A Preview of the Television Video and Audio – a Ready Reference for Engineers Course

About the Author

Randy Hoffner is retired from ABC Broadcast Operations and Engineering West Coast, where he held the position of Manager of Technology and Quality Control. Prior to 2006, Hoffner spent a number of years at ABC BOE New York, where he worked on advanced television projects including HDTV and early film-to-720p video transfers, accomplished using data telecine transfer techniques.

Before joining ABC, Hoffner spent 15 years at NBC, New York, where he worked on technology development projects, including early HDTV work, and held the positions of Director of Research and Development and Director of Future Technology. He also played a large role in the transition of the NBC Television Network and local NBC television stations to stereo broadcasting. He twice received the NBC Technical Excellence Award.

Hoffner is a Fellow of SMPTE, and has served the Audio Engineering Society as Convention Chairman and Vice President, Eastern Region, US and Canada. He has authored a number of technical papers presented at SMPTE and AES conventions, and published in their respective Journals. For many years he has written a column, "Technology Corner", for the trade publication TV Technology.

Introduction

The purpose of this course is to give the television engineer a solid grounding in the various aspects of video and audio for television, and to serve as a ready reference to the pertinent standards. This course provides an overview of video and audio for television, from the dawn of analog television broadcasting to today's digital television transmission. It is specifically aimed at television in the U.S., although, where appropriate, other world television systems are mentioned.

Who Will Benefit

This course is for anyone who needs a good grounding in television video and audio, from the relative newcomer to the field who needs to learn all about how television video and audio work, to the experienced engineer who needs to know something about digital. It also serves as a ready reference for the working engineer.

Chapter Breakdown

1. Introduction
2. NTSC Video and Transmission
3. Component Video, Analog Resolution and Aspect Ratio
4. Digital Video
5. Digital Scanning Formats
6. Other SD and HD Video Characteristics
7. Baseband Signal Interfaces
8. Video Compression for DTV
9. Video Storage
10. Analog Audio
11. Digital Audio
12. Digital Audio Compression for Television
13. Audio Storage
14. Summary

Enrollment Information

SBE Member Price: $99

Non-Member Price: $149

NTSC Video and Transmission

When U.S. television was in the experimental stage, before World War II, various scanning formats were tried. The earliest experimental formats used by RCA were progressively scanned, i.e., the camera scanned an image line-by-line: Line 1, Line 2, Line 3...; so that entire video frames were scanned in sequence, as opposed to interlaced scanning, as is used in NTSC. While this would seem to be a perfectly logical way to scan a picture, there was the problem of flicker. In order to transmit sufficient scanning lines to provide adequate spatial resolution within the assigned television broadcast channel, the frame rate had to be limited to about 30 frames per second, which, when progressively scanned, provide 30 large-area light flashes per second. This, however, is below the threshold of perceptible flicker, which, for a bright television screen in a brightly lit room, is something over 50 large-area light flashes per second. So interlaced scanning was tried. In this scheme, each frame is divided into two sequentially-scanned half-frames, or fields, each field containing every other scan line. Field 1 contains lines 1,3,5,...525 in the continuous sequence of lines that make up the total frame, while Field 2 contains lines 2,4,6...524 in the continuous sequence of lines that make up the total frame. It must be noted that in the NTSC standard, lines are numbered sequentially within each of the two fields, not with odd numbers in one field and even numbers in the other field. That is, Field 1 contains lines numbered 1 through 262 ½, while Field 2 contains lines numbered 262 ½ through 525. When these sequentially-scanned pairs of fields are viewed, the human visual system integrates them into full pictures. 60 fields per second provide 60 large-area light flashes per second, comfortably above the threshold of perceptible flicker. Interlaced scanning is used in all the analog television systems of the world: NTSC, PAL and SECAM.

