Video

6.1 Introduction

After text, image, graphics and audio, the next element to be included in a multimedia presentation is video. Video is a combination of a sequence of images and associated audio. Hence, in most cases concepts and theories applicable for images and audio are also applicable for video. In this chapter, the term "video" is used to mean "motion video"—the media type consisting of images and sound, as in "video-CD" instead of indicating the visual signals sent to the monitor from a processing system as in "video adapter". One of the major difficulties in incorporating video in presentations is to deal with its size. Since it typically consists of a large number of images as well as audio tracks, sizes of video files can span several GBs of disk space. So, compressing video files is essential and only after the advent of efficient compression software and hardware, and powerful processors did the use of video in PC based applications gain momentum. Initially, video was recorded in analog form using analog video cameras. Later on, these were replaced by electronic sensors called Charge Coupled Devices (CCD) in digital video cameras. Color video cameras produce three separate signals corresponding to red, green, and blue components of the incident light using optical splitters. Due to technical problems of transmitting these separate color signals over large distances as in television transmission, another form of the signal called **composite video** was mostly used where the signals were transformed into a luminance (brightness) and chrominance (color) information and combined on a single cable or channel for transmission. Compositing video signals also had another advantage—the option of using chroma sub-sampling whereby color information is reduced to decrease the bandwidth of the transmitted signal. To standardize these parameters, a number of standards like ITU-R (CCIR), ATSC, DVB, and SMPTE have been developed. Each of these systems use different color spaces to represent the luma and chroma information, which can be derived through various linear combinations from *RGB* signals. Often a video-editing software is used to modify and enhance video quality by manipulating adjustable parameters like brightness, contrast, color, frame dimensions, and frame rate, and these can subsequently be stored using specific file formats like AVI, MOV, or MPEG. These topics have been discussed in detail in the following sections.

6.2 Motion Video • •

Motion video (or simply video) is a combination of image and audio. It consists of a set of still images called **frames** displayed to the user one after another at a specific speed, known as the **frame rate** measured in number of frames per second (fps). If displayed fast enough, our eye cannot distinguish the individual frames, but because of persistence of vision, merges the individual frames with each other thereby creating an illusion of motion. The frame rate should range between 20 and 30 for perceiving smooth realistic motion. Audio is added and synchronized with the apparent movement of images. See Fig. 6.1. The recording and editing of sound has long been in the domain of the PC. Doing so with motion video has only recently gained acceptance. This is because of the enormous file size required by video. For example, one second of 24-bit, 640 × 480 mode video and its associated audio requires 27 MB of space, viz., (640 × 480 pixels) × (24 bits/pixel) × (30 frames/second). Thus, a 20-minute clip fills up 32 GB of disk space. Moreover, it requires processing at 30 MB/s. The only solution to this problem is to compress the data, but compression hardware and software were very expensive in the early days of video editing. As a result, video was played in very small-sized windows of 160 × 120 pixels, which occupied only 1/16 th of the total screen. It was only after the advent of the Pentium-II processor coupled with cost reduction of video compression hardware, that full-screen digital video finally became a reality.

Motion video is conceptually similar to but physically different from motion picture. **Motion picture** is recorded on celluloid film and displayed in cinema theaters by projecting on a screen, whereas motion video is represented in the form of electrical signals as an output from video cameras. These signals can either be recorded on magnetic media like video cassettes and played back using a VCP, or transmitted directly as TV broadcast signals and received by a television set. Film playback is usually at 24 fps, while for video signals the playback rate varies from 25 fps for PAL standard to 30 fps for NTSC standard. Motion video is also conceptually similar to **animation**, the difference being that while video represents a sequence of real-world images captured by a movie camera, which depict an event that physically took place in reality, an animation sequence consists of images drawn by artists, using pen or paper, or software. So the events do not depict any real sequence of events taking place in the physical world. Animation is discussed in the next chapter.

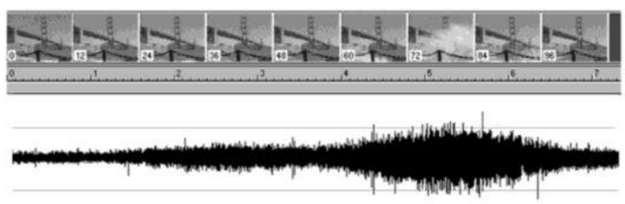


Fig. 6.1 *Video is a collection of frames along with associated sound*

Further Readings

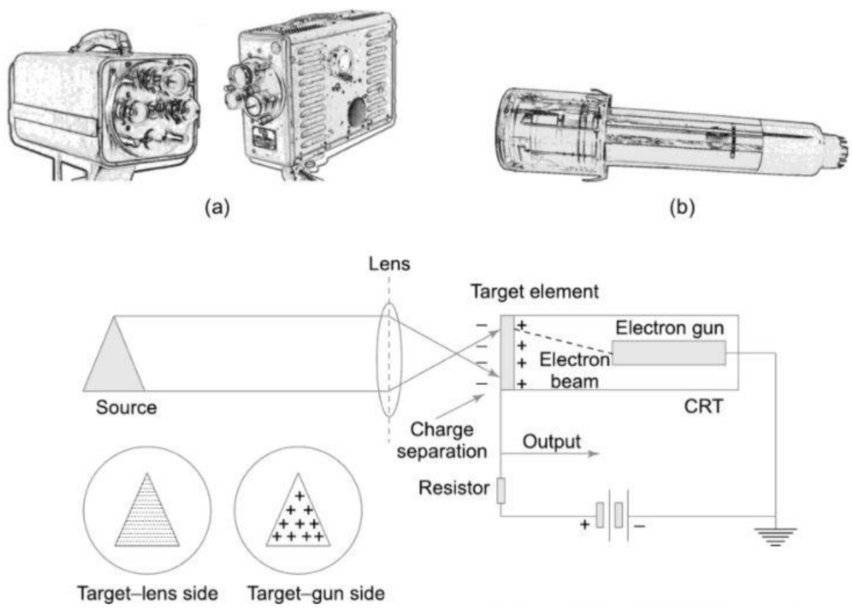
- Programmer's Guide to Video Systems [*http://lurkertech.com/lg/video-systems/*]
- Charles Bensinger, "The Video Guide", 2nd edition, Video-info Publications, 1981 [*http://videopreservation.conservation-us.org/vid_guide/*]

6.3 Analog Video Camera • •

Video recording was initially done using analog equipment. Analog video cameras are used to record a succession of still images and then convert the brightness and color information of the images into electrical signals. These signals are transmitted from one place to another using cables or by wireless means and in the television set at the receiving end, these signals are again converted to form the images. The tube-type analog video camera is generally used in professional studios and uses electron beams to scan in a raster pattern. See Fig. 6.2(a).

6.3.1 Monochrome Video Camera

The essential components of an analog video camera consists of a vacuum tube containing an electron gun, and a photo-sensitive semiconductor plate, called **target element,** in front. See Fig. 6.2(b). A lens in front of the target focuses light from an object on to the target. The positive terminal of a battery is connected to the lens side of the target while the negative terminal is attached to the cathode of the electron gun. See Fig. 6.2(c).



