SECTION 21

VIDEO AND FACSIMILE SYSTEMS

Although much of this section describes basic video and facsimile technologies that have not changed over the years, newer material is also included. For example, international agreement was reached recently on the use of 1920 × 1080 as a common image format for high-definition (HD) production and program exchange. The 1920 × 1080 format has its roots in the CCIR (Consultative Committee in International Radio) sampling standard and brings international compatibility to a new level.

Set-top boxes and high-definition or digital-ready TV sets will be the mechanism that brings digital technology to the consumer for the next several years as the transition from analog to digital takes place. In the United States, three modulation techniques have become "standards" in a particular application: vestigial sideband (VSB) for terrestrial, quadrature amplitude modulation (QAM) for cable, and quaternary phase-shift keying (QPSK) for direct-to-home satellite.

With Internet facsimile, store-and-forward facsimile occurs when the sending and receiving terminals are not in direct communication with one another. The transmission and reception takes place via the store-and-forward mode on the Internet using Internet e-mail. In this mode, the facsimile protocol "stops" at the gateway to the Internet. It is reestablished at the gateway leaving the Internet. Real-time facsimile is covered by Recommendation T.38 approved by the International Telecommunications Union, Telecommunications (ITU-T) in 2002. R.J.

In This Section:

On the CD-ROM:

The following is reproduced from the 4th edition of this handbook:
"Television Cameras," by Laurence J. Thorpe.

CHAPTER 21.1
TELEVISION FUNDAMENTALS AND STANDARDS

James J. Gibson, Glenn Reitmeier

INTRODUCTION

This chapter summarizes analog and digital television signal standards and the principles on which they are based.

The technical standards for color television developed in 1953 for the United States by the National Television System Committee (NTSC) are described on a few pages in the rules of the Federal Communications Commission (FCC Rule Part 73). The rules only specify the radiated signal in sufficient detail for a receiver manufacturer to produce receivers, which convert this signal into a television picture with sound. This traditional approach to formulating standards leaves implementation to competitive forces. Since 1953 many international standards and recommended practices have evolved. A similar philosophy was used in the FCC's adoption of the Advanced Television System Committee (ATSC) digital television standards in 1996.

All color television standards are based on the same principles:

- The psychophysics of the human visual system (HVS).
- Picture-signal conversion by sampling/display, at field rate, of three primary colors in a flat rectangular dynamic picture on a raster of horizontal scan lines, scanned from left to right and top to bottom.
- The signals are conveyed as three components: one luminance signal, which essentially provides brightness information, and two chrominance signals which essentially provide hue and color saturation information.
- For radio frequency transmission these three signals and audio signals are multiplexed to form a single r.f. signal, which occupies a channel in the frequency spectrum.

Some of these principles are illustrated in Fig. 21.1.1, which shows a block diagram of a standard analog television system for terrestrial broadcasting. The figure shows that the video and audio signals are multiplexed separately to form *composite video and audio signals*, which are subsequently delivered to separate picture and sound transmitters generating signals, which are diplexed to form a radiated signal that occupies a 6, 7, or 8 MHz band in the radio frequency spectrum from 40 to 900 MHz. This is the usual practice in broadcasting of analog television signals in the NTSC, PAL (Phase Alternation Line), and SECAM (Séquentiel à mémoire) systems. These systems, which are compatible with black and white reception, use frequency division multiplex: The chrominance signals are bandlimited and modulate one or two subcarriers that are "inconspicuously" added to (multiplexed with) the luminance signal. Besides analog television terrestrial broadcast standards, there are many standards for analog television production, storage, and distribution (terrestrial, satellite, and cable) developed by several organizations. In analog television there are two picture scanning standards specified as N/F_v where N = total number of scanning lines and F_v = number of picture fields per second. These standards are 625/50 and 525/60 (including

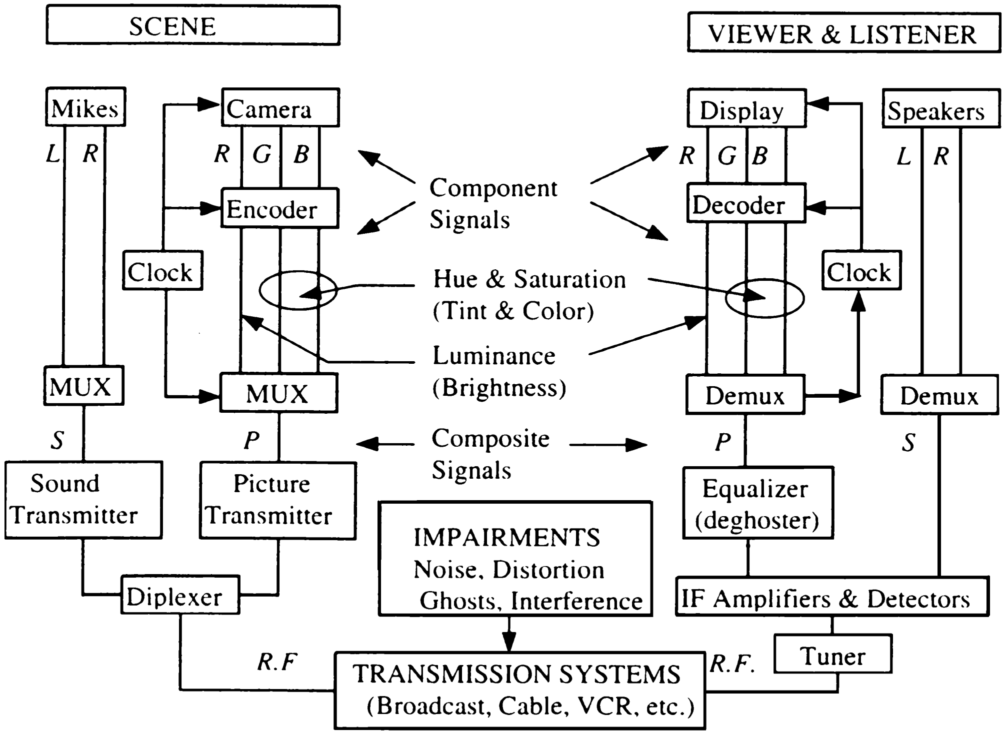


FIGURE 21.1.1 Functional block diagram of a standard analog television broadcast system. Some systems may multiplex sound and picture before delivery to a transmitter.

525/59.94). There are three basic *composite color video signal standards* that carry all the color information in one signal: NTSC, PAL, and SECAM. In addition there are nine standards (with variations) describing the radiated signal carrying the composite color video signals with various types of r.f. modulation and bandwidths and with various standards for audio signals and audio signal modulation. These nine standards are referred to as B, G, H, I, K, K1, L, M, N. Only system M is of type 525/60.

In digital television (DTV), the three component signals and audio are sampled, digitized, and data compressed to eliminate redundant and psychophysically irrelevant data. Digital signals use available spectrum more effectively than analog signals. Digital television standards have been developed with consideration for flexibility, extensibility (for future applications), and interoperability with other systems for information production and distribution (e.g., computers).

In 1982 the Radio Consultative Committee of the International Telecommunications Union (ITU-R), formerly called the International Radio Consultative Committee (CCIR) adopted an international digital television component standard, ITU-R Recommendation 601. In this chapter this important standard is referred to by its popular old designation: CCIR601. This standard was primarily intended for production and for tape recording (SMPTE format D-1), but is now used in many applications, including DTV transmission.

In digital television (DTV), including high-definition television (HDTV), digital techniques are used for video compression, data transport, multiplexing, and r.f. transmission. DTV promises to be more than television, in the sense that it delivers to homes a digital channel with high data rate which may carry more than high quality pictures and sound. Of particular importance are the standards developed by the Moving Picture Expert

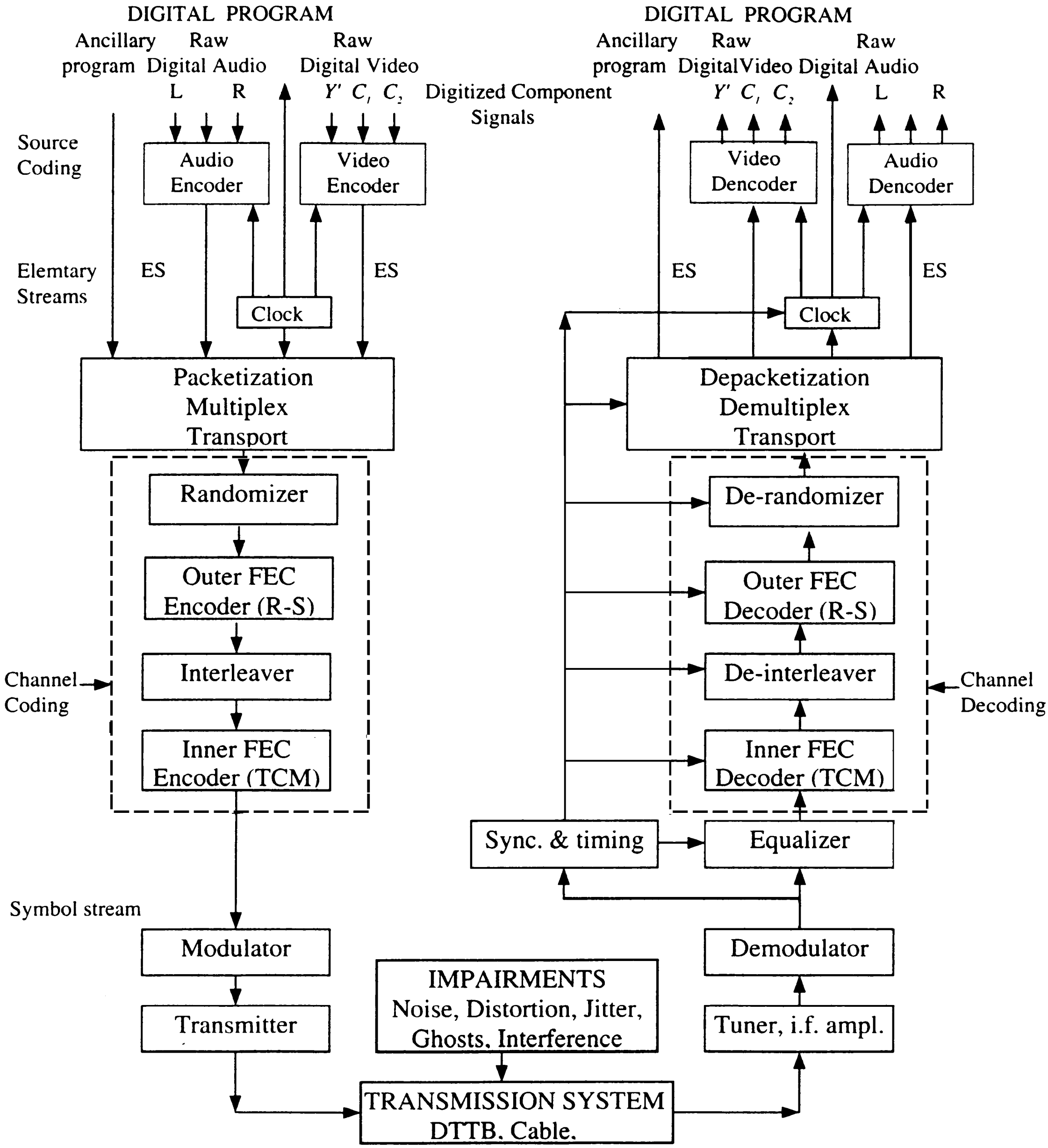


FIGURE 21.1.2 Functional block diagram of a standard television broadcast system. Different modulation techniques and signal bandwidths are used for different transmission media.

Group (MPEG) of ITU. MPEG standards are quite flexible, but include specific approaches to television data compression and packetization. MPEG standards have been adopted by the International Standards Organization (ISO) and the International Electro-technical Committee (IEC). Based on the MPEG-2 format, the FCC Advisory Committee on Advanced Television Systems (ACATS) provided oversight for the development of a "Grand Alliance," a system for high-definition digital television (HDTV) for North America. The Grand Alliance itself was comprised of AT&T, General Instrument, Massachusetts Institute of Technology, Philips, Sarnoff, Thomson, and Zenith, which joined forces to forge a single best-of-the-best system from four competing system proposals. The Grand Alliance system is the basis of the ATSC digital television standard and the FCC digital broadcast standards for the United States adopted in 1996.

Figure 21.1.2 shows basic functional blocks of the "Grand Alliance" HDTV system. The basic principles listed above for analog systems still apply to this digital system. The video and audio components are basically the same as in the analog system shown in Fig. 21.1.1, but source coding (removal of irrelevant and redundant data), ancillary signals, multiplexing, transport, packetization, channel coding (for error management), and modems are entirely different.

In both analog and digital systems, the relation between the picture, as observed by a viewer, and the signals is not a simple one. Signal quality measures are not easily related to the subjective quality of pictures. In fact, there are very few objective measures of picture quality, while there are many objective measures of signal quality and signal tolerances. The selection of television standards, including recent HDTV standards, are based on subjective picture quality evaluations. Relations between signal quality and picture quality in NTSC, PAL, and SECAM are fairly well known. Numerous measurement and monitoring techniques using test signals have been developed for these signals. Relations between objective measures of signal quality and picture and sound quality are not as well correlated in digital television and HDTV.

Ongoing work on digital TV standardization is carried out by numerous committees. ATSC continues to develop its terrestrial transmission standard. The Society of Cable Telecommunications Engineers (SCTE) develops standards for cable transmission. The Digital Video Broadcast (DVB) group also continues to develop its standards for terrestrial (DVB-T), cable (DVB-C), and satellite (DVB-S) transmission. The Society of Motion Picture and Television Engineers (SMPTE) is involved in standards for related professional production equipment. The Consumer Electronics Association (CEA) establishes industry standards for consumer equipment.

PHOTOMETRY, COLORIMETRY, AND THE HUMAN VISUAL SYSTEM

Radiance and Luminance

The HVS is sensitive to radiation over a 2:1 bandwidth of wavelengths extending from 380 to 760 nm, i.e., from extreme blue to extreme red. When adapted to daylight it is most sensitive to green light at 555 nm. As the wavelength of monochromatic light departs from 555 nm, the *radiance* (radiation from a surface element in a given direction defined by the angle Θ from the normal and measured in watts/steradian per unit projected area = the actual surface of the radiating element times cosΘ) must be increased for constant perception of brightness. The International Commission on Illumination (CIE) has standardized the response versus wavelength of the HVS $\bar{y}(\lambda)$ of a standard observer adapted to daylight vision (photopic vision). Figure 21.1.3 shows $\bar{y}(\lambda)$ versus λ. Luminance, sometimes referred to as brightness and measured in cd/m^2, *Candelas per projected area in m*^{2}, is defined as

$$Y = 680\int_0^\infty E(\lambda)\bar{y}(\lambda)d\lambda \quad \text{cd/m}^2 \qquad E(\lambda) = \text{spectral density of radiance in (W/nm)/m}^2 \tag{1}$$

where λ is the wavelength in nanometers. Older units for luminance are: a footlambert (ft.L) = 3.42626 cd/m^2 and a millilambert (mL) = 3.18310 cd/m^2.

A surface is *diffuse* at a wavelength λ if $E(\lambda)$ is independent of the direction Θ. The face plate of a TV tube is essentially diffuse for all wavelengths of visible light, but this may not be the case for a projection screen. Thus the luminance of the face-plate of a picture tube is roughly independent of the direction from where it is seen. Typically the peak luminance levels of picture tubes is about 120 cd/m^2, but bright displays may have luminance levels exceeding 250 cd/m^2. The luminance of bright outdoor scenes may well exceed 10,000 cd/m^2,

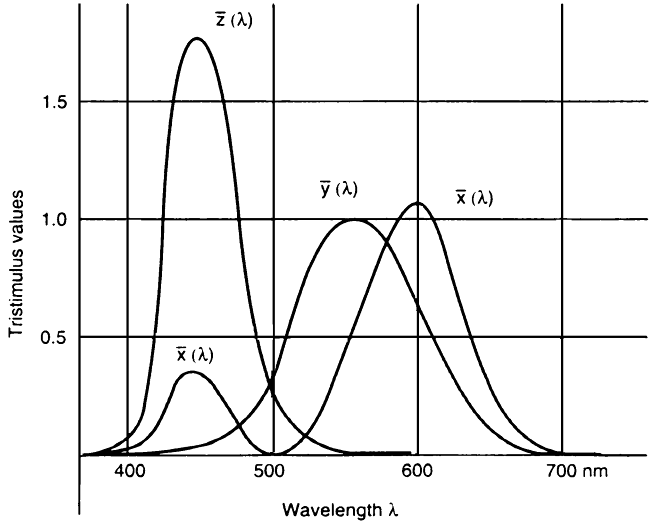


FIGURE 21.1.3 Tristimulus values of the CIE nonphysical primaries, $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$. Note that $\bar{y}(\lambda)$ is the CIE standard luminosity function.

while a motion picture screen may be 30 cd/m^2. *Equal Energy Radiation*, $E(\lambda)$ = constant, appears to have the color "white." Corresponding radiometric and photometric units are:

Radiant flux	Watt	Luminous flux	Lumens
Irradiance	Watt/m^2	Illuminance	Lux = lumens/m^2
Radiant intensity	Watt/steradian	Luminous intensity	Candela = lumens/steradian
Radiance	(Watt/steradian)/m^2	Luminance	Candela/m^2

At $\lambda = 555$ nm there are 680 lm/w. Without a color filter, the illuminance I of a target in a camera or of the retina of an eye is proportional to the luminance Y of the scene (theoretically $I \cong \pi Y/4(\text{f-number of the lens})^2$.

Contrast and Resolution

The HVS is sensitive to relative changes in luminance levels. The just noticeable difference (JND) between the luminance $Y + \Delta Y$ of a small area and the luminance Y of a surrounding large area can be expressed by the ratio

$$F = \Delta Y/Y \tag{2}$$

where F is the *Fechner ratio*.

The Fechner ratio is remarkably constant over the large range of the high luminance levels used in TV displays. It ranges from 1 percent for $Y > 10$ cd/m^2 up to 2 percent at $Y = 1$ cd/m^2. Assuming a constant Fechner ratio, the number of distinguishable small area gray levels between a "black" level Y_b and a highlight "white" level Y_w is

$$n \cong (\ln Y_w/Y_b)/F \qquad \text{or} \qquad Y_w/Y_b \cong \exp(nF) \cong (1 + F)^n = \text{contrast} \tag{3}$$

For example, for n = 255 levels (8-bit quantization of log luminance) and F = 1.5 percent, the contrast is 45. With 9 bits (511 levels @ 1 percent) the contrast is 165. Contrast rarely exceeds 50:1 in TV displays owing to ambient light and light scattered within the picture tube. Thus, due to the contrast available in consumer displays, DTV systems use 8-bit luminance (and chrominance) signals.

More relevant test patterns for measuring the contrast sensitivity of the HVS are luminance sinewave gratings added to a uniform background luminance level. The contrast ratio of a sinewave grating, also referred to as *modulation*, is defined as:

$$m = (Y_{maz} - Y_{min})/(Y_{max} + Y_{min}) = (\text{luminance amplitude/average luminance}) \quad (4)$$

The just perceptible modulation (JPM) of a stationary sinewave grating is sensitively dependent on a large number of factors: the spatial frequency v of the grating expressed in cycles per degree (cpd), the average luminance level, the orientation of the wavefront, the adaptation of the pupil of the eye, and ambient factors. If the line from the eye to the observed point is perpendicular to the picture, v cycles per degree can be converted into k cycles per picture height (cph) by multiplication with 57.3/D, where D is the viewing distance in picture height. The most sensitive m is 0.3 percent at about 3.5 cpd = 200/D cph and increases to 100 percent at about 35 cpd = 2000/D cph (500 cph @ D = 4). This upper limit can be taken as an estimate of the resolution limit of the HVS. TV systems have been designed with HVS characteristics in mind. Analog and standard definition (SD) digital systems assume D = 6; therefore, the HVS resolution limit is approximately 383 cph. HDTV systems assume a larger display or closer viewing, D = 2; therefore, the HVS resolution limit is approximately 1000 cph.

The ratio of the perceived to the displayed modulation is the modulation transfer function (MTF) of the HVS. It rolls off with spatial frequency depending on a number of parameters including the adaptation of the pupil of the eye. It is down about 6 dB at 750/D cph for a luminance of 140 cd/m^2. One performance measure of a TV system is its MTF = ratio of displayed m(k) to input m(k) versus spatial frequency for various orientations of the wavefront of the grating.

The eye has high resolution only within a sustained angle of about 2°. For D = 6 this is 2.5 percent of a TV display area. The remaining area is viewed with low resolution rod vision which is sensitive to flicker.

Gamma "Pre-Correction"

Since the HVS is sensitive to contrast, it can tolerate more noise and larger luminance errors in bright areas than in dark areas. Since noise and errors are added to the luminance signal during transmission, the visibility of the added noise is reduced if the luminance signal is compressed at the source and expanded in the receiver. A constant Fechner ratio suggests logarithmic compression and exponential expansion of the luminance, a proposal that was recommended by the NTSC in 1941 and is still allowed in the FCC rules for black and white TV. In the early days of black and white TV, it was found, however, that the picture tube itself acted as a good and inexpensive expander. The displayed luminance Y on the face plate of a picture tube in response to an applied signal voltage V is approximately equal to:

$$Y = \text{const.} \times V^{\gamma} \quad (5)$$

where γ is *gamma* ranges from 2 to 3. Assuming that the output signal Y from a TV camera is proportional to the illuminance of the target (a fair assumption for modern cameras), the luminance signal Y delivered by such a linear camera is compressed right at the output of the camera to yield a *gamma "corrected" luminance signal*, which, in black and white TV, is approximately proportional to $Y^{1/\gamma}$. In DTV systems, the same practice continues to be employed, both for reducing the visibility of compression-related artifacts as well as maintaining compatibility with legacy analog systems, which is an economic consideration in dual digital/analog receivers. Table 21.1.1 shows standards for gamma and modifications of Eq. (5) in recent standards.

Flicker

The human visual system is quite sensitive to large area luminance fluctuations at frequencies below 100 Hz (flicker). The critical flicker frequency f_f Hz (threshold of perception) is determined by on-off modulation of the luminance of a large area. The critical frequency increases, according to the "Ferry-Porter law," in proportion to the logarithm of the highlight Y cd/m^2 of the luminance:

$$f_f = 30 + 12.5 \log_{10}(Y \text{ cd/m}^2) \quad \text{Hz} \quad (6)$$

TABLE 21.1.1 Standards for Electro-Optical Transfer Characteristic: $V = F(L)$, $L = F^{-1}(V)$ Example: $R' = F(R = R)$ in "gamma corrector," $R = F^{-1}(R')$ in display

SYSTEM	γ (gamma)	V_0(bias)	k(slope)	L^*	V^*
NTSC	2.2	0	0	0	0
PAL/SECAM	2.8	0	0	0	0
TU R-709, SMPTE170M, SMPTE274M	1/0.45	0.099	4.5	0.018	0.081
SMPTE 240M	1/0.45	0.1115	4	0.0228	0.0912

V = *normalized electric signal level* and L = normalized optical stimulus level (at Maximum level $V = L = 1$ by definition):
$V = F(L) : V = kL$ for $0 \le L < L^*$ and for $L^* \le L \le 1$ $V = (1 + V_o)L^{1/\gamma} - V$
$L = F^1(V) : L = V/k$ for $0 \le V < V^* = kL^*$ and for $V^* \le V \le 1$ $L = [(V + V_o)/(1 + V_o)]^\gamma$

The result varies with individuals and with the adaptation of the eye. It shows that f_f increases by about 12.5 Hz when Y increases by a factor of 10. For example, f_f is 48, 50, and 60 Hz for a peak luminance Y = 25.4, 36.6, and 227.6 cd/m^2. The HVS is more sensitive to flicker for light entering the eye at an angle (rod vision). In motion pictures the picture frame rate is 24 pictures per second. To avoid flicker at acceptable luminance levels, every picture is shown at least twice, when the film is projected (this is referred to as double-shuttering). In television the frame rate is 25 Hz in most of the world and 30 Hz in the United States, Japan, and some other countries. In all analog TV systems, large area flicker rates are doubled by using two fields with interlaced scan lines to complete a frame.

DTV systems decouple the display frame rate and the transmitted frame rate. DTV systems generally support the legacy analog frame rates and interlaced scanning formats, and additionally provide new capabilities for progressive scan formats and film-related frame rates.

Color Space, Color Primaries, and CIE Chromaticities

It is a fact that a wide range of colors can be reproduced with only three sources of light, e.g., red, green, and blue. The basic laws of *colorimetry*, the science devoted to color vision are:

- The HVS can only perceive three attributes of color: *brightness, hue*, and *saturation*, more precisely defined as *luminance, dominant wavelength*, and *purity*.
- Colors can be represented as vectors in a three-dimensional linear vector space referred to as *color space*. Colors add and subtract as vectors, but a distance in color space is not a measure of perceived difference in colors. In color space the luminance component $Y \ge 0$ traditionally points up from a horizontal chrominance plane $Y = 0$.

Figure 21.1.4 shows a color vector $\mathbf{A} = Y\mathbf{W} + \mathbf{C}$ as the sum of a white vector $Y\mathbf{W}$ with the same luminance (brightness) Y as **A** and a chrominance vector **C** in a constant luminance plane with an orientation related to hue (tint) and a magnitude related to saturation (color).