The Birth of NTSC

In 1936, two committees of the Radio Manufacturers Association, now known as the Consumer Electronics Association, proposed that U.S. television channels be standardized at a bandwidth of 6 MHz, recommending a 441-line, interlaced, 30 frame-per-second (fps) television system. This system proposed double-sideband, amplitude-modulated transmission, the same modulation system used in AM radio, which modulation system limited the video bandwidth to 2.5 MHz. Because double-sideband amplitude modulation contains two identical sets of sidebands, upper and lower, it can be thought of as a 100 percent redundant system; all the information is contained in either the upper or the lower sidebands, so a single set of sidebands is sufficient to convey it. In 1938, this proposal was amended to use vestigial sideband (VSB), rather than double-sideband modulation. In the VSB system, only the upper sidebands, those above the carrier frequency, plus a small segment, or vestige, of the lower sidebands, are transmitted. VSB raised the transmitted video bandwidth capacity to 4.2 MHz, greatly increasing the spatial resolution capability of the transmission system. In 1941, the first National Television Systems Committee (NTSC) adopted the 525-line, 30 fps, VSB system.

In 1953, the second National Television Systems Committee adopted the compatible NTSC color television system that is used for analog television transmission in the U.S. and a number of other countries in the world, and, in fact, provides the basis for the PAL television system as well. The NTSC system, used in a 6 MHz television channel, permits the transmission of a limiting horizontal resolution of about 333 TV lines per picture height (TVL/PH), a frame composed of 525 total lines, and a frame repetition rate of 29.97 per second. For more about resolution, see Chapter 2, Component Analog Video, Analog Resolution, and Aspect Ratio.

Monochrome NTSC

Let's look first at the parameters of monochrome NTSC. The monochrome NTSC signal is composed of 525 scanning lines per frame. Each frame is composed of two sequentially-transmitted fields, or half-frames, each field containing 262 1/2 lines. Of the 525 lines, an NTSC frame contained 486 active, picture-carrying lines. The remaining 39 lines comprised the vertical blanking interval (VBI), an interval when the raster is blanked in order to conceal, in a CRT, the electron beam's travel from the bottom of the screen back up to the top, to start scanning another field. As the power line frequency in the U.S. is 60 Hz, the field rate of NTSC was also set at 60 Hz, in order to minimize the appearance of power line hum in the picture. A frequency-modulated aural subcarrier was added, located 4.5 MHz above the visual carrier. The arithmetic of monochrome NTSC:

- 525 lines per frame; 262 ½ lines per field; 486 active lines
- Horizontal scan frequency: $f_H = 525/2 \times f_V = 262.5 \times 60 = 15750.00$ Hz, where f_V is the vertical frequency or frame rate
- Visual carrier located 1.25 MHz above the lower band-edge of the 6 MHz channel; aural carrier located 4.5 MHz above the visual carrier

- Aural carrier frequency deviation ±25 kHz

NTSC Compatible Color

By means of a color subcarrier, color information was compatibly added to the monochrome NTSC signal. If we examine a monochrome NTSC signal with a spectrum analyzer, we find that the energy it contains is not spread smoothly across its frequency spectrum. It is, rather, concentrated in "clumps", spaced 15,750 Hz apart across the occupied spectrum. The spectral "spaces" between luminance energy clumps are virtually empty. So in order to add chrominance information to the signal, a color subcarrier was added, at a frequency that caused it to be interlaced into the luminance signal so that its energy clumps fall precisely into the "holes" between luminance energy clumps. In order to make this work, some adjustment had to be made to the frequencies of the NTSC signal. The color subcarrier uses double-sideband suppressed carrier modulation, so that when no chrominance information is entering the subcarrier generator, the subcarrier energy disappears.