(c)

Fig. 6.2 *(a) Tube-type analog video cameras (b) Vacuum tube (c) Schematic diagram of the tube-type camera*

The target element is photo-sensitive and behaves like an insulator in the absence of light. With absorption of energy caused by light striking the target from the source through the lens, electrons within the target acquire sufficient energy to take part in current flow. The negative electrons migrate towards a positive potential applied to the lens side of the target. The vacant energy states left by the liberated electrons, called holes, migrate towards the electron-gun side of the target. Thus, a **charge pattern** appears on the gun-side of the target that is most positive where the brightness of the source is the greatest. The charge pattern is sampled point-by-point by a moving **electron beam**, which originates from the electron gun in the tube. The beam scans the charge pattern in the same way a raster is produced in a monitor but approaches the target at a very low velocity. The velocity of the electrons are reduced by a **decelerating grid** within the CRT, which ensures that the electrons in the beam do not have sufficient energy to hit the target. However, during the raster-scan process when the beam is directly above the positive charge pattern, the additional **electrostatic attractive force** between the negative electrons and the positive holes imparts sufficient energy to an electron to combine with a hole and neutralize its charge. In other words, the beam deposits just enough carriers to neutralize the charge pattern formed by the holes, while excess electrons are turned away from the target.

It is well known that in an electrical circuit, the total number of electrons that start from a current source must return back to it, i.e., charge cannot be accumulated in circuit elements like a resistor. In this case too, the exact number of electrons that combines with the holes within the CRT must flow out in the external circuit and back to the battery, since the battery and CRT are connected in series. This current flowing across a load resistance forms the **output signal voltage** of the tube. Variants of essentially the same principle are known by various names registered by respective manufacturing companies, viz., orthicon, vidicon, plumbicon, saticon, trinicon, etc.

Further Readings

- Vidicon @ Britannica Online Encyclopedia [*http://www.britannica.com/EBchecked/topic-art/627966/60165/Vidicon-television-camera-tube-The-image-of-the-scene-is*]
- Camera Tubes @ The Cathode Ray Tube site [*http://members.chello.nl/~h.dijkstra19/page4.html*]
- TV Camera tubes [*http://www.earlytelevision.org/camera_tubes.html*]
- RCA TV equipment archive [*http://www.oldradio.com/archives/hardware/TV/RCA-TV.htm*]
- Museum of the Broadcast TV Camera [*http://www.tvcameramuseum.org/*]
- Broadcast cameras [*http://www.marcelstvmuseum.com/photoalbum2.html*]
- Museum of Extinct Video Cameras [*http://www.labguysworld.com/VTR-Museum_002.htm*]
- Video Camera Tube @ Wikipedia [*http://en.wikipedia.org/wiki/Video_camera_tube*]

6.3.2 Color Video Camera

In a tube-type video camera, the output voltage signal is proportional to the intensity of the light falling on the target element. Greater the light intensity, more is the charge separation and larger will be the current flowing across the external resistor. This signal when fed to a monitor will generate a grey-scale image where bright portions of the source will be reflected by white areas and dark portions by black areas. To capture the color of the source, the white light from the source is split into red, green, and blue components using optical splitters like glass prisms or dichroic mirrors. A **dichroic mirror** passes one wavelength and rejects other wavelengths. The primary colored lights pass through color filters to form highly precise primary color images which are converted into video signals by a system of three camera tubes, one for each primary component. See Fig. 6.3. Each camera tube generates a signal voltage proportional to the respective color intensity received by it. This generates the three color signals—the *R* signal, the *G* signal and the *B* signal, the voltage levels of which

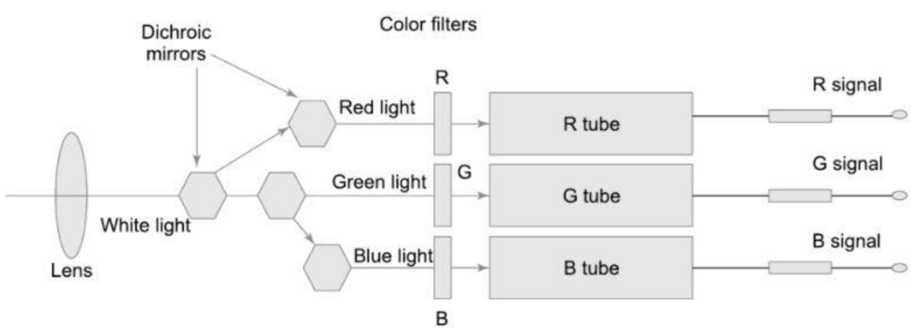


Fig. 6.3 *Analog signal outputs from a color video camera*

are proportional to the intensity of the corresponding colored light falling on the specific tube. To form an image on a screen, the three color signals are fed to a color monitor, where each signal activates the corresponding electron gun of the monitor to generate corresponding colored lights. The combination of the three colored lights regenerates the color image on the monitor screen.

Further Readings

- Charles Bensinger, "The Video Guide", chap1 [*http://videopreservation.conservation-us.org/vid_guide/1/1.html*]
- Charles Bensinger, "The Video Guide", chap2 [*http://videopreservation.conservation-us.org/vid_guide/2/2.html*]
- Charles Bensinger, "The Video Guide", chap3 [*http://videopreservation.conservation-us.org/vid_guide/3/3.html*]
- Charles Bensinger, "The Video Guide", chap4 [*http://videopreservation.conservation-us.org/vid_guide/4/4.html*]

6.4 Analog Video Signal Representation • •

6.4.1 Problems in Transmitting Color Signals

We have seen above that a color video camera produces three color signals corresponding to the *R*, *G*, *B* components of the color image. These signals must be combined in a monitor to reproduce the original image. Such a scheme is suitable when the monitor is close to the camera, and three cables could be used to transmit the signals from the camera to the monitor. However, when the displaying monitor is at a far-off distance from the camera, for example, in the case of TV transmission, technical difficulties are encountered in transmitting these signals in the original *R*, *G*, *B* format. First, it requires three **separate cables** or wires or channels, which increases the cost of the set-up for large distances involving several kilometers. Second, it was found to be difficult to transmit the signals at **exact synchronism** with each other so that they arrive at exactly the same instant at the receiving end. Usually, a small amount of phase difference was introduced as the signals travelled along the transmission lines, which produced distortions in the image at the receiver. Third, for TV signals, the transmission scheme had to **adapt** to the existing monochrome TV transmission set up, i.e., the same signals would need to produce a monochrome image on a B/W (black and white) TV set and a color image on a color TV set. Due to these reasons, instead of the *RGB* format, color signal transmission is done using a new format called *YC* format to address all these issues. *Y* stands for the luminance or brightness part of the information, while *C* stands for the chrominance or color part of the information. The principle of the *YC* format was based on the human visual perception of color and is described below.