Based on experiments with hundreds of observers matching colors composed of monochromatic light of different wavelengths, the CIE standardized X- and Z-axes at right angles in the chrominance plane, such that all color vectors have nonnegative components X, Y, Z, and that X = Y = Z for equal energy light. Thus, basis-vectors **X, Y, Z** with unit length along the *XYZ*-axes are *artificial or nonphysical primary vectors*. The X, Y, Z components of monochromatic colors all having the same radiance are $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ (Fig. 21.1.3). Given a color with a radiance having a spectral power density $E(\lambda)$ (radiance per nm), the relative X, Y, Z components are

$$X = \int_0^\infty E(\lambda)\bar{x}(\lambda)d\lambda \qquad Y = \int_0^\infty E(\lambda)\bar{y}(\lambda)d\lambda \qquad Z = \int_0^\infty E(\lambda)\bar{z}(\lambda)d\lambda \tag{7}$$

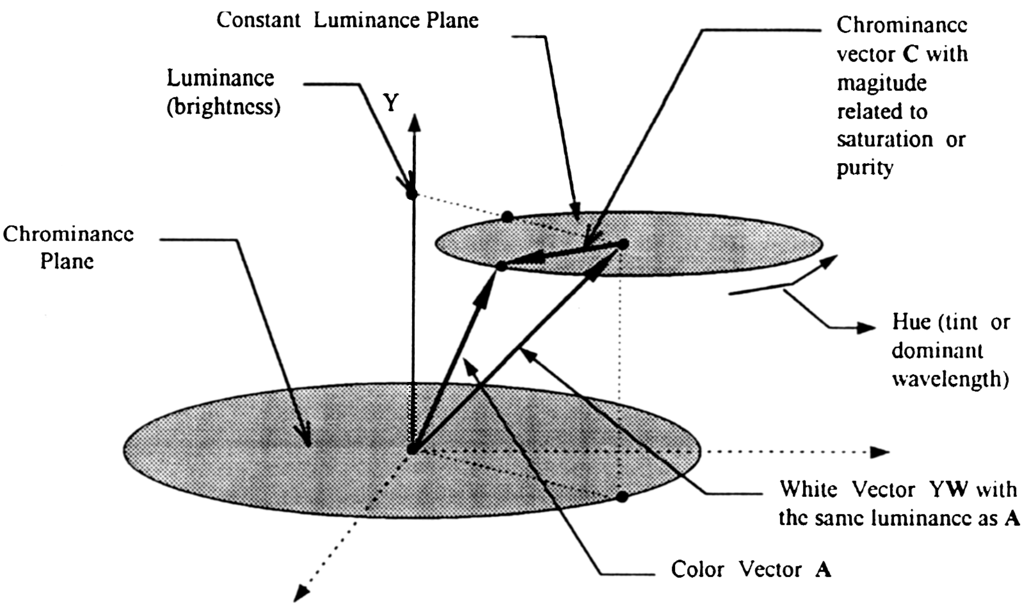


FIGURE 21.1.4 Representation of a color as a vector **A** with three attributes luminance (brightness), hue (tint), and saturation. **A** has a vertical luminance component Y. **A** is the sum of a white vector **YW**, which has the same luminance as **A**, and a chrominance vector **C**, which is in the constant luminance plane.

Since X = Y = Z for equal energy white (E = const.), the areas under the curves $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ are the same. Clearly lights with different spectral content E(λ) can have the same color. The direction of a color vector is specified by the point [x, y, z] where the vector or its extension penetrates the unit plane X + Y + Z = 1.

$$[x, y, z] = [X, Y, Z]/(X + Y + Z) \tag{8}$$

thus x + y + z = 1 and Y/y = X/x = Z/z = X + Y + Z = "gain."

The coordinates x, y, z are referred to as the *CIE chromaticities* of the color. Clearly z = 1 – x – y is redundant. It is common and practical to specify a color vector **A** by chromaticities *x* and *y* and luminance Y as

$$\mathbf{A} = [X, Y, Z] = (Y/y)[x, y, 1 - x - y)] \tag{9}$$

Figure 21.1.5 shows the CIE color space and the CIE chromaticity plane penetrated by color vectors or their extensions. Figure 21.1.6 shows the horseshoe-shaped locus of the xy-chromaticities of monochromatic light $[x(\lambda), y(\lambda)] = [\bar{x}(\lambda), \bar{y}(\lambda)]/(\bar{x}(\lambda)+\bar{y}(\lambda)+\bar{z}(\lambda)]$. All realizable chromaticities are within the horseshoe. The magenta colors along the straight line from extreme blue to extreme red are not monochromatic, but are obtained as red-blue mixtures. Areas in the CIE chromaticity chart of just noticeable differences in chromaticities (JNDs) are shaped like ellipses which are larger in the green area than in the red area, which in turn are larger than in the blue area. The number of chromaticities (color vector directions) required for perfect color reproduction has been estimated to range from 8000 (13 bits) to 130,000 (17 bits). The number of chromaticity JNDs within the color gamuts of current display devices has been estimated to be about 5000. In television standards all color vectors are defined in CIE color space.

The CIE components, X, Y, Z are not practical. Television standards specify realizable red, green, and blue *primary colors* **R, G, B**, which are not in the same plane. In terms of these primaries a color is represented as a vector

$$\mathbf{A} = [X, Y, Z] = R\mathbf{R} + G\mathbf{G} + B\mathbf{B} = Y\mathbf{W} + C \tag{10}$$

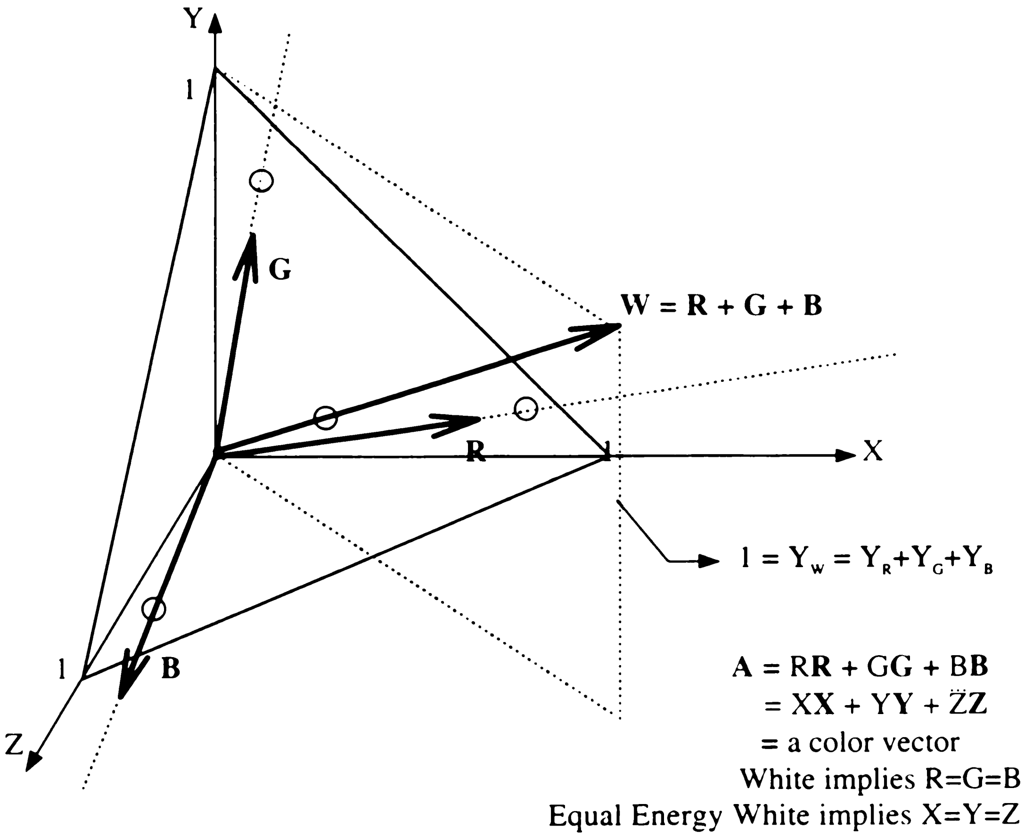


FIGURE 21.1.5 CIE color space with CIE chromaticity plane $X + Y + Z = 1$. The CIE chromaticities of a color vector are the coordinates x, y, and z = 1 − x − y, where the vector or its extension penetrates the CIE chromaticity plane. These points are shown for a set of **R, G, B** basis vectors and the associated white vector **W**, which by definition has a luminance of unity. **X, Y, Z** are unit vectors along the X, Y, Z axes.

where R, G, B are *tristimulus values* = quantity of each primary color in the mixture. The luminance Y of **A** multiplies a "reference *white vector*" **W** = **R** + **G** + **B**, which, by definition, has a luminance component normalized to unity ($Y_w = 1$). Thus *chrominance vector* **C** is in a *constant luminance plane* ($Y_c = 0$). The luminance components of **R, G**, and **B** are Y_r, Y_g, and Y_b, respectively. Consequently:

$$Y = Y_r R + Y_g G + Y_b B = \text{the luminance component of } \mathbf{A} \quad (11a)$$

$$1 = Y_r + Y_g + Y_b = \text{normalized luminance component of the white vector } \mathbf{W} \quad (11b)$$

$$\mathbf{W} = \mathbf{R} + \mathbf{G} + \mathbf{B} = \text{white vector with luminance component } Y_w = 1 \quad (11c)$$

$$C = M(R - Y) + N(B - Y) = \text{chrominance vector with luminance component } Y_c = 0 \quad (12a)$$

$$\mathbf{M} = R - (Y_r/Y_g)\mathbf{G} = \text{the R} - \text{Y basis vector in the chrominance plane} \quad (12b)$$

$$\mathbf{N} = \mathbf{B} - (Y_b/Y_g)\mathbf{G} = \text{the B} - \text{Y basis vector in the chrominance plane} \quad (12c)$$

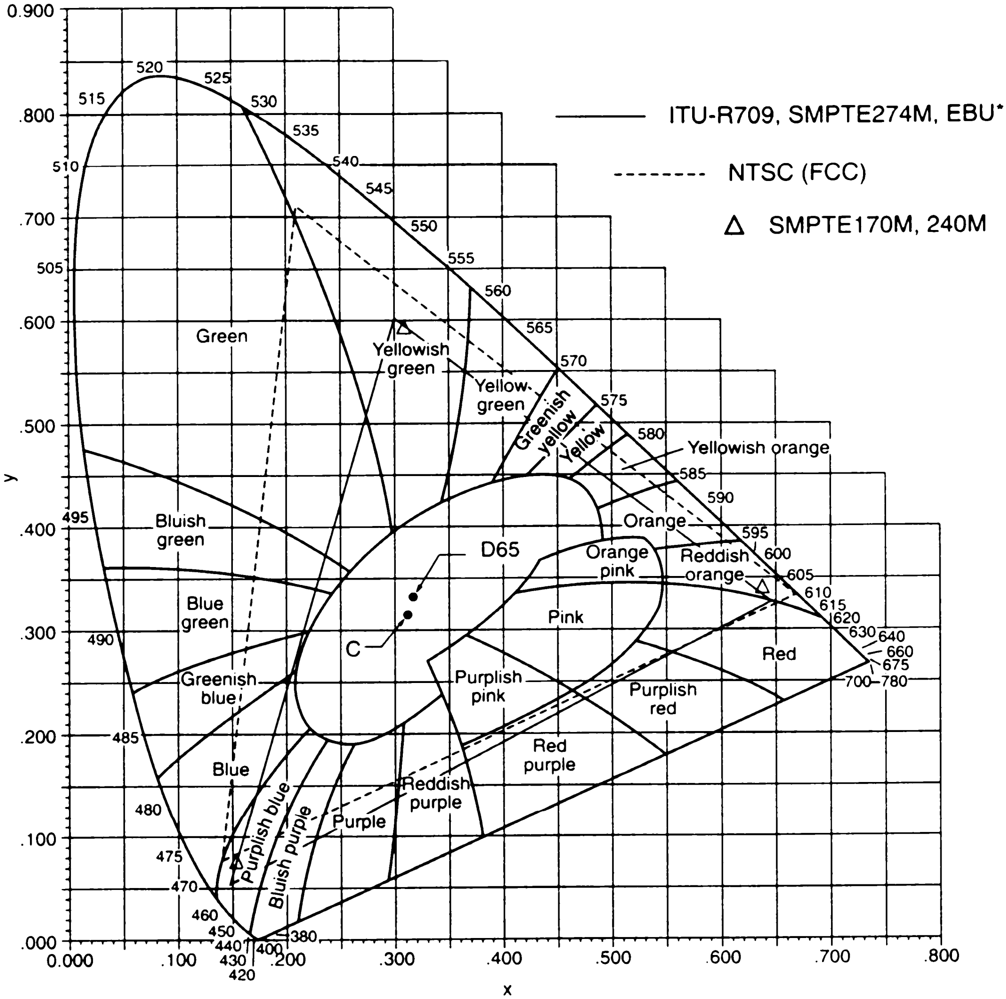


FIGURE 21.1.6 The *xy*-chromaticity diagram of the CIE system. Also shown are television standard rgb chromaticities and white illuminants. *The EBU *x*-coordinate for green is 0.29.

A white vector implies **C** = 0 and R = G = B = Y. Equation (12) is derived from Eqs. (10) and (11). Figure 21.1.5 shows vectors **R, G, B**, and **W**. Figure 21.1.7 shows the vectors **M** and **N**. Given the X, Y, Z components of the primaries **R, G, B** the R, G, B tristimuli can be related to the CIE tristimuli X, Y, Z by a matrix P as:

$$\begin{vmatrix} X \\ Y \\ Z \end{vmatrix} = P \begin{vmatrix} R \\ G \\ B \end{vmatrix} = \begin{vmatrix} X_r & X_g & X_b \\ Y_r & Y_g & Y_b \\ X_r & Z_g & Z_b \end{vmatrix} \begin{vmatrix} R \\ G \\ B \end{vmatrix} \qquad (13)$$

RGB components of two different basis systems (e.g., NTSC and PAL) specified by matrices P_1 and P_2 are related by a matrix $P_1P_2^{-1}$ where P^{-1} is the inverse of P. In television standards the **R, G, B** primaries

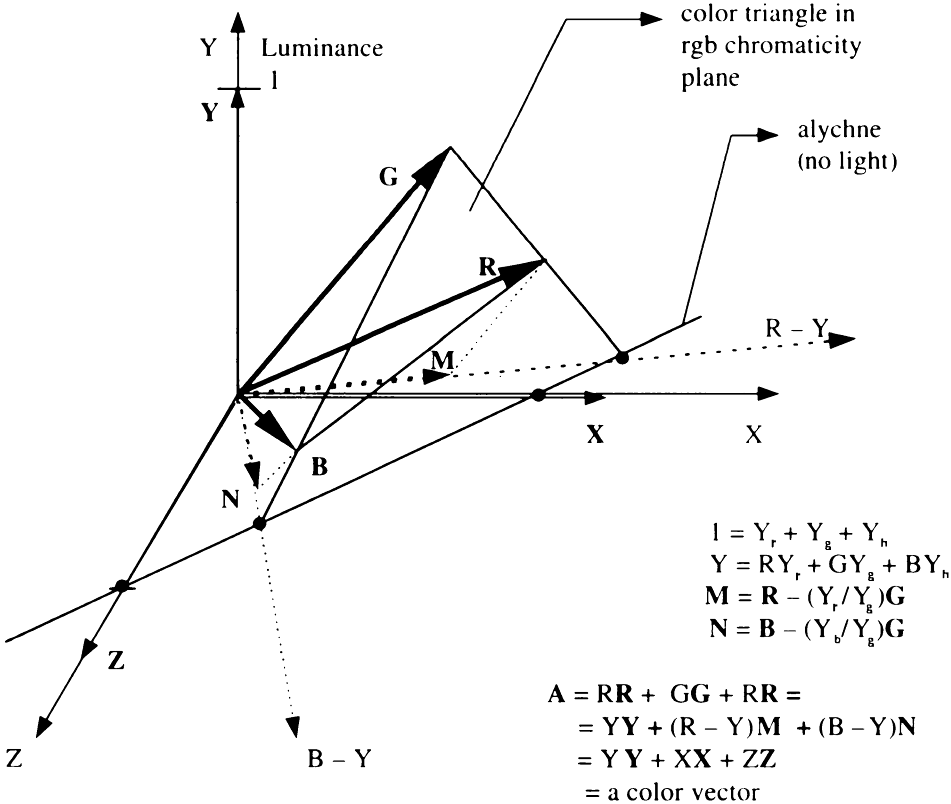


FIGURE 21.1.7 The rgb chromaticity plane goes through the tips of the basis vectors **R, G,** and **B**. It intersects the chrominance plane along the alychne. Also shown are the CIE basis vectors **X, Y,** and **Z** and the basis vectors used in television **Y, M,** and **N**. The chrominance values defined by the chrominance basis vectors **M** and **N** are the "color difference" values M = R – Y and N = B – Y.

are specified by their CIE chromaticities and by the chromaticity of **W** (illuminant). The luminance values of the primaries Y_r, Y_g, Y_b can be derived from these eight standardized chromaticities by inverting the matrix **P** for the white vector $R = G = B = Y_w = 1$ and noting that $X/Y = x/y$ and $Z/Y = z/y$ and that $Y_r + Y_g + Y_b = 1$:

$$\begin{vmatrix} X_w \\ Y_w \\ Z_w \end{vmatrix} = \begin{vmatrix} x_w/y_w \\ 1 \\ z_w/y_w \end{vmatrix} = \begin{vmatrix} x_r/y_r & x_g/y_g & x_b/y_b \\ 1 & 1 & 1 \\ z_r/y_r & z_g/y_g & z_b/y_b \end{vmatrix} \begin{vmatrix} Y_r \\ Y_g \\ Y_b \end{vmatrix} \qquad (14)$$

Given the Y values of the primaries, the X and Z values can be determined by Eq. (9). Table 21.1.2 lists the xy-chromaticities and the luminance values of the primaries in standard TV systems. Figure 21.1.6 shows the NTSC, ITU-R709 (SMPTE274M), EBU, and the SMPTE170M (SMPTE240M) chromaticities in the CIE diagram. Television programs are produced with the assumption that display devices have (phosphors with) these chromaticities. The chromaticities of the primaries of a display device can be marked as corners of a triangle in Fig. 21.1.6. Since the tristimuli R, G, B are nonnegative (see Figs. 21.1.5 and 21.1.7, only colors with chromaticities within the triangle can be displayed.

TABLE 21.1.2 Chromaticities and Luminance Components in Standard Television Systems. White Illuminant (last row) is C in NTSC, Otherwise D_{65}. Note that ITU-R709 is Almost the Same as EBU, and That NTSC Accommodates More Green Colors

	NTSC (FCC) System M			ITU-R709 & SMPTE274 EBU* (PAL* & SECAM*)			SMPTE 170M & 240M		
	x	y	Y	x	y	Y	x	y	Y
Red	0.67	0.33	0.299	0.64	0.33	0.2125 0.222*	0.630	0.340	0.212
Green	0.21	0.71	0.587	0.30 0.29*	0.60	0.7154 0.707*	0.310	0.595	0.701
Blue	0.14	0.08	0.114	0.15	0.06	0.0721 0.071*	0.155	0.070	0.087
White	0.3101	0.3162	1	0.3127	0.3291	1	0.3127	0.3291	1

In television the key components are R – Y, Y, and B – Y of the **M, W, N** primaries. They are related to the tristimuli *R*, *G*, *B* by the matrix *H*:

$$\begin{vmatrix} R - Y \\ Y \\ B - Y \end{vmatrix} = H \begin{vmatrix} R \\ G \\ B \end{vmatrix} = \begin{vmatrix} 1 - Y_r & -Y_g & -Y_b \\ Y_r & Y_g & Y_b \\ -Y_r & -Y_g & 1 - Y_b \end{vmatrix} \begin{vmatrix} R \\ G \\ B \end{vmatrix}, \qquad \begin{vmatrix} R \\ G \\ B \end{vmatrix} = \begin{vmatrix} 1 & 1 & 0 \\ -Y_r + Y_g & 1 & -Y_b - Y_g \\ 0 & 1 & 1 \end{vmatrix} \begin{vmatrix} R - Y \\ Y \\ B - Y \end{vmatrix} \quad (15)$$

Two colors are said to be *complementary* if some weighted mixture of them results in a white color **WW**. Saturated complementary colors to the primaries **R, G, B** are Cyan **CY** = **G** + **B,** Magenta **MA** = **B** + **R** and Yellow **YE** = **R** + **G**. If a primary has components M, Y, N, the saturated complementary color has components –M, 1 – Y, –N. The RGB and MYN values of these basic colors in NTSC, ITU-R709 & SMPTE274M, and SMPTE240M are listed in Table 21.1.3. The EBU colors in PAL & SECAM differ only slightly (third decimal) from ITU-R709. These saturated colors can be seen on TV as color bars.

The rgb-Chromaticities and the Color Triangle

Figure 21.1.7 shows a plane connecting the tips of the primary vectors **R, G, B**. This plane is referred to as the rgb-chromaticity plane, not to be confused with the CIE xyz-chromaticity plane. A color vector **A** = R**R** + G**G** + B**B** = (R + G + B)(r**R** + g**G** + b**B**) has rgb-chromaticities = [r,g,b] = [R,G,B]/(R + G + B). Clearly g = 1 – r – b is redundant. An rgb-chromaticity vector is conveniently represented as **G** + *r*(**R** – **G**) + b(**B** – **G**).

TABLE 21.1.3 Values of Y, M = R – Y and N = B – Y for Saturated Colors

		NTSC			ITU-R709 & SMPTE 274M			SMPTE 240M & 170M		
COLOR	R G B	Y	R – Y	B – Y	Y	R – Y	B – Y	Y	R – Y	B – Y
White	1 1 1	1	0	0	1	0	0	1	0	0
Yellow	1 1 0	0.886	+0.114	–0.886	0.928	+0.072	–0.928	0.913	+0.087	–0.913
Cyan	0 1 1	0.701	–0.701	+0.299	0.787	–0.787	+0.213	0.788	–0.788	+0.212
Green	0 1 0	0.587	–0.587	–0.587	0.715	–0.715	–0.715	0.701	–0.701	–0.701
Megenta	1 0 1	0.413	+0.587	–0.587	0.285	+0.715	+0.715	0.299	+0.701	+0.701
Red	1 0 0	0.299	+0.701	–0.299	0.213	+0.787	–0.213	0.212	+0.788	–0.212
Blue	0 0 1	0.114	–0.114	+0.886	0.072	–0.072	+0.928	0.087	–0.087	+0.913

The rgb-chromaticity of the white vector **W** is r = g = b = 1/3. Figure 21.1.7 also shows that the **GR** and **GB** planes intersect the chrominance plane along the M = R − Y and N = B − Y axes. The xyz chromaticities are nonlinearly related to the rgb chromaticities. The relations are easily determined with Eq. 13, given [R,G,B] = [r,g,b] *or* [X,Y,Z] = [x,y,z]. The figure also shows a line, the *alychne* (no light), which is the intersection of the rgb-chromaticity plane and the chrominance plane Y = 0.

The tips of the **R, G,** and **B** primary vectors are the corners of the ***color triangle***. The rgb-chromaticities of the saturated complementary colors, cyan, magenta, and yellow, are midpoints on the sides of the triangle. The color triangle can be mapped to take more convenient shapes, for example, in a right angle or coordinate system. Figure 21.1.6 shows color triangles for some television standards first polar-projected onto the CIE chromaticity plane and subsequently parallel projected onto CIE xy-chromaticity diagram. Display devices using the standard chromaticities can only display colors with chromaticities inside the color triangle (nonnegative R, G, B tristimuli).

Relation Between Chromaticities and Chrominance-to-Luminance Ratios

Since R − Y, Y, B − Y are related to signals in TV transmission, it is of interest to relate the ratios R/Y and B/Y to the xyz (CIE) chromaticities. In the CIE chromaticity diagram (Fig. 21.1.6) all lines R/Y = constant go through the intersection of the projection of the green-blue side (R/Y = 0) of the color triangle with the x-axis and all B/Y = constant lines go through the intersection of the projection of the green-red side (B/Y = 0) of the color triangle with the x-axis. Linear scales of R/Y and B/Y can be made on any line parallel with the x-axis. The scales are easily calibrated since lines R/Y = 1 and B/Y = 1 for lines going through the illuminant white point. The intersection between the pivoting R/Y and B/Y lines determines the chromaticity. Errors in the (R − Y)/Y and (B − Y)/Y ratios cause errors in the slope of these pivoting lines and may require the display device to produce negative light if they intersect outside the color triangle. Chroma noise and incorrect chroma gain ("color" on TV sets) may cause such errors. Errors in the chrominance cause large chromaticity errors when Y is small, which it is in the blue-magenta-red areas, especially in dark areas.

Implications of Chromaticity Standards

Television standards specify the chromaticities that are expected to be used in the display devices at the receivers. Receiver manufacturers do not have to meet these standards, and in fact, in the United States, more efficient phosphors are used, which do not have the standard FCC chromaticities. The SMPTE has recommended other chromaticities to be used at the source which better match the more efficient phosphors now used in most receivers. Various chromaticity standards are shown in Table 21.1.2. SMPTE 170M is proposed as NTSC broadcast studio standards and ITU-R709 (equal to SMPTE 274M) and 240M for DTV standard. White illuminant D_{65} is becoming universal. Figure 21.1.6 shows that the "old" NTSC gamut of colors (triangle) is substantially larger in the green-cyan area than the gamuts of the newer chromaticities. However, the gamuts of the new phosphors cover almost the gamut of paints, pigments, and dies, and there are few complaints about the green color. Future displays may provide larger gamuts and force a change in *de facto* standards.

