扫描偶数行形成另一幅子图像, 称为偶数场(简称偶场)。隔行扫描示意图如图 3-10 所示。

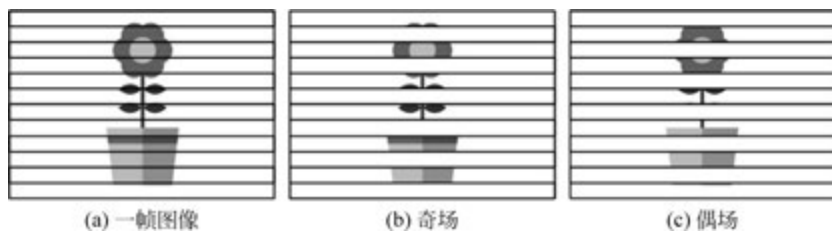


图 3-10 隔行扫描示意图

由于隔行扫描视频系统中, 一帧图像分为两场, 因此场频 f_{V} 是帧频 f_{F} 的 2 倍。例如, 我国采用的标清和高清电视系统均为隔行扫描, 场频为 50Hz, 帧频为 25Hz。根据式 (3-10), 隔行扫描方式的带宽只需逐行扫描系统带宽的一半。同时, 50Hz 的场频, 也能基本满足人眼视觉特性中对融合频率和临界闪烁频率的要求。因此, 隔行扫描的优点是在保证图像分解力不甚下降和画面无大面积闪烁的前提下, 使图像信号的带宽减小一半。

相应的,在隔行扫描的一场中,隔一行扫描一行,电子束从上到下的全过程被称为场扫描正程(简称场正程),时长用 T_{Vt} 表示;电子束从下到上的全过程被称为场扫描逆程(简称场逆程),时长用 T_{Vr} 表示。场周期和场频分别满足式(3-14)和式(3-15)。

$$T_V = T_{Vt} + T_{Vr} \quad (3-14)$$

$$f_V = 1/T_V \quad (3-15)$$

无论是逐行扫描还是隔行扫描,统一将水平和垂直方向上正程期间扫描的像素称为有效像素。我国模拟电视系统采用 625 行隔行扫描制,帧频设定为 25Hz,场频为 50Hz。根据式(3-8),行频为 15 625Hz。根据式(3-5),帧周期为 40ms,而场周期为帧周期的一半,为 20ms。场逆程系数(β_V)为 0.08,根据式(3-5)(该公式在隔行扫描中写为 $\beta_V = T_{Vr}/T_V$),场逆程时间为 1.612ms。根据式(3-2),行周期为 64 μ s。行逆程系数(α)为 0.1875,行逆程为 12 μ s。

2. 隔行扫描的主要缺陷

隔行扫描技术早期在电视系统中广泛应用,但也存在显著缺点:一是行间闪烁,当图像有细密水平条纹时,亮暗线分布在奇偶行,导致亮线每秒仅出现 25 次,低于临界闪烁频率,观众会感到明显闪烁;二是并行现象,包括因行场扫描频率不匹配导致的真实并行,使图像垂直清晰度减半,以及物体垂直运动速度与场频同步时产生的视在并行,造成图像清晰度下降;三是边沿锯齿现象,物体水平高速运动时,因隔行传送时间差,会导致图像垂直边缘出现锯齿。尽管有这些缺陷,但受限于早期成本、带宽和技术,隔行扫描仍被全球广播电视系统采用。然而,随着数字时代和视频技术的进步,逐行扫描已成为主流,特别是在流媒体和 4K、8K 超高清标准中,隔行扫描技术已被摒弃。

3. 图像显示格式及扫描格式表示

数字电视画面的图像显示格式通常指的是图像在水平方向和垂直方向上的有效像素数。在我国,标准清晰度数字电视(Standard Definition Television, SDTV)的图像显示格式为 720×576 像素,即每帧图像在水平和垂直方向上的有效像素数分别为 720 和 576。SDTV 采用隔行扫描方式,通常会标注场频用 $720 \times 576/50i$ 表示,其中“i”代表隔行(interlaced)扫描,场频为 50Hz。

我国高清晰度电视(High Definition Television, HDTV)的图像显示格式为 1920×1080 像素,同样采用隔行扫描方式,标注场频后用 $1920 \times 1080/50i$ 表示,场频为 50Hz。

我国超高清电视(Ultra High Definition Television, UHD TV)按分辨率则分为 4K HDTV 和 8K UHD TV 两种,均为逐行扫描,且帧频有 50Hz、100Hz、120Hz 三种。举例来说,UHD TV 的图像显示格式,标注帧频后可以用 $3840 \times 2160/50p$ 、 $7680 \times 4320/100p$ 表示,其中“p”表示逐行(progressive)扫描,帧频分别为 50Hz 和 100Hz。

3.1.4 模拟黑白全电视信号

数字电视系统受到了模拟电视制式和参数的影响,为了更好地理解数字电视系统的参数,有必要了解模拟电视信号的组成及相关参数。在模拟黑白电视系统中,传输的是所谓的黑白全电视信号,该信号由三部分组成:图像信号、复合消隐信号以及复合同步信号。

1. 图像信号

图像信号是摄像机通过光电转换和扫描把活动景物变成随时间变化的电信号,仅携带

了景物的明、暗(白、黑)信息。摄像机拍摄的图像及其对应的黑白全电视信号如图 3-11 所示,该图展示了摄像机拍摄的 4 条灰阶图像及其对应的黑白全电视信号的一行波形。根据我国标清模拟电视的标准,黑白全电视信号的行周期为 $64\mu\text{s}$,其中 $52\mu\text{s}$ (行正程)是图像信号,剩余的 $12\mu\text{s}$ 用于传输行消隐信号,行消隐信号中还嵌入了行同步信号。该图像中,亮度越高的区域对应的信号电平也越高。在白色和灰色衔接的地方,由于亮度发生剧烈变化,相应的图像信号电平也会出现突然改变,呈现出脉冲特性。这种图像信号中的上下跳变部分,与信号频谱中的高频成分相对应,反映了图像内容中的细节部分。

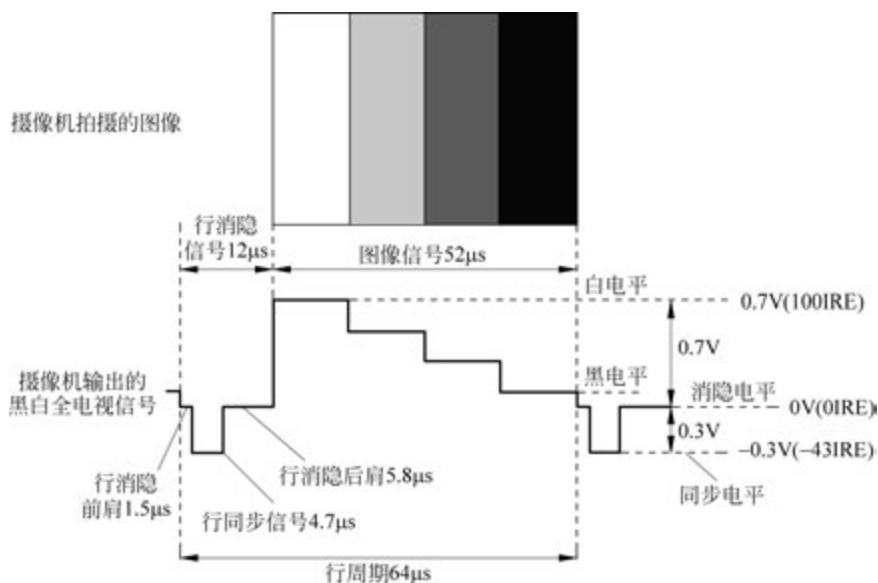


图 3-11 摄像机拍摄的图像及其对应的黑白全电视信号

在图像信号的电平分配方面,景物中最亮部分的图像信号电平称为白电平,幅值被设定为 0.7V ,代表了黑白全电视信号的峰值电平。相对的,景物中最黑部分的图像信号电平被称为黑电平,通常比 0V 稍高一点。图 3-11 所示的这种白电平高、黑电平低的图像信号称为正极性图像信号。在摄像机、录放机等视频设备的接口,规定使用峰-峰电平值为 $0.7\text{V}_{\text{P-P}}$ 的正极性信号。相反,黑电平高、白电平低的图像信号被称为负极性图像信号,多用于发射机。这是因为电视画面内容通常比较明亮,使用电平较低的负极性图像信号有助于节省发射功率。为了方便表达和计算,视频信号经常使用无线电工程学会单位(Institute of Radio Engineers, IRE)来标注视频信号的幅度, $0.7\text{V}_{\text{P-P}}$ 的范围被定义成 $0\sim 100\text{IRE}$,即从视频黑电平到视频白电平之间的范围被平分 100 份,是一种相对值的表示。

2. 复合消隐信号

复合消隐信号由行、场消隐信号构成,分别标识了电子束水平回扫和垂直回扫的逆程时间,以便停止发射电子束。图 3-11 中图像信号左侧的部分就是行消隐信号,其所处的电平被称为消隐电平。消隐电平被设定在 0V ,这样能使电子束截止。对于正极性信号而言,消隐电平比黑电平稍低,也就是说,它比黑电平“更黑”。消隐电平和黑电平的差被称为消隐电平提升,通常被设置为 $0\sim 50\text{mV}$ 。场消隐信号,处于奇数场和偶数场交替的位置,电平也设定在 0V 。我国标清模拟电视的标准中,行周期为 $64\mu\text{s}$,行消隐信号的脉冲宽度为 $12\mu\text{s}$,

根据式(3-3)可计算出行逆程系数 α 为 0.1875; 场周期为 20ms, 场消隐信号的脉冲宽度为 $25T_H + 12\mu s = 1612\mu s$, 场逆程系数 β 为 0.08。复合消隐信号的周期为 40ms, 其频率与帧频一致。

3. 复合同步信号

复合同步信号由行、场同步信号构成, 分别指电子束行、场扫描逆程的开始时刻, 控制电子束扫描的行、场同步。在正极性的黑白全电视信号中, 图像信号占据正值电平, 消隐信号位于 0 电平, 行、场同步电平则被设置为 $-0.3V$ (即 -43IRE)。为避免同步信号影响到图像信号, 负电平的行、场同步信号被分别叠加在行、场消隐信号上。我国模拟标清黑白电视信号的行同步信号宽度为 $4.7\mu s$, 场同步信号宽度为 $160\mu s (2.5T_H)$ 。为了在场同步期间不丢失行同步, 以及考虑到奇场和偶场场同步信号的一致性, 使得复合同步信号的波形较为复杂。

即使是在数字高清演播室, 目前仍然会采用图像信号幅度为 0 的模拟黑场 (Black Burst, BB) 信号作为外同步基准信号, 同步锁定系统中的所有设备, 可见模拟视频信号仍有用武之地, 存在于数字, 尤其是数模混合视频系统当中。

3.2 图像传感器

摄像机使用图像传感器, 将光信号转换为电信号, 也称为摄像机器件或感光器件。其质量与体积直接决定摄像机的整体性能、尺寸及调节方式, 是摄像机内部最为关键的组件。根据使用感光器件的数量, 摄像机可分为多片式与单片式。例如, 彩色电视摄像机通常采用三片感光器件, 即三片式摄像机; 而数字电影摄影机则多采用单片感光器件, 即单片式摄像机。

图像传感器主要有电荷耦合器件 (Charge Coupled Device, CCD) 和互补金属氧化物半导体 (Complementary Metal-Oxide-Semiconductor, CMOS) 两种类型。目前 CCD 已经逐步被 CMOS 替代, 鉴于 CCD 的基本单元结构与 CMOS 一致, 本节首先介绍 CCD。

3.2.1 CCD

1. CCD 的结构与工作原理

CCD 是一种利用电荷量表示光强大小, 并通过耦合方式传输电荷信号的传感器元件。其核心构成单元是 MOS 电容器, MOS 结构示意图如图 3-12 所示, 它能够完成光电转换、电荷存储以及电荷传输的功能。

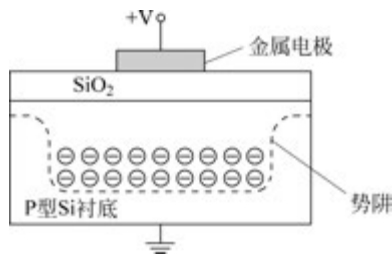


图 3-12 MOS 结构示意图

CCD 芯片中的 MOS 电容器通常采用三层结构。最下层为 P 型 (或 N 型) 半导体作为衬底; 在其上生长一层厚度为 $10 \sim 150\text{nm}$ 的 SiO_2 作为绝缘层; 在 SiO_2 层上面蒸镀一层金属, 作为电极 (连接栅极)。

硅具有四价共价键结构。当掺杂如砷、磷等五价元素时, 这些元素会取代硅原子的位置, 并因此多出一个自由电子, 从而增强材料的导电性, 从而形成 N 型 (Negative, 负) 半导体硅。相反, 如果掺杂的是硼等三

价元素,它们同样占据硅原子的位置,但此时会缺少一个电子,从而形成空穴。在适当的光、热激发条件下,其他电子可以被激发并填补这些空穴,形成连续的电子填补和空穴传导过程,这种结构就是P型(Positive,正)半导体硅。