Further Readings

- Video Basics [*http http://pdfserv.maxim-ic.com/en/an/AN734.pdf*]

6.4.2 Color-perception Curve

All objects that we observe are focused sharply by the lens system of the human eye on the retina. The **retina** which is located at the back of the eye has light-sensitive cells which capture the visual sensations. The retina is connected to the optic nerve that conducts the light stimuli to the optical centre of the brain. According to the theory formulated by Helmholtz, the light-sensitive cells are of two types—rods and cones. The **rod cells** provide brightness sensation and thus perceive objects in various shades of gray, from black to white. The **cone cells** are sensitive to color and can broadly be classified into three different groups. One set of cones detect the presence of blue color, the second set perceives red color and the third is sensitive to

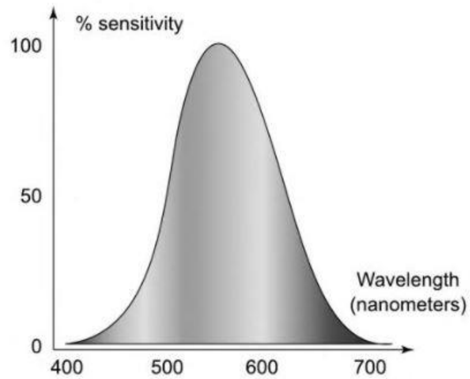


Fig. 6.4 *Human color-perception curve*

the green shade. The combined relative luminosity curve showing relative sensation of brightness is produced by individual spectral colors. See Fig. 6.4. It is seen from the plot that the **sensitivity** of the human eye is greatest for the green-yellow range decreasing towards both the red and blue ends of the spectrum. Any color other than red, green, and blue excites different sets of cones to generate a cumulative sensation of that color. White color is perceived by the additive mixing of the sensations from all the three sets of cones. Based on the spectral response curve and extensive tests with a large number of observers, the relative intensities of the primary colors for color transmission, e.g., for color television, has been standardized. The reference white for color television transmission has been chosen to be a mixture of 30% red, 59% green and 11% blue. These percentages are based on the light sensitivities of the eye to different colors. Thus, one **lumen** (lm) of white light = 0.3 lm of red + 0.59 lm of green + 0.11 lm of blue = 0.89 lm of yellow + 0.11 lm of blue = 0.7 lm of cyan + 0.3 lm of red = 0.41 lm of magenta + 0.59 lm of green.

6.4.3 Luminance and Chrominance

The *RGB* model is used mainly in color image acquisition and display. In color signal processing, including image and video compression, however, the luminance–chrominance color system is more efficient and hence widely used. This has something to do with color perception of the **Human Visual System** (HVS). It is known that the HVS is more sensitive to green than red and the least sensitive to blue. An equal representation of red, green, and blue leads to inefficient data representation when the HVS is the ultimate viewer. Allocating data only to the information that the HVS can perceive can make video coding more efficient. Thus, video signals are expressed in terms of luminance–chrominance instead of *RGB*. The **luminance** component, describes the variation of perceived brightness by the HVS in different portions of the image without regard to any color information, e.g., an image with only the luminance information would be a grayscale image similar to that seen on a monochrome TV set. The luminance component is denoted by *Y*. In colorimetry, luminance indicates the brightness of a light source. Brightness is defined by the International Commission on Illumination (CIE) as the attribute of visual sensation according to which an area appears to emit more or less light. Because the brightness perception is quite complex, the CIE defined a more measurable quantity called luminance *Y* whose magnitude is proportional to power (radiance) of the light source and hence it is similar to intensity. The unit for luminance is actually candela per square meter (cd/m^2) but it is often normalized to the range 1 to 100 with respect to the luminance of a **white reference**. For example, a studio monitor has a white reference of 80 cd/m^2 and $Y = 1$ refers to this. Human vision has a non-linear response to brightness: a source having a luminance of 18% of a reference luminance appears to have a brightness of 50%. The perceptual response to luminance is called **lightness**. It is denoted as L^* and is defined by CIE as in equation (6.1):

$$L^* = 116\left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} - 16, \quad \frac{Y}{Y_n} > 0.008856 \tag{6.1}$$

Here, Y_n is the luminance of the white reference. An **absolute color space** is defined in terms of colors which are colorimetric parameters without being dependent on changing external factors, e.g., CIE *XYZ* and *sRGB* are absolute color spaces while *RGB* is a generic color space.

The **chrominance** component describes the variation of color information in different parts of the image without regard to any brightness information. It is denoted by *C* and consists of two sub-components: **hue** (*H*) which is the actual name of the color, e.g., red, and **saturation** (*S*) which denotes the purity of the color, i.e., how much gray is mixed with the original color, e.g., bright red, dull red, etc. An image with only the *C* component would consist of flat areas of color without any brightness/contrast variations, i.e., no shading. Chrominance is defined by CIE in terms of **chromacity diagram** and *R*, *G*, *B* primaries. See Chapter 3 for further details.

An image may therefore be thought to be composed of two separate portions, a luminance component and a chrominance component, which when superimposed on each other produce the final image that we see. See Fig. 6.5. The lighter portions in the luminance component indicate regions with higher brightness, while the lighter portions in the chrominance component indicate regions with higher saturation of colors.

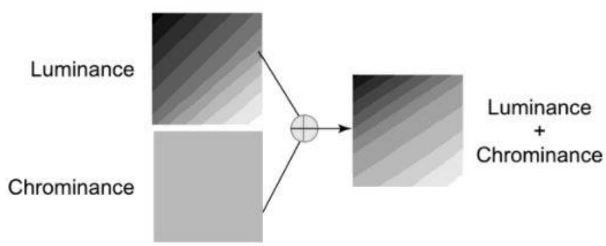


Fig. 6.5 *Luminance and chrominance components (Refer Plate 8)*

6.4.4 Luma and Chroma

When dealing with video signals, luminance is used to denote the brightness of the video image. If three colored light sources (red, green, blue) have the same power or radiance then green will appear the brightest of the three, red will appear less bright and blue the darkest, because of the luminous sensitivity curve as shown in Fig. 6.4. Luminance is, therefore, calculated as a weighted sum of *R, G, B* values. As per ITU-T Rec. 601 (1982), luminance *Y* is related to *R, G, B* values according to equation (6.2):

$$Y = (0.299)R + (0.587)G + (0.114)B \tag{6.2}$$

To take into consideration the **non-linear response** of cathode ray tubes, in practice, a non-linear transformation called **gamma correction** is applied to each of the *R*, *G*, *B* values and then a weighted sum of the non-linear components is computed to form a non-linear representative of luminance. The resulting component is called **luma** and, to differentiate it from luminance, it is denoted as Y' (Y-prime). Similarly, the non-linear (gamma compensated) components of R, G, B are denoted with prime symbols R', G', B' to differentiate them from linear components without prime symbols. According to ITU-T Rec. 601, luma is related to non-linear R', G', B' as per equation (6.3):

$$Y' = (0.299)R' + (0.587)G' + (0.114)B' \tag{6.3}$$

The color information of the video information is transmitted separately from the luminance information. Color is represented as two **color-difference** components each having no contribution to brightness. These color-difference components are called **chroma**, and are designated as C_b and C_r. The easiest way to remove brightness information from the color information is to subtract the former from the latter, as given in equation (6.4).

$$\begin{aligned} C_b &= B' - Y' \\ C_r &= R' - Y' \end{aligned} \tag{6.4}$$

Note 6.1: When referring to luma and chroma, as in equations (6.3) and (6.4), in most cases the prime symbols are omitted. Since no practical image-coding systems work with linear components, it might be safe to write the components without the prime symbols; however, it should be noted that these actually represent the non-linear gamma-compensated components and not the linear components as given by equation (6.2). In subsequent sections for the sake of simplification, the non-linear components will be denoted without the prime symbols. Linear components, if necessary, will be explicitly stated.