The implications of chromaticity standards or of *de facto* standards are that in production of television programs it is expected that most displays in the receivers have luminous phosphors and a reference white illuminant with chromaticities close to the standards. Cameras usually have sensors with nonstandard chromaticities. The output signals from a camera can be matrixed to conform with a large gamut of chromaticity standards. A cameraman can adjust this matrix and process the output signals (linearly as well as nonlinearly) to produce an artistically desirable picture on the faceplate of displays with standard or *de facto* standard primaries. The chromaticities listed in television standards are not enforceable, but are taken as guidelines for designing receivers and displays.

Progress in Psychophysics of the Human Visual and Aural Systems

Current television standards are based on an understanding of psychophysics which dates back a number of decades. Although current standard analog television systems are clever, economical, and robust, it is clear that they make inefficient use of channel capacity. In a 6-MHz transmission channel, over 90 percent of the energy of the visual information occupies a bandwidth of a few hundred kHz. With the advent of digital television, methods for effective *reduction of irrelevancy and redundancy* in communication of moving pictures and audio benefit from a new technology basis in digital compression.

PICTURE SIZE, SCANNING PROCESS, AND RESOLUTION LIMITS

Picture Size and Scanning Raster

In analog TV systems, the transmitted signal was designed to be directly related to the image representation of the camera and the display. The white rectangular area in Fig. 21.1.8 shows the *active area* of the picture, i.e., the area containing pictorial information. The active area is assumed to be flat. The height of the active area is the length unit ph (picture height) and the width A ph of the active area is the *aspect ratio*. In current standard television systems the aspect ratio is 4/3 and in HDTV it is 16/9. The viewing distance D is commonly assumed to be 6 to 8 ph in standard TV and 3 to 4 ph in HDTV.

Also shown in Fig. 21.1.8 is an imaginary *total picture area* that includes inactive areas representing durations used in the scanning process for synchronization and retrace blanking. These durations are referred to as the *vertical blanking interval VBI* and the *horizontal blanking interval HBI*. The *total vertical height* including the VBI is $1/\eta_V$ and the *total horizontal width* including the HBI is A/η_H where η_v and η_H are vertical and horizontal scanning efficiencies. They are shown in Table 21.1.4 for analog TV systems and their standard definition DTV equivalents.

In all television systems the total picture is scanned along horizontal *lines* from left to right while the scan lines progress downward from the top of the picture. The duration of a horizontal scan line including the duration of the HBI is $H = 1/F_H$, where F_H = line rate in Hz. The duration of a *field* including the VBI is $V = 1/F_v$, where F_v = field rate in Hz. In television standards a scanning system is specified as N/F_v where $\eta_v N = N_0$ = number of nonoverlapping lines displayed in the visible picture area (active lines). In *progressive scan* ("proscan" or "1:1 scan") displayed lines in successive fields are co-located (overlap) and consequently the total number of scan lines is $N = F_H/F_V = V/H$. In 2:1 *interlaced* scan, displayed lines in even numbered fields are interlaced with the lines in odd-numbered fields and consequently $N = 2F_H/F_V = 2V/H$. When referring to "interlaced scan" it is generally assumed to be 2:1 interlaced. Multiple interlace without storage causes visible

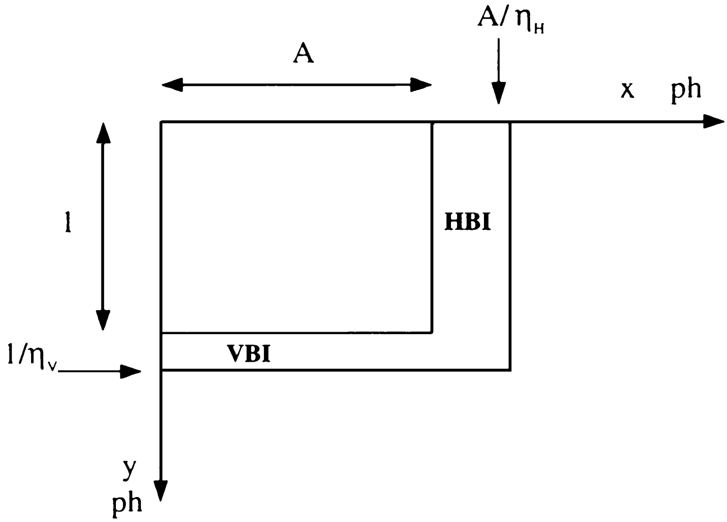


FIGURE 21.1.8 Active picture area and the total area that includes the horizontal and vertical blanking intervals. η_v and η_H are scanning efficiencies.

TABLE 21.1.4 Scanning Parameters and Scanning Efficiencies in Interlaced Analog TV Systems and Their Digital Equivalents

	Analog		Digital CCIR601	
	525/60	625/50	525/60	625/50
A	4/3	4/3	4/3	4/3
N	525	625	525	625
F_V Hz	60/1.001	50	60/1.001	50
F_H Hz	15,750/1.001	15,625	15,750/1.001	15,625
η_v	0.92	0.92	480/525	576/625
η_H	0.8285	0.8125	720/858	720/864
$N_o = N\eta_v$	483	575	480	576
$V_v = F_v/\eta_v$ ph/s	65.152	54.348	65.559	54.253
$V_H = F_H A/\eta_H$ ph/s	25,321	25,641	25,000	25,000
$V_T = F_v/N_o$ ph/s	0.1241	0.0870	0.1249	0.0868

line-crawl. In interlaced scan *one frame comprises an odd-numbered field followed by an even-numbered field*. The frame rate is $F_V/2$ Hz. In proscan a frame and a field is the same thing. Automatic interlace can be obtained if N = number of lines per frame is odd, as shown in Fig. 21.1.9. Interlaced scan can be achieved with an even number of lines per frame and proscan can be achieved with an odd number of half lines per field. The PAL/SECAM systems are 625/50(2:1) systems and the NTSC system is a 525/59.94(2:1) system or simply as a 525/60(2:1) system (although the field rate is exactly 60/1.001 Hz). Figure 21.1.9 shows the scanning raster for the interlaced 525/60 system. The scanning raster shown in Fig. 21.1.9 covers the total area including the blanking intervals. Interlaced scan has been used in all analog television systems to achieve good vertical resolution and little large area flicker, given available transmission bandwidth.

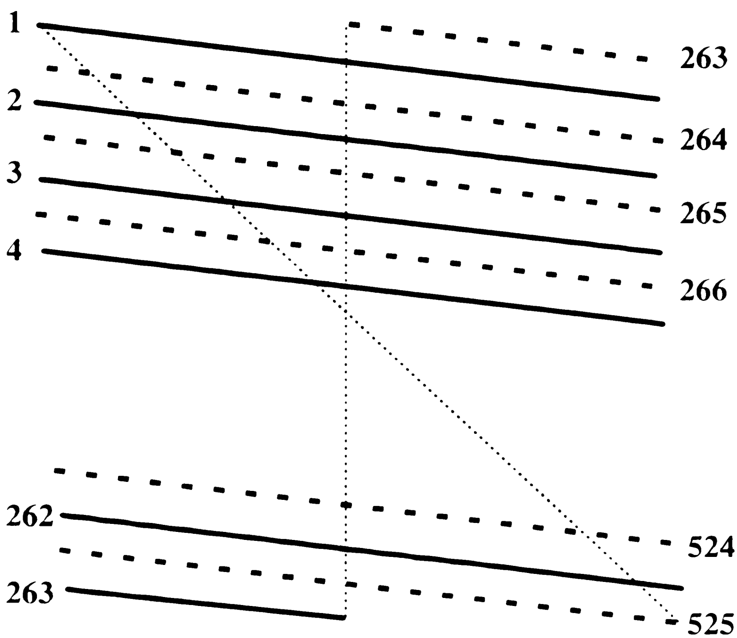


FIGURE 21.1.9 Interlaced scanning raster covering the total picture (including blanking intervals) for a 525/60 system (NTSC). Retrace during horizontal blanking is not shown.

Table 21.1.4 shows scanning parameters for various raster formats including the number of visible scan lines $N_o = \eta_v N$, the *vertical scanning velocity* $v_V = F_V/\eta_v$ ph/s, and *the horizontal scanning velocity* $v_H = F_H A/\eta_H$ ph/s in interlaced scan (v_H must be doubled for proscan). Velocities of moving objects in the picture are conveniently expressed relative to the reference velocity $v_T = F_V/N_o$ ph/s (= one line-spacing per field). In digitized systems η_H = ratio of active samples per line to total samples per line.

The Video Signal Spectrum. The dynamic image on the target of a TV camera is a three-dimensional message: two spatial and one temporal. In the scanning process this message is converted into a one-dimensional signal. Consider first a stationary image having a property, e.g., a response to exposure, which can be specified as a function g(x,y) defined over the total picture area shown in Fig. 21.1.8. This is a rectangular area of width $A/\eta_H = v_H/F_H$ and a height $1/\eta_V = v_V/F_V$. The scanning beam can be represented as progressing with time along a straight line $x = v_H t$ and $y = v_v t$ over the periodically repeated image as shown in Fig. 21.1.10. The figure shows an interlaced scan with five lines per frame. A two-dimensional Fourier series component of the periodically repeated function g(x,y) is sampled by the scanning beam to yield a one-dimensional signal:

$$C_{mn}\cos 2\pi[(mF_H/v_H)x + (nF_V/v_V)y + c_{mn}] = C_{mn}\cos 2\pi[(mF_H + nF_V)t + c_{mn}] \qquad (16)$$

where n = 0, ±1, ±2, when m = 1,2,3,…, and n = 0,1,2,3,…, when m = 0. A Fourier component in the picture is a sinewave grating with a constant phase wavefront perpendicular to the spatial frequency vector $\mathbf{k}_{mn} = [(mF_H/v_H), (nF_V/v_V)]$ = *cycles per picture height* (cph). In the scanning process this grating generates a spectral component in the signal at the discrete frequency $mF_H + nF_V$ Hz. A grating that cannot be expressed with m and n as integers has "kinks" at the borders of the total pictures shown in Fig. 21.1.10 and consists of several gratings of the type shown in Eq. (16).

Figure 21.1.11 shows a part of the spectrum of an interlaced system with N = 9 lines. Spectral components for n > 0 interleave with components for n < 0. The figure shows that a frequency component can be generated in more than one way (*aliasing*) if |n| > N/2. Thus, as expected, the highest vertical spatial frequency that can be conveyed without aliasing is $N_o/2$ cph. Aliasing shows in certain pictures: venetian blinds, striped shirts, resolution test-patterns. If the horizontal scan frequency is doubled to $2F_H$ (progressive scan with *N* lines per field), the spectral components shown as dotted lines in Fig. 21.1.11 disappear. The components are spaced F_v in proscan and $F_v/2$ interlaced scan, but aliasing occurs in both systems if |n| > N/2. Optical defocusing and vertical defocusing of the scanning beam reduces aliasing at the expense of sharpness.

As the picture changes with time the parameters C_{mn} and c_{mn} in Eq. (14) become functions of time. As a consequence the frequency components of the signal shown in Fig. 21.1.11 develop sidebands just as amplitude and phase modulated carriers. A sideband with a frequency $f = mF_H + nF_V + f_{mn}$ could have been generated by the sinewave grating defined by Eq. (16) with C_{mn} = constant and $c_{mn} = f_{mn}t$ + constant. This moving

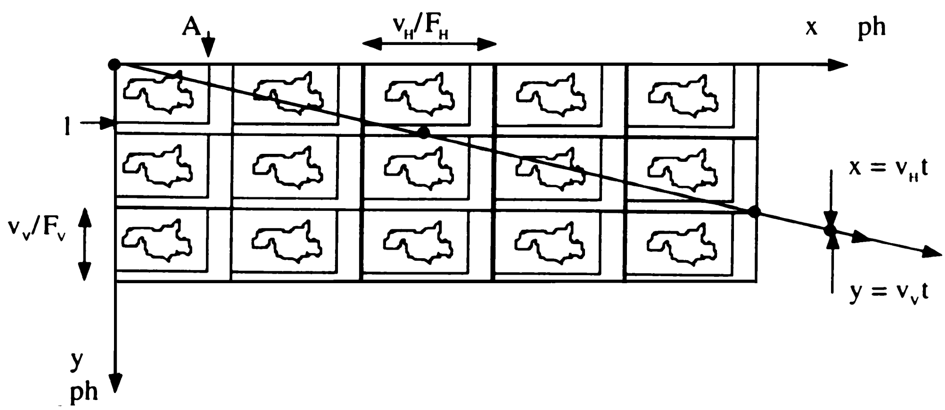


FIGURE 21.1.10 Interlaced scan of a stationary picture with *N* = 5 lines per frame.

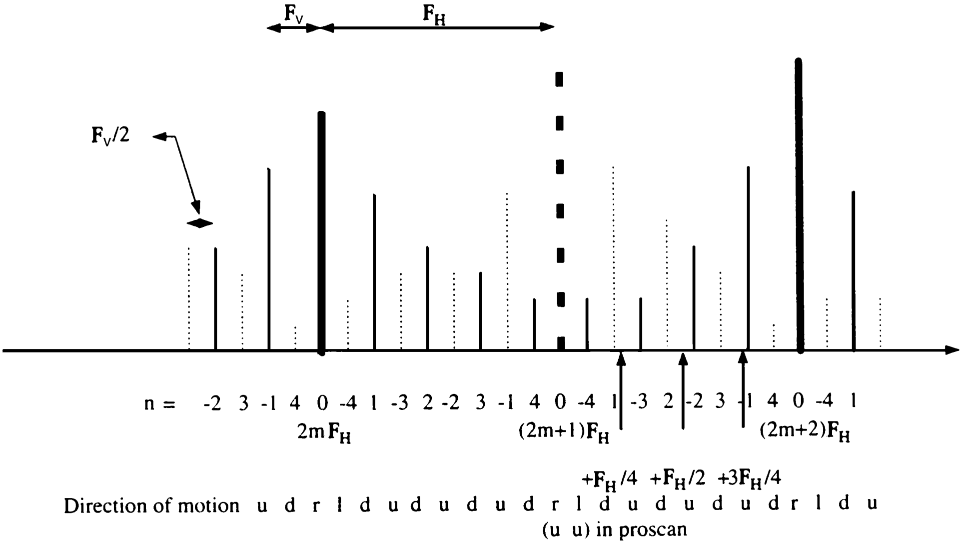


FIGURE 21.1.11 Part of the frequency spectrum generated by interlaced scan of a stationary picture with N = (4p + 1) = 9 lines per frame. In proscan at line-rate $2F_H$, dotted line components disappear. Motion directions up, down, left, right are for sidebands associated with "carriers" at the endpoints of the $F_V/2$ intervals.

grating causes the amplitude at every active point x, y in the picture to oscillate with a frequency f_{mn}. The constant phase front of the grating moves in *a direction opposite to* $\mathbf{k}_{mn}$ *with a phase velocity*

$$v_{mn} = f_{mn}/|\mathbf{k}_{mn}| = f_{mn}\,\lambda_{mn} \quad \text{ph/s} \quad \lambda_{mn} = \text{wavelength in ph} \tag{17}$$

A component at a frequency f in the video signal can, however, be generated by a number of moving sinewave gratings with different velocities and spatial frequencies, and it is up to the viewer to interpret and choose between $f = mF_H + nF_V + f_{mn}$ and $f = pF_H + qF_V + f_{pq}$. More often than not this ambiguity causes little confusion, because a moving object generates sidebands to a large number of carriers, including low-frequency carriers, with the consequence that the viewer will choose a good fit to his or her expectation of the motion. This psychophysical phenomenon is also taken advantage of in a 24 frame per second motion picture (aliasing: wagon wheels appearing as rotating the wrong way). In television, however, the confusion is accentuated by the scanning raster.

Figure 21.1.11 indicates perceived direction of motion (up, down, left, right) of a spectral component, under the assumption that the viewer interprets it as a sideband to either the nearest carrier to the left or the nearest carrier to the right. The motion designation (u, d, l, r) of these frequency intervals of width $F_V/2$ remain in proscan. Viewers tend to choose a carrier with the lowest spatial resolution, i.e., with the lowest value of $|n|$. As the velocity of a grating increases and $|f_{mn}|$ goes through $F_V/2$ the grating may appear to reverse direction of motion. This may cause visible aliasing effects, e.g., when busy pictures are panned. However, confusion is not likely if $|\mathbf{k}_{mn}| < (dF_V/2)/|v_{mn}|$ ph/s, where d = 1 in proscan and $d = 1 - 2|n|/(N - 1)$ in interlaced scan. Thus, for a given velocity, the highest nonaliasing spatial frequency depends only on d times the frame rate, which is the same for HDTV and standard TV. For example, for d = 1 and v = 1 ph/s, the maximum nonaliasing spatial frequency is 30 cph in N/60 systems and 25 cph in N/50 systems. Motion is handled better in progressive scan (d = 1) than in interlaced scan (d < 1). If C_{mn} varies with time but c_{mn} is constant there is level variation but no motion.

Resolution

Resolution is expressed in terms of the highest spatial frequencies k_{max} in *cycles per (picture) height* (cph), which can be conveyed by the television system and displayed given a scanning process and a transmission bandwidth. Sometimes resolution is still expressed in the old unit "TV lines," which should mean $2k_{max}$ cph. Sometimes horizontal resolution is expressed in cycles or in TV lines per picture width, which is misleading. Sometimes it means number of "perceived" lines of a square wave test signal. Specification of resolution in terms of TV lines and measurements of resolution with square waves should be avoided. Square waves may have lost some harmonics.

In the vertical direction, stationary information is sampled at N_o active lines per picture height. The corresponding Nyquist bandwidth $N_o/2$ cph is theoretically the maximum *vertical resolution* for stationary pictures. The vertical resolution in the NTSC system is $k_v = N_o/2 = 241.5$ cph. The "perceived" vertical resolution of stationary pictures is less than $N_o/2$ because of many factors: aliasing, source processing, scanning spot sizes, contrast, brightness, ambient conditions, test signals, viewing distance, line flicker, luminance-chrominance cross talk, and confusion caused by the display of the scanning raster (display of high spatial frequency repeat spectra). The combination of all these factors on perceived vertical resolution is sometimes expressed in terms of some perceived vertical resolution $KN_o/2$, where K = *Kell factor* < 1. The Kell factor, which has been quoted to range from 0.7 to 1. must be used with many qualifications and a great deal of caution. It is not a basic system parameter. It should preferably be avoided. In interlaced scan *line flicker* can be very disturbing in busy slow-moving pictures. In some receivers line flicker is eliminated by converting interlaced scan into progressive scan ("line doubling") by motion-adaptive interpolation.

The *horizontal resolution* of stationary pictures is $k_H = B/v_H$ cph, where B is the highest frequency in Hz of the information, which can theoretically be conveyed and which modulates "the scanning beam." For example, in the NTSC system the highest frequency of the luminance signal is B = 4,200,000 Hz. Consequently the horizontal resolution of luminance in NTSC is k_H = 4,200,000/25,321 = 166 cph. The "perceived" horizontal resolution is less than k_H depending on many factors: video filter transfer functions, luminance-chrominance cross talk, type of test signal, contrast, and so forth. The horizontal resolution of the NTSC chrominance signals are $k_I = 52$ and $K_Q = 24$ cph, which are much less than the vertical chrominance resolution (241.5 cph).

When the NTSC system was developed it was observed that horizontal and vertical *chrominance resolution* can be reduced to about half the luminance resolution without significantly reducing perceived quality. This observation was crucial for the success of NTSC.

Table 21.1.5 shows the theoretical maximum bandwidth and resolution of luminance and chrominance in various systems. In the analog TV systems chrominance resolution is significantly reduced in the horizontal

TABLE 21.1.5 Potential Resolutions in cph in the Interlaced Systems Listed in Table 21.1.2 and the HDTV Common Usage Format (The Digital CCIR System is a 4:2:2 System with a Luminance Sample Rate 13.5 MHz and Chrominance Sample Rate 6.75 MHz. In the 1125/60 System, Luminance is Sampled at 74.25 MHz and in the 1250/50 System Luminance is Sampled at 72 MHz Nyquist Maximum Frequencies and Resolutions Are Shown for all Sampled Systems)

		Analog (A = 4/3)		Digital CCIR 601[4]		Digital HDTV (A = 16/9)		
		NTSC	PAL	525/60	625/50	750/60[6]	1125/60	1250/50
M_0/N_0		443/483	520/575	720/480	720/576	1280/720	1920/1080	2048/1152
Max.Hor.Y Freq. in MHz		4.2	5[3]	6.75[4]	6.75[4]	37.125	37.125	36
Max.Hor R – Y Freq. in MHz		1.3[1]	1.3	3.375[4]	3.375[4]	18.5625	18.5625	18
Max.Hor B – Y Freq. in MHz		.6[2]	1.3	3.375	3.375	18.5625	18.5625	18
Y′	Hor.	166	195[3]	270[4]	270[4]	360	540	576
Y	Vert.	241.5	288	240	288	360	540	576
R – Y	Hor.	51.5[1]	50.5	135[4]	135[4]	180	270	288
B – Y	Hor.	24[2]	50.5	135[4]	135[4]	180	270	288
R – Y, B – Y	Vert.	241.5[1,2]	288*	240	288	180	270[5]	288[5]

(1) Applies to I = (R – Y)cos33° – (B – Y)sin33° (2) Applies to *Q* = (R – Y)sin 30° + (B – Y)cos33° (3) 5.5 MHz in PAL/I and 6 MHz in SECAM. The corresponding max resolutions of Y are 214.5 and 234 cph, respectively. *144 in SECAM. (4) *Standard CCIR601* luminance and chrominance bandwidths are 5.75 and 2.75 MHz respectively and A = 4/3. Corresponding resolutions are *230 and 115 cph*. (5) In the HDTV systems the resolution of R – Y and B – Y is half the luminance resolution horizontally as well as vertically. (6) Progressive scan.

direction, but it is not reduced in the vertical direction. In digital TV with rectangular pixels the maximum resolution is the Nyquist bandwidth in cph determined by the horizontal and vertical distances between neighboring samples. The meaning of 4:2:2 is that for every four samples of the luminance signal Y along a horizontal line there are two samples of B – Y and R – Y. In 4:2:2 systems the chrominance resolution is half the luminance resolution in the horizontal direction. In 4:2:0 systems the chrominance resolution is half the luminance resolution both horizontally and vertically. Tables 21.1.4 and 21.1.5 show a 4:2:2 version of CCIR601, while HDTV systems are 4:2:0 versions. The resolutions shown in Table 21.1.5 can be related to the luminance resolution of the HVS, which is at most 250 cph and 500 cph at viewing distances of 8 and 4 ph, respectively. It should be emphasized that the resolutions shown in Table 21.1.5 are theoretical maxima. The resolutions of the pictures displayed by a receiver depend on many factors and are usually much lower than the maximum resolutions shown in Table 21.1.5. One reason is that the signals conveyed by the system are not proportional to the tristimuli Y, R – Y, B – Y. In all systems the tristimuli are nonlinearly compressed before matrixed and bandlimited to form video signals Y', $B' - Y'$, and $R' - Y'$. Whatever the resulting resolutions are in cph, the displayed resolution in cycles per visible height is usually lower owing to the overscan needed to allow for deflection tolerances.

In summary, the critical resolutions in cph, given the number of visible scan lines N_o, a maximum horizontal frequency B Hz, a grating phase velocity υ, and a viewing distance D ph are:

$$k_{max}\,(\text{hor.}) = B/v_H, \qquad k_{max}\,(\text{vert.}) = N_o/2, \qquad k_{max}(\text{move}) = dF_v/2v, \qquad k_{max}(\text{HVS}) \cong 2000/D \tag{18}$$

where d = 1 in proscan and $1 - 2|k_y|/N_z$ in interlaced scan (k_y = vertical spatial frequency of grating in cph).

Table 21.1.5 summarizes the static resolutions of the interlaced systems listed in Table 21.1.4 and the ATSC (Grand Alliance) HDTV system formats. The IQ chrominance axes in NTSC were chosen because the HVS can better resolve chrominance detail in the I direction (approximately red-cyan) than in the Q direction (approximately blue-yellow).

Standard Frequencies in Color Television Systems

Figure 21.1.12 shows that the key frequencies used in worldwide color television systems are related to a common frequency = 2.25 MHz. Multiples of this frequency are used as international standards for sample rates in HDTV and in the digital CCIR601 standard used in production and in the professional tape recording in the D1 format. The horizontal line rate in NTSC is F_H = 2.25/143 MHz and in PAL and SECAM it is F_h = 2.25/144 MHz. The field rate in NTSC is 60/1.001 Hz and in PAL and SECAM it is 50 Hz. ATSC Grand Alliance HDTV formats when operating with 60 Hz field rates have frequencies related to 2.25 MHz, but when operating with 60/1.001 they are related 2.25/1.001 MHz.

In the NTSC, PAL, and SECAM systems, which are *compatible* with black and white reception, most of the chrominance information R – Y and B – Y is conveyed by modulating one or two subcarriers to form a *chrominance subchannel*, which is added to (multiplexed with) the signal that carries most of the luminance information Y. The resulting signal is a *composite video signal*. In the FCC rules the NTSC color subcarrier frequency is specified to be f_c = 315/88 MHz = 227.5 F_H = 3.58…MHz. In PAL the color subcarrier is specified to be f_c = 283.75F_H + F_v/2 = 4,433,618,75 MHz. SECAM operates with two frequency modulated subcarriers: f_{ob} = 4.25 MHz and f_{or} = 4.40625 MHz. The subcarriers are chosen to provide acceptable compatibility with black and white reception and to minimize visible cross talk between luminance and chrominance in color reception.