在CCD中,每个电极分别对应一个CCD单元。当在金属电极和最下层硅片底部的欧姆接触点(金属与半导体之间形成的一种特殊接触界面,阻抗接近于零)之间施加一个偏置电压 V (金属电极为正电压)时,会在电极下面的衬底内形成耗尽层,即一个没有空穴的、带负电荷的区域。栅极电压的大小会直接影响耗尽层的厚度,电压越高,耗尽层就越厚。当耗尽层达到一定深度时,就能形成一个能够捕获电子的陷阱,称为电子势阱(简称势阱)。

当光子投射到MOS电容器上时,光子会穿过透明的氧化层,进入P型Si衬底,衬底中处于价带的电子将吸收光子的能量而跃入导带,形成电子-空穴对。这些电子-空穴对在外加电场的作用下,分别向电极两端移动,形成光生电荷。这些电荷将储存在由电极造成的势阱中,形成电荷包,从而实现光电转换和电荷的存储。势阱的深浅与电压大小有关,电压越高,势阱越深。电荷的电量与摄入CCD单元的进光量有关,进光量由光照时间和光照强度共同决定。

在CCD摄像器件感光面中,有序排列着百万、千万个电极(相当于像素),每个电极下都有一个相同深度的电子势阱。当景物成像在感光面时,光图像对应到每个CCD像元的进光量,产生了与进光量相对应的电荷包,这样,一幅光图像就转变成了对应的电荷图像。

2. CCD的光电转换

灵敏度是衡量图像传感器对光信号响应能力的一个关键指标,它定义为在一定光谱范围内,单位曝光量下传感器输出的信号电压,单位为 $V/\text{lux} \cdot \text{s}$ 。这一参数综合体现了图像传感器的光电转换能力,是评价图像传感器性能的重要标准。

为了提升CCD图像传感器的灵敏度,制造商通常会采取提高CCD开口率的策略。开口率是指CCD器件中,一个像素能够感光的面积占单位像素平均面积的比例。开口率越大,意味着感光面积越大,器件灵敏度也越高。

前照式CCD图像传感器如图3-13所示。CCD上存在金属电极,甚至可能覆盖有遮光膜,由于电极等遮光物的遮挡,光线只能由电极缝隙射入,导致CCD的实际感光面积受限。为了增大CCD的开口率,让更多的光线能够射入CCD,制造商通常会在CCD上方安装一层微透镜阵列(也被称为片上微透镜)。这层微透镜阵列是在器件芯片的常规制作工序完成后,利用光刻技术在每个像元的表面直接制作而成的微型光学透镜,它能够对入射光进行会聚,从而有效扩大每个像素的受光面积。

为了克服电极遮挡的问题,图像传感器也有被设计成背照式。在背照式设计中,光线从传感器的背面通过,这样就能够避免正面照射时多晶硅电极及一些不必要的吸收,使更多的光子进入。不过,背照式设计要求衬底材料非常薄,以确保在可见光范围内都能保持良好的光电响应特性。

另外,只要势阱足够深,MOS电容器的曝光时间越长,入射的光子就越多,从而转换出来的电荷也就越多。通常,曝光时间会被设置为略小于帧周期(40ms)或场周期(20ms)。

3. 电荷的存储

光电转换过程产生的电荷包被存储在CCD的基本单元——MOS电容器的势阱(耗尽层)中。MOS电容器存储电荷的能力是有限的,一旦电荷量超出势阱的深度,就会发生电荷

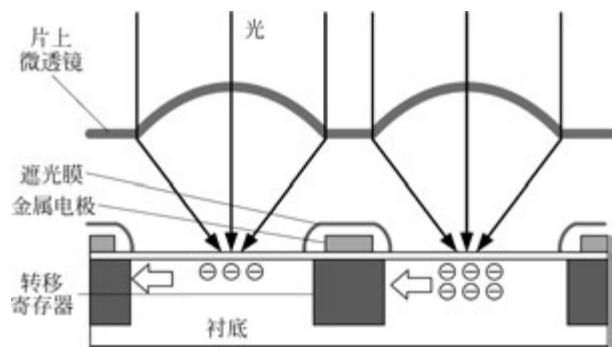


图 3-13 前照式 CCD 图像传感器

溢出,导致图像中出现“高亮点开花或拖彗尾”现象。这一现象会影响 CCD 图像传感器的动态范围。动态范围定义为信号容量与噪声之比,其值越大,意味着 CCD 芯片在采集对比度差异显著的场景时表现越出色。

4. 电荷的转移

在场正程期间,CCD 需要完成光电转换和电荷包存储的任务。当一场电荷像生成后,必须在场逆程期间将所有电荷包在微秒级的时间内迅速转移出去。CCD 转移电荷包的方式可以通过控制相邻 MOS 电容器的栅极电压,从而改变相邻势阱的深度,引导电荷从浅势阱向深势阱移动。CCD 电荷转移示意图如图 3-14 所示,当施加高电位时,其电极下方的势阱变深,而 0 电位电极下方则无法存储电荷。

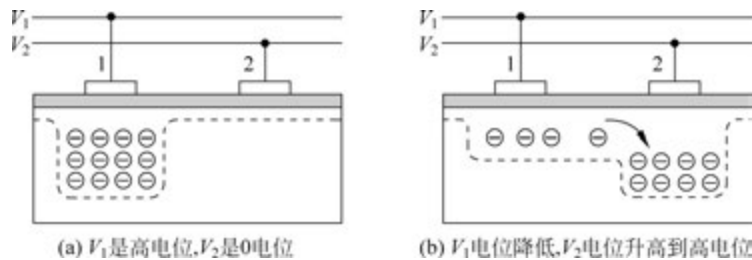


图 3-14 CCD 电荷转移示意图

在曝光时间内,MOS 电容器生成并积累了一定的电荷包于较深的势阱中。通过调整相邻栅极电压,可以改变相邻 CCD 单元势阱深度,使得电荷被“挤”到变深的势阱中,实现电荷的转移。注意,电荷转移的过程是一个模拟信号处理过程,而非简单的有无电荷的“0”或“1”数字概念。

在实际应用中,CCD 记录的电图像是二维的,最终需要将所有电荷包进行水平和垂直方向的转移。针对不同形式的电荷转移需求,主要有二相、三相、四相时钟驱动三种方式。

5. CCD 图像传感器的主要类型

CCD 图像传感器主要分为帧间转移式(Frame Transfer, FT)、行间转移式(Interline Transfer, IT)以及帧行间转移式(Frame-Interline Transfer, FIT)三种基本类型。这些类型在电荷移动方式和器件结构上存在差异,但共同点是垂直移位寄存器均采用四相时钟驱动,而水平移位寄存器则采用二相时钟驱动。为防止图像过亮导致的“开花”现象,这些 CCD 的各感光单元一侧通常安装有溢流控制门和溢流阱(或称为溢流沟道),用以吸收势阱中溢出

的电子。

6. CCD 曝光时间的控制

在拍摄高速运动景物时,为了获取更清晰的图像,通常需要减少曝光时间。对于采用 CCD 图像传感器的摄像机而言,可通过两种主要技术来实现曝光时间的减少:机械快门和电子快门。

(1) 机械快门。机械快门是一种物理式的开关装置,通过驱动机械叶片遮挡或不遮挡感光单元来调节曝光时间。在摄像机(包括电影摄影机)中,机械快门通常是一种可旋转的叶子板,每次曝光时旋转一圈(360°),通过控制叶子板的开角大小来改变曝光时间。单反相机则通过齿轮、凸轮、弹簧等装置,或者电磁手段来控制快门叶片的开启和闭合。然而,机械快门存在占用空间体积大、内部结构复杂、调光速度缓慢等缺点,且可能产生噪声或震动,影响拍摄效果。

(2) 电子快门。电子快门是一种通过控制图像传感器的像素电荷积累来实现曝光控制的技术。对于场频为 50Hz 的摄像机,其 CCD 图像传感器在不采用电子快门时,电荷积累时间约为 $1/50\text{s}$ 。为了降低曝光时间,CCD 图像传感器将这 $1/50\text{s}$ 分为两段:无效电荷积累时间(t_1)和有效电荷积累时间(t_2)。在 t_1 时间段内积累的电荷被释放,不形成信号输出;而在 t_2 时间段内积累的电荷则形成信号被输出。

常见的电子快门速度挡有关闭、 $1/50\text{s}$ 、 $1/100\text{s}$ 、 $1/250\text{s}$ 、 $1/500\text{s}$ 、 $1/1000\text{s}$ 、 $1/2000\text{s}$ 等。有时,为了拍摄夜景等长时间曝光的场景,快门时间可能设置在 1s 以上。需要指出的是,在使用高速电子快门时,由于电荷的有效积累时间短,电荷包内电荷的数量少,输出图像信号的幅度小,信噪比下降,图像噪点多。为了保证输出视频信号有足够的信噪比,电子快门需在高照度下才适宜应用,快门时间越短,需要景物的照度越高。

7. CCD 图像传感器的特性

CCD 图像传感器作为一项成熟的技术,曾在数字视频领域占据重要地位。CCD 摄像机的优点包括长寿命、简便的维护流程、高灵敏度、低惰性(有效避免高照度下的拖尾和开花现象)、耐灼伤特性、无画面几何畸变、良好的抗震性能、不受磁场干扰影响,以及可配备电子快门等功能。相较于目前广泛应用的 CMOS 图像传感器,CCD 图像传感器在像元结构上更为简单,拥有更高的开口率,且其不均匀噪声更低。然而,CCD 技术也存在一些固有的局限性。其信号读出方式相对固定,导致读取速度受限,同时功耗和所需电压较高。这使得 CCD 技术在面对 4K、8K 等超高清传感器需求时,难以提供稳定的支持。此外,CCD 的集成度较低,摄像机制造过程中还需增设大量的驱动电路、控制电路、模数转换等辅助电路,这无疑增加了制造的复杂性和成本。

3.2.2 CMOS

CMOS 图像传感器(CMOS Image Sensor,CIS)作为现代数字成像技术的核心组件,通常由光电器件(如光电二极管或光电晶体管)、输出电路及信号处理电路构成。在 CMOS 图像传感器设计中,光电二极管被广泛用作光电器件,其工作原理是基于 P-N 结的光电效应。

当光线入射到光电二极管时,光子的能量会传递给二极管的硅原子。若这一能量超过硅的带隙能量,价带中的电子将获得足够的能量跃迁到导带,进而形成电子-空穴对。由于光电二极管耗尽区内存在强内建电场,当电子-空穴对在这一区域产生时,它们将在电场力

的作用下分别向 PN 结的 N 端和 P 端漂移,这一过程将产生光电流。值得注意的是,光信号越强,光电二极管产生的光电流越大,因此,光电流的信息直接反映了入射光的强度信息。与 CCD 的光电转换原理相似,CMOS 同样利用了硅的光电效应。然而,两者在数据传输方式上存在显著差异,这也是 CMOS 图像传感器得以广泛应用的重要原因之一。

1. CMOS 图像传感器的结构

CMOS 图像传感器结构示意图如图 3-15 所示。其成像区由像素阵列构成,负责接收通过摄像机镜头投射进来的光图像。在场正程(帧正程)期间,每个像素单元进行光电转换并积累电荷。随后,在场消隐(帧消隐)期间,行选择逻辑电路会按照从上到下的顺序,依次选中某一行施加特定电压,以控制该行像素的信号输出。这一选择过程会根据逐行扫描或隔行扫描的需要来确定要输出的行。同时,列选择单元也会对从左到右相应位置的像元施加特定电压,实现精确控制每个像素信号的输出。像素信号的读出过程受到定时与控制单元的管理。通常,每个像素信号会通过其所在列的信号总线(简称列总线)传输到信号处理单元(如模拟信号处理单元、模数转换单元等),进行电荷积分、增益控制、采样与保持、双相关采样与固态噪声抑制、数字化等处理。具体的结构和信号处理流程会根据 CMOS 芯片的种类不同而有所不同。

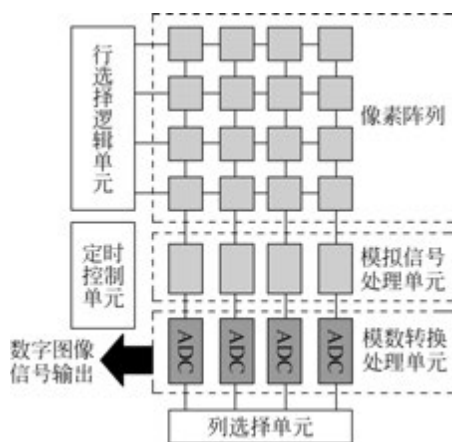


图 3-15 CMOS 图像传感器结构示意图

从这里可以看出,CMOS 的电荷传输更加灵活,它采用逐行传输的方式,能够实现像素级的寻址和控制。这与 CCD 的工作方式有所不同,CCD 是在驱动电路的控制下,整个像面上的所有像元同时进行电荷的水平或垂直移位转移。