From equations (6.3) and (6.4), omitting the prime symbols, we have

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B \\ C_b &= B - Y \\ C_r &= R - Y \end{aligned} \tag{6.5}$$

From equation (6.5),

$$\begin{aligned} C_b &= B - (0.299R + 0.587G + 0.114B) = -\,0.299R - 0.587G + 0.886B \\ C_r &= R - (0.299R + 0.587G + 0.114B) = 0.701R - 0.587G - 0.114B \end{aligned}$$

From the above, the conversion from RGB values to YC values are defined in equation (6.6):

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B \\ C_b &= -\,0.299R - 0.587G + 0.886B \\ C_r &= 0.701R - 0.587G - 0.114B \end{aligned} \tag{6.6}$$

From equation (6.5), we get $R = C_r + Y$ and $B = C_b + Y$. Putting the values of R and B in the expression for Y, we get an expression for the third component G:

$$\begin{aligned} Y &= 0.299(C_r + Y) + 0.587G + 0.144(C_b + Y) \\ \rightarrow G &= Y - 0.5093C_r - 0.1942C_b \end{aligned}$$

Combining the values of R, G, B in terms of Y, C_b, C_r we obtain the inverse relations of RGB values from YC values as follows.

$$R = Y + C_r$$
$$G = Y - 0.5093C_r - 0.1942C_b$$
$$B = Y + C_b \tag{6.7}$$

Also note that the color-difference signals equal zero when white or gray shades are being transmitted. This is illustrated by the calculation below:

For any gray shade (including white), let $R = G = B = v$ volts. Then $Y = 0.299\ v + 0.587\ v + 0.144\ v = v$. Thus, $(R - Y) = v - v = 0$ volt, and $(B - Y) = v - v = 0$ volt.

Further Readings

- Color FAQ [*http://www.poynton.com/notes/colour_and_gamma/ColorFAQ.html*]

6.4.5 Generating *YC* signals from *RGB*

The *RGB* output signals from a video camera are transformed to *YC* format using electronic circuitry before being transmitted. At the receiving end for a B/W TV, the *C* component is discarded and only the *Y* component is used to display a grayscale image. For a color TV, the *YC* components are again converted back to *RGB* signals which are used to drive the electron guns of a CRT. See Fig. 6.6.

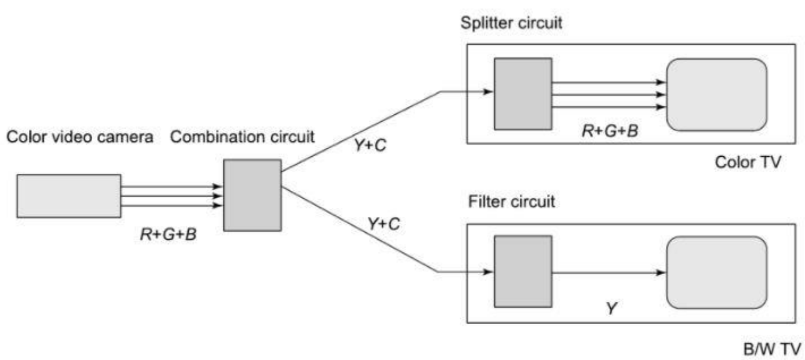


Fig. 6.6 *Transmission of TV signals in the YC format*

The combination circuit in Fig. 6.6 uses equation (6.6) to convert *R, G, B* signals to Y, C_b, C_r format. The splitter circuit at the receiving end of a color TV uses equation (6.7) to convert *YC* signals to *RGB* signals. The filter circuit of a B/W TV set only uses the *Y* value of equation (6.6) to generate a grayscale image, while discarding the C_b and C_r values.

Physically, the *Y* signal is obtained by passing the *R, G, B* current signals from the respective camera tubes through a resistor bridge. In the arrangement shown in Fig. 6.7, an impedance of 30 kΩ is connected in parallel to each of the signals *R, G* and *B* output from the tube camera. Hence, the output is a combination of {30/(70 + 30)}*R*, {30/(20 + 30)}*G*, {30/(270 + 30)}*B*, i.e., 0.3*R* + 0.6*G* + 0.1*B* approximately.

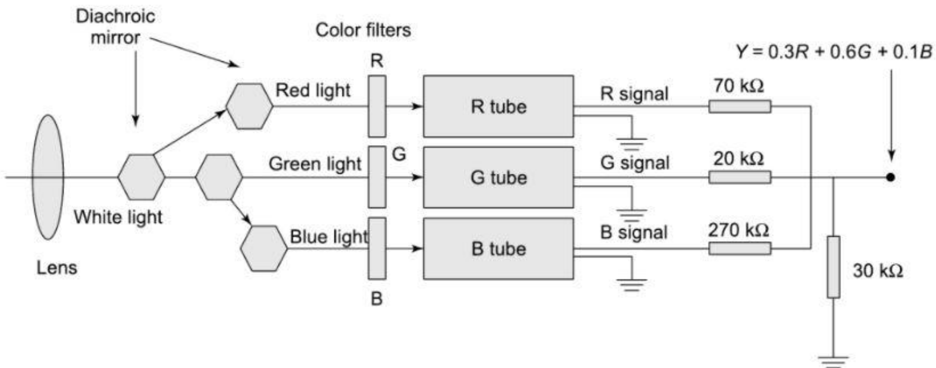


Fig. 6.7 *Generation of Y from R, G, B signals*

The color-difference signals are generated by inverting Y using an inverter and adding the inverted signal separately to R and B to obtain $(R - Y)$ and $(B - Y)$ as shown in Fig. 6.8.

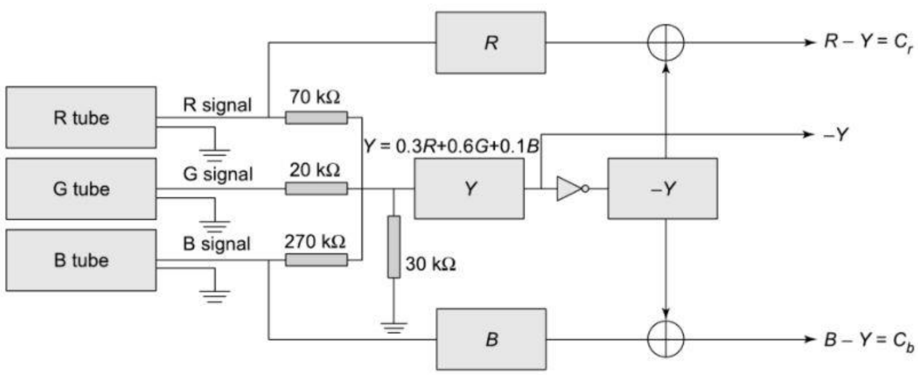


Fig. 6.8 *Generation of C_r and C_b from Y, R, G, B signals*

6.4.6 Chroma Sub-sampling

Conversion of *RGB* signals into *YC* format also has another important advantage of utilizing less bandwidth through the use of chroma sub-sampling. Studies on visual perception of the eye have shown that the human eye is less sensitive to color information than to brightness information. This means that small differences in color information are ignored by the eye. This limitation can be exploited to transmit reduced color information as compared to brightness information, a process called **chroma sub-sampling**, and save on bandwidth requirements. Chroma sub-sampling is indicated by a set of numbers that denote the amount of luminance and chrominance information transmitted from the video camera to the TV receiver set. It specifies that when the signal is converted into an image on the TV screen, out of a set of pixels containing luminance information, how many corresponding pixels contain chrominance information. The sub-sampling scheme is commonly expressed as a **three-part ratio,** viz., *A:B:C*. The numbers describe the number of luminance and chrominance samples in a window region *W* on the screen that is *A* pixels wide and 2 pixels high. The first number *A* denotes the number of pixels containing luminance information, along the first row of *W*, and is usually 4. The second number *B* denotes the number of pixels containing chrominance information along the first row of *W*, and the third number *C* denotes the number of pixels containing chrominance information along the second row of *W*. The window *W* itself can be slid over the screen without changing the meaning of the numbers. See Fig. 6.9. The reduction in color information helps reduce bandwidth of the transmitted signal. There can be different schemes of chroma sub-sampling as described below.