The original choice for the color subcarrier frequency was f_{SC} = 455/2 x f_H = 3.583125 MHz exactly. This generated a visible beat between the 4.5 MHz sound carrier and f_{SC} at approximately 921 kHz. It was decided that this beat would be less visible if it were an odd multiple of $f_H/2$. As it was not wished to move the sound carrier frequency, this relationship was realized by moving fSC down by 0.1 percent. In order to maintain the 455/2 relationship between f_H and f_{SC} and the 525/2 relationship between f_H and f_V, the horizontal and vertical frequencies also had to be moved down by 0.1 percent. Here is the full mathematical derivation:

$f_H = 525/2 \times f_V$
$f_{SC} = n/2 \times f_H$, where n is an odd integer
$4.5 = f_{SC} = m/2 \times f_H$, where m is an odd integer
$4.5 = (n/2 \times f_H) = m/2 \times f_H$
$4.5/f_H = n/2 + m/2$
$9.0/f_H = n + m$

If n = 455 and m = 117
$F_H = 9.0/572 = 15{,}734.26...$ Hz

$f_V = f_H/(525/2) = 59.94...$ Hz
$f_{SC} = 455/2 \times f_H = 3.5795454...$ MHz[1]

This is where the fractional field and frame rates, and thus the requirements for drop frame time code and a clever method to distribute digital audio frames over video frames, came from.

Figure 1 below depicts a color NTSC signal as transmitted. This is a diagrammatic representation, illustrating how the various components of the transmitted signal fit into the transmission channel.

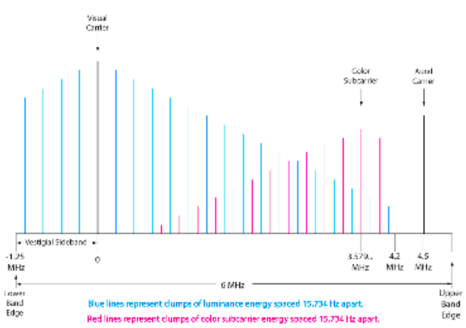


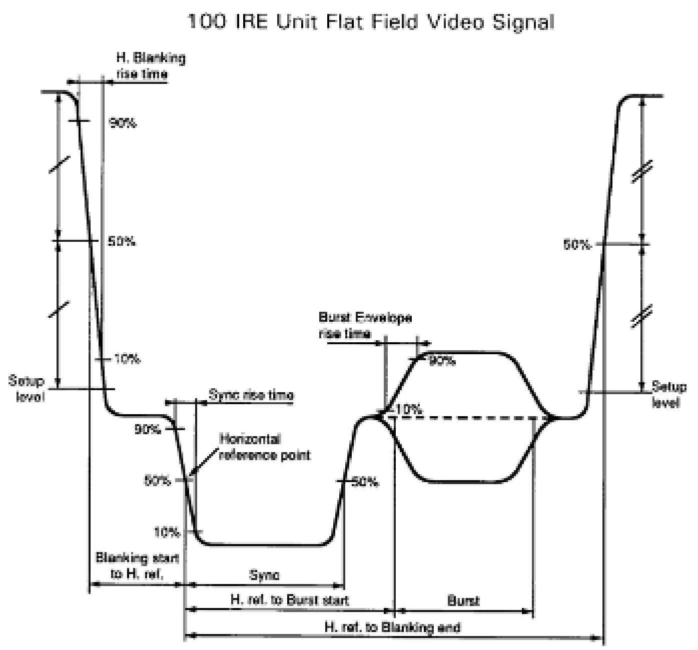
Figure 1
NTSC Color Broadcast Signal

Enlarge

The visual carrier, marked "0" on the horizontal frequency axis, is located 1.25 MHz above the lower band edge of the 6 MHz channel, the aural carrier is located 4.5 MHz above the visual carrier, and the color subcarrier is located 3.579... MHz above the visual carrier. It must be pointed out that while there is a red vertical line marking the location of the color subcarrier, the subcarrier is in fact suppressed. The blue vertical lines represent the clumps of luminance energy, spaced 15,734... Hz apart along the frequency axis. The red vertical lines represent the clumps of color subcarrier energy, also spaced 15,734... Hz apart, and interleaved between the luminance clumps. Of course, this is a representational diagram to illustrate the components of the signal. There are many more energy clumps than are represented in this diagram. The luminance and chrominance components occupy the spectral space between the visual carrier and the point 4.2 MHz above the visual carrier. The space from 4.2 MHz above the visual carrier to the upper band edge is occupied by the aural carrier sidebands. The vestigial sideband occupies the spectrum space from the visual carrier to the lower band edge of the channel.