与 CCD 类似,为了提高灵敏度,CMOS 图像传感器也采用了片上微透镜技术来增大开口率,并有针对性 CMOS 的背照式工业设计。对于单片式摄像机,还需要在传感器上增加彩色滤镜矩阵。前照式与背照式 CMOS 图像传感器结构对比示意图,如图 3-16 所示。在前照式 CMOS 图像传感器结构中,光线首先通过片上微透镜和滤色片,然后穿过布线层,最终达到感光层。由于布线层的遮挡,光线进入感光层的量有所减少。而背照式 CMOS 图像传感器则将感光层置于布线层之上,光线通过片上微透镜和滤色片后直接到达感光层,这样便提高了开口率。

2. CMOS 图像传感器的发展与分类

CMOS 图像传感器的发展经历了三个主要阶段:无源像素图像传感器(Passive Pixel Sensor,PPS)、有源像素图像传感器(Active Pixel Sensor,APS)和数字像素图像传感器(Digital Pixel Sensor,DPS)。

(1) CMOS-PPS 无源像素图像传感器。CMOS-PPS 是 CMOS 图像传感器的一种早期形式,出现时间早于 CCD。CMOS-PPS 结构示意图如图 3-17 所示,每个像元包含一个 PN 结光电二极管和一个 MOS 开关管。在场正程曝光期间,光电二极管进行光电转换并积累电荷,而开关管处于关断状态。在场消隐期间,通过行选择和列选择线路控制,特定一行的某个像元的开关管导通,将光电二极管存储的电荷传送至列总线,由末端的积分放大器转换

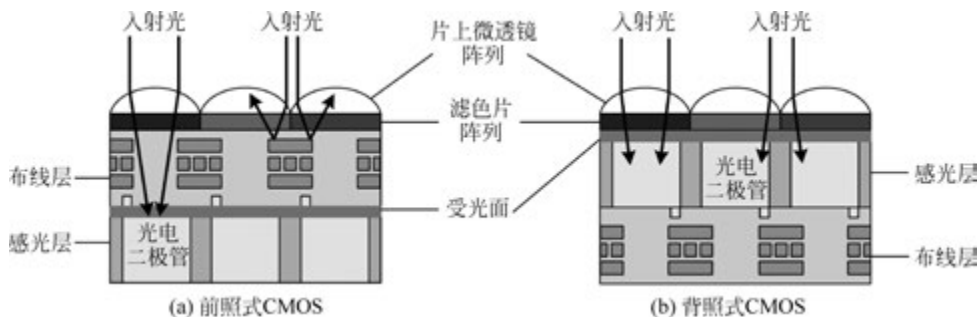


图 3-16 前照式与背照式 CMOS 图像传感器结构对比示意图

为电压信号输出。尽管 CMOS-PPS 具有结构简单、开口率高的优点,但由于缺乏内部有源器件,信号传输速度慢且结构噪声大,限制了高密度像素阵列的发展。

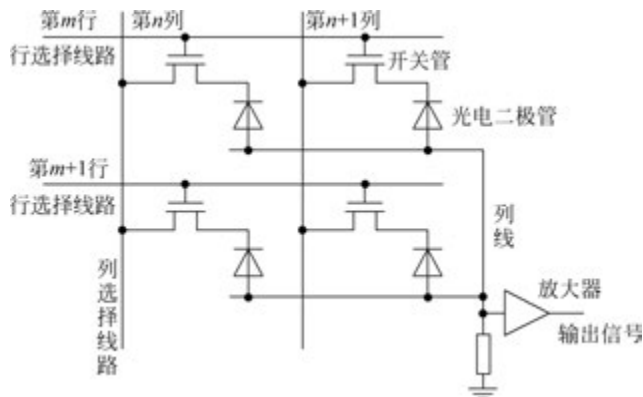


图 3-17 CMOS-PPS 结构示意图

(2) CMOS-APS 有源像素图像传感器。CMOS-APS 像元基本结构示意图如图 3-18 所示,每个像元内部都包含了多个晶体管(图中采用了 3 个晶体管,属于 3T 结构),典型的有光电二极管、源极跟随器、开关管(选通管)和复位管等。源极跟随器的出现增加了 CMOS 的驱动能力,它与列总线的负载电流源组成放大器结构,这样列总线电容的大小就不会对像素内的信号产生影响。因为这个结构,像素面阵的大小不再影响传感器的光电性能,噪声更小,读出信号速度更快。

在场正程启动前,复位管先对光电二极管进行复位,清除其残留电荷。接着,在场正程期间,光电二极管负责光电转换并存储电荷。进入场消隐期,电荷通过源极跟随器缓冲后输出,此时行选通管导通,使电荷得以通过列总线传输。随后,行选通管关闭,复位管再次复位光电二极管,为下一场正程的光电转换和电荷积累做准备。

在实际应用中,CMOS-APS 采用双相关采样电路来消除固态噪声。固态噪声是由传感器内部缺陷导致的空间分布噪声,具有不随时间变化的特性,会在所有帧中形成稳定的噪声图样。双相关采样电路通过计算电荷积累后像元输出的电压与曝光前该像元输出的电压之差,有效去除像元的固态噪声。

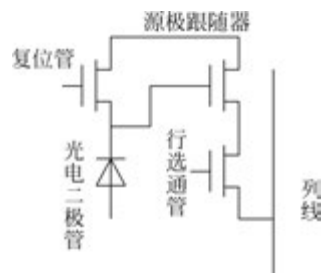


图 3-18 CMOS-APS 像元基本结构示意图

因此,CMOS-APS 不仅适用于大规模集成电路开发,还能提供比 CMOS-PPS 质量更高的图像信号。此外,由于 CMOS-APS 中的有源放大器仅在读出时工作,其功耗相较于 CCD 显著降低。然而,由于像元中晶体管数量的增加,CMOS-APS 的开口率相较于 CMOS-PPS 有所下降,典型值为 20%~30%,与 IT CCD 相当。为了提升开口率,这类 CMOS 芯片通常会配置片上微透镜阵列,可将开口率提高 2~3 倍。由此可见,CMOS-APS 的结构更为复杂,对制造工艺的要求也更高。

(3) CMOS-DPS 数字图像传感器。目前图像传感器普遍集成了模数转换器(ADC),这一趋势得益于数字技术的成熟发展,使得数字接口成为设备与系统间连接的主流。采用数字图像传感器制造摄像机,不仅能显著简化内部结构,还有助于设备的小型化。

在 CMOS 芯片上设计 ADC 主要有三种模式:芯片级 ADC、像素级 ADC 和列级 ADC。

① 芯片级 ADC 对整个 CMOS 芯片像素矩阵的信号电压进行统一数字化处理,其特点是功耗高、复杂度低、噪声水平低,但处理速度受限于单一 ADC,相对较慢。不过,其开口率较大,是一个显著优势。芯片级 ADC 拓扑示意图如图 3-19 所示。

② 像素级 ADC 则为每个像元配备一个 ADC,在像元内部完成电荷积累、放大和模数转换等工作。由于集成了更多器件,这类 ADC 开口率低、复杂度高、功耗低,但噪声水平增加,需配合噪声消除电路使用。其处理速度非常高,因为每个像素并行完成模数转换。像素级 ADC 拓扑示意图如图 3-20 所示。

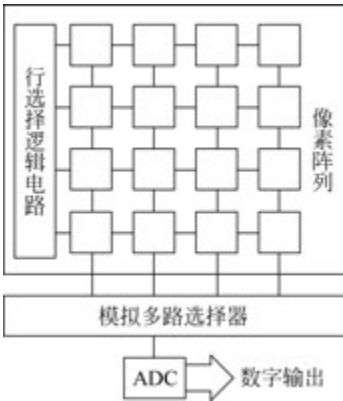


图 3-19 芯片级 ADC 拓扑示意图

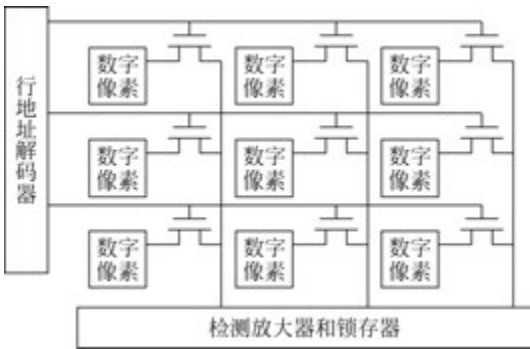


图 3-20 像素级 ADC 拓扑示意图

③ 列级 ADC 则在每列列总线末端独立安装一个 ADC,对该列信号进行数字化处理,如图 3-15 所示。它是芯片级 ADC 和像素级 ADC 在速度、功耗及复杂度上的折中选择。大部分图像传感器采用列级 ADC 结构,既保证了较快的处理速度,又在一定程度上控制了功耗和复杂度。

CMOS 图像传感器中 ADC 的特点对比,详见表 3-1。

表 3-1 CMOS 图像传感器中 ADC 的特点对比

	开口率	功耗	噪声	速度	复杂度
片级 ADC	高	高	低	低	低
列级 ADC	中	中	中	中	中
像素级 ADC	低	低	高	高	高

3. CMOS 快门模式

(1) 卷帘式快门。从 CMOS 的工作原理中知道,像素信息是在行选择逻辑的控制下逐行输出的,当某一行的信号正在被输出时,其下方的行仍在继续执行光电转换任务。为了保证每行像素单元都能获得相同的曝光时长,CMOS 图像传感器采用了从顶端到底端的逐行复位机制。在这种机制下,每一行复位后立即开始光电转换和电荷积累,经过一段固定的曝光时间(如行正程时间)后,电荷按行读出,这一过程同样是逐行进行的。这种工作模式被称为“卷帘式快门”,或者称为“滚动式快门”。卷帘式快门模式下,每帧(或每场)图像的曝光起始和结束时刻会沿着垂直方向逐行发生变化。当拍摄对象或摄像机镜头快速移动时,这种逐行曝光的机制可能会引发图像变形,该现象被称为“果冻效应”。

CMOS 果冻效应示意图如图 3-21 所示。以画面中存在快速水平右移的垂直条纹为例,由于不同行的曝光时刻条纹位置各异,最终输出的图像中,垂直条纹会呈现为上端向左倾斜、下端向右倾斜的形态。具体来说,在初始的 t_1 时刻,条纹位于图像的左侧,此时传感器的最上方行开始曝光。经过短暂的时间间隔到达 t_2 时刻,垂直条纹已移动到画面中间,而中间行正好进行曝光。进一步地,在 t_3 时刻,条纹继续移动到画面偏右侧,此时下方行开始曝光。因此,由于逐行曝光,原本垂直的条纹在输出图像中呈现出倾斜形态,如图 3-21(b) 所示。

类似地,当拍摄对象是一个从上到下快速运动的小球时,也会发生类似的变形,如图 3-21(c) 所示。在 t_1 时刻,小球位于图像上方,且最上方所在行正好曝光。随后,在 t_2 时刻,小球移动到画面中间,中间行进行曝光。最后,在 t_3 时刻,小球到达图像下方,此时下方

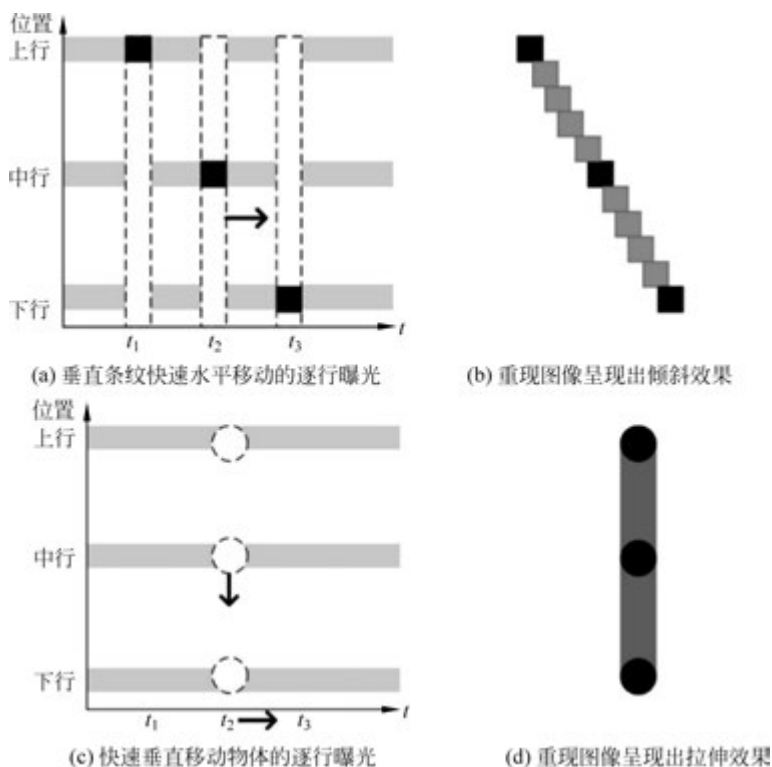


图 3-21 CMOS 果冻效应示意图

行开始曝光。由于每行曝光时刻的独立性,最终输出的图像中,小球的出现范围覆盖了从上到下的多个位置,给人一种被拉长的视觉效果,如图 3-21(d)所示。

(2) 全局式快门。在高速摄影领域,如捕捉子弹发射或闪电瞬间等场景,卷帘式快门的 CMOS 器件可能因严重的失真而不适用于。此时,另一种类型的快门模式——“全局式快门”(Global Shutter)或“全域式快门”显得尤为重要。全局式快门的 CMOS 器件特别之处在于其集成了存储区。在拍摄开始前,整个像素阵列会统一复位,确保所有像元能够同时开始并结束曝光。曝光结束后,数据会在场消隐期间被统一送入存储区。随后,在下一场的场正程期间,信号再从存储区按照特定顺序逐行输出。这种工作机制使得最终图像的效果与 CCD 图像传感器相似,有效避免了果冻效应。