The peculiar frequencies in NTSC resulted from a slight modification of the monochrome standards (60 Hz and 15.75 kHz) for the purpose of reducing the visibility of a beat between the color subcarrier and the 4.5 MHz sound carrier. F_H was chosen to be 4.5/286 MHz to make the beat fall at an odd multiple of F_H/2 ((286 – 227.5)F_H = 920 kHz). This is how the 2.25 MHz frequency came about. In digital processing of NTSC and PAL signals, including professional recording in the *D*2 format as well as in signal processing in consumer products, 4f_c is often used as the sample rate. This sample rate exceeds the Nyquist rate somewhat in most applications, but it is readily available and phase-locked to F_H and F_v. It is particularly well suited for sampling composite video signals because it performs the function of a synchronous detector separating the in-phase and quadrature chrominance component which modulate the color subcarrier.

Some frequencies used or proposed for use in television are not well related to the frequencies shown in Fig. 21.1.12. One is 24 frames per second used in motion pictures. In NTSC broadcasting the movie frame rate

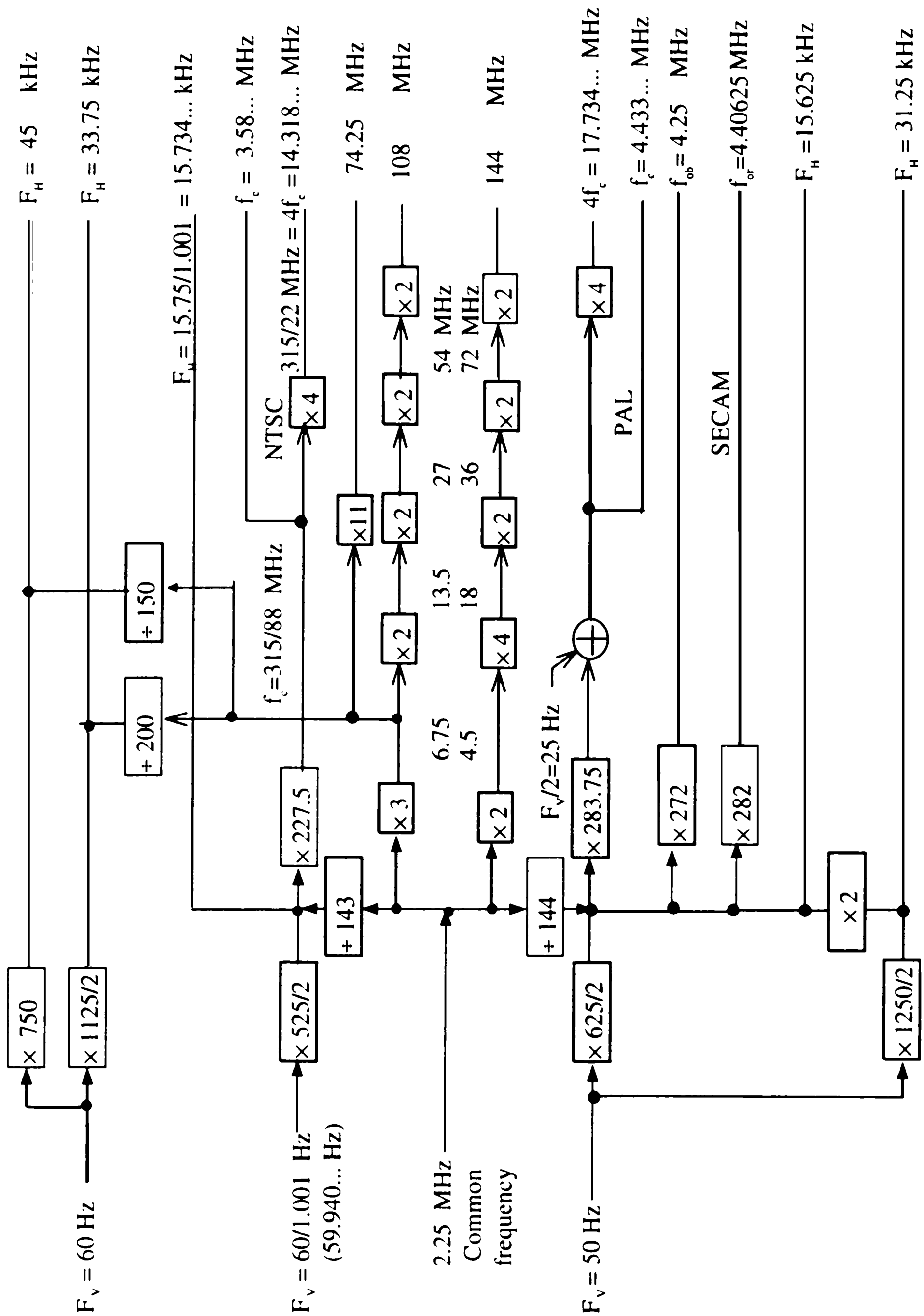


FIGURE 21.1.12 Relations between frequencies in standard analog and digital television systems.

is first reduced to 24/1.001 Hz. This is followed by a so-called 3-2 pull-down used to fit four movie frames to 10 NTSC fields, alternating three fields and two fields per movie frame. In Europe it has been common to speed up the movie by a ratio 25/24, but this increases the pitch of the sound unless digital pitch-preserving nonlinear time base techniques are used to accomplish the audio speedup. With digital techniques, better frame conversions can now be made. Other frequencies that are not well related to 2.25 MHz are: digital audio sample rates 48 and 44.1 kHz (8/375 and 49/2500 times 2.25 MHz).

ANALOG VIDEO SIGNALS

Gamma Pre-Correction

Three signals R, G, and B emerge from the camera amplifiers and associated signal processing circuits (aperture correction, peaking, correction for camera nonlinearity, and so forth). These signals are assumed to be proportional to the tristimulus values R, G, B defined by standard or *de facto* standard chromaticities and the illuminant of receiver display devices (see Photometry, Colorimetry, and the Human Visual System in this chapter). These signals and the tristimuli are all normalized so that

$$R = \mathrm{R} \quad G = \mathrm{G} \quad B = \mathrm{B} \quad \text{and} \quad R = G = B = 1 \text{ for reference white highlight} \tag{19}$$

In television standards the signals R, G, B are usually denoted E_r, E_g, E_b to distinguish them from tristimulus values. Simplified notations are used in what follows: italics for signals and romans for optical tristimulus values.

True luminance and chrominance signals are:

$$Y = \mathrm{Y}_r R + \mathrm{Y}_g G + \mathrm{Y}_b B \quad M = R - Y \quad N = B - Y \tag{20}$$

$$R = G = B = Y \text{ for white with luminance } \mathrm{Y} = Y \tag{21}$$

The signals M, Y, N are related to the signals R, G, B by the matrix H [Eq. (5) and Table 21.1.3]. It is noted that only the luminance values Y_r, Y_g, Y_b appear in this matrix. Tristimuli components X and Z as well as chromaticities, while implicit in the significance of M, Y, and N in electro-optical conversion in the display are of no concern in video signal analysis.

The original intent of the NTSC was to convey the "true" signals M, Y, N over separate channels. This was referred to as the *constant luminance principle*, because all the luminance would be conveyed by the luminance signal. Reasons for this approach are:

- compatibility with black and white receivers
- luminance information requires more picture detail than chrominance
- noise is more visible in luminance than in chrominance. To minimize the visibility of noise added in transmission, the luminance signal was companded in black and white television using the nonlinear transfer characteristic of the picture tube as expander in the receivers. The idea of using the picture tube as an expander has also been adopted in color television

To display a red color with a tristimulus value R, a nonlinear display must be driven by a "*gamma corrected*" signal R' which is a function of R. It is assumed that green and blue have the same electro-optical transfer functions and must be driven by "gamma corrected" signals

$$R' = \mathrm{F(R)} \quad G' = \mathrm{F(G)} \quad B' = \mathrm{F(B)} \quad \text{where} \quad \mathrm{F}(0) = 0 \quad \text{and} \quad \mathrm{F}(1) = 1 \tag{22}$$

assuming display systems with inverse electro-optical transfer functions

$$\mathrm{R} = \mathrm{F}^{-1}(R') \quad \mathrm{G} = \mathrm{F}^{(1}(G') \quad \mathrm{B} = \mathrm{F}^{-1}(B') \quad \text{where } \mathrm{F}^{-1}(0) = 0 \quad \text{and} \quad \mathrm{F}^{-1}(1) = 1 \tag{23}$$

In NTSC and PAL/SECAM standards, these functions are as in black and white television:

$$R' = R^{l/\gamma}, \quad G' = G^{l/\gamma} \quad \text{and} \quad B' = B^{l/\gamma} \tag{24}$$

In NTSC, $\gamma = 2.2$ ("but not enforced"). In PAL and SECAM, the recommendation is $\gamma = 2.8$. Recent standards are listed as electro-optical transfer characteristics in Table 21.1.1.

The nonlinear transfer characteristics of picture tubes usually differ from the production standards $L = F^{-1}(V)$ listed in Table 21.1.1. It is often assumed that they are close to $L = V^{2.5}$.

The Luminance Signal

In all current analog and digital standard television systems the gamma-corrected primary signals R', G', B' are matrixed in a matrix T to form signals for transmission which are proportional to:

$$Y' = T_R R' + T_G G' + T_B B' \qquad M' = R' - Y' \qquad N' = B' - Y' \tag{25a}$$

$T_R + T_G + T_B = 1$. When the color is gray $R' = G' = B' = Y'$ and $M' = N' = 0$. Conversely

$$G' = Y' - (T_R/T_G)M' - (T_B/T_G)N' \qquad R' = Y' + M' \qquad B' = Y' + N' \tag{25b}$$

Y' is called the *luminance signal*, while signals proportional to M' and N' are called *chrominance signals*. In standard definition television (SDTV), including NTSC, PAL, SECAM, CCIR601, and SMPTE170M, the transmission coefficients T_R, T_G, T_B, are equal to the luminance components Y_r, Y_g, Y_b of the NTSC primaries (Table 21.1.2), i.e.,

$$Y' = 0.299R' + 0.587G' + 0.114B' \tag{25c}$$

The coefficient should *not* be rounded to two decimals as they unfortunately are in the official FCC standards. In HDTV, the transmission coefficients are equal to the luminance components of the primaries associated with the system (Table 21.1.2), i.e.,

$$Y' = 0.2125R' + 0.7154G' + 0.0721B' \ldots \text{SMPTE274M and ITU-4 709} \tag{25d}$$

$$Y' = 0.212R' + 0.701G' + 0.087B' \ldots \text{SMPTE240M} \tag{26e}$$

While the signal Y' is traditionally referred to as the luminance signal or just "the luminance" it is not a function of true luminance Y only. This happens only when the color is gray. For all other colors, part of the true luminance information is conveyed by the chrominance signals. When the chrominance (color) in a color receiver is turned off, the displayed luminance in the resulting black-and-white picture is less than in the color picture in all but originally gray areas. The ratio of the displayed luminance when the chrominance signals are turned off (compatible monochrome TV) to the displayed luminance when they are turned on can be taken as a viewer's perception of how much true luminance is conveyed by luminance signal Y'. This ratio is $G(Y')/Y$, where $G(V)$ is the transfer characteristic of the display. The ratio can be calculated given $G(V)$, Eqs. (20) and (25), and parameters for various systems given in Table 21.1.2. The ratio becomes smaller with increased saturation of the colors, and becomes exceptionally small for saturated blue. For a display with $G(V) = V^{2.5}$ the ratio for saturated blue is 3.8 percent in NTSC, 6.2 percent in PAL/SECAM, and 2 percent in ITU-R709. For $G(V) = F^{-1}(V)$ according to Table 21.1.1, corresponding numbers are 7.4, 3.2, and 22.2 percent. While saturated blue is an extreme case and the ratios are sensitive to the transfer function of the display at low signal levels, it is clear that a significant amount of luminance information is conveyed by the chrominance channels. Since the chrominance channels have less bandwidth, and in HDTV also less vertical resolution than the luminance channel, luminance resolution is lost, particularly in saturated colors. Another consequence is that noise (including various coding and transmission errors and defects) in the chrominance channels is displayed as luminance noise. The HVS is more sensitive to luminance noise than to chrominance noise. These effects are consequences of the violation of the constant luminance principle caused by pre-gamma correction of the R, G, B primaries.

The Chrominance Signals

The chrominance signals, also referred to as the *color difference signals*, are proportional to $M' = R' - Y'$ *and* $N' = B' - Y'$ and defined in Eq. 25*a*, *b*, *c*, *d*, and *e*. The N' blue-yellow range $\pm(1 - T_B)$ is larger than the M'

TABLE 21.1.6 Chrominance Scale (gain) Factors K_R and K_B Multiplying the Basic Chrominance Signals $M' = R' - Y'$ and $N' = B' - Y'$ to Yield Transmitted Chrominance Signals

SYSTEM	$K_R(R' - Y')$	$K_B(B' - Y')$	K_R	K_B
NTSC, PAL, SMPTE170M	V	U	0.877 (1/1.14)	0.493 (1/2.03)
SECAM*	D_R	D_B	– 1.902	1.505
CCIR601†	C_R	C_B	0.713 (.5/.701)	0.564 (.5/.886)
ITU-R709, SMPTE274M	P'_R	P'_B	0.6349 (.5/.7875)	0.5389 (.5/.9279)
SMPTE240M	E'_{PR}	E'_{PB}	0.6345 (.5/.788)	0.5476 (.5/.913)

*Normalized frequency deviations for FM modulation.
†Includes all SDTV systems when luminance and chrominance components are digitized. See MPEG standards Rec. ITU-T H262.

red-cyan range $\pm(1 - T_R)$. In transmission, the chrominance signals $M' = R' - Y'$ *and* $N' = B' - Y'$ are multiplied by scale (gain) factors K_R and K_B. Scale factors of digitized chrominance signals are $K_R = (1 - T_R)/2$ and $K_B = (1 - T_B)/2$ to yield scaled chrominance signal ranges of ±0.5, equal the unity range of the luminance signal Y'. Standard *chrominance scale factors* are shown in Table 21.1.6.

In NTSC and PAL the chrominance signals can also be represented by a vector

$$C = [U, V] = [0.493(B' - Y'), 0.877(R' - Y')] = [C\cos(c), C\sin(c)] \tag{26}$$

$$C = (U^2 + V^2)^{1/2} \text{ amplitude} \qquad \text{and} \qquad c° = \arctan(V/U) = \text{phase in degrees}$$

The chrominance vector $\boldsymbol{C}$ is shown in Fig. 21.1.15. The vertical V-axis is in a "reddish" direction and the horizontal U-axis, in a "bluish" direction, is the reference direction of zero phase. The chrominance vector is displayed in an instrument called a *vector scope* as illustrated at the bottom of Fig. 21.1.16. Table 21.1.7 shows Y', M', N', V, U, C, and $c°$ for white and for the saturated colors yellow, cyan, green, magenta, red, and blue. These are the colors shown in the ubiquitous color bars. While the color bar signals are the same in NTSC, PAL, and SECAM standards, the colors vary from system to system as well as within each system depending on the chromaticities of the display and adjustments of primary levels (gains).

In PAL the chrominance signals $V = .877(R' - Y')$ and $U = .493(B' - Y')$ are both bandlimited to 1.3 MHz. Because of lack of available transmission bandwidth in NTSC the V and U signals are further matrixed in a rotational network (33°) yielding the transmitted signals $I = .839V - .545U$ and $Q = .545V + .839U$. The I and Q axes are shown in Figs. 21.1.15 and 21.1.16. The NTSC standards specify the maximum I-frequency to be 1.3 MHz and the maximum Q-frequency to be 0.6 MHz. The reason for this choice is that the HVS has better resolution for colors along the I-axis (red-cyan) than for colors along the Q-axis (blue-yellow). More often than not I and Q both roll off toward a max. frequency < 0.6 MHz.

TABLE 21.1.7 Luminance and Color Difference Signals for Saturated Colors (Color Bar Signal) in SDTV. The Sum of Complementary Colors (180° Apart) is White

COLOR	R' G' B'	Y'	$R' - Y'$	$B' - Y'$	V	U	C	$c°$
White	1 1 1	1.000	0	0	0	0	0	0
YEllow	1 1 0	0.886	+0.114	−0.886	+0.100	−0.437	0.448	+167.11
CYan	0 1 1	0.701	−0.701	+0.299	−0.615	+0.147	0.632	−76.56
Green	0 1 0	0.587	−0.587	−0.587	−0.515	−0.289	0.590	−119.30
MAgenta	1 0 1	0.413	+0.587	+0.587	+0.515	+0.289	0.590	+60.70
Red	1 0 0	0.299	+0.701	−0.299	+0.615	−0.147	0.632	+103.44
Blue	0 0 1	0.114	−0.114	+0.886	−0.100	+0.437	0.488	−12.89

Video Component Signals

While the primary signals *R, G, B* and the precorrected signals *R′, G′, B′* are video component signals, what is usually meant by *video component signals* is the set consisting of the luminance signal *Y′* and the chrominance signals $K_R(R' - Y')$ and $K_b(B' - Y')$. Figure 21.1.13 illustrates an NTSC transmission system of *video component*

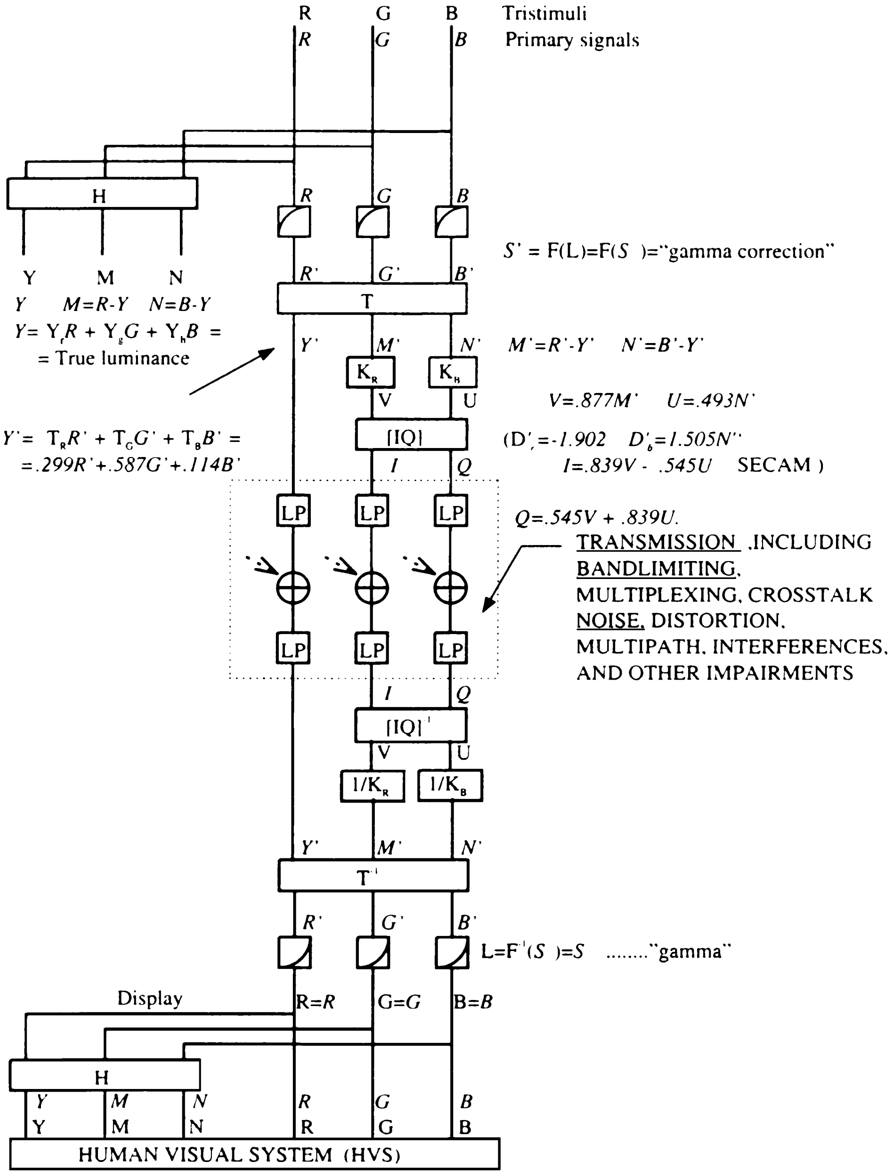


FIGURE 21.1.13 NTSC video transmission system. In PAL and SECAM drop the IQ matrices. In SECAM replace *V* by D'_r and *U* by D'_b. In digital systems, K_R and K_B are different. In HDTV the T matrices are different but, as in NTSC, the H and T matrices are identical.

signals. By dropping the IQ and IQ^{-1} matrices and ignoring numerical illustrations of K_R, K_B, T_R, T_G, and T_B it is valid for all television systems. All SDTV systems use the same T-matrix (NTSC). The T and H matrices are identical in NTSC as well as in HDTV systems. The H matrices, which are not in the transmission path, only show conversion from R, G, B to true luminance and chrominance Y, M, N. The block with dotted borders shown in Figure 21.1.13 represents bandlimiting and impairments of the component signals in transmission. In digital television the block includes data compression. A/D and D/A conversions, contributing with quantization and compression errors, are usually also inside this block. If this block is bypassed, the transmission is perfect if the processes at the transmitter have corresponding inverse processes at the receivers. Many proposals have been made to "correct" for luminance lost in the narrowband chrominance channels. Since the early 1950s there have also been proposals to replace Y' [see Eq. (25)] with the compressed luminance signal F(Y) [see Eqs. (20) and (22)], thereby conveying all the luminance in the luminance channel. No such *constant luminance* standards have been adopted, partly because it would not be compatible with current receivers and production standards, and partly because of added receiver costs for nonlinear expansion.

The NTSC Composite Video Signal, the HBI, and the Color Bar Signal

In NTSC, PAL, and SECAM the video component signals are multiplexed to form a single signal. In NTSC and PAL the chrominance signals modulate *a color subcarrier* which is "inconspicuously" added to the luminance signal to form a *composite video signal*. SECAM uses two frequency modulated subcarriers. In NTSC there is only one *color subcarrier* at a frequency $f_c = 227.5 \times F_H = 315/88$ MHz $\cong$ 3.58 MHz. It is amplitude modulated in-phase and in quadrature by the chrominance signals (QAM) and added to the luminance signal to form the *composite NTSC signal*:

$$S = Y' + U\sin(\omega_c t) + V\cos(\omega_c t) = Y' + Q\sin(\omega_c t + 33°) + I\cos(\omega_c t + 33°) = Y' + C\sin(\omega_c t + c°)$$

$$V = .877(R' - Y') \qquad U = .493(B' - Y') \qquad I = .839V - .545U \qquad Q = .545V + .839U$$

$$C = (U^2 + V^2)^{1/2} = (Q^2 + I^2)^{1/2} \qquad C° = \arctan(V/U) = 33° + \arctan(I/Q) \tag{27}$$

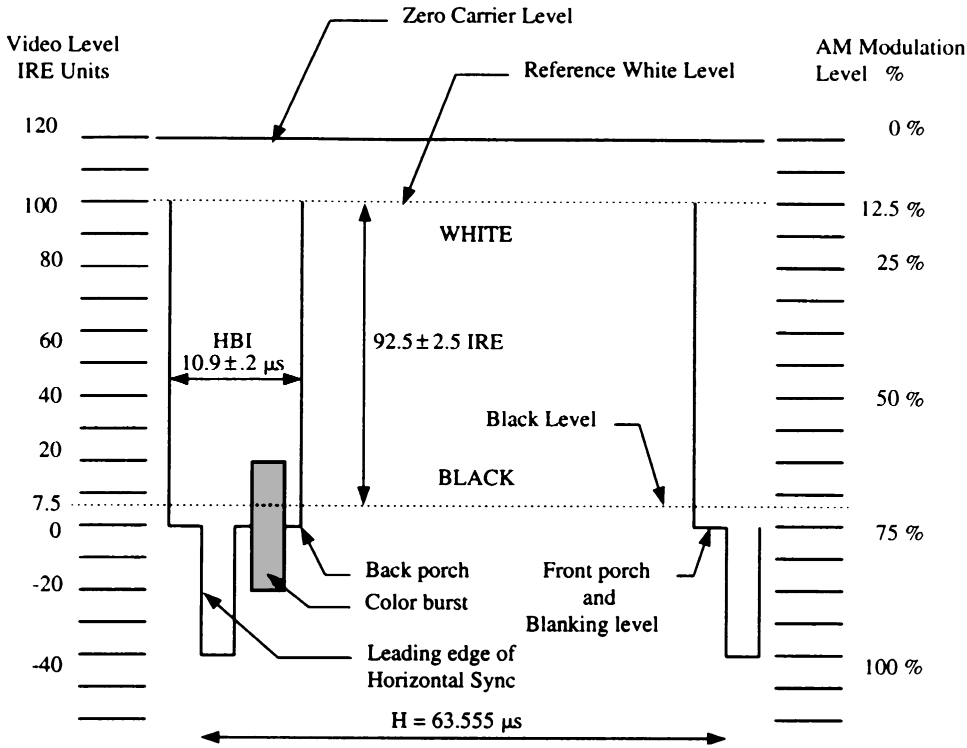
The vector diagram in Fig. 21.1.15 illustrates the instantaneous level of the color subcarrier, given the chroma vector [U,V]. A reference phase (c = 180°, "yellow") for synchronous detection is transmitted during the horizontal blanking interval by a short *burst* (8 to 11 cycles) of the color subcarrier (see Figs. 21.1.14, 21.1.16, and 21.1.17. The subcarrier components are [0, V] when $\omega_c t = 0$, [U,0] when $\omega_c t = 90°$; and (–0.4,0) for the burst at $\omega_c t = 270°$. As time progresses the instantaneous NTSC color subcarrier amplitude progresses in the order $V, U, -V, -U$ or $I, Q, -I, -Q$, as illustrated in Fig. 21.1.15.