4:2:2 Indicates that within a sliding window *W* on the screen 4 pixels wide by 2 pixels high, there are 4 pixels containing *Y* information along its first row, 2 pixels of *C* information along its first row, and 2 pixels of *C* information along its second row. Essentially, this means that while all pixels contain brightness information, only half of the pixels contain color information. See Fig. 6.9(b).

4:1:1 Indicates that within a sliding window *W* on the screen 4 pixels wide by 2 pixels high, there are 4 pixels containing *Y* information along its first row, 1 pixel of *C* information along its first row, and 1 pixel of *C* information along its second row. Essentially, this means that while all pixels contain brightness information only one-fourth of the pixels contain color information. See Fig. 6.9(a).

4:4:4 Indicates that within a sliding window *W* on the screen 4 pixels wide by 2 pixels high, there are 4 pixels containing *Y* information along its first row, 4 pixels of *C* information along its first row, and 4 pixels of *C* information along its second row. Essentially, this means that all pixels contain brightness and color information and hence no reduction in color information. See Fig. 6.9(c).

4:2:0 Indicates that within a sliding window *W* on the screen 4 pixels wide by 2 pixels high, there are 4 pixels containing *Y* information along its first row, 2 pixels of *C* information along its first row, and 0 pixel of *C* information along its second row. Essentially, this means that while all pixels contain brightness information only one-fourth of the pixels contain color information. See Fig. 6.9(d). Note that the amount of information loss is double that of the 4:2:2 scheme and comparable to the 4:1:1 scheme.

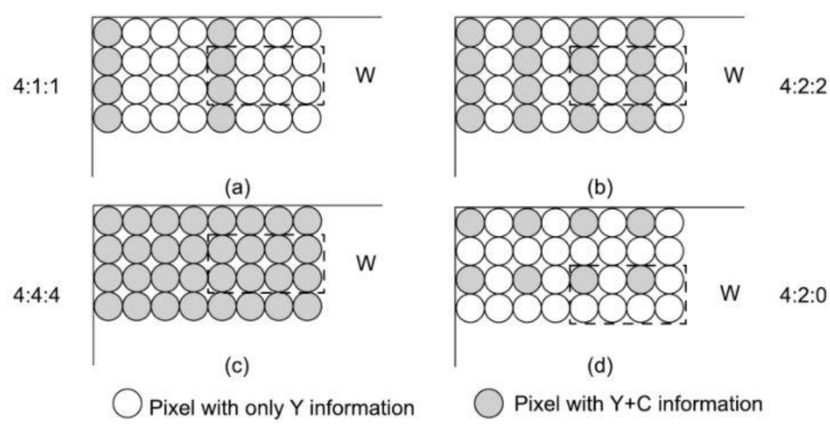


Fig. 6.9 *Chroma sub-sampling schemes: (a) 4:1:1 (b) 4:2:2 (c) 4:4:4 (d) 4:2:0*

6.4.7 Video Connectors

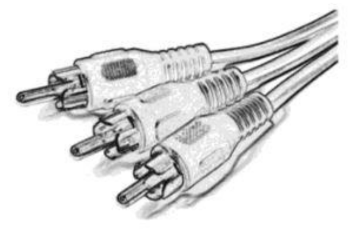


Fig. 6.10 *Component video connectors*

Component Video

This refers to a video signal that is stored or transmitted as three separate component signals. The simplest form is the collection of *R, G* and *B* signals that usually form the output of analog video cameras. Three separate wires and connectors are usually used to convey such signals from the camera to another device for storage or playback. In situations where the same set of signals are required to drive both a color monitor as well as a B/W monitor, as also to reduce signal bandwidth, *R, G, B* signals are replaced by Y, P_b, P_r signals, also delivered along three separate wires. The connectors used are typically RCA connectors, typically colored green (Y), blue (P_b) and red (P_r). See Fig. 6.10. The conversion between *RGB* and Y, P_b, P_r formats is governed by equations (6.11) and (6.12).

Fig. 6.11 *Composite video connector*

Composite Video

For ease in signal transmission, specially TV broadcasting, as also to reduce cable/channel requirements, component signals are often combined into a single signal which is transmitted along a single wire or channel. This is referred to as composite video. Usually, the luma–chroma format, Y, C_b, C_r is used for composite video transmission when dealing over short distances, although YIQ or YUV, depending on whether the broadcasting standard is NTSC or PAL respectively, when sent over a single channel can also be regarded as composite video. In this case, the total bandwidth of the channel is split into separate portions and allotted for the luminance and chrominance parts. Since the human eye is more sensitive to luminance changes than color changes, luminance is allotted a greater bandwidth than the chrominance parts. Composite video is transferred from one device to another through a single RCA jack, usually colored yellow, e.g., from a VCR/VCP to a TV. In some cases, a SCART or BNC connector is used instead. Since a single wire carries different types of signals, a certain amount of crosstalk or interference is introduced which leads to a slight degradation in video quality compared to component format. However, since a single cable is used, this leads to cost savings. See Fig. 6.11.

S-Video

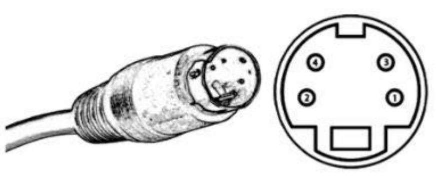


Fig. 6.12 *S-video connector and pins*

Super-video (S-video) is an analog video signal format where the luminance and chrominance portions are transmitted separately using multiple wires instead of the same wire as for composite video. The picture quality is better than that of composite video because of reduced interference but the cable is more expensive, and is usually found in high-end VCRs and capture cards. The connector used is a 4-pin mini-DIN connector with 75-ohm termination impedance. See Fig. 6.12. The pin assignments are as follows:

1	2	3	4
Y Ground (GY)	C Ground (GC)	Luminance (Y)	Chrominance (C)

SCART

SCART (Syndicat des Constructeurs d'Appareils Radiorecepteurs et Televiseurs) is a French standard of a 21-pin audio and video connector. It can be used to connect VCRs, DVD players, set-top boxes, game systems and computers to television sets. SCART attempts to provide a standardized connector containing all the signals for audio and video applications across different manufacturers. SCART compatible devices have multiple connectors that can be used for daisy-chaining purposes. The signal levels are around 1 volt, and so they are not much influenced by noise. It also supports bi-directional communication between connected devices. Some of the drawbacks of SCART are that it cannot carry both S-video and *RGB* signals at the same time, it cannot transmit surround sound formats and can only transmit analog signals, not digital. It has a few physical drawbacks too. SCART connectors are non-locking and may become loose or fall off, cheap SCART connectors are fragile and prone to broken pins, and maximum cable length is 10 m to 15 m. Properly manufactured SCART connectors use coaxial cables to transmit audio/video signals, however cheaper versions may use plain wires resulting in degraded image quality. See Fig. 6.13.