实现全局式快门需要在 CMOS 图像传感器上增设存储区域,以此实现电荷积累与读取过程的独立进行,从而有效防止果冻效应的发生。此外,电荷积累时间可以根据拍摄需求进行灵活调整,而不会影响到信号的读取。然而,电路元件的增加会降低 CMOS 图像传感器的开口率,增加噪声,并对制造工艺提出更高要求,从而大幅增加成本。

4. CMOS 图像传感器的特性

CMOS 图像传感器与 CCD 图像传感器在感光并产生信号的基本流程上具有相似性,但它们在数据传送方式上存在着显著的差异。CMOS 图像传感器的读出方式更为灵活多变,易于控制,且读取速度快,提供了多样化的读取模式。与 CCD 图像传感器相比,CMOS 图像传感器通常具有更简单且更低的工作电压,因此其功耗也相对较低。

此外,CMOS 图像传感器基于 CMOS LSI 的制造工艺,与现有的 CMOS 集成电路技术在基础工艺和制造设备上具有高度的兼容性,这使其集成度更高,成本更低,非常适合用于制造 4K、8K 等超高清感光器件。配合数字像素传感器(Digital Pixel Sensor,DPS)技术,CMOS 图像传感器还能够实现 HDR 成像。例如,通过采用不同的曝光方式来采集同一场景的多幅图像,并在片外对这些图像进行处理,最终将它们重新组合成一幅具有高动态范围的图像。

尽管 CMOS 图像传感器在动态范围和功耗方面仍有进一步改进的空间,但其在技术上的诸多优势已经使其在诸多应用中取代了 CCD 图像传感器。然而,值得注意的是,CMOS 图像传感器在每个像元中内置了相对复杂的电荷信号读出电路,这可能会引入更多的噪声。同时,在工艺和像元尺寸相近的情况下,CMOS 传感器的开口率相较于 CCD 会更低,这会对灵敏度产生一定的影响。

3.3 三基色信号的产生

3.3.1 彩色摄像机的基本构成

彩色摄像机系统结构示意图如图 3-22 所示,主要由光学镜头、滤色片、分光系统、图像传感器、视频处理放大器、编码器等部分组成。光学镜头将景物光线汇聚,随后,该光线通过分光系统被分解成红、绿、蓝三种基色光。三种色光进而在图像传感器上进行光电转换,变为三种基色电信号。为了让图像色彩更加逼真,亮度层次更加合理,或达到一定的艺术效果,图像传感器输出的图像电信号需要经过视频处理放大器的进一步处理。最后,为了满足摄像机输出接口的特性,图像信号还需要在编码器中进行编码,转换成适合传输的信号格

式。下面分别介绍各个模块。

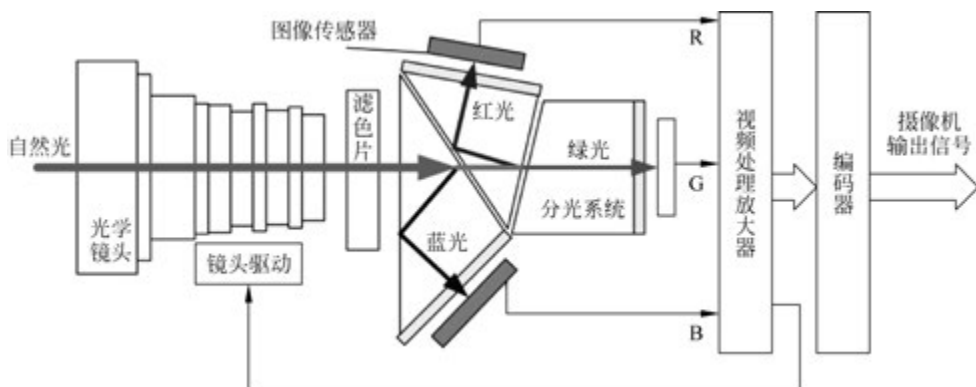


图 3-22 彩色摄像机系统结构示意图

1. 光学镜头

(1) 光学镜头的组成。光学镜头主要负责将被摄景物反射或透射的光线汇聚并聚焦在图像传感器上,以确保清晰成像。镜头的可调节部分包括聚焦环、变焦环、光圈环、倍率镜、微距镜以及后焦距调节环等。

调整聚焦环能使特定距离的拍摄物清晰成像在图像传感器上,环上所标记的数字代表能清晰成像的物距。

变焦环的作用是调整镜头的焦距,从而改变画面的构图和视角。

光圈环的作用是控制镜头的进光量,并调节景深,即能清晰成像的景物深度范围。大光圈,摄像机输出的图像会更亮,但景深会减小。

倍率镜能以一定的倍数扩大镜头的焦距,扩展拍摄范围。例如,加入 2 倍($\times 2$)的倍率镜后,原镜头焦距 9~143mm 可变为 18~286mm。

微距镜,又叫超近摄镜,专门用于拍摄因物距过近而无法清晰成像的景物,它能在更短的物距下实现清晰拍摄。

后焦距调节环用于调整镜头的后焦距,即镜头最后一个透镜表面到成像面之间的距离,确保变焦过程中聚焦环锁定的景物能一直清晰成像。如果后焦距调节不准确,就可能导致聚焦不清晰的问题。例如,在长焦端拍摄时某景物可能呈现清晰,但当变焦至短焦端时,该景物却可能变得模糊。

对于在长焦和广角端容易出现焦点漂移的 4K 镜头,还特别加入了后主镜组,以补偿焦点漂移,确保图像质量。

(2) 光学镜头的透射特性。自然光是由不同波长的单色光组合而成。当光线穿越摄像机光学镜头的镜片组时,会经历反射、透射、折射、散射等物理过程。对于不同波长的光线而言,在同一光学镜头光轴上,其透过率也有可能展现出明显差异。为了描述摄像机光学镜头的透射特性,业界引入了光谱透过率曲线的概念,该曲线展现了不同波长的光穿越摄像机光学镜头光轴时所表现出的透过能力的差异。光学镜头光谱透过率曲线示例如图 3-23 所示,水平坐标轴代表光谱波长,以符号 λ 表示,其计量单位为纳米(nm);垂直坐标则反映了光学镜头的透过率,用符号 τ 标记,并以百分数的形式表达。通过运用先进的镀膜技术以及选配不同材质的镜片,各类镜头的光谱透过率曲线也有所不同。一般而言,在可见光波段范

围,这些曲线往往表现得相对平缓。

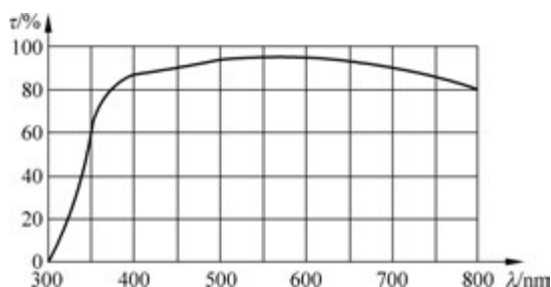


图 3-23 光学镜头光谱透过率曲线示例

2. 滤色片

在摄像机光学镜头的后方,通常会装配中性滤色片或色温校正片。

中性滤色片(Neutral Density filter, ND),简称 ND 滤镜,作为摄像机上最常见的滤色片,它对所有波长的光有相同的透过率。由于 ND 滤镜对光线色彩不产生改变,同时保持光线色温的原貌,因此又称作“灰片”。与调整光圈相比,使用 ND 滤镜可在不改变景深的前提下,减少入射光的光通量。市场上常见的摄像机 ND 滤镜通常采用多档位设计,包括 Clear (无 ND 滤镜状态)、1/4 ND、1/16 ND、1/64 ND 等。这些档位代表了滤镜对光线的透过比例,比如 1/4 ND 意味着光线透过率为 25%。在高端摄像机中,已引入基于电子控制的连续可调 ND 滤镜,实现了透光率的精确无级调节。

由于早期摄像机的白平衡电路增益调节范围不够,当光源的色温偏差较大时,需借助色温校正片(Color Compensating filter, CC)来辅助实现白平衡(见 3.3.2 节)校正。在电视摄像领域,通过使用特定色温值的色温校正片来补偿光源色温与基准色温之间的偏差。实际操作中,首先利用色温校正片对光源色温进行初步补偿,再调整白平衡设置进行色温的精细校正。然而,随着摄像机技术的不断进步,特别是增益调节范围的显著扩大,现代摄像机大多已摒弃了色温校正片的物理装置,转而全面采用电子手段来完成白平衡的精确调节。

此外,部分摄像机在其图像传感器的前端配备了光学低通滤波器(Optical Low-Pass Filter, OLPF),旨在滤除光学信号中的高频成分,以减轻图像传感器对光图像进行采样时因高频成分混叠所导致的摩尔纹。另外,某些摄像机还特别安装了红外线截止滤波器,目的在于防止红外线对图像成像的干扰。

3. 分光系统

分光系统的作用是将景物的彩色光线按照特定的光谱响应分解为三路基色光(即红、绿、蓝三基色图像),并精准投射到相应的图像传感器光敏面上。这一过程通常依赖于分光棱镜或滤色片来实现光线的有效分解。

(1) 分光棱镜分光。分光棱镜的设计基础是薄膜干涉原理,它通常是由三块(如图 3-22 所示)或四块棱镜贴合而成,四块分光棱镜结构示意图如图 3-24 所示。在这些棱镜的分光面上,蒸镀有具有一定厚度和折射率的介质薄膜。该薄膜通常具有较大折射率,又被称为干涉膜或分色膜。

分光棱镜的设计是通过对光线的折射、反射及光程差的控制,使某些波长的光线反射出来,另一些波长的光线透射过去,从而实现分光。以图 3-24 为例,当景物反射的自然光通过

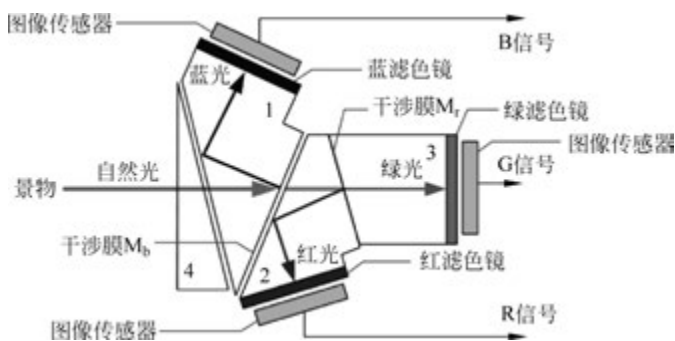


图 3-24 四块分光棱镜结构示意图

光学镜头透射到分光棱镜 1 时,棱镜 1 表面镀的干涉膜 M_b 会将其中的蓝光反射出来。经蓝滤色镜进一步提高纯度后,蓝光被投射到相应的图像传感器上进行成像。而棱镜 1 透过的光线则会射到棱镜 2 的干涉膜 M_r 上,此时红光被反射出来。剩下的绿光则继续进入棱镜 3。与蓝光相似,红光和绿光也会经过相应颜色的滤色镜提纯后,进入图像传感器进行成像。

薄膜干涉原理示意图如图 3-25 所示。假设干涉膜的折射率为 n_0 ,厚度为 d 。入射光从折射率为 n_1 的玻璃射到干涉膜表面 1 的 A 点,入射角为 α ,之后一部分光反射,形成光路 F_1 ; 剩余的光线则折射到干涉膜表面 2 的 B 点,并再次发生折射和反射。当第二次反射的光折回到表面 1 的 C 点时,一部分光被折射到玻璃中,形成光路 F_2 ; 而另一部分则再次被反射。因为薄膜的厚度非常薄,光束 F_1 和 F_2 会发生相互干涉,它们之间的光程差 δ 计算如下:

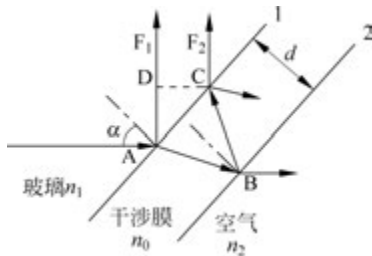


图 3-25 薄膜干涉原理示意图

$$\delta = (\overline{AB} + \overline{BC})n_0 - \overline{AD}n_1 = 2d\sqrt{n_0^2 - n_1^2 \sin^2 \alpha} \quad (3-16)$$

这一光程差是决定干涉现象的关键因素。当 δ 等于光的波长时,光束 F_1 和 F_2 相位相同,它们会互相加强,从而实现该波长光的全反射。当 δ 等于光的半波长时,光束 F_1 和 F_2 相位相反,它们会相互抵消,从而实现该波长光的全透射。