The signal S occupies active timeslots along horizontal scan lines and is bordered by horizontal and vertical blanking signals as shown in Figs. 21.1.14, 21.1.16, and 21.1.17 which also show the timing signals: horizontal and vertical sync pulses as well as the color subcarrier burst. Video signal levels in NTSC are specified in IRE units. The peak of the sync pulses is at –40 IRE, the blanking level at 0 IRE, and the reference white level at 100 IRE. The black level "setup" is at +7.5 IRE. Figure 21.1.14 includes a table with specification of durations, levels, and tolerances for NTSC and most PAL standards. Figure 21.1.14 also shows that the composite signal "inversely" modulates the main carrier of a broadcast transmitter with zero carrier (0 percent) at the whiter-than-white level of 120 IRE and 75 percent at blanking level and 100 percent at – 40 IRE (peak of sync).

Table 21.1.7 shows that the peak composite signal $Y' + C$ for saturated yellow, cyan, and green would exceed reference white level and overmodulate a broadcast transmitter. Similarly the lower level $Y' - C$ would drop significantly below blanking level, which may cause sync problems. To transmit the composite signal S within tolerable limits it must be reduced to $g_s S$, where g_s is a gain factor <1:

$$\text{NTSC video signal} = 7.5 + g_s \times 92.5 \times [Y' + C\sin(\omega_c t + c)] \qquad \text{IRE units.} \tag{28}$$

In NTSC and PAL/BG the color bar signal as broadcast is at 75 percent amplitude (g_s = 0.75).

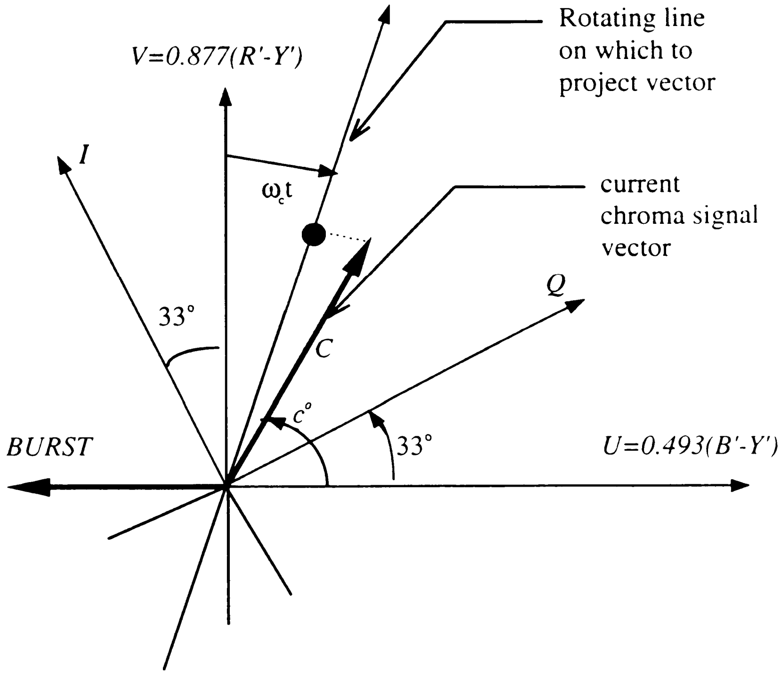


	NTSC	PAL/BDGHI
H	63.555 µs	64 µs
HBI	10.9 ±.2 µs	12 ± .3 µs
Sync pulse	4.7 ± .2 µs	4.7 ± .2 µs
Front porch	1.75 ± .48 µs	1.5 ± .3* µs
Burst	9 ± 1 f_c-cycles	10 ± 1 f_c-cycles
Burst level	40 IRE peak-peak	0.3 Volts peak-peak
Black level	7.5 ± 2.5 IRE	.01 ± .01 Volts+

*1.65 ± 0.1 in PAL/I.
+ In PAL 100 IRE = 0.7 Volts. Sync pulse is 0.3 Volts and modulation levels at white reference level are 10 percent in BDG and 20 percent in I.

FIGURE 21.1.14 Signal levels along a horizontal line in NTSC. Levels are shown in IRE video level units and in r.f. modulation levels. NTSC and PAL specifications are shown in the table. Modulation is negative in the sense that the white reference level is at a low carrier level and peak of sync is at peak r.f. power level.

Figure 21.1.16 shows a horizontal line signal in NTSC with color bars at 75 percent. The peak signal just about touches the 100 IRE white reference level, but drops to −16 IRE in blue and red (peak of sync is at −40 IRE). The figure also shows the vector diagram (vector scope) of the chrominance components and the burst. In PAL the color bar signals have ratio C/Y' as in NTSC but there is no 7.5 IRE black level setup. In PAL the phase is $+c°$ and $-c°$ on alternate lines. In PAL/BG the EBU-color bars at 75 percent are



$$\text{NTSC Color subcarrier signal} = U\sin(\omega_c t) + V\cos(\omega_c t) = Q\sin(\omega_c t + 33°) + I\cos(\omega_c t + 33°)$$
$$= C\sin(\omega_c t + c)$$

FIGURE 21.1.15 Vector diagram showing quadrature modulation of the NTSC color subcarrier. With time the signal progresses in order I, $Q - I$, $-Q$.

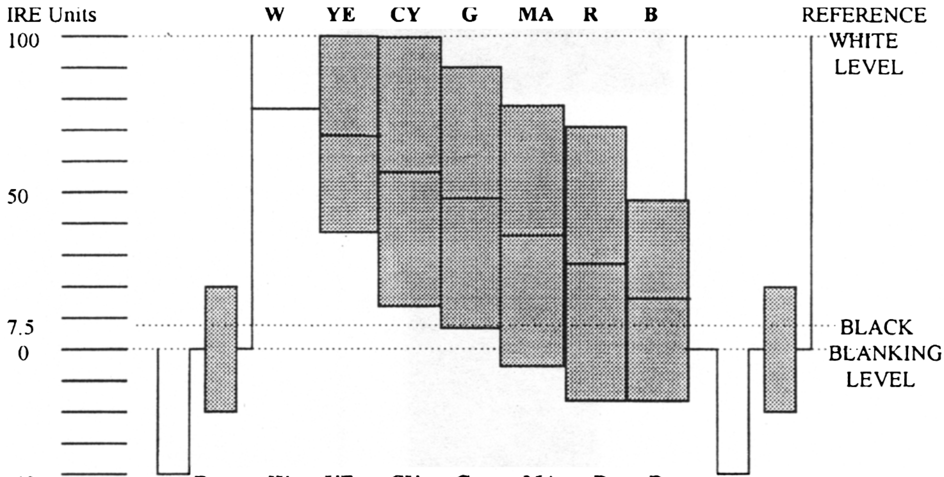
$0.3 + 0.7 \times 0.75 \times S$ volts, except that white is shown at 100 percent, i.e., at 1 V. In PAL/I they are at 100 percent amplitude and 95 percent saturation (V and U multiplied by 0.95). If saturated colors are avoided in programs, the gain g_s could be larger than 0.75, but this presents a risk for overmodulation, which causes noise in the received audio. In compatible black-and-white reception this gain reduction results in a 2.5-dB loss in signal-to-noise ratio.

The PAL Composite Video Signal

In NTSC a phase error with respect to the burst reference phase used to be a problem with early tape recorders. Adding a phase error ϕ to $\omega_c t$ in Eq. (24) causes the synchronously detected U and V outputs to be:

$$U(\phi) = U\cos\phi - V\sin\phi, \qquad V(\phi) = V\cos\phi + U\sin\phi \tag{29}$$

This UV cross talk changes the tint because of the $\sin\phi$ contributions. To abate this effect, the PAL system alternates the polarity of V and $V(\phi)$ on successive lines in a field (by switching during the HBI) and on the same line after two frames by choosing a subcarrier frequency that has an odd number of cycles in four frames = a PAL color cycle. Equation (29) shows that this is equivalent to altering the sign of the $\sin\phi$ contribution to the error. These chroma errors may then be averaged and “canceled” in the HVS. However,



	Burst	W	YE	CY	G	MA	R	B	
$Y'+C$	20	77	100	100	89	77	72	46	IRE
Y'	0	77	69	56	48	36	28	15	IRE
$Y'-C$	-20	77	38	12	7	-5	-16	-16	IRE
C	20	0	31	44	41	41	44	31	IRE
c^o	180		167	-77	-119	61	103	-13	DEG

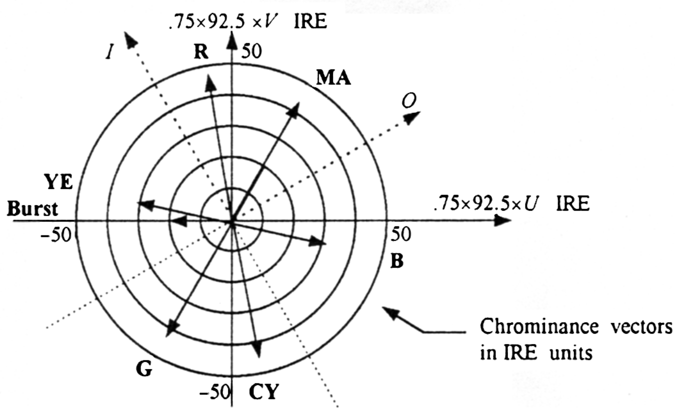


FIGURE 21.1.16 NTSC Color bars at 75 percent amplitude (EIA standard RS 189-A). See Table 21.1.8 for U, V, $B'-Y'$, $R'-Y'$, Y', R', G', B'.

the alternating errors tend to be visible as crawling ("Hanover") bars, partly due to the nonlinearity of the display. This problem is solved in "standard PAL" by canceling the sinϕ error with the use of a comb filter that averages luminance and chrominance signals on successive lines. The price is reduced vertical resolution. In PAL the phase $c°$ of the chroma signal vector **C** shown in Fig. 21.1.15 alternates sign on every other line. To identify A-lines with chroma vector (U, V) and B-lines with chroma vector (U, V) the burst swings from a phase $c = 135°$ on A-lines to $-135°$ on B-lines. The reference phase $c = 180°$ is derived by averaging burst on A and B lines. The color subcarrier progresses with time in the order V, U, $-V$, $-U$ on A-lines and in the order U, V, $-U$, $-V$ on B-lines.

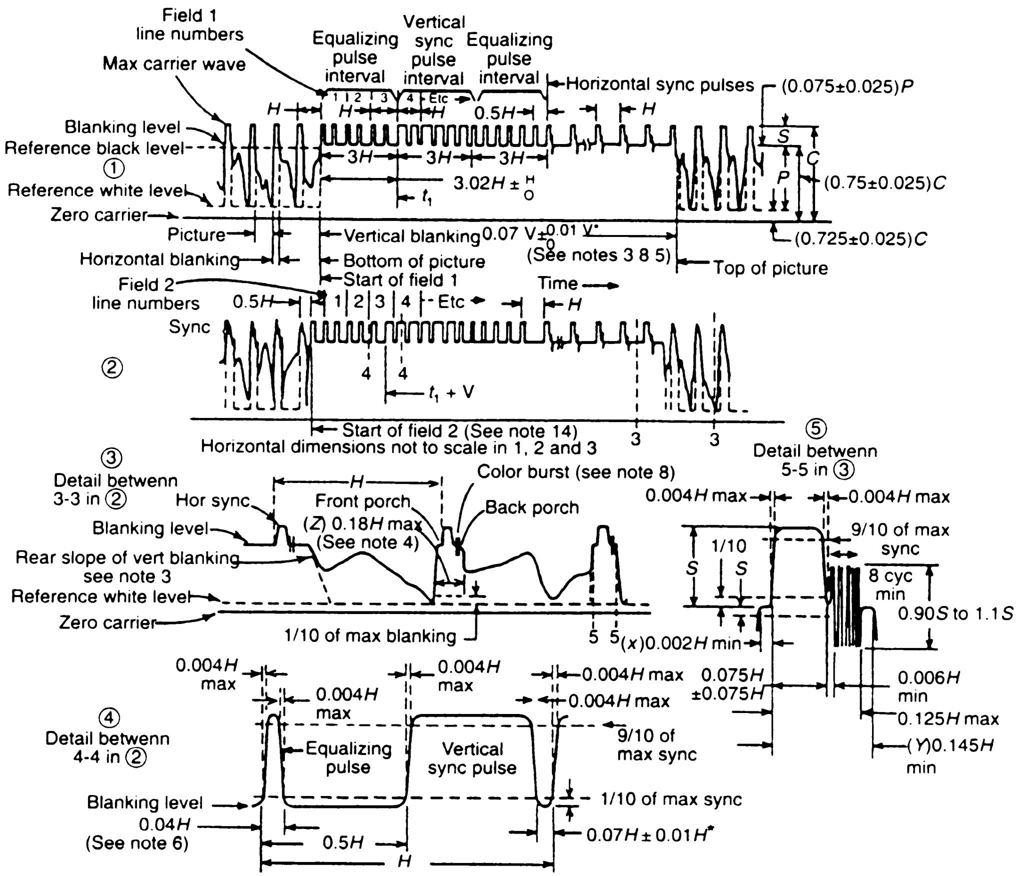


FIGURE 21.1.17 National Television System Committee color synchronizing-signal waveform (FCC Standard).

Notes: (1) H = time from start of one line to start of next line. (2) V = time from start of one field to start of next field. (3) Leading and trailing edges of vertical blanking should be complete in less than $0.1H$. (4) Leading and trailing slopes of horizontal blanking must be steep enough to preserve minimum and maximum values of $(x + y)$ and (z) under all conditions of picture content. *(5) Dimensions marked with an asterisk indicate that tolerances are permitted only for long time variations and not for successive cycles. (6) Equalizing pulse area shall be between 0.45 and 0.5 of the area of a horizontal synchronizing pulse. (7) Color burst follows each horizontal pulse but is omitted following the equalizing pulses and during the broad vertical pulses. (8) Color burst to be omitted during monochrome transmission. (9) The burst frequency shall be 3.579545 MHz. The tolerance on the frequency shall be ±10 Hz with a maximum rate of change not to exceed 0.1 Hz/s. (10) The horizontal scanning frequency shall be 2/455 times the burst frequency. (11) The dimensions specified for the burst determine the times of starting and stopping the burst but not its phase. The color burst consists of amplitude modulation of a continuous sine wave. (12) Dimension P represents the peak excursion of the luminance signal from blanking level but does not include the chrominance signal. Dimension S is the synchronizing amplitude above blanking level. Dimension C is the peak carrier amplitude. (13) Start of field 1 is defined by a whole line between the first equalizing pulse and the preceding H sync pulses. (14) Start of field 2 is defined by a half line between the first equalizing pulse and the preceding H sync pulses. (15) Field 1 line numbers start with the first equalizing pulse in field 1. (16) Field 2 line numbers start with the second equalizing pulse in field 2.

Composite Signal Spectra: NTSC, PAL, and SECAM III

In NTSC the color subcarrier frequency is $f_c = 455 \times (F_H/2) \cong 3.58$ MHz. FCC rules specify f_c as $(63/88) \times 5$ MHz ± 10 Hz. This frequency is located exactly between two horizontal frequency components and exactly between vertical components $n = \pm (N - 1)/4 = \pm 131$ marked as $+ F_H/2$ in the spectrum shown in Fig. 21.1.11.

In the vertical direction the subcarrier alternates phase on adjacent lines in a field and in the temporal direction it alternates phase on adjacent frames, completing a *color-cycle* in the duration of two frames (1/15 s or 238,875 cycles). In the picture the color subcarrier appears as an interference pattern of dots seen to move slowly upward with a velocity $v_T \cong 1/8$ ph/s, a phenomenon that can be seen on some black-and-white sets and in the green-magenta transition in color bars. For stationary pictures, spectral components of the modulated subcarrier interleave with the spectral components of the luminance, but the luminance and chrominance spectral components are now only $F_V/4$ apart as can be seen in Fig. 21.1.11. Luminance-chrominance cross talk is unavoidable and deteriorates with motion. High-frequency luminance components show up as low-frequency components in the synchronously detected chroma signals *U* and *V* or *I* and *Q*. Luminance cross talk into chrominance, referred to as *cross-color*, appears as undesirable color patterns. Chrominance cross talk into the luminance, referred to as *cross-luminance*, shows up as moving luminance high frequency patterns (dots, gratings, and so forth). Since the luminance energy is primarily concentrated at even multiples of $F_H/2$ and the chrominance energy is concentrated at odd multiples of $F_H/2$, the cross talk can be reduced by passing the frequency band with the shared spectrum through a 1H-delay comb filter. The comb filter partially separates luminance from chrominance, but also averages luminance and chrominance on successively transmitted lines, which are two lines apart in interlaced scan. The price is "tolerable" reduction of diagonal luminance resolution, tolerable reduction of vertical chrominance resolution, and *hanging dots*, which occurs when the chrominance information changes significantly from one line to the next, and some other artifacts. The picture quality is usually improved by a comb filter, particularly with some cooperative processing at the source. Chroma-luma separation can be significantly improved with more sophisticated techniques.

In PAL the color subcarrier has a frequency $f_c = (1135/4)F_H + F_V/2 = 4{,}433{,}618.75$ Hz (709,379 cycles in four frames). The PAL composite signal is

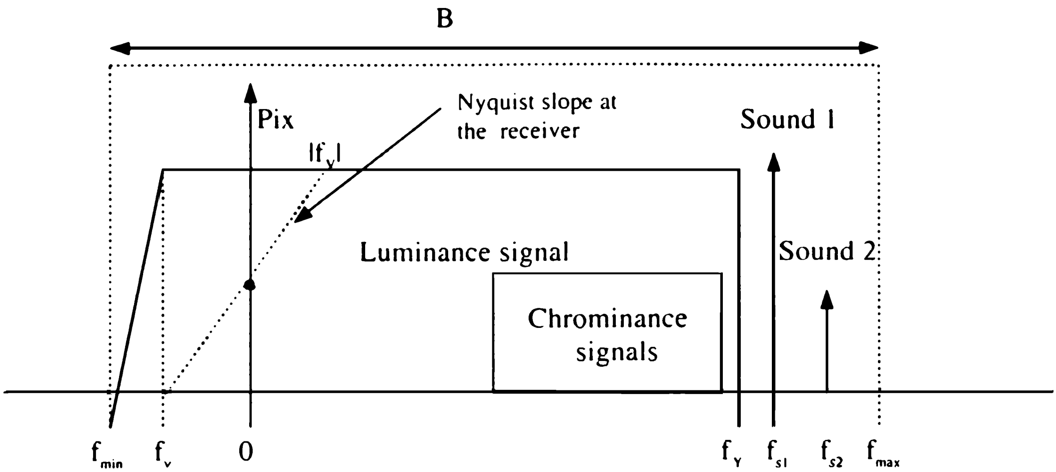
$$S = Y' + U\sin(\omega_c t) \pm V\cos(\omega_c t) = Y' + C\sin(\omega_c t \pm c) \tag{30}$$

The *V*-subcarrier is multiplied by a square wave with a frequency $F_H/2$. As a consequence, *V* modulates a number of subcarriers at frequencies $f_c \pm (2p + 1)(F_H/2)$ having amplitudes diminishing in proportion to $1/(2p + 1)$. The PAL color subcarriers appear as gratings which would move up slowly if it were not for the added 25 Hz. Shifting them all up in frequency by 25 Hz turn them into less visible gratings which appear to move three times faster downwards. A complete color cycle takes a duration of four frames = 0.16 s (709,379 cycles of the color subcarrier). An important feature of the PAL system is that the chrominance signals *U* and *V* can be partially separated with a comb filter.

In the SECAM III system (625/50) the chrominance signals U and V are conveyed on alternate lines (line sequential chroma). A 1H-delay line is used to make the U and V signals available simultaneously. Since there are an odd number of lines per frame, a line carrying U on one frame carries V in the next frame. The pre-emphasized U signal frequency modulates a subcarrier at a frequency $f_{ob} = 272\ F_H = 4.25$ MHz and the pre-emphasized V signal frequency modulates a carrier at a frequency $f_{or} = 282\ F_H = 4.40625$ MHz. The sidebands of the carriers are also pre-emphasized, and the polarity of the subcarriers (not the polarity of U and V) are altered in a sequence chosen to minimize their visibility. Flag signals are transmitted during vertical and horizontal blanking intervals to inform the receiver about parameters needed for proper detection. Specifications of the PAL and SECAM III systems are summarized in Fig. 21.1.18.

The Vertical Blanking Interval (VBI)

Figure 21.1.17 shows specifications for the vertical blanking interval in NTSC with the FCC designations of line numbers. The VBI occupies 21 lines on field 1 and 22 on filed 2. Three lines of equalizing pulses are followed by three lines of vertical sync pulses and another three lines of equalizing pulses, all needed for reliable interlaced scan in adverse reception conditions. Synchronization pulses are followed by line 10 through line 22. Most of these lines can be used compatibly (without unblanking receivers) for transmission of auxiliary information including test signals, data transmission (teletext, ceefax, antiope), and reference signals to improve reception. Each line is worth the equivalent of continuous transmission with 13-kHz bandwidth covering large areas with excellent signal-to-noise ratios. Most lines have been allocated for



	NTSC	PAL/BGH	PAL/DHK	PAL/I	SECAM III/L
B MHz	6	7(B) 8(G)	8	8	8
f(min) MHz	−1.25	−1.25	−1.25	−1.25(4)	−1.25
f(max) MHz	4.75	5.75	6.75	6.75	6.75
f(Y) MHz	4.2	5	5	5.5	6
f(v) MHz	−0.75	−0.75 (−1.25 H)	−0.75	−1.25	−1.25
f(s1) MHz	4.5	5.5	5.5	6	6.5
f(s2) MHz		5.7421875		6.552	
Luminance modulation	White Level 12.5 ± 2.5%	White Level 10 to 12.5 %	White Level 10 to 12.5 %	White Level 20 ± 2 %	Peak of sync < 6%
Audio 1 level rel max.pix dB	BTSC-Stereo FM, Δf = 0.073 −10 to −7 (3)	Mono FM, Δf = 0.050 −13 to −10	L + R (Mono) FM, Δf = 0.050 −13	(Mono) FM, Δf = .050 −10 to −7	Mono AM −10
Audio 2 level rel max pix dB			R FM, Δf = 0.050 −20	NICAM728 DPSK Stereo −20	
Chroma 1	0.5 MHz Q inphase on f_c = 3.58 MHz	0.5 MHz U inphase on f_c = 4.43 MHz	1.3 MHz U inphase on f_c = 4.43 MHz	1.3 MHz U inphase on f_c = 4.43 MHz	1.3 MHz U FM on f_c = 4.25 MHz
Chroma 2	1.3 MHz I in in quadrature on f_c = 3.58 MHz*	1.3 MHz ± V in quadrature on f_c = 4.43 MHz	1.3 MHz ± V in quadrature on f_c = 4.43 MHz	1.3 MHz ±V in quadrature on f_c = 4.43 MHz	1.3 MHz V FM, on f_c = 4.40 MHz (1)
Picture i.f. Freq. MHz	45.75 (2)	38.9 (2)	38.9 (2)	38.9 (2)	32.7 (2)

*DSBAM up to 0.5 MHz and SSB-Lower Sideband to 1.3 MHz. (1) For detailed standards of SECAM preemphasis, deviations, and level control see Ref. 1. (2) Usual practice. I. F.-spectrum is flipped, i.e., sound is below pix. (3) Cable TV practice in the United States is −15 dB. (4) Roll-off actually spills over into lower adjacent channel.

FIGURE 21.1.18 Some r.f. transmission standards for NTSC, PAL, and SECAM.

specific purposes and with specified broadcast standards. In NTSC there is an FCC standard, *Teletext*, for broadcasting a digital ancillary signal with a payload of 33 bytes per line. Teletext can be used for many purposes, including captions.

One reference signal, a *ghost cancelling reference* (GCR) intended to correct for multipath distortion in television reception, has been adopted by the FCC as a standard in the United States. Basically the receiver adapts its r.f. transfer characteristics to correct detectable errors in the GCR. This system equalizes the transfer characteristic, not only for the effects of multipath but also for other transmission errors. The GCR adopted by the FCC appears as a frequency sweep (chirp) with carefully designed leading and trailing edge transients.

THE AUDIO SIGNALS

Monophonic Sound Broadcasting

In NTSC, PAL, and SECAM broadcasting, a "sound carrier" is modulated by an audio signal and added to the "picture carrier," which is modulated by the composite video signals. The summation is usually done at a high power level in a passive linear hybrid circuit called a diplexer (Fig. 21.1.1). In NTSC and PAL the sound carrier is frequency modulated, in SECAM/L it is amplitude modulated. In NTSC the ratio of the peak picture carrier to the sound carrier is 7 to 10 dB (15 dB in CATV). As in FM radio broadcasting, the audio signals are pre-emphasized (75 μs in NTSC, 50 μs in all other systems) and limited to a bandwidth from 40 to 15,000 Hz. The frequency deviation in the NTSC monophonic standard is 25 kHz, ensuring excellent audio quality even when the picture is quite degraded by interference.

Stereophonic Sound Broadcasting

In the United States a standard (BTSC) was adopted by the FCC in 1984 for multiplexing two analog stereo signals, L + R and L − R, to form a composite "stereo baseband signal." The system is also referred to as MTS (multichannel television sound) and is described in the chapter on broadcasting.