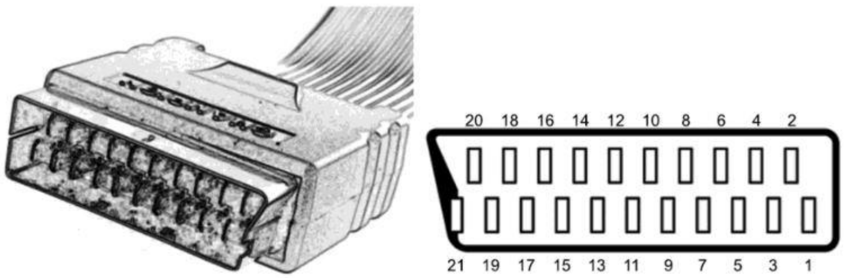


Fig. 6.13 *SCART connector and pins*

6.5 Television Systems • •

6.5.1 NTSC

NTSC (National Television Systems Committee) is a television broadcasting system used in a number of countries including Korea, Japan, Canada, North America, parts of South America, Mexico, and the Caribbean islands, created by an industry-wide standardization body. It was set up in 1940 by the Federal Communication Commission (FCC) in the United States to establish a nationwide standard for black-and-white TV transmission. It specified a standard using 525 horizontal lines, 30 frames per second, 2 interlaced fields per frame (also called 2:1 interlacing), 262.5 lines per field, 60 fields per second, an aspect ratio of 4:3 and frequency modulation for the audio signal (frequency modulation is a form of modulation which represents information as variations in the instantaneous frequency of a carrier wave). In 1950, the Committee was reconstituted to establish a standard for color TV transmission, which would be at the same time compatible with the existing format for B/W TV. As per the recommendations of the committee, color information was added to the black-and-white image by adding a color sub-carrier to the video signal. (A **sub-carrier** is a separate analog or digital signal carried on a main transmission signal, which contains some extra information. TV signals are transmitted with the luminance part as the main signal and the chrominance part as the sub-carrier.) Due to certain technical considerations, the addition of the color sub-carrier also required a slight reduction in the frame rate from 30 fps to 29.97 fps. Due to this, a discrepancy between the real time and the time stamped on the video was introduced, leading to the concept of **Non-drop frame mode** and **Drop-frame mode**. NTSC uses a chroma sub-sampling scheme of 4:2:2. Although, each frame in an NTSC video consists of 525 horizontal lines, however, only 480 of them are actively used for generating the picture on the screen, the rest being used for synchronization and vertical retrace. In color transmission, the luminance signal *Y* takes the place of the original monochrome signal used for B/W TVs. It is derived from the gamma-corrected red, green and blue signals, as given by equation (6.6) and reproduced below.

$$Y = 0.299R + 0.587G + 0.114B$$

Since the human eye is most sensitive to variations in luminance, the *Y* signal is allowed a bandwidth of 0–4.2 MHz. The chrominance part of the signal is represented by two 3.58 MHz sub-components known as *I* (in-phase) and Q (quadrature). The *I* and Q signals are amplitude modulated in quadrature (QAM) onto the sub-carrier wave. The *I* signal is in phase with the carrier wave and the *Q* is in quadrature, i.e., 90 degrees out of phase with the sub-carrier signal. I physically represents the orange-cyan axis (positive *I* is orange, negative *I* is cyan) and is bandlimited to 1.3 MHz while *Q* physically represents the magenta-green axis (positive *Q* is magenta, negative *Q* is green) and its bandwidth is limited to 0.6 MHz. The hue part of a given point on the screen is defined by the phase of the sub-carrier at that point. Since *I* and *Q* are clearly phase-sensitive, some sort of phase reference must be supplied. This reference is supplied after each horizontal scan and included at the back portion (called "back porch") of the horizontal sync pulse, and is known as the "color burst".

I and *Q* can be defined in terms of *R, G* and *B* (from equation 6.5) as shown in equation (6.8):

$$I = 0.74C_r - 0.27C_b = 0.74(R - Y) - 0.27(B - Y) = 0.597R - 0.274G - 0.321B$$
$$Q = 0.48C_r + 0.41C_b = 0.48(R - Y) + 0.41(B - Y) = 0.211R - 0.523G + 0.311B \tag{6.8}$$

NTSC has the following variants:

- NTSC-M uses 525 lines/frame, 29.97 frames per second and a color sub-carrier of 4.43 MHz, instead of 3.58 MHz. NTSC-M is also called NTSC 4.43.
- NTSC-J used in Japan is same as NTSC-M except for the fact that in NTSC-J, the black level and blanking levels are identical, while in NTSC-M, the black level is slightly higher than blanking level.

6.5.2 PAL

PAL (Phase Alternation Lines—sometimes the terms Alternate or Alternating are also used) is a TV broadcasting standard used in Europe, Asia, Africa, and Australia. PAL was developed by Walter Bruch at Telefunken in Germany, and first introduced in 1967. French electronics manufacturer Thomson later bought Telefunken. The PAL system uses 625 horizontal lines at 25 frames per second, 2:1 interlacing with 2 fields per frame, 4:3 aspect ratio and a chroma sub-sampling scheme of 4:2:2. Similar to NTSC, PAL also uses a luminance component *Y* derived from equation (6.6) and reproduced below:

$$Y = 0.299R + 0.587G + 0.114B$$

The chrominance part of the signal is represented by two 4.43 MHz sub-components known as *U* (in-phase) and *V* (quadrature). The *U* and *V* signals are amplitude modulated in quadrature onto the sub-carrier wave. Both *U* and *V* are bandlimited to 1.3 MHz each, while the composite video signal may have bandwidths of 4.2, 5, 5,5 or 6 MHz depending on the specific standard used.

U and *V* can be defined in terms of *R, G* and *B* (from equation 6.5) as shown in equation (6.9):

$$U = 0.492C_b = 0.492(B - Y) = -0.147R - 0.289G + 0.436B$$
$$V = 0.877C_r = 0.877(R - Y) = 0.615R - 0.515G - 0.1B \tag{6.9}$$

PAL has a number of variants: PAL-B, PAL-G, PAL-I, PAL-D, PAL-N, PAL-Nc all use the 625/50 systems with active lines 576 (576i). The differences are as follows:

- PAL-B uses video bandwidth of 5 MHz, sound carrier at 5.5 MHz and channel bandwidth of 7 MHz.
- PAL-G uses video bandwidth of 5 MHz, sound carrier at 5.5 MHz and channel bandwidth of 8 MHz.
- PAL-I uses video bandwidth of 5.5 MHz, sound carrier at 6 MHz and channel bandwidth of 8 MHz.
- PAL-D uses video bandwidth of 6 MHz, sound carrier at 6.5 MHz and channel bandwidth of 8 MHz.
- PAL-N uses video bandwidth of 5 MHz, sound carrier at 5.5 MHz and channel bandwidth of 6 MHz.
- PAL-Nc uses video bandwidth of 4.2 MHz, sound carrier at 4.5 MHz and channel bandwidth of 6 MHz.
- PAL-M uses video bandwidth of 4.2 MHz, sound carrier at 4.5 MHz and channel bandwidth of 6 MHz, but the 525/60 system with active lines 480 (480i).
- In recent years, a variant called PAL-60 has been developed which uses the color coding of PAL but the number of scan lines (525) and refresh rate of NTSC (60). Strictly speaking, however, the refresh rate is not 60 Hz but 59.94 Hz.

Note 6.2: During signal propagation in NTSC, the phase relationships associated with the color signals are liable to drift, resulting in incorrect hues at the receiver set. This cannot be controlled automatically, so a manual "hue" control knob is included in NTSC TV sets. In addition, crosstalk between the two color signals resulting from poor separation during decoding also leads to incorrect hues. For this reason, NTSC is sometimes jokingly referred to as "Never The Same Color". In PAL, the name "phase alternating lines" describes the way that the phase part of the color information is reversed with each line, which automatically helps correct phase errors in the transmission of the signal by canceling them out. PAL is, therefore, sometimes jokingly referred to as "Perfect At Last".