在实际应用中,分光棱镜的结构设计会根据特定需求进行调整,不同的分光棱镜可能会具有不同的分光次序。

分光棱镜将自然光分解成红、绿、蓝三种色光,这三种色光随后被独立的图像传感器捕捉。因此,图像传感器也需要安装三片。三片式摄像机通过合理地在光路上镶嵌好三片独立的图像传感器,能够有效避免彩色摩尔纹现象,进而获得高质量的彩色图像。然而,由于分光棱镜占据了一定的物理空间,三片式摄像机在结构设计上要求在镜头和传感器之间保持足够的距离(后焦距),以确保分光棱镜的妥善安装。这一特定要求相应地限制了摄像机整体尺寸的小型化潜力。

(2) 单层滤色片分光。在相机和电影摄影机的设计中,由于后焦距的限制,往往没有足够的空间来安装分光棱镜。因此,这类摄像系统通常采用一种在单一图像传感器上交错部署红、绿、蓝三种颜色滤色片阵列的方法来实现分光。具体而言,图像传感器的表面被精细划分成了数以百万至千万的像元,而每个像元上面都镶嵌有一片特定颜色的滤色片单元。

单层滤色片分光大都采用棋盘格式滤色膜阵列,目前常见的是拜尔滤色片(Bayer filter mosaic),拜尔滤色片像元结构示意图如图 3-26 所示,该技术是由伊士曼柯达公司的美国科学家 Bryce Bayer 在 20 世纪 70 年代中期发明的。拜尔阵列的红(R)、绿(G)和蓝(B)滤色片单元按照像素正交排列,每一行上只有两种滤色单元,或者是 G、B,或者是 G、R。因为人眼对绿光最敏感,因此 G 滤色单元光的数量是 R 或 B 的两倍,也就是说,所有的滤色单元中, G 单元占 50%,R 和 B 各占 25%。因此,拜尔滤色片的模式也被称为 RGBG、GRGB 或者 RGGB。

索尼公司开发的 8K CMOS 图像传感器,为了提高其图像传感器的表面利用率,采用了 Q67 滤色片结构。Q67 滤色片像元结构示意图如图 3-27 所示,相当于将像素正交排列的拜尔滤色片旋转了 45° 。

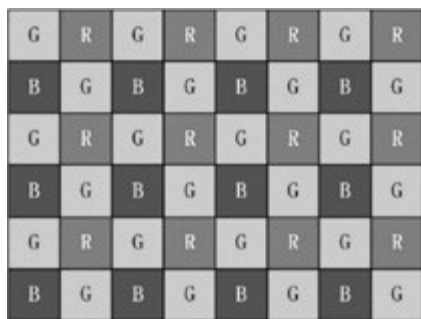


图 3-26 拜尔滤色片像元结构示意图

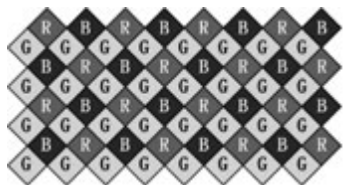


图 3-27 Q67 滤色片像元结构示意图

基于滤色片分光的摄像机,主要采用内插算法来重建全分辨率的 RGB 图像信号。以最基础的双线性插值法为例,针对当前像素缺失的两个基色分量,提取其 8 邻域中对应的 4 个同色基色像元,通过计算这 4 个邻域像素的平均值进行估算与填补,从而合成完整的全分辨率 RGB 图像。

由滤色片分光的单片式摄像机不需要考虑尺寸庞大的分光棱镜,这使得摄像机结构得以更加小巧和轻便。不过,由于采用了红、绿、蓝滤色阵列,三种基色信号的采集位置存在差异,这可能会在拍摄图像高频信息中引入彩色摩尔纹,进而对图像质量造成一定影响。

(3) 分层感光技术。每一像素单元包含有上下三层感光单元,分别负责吸收不同波长的光线,同时图像传感器像素元件设置在不同深度,从而实现在同一像素位置同时捕获 R、G、B 三基色数据。这种设计使得传感器避免了因 R、G、B 像素位置差异产生的彩色摩尔纹,也不需要配置尺寸庞大的分光棱镜,减少了对摄像机外形设计的影响。

2020 年,NHK 利用分层感光技术,开发了一款基于有机薄膜的图像传感器。NHK 开发的基于有机薄膜的图像传感器结构示意图如图 3-28 所示。该传感器在硅基板上集成了用于红光成像的 CMOS 图像传感器,其上叠加了两层有机薄膜(有机层),分别用于拾取蓝光和绿光。为了读取并输出有机层产生的电信号,有机层表面还配置了薄膜晶体管电路层。同时,为了保护整个图像传感器的稳定性和耐用性,在最上层安装了玻璃基板。当入射光线进入传感器时,它首先穿透玻璃基板,到达第一层有机层。该层仅吸收蓝光成分,并将其转换为 B 信号,同时允许绿光 and 红光成分穿透。附近的薄膜晶体管电路读取 B 信号并将其输出。接着,第二层有机层仅吸收绿光成分,并将其转换为 G 信号,同样地,其一侧的薄膜晶

体管电路读取该信号并将其输出。最后,红光成分由底层 CMOS 图像传感器采集,并转成 R 信号输出。

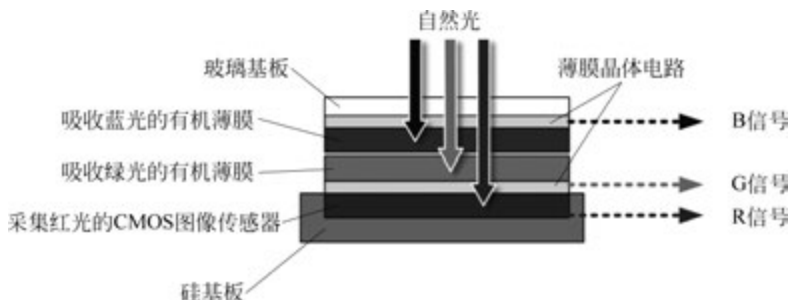


图 3-28 NHK 开发的基于有机薄膜的图像传感器结构示意图

(4) 分光特性。摄像机分光系统的分光特性,通常是以光线从分光系统透射出来的色光的相对透过率(归一化分光响应)为纵轴,以可见光谱波段为横轴来直观展现的。分光系统分光特性曲线示意图如图 3-29 所示。分光系统将自然光分解为蓝、绿、红三个独立的光路通道,每一通道对光谱中不同波长的光线展现出独特的分光响应,三条曲线分别对应了三个光路对不同波长光的分光响应。不同类型的分光系统,分光特性曲线会有细节上的差异。

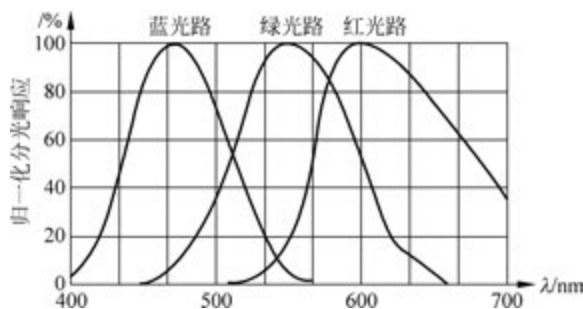


图 3-29 分光系统分光特性曲线示意图

4. 图像传感器

CMOS 对光谱的感光范围不但覆盖了人眼视觉系统所能感知的整个可见光波段,更延伸至近红外区域,最远波长可达 1100nm。CMOS 图像传感器的归一化光谱灵敏度曲线与人眼相对视敏函数曲线的对比示意图,如图 3-30 所示。在可见光范围内,CMOS 的光谱灵敏度展现出相对平坦的分布特性。为了避免人眼不可见的红外线光干扰 CMOS 成像,则需要增加红外线截止滤波器。

5. 视频处理放大器

由于摄像机光学系统和图像传感器的性能都不是最理想的,摄像机经过光电转换生成的电信号会有很多缺陷,诸如信号电平不够高、亮度层次不合理,以及色彩不够逼真等。为此,摄像机的视频处理放大器承担起对图像信号进行精细化调整的重任,这包括但不限于信号增益(电平放大)、黑平衡、白平衡、伽马校正、彩色校正和细节校正等。

当采用 CCD 作为图像传感器的数字摄像机时,由于其输出为模拟信号,视频处理放大器就需要由模拟处理部分、AD 转换器和数字处理部分组成。而采用 CMOS-DPS 图像传感

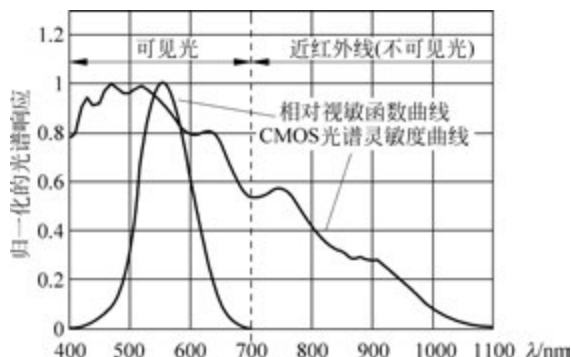


图 3-30 CMOS 图像传感器的归一化光谱灵敏度曲线与人眼相对视敏函数曲线的对比示意图

器的摄像机,AD 变换、部分增益处理和去噪是在图像传感器中完成的,CMOS 输入视频处理放大器的信号是数字视频信号。因此,此类视频处理放大器倾向于采用全数字化设计,不仅显著提升了系统的稳定性和可靠性,还实现了参数控制的精细化与用户自定义设置的灵活性,同时促进了设备的高度集成化。

视频处理放大器内部的处理流程复杂多样,部分处理基于红、绿、蓝三基色信号展开,而另一部分则需依赖于如亮度和色差信号等其他类型的基色信号。因此,还需要矩阵运算模块将红、绿、蓝三基色信号转换为其他基色信号。

综上所述,视频处理放大器在摄像机视频系统中扮演着至关重要的角色,其设计与优化直接关系到摄像机输出的最终图像质量的高低。

3.3.2 白平衡与黑平衡

在数字视频拍摄过程中,光线条件的变化,尤其是色温的变化,对画面色彩有着至关重要的影响。实际的拍摄环境往往具有一定的色温,带有特定的光色,这会导致摄像机捕捉到的画面带有一定的色彩偏移。相同的拍摄对象在不同的色温环境下,如人从室内转移到室外,不同色温环境下拍摄的画面交替出现时,会有明显的色彩跳跃,影响观看体验。此外,由于摄像机器件特性的固有差异,同一场景在不同摄像机的记录下,也会呈现出不同的色彩表现,这个问题在多机位拍摄中尤为突出。为了减少画面色彩不统一的问题,确保拍摄到的基准白色和黑色能够准确还原,即还原拍摄目标自身固有的颜色,专业摄像机需要进行白平衡和黑平衡的调整。

1. 白平衡

由色度学原理可知,等量的 R、G、B 三基色可以混合出白色,因此 R、G、B 三基色图像信号相等时,对应的就是黑白图像。当摄像机在拍摄白色对象时,输出的 R、G、B 三个基色信号电平相等,就称为摄像机白平衡;而当 R、G、B 三个基色激励信号等幅时,显示器得到的是显示参考白的白色或灰度图像,称为显示器白平衡。由于显示器白平衡一般是固定不允许调整的,通常白平衡就特指摄像机白平衡。白色对象处于不同色温环境下时,会表现出色温的颜色,但可以通过白平衡的调整使得输出的 R、G、B 三个基色信号电平相等,从而在显示器端恢复白色对象的本身固有的颜色(白色)。这一过程主要通过调整视频处理放大器中基色信号放大器的增益来实现。白平衡的调整方法主要分为自动和手动两种。

自动白平衡(Auto White Balance, AWB)操作简便,只需在景物前方放置一张反射率为90%的标准白卡(或纯白色纸张、衣物等替代物),调整焦距,使白卡占据画面70%以上,并合理设置曝光(达到一定的信号电平),然后按下摄像机的AWB按键,即可完成调整,使R、G、B三路信号电平一致,重现纯白色画面。然而,自动白平衡调整不够精细,即使是同型号的摄像机,经过自动白平衡调整后,仍然存在微小的颜色差异,在高质量节目制作中,这种差异可能无法接受。

因此,在更高要求的视频应用中,通常会采用手动白平衡。这需要使用专业测试卡、监测设备(如示波器),并由视频工程师手工调整R、G、B信号的增益旋钮,以实现更为精确的白平衡。手动白平衡虽然操作烦琐,但能提供更高的色彩一致性,确保拍摄到的“白”更加准确。

显然,摄像机在一定色温环境下调整好白平衡后,转移到更低色温环境下的拍摄画面会偏黄,高色温环境会偏蓝。

2. 黑平衡

在高质量节目制作,如演播室节目制作中,黑平衡的调整通常先于白平衡进行。调整黑平衡的目的是确保摄像机在拍摄纯黑物体(或关闭光圈)时输出的红、绿、蓝电平在监视器上呈现为纯黑色画面。