In Germany (PAL/B,G) an additional r.f. carrier, 7 dB weaker than the main sound carrier, has been added at a frequency $15.5F_H \cong 242$ kHz above the main sound carrier, to transmit the right signal R. In this dual-sound-carrier system the compatible monophonic signal $M = L + R$ is transmitted on the main audio carrier.

In terrestrial broadcasting in Great Britain (PAL/I) an additional r.f. carrier, 10 dB weaker than the main sound carrier has been added at a frequency about 552 kHz above the main sound carrier, to convey digital stereophonic sound referred to as NICAM728 (*near instantaneous companding audio multiplex* @ 728 kb/s). To provide 15 kHz stereo the L and R signals are sampled at a 32-kHz rate with 14 bits per sample. Companding reduces the required number of bits per sample to be conveyed to 10. The 14 bits are restored in the receiver. In terrestrial broadcasting, bit-pairs, one bit from each channel, modulate the r.f. carrier in *differential quadrature phase shift keying* (DPSK) by shifting the phase of the carrier from the current phase by a multiple (0, 1, 2, 3) of 90°. NICAM was also the preferred digital format for D-MAC satellite broadcasting in Great Britain. Some SDTV television audio transmission standards are summarized in Fig. 21.1.18.

R.F. TRANSMISSION SIGNALS

Basics and Status

Video signals are transported to consumers by carriers having frequencies in various bands in the r.f. spectrum. A modulated carrier has the form

$$z(t) = x(t)\cos(\omega_c t) - y(t)\sin(\omega_c t) = E(t)\cos(\omega_c t + p(t)) \tag{31}$$

where x = in-phase signal
y = quadrature signal
$E = (x^2 + y^2)^{1/2}$ = envelope
$p = \arctan (y/x)$ = phase
$f = (1/2\pi)(dp/dt)$ = frequency deviation

In double sideband amplitude modulation (DSBAM) x is linearly related to the video signal while $y = 0$. In digital quadrature amplitude modulation (QAM) both x and y take a number of discrete levels. In phase shift keying (PSK), p takes discrete levels. In frequency modulation (FM) f carries the video information. In PSK and FM, E carries no information and is approximately constant. In orthogonal frequency division multiplex (OFDM), a vast number of carriers transport a video signal. In fiber optic transmission, the carrier is light, not always with a single frequency.

Television is broadcast to a large number of viewers by three transmission technologies: terrestrial broadcasting, cable, and satellite. Terrestrial broadcasting includes a microwave multipoint distribution system (wireless cable). Cable broadcasting includes all technologies that can deliver television into homes via some conduit. In satellite communication FM is used for analog video transmission and usually PSK (or differential PSK) for digital video transmission. While satellite broadcasting of digital television is rapidly becoming popular, open (i.e., not proprietary) signal standards are still under development. Detailed specifications of signals used in digital television satellite broadcast systems and in digital video disks (DVD) are proprietary at the time of this writing. A major effort to develop an international digital video broadcast standard is pursued by several standards organizations, in particular by the DVB group of the EBU. It is beyond this section to discuss various standards for point-to-point transmission of video, e.g., satellite and broadband digital networks (BISDN). In what follows the discussion is limited to standards for analog and digital terrestrial broadcasting in the VHF and UHF television broadcast bands (54 to 890 MHz in the United States), where the channel bandwidth constraints (6, 7, or 8 MHz) have had a major impact on video signal standards. Cable TV systems operate with essentially the same r.f. signals.

R.F. Signals in Analog Terrestrial and Cable TV Broadcasting

All high-power television stations radiate a signal consisting of a picture carrier that has a frequency close to the lower band limit of the channel and one or two sound carriers with frequencies close to the upper band limit of the channel as shown in Fig. 21.1.18. The picture carrier is amplitude modulated by the composite video signal, but only a part of the lower sideband is transmitted. This is referred to as a *vestigial sideband amplitude modulation*. VSBAM is usually generated by passing a DSBAM signal through a vestigial sideband filter. Figure 21.1.18 illustrates the bandpass characteristic of the radiated VSBAM picture signal. In the i.f. amplifiers of receivers, the signal is transferred through a “Nyquist” filter, which in theory is phase-linear and has an antisymmetric amplitude (not necessarily linear) response around the picture carrier frequency (Fig. 21.1.18). The output from the Nyquist filter consists of an in-phase signal x amplitude modulated by the desired video and a quadrature signal y amplitude modulated by a video frequency component delayed 90° (Hilbert transform of video). With a perfect Nyquist filter synchronous detection yields undistorted video, while envelope detection contains some quadrature distortion. Ghosts also contain the quadrature signal. For these reasons the Nyquist slope must not be too steep. VSBAM wastes almost half the transmitter power to reduce radiated bandwidth. In the United States some low-power TV stations are allowed to operate with DSBAM.

Figure 21.1.18 shows the radiated spectrum of the NTSC, PAL, and SECAM signals including the sound carrier(s). Also shown are a number of r.f. signal standards. The NTSC sound carrier frequency is 4.5 MHz (286 F_H) above the picture carrier frequency and is frequency modulated by the BTSC (MTS) stereo baseband signal. The power of the NTSC sound carrier is 7 to 10 dB weaker (in Cable TV 15 dB weaker) than the power of the picture carrier at peak of sync (peak power). The radiated power of an analog television signal is always specified by the peak power of the picture carrier. Modulation levels at peak of sync, blanking, and white level are shown in Fig. 21.1.14. Broadcast technology, recommended practices, and FCC rules are included in the section on broadcast transmitters.

Transmission Impairments of Analog Television Signals

Major impairments are:

- Random noise and man-made noise (including impulse noise)
- Linear distortion including multipath distortion (ghosts)
- Interference from other broadcast stations (including television co-channel interference, adjacent channel interference, and at UHF, interference from taboo channels)
- Jittery local oscillators
- Nonlinear distortion (including crossmodulation, triple beats)

Random noise ("snow"), caused primarily by receiver noise, is the principal impairment limiting the area, which can be covered by a broadcast station or the number of homes that can be reached from a cable television headend. Reception impaired by random noise is specified by the carrier-to-noise ratio (CNR) defined to be the ratio in dB of carrier peak power to total noise power in an r.f. bandwidth = B MHz. In NTSC terrestrial broadcasting the practice is to specify B = 6 MHz (TASO definition) and in cable television the practice is to specify B = 4 MHz (NCTA definition). Thus

$$\mathrm{CNR(NCTA)} - \mathrm{CNR(TASO)} = 1.76 \quad \mathrm{dB} \tag{32}$$

Video signal-to-noise ratio, unweighted for subjective perception of noise, is defined as the ratio in dB of the white-minus blanking level to rms noise in the luminance bandwidth. In NTSC the unweighted SNR is related to the CNR as

$$\mathrm{SNRU} = \mathrm{CNR(TASO)} - 5.3 = \mathrm{CNR(NCTA)} - 7.06 \ \ \mathrm{dB} \tag{33}$$

With noise weighted for subjective perception (CCIR), the NTSC weighted SNR is

$$\mathrm{SNRW} = \mathrm{SNRU} + 7.6 \quad \mathrm{dB} \tag{34}$$

CNR(TASO) is related to the field strength E dBμV/m (dBu) at the location of the receiving antenna:

$$\mathrm{CNR(TASO)} = E + G - F - 20 \log_{10}(f \ \mathrm{MHz}) + 31.1 \ \mathrm{dB} \tag{35}$$

where G is the antenna gain minus cable losses and F is the system noise figure, both in dB.

The FCC has based determination of "grade B" noise-limited coverage on a just barely acceptable CNR(TASO) = 28.2 dB to be exceeded 90 percent of the time at 50 percent of the locations. To meet this performance with FCC "planning factors" (G, F, and a 90 to 50 percent time factor), the field strength on the grade B contour must exceed E(50/50) shown in Table 21.1.8 50 percent of the time at 50 percent of the locations.

Determination of E(50/50) and the 90 to 50 percent time factor (50/10) – E(50/50) is discussed in the section on broadcast transmitters.

The effects of an r.f. impulse (*impulse noise*) in the received composite video is a pulse with a shape, amplitude, and ringing which is determined by the video bandwidth of the receiver and the phase of the carrier at the instant of the impulse. Impulse responses are colorful, and chroma streaks can be long because of the narrow chrominance bandwidth.

Co-channel interference shows up as crawling venetian blinds. Their visibility for a given relative level of the interfering channel can vary by as much as 20 dB depending on the difference between the frequencies of

TABLE 21.1.8 Grade-B Field Strength at 50 Percent of the Locations 50 Percent of the Time 9 m Above Ground

	Low VHF (54–88 MHz)	High VHF (174–216 MHz)	UHF (470–806 MHz)
Grade-B E(50/50) dBμV/m	47	56	64

the co-channels. A precise visual carrier frequency offset of ±10010 Hz is cooperatively used by many NTSC broadcasters to minimize the visibility of co-channel interference.

Multipath transmission has two effects: (1) it changes the amplitude of the carrier which carries the desired signal and consequently the CNR and (2) it causes linear distortion of the video signal which can be characterized by an r.f. to video transfer function. Because VSBAM transmission also transmits the Hilbert transform of the video signal on a quadrature carrier, part of the ghost as it appears in the picture may consist of this quadrature signal. The transfer function varies rapidly with the relative carrier-phase of the ghosts. This is noticeable if a ghost is a reflection from an airplane (*airplane flutter*). The transfer characteristic can be equalized with a *deghoster* operating with a ghost cancelling reference signal (GCR) in the VBI. The standard GCR for NTSC is discussed in the sections on television transmitters and receivers. The carrier level, and consequently CNR, can be significantly affected by ghosts, not the least caused by reflections from nearby clutter, i.e., buildings, hills, trees, and so forth. A deghoster can straighten out the transfer characteristic, but it cannot improve degradation of CNR by ghosts.

There are numerous other impairments of the video signal, some caused in the r.f. path, some in the transmitter, and some in the receiver. Distortion, distortion specifications, measurement techniques, performance standards, and vertical interval test signals (VITS) for measurements of impairments are discussed in the chapter on broadcast transmitters.

Cable TV introduces impairments resulting from nonlinear distortion in broadband amplifiers. The most significant consequences of intermodulation between the many r.f. signals passed through a cascade of broadband amplifiers, are beats between three channels (triple beats), which primarily cause streaks in the picture, and cross-modulation, whereby a carrier is modulated by video conveyed by other carriers. Progress in amplifier design has made possible the distribution of the sum of more than 50 standard VSBAM television signals on a cable with imperceptible cross-modulation and triple beat interference levels in the video below −55 dB relative to 100 IRE. For a detailed discussion see the section on Cable TV.

DIGITAL VIDEO SIGNALS

Sampling and Quantizing Video Signals

In theory, a dynamic picture is sampled at lattice points in a three-dimensional space, two spatial and one temporal, and at each point (pixel or pel), three digital values specifying a color are sampled. In practice this process is often accomplished by sampling and quantizing analog component video signals. Conventional practice is to align samples along vertical lines ("rectangular" pels) in identical locations on successive frames.

Of key importance is where in the video transmission path, shown in Fig. 21.1.13, the A/D conversion takes place, what signals are sampled, and how noisy they are. Other key parameters are: sample rate, number of quantization levels, and anti-alias filtering preceding A/D conversion. A camera with CCD or CMOS sensors may be the first device to sample the image along a line. The number of reliably distinguishable levels is limited to a few hundreds by camera noise. Aliasing in CCD cameras in the spatial directions can be partially controlled by sensor size and defocusing, and in the temporal direction by the integration time. The data sampled on the faceplate of a CCD camera are, however, converted to analog signals, which are subject to very sophisticated signal processing. At the other extreme, digitization of the composite signal occurs far down the transmission path. While the sampled composite signal is free of horizontal aliasing, it is oversampled and impaired by luminance-chrominance cross talk, bandwidth limitations, and other artifacts that cannot be corrected. Composite digital signals may be useful and inexpensive for some NTSC and PAL applications, but repeated A/D and D/A conversion cause significant degradation of the signals.

Composite NTSC and PAL Signals

NTSC or PAL composite signals are almost always sampled at a sample rate equal to four times the color subcarrier frequency ($4 \times f_c$.). Figure 21.1.15 shows that if the samples are taken at nulls of the in-phase and quadrature components of the color subcarrier, the sequence of sampled values in NTSC are $Y' + I$, $Y' + Q$, $Y' - I$, $Y' - Q$, . . . In PAL they are $Y' + V$, $Y' + U$, $Y' - V$, $Y' - U$, . . . on A-lines and $Y' - V$, $Y' + U$, $Y' + V$,

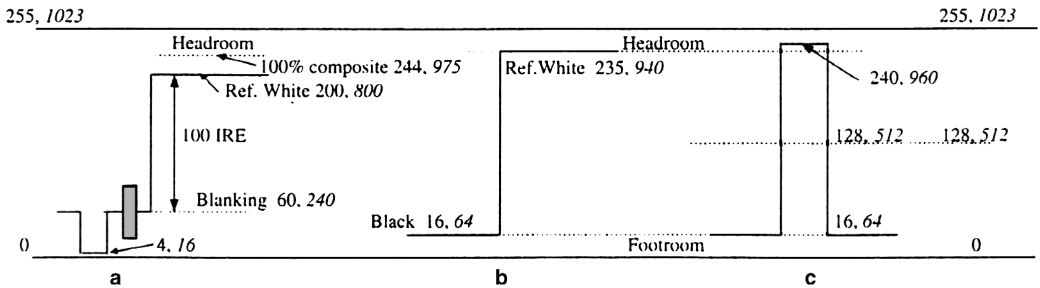


FIGURE 21.1.19 Digitized levels for composite NTSC and for components in CCIR601. Standard is for 8- and 10-bit quantization. 10-bit quantization levels are shown in italics.
a. Composite NTSC
b. Luminance component *Y′* in CCIR601
c. Chrominance components in CCIR601

Y′ − *U*, . . . on B-lines. There are 910 samples on a NTSC line and 1135 on a PAL line. Chrominance and luminance can be separated in digital circuits with inputs from adjacent lines in a field.

A typical NTSC scale for 256 levels (8 bits) and 1024 levels (10 bits) is shown in Fig. 21.1.19*a* (SMPTE259M). According to this scale some levels are reserved for headroom and sync. In the 8-bit version, only a little more than 7 bits are available for 100 IRE units (blanking to white). Quantization causes distortion of the signal which sometimes appears in the picture as *quantization contours or puddles*. One way to reduce the visibility of contours is to "dither" the least significant bit with random noise. In busy picture quantization may have noiselike appearance. These artifacts are visible in 8-bit NTSC, so 10 bits per sample is commonly used in professional applications, implying a data rate of 143 Mb/s requiring a transmission bandwidth of about 80 MHz.

The adoption in 1982 of an international standard, ITU-R Recommendation 601 (CCIR601), for component digital television, was a major milestone in the development of digital television. The D1 tape format and numerous other applications are based on the CCIR601 standard. Television studio equipment is now almost all digital, and digital interfaces between equipments are almost always based on CCIR601.

TABLE 21.1.9 ATSC Format Interoperability Considerations

	Format	Aspect Ratio		Frame Rate			
	1920 x 1080 (square pixels)	16:9			60 I	30 P	24 P
	1280 x 720 (square pixels)	16:9		60 P		30 P	24 P
TV	704 x 480 (CCIR 601)	16:9	4:3	60 P	60 I	30 P	24 P
	640 x 480 (square pixels)		4:3	60 P	60 I	30 P	24 P
	Computer	Film	TV	Computer	TV	Computer	Film

Supported frame rates include both 60.0 and 59.94 Hz related rates

COMPONENT DIGITAL VIDEO SIGNALS

Digitization of composite analog signals is still common, e.g., in the D2 and D3 tape recording formats, in various production systems, e.g., for time base correction, frame synchronization, editing, as well as for video processing in consumer products. Advantages of digital technology: perfect multigeneration reproducibility; precise time and level controls, digital storage, and signal processing; data compression; easy manipulation by computers, multimedia; incorporation in digital data transport packets for packet switching (ATM), broadband digital communication (BISDN); error control; and so forth. Recently, international agreement has been reached on the use of 1920 × 1080 as a common image format for high-definition production and program exchange. The 1920 × 1080 format has its roots in the CCIR sampling standard, and brings international compatibility to a new level. Table 21.1.9 shows the picture sensing characteristics of the 1920 × 1080 format for a variety of frame rates.

HIGH-DEFINITION TELEVISION (HDTV)

Fundamentals and Picture Quality

The fundamentals of HDTV are the same as the fundamentals of television in general, which have been reviewed in the previous sections of this chapter. From the viewers point of view, HDTV differs from "standard" definition TV (SDTV) by a significantly improved viewing experience in many respects:

1. Sharper, higher resolution images. By design, HDTV is intended to be as good as the resolution of the human visual system at a viewing distance of four picture heights (D = 4 ph). A large screen and a viewing distance of D < 4 stimulates viewer involvement in the scene. Compared to standard definition TV, HDTV increases horizontal luminance resolution by a factor of ≅ 3, vertical luminance resolution by a factor of ≅ 2, and horizontal chrominance resolution by a factor of ≅ 5.
2. A wider, cinemalike picture with an aspect ratio of 16/9 instead of 4/3.
3. High-quality surround sound audio that produces an immersive, theaterlike experience.
4. A substantial reduction in picture artifacts, including the cross-luminance and cross-chrominance artifacts of analog color systems. In addition, digital HDTV eliminates the analog transmission impairments such as noise, ghosts, distortion, and interference from other stations.

Also, a digital HDTV channel has all the potentials of digital TV in general and can offer many information services to the consumer. A packetized data format with headers and descriptors provide flexibility, interoperability, and extensibility.

HDTV, however, suffers from some imperfections:

1. Although HDTV supports multiple raster formats, each one has certain trade-offs among spatial and temporal resolution. Interlaced scan formats at 30, 30/1.001, or 25 Hz and progressive scan formats at the film rates of 25 or 24 Hz continue to have the same motion-related artifacts as today's analog TV systems and film. Also, in interlaced scan formats, line flicker may be visible at viewing distances D < 3.3 ph. While HDTV progressive scan formats at 60, 60/1.001, and 50 Hz improve motion rendition, they do so at the expense of horizontal and vertical spatial resolution.
2. Large area flicker may be perceived at viewing distances < 4 ph at typical TV brightness. At D = 4 ph the width of the picture is sustained by a 25° viewing angle, which means that a part of the picture is seen by flicker-sensitive rod vision. While a 60-Hz flicker may be acceptable, 50 Hz may not be acceptable and may require frame doubling as in motion pictures.
3. In digital HDTV over narrowband channels, there may be some visible video compression and quantization artifacts.
4. In DTTB of HDTV, there is the "cliff effect": Either a virtually perfect picture is received or no picture is received.

HDTV Production Standards

Efforts to establish national and international technical standards for production, distribution, and transmission of HDTV have been going on ever since Japan Broadcasting Corporation (NHK) started development of the analog 1125/60 HDTV system (MUSE) in the 1970s. In the mid-1980s the EBU proceeded with a project, referred to as "Eureka," to develop European 1250/50/2:1 and 1250/50/1:1 production standards. In the United States, the Society of Motion Pictures and Television Engineers (SMPTE) has proposed production standards for an 1125/60 high-definition production system, referred to as the SMPTE274M and SMPTE296M, which have been adopted by ATSC, while SMPTE240M is used in Japan.

Analog HDTV Broadcast Standards

The first operational system was the analog 1125/60 MUSE system ("multiple sub-Nyquist sampling encoding"), which has been adopted in Japan as a standard for HDTV satellite direct broadcasting (DBS). In the MUSE system, video bandwidth requirements are reduced from 21 to 8.1 MHz by a technique of horizontal and temporal subsampling, whereby diagonal and motion resolution is traded off to improve horizontal luminance resolution. Four fields are required to complete a picture with all details. Some motion compensation is used. Chrominance is also subsampled and the chrominance signals C_r and C_b are time compressed by factors of 4 and 5, respectively, and transmitted on alternate lines in the horizontal blanking interval which occupies about 20 percent of the time. This "time compressed integrated" (TCI) signal is converted into an analog signal that modulates the satellite FM transmitter.

In Europe an analog DBS system, HD-MAC, was developed based on a 1250/50/2:1 production standard. HD-MAC has been abandoned in favor of digital television systems. HD-DIVINE developed by a Scandinavian consortium for DTTB was demonstrated in 1992.

In the United States and Canada, the Advanced Television Systems Committee (ATSC) was formed in 1982 by the television industry to coordinate standards for advanced television systems. In 1987 a number of committees were formed to study and tests proposed systems for DTTB of HDTV. Originally over a dozen analog systems were proposed and many transmission methods were considered. But in June of 1990, a landslide move toward digital television terrestrial broadcasting began when General Instrument Corporation announced its DigiCipher all digital system. Others soon proposed all digital systems and in 1992 and 1993, four digital systems and Narrow MUSE, were tested at the ACATS' Advanced Television Test Center and by Cable Television Laboratories. Subjective tests were conducted in Canada by the Advanced Television Evaluation Laboratory (ATEL). The tests showed many advantages of the digital systems, which all performed better than Narrow-MUSE. In May 1993, the four proponents of digital systems formed a *Grand Alliance* which would develop a system combining the best features of each system. In 1994 and 1995, the Grand Alliance system was a subject of extensive tests (objective, subjective, and field tests), that validates its performance as a "best-of-the-best" digital system. The Grand Alliance system formed the basis for the ATSC standard (which added standard-definition formats) and the FCC transmission standard (which requires the use of the ATSC standard, with the exception that no picture format requirements whatsoever are specified by the FCC).

A better, but more expensive choice in terms of data rate, is A/D conversion of component video signals components. Usually the components are the pre-gamma corrected luminance signal Y' and the scaled chrominance signals $C_r = K_r(R' - Y')$ and $C_b = K_b(B' - Y')$ where $K_r = 0.5/(1 - T_R)$ and $K_b = 0.5/(1 - T_b)$. However, the choice can also be the pre-gamma corrected primaries R', G', and B' or the uncorrected primary signals R, G, and B. In the latter case the quantization must be at least 10 bits because of the high gain of the gamma-correcting circuit at low signal levels and the sensitivity of the HVS to contrast ratios. A quantum step at low luminance levels is more perceptible than at high luminance levels.

CCIR601 is a family of compatible international standards for component television. Figures 21.1.19*b*, 21.1.19*c*, and 21.1.20 summarize some key aspects of CCIR601. The common ground of 525/59.94 and 625/50 systems is that the duration of a TV line is about the same and that at a 13.5-MHz sample rate there is an integral number of sample points per line in all standard analog systems: $6 \times 143 = 858$ in 525/59.94 systems and $6 \times 144 = 864$ in 625/50 systems, as shown in Fig. 21.1.20. In both systems there are *720 active samplepoint locations* per line. Along a line the active samples are preceded and followed by sequences of digital timing

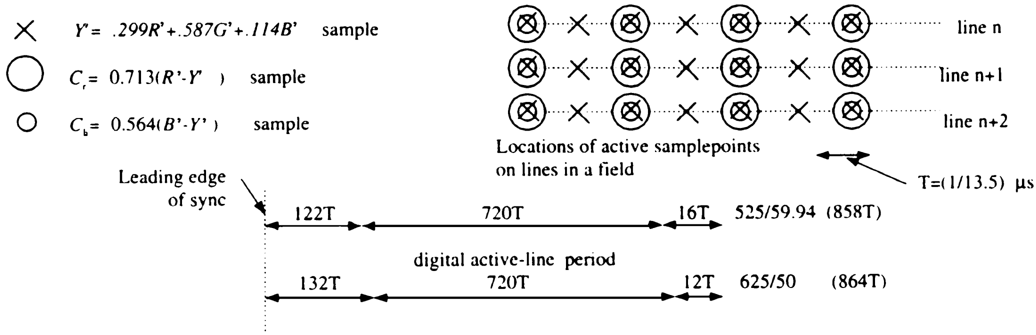


FIGURE 21.1.20 Sample point locations in 4:2:2 CCIR601 (ITU-R Rec. 601) for component SDTV digital television systems 525/59.94 and 625/50. Total samples per second = 27,000,000. In 4:2:0 the density of chrominance samples along a vertical line is reduced by a factor of 2 (see text). In 4:4:4 there are chrominance samples at every samplepoint location.

and ancillary data. This inactive period coincides with the horizontal blanking interval in all analog standard systems.

The CCIR601 4:2:2 format (family member) is summarized in Fig. 21.1.20. In a sequence of four Y' samples along a horizontal line there are two $C_r = K_R(R' - Y')$ samples and two $C_b = K_B(B' - Y')$ samples which are both co-located with every other Y' sample. The total number of active sampled values per line is thus 1440. The Nyquist bandwidths are 6.75 MHz for luminance and 3.875 MHz for chrominance. The CCIR601 standard specifies a 5.75 MHz low pass cutoff for luminance and a 2.75 MHz cutoff for chrominance. CCIR601 4:2:2 standard specifies either 8 or 10 bits per sampled value. The luminance and chrominance signal level as well as headroom and footroom are shown in Figs. 21.1.19*b* and 21.1.19*c*. The scale factors K_R and K_B (Table 21.1.6) are chosen to fit the chrominance signals C_r and C_b at 100 percent amplitude in the range 128 ± 112 (512 ± 448), which is about the same as the range of the luminance level.