6.5.3 SECAM

SECAM (Sequential Couleur Avec Memoire), French for "Sequential Color with Memory", is a TV broadcasting standard used in France, Russia, and the Middle East. A team led by Henri de France, working at Compagnie Francaise de Television (later bought by Thomson) invented SECAM. Like PAL, SECAM also uses 625 horizontal lines per frame and 25 frames per second. However, the fundamental difference between NTSC/PAL and SECAM is that while the former transmits two color signals simultaneously, the latter transmits only one color difference signal at a time. It gets the information about the other color difference signal from the preceding line transmitted, which is stored in the memory inside the receiving set. Because of this, SECAM is free of the color artifacts present in NTSC and PAL. Another notable difference between NTSC/PAL and SECAM is that while the former uses amplitude modulation to encode color information, SECAM uses frequency modulation to encode chrominance information onto the sub-carrier. Similar to NTSC and PAL, SECAM also uses a luminance component Y derived from equation (6.6). The chrominance part of the signal is represented by two sub-components D_b and D_r, which are bandlimited to 1.3 MHz each, while the composite video signal may have bandwidths of 5 or 6 MHz depending on the specific standard used. D_b and D_r can be defined in terms of R, G and B (from equation 6.5) as shown in equation (6.10):

$$\begin{aligned} D_b &= 1.5C_b = 1.5(B - Y) = -\,0.450R - 0.833G + 1.333B \\ D_r &= -1.9C_r = -1.9(R - Y) = -\,1.333R + 1.116G + 0.271B \end{aligned} \tag{6.10}$$

6.6 Video Color Spaces • •

6.6.1 YP_bP_r

YP_bP_r is a color space used in connection with analog component video. The first row represents luma coefficients and these sum to unity, while the second and third rows represent color difference coefficients and these sum to zero.

Conversion from RGB to YP_bP_r

$$\begin{bmatrix} Y \\ P_b \\ P_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.5 \\ 0.5 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{6.11}$$

Conversion from YP_bP_r to RGB

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.402 \\ 1 & -0.344 & -0.714 \\ 1 & 1.772 & 0 \end{bmatrix} \begin{bmatrix} Y \\ P_b \\ P_r \end{bmatrix} \tag{6.12}$$

Range of values: $0 \le R, G, B \le 1$, $0 \le Y \le 1$, $-0.5 \le P_b, P_r \le +0.5$

6.6.2 YC_bC_r

YC_bC_r is a color space used in connection with analog composite video. Like YP_bP_r, the first row represents luma coefficients and sum to unity, while the second and third rows represent color difference coefficients and these sum to zero.

Conversion from RGB to YC_bC_r

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.299 & -0.587 & 0.886 \\ 0.701 & -0.587 & -0.114 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{6.13}$$

Conversion from YC_bC_r to RGB

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -0.1942 & -0.5094 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} \tag{6.14}$$

Range of values: $0 \le R, G, B \le 255$, $0 \le Y \le 1$, $-0.886 \le C_b \le +0.886$, $-0.7 \le C_r \le +0.7$

In order to normalize values of R, G, B, Y, C_b, C_r to the range [0, 255] for digital video applications, an alternative form of conversion is used:

Conversion from RGB to YC_bC_r (Digital)

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.5 \\ 0.5 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{6.15}$$

Conversion from YC_bC_r to RGB (Digital)

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.4 \\ 1 & -0.343 & -0.711 \\ 1 & 1.765 & 0 \end{bmatrix} \begin{bmatrix} Y \\ C_b - 128 \\ C_r - 128 \end{bmatrix} \tag{6.16}$$

Range of values: $0 \le R, G, B \le 255$, $0 \le Y, C_b, C_r \le 255$

6.6.3 YUV

The *YUV* color space is used in connection with PAL television broadcasting standard.

Conversion from RGB to YUV

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.1 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{6.17}$$

Conversion from YUV to RGB

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.14 \\ 1 & -0.395 & -0.581 \\ 1 & 2.032 & 0 \end{bmatrix} \begin{bmatrix} Y \\ U \\ V \end{bmatrix} \tag{6.18}$$

Range of values: $0 \le R, G, B \le 1$, $0 \le Y \le 1$, $-0.436 \le U \le +0.436$, $-0.615 \le V \le +0.615$

6.6.4 YIQ

The *YIQ* color space is used in connection with NTSC television broadcasting standard.

Conversion from RGB to YIQ

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.321 \\ 0.211 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{6.19}$$

Conversion from YIQ to RGB

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0.956 & 0.621 \\ 1 & -0.272 & -0.647 \\ 1 & -1.107 & 1.705 \end{bmatrix} \begin{bmatrix} Y \\ I \\ Q \end{bmatrix} \tag{6.20}$$

Range of values: $0 \le R, G, B \le 1$, $0 \le Y \le 1$, $-0.596 \le I \le +0.596$, $-0.523 \le Q \le +0.523$

6.6.5 YD_bD_r

The YD_bD_r color space is used in connection with SECAM television broadcasting standard.

Conversion RGB to YD_bD_r

$$\begin{bmatrix} Y \\ D_b \\ D_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.450 & -0.883 & 1.333 \\ -1.333 & 1.116 & 0.217 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{6.21}$$

Conversion from YD_bD_r to RGB

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0.000092304 & -0.525912631 \\ 1 & -0.129132899 & 0.267899328 \\ 1 & 0.664679059 & -0.000079203 \end{bmatrix} \begin{bmatrix} Y \\ D_b \\ D_r \end{bmatrix} \tag{6.22}$$

Range of values: $0 \le R, G, B \le 1$, $0 \le Y \le 1$, $-1.333 \le D_b \le +1.333$, $-1.333 \le D_r \le +1.333$

Solved Examples

Example 6.1 **Consider a TV camera where the maximum intensity of a color signal is represented by 1 volt. An unsaturated magenta signal is formed by mixing 70% *R*, 20% *G* and 60% *B*. What is the luminance output voltage for the signal? What would this value if the magenta color is saturated.**

Maximum output value of a color signal is assumed to be 1 volt.

For unsaturated magenta, R = 0.7 volt, G = 0.2 volt, B = 0.6 volt.

From equation (6.2), Y = 0.3(0.7) + 0.59(0.2) + 0.11(0.6) = **0.394 volt**.

Since 0.5 volt represents a middle gray shade, the unsaturated magenta color would be represented by a dark gray shade (39.4% gray)

Saturated magenta consists of 100% red and 100% blue, i.e., R = 1 volt, B = 1 volt.

From equation (6.2), Y = 0.3(1) + 0.1(1) = **0.4 volt** (40% gray)

Example 6.2 **Calculate the bit rate and memory required to store a 1-hour movie in NTSC formats, assuming 4:2:2 chroma-subsampling, according to CCIR-601 recommendations.**

NTSC format has 720 × 480 pixels at 30 frames/second.

Since sub-sampling scheme is 4:2:2, in one frame there are 720 × 480 pixels of *Y* and 360 × 480 pixels each of C_b and C_r.

As per CCIR-601, sampling rate is 13.5 MHz for *Y* and 6.75 MHz for each of *C* components, with 8 bits per sample.

Hence, bit rate is $[13.5 \times 10^6 + 2(6.75 \times 10^6)] \times 8$ = **216 Mbps**

Memory required per line: (24 × 360) + (8 × 360) = 11520 bits or 1440 bytes

Memory required per frame each made of 480 lines: 1440×480 = 691.2 KB

Memory required to store 1 hour of movie: 691.2×30 frames/s $\times$ 3600 s = **112 GB**.