黑平衡的调整同样分为自动和手动两种方法。自动黑平衡(Auto Black Balance, ABB)只需按下相应按键,摄像机会自动关闭光圈并完成R、G、B三路电平一致性的调节。而手动黑平衡则需要在专业监测设备的辅助下,由视频工程师手工调整R、G、B信号的增益旋钮,以实现精确的黑平衡。

3.4 色度匹配

在理想的视频系统中,视频信号历经采集(光电转换)、传输和重现(电光转换)所有步骤后,显像端重现的色光,应该与景物反射的色光(或者说进入摄像机的色光)外貌相同(即色度坐标相同),这样才能保证视频系统正确重现了景物的颜色,这就是色度匹配。

为了更好地理解色度匹配,首先需要分析整个视频传输流程中,每个关键环节对光谱特性的具体影响。视频系统信号光谱特性传输示意图如图3-31所示,视频传输流程从光源开

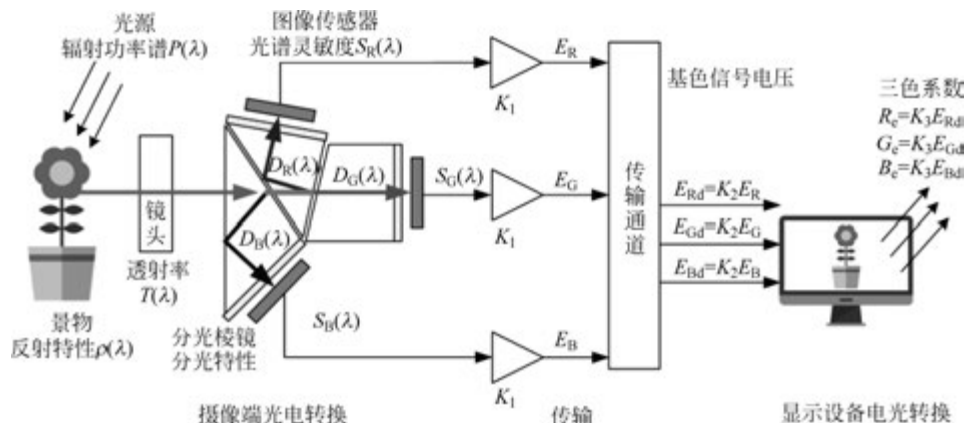


图 3-31 视频系统信号光谱特性传输示意图

始,辐射功率谱为 $P(\lambda)$ 的光源照射到被摄场景上,随后被场景中的景物反射。景物的反射特性表示为 $\rho(\lambda)$,因此,反射出的光线辐射功率谱为 $P(\lambda) \rho(\lambda)$ 。此光线再通过透射率为 $T(\lambda)$ 的镜头透射后,辐射功率谱进一步变为 $P(\lambda) \rho(\lambda) T(\lambda)$ 。

接着,光线进入分光特性为 $D_R(\lambda)$ 、 $D_G(\lambda)$ 、 $D_B(\lambda)$ 的分光棱镜,被分解为 R、G、B 三路光像信号输出。随后,这三路光信号分别投射至对应的图像传感器上进行光电转换,其光谱灵敏度分别为 $S_R(\lambda)$ 、 $S_G(\lambda)$ 、 $S_B(\lambda)$ 。

最后,摄像机的放大器将这三个基色信号处理并放大成适合传输的电压信号,假设三路放大器输出电压与输入电流的转换比例均为 K_1 ,则生成的三路基色电压 E_R 、 E_G 、 E_B 与景物光色之间的数学关系,可由式(3-17)来描述。该式揭示了视频系统采集端由光谱特性到生成电信号的对应关系。

$$\begin{cases} E_R = K_1 \int_{380}^{780} P(\lambda) \rho(\lambda) T(\lambda) D_R(\lambda) S_R(\lambda) d\lambda \\ E_G = K_1 \int_{380}^{780} P(\lambda) \rho(\lambda) T(\lambda) D_G(\lambda) S_G(\lambda) d\lambda \\ E_B = K_1 \int_{380}^{780} P(\lambda) \rho(\lambda) T(\lambda) D_B(\lambda) S_B(\lambda) d\lambda \end{cases} \quad (3-17)$$

其中,将镜头的透射特性 $T(\lambda)$ 、分光系统的分光特性 $D(\lambda)$ 、图像传感器的光谱灵敏度 $S(\lambda)$ 的乘积定义为摄像机的综合光谱响应特性,并由式(3-18)所示的 $\bar{r}_o(\lambda)$ 、 $\bar{g}_o(\lambda)$ 、 $\bar{b}_o(\lambda)$ 分别表示摄像机三路基色信号的综合光谱响应特性函数,则式(3-17)可简化为式(3-19)。

$$\begin{cases} \bar{r}_o(\lambda) = T(\lambda) D_R(\lambda) S_R(\lambda) \\ \bar{g}_o(\lambda) = T(\lambda) D_G(\lambda) S_G(\lambda) \\ \bar{b}_o(\lambda) = T(\lambda) D_B(\lambda) S_B(\lambda) \end{cases} \quad (3-18)$$

$$\begin{cases} E_R = K_1 \int_{380}^{780} P(\lambda) \rho(\lambda) \bar{r}_o(\lambda) d\lambda \\ E_G = K_1 \int_{380}^{780} P(\lambda) \rho(\lambda) \bar{g}_o(\lambda) d\lambda \\ E_B = K_1 \int_{380}^{780} P(\lambda) \rho(\lambda) \bar{b}_o(\lambda) d\lambda \end{cases} \quad (3-19)$$

通常摄像机镜头的透射特性曲线(见图 3-23)和图像传感器的光谱灵敏度曲线(见图 3-30)都比较平坦。因此,摄像机的综合光谱响应特性主要由分光系统的分光特性函数 $D(\lambda)$ 决定。

接收端显示器接收到三基色电压信号 E_R 、 E_G 、 E_B 后,将其放大转换为用于显示器件发光的激励电压信号 E_{Rd} 、 E_{Gd} 、 E_{Bd} 。随后,该激励电压信号激发显示器发出一定强度、三色系数分别为 R_e 、 G_e 、 B_e 的显像三基色光。假设在这一过程中的两次转换里,三个基色通道的特性均表现一致且为线性,设转换系数分别为 K_2 、 K_3 (实际中,调节三个基色通道的电信号放大系数,实现假设条件),则显示器实际发出的显像三基色的三色系数 R_e 、 G_e 、 B_e 与接收端收到的三基色电压 E_R 、 E_G 、 E_B 的数学关系为

$$\begin{cases} R_e = K_3 E_{Rd} = K_2 K_3 E_R \\ G_e = K_3 E_{Gd} = K_2 K_3 E_G \\ B_e = K_3 E_{Bd} = K_2 K_3 E_B \end{cases} \quad (3-20)$$

结合摄像机采集端的信号转换,将式(3-19)代入式(3-20),可以得到从光源特性 $P(\lambda)$ 、场景反射特性 $\rho(\lambda)$,再经摄像采集、视频传输,最后显像重现三基色光的数学建模表达式如式(3-21)所示。

$$\begin{cases} R_e = K_1 K_2 K_3 \int_{380}^{780} P(\lambda) \rho(\lambda) \bar{r}_o(\lambda) d\lambda \\ G_e = K_1 K_2 K_3 \int_{380}^{780} P(\lambda) \rho(\lambda) \bar{g}_o(\lambda) d\lambda \\ B_e = K_1 K_2 K_3 \int_{380}^{780} P(\lambda) \rho(\lambda) \bar{b}_o(\lambda) d\lambda \end{cases} \quad (3-21)$$

假设显像三基色混色曲线为 $\bar{r}_e(\lambda)$ 、 $\bar{g}_e(\lambda)$ 、 $\bar{b}_e(\lambda)$,则通过式(3-22),可计算混配出辐射功率谱为 $P(\lambda)\rho(\lambda)$ 的色光所要求的理想三系数为 R'_e 、 G'_e 、 B'_e 。

$$\begin{cases} R'_e = \int_{380}^{780} P(\lambda) \rho(\lambda) \bar{r}_e(\lambda) d\lambda \\ G'_e = \int_{380}^{780} P(\lambda) \rho(\lambda) \bar{g}_e(\lambda) d\lambda \\ B'_e = \int_{380}^{780} P(\lambda) \rho(\lambda) \bar{b}_e(\lambda) d\lambda \end{cases} \quad (3-22)$$

为了实现显像端重现的色光与景物反射的色光外貌相同,即色度匹配,需满足式(3-23)所示关系。

$$\begin{cases} R_e = R'_e \\ G_e = G'_e \\ B_e = B'_e \end{cases} \quad (3-23)$$

将式(3-21)与式(3-22)代入式(3-23)的两端,可得

$$\begin{cases} \bar{r}_e(\lambda) = K_1 K_2 K_3 \bar{r}_o(\lambda) = K \bar{r}_o(\lambda) \\ \bar{g}_e(\lambda) = K_1 K_2 K_3 \bar{g}_o(\lambda) = K \bar{g}_o(\lambda) \\ \bar{b}_e(\lambda) = K_1 K_2 K_3 \bar{b}_o(\lambda) = K \bar{b}_o(\lambda) \end{cases} \quad (3-24)$$

其中, $K = K_1 K_2 K_3$ 。式(3-24)表明,要想实现色度匹配,摄像机红、绿、蓝三路综合光谱响应特性曲线 $\bar{r}_o(\lambda)$ 、 $\bar{g}_o(\lambda)$ 、 $\bar{b}_o(\lambda)$ 必须正比于显像三基色的混色曲线 $\bar{r}_e(\lambda)$ 、 $\bar{g}_e(\lambda)$ 、 $\bar{b}_e(\lambda)$ 。也就是说,显像混色曲线就是理想的摄像综合光谱响应曲线。由于镜头的透射率特性和图像传感器的光谱灵敏度特性曲线都比较平缓,综合光谱响应特性主要受到分光系统的分光特性影响。这意味着摄像机的分光系统的分光特性需与显像混色曲线的形状一致。

摄像机分光系统响应曲线与显像三基色混色曲线的对比如图 3-32 所示。可以观察到摄像机分光系统响应曲线与显像三基色混色曲线的主瓣曲线相似,但显像混色曲线中还有负次瓣和正次瓣曲线。因此,摄像机还需进行彩色校正处理,从而在光电转换的综合效果上弥补摄像机分光系统响应曲线与显像三基色混色曲线不一致的不足,使整个视频系统接近色度匹配。

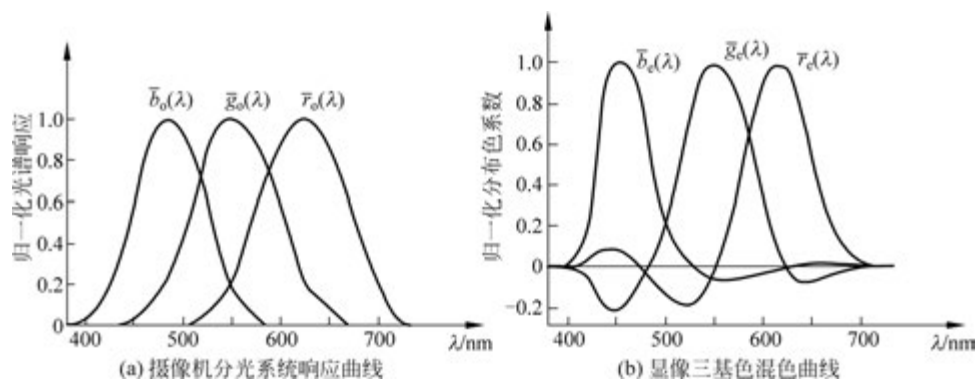


图 3-32 摄像机分光系统响应曲线与显像三基色混色曲线的对比

3.5 视频系统的 γ 特性

3.5.1 视频系统的非线性

对视频系统来说,保证拍摄场景和显示重放的亮度感知一致性非常重要。视频系统信号传输示意图如图 3-33 所示,该图描绘了图像光、电信号在视频系统中的整个处理流程:从采集阶段开始的光电转换,经由传输阶段的电信号转换,直至最终在重现阶段完成电光转换。具体而言,这一过程始于将场景亮度(记作 B_S)转换为摄像机输出的电压信号(U_S),随后该电压信号在传输系统中被进一步转换为重现端接收的电压信号(U_P),最后,显示设备将这一电压信号转换成重现亮度(B_P)。视频系统需要确保所重现的画面亮度 B_P 与拍摄场景 B_S 给人带来的亮度层次感觉一致。

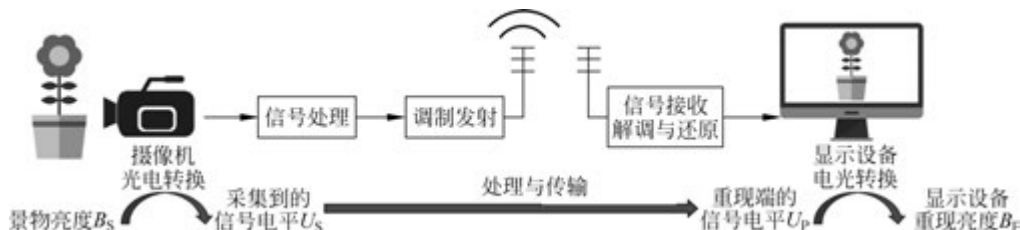


图 3-33 视频系统信号传输示意图

早期电视系统,受 CRT 显像管的特性局限,图像信号电平与显示器重现的亮度之间是非线性转换关系。所谓线性,就是输出信号幅度与输入信号幅度之间呈现等比例的线性关系,其特性为过原点的一条直线,表达函数为线性的一次方程。非线性,就是输出信号幅度与输入信号幅度之间呈现非直线的曲线关系,表达函数为非线性方程。