In the CCIR601 4:4:4 format chrominance samples are co-located with every luminance sample. In the CCIR601 4:2:0 format the density of chrominance samples along a vertical line is reduced by a factor of 2 usually by taking some weighted average of 4:2:2: chrominance samples. In interlaced scan the so-derived 4:2:0 chroma samples appear to be uniformly interlaced in successive fields. The vertical chroma resolution in 4:2:0 is a factor of 2 less than in 4:2:2 or 4:4:4. The "raw" data rate of digitized video is high: 115 Mb/s for composite NTSC, up to 270 Mb/s for 4:2:2 CCIR601, and about 1 Gb/s for HDTV. By removing *statistical redundancy* and taking advantage of psychophysics to remove *irrelevant information*, the data rate of high quality HDTV has been reduced by a factor of 50 to about 20 Mb/s, which can be conveyed over a 6 MHz television broadcast channel.

With digitization of audio and video signals, television is becoming an integral part of a much broader field of digital information production and distribution. Digital standards strive to develop formats which provide high-quality data-compressed television and audio with considerations to *flexibility* of use and of receiver standards, *extensibility* to various formats, and *interoperability* with other applications, e.g., computers and data communications. In order to meet these objectives, a layered digital system architecture approach has been adopted as the basis for all DTV standards. This approach is loosely analogous to the ISO Open System Interconnect layered model for data communications. An international video compression standard, MPEG-2, provides the flexibility required to implement the three uppermost layers of the DTV system architecture. This standard was developed by the Motion Picture Expert Group of ITU and adopted by the International Standards Organization (ISO) and the International Electrotechnic Committee (IEC) in 1993. This accomplishment is the second major breakthrough toward international standards for digital television.

MPEG has a standard syntax (language) and protocol (rules) allowing for a great variety of inputs and for compatibility with many applications, i.e., not just one television standard or ancillary program. For example, there are "flags" specifying aspect ratio, samplepoint format, scanning standard, chromaticities, electro-optical transfer characteristics, sample rates, and so on. *Digital Television Standard* is referred to as the *ATSC*

TABLE 21.1.10 Picture Scanning Characteristics of the 1920 × 1080 High-Definition Common Image Format (ITU-R BT. 709-3)

Item	Parameter	System values									
		60/P	60/1	30/PsF	30/P	50/P	50/I	25/PsF	25/P	24/PsF	24/P
6.1	Order of sample presentation in a scanned system	Left to right, top to bottom For interlace and segmented frame systems, 1st active line of field 1 at top of picture									
6.2	Total number of lines	1125									
6.3	Field/frame/segment frequency (Hz)	60 (60/1.001)			30 (30/1.001)	50			25	48 (48/1.001)	24 (24/1.001)
6.4	Interlace ratio	1:1	2:1	1:1	1:1	1:1	2:1	1:1	1:1	1:1	1:1
6.5	Picture rate (Hz)	60 (60/1.001)	30 (30/1.001)			50	25			24 (24/1.001)	
6.6	Line frequency* (Hz)	67500 (67500/1.001)	33750 (33750/1.001)			56250	28125			27000 (27000/1.001)	
6.7	Samples per full line – R, G, B, Y $-C_B, C_R$	2200 1100				2640 1320				2750 1375	
6.8	Nominal analog signal bandwidths† (MHz)	60	30			60	30				
6.9	Sampling frequency $-R, G, B, Y$ (MHz)	148.5 (148.5/1.001)	74.25 (74.25/1.001)			148.5	74.25			74.25 (74.25/1.001)	
6.10	Sampling frequency‡ $-C_B, C_R$ (MHz)	74.25 (74.25/1.001)	37.125 (37.125/1.001)			74.25	37.125			37.125 (37.125/1.001)	

*The tolerance on frequencies is ± 0.001 percent.
†Bandwidth is for all components.
‡C_B, C_R sampling frequency is half of luminance sampling frequency.

TABLE 21.1.11 Some Key Transmission Parameters of the ATSC System for DTTB of HDTV

Transmission parameters	Terrestrial mode
Modulation method	VSB
Channel bandwidth MHz	6
Symbol rate megasymbols/second	10.76
Excess bandwidth	11.5%
Bits per symbol	3
Reed-Solomon FEC	T = 10 (208, 188)
Trellis coding	2/3 rate
Segment length symbols	836
Sync symbols per segment	4
Frame sync	1 per 313 segments
Payload data rate Mb/s	19.3
CNR threshold dB	14.9
Transport parameters	
Multiplex technique	MPEG-2 systems layer
Packet size bytes	188
Packet header bytes	4 including sync

system for HDTV. The ATSC standards accommodate one interlaced and one proscan HDTV format, shown in Table 21.1.10. It accommodates a number of frame rates: 60, 30, and 24 frames per second as well as these rates divided by 1.001. The total samples per line shown in Table 21.1.10 are for 60 frames per second proscan (format 1) and for 30 frames per second interlaced (format 2) and imply a 74.25 MHz (74.25/1.001 MHz) sample rate in accord with the SMPTE274M and SMPTE296M production standards. The ATSC standard also accommodates two SDTV formats with 480V/704-H and 480V/640-H active luminance samples per frame and a standard 13.5 MHz sample rate for 60/1.001 field per second, as in CCIR601. The ATSC has adopted the Dolby AC-3 digital audio system, while European standards for digital TV may lean toward MPEG2 for audio. The ATSC system uses VSB for transmission.

Some of the key transmission parameters for the ATSC system are shown in Table 21.1.11. MPEG-2 has been rapidly and enthusiastically adopted for many applications besides DTTB, e.g., satellite broadcasting, digital video disc (DVD). Digital satellite systems (DSS) are in operation worldwide. In the United States DIRECTV, for home reception with an 18 in. dish, provides 150 television channels.

MPEG-2 Format for Video Compression

Source coding or compression of audiovisual data has two objectives: removal of redundancy (*entropy coding*) and reduction of irrelevancy (*perceptual coding*). The first objective is to take advantage of statistical correlation in the bit stream to increase the information (entropy) per transmitted bit. There is no redundancy in a picture if all samples are statistically independent as they are in random noise, while there is a lot of redundancy in a patch of blue sky. Statistical coding requires buffering to absorb a variable information rate. The second objective is to eliminate information which cannot be perceived by human observers. While a significant amount of information which is irrelevant to human observers has been removed in standard analog television (number of lines and fields, relative bandwidth of components, and so forth.), further elimination of imperceptible information is possible with digital techniques. Removal of redundancy and reduction of irrelevancy, while having different objectives, are not independent of each other or of the constraints imposed by the transmission channel.

Data compression obviously increases the information carried by remaining bits, which must be conveyed with a very low error rate. If at most one error per minute can be tolerated in a 20 Mb/s HDTV signal, the error rate must be less than one in a billion. A channel coder encodes the source coded bit-stream to minimize transmission errors given the characteristics of the channel.

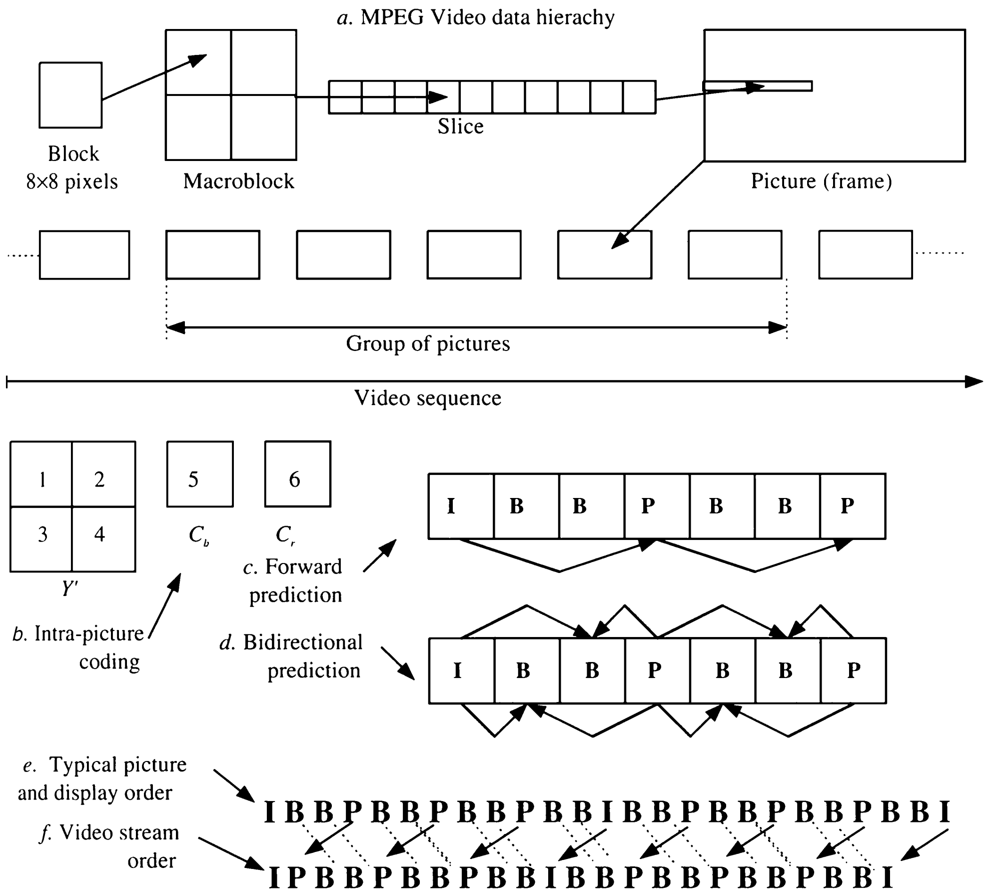


FIGURE 21.1.21 Basic principles of video compression coding. I frames are intraframes coded only with DCT (see Fig. 21.1.23). In forward predicted frames *P* and bidirectionally predicted frames B, the difference between actual data and motion predicted data is DCT coded. Data in DCT "frequency" space are adaptively quantized and zig-zag scanned with run-length coding. From S. N. Baron and W. R. Wilson (1994).

Figure 21.1.21 illustrates the basic elements of the MPEG-2 compression format. A video sequence consists of a series of pictures. A picture is usually a progressively scanned or interlaced scanned frame, but it may also be a single interlaced field. The pictures are organized into groups of pictures (GoPs), which are a basic unit for compression (Fig. 21.1.22*a*). Within a GoP, the first picture is coded entirely with spatial, or intraframe coding, and is referred to as an I-picture (I-frame). Subsequent pictures in the GoP are predicted using block-based motion estimation techniques. P-pictures are predicted from preceding I- or P-pictures using forward motion compensation. B-pictures are bidirectionally predicted from adjacent I- and/or P-pictures.

All pictures are composed of Y, U, and V pixels in a 4:2:0 sampling grid organized as 8 × 8 blocks, as shown in Fig. 21.1.22*b*. A macroblock consists of the four luminance blocks and two chrominance blocks covering the same spatial area of the picture, and macroblocks are organized into slices. A macroblock is the basic unit for motion compensation, and each macroblock in a P- or B-picture may have motion vector information indicating its prediction in 1/2 pixel components from the adjacent reference frames.

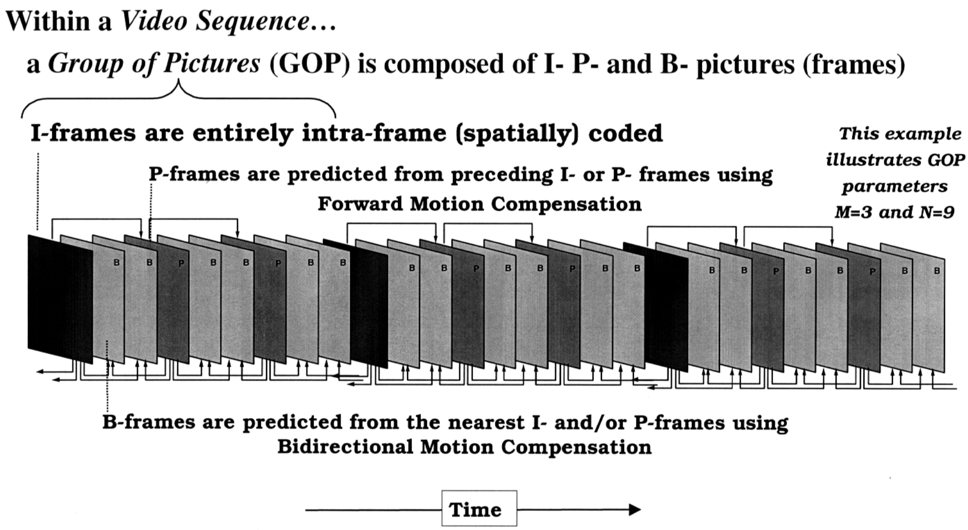


FIGURE 21.1.22 (*a*) MPEG Group of Pictures.

Each 8 × 8 block is encoded with the DCT. DCT coefficients are quantized and scanned in zigzag order.

Two methods of data compression are used in MPEG-2: (1) intra-picture transform coding using discrete cosine transform (DCT), adaptive quantization, and variable length coding (VLC) and (2) motion compensation for prediction of P and B pictures. Motion vectors together with the intra-frame transform coded difference between predicted and actual picture is transmitted.

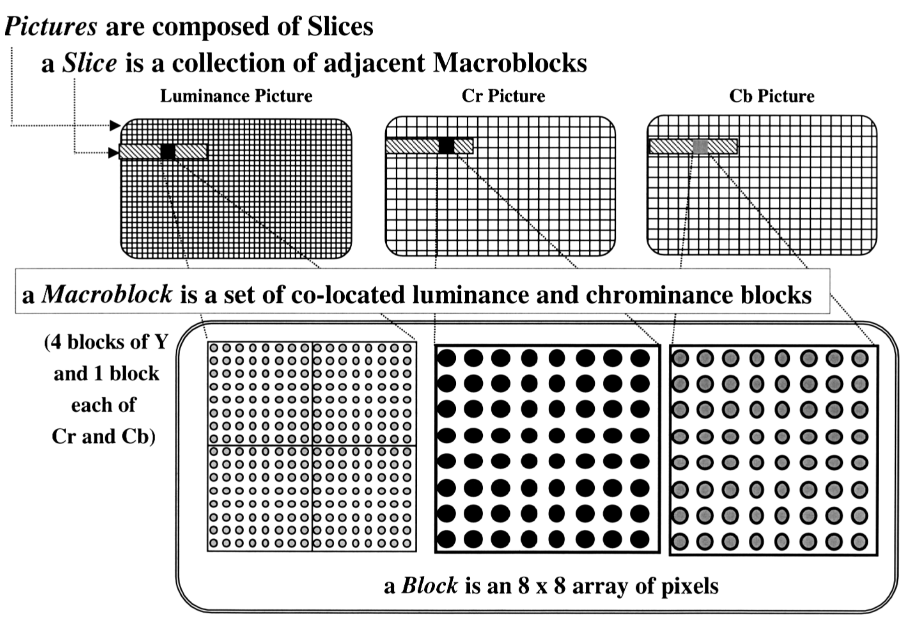
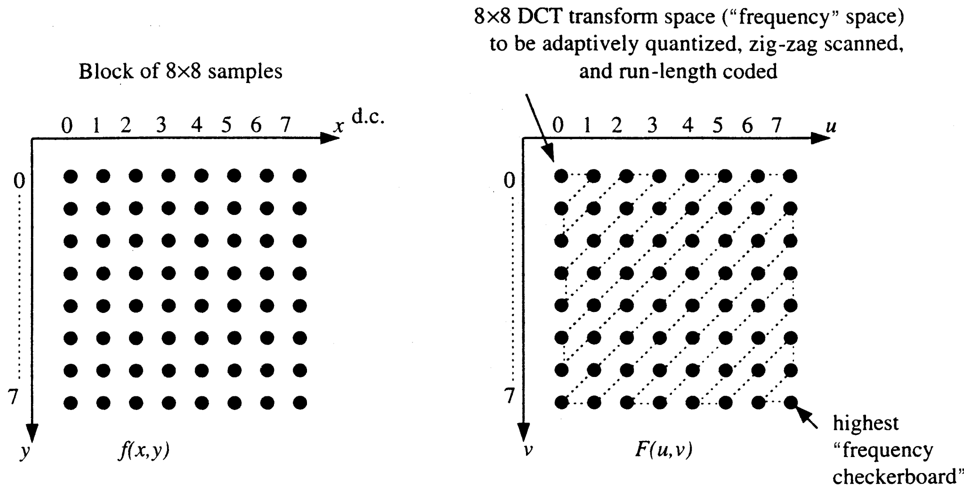


FIGURE 21.1.22 (*b*) Slices, Macroblocks, and Blocks.



The N × N two-dimensional discrete Fourier transform (DCT) is defined as:

$$F(u,v) = \frac{2}{N} C(u)C(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cos\frac{(2x+1)u\pi}{2N} \cos\frac{(2y+1)v\pi}{2N}$$

where $x, y = 0, 1, 2, \ldots, N-1$ are discrete spatial coordinates in the pixel domain

and $u, v = 0, 1, 2, \ldots, N-1$ are discrete coordinates in the transform domain

and

$$C(u), C(v) = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } u, v = 0 \\ 1 & \text{otherwise} \end{cases}$$

The inverse discrete Fourier transform (IDCT) is defined as:

$$f(x,y) = \frac{2}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u)C(v)F(u,v) \cos\frac{(2x+1)u\pi}{2N} \cos\frac{(2y+1)v\pi}{2N}$$

FIGURE 21.1.23 The discrete cosine transform (DCT). Many transform components have quantization level = zero. If *x* and *y* were continuous (which they are not) each transform component would be a product of a vertical and a horizontal cosine grating. From ISO/IEC 1.3818-2 Rec. H262.

The DCT transform is applied to an 8 × 8 block of samples *f*(*x*,*y*) as shown in Fig. 21.1.23. In the transform domain the information is represented by an 8 × 8 array of real "frequency" components *F*(*u*,*v*). A component *F*(*u*,*v*) can be considered as the amplitude of an interference pattern consisting of the product of a horizontal and a vertical cosine wave (grating) forming a "checker-board" with soft-edged rectangular "squares." The pattern with the smallest "squares" is represented by the component in the lower right corner in the frequency domain shown in Fig. 21.1.23, and the biggest "square" is represented by the dc component in the upper-left corner. The pels can be considered as samples of a superposition of these interference patterns. It must be kept in mind that the pixels *f*(*x*,*y*) are *not* meant to represent a continuous function periodically mirrored in the *x* and *y* axis and band limited to the Nyquist bandwidth of the samplepoint lattice, and that a "frequency" component *F*(*u*,*v*) does *not* specify a sine-wave grating. The DCT is strictly a transformation of an $N \times N$ array of numbers into another $N \times N$ array of numbers using Fourier techniques.

The DCT is remarkably effective in removing redundancy in typical pictures. When the pels $f(x,y)$ in a block are highly correlated, a large number of coefficients $F(u,\upsilon)$ in the transform domain are zero or very small, particularly the high-frequency components. The components in the transform domain are quantized. While the HVS is sensitive to quantization of low-frequency components, it tolerates large steps for high frequencies. An 8×8 *quant matrix* specifies the quantization steps to be associated with the array of $F(u,v)$ components. With quantum steps increasing with frequency, a large number of high-frequency components are quantized to level zero. MPEG allows for adapting the quant matrix for each block for an optimum trade-off between quality and available buffer memory. The quant matrix is specified in a header. The components are scanned in a zigzag pattern as shown in Fig. 21.1.23, generating a stream of symbols with bursts of zeros. The scanned data are subsequently variable length coded in run-amplitude pairs, specifying a nonzero component and an associated run length of zero coefficients. As in the Morse code, shorter codes are used for common pairs and longer for less common pairs. These symbols are delivered for transmission. DCT is the only technique used for I-picture data compression.

Motion compensation is used for P- and B-picture data compression. Before intra-frame coding (DCT), samples in blocks of P or B pictures are subtracted from predicted samples. The intra-frame encoded data of these "differential" pictures is transmitted together with data, *motion vectors*. Motion vectors are determined by matching a macroblock in the actual picture with macroblocks in a nearby area in the reference pictures (the reference picture for a P picture is the previous P or I picture and the reference pictures for B pictures are nearby I and P pictures). MPEG has standardized the vector range and how to specify macroblock motion vectors, but not how they are determined. After motion compensation, the pels of the P and B frames are highly decorrelated. This makes the subsequent DCT and quantization processes extremely efficient. In the atypical case where a predicted macroblock has higher entropy than the original image macroblock, intra-coding may still be used in P + B pictures. Given the capacity of the channel and the buffer memory, increased range and more accurate block prediction can reduce the call for quant matrices with course quantization. The consequence is compatible improvement of picture quality. Buffering the data for adaptive quantization is a sophisticated process. The outputs of video and audio data compression encoders are called program elementary streams (ES). MPEG classifies video ESs compression types, called *profiles*, intended for various applications. The profiles are called *simple, main, snr, spatial, and high*. MPEG also identifies various *levels* of video ESs characterized by the maximum number of active luminance samples per second. They are

Low:	$352\text{H} \times 288\text{V} \times 30\text{Hz}$	Main: $720\text{H} \times 576\text{V} \times 30\text{Hz}$
High-1440:	$1440\text{H} \times 1152\text{V} \times 60\text{Hz}$	High: $1920\text{H} \times 1152\text{V} \times 60\text{Hz}$

The two most important are main profile at main level (MP@ML) used in standard definition DTV system such as DVB and main profile at high level (MP@HL) and in high definition DN systems such as the ATSC standard.

Digital Audio for DTV

Just as analog video signals are digitized by sampling and quantization, analog audio signals are similarly converted into digital audio. A *de facto* international standard (AES/EBU) for high-quality digital audio is 48,000 samples per second and 16 bits per sample. Other standard sample rates are 44.1 kHz (compact disc) and 32 kHz. And just as digital video pels are compressed to discard redundant and imperceptible visual information, audio can also be compressed. By taking advantage of the psychophysics of the Human Aural System the data rate per audio channel can be reduced for transmission from $48 \times 16 = 768$ kb/s to substantially less than 128 kb/s with virtually imperceptible degradation. While there is an MPEG-2 audio standard, the Dolby AC-3 system, used in cinema theaters, has been adopted for the ATSC system for HDTV in North America. Both systems use *psychoacoustic coders* and deliver high-quality audio. The ATSC "Dolby Digital" system delivers five 20 kHz audio channels and a low-frequency (3 to 120 Hz) enhancement channel (LFE). This package of "5.1 channels" requires a maximum digital channel capacity of 384 kb/s.

Strings of digitized samples are converted from the time domain to the frequency domain, where spectra in 24 "critical bands" are processed. The basis for data compression is that in the vicinity of strong components, relatively weak components can be ignored without perceptible degradation (masking). The frequency components are coded and packetized. There are six levels of services including surround sound (complete main = CM, music and effects = CE), sound for visually impaired (VI) or hearing impaired (HI), as well as dialogue (D), emergency (E), and voice over (VO). Principles of *psychoacoustic coders* are discussed in the chapter on audio.

The ES data are put as payload in *packets*, like merchandise in freight cars to be assembled to form a train and referred to as *Packetized Elementary Streams* (*PES*). A packet has a header including a label (packet ID, *PID*) and timing information. The packet also includes an *adaptation field*, including program clock, reference addresses, and descriptors for how the payload is to be used. In the MPEG-2 transport system a packet contains 188 bytes. Just as MPEG-2 compression is highly flexible, MPEG-s transport contains flags for local program insertion points, encryption for controlled access (CA), multimedia, interactive data, and many other descriptors.

It is one of the most important attributes of digital television that the information can be packetized and transmitted in a data communications format which is interoperable with other applications, and can be distributed not only by broadcasting (terrestrial, satellite, cable), but over various data communication networks, for example, as *Asynchronous Transfer Mode* (ATM) using communication over BISDN. BISDN includes fiber optics transmission. ATM uses fixed length packets (53 bytes) with a header containing descriptors as well as addresses for switching. The convergence of broadcasting, multimedia, telecommunications, data communications, and computer applications presents complex compatibility and standardization problems which at present are addressed by a number of international standardization committees. Digital television is now part of the broader industry of information production and distribution. MPEG is part of this activity, and the various adaptation layers have been designed to provide interoperability, so that MPEG-2 transport streams may be easily layered on top of other digital communication systems.

Digital Television Terrestrial Broadcasting (DTTB)

Considerations driving the selection of digital modulation approaches for DTTB include striking an appropriate balance among *quality*, *coverage*, *and compatibility* as determined by public policy priorities of various countries:

- Reception of high-quality HDTV and audio
- Good coverage of population and area with acceptable reliability
- Acceptable levels of DTTB interference into reception of existing services, particularly reception of standard analog television signals (NTSC, PAL, SECAM) (Fig. 21.1.18)

In the United States the FCC has only made available the existing TV broadcast bands for DTTB of HDTV, allocating an additional 6-MHz TV channel for DTV to every existing broadcast station to be radiated from an antenna which is close to the NTSC antenna of the associated station. The consequence is that current rules for station allocations, including so-called UHF taboos, must be changed to accommodate about 1700 new DTTB channels in the TV broadcast bands, primarily at UHF, with tolerable interference from NTSC transmitters and acceptable DTTB interference into existing NTSC reception. The FCC's channel allocation plan is based on extensive measurements, field tests, and computer modelling. Thus DTV coverage is primarily limited by interference into existing NTSC service, so achieving the required data rate (approximately 20 Mb/s) at the lowest possible power is paramount.

CNR in DTTB is defined as the ratio in dB of *average power* to total noise power in the channel. With the USB modulation used in the ATSC standard, the threshold of visibility of errors caused by random noise occurs at CNR $\cong$ 15 dB. At about 1 dB below this threshold the picture is not useable. Random noise is only one of a number of possible transmission impairments (see previous section). In contrast with broadcasting of analog video, impairments in proposed DTTB systems add up toward a threshold at which reception abruptly changes from perfect to useless. This abrupt coverage limit in DTTB is referred to as the "cliff effect," and is similar to the effect in FM radios.