Example 6.3 **An NTSC encoded video clip has a frame size of 720 × 480 pixels and is digitized using a bit depth of 8 bits for each of Y, C_b and C_r, and a chroma sub-sampling scheme of 4:2:2. Calculate the file size of 1 minute of the video clip and the total time taken for it to be transmitted over a 2 Mbps transmission line.**

The frame size of the NTSC encoded video is 720 × 480 pixels, and frame rate is 30 fps.

Chroma sub-sampling scheme is 4:2:2, and bit depth is 8 bits for each of Y, C_b and C_r.

In a single line of the video frame, 360 pixels have 24 bits ($Y + C_b + C_r$) and 360 pixels have 8 bits (Y).

Memory required per line = $(24 \times 360) + (8 \times 360)$ = 11520 bits = 1440 bytes

Memory required per frame = 1440×480 = 691200 bytes = 11520×480 = 5529600 bits

Memory required to store one second of the video = 5529600×30 frames/second = 165888000 bits = 19.77 MB

Memory required to store one minute of the video = 165888000×60 = 9953280000 bits = **1186.2 MB**

Time taken to transmit this over a 2 Mbps line = (9953280000/2000000) second = 4976.64 second = **1.38 hours**

Example 6.4 **Repeat the previous problem if the chroma sub-sampling of scheme 4:2:0.**

The frame size of the NTSC encoded video is 720 × 480 pixels, and frame rate is 30 fps.

Chroma sub-sampling scheme is 4:2:0, and bit depth is 8 bits for each of Y, C_b, and C_r.

There are two types of video lines: those with C information and those without C information

For the first type, 360 pixels have 24 bits ($Y + C_b + C_r$) and 360 pixels have 8 bits (Y)

Memory required per line = $(24 \times 360) + (8 \times 360)$ = 11520 bits = 1440 bytes

For the second type, 720 pixels have 8 bits (Y)

Memory required per line = (8×720) = 5760 bits = 720 bytes

Memory required per frame = $(1440 \times 240) + (720 \times 240)$ = 518400 bytes (since there are 240 lines of each type) = 4147200 bits

Memory required to store one second of the video = 518400×30 frames/second = 15552000 bytes = 14.83 MB = 124416000 bits

Memory required to store one minute of the video = 14.83×60 = **889.8 MB** = 7464960000 bits

Time taken to transmit this over a 2 Mbps line = (7464960000/2000000) seconds = 3732.48 seconds = **1.04 hours**

Solved Examples

Example 5.1 **An audio signal is digitized at a sampling rate of 44.1 kHz, a bit depth of 16 and in stereo mode. Calculate the space occupied by 1 minute of the audio and its data rate for playback.**

Here, F = 44100 Hz, b = 16, c = 2, T = 60 seconds

From equation (5.7),

Data rate $D = F.b.c$ = (44100 samples/seocond)(16 bits/sample)(2 channels) = 1411200 bps = **1.4 Mbps**

From equation (5.9),

File size $S = D.T$ = (1411200 Mbits/second)(60 second) = 84672000 bits = **10.09 MB**

(This forms the general rule of thumb that each minute of uncompressed audio takes up approximately 10 MB for audio-CD quality.)

Example 5.2 **An audio signal is recorded on a computer in stereo mode using a bit depth of b such that its data rate is R KB/s. Find an expression for the average frequency of the signal.**

Let the average frequency be f and sampling rate be F.

From equation (5.6), $F = 2f$.

Since the audio is in stereo mode, number of channels $c = 2$

From equation (5.7), data rate $D = F.b.c = (2f)(b)(2)$ bits/second $= (2f)(b)(2)/(8 \times 1024) = fb/2048$ KB/s

Equating to R, we get f **= 2048*R*/*b***

Example 5.3 **An audio clip has a duration of 8 minutes. The highest frequency in the sound wave is 15 kHz. This is to be sampled using 8 bits per sample and in stereo mode. Estimate the minimum data rate in KB/s required to playback the digital file, and the file size in MB.**

Here, f = 15 kHz, b = 8, c = 2, T = 8 minutes.

From equation (5.6), sampling rate $F = 2f$ = 30 kHz

From equation (5.7), data rate $D = F.b.c = (30000 \times 8 \times 2)/(8 \times 1024)$ = **58.6 KB/s**

From equation (5.9), file size $S = F.b.c.t = (30000)(8)(2)(8 \times 60)/(8 \times 1024 \times 1024)$ = **27.46 MB**

Problem 2.1

An analog signal containing components with frequency values ranging from 50 Hz to 5 kHz. is to be sampled. Determine the sampling frequency and the bandwidth of the band-limiting filter.

As per the Nyquist theorem, the sampling rate must be at least twice the bandwidth or the highest frequency component of the source signal.

Hence, sampling rate = 2 × 5 kHz = **10 kHz.**

The bandwidth of the band-limiting filter is from **0 Hz** to **5 kHz.**

Problem 2.2

Repeat Problem 2.1 if the signal is transmitted over a communication channel with a bandwidth from 100 Hz to 4.5 kHz.

In this case, since the bandwidth of the transmission channel is smaller than that of the source signal, the sampling rate = 2 × 4.5 KHz = **9 kHz.**

The bandwidth of the band-limiting filter is from **0 Hz** to **4.5 kHz.**

Problem 2.3

An analog audio signal has a dynamic range of 50 dB. Determine the magnitude of the quantization noise relative to the minimum signal amplitude if the quantizer uses (a) 3 bits (b) 16 bits

We know dynamic range

$$R = 20\log_{10}(S_{max} / S_{min})\text{ dB}$$

Quantization noise $e = \pm 1/2 \times \dfrac{S}{2^n} = \pm 1/2 \times \dfrac{2S_{max}}{2^n} = \pm \dfrac{S_{max}}{2^n}$

Substitution gives us $50 = 20\log_{10}(S_{max}/S_{min})$ dB, or $(S_{max} / S_{min}) = 316.23$

The minimum signal amplitude $S_{min} = S_{max}/316.23$

(a) $n = 3$ bits. Hence, $e = \pm S_{max}/8$. Here we see that the noise is greater than the minimum signal amplitude and hence, is unacceptable.

(b) $n = 16$ bits. Hence, $e = \pm S_{max}/65536$. Here noise is lower than the minimum signal amplitude and hence, acceptable.

Problem 3.3

A monitor has a pixel addressability of 800 × 600 and a color depth of 24 bits. Calculate the minimum amount of display memory required on its adapter card to display an image on the screen.

Each pixel requires 24-bits to be stored. Hence, the entire screen requires a total of $800 \times 600 \times 24$ bits of data to be generated which much be stored in the display memory. Converting this to megabytes and rounding to the next higher integer, the amount of display memory required is $(800 \times 600 \times 24)/(1024 \times 1024 \times 8) \approx$ **2 MB**

Problem 5.1

A GIF image occupies a rectangular area of A inch by B inch on a monitor screen. The resolution of the monitor is C dpi. What is the file size of the image in KB?

The dimensions of the image is A inch × B inch.

The monitor resolution is C dpi.
The total number of pixels in the image is AC × BC.
A GIF image has a color depth of 8 bits.
Hence, storage size of the image = $AC \times BC \times 8$ bits = $(AC \times BC \times 8)/(8 \times 1024)$ KB = **ABC^2/ 1024 KB**

Problem 5.2

Consider an image of dimensions 640 × 480 and color depth of 16 bits. If this image is to be transmitted along a 56 kbps line from a Web server, calculate how long it would take before the entire image is visible on the screen.

The size of the image is $640 \times 480 \times 16$ bits = 4915200 bits
To be transmitted along a 56 kbps line would take (4915200/56000) = **87.78 seconds**.