摄像机端的光电转换(B_S 到 U_S)、处理与传输端的信号转换(U_S 到 U_P)、显示设备端的电光转换(U_P 到 B_P),均可由三个幂函数公式来精确表征,即式(3-25)、式(3-26)和式(3-27),其中 K_1 、 K_2 、 K_3 是放大倍数,调节着信号的强度;而 γ_1 、 γ_2 、 γ_3 则决定了信号转换过程中的线性或非线性特征。 γ 读作“伽马”(Gamma),通常人们用伽马值的大小来描述信号转换的非线性特性。

$$U_S = K_1 B_S^{\gamma_1} \quad (3-25)$$

$$U_P = K_2 U_S^{\gamma_2} \quad (3-26)$$

$$B_P = K_3 U_P^{\gamma_3} \quad (3-27)$$

联立式(3-25)、式(3-26)、式(3-27),可以推导出视频系统从输入到输出的整体转换特性模型,如式(3-28)所示。

$$B_P = K_3 K_2^{\gamma_3} K_1^{\gamma_2 \gamma_3} B_S^{\gamma_1 \gamma_2 \gamma_3} = K B_S^\gamma \quad (3-28)$$

其中,放大系数 $K = K_3 K_2^{\gamma_3} K_1^{\gamma_2 \gamma_3}$, $\gamma = \gamma_1 \gamma_2 \gamma_3$,称为系统 γ 。

若式(3-28)中的 γ 值取 1,则视频系统中被拍摄的场景亮度和图像重现的亮度呈现线性关系,从而保证观众看到的画面亮度更加逼真,实现无失真的视觉效果。

然而,早期电视系统采用的 CRT 显示器,其电光转换特性具有固有的非线性物理特性, γ_3 大于 1,CRT 显像管电光转换的非线性特性曲线如图 3-34 所示。在视频传输系统中,对图像电信号的处理往往是线性的增益处理;同时,摄像机内部广泛采用的 CCD 或 CMOS 图像传感器,其光电转换特性也呈现出线性特征。因此,在上述信号转换过程中, γ_1 和 γ_2 都取值为 1。由式(3-28)可知,视频系统的 γ 值大于 1,如果没有任何处理,视频系统就会引入非线性失真,对图像带来影响。

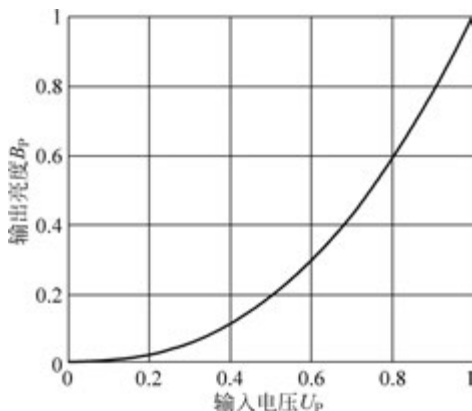


图 3-34 CRT 显像管电光转换的非线性特性曲线

3.5.2 非线性对图像的影响

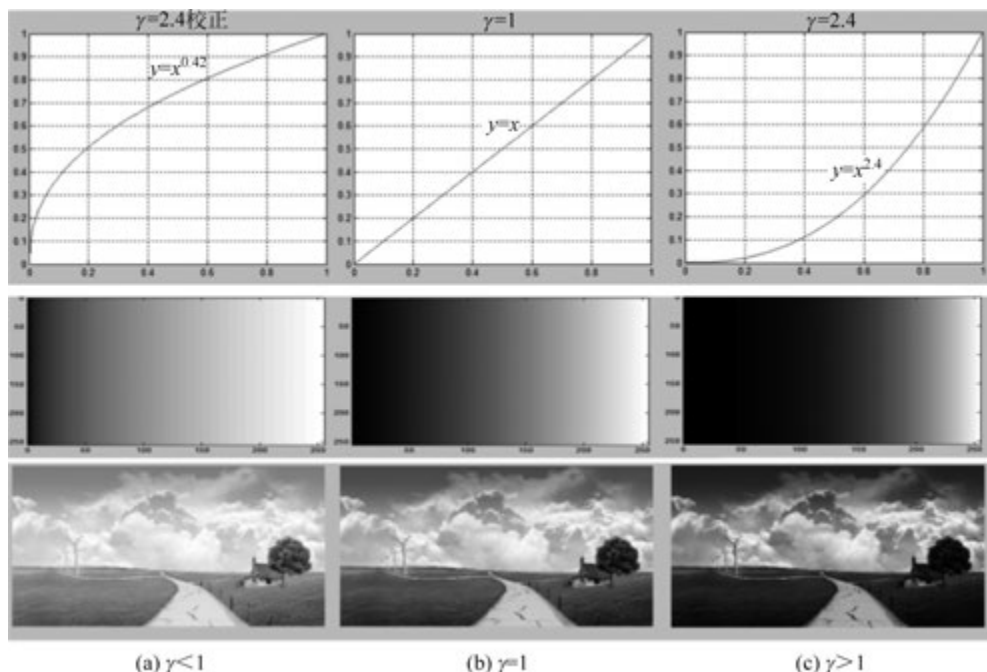
1. γ 对图像亮度的影响

当 γ 值不等于 1 时,视频系统可能会引入亮度失真,这种失真在 γ 为常数的情况下被称为均匀性亮度失真。

考虑一个特定的场景:一幅从左至右渐变,亮度由低至高的黑白图像。 γ 值对图像亮度影响的示意图如图 3-35 所示。当 $\gamma=1$ 时,视频系统能够精确重现原始景物的亮度信息,此时亮度转换曲线呈现为一条直线,如图 3-35(b)所示,确保了画面亮度的真实还原。

当 $\gamma>1$ 时,亮度转换曲线的形态发生变化,表现为中部向下凹陷,见图 3-35(c)。而曲线的两端仍与 $\gamma=1$ 时保持一致。这意味着,尽管图像的最亮和最暗部分亮度不受影响,但中间亮度区域普遍低于 $\gamma=1$ 时的水平。具体来说,原始场景中亮度低于 50% 的区域,在重现图像中亮度被显著压缩(即“黑压缩”),导致暗部细节丢失;相反,亮度高于 50% 的区域则获得更多亮度层次(即“白扩张”),形成所谓的均匀性白扩张效应。这种处理增强了画面的反差,令图像的亮暗对比更鲜明,但同时整体图像的平均亮度会下降。

反之,若 $\gamma<1$,亮度转换曲线则向上凸起,如图 3-35(a)所示,同样保持两端亮度不变,但中间区域的亮度再现值高于 $\gamma=1$ 时的情况。此时,低亮度区域被赋予了更多的亮度层次(即“黑扩展”),而高亮度区域则经历亮度层次的压缩(即“白压缩”),此现象称为均匀性白压缩。这种处理方式降低了画面的亮度反差,使得整体图像的平均亮度提升。

图 3-35 γ 值对图像亮度影响的示意图

2. γ 对图像色彩的影响

在彩色视频系统中,设景物色光的三系数为 R_0, G_0, B_0 ,重现色光的三系数为 R_d, G_d, B_d ,假设视频系统对各三色通道的放大系数 k 和 γ 特性均保持一致,那么它们之间的关系如式(3-29)所示。

$$\begin{cases} R_d = kR_0^\gamma \\ G_d = kG_0^\gamma \\ B_d = kB_0^\gamma \end{cases} \quad (3-29)$$

已知,颜色的色度由三系数的比例决定,不受系数 k 的影响,为简化分析,设 $k=1$ 。

在 $\gamma=1$ 时,景物色光的三系数能够无失真地通过视频系统,确保重现色光的色度与原始景物保持一致。而当 $\gamma \neq 1$ 时,除了可能引入亮度失真外,还可能对图像的色彩表现产生影响,这种彩色失真的具体表现取决于图像的颜色构成。

对于黑、白、灰这类三系数相等的消色,即使 $\gamma \neq 1$,重现色光的三系数仍然保持相等,颜色仍然是没有色调的消色。

对于红、绿、蓝三基色或其对应的三补色青、品、黄,由于它们的三系数中仅包含 0 和 1,经过 $\gamma \neq 1$ 的视频系统处理后,重现色光三系数保持不变,色度不变。

对于非饱和的三基色或其补色,其三系数比值可表示为一个或两个 1,其余为小于 1 的正实数比值。例如非饱和黄色 $\left(\frac{5}{12} : \frac{5}{12} : \frac{1}{6}\right)$,三系数比值等价于 $1 : 1 : \frac{2}{5}$ 。当 $\gamma > 1$ 时,以 $\gamma=2$ 为例,重现色光的三系数比例变为 $1 : 1 : \frac{4}{25}$,原本占比少基色光成分重现时会变得更少,因此相当于白光成分在减小,饱和度增加,而色调维持黄色不变。相反,当 $\gamma < 1$

时,非饱和的三基色或其补色,重现色光的饱和度会降低,色调保持不变。

对于饱和度为 100% 且非三基色和其补色,如橙色 $\left(\frac{3}{4} : \frac{1}{4} : 0\right)$,其原始三色系数比值为 3 : 1 : 0,由于红色系数的成分最高,所以是偏红的橙色。当 $\gamma > 1$ 时,以 $\gamma = 2$ 为例,三色系数比例调整为 9 : 1 : 0,饱和度保持不变为 100%,但色调中原本成分较多的红色所占比重会增加更多。从该示例可以看出,当 $\gamma > 1$ 时,重现色光的三色系数比例会发生变化,饱和度仍维持不变,但色调变化的结果是改变原始色光中占比,色光中原本占比较大的颜色(本例中为红色)会变得更加显著。相反地,当 $\gamma < 1$ 时,系统重现色光的色调会缩小原始色光间的占比差距。

综上所述,可以用颜色在麦克斯韦彩色三角形的色度坐标变化来更好地描述颜色受 γ 的影响:颜色的色度坐标处于三角形的顶点(三基色)、底边中点(三补色)和重心(消色),不受 γ 值影响;当 $\gamma > 1$ 时,其他颜色的重现色光的亮度会呈现均匀性白扩张,色度坐标朝着三角形的底边或顶点移动,饱和度增加,色调向三色系数最大值的方向移动;当 $\gamma < 1$ 时,重现色光的亮度出现均匀性白压缩,饱和度降低,色调会远离三色系数最大值的方向移动。

因此,实际视频系统的系统 γ 值往往都稍大于 1,为 1~1.2,这样使得重现视频图像的饱和度更大,画面更加鲜艳,观看体验更佳。在国际电信联盟 2008 年发布的 ITU-R BT. 2129 报告书“用户对周围 HDTV 节目制作环境中主监视器的平板监视器的需求”中提到,消费者观看环境的端到端系统 γ 值大约为 1.2。

3. γ 的非线性转换函数表示

为了更好地描述 γ 对视频系统的影响,相关电视系统视频标准中将图 3-33 所示的视频系统中的光到电的转换特性用电光转换函数(Opto-Electrical Transfer Function, OETF)描述,将显示器电到光的转换特性用电光转换函数(Electro-Optical Transfer Function, EOTF)描述,将系统光到光的转换特性用光光转换函数(Opto-Optical Transfer Function, OOTF)描述,并且有 $OOTF = OETF \times EOTF$ 。实际视频标准中,为了准确规范 OETF 或 EOTF,往往会用一个较为复杂的函数来描述非线性转换特性。由于非线性整体特性仍然近似于一个 γ 值幂函数,因此有时也会直接用具体的 γ 值代称对应的 OETF 或 EOTF,但还是注意区分这两者的不同。一般来说,不讨论视频传输系统中的电电转换特性,即假设 $\gamma_2 = 1$ 。

值得注意的是,在过去的一段时间里,电视行业并没有统一的标准来规范摄像机端 OETF 的 γ 值。为了提供更大的灵活性和便利性,彩色摄像机中 OETF 的 γ 值通常可以在一定范围内进行调整, γ 值范围一般为 $\frac{1}{2.4} \sim \frac{1}{2.0}$ 。从 2007 年国际电信联盟发布的标准清晰度电视规范 BT. 601-6 开始,以及在高清晰度 BT. 709 标准中,典型的制作实践中的 OETF 转换特性曲线的数学表达式使用了指数为 0.45 的幂函数,但是结合比例因子和偏移量后,该曲线逼近于指数为 0.5(1/1.956)的幂函数。

3.5.3 感知均匀编码与 γ 特性

摄像机端 γ 校正电路的引入,即 OETF,最初是为了补偿 CRT 显示器的 EOTF 非线性。这种校正电路的添加是一种“被动”适应。随着 CRT 显示器的淘汰,现代平板显示器