The radiated DTTB signal is noiselike and has a peak-to-average power ratio of 8 to 10 dB. Co-channel DTTB interference into NTSC appears essentially as random noise. Comparing DTTB coverage with NTSC grade-B coverage on the same channel and with the same receiving antenna, transmission line and receiver noise figure as in NTSC reception, the required effective radiated *peak power* in DTTB of HDTV could be 3 to 5 dB less than NTSC *peak power* radiating from the same antenna. While HDTV quality of DTTB is far better than NTSC grade-B quality, statistics of field strength variability with time and location is critical considering the "cliff effect" in DTTB. DTTB reception reliability can, however, be improved by as much as 10 dB with a better receiving antenna and an antenna amplifier.

Modulation Methods for Digital TV

Four modulation methods (see Eq. 31) widely used for r.f. transmission of digital television:

- PSK: Phase shift keying may be preferred in digital TV satellite broadcasting
- QAM: Quadrature amplitude modulation in digital cable television system
- VSB: Vestigial sideband is used in the ATSC standard for terrestrial DTV broadcasting
- COFDM: Coded orthogonal frequency division multiplex is used in the DVB-T standard for terrestrial DTV broadcasting

In QAM, a carrier in the middle of the channel is amplitude modulated double sideband in-phase by one sequence of digital symbols and in quadrature by another sequence of symbols. In a 2.5-MHz Nyquist bandwidth 5 megasymbols per second can be transmitted in the I channel as well as in the Q channel. Thus, a 5-MHz channel can convey 10 megasymbols per second. To reduce extended ringing and intersymbol interference, the spectrum is rolled off antisymmetrically around the Nyquist band edges. This calls for some extra bandwidth. Figure 21.1.24 illustrates typical roll-off for transmission of about 10 megasymbols per second over a 6-MHz television channel.

The quantized symbols in the I and Q channels are coincident, and consequently the amplitude and phase of the carrier at symbol detection time can be represented by a point in a constellation of points in the *xy-plane* representing all possible transmitted vectors as shown in Fig. 21.1.24. If the number of constellation points is K, the signal is referred to as K-QAM (K = 64 for 3 bits/symbol and 16 for 2 bits/symbol). Figure 21.1.24 shows that each constellation point is associated with a "decision" cell. An error occurs if transmission impairments at decision time moves a signal vector into the wrong cell. The distance from a constellation point to the border of its cell represents the allowed "budget" for transmission impairments. Figure 21.1.24 shows an approximate relation between the error rate and the rms level σ of Gaussian noise (expressed in terms of the distance from a constellation point to the border of its box). For a typical error rate of 10^{-3}, $\sigma =$ 0.35 or −11 dB. The figure also shows the carrier level for $(2N)^2$-QAM. For 16QAM the required CNR for an error rate of 10^{-3} is seen to be about 18 dB. Channel coding (see below) can improve this relation to an error rate of 10^{-9} out of the channel decoder for a CNR of 15 dB at the input of the demodulator. However, other impairments leave only a part of the error budget to random noise. Phase jitters are particularly damaging to the peripheral vectors. Carrier phase and symbol phase recovery in the demodulator are also critical and complex functions.

VSB is quite similar to *QAM*. In fact if *I* and *Q* symbols in *QAM* are interleaved in time rather than coincident, they can be detected with a synchronous carrier at the Nyquist band-edges as well as in the center of the band. In VSB detection, the symbols arrive in a sequence I,Q, –I, –Q, I, Q, . . . , i.e., at twice the QAM rate but with half the number of bits per symbol. Thus, for a given data rate, the required CNR for a given error rate is about the same as in QAM. As in standard analog television VSB transmits a quadrature signal that eliminates a sideband but doubles the average power. Peaks in the transmitted signal are caused by the quadrature signal. The peak-to-rms ratio is, however, about the same as in QAM (8 to 10 dB). The antisymmetric roll-off around the carrier reduces peaks, ringing, and intersymbol interference. A residual carrier pilot tone is also used. ATSC standard has to improve robustness and reduce lock-in time.

While VSBAM is an extreme time division multiplex system, OFDM is an extreme frequency division multiplex system. OFDM is an entirely different approach to over-the-air transmission of digital data. OFDM uses hundreds of narrowband channels equally spaced along the frequency axis. Each carrier is modulated (QAM)

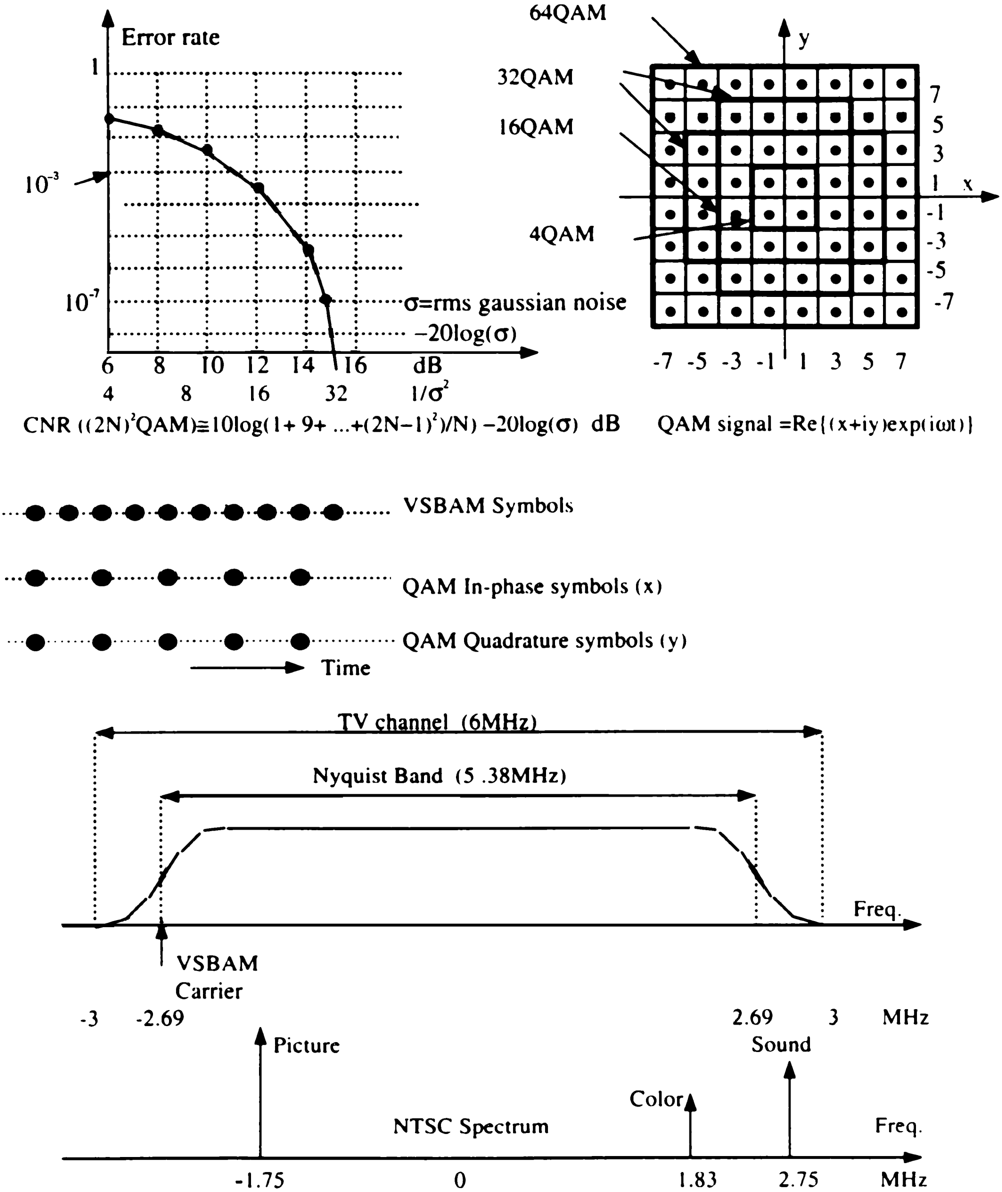


FIGURE 21.1.24 QAM signal vector constellation points and CNR. Spectra VSBAM, DTTB broadcasting in the United States in the presence of co-channel NTSC radiation.

by digital symbols in such a way that the spectrum of the channel has a sinx/x response with nulls at the carrier frequencies of the other channels, as illustrated in Fig. 21.1.25. As a consequence these channels are "orthogonal" along the frequency axis and do not interfere with each other. Figure 21.1.25 shows that each QAM symbol occupies a long period of time T. In the frequency domain the symbols are spaced F = 1/T Hz. There is a period D between symbols in the time domain to guard against intersymbol interference due to ghosts.

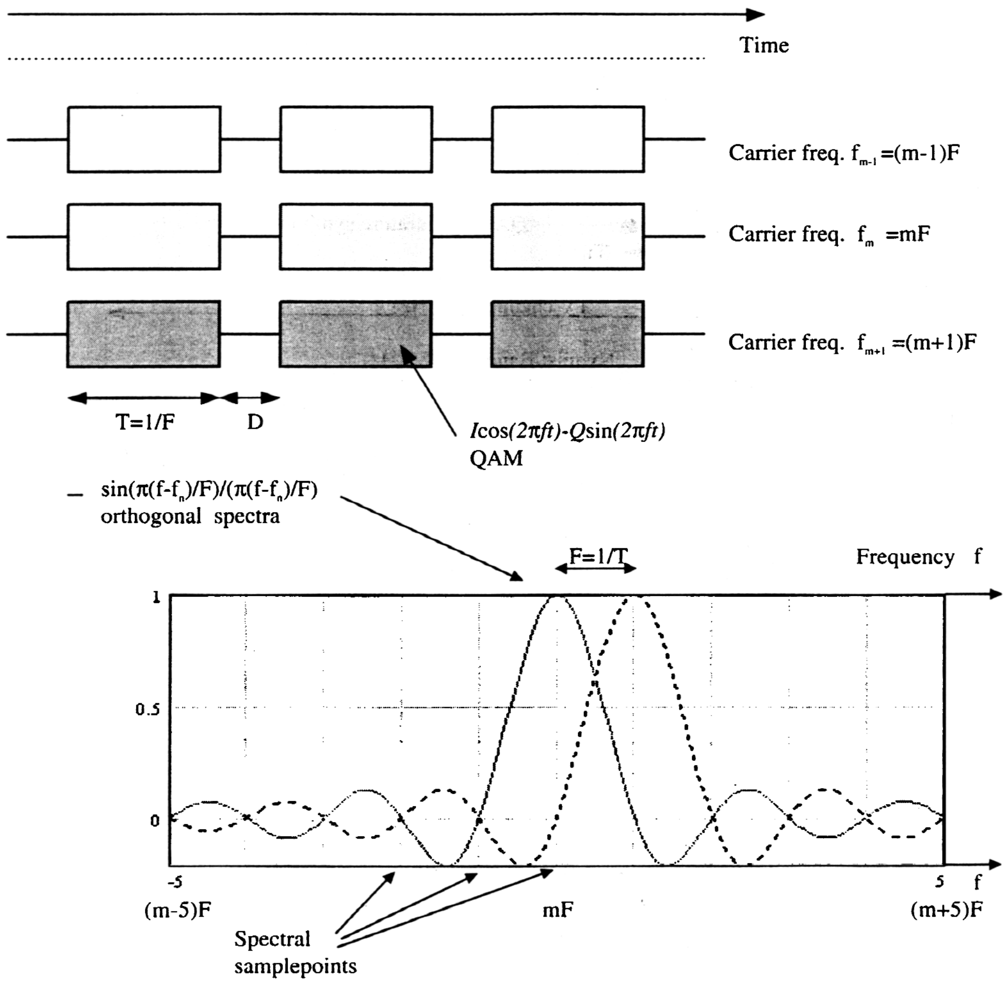


FIGURE 21.1.25 OFDM (orthogonal frequency division multiplex) time-frequency representations. Temporal guardband D is chosen to minimize intersymbol interference due to multipath.

Each symbol is not distorted by a ghost, but the level of the symbol is affected by ghosts. With an equalizer, the CNR of different channels may be different. The terrestrial DVB standard (D1 3-T) includes two OFDM formats: 2k with 1705 carriers and 8k with 6817 carriers. In the 8k format the guard interval D is scalable.

With coded OFDM (*COFDM*) lost information in different channels can be managed. While OFDM is tolerant of impulse noise, a disturbance occupying a part of the spectrum may wipe out a number of channels. For broadcasting in the presence of strong interference from standard analog television transmitters, it has been proposed that OFDM channels coinciding with bands around the picture, sound, and color subcarrier of an analog TV co-channel signal be left vacant (Fig. 21.1.24). Other methods using error correction designed for this type of interference are also possible. COFDM can be used for same channel cellular digital broadcasting of identical information. Co-channel interference is then equivalent to a ghost. The time guard D has a significant impact on the geographic spacing of the cell. In detection of OFDM, the symbols transmitted during a period T are derived by a spectral analysis of a fast Fourier transform of the temporal signal received during the period.

All transmission systems used for DTTB take the packetized bitstream through a channel coder before it is delivered to a modulator. Since the source coders have eliminated a large amount of redundancy in the video and audio programs, the bitstream feeding *the channel coder* is extremely intolerant to errors. The channel coder adds redundancy in the form of *forward error correction* (FEC) and possibly by *trellis coding* as illustrated in Fig. 21.1.2. The purpose of FEC is to correct transmission errors at the receiver. The Reed-Solomon block FEC can very effectively correct random errors. In the Grand Alliance system 20 parity bytes are added to the 188 byte packet to correct for up to 10 errors (t = 10 R-S code). The price is a 10 percent increase in required channel capacity (code rate R = 188/208). However, since errors induced by RF interference are generally burst errors, the data interleaving technique is used to regroup data, so that a burst error in transmission is dispersed to appear as random errors spread across many Reed-Solomon FEC blocks, each of which can correct the random error. The purpose of trellis-coding is to prevent errors from occurring, particularly errors caused by random noise, or conversely, to reduce the required CNR for a given error rate. In trellis-coded transmission, fewer symbols are used than the number of symbols that the modulator can generate, i.e., more bits per symbols are available than are transmitted. In the decoder, hard decisions are not made on a symbol by symbol basis, but detected levels of a number of symbols in a sequence form a vector (soft data) which is compared with vectors which possibly could have been transmitted (allowed vectors). Using "soft" data from several symbols and cleverly selected allowed vectors, the required CNR for a given error rate can be reduced.

BIBLIOGRAPHY

"CCIR and FCC TV Standards," a condensed pamphlet of worldwide SDTV standards, Rohde & Schwarz, P.O. 801469, D-8000 München 80, Germany.

Hutson, G. H., and P. J. Shepderd, "Colour Television," 2nd ed., McGraw-Hill, 1990. Emphasis on European standards.

Jordan, E. C. (ed.) "Reference Data for Engineers: Radio, Electronics, Computer, and Communications," 8th ed., Howard W. Sams (McMillan), 1993. Includes numerous references to TV standards and FCC F(50/50) charts. *Journal of the SMPTE* publishes many excellent papers on Television Fundamentals and Standards. The papers include numerous references. Since 1994 the journal regularly publishes tutorials. There are a number of tutorials on digital television by S. Baron and from 3/95 to 3/96 by D. Strachan.

Pearson, D. E. "Transmission and Display of Pictorial Information," Pentech Press, 1975.

Poynton, C. "A Technical Introduction to Digital Video," John Wiley & Sons, 1996.

Pritchard, D. H., and J. J. Gibson, "Worldwide Color Television Standards—Similarities and Differences, *SMPTE J*., 89: 111–120, February 1980. Also in Rzeszewski, *Color Television.*

Rzeszewski, T. "Color Television," IEEE Press, 1993.

Symes, P. D., "Video Compression," McGraw-Hill, 1998.

TV Technology publishes many timely papers, including informative tutorials on advanced TV by M. Weiss, C. Rhodes, and others.

Whitaker, J. C. "DTV: The Revolution in Digital Video," 2nd ed., McGraw-Hill, 1999.

Whitaker, J. C., and K. B. Benson (eds.) "Standard Handbook of Video and Television Engineering," 3rd ed. (formerly the Television Engineering Handbook), McGraw-Hill, 2000.

Standards

DVB Standards are Published by ETSI, at http://www.etsi.org in the "Publications Download Area."

FCC Rules, Part 73, Broadcast Services, Federal Communications Commission, Washington, D.C.

ITU-R Volume XI. Part 1. *Broadcasting Service-Television*. Report 624-3: *Characteristics of Television Systems*. Report 476-2*: Colorimetric Standards in Colour Television*. Recommendation ITU-R 601-1: *Encoding Parameters of Digital Television for Studios*, also in *SMPTE J.* 90–10, October 1981, *SMPTE* 103–12, December 1994 (SMPTE125M) and *EBU Review* #187 June 1981.

MPEG-2, Recommendation ITU-T H.222.0, (ISO/IEC 13818-1) Generic Coding of Moving Pictures and Associated Audio: SYSTEMS.

MPEG-2, Recommendation ITU-T H.262, (ISO/IEC 13818-2) Generic Coding of Moving Pictures and Associated Audio: VIDEO.

MPEG-2. *ISO/IEC*. Generic Coding of Moving Pictires and Associated Audio" *International Standard 13818-2, Recommendation H.262, (1995E)*. ISO/IEC Copywrite Office, Case Postale 56, CH1211 Geneva 20, Switzerland. ATSC Digital Television Standard and Guide to the Use of ATSC Digital Television Standard, documents A/53 and A/54 published by ATSC in 1995. ATSC, 1700 K Street NW, Suite 800, Washington, D.C.

Recommendation ITU-R BT.470-6, Conventional Television Systems.

Recommendation ITU-R BT.601-5, Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-Screen 16:9 Aspect Ratios.

Recommendation ITU-R BT. 656-4, Interfaces for Digital Component Video Signals in 525-line and 625-line Television Systems Operating at the 4:2:2 Level of Recommendation ITU-R BT. 601 (Part A).

Recommendation ITU-R BT.709-3, Parameter Values for the HDTV Standards for Production and International Programme Exchange.

Recommendation ITU-R BT.1381, SDI-Based Transport Interface for Compressed Television Signals in Networked Television Production Based on Recommendations ITU-R BT.656 and ITU-R BT.1302.

SMPTE 170M, Composite Analog Video Signal—NTSC for Studio Applications.

SMPTE 310M, Synchronous Serial Interface for MPEG-2 Digital Transport Streams.

SMPTE 314M, Data Structure for DV Based Audio, Data and Compressed Video—25 Mb/s and 50 Mb/s.

SMPTE proposed Standards are published in the *Journal of the SMPTE*. SMPTE 240M-1988. *Standard for-1125/60 high-definition production*. September 1989. SMPTE274M *Standard* for *1020 × 1080 Scanning*, October 1994. SMPTE296M *Standard for 1280 × 720 Scanning*, May 1996.

Audio

Gibson, J. J., *J. Audio Engineering Society*, September 1986. p. 647–660. Reviews the FCC-BTSC Standard (FCC Bulletin OET60) for and the performance of Television Multichannel Sound (MTS). Includes FCC specifications in FCC Bulletin OET-60.

Digital Television and HDTV

Advanced Digital HDTV System Description, Sarnoff, Thomson, Philips and NBC submission to the Advisory Committee on Advanced Television Service, 1992.

Anaslassiou, D. (ed.) "Special Issue on HDTV," *IEEE Communications Magazine*, Vol. 34, No. 6, June 1996.

ATVA-Progressive System Description, MIT and General Instruments submission to the Advisory Committee on Advanced Television Service, 1992.

Baron, S. N. "An overview of the DTTB model," *J. SMPTE*, Vols. 103–105, May 1994.

Baron, S. N. "Digital television scalability and interoperability, *J. SMPTE*, Vol. 103, No. 10, October 1994, p. 673.

Baron, S. N. and W. R. Wilson, "MPEG Overview," *J. SMPTE*, Vol. 103 No. 6. June 1994.

Beyers, B. W. "Digital Television: Opportunities for Change," *IEEE Transactions on Consumer Electronics*, Vol. 38, No. 1, February 1992.

Brandenburg K. and G. Stoll. "ISO-MPEG-1 Audio: Ageneric Standard for Coding of High Quality Digital Audio," *J. Audio Engineering Society*, Vol. 42, No. 10, October 1994.

Bretl, W., et al. "VSB Modern Subsystem Design for Grand Alliance Digital Television Receiver," *IEEE Transactions on Consumer Electronics*, Vol. 41, No. 3, August 1995.

Bryan, D. "QAM for Terrestrial and Cable Television," *IEEE Transactions on Consumer Electronics*, Vol. 41, No. 3, August 1995.

Chopra A., et al. "The HDTV Broadcast Standard: Subjective Evaluation of Proponent Systems," *SMPTE J.*, July 1998.

Cugnini A., et al. "MPEG-2 Decoder for the Digital HDTV Grand Alliance System," *IEEE Transactions on Consumer Electronics*, Vol. 41, No. 3, August 1995.

DigiCipher System Description, General Instruments submission to the Advisory Committee on Advanced Television Service, 1992.

Digital HDTV Grand Alliance System Description, Grand Alliance (AT&T, General Instruments, Sarnoff, Thomson, Philips, MIT and Zenith) submission to the Advisory Committee on Advanced Television Service, 1994.

"Digital Television in the U.S.," *IEEE Spectrum*, Vol. 4, No. 5, March 1999.

Digital Spectrum Compatible HDTV System Description, Zenith and AT&T submission to the Advisory Committee on Advanced Television Service, 1992.

Hopkins, R. "Digital Terrestrial HDTV for North America: The Grand Alliance HDTV System," *IEEE Transactions on Consumer Electronics*, Vol. 40, No. 3, August 1994.

Joseph, K., et al. "Prioritization and Transport in the ADTV Digital Simulcast System," *IEEE Transactions on Consumer Electronics*, Vol. 38, No. 3, August 1992.

Kraus, J. A. "Source Coding, Channel Coding and Modulation Techniques Used in the DigiCipher System," *IEEE Transactions on Broadcasting, Special Issue on HDTV Broadcasting*, Vol. 37, No. 4, December 1991.

Kraus, J. A. "Source Coding, Channel Coding and Modulation Techniques Used in the ATVA-Progressive System," *IEEE Transactions on Broadcasting, Special Issue on HDTV Broadcasting*, Vol. 37, No. 4, December 1991.

Krivocheev, M. edited by S. N. Baron," The First Twenty Years of HDTV: 1972–1992," *J. SMPTE*, Vol. 102, No. 10, October 1993. (Includes many references to international standards committees).

Luplow, W. "Digital Spectrum Compatible HDTV," *HD World Review* (ISSN 1055-6931) Vol. 2, No. 4, Fall 1991.

Luplow, W., et al. "Source Coding, Channel Coding and Modulation Techniques Used in the Digital Spectrum-Compatible HDTV System," *IEEE Transactions on Broadcasting, Special Issue on HDTV Broadcasting*, Vol. 37, No. 4, December 1991.

Lyons, P. "Grand Alliance Prototype Transport Stream Encoder Design and Implementation," *IEEE Transactions on Consumer Electronics*, Vol. 41, No. 3, August 1995.

Mailhot J., et al. "The Grand Alliance HDTV Video Encoder," *IEEE Transactions on Consumer Electronics*, Vol. 41, No. 3, August 1995.

Miller, H. N. "A Market Perspective on the HDTV Interoperability Standards Question," *IEEE Transactions on Consumer Electronics*, Vol. 38, No. 1, February 1992.

Netravali, A. B. Haskell, "Digital Pictures—Representation and Compression," Plenum Press, 1988.

Reitmeier, G. "The U.S. Digital television Standard and Its Impact on VLSI," *J. VLSI SIGNAL PROCESSING SYSTEMS for Signal, Image and Video Technology*, Vol. 17, No. 2/3, pp. 281–290, November 1997.

Reitmeier, G., and C. Basile "Advanced Digital Television," *HD World Review* (ISSN 1055-6931) Vol. 2, No. 4, Fall 1991.

Reitmeier, G., et al. "Source Coding, Channel Coding and Modulation Techniques Used in the ADTV System," *IEEE Transactions on Broadcasting, Special Issue on HDTV Broadcasting*, Vol. 37, No. 4, December 1991.

Sablatash, M. "Transmission of all Digital Advanced Television: State of the Art and Future Reflections," *IEEE Transactions on Broadcasting*, Vol. 40, No 2, June 1994.

Sgrignoli, G. "Field Test Results of the ATSC VSB Transmission System," *J. SMPTE,* Vol. 106, p. 601, September 1997.

Sgrignoli G., et al. "VSB Modulation Used for Terrestrial and Cable Broadcasts," *IEEE Transactions on Consumer Electronics*, Vol. 41, No. 3, August 1995.

"Special issue on consumer electronics," *Proc. IEEE*, Vol. 82–84, April 1994.

"Special issue on digital television," *IEEE Communications Magazine*, 35–5, May 1994.

"Special Issue on Digital Television," *IEEE Spectrum*, Vol. 32, No. 4, April 1995.

"Special Issues on Digital Television," *Proc. IEEE*, Vol. 83, No. 6, June 1995; Vol. 83, No. 7, July 1995.

Wu, Y., et al. "Orthogonal Frequency Division Multiplex: A Multi-Carrier Modulation Scheme," *IEEE Transactions on Consumer Electronics*, Vol. 41, No. 3, August 1995.