或投影设备 EOTF 的 γ 值基本上都是 1, 是否可以重新设计视频系统, 将 OETF 和 EOTF 的 γ 值都定为 1, 使视频系统对信号的处理更加简单呢? 下面讨论这一问题。

第 1 章讨论了人眼的视觉特性, 人眼在较低亮度环境下更为敏感, 能够感知小的亮度差异。在较高亮度环境下, 可感知的最小亮度差异则增大。人眼的对比度灵敏度阈 δ 描述了这一特性, 即人眼能分辨的最小亮度与基础亮度的比值, 在一定的亮度区间内, 趋近于一个常数。

假设一显示器的最高亮度为 100nit, 最低亮度为 1nit, 对比度为 100 : 1。若 $\delta = 0.01$, 则在最低亮度 1nit 的观察环境下, 人眼能识别 1nit 的 1%, 即 0.01nit 的亮度级差。如果采用线性编码, 以 0.01nit 为固定亮度级差, 为该显示器提供的数字视频信号需要为 $100/1\% = 10\,000$ 个亮度层次进行编码, 所需的量化比特数高达 14bit ($2^{13} = 8192 < 10\,000$)。尽管这种方法能确保暗部细节的量化噪声不被察觉, 但在高亮部分, 由于人眼无法分辨过多的细节, 这些数据实际上是被浪费的。而且, 编码后的数据量庞大, 占用了过多的传输带宽。相反, 在 100nit 的最高亮度环境下, 人眼只能识别 100nit 变化 1%, 即 1nit 以上的亮度级差, 此时只需对 100 个亮度层次进行编码, 使用 7bit 量化即可, 这样虽然可以减少数据量, 节省传输带宽, 但由于亮度层次划分不够细致, 导致低亮度区域亮度跳变, 人眼会察觉到明显的伪轮廓, 影响画面质量。

第 2 章的亮度层次一节知道, 在一定亮度对比度 C 变化范围内, 可以按照对比度灵敏度阈 δ , 将该亮度范围以一级级的亮度级差划分为若干亮度层次。按照第 1 章式(1-7)计算亮度层次的方法, 代入 $\delta = 1\%$ 和 $C = 100 : 1$, 需要 463 级亮度层次 9bit 编码才能表示。能否进一步降低量化比特数呢? 注意到, 这里 δ 采用了固定的 1% 取值, 而由图 1-13 所示的对比度灵敏度分布可知, 在低亮度和高亮度区域, δ 都会增大, 人眼相对不那么敏感。采用固定精细的 δ 取值, 会错误放大所需的亮度层次。最理想的情况是, 只对人眼实际能够分辨的亮度层次编码即可, 而这些亮度层次在人眼的感知上是均匀变化的, 这就需要所谓的感知均匀编码。

已知人类所有的感觉器官的感知特性与对应的实际物理量之间都是非线性映射关系, 并且这种对应关系都呈现幂函数(有些文献描述为对数)特性。例如, 人对重量的感知, 50kg 杠铃和 100kg 杠铃, 同样再加重 1kg, 人们会感受到后者会重很多, 这就是人的重量感知和实际重量之间是非线性对应关系。又如人的咸度味觉, 饭菜加一小勺盐觉得不咸, 如果再加一勺盐可能就咸过头了, 这就是人的咸度味觉感知和盐量是非线性对应关系。人眼的视觉也是如此, 在较暗环境下, 增加一点尼特的亮度, 人眼就能感知到明显的变化, 但在明亮环境中, 同样的亮度增幅, 可能察觉不出来。因此, 人眼感知均匀编码就是对人眼感知的亮度编码, 而不是实际的物理亮度尼特数值或相对亮度进行编码。

如何实现感知均匀编码呢? 这需要借助 CIE 标准化的明度 L^* 函数, 即主观亮度感知特性, 其描述了相对亮度下人眼主观亮度的感知特性, 即随着相对亮度的变化(0~1), 人眼主观亮度感知的实际变化情况。因此, L^* 曲线就是感知均匀编码的特性曲线, 如图 3-36 所示, 它给出了如何将物理相对亮度值转换为人眼感知均匀编码的非线性映射关系。

图 3-36 中, 水平坐标是相对亮度, 取值为 0~1, 与一定对比度范围内的光强成正比, 0 对应最低亮度; 1 对应最高亮度。垂直坐标表示人眼的主观亮度 L^* , 其曲线数值的变化更精确地表明了人眼感知亮度的均匀变化和相对亮度的变化间的非线性映射关系, 如当相对

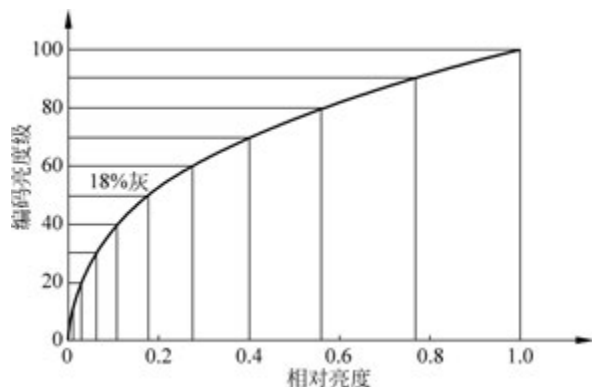


图 3-36 感知均匀编码的特性曲线

亮度为18%的暗灰时，人眼感知的是50%的中灰。此外，从图中也能揭示“低亮度区域的微小变化和高亮度区域的较大变化，人眼感知的亮度极差是一样的”这一结论，并且 δ 是动态变化的。

实践证明，按照图3-36纵轴标定的主观亮度 L^* 进行编码，即可实现感知均匀编码，从而节省编码比特数，实现8bit编码满足 $C=100:1$ 的编码需要。

为了实现感知均匀编码，需要对摄像机采集到的视频信号进行基于感知均匀编码的特性曲线的非线性变换，那么显示器接收视频信号重现图像时，还需要进行非线性的反变换。但令人惊讶的是，感知均匀编码的特性曲线非常近似于 $\gamma=1/2.4$ 的幂函数非线性曲线（可以验证： $L^*(0.18) \approx 0.495$ ，而 $(0.18)^{1/2.4} \approx 0.487$ ，两者非常接近），恰好就是显示器的固有特性 $\gamma=2.4$ 的幂函数非线性曲线的逆变换。因此现有摄像机的OETF函数，恰好就完成了感知均匀编码的非线性变换处理，可谓是惊人的巧合。

综上所述，在早期的CRT显示器时代，摄像机中的伽马校正(OETF函数)弥补了显示器电光转换的非线性，而在数字视频系统中，EOTF实现的感知均匀编码的非线性处理依然至关重要，确保了视频数值与人眼亮度感知的一致性，同时节约了数据处理资源和传输带宽。因此，摄像机的OETF即使不是因为CRT显示器的巧合而“被动”采用，也会因感知均匀编码的需要而“主动”发现，这一点可以在电影技术的伽马校正实践中得到验证。

3.5.4 视频系统 γ 特性标准

1. ITU-R BT. 1886 建议书中的EOTF

在CRT显示器时代，显示器 γ 值主要取决于显像管自身的物理特性，由于不同厂商生产的显像管产品 γ 值存在差异，电视行业难以形成统一的显示器 γ 标准。随着显像管退出市场，而新型平板显示器的电光转换特性普遍为线性，为了规范平板监视器 γ 特性，国际电信联盟于2011年发布了ITU-R BT. 1886建议书，定义了高清电视节目制作中使用的显示器应遵循的一个基于指数函数的参考EOTF公式，以使新型显示器能够模拟CRT的特性，从而保证节目制作和交换过程中图像的一致性。这一规范被电视、计算机、智能手机和平板、投影机显示设备广泛采纳。

ITU-R BT. 1886建议书的参考EOTF公式如下：

$$L = a(\max[(V+b), 0])^\gamma \quad (3-30)$$

式(3-30)中, L 代表屏幕亮度(单位为 nit 或 cd/m^2)。 V 是输入视频信号的标准化电平(黑色信号对应 $V=0$, 白色信号对应 $V=1$)。根据 ITU-R BT. 709 建议书, V 可通过式(3-31)从 10bit 量化数字高清信号中的亮度数字电平 D 推算出来。 γ 为幂函数指数, 取值为 2.40。 a 是用户增益变量, 相当于传统电视机的对比度控制。 b 是用户黑电平提升变量, 相当于传统电视机上的亮度控制。

$$V = \frac{D - 64}{876} \quad (3-31)$$

当 $V=0$ 时, 显示器的亮度为黑色亮度 $L=L_B=ab^\gamma$; 当 $V=1$ 时, 显示器的亮度为白色亮度 $L=L_W=a(1+b)^\gamma$ 。由此可推导出 a 和 b 的值:

$$a = (L_W^{\frac{1}{\gamma}} - L_B^{\frac{1}{\gamma}})^\gamma \quad (3-32)$$

$$b = \frac{L_B^{\frac{1}{\gamma}}}{L_W^{\frac{1}{\gamma}} - L_B^{\frac{1}{\gamma}}} \quad (3-33)$$

ITU-R BT. 1886 建议的 EOTF 特性与 $L=V^{2.4}$ 有一定的差异。当 $L_B=0$ 时, 由式(3-32)可得 $a=L_W$, 由式(3-33)可得 $b=0$, 此时 EOTF 特性与 $L=V^{2.4}$ 一致。当 L_B 值很低时, 两者非常接近。

2. ITU-R BT. 709 建议书中的 OETF

ITU-R BT. 709 是国际电信联盟发布的一项建议书, 规定了高清电视系统的图像特性、信号格式、数字表示、图像扫描特性、同步信号等格式和参数, 旨在为高清电视节目制作和国际交换提供统一的参数标准, 确保不同系统之间的兼容性和高质量转换。该建议书规定的高清电视信号源 OETF 表达式见式(3-34)。

$$V = \begin{cases} 4.5L, & 0 \leq L < 0.018 \\ 1.099L^{0.45} - 0.099, & 0.018 \leq L \leq 1 \end{cases} \quad (3-34)$$

其中, 场景亮度为 L (相对亮度, 取值为 $0 \sim 1$), 相应的电信号电平为 V 。该特性分为两部分: 亮度 $L \geq 0.018$ 时, 采用指数为 0.45 的幂函数(非线性特性); 该函数乘以 1.099 (比例因子) 减去 0.099 (偏移量) 是为了降低暗部亮度以减轻屏幕反光的影响, 但该幂函数在原点附近变得垂直, 使得在黑色部分有无限大增益, 产生大量噪波, 因此, 亮度 L 低于 0.018 时采用斜率为 4.5 的线性特性。该方程虽然包含指数为 0.45 的幂函数, 但比例因子和偏移量使整个方程逼近指数为 0.5 的幂函数。

数字视频分量信号在量化之前的模拟电信号, 都需要经过 OETF 转换函数, 以更好地保留低亮度区域的细节, 避免信息丢失, 符合人眼对亮度的感知特性。这些规定对于确保高清电视信号的正确处理和高质量图像显示至关重要。

3. sRGB 标准中的 EOTF 和 OETF

sRGB 颜色空间与 ITU-R BT. 709 标准一致, 并通过设备无关的颜色定义来确保颜色信息的正确还原。sRGB 的标准化工作由国际电工委员会完成, 形成了 IEC 61966-2-1 标准。根据这一标准, 显示特性曲线 EOTF 的伽马值为 2.2, 这与传统 CRT 的统计值一致。在 sRGB 的编码过程中, 采用了以下非线性的 OETF 转换方程:

$$V = \begin{cases} 12.92L, & 0 \leq L \leq 0.003\,130\,8 \\ 1.055L^{\frac{1}{2.4}} - 0.055, & 0.003\,130\,8 < L \leq 1 \end{cases} \quad (3-35)$$

与式(3-34)相似,式(3-35)也分为线性和非线性两个部分,线性部分的使用原因相同。

sRGB 的 OETF 特性使用了指数为 $\frac{1}{2.4}$ 幂函数,并通过增加比例因子和偏移量,使得该函数

逼近指数为 $\frac{1}{2.2} \approx 0.45$ 的幂函数。

4. 视频系统的 OOTF

在视频系统中,OOTF 的考量尤为重要。如果采集端采用高清 ITU-R BT. 709 标准 OETF($\gamma \approx 0.5$),而显示端采用 ITU-R BT. 1886 标准的 EOTF($\gamma = 2.4$),则系统的 γ 值为 $0.5 \times 2.4 = 1.2$ 。由于系统 γ 值大于 1,显示端重现的画面相较于原始场景,颜色更加鲜艳,更符合人眼的视觉偏好。

当采集端和显示端都遵循 sRGB 标准时,则 OETF 的 γ 值为 $\frac{1}{2.2}$,而 EOTF 的 γ 值为 2.2,此时系统 γ 为 1,保证了亮度层次与色彩的一致性。若 ITU-R BT. 709 视频内容($\gamma \approx 0.5$)在计算机显示器($\gamma = 2.2$)上显示,则系统 γ 值为 1.1。由于 sRGB 标准广泛应用于各类媒体设备,如计算机显示器,比 1.2 更低的系统 γ 值更适合于高显示亮度、明亮环境、对比度差的典型 sRGB 显示环境(因为 γ 值越小,平均亮度就越高)。



习题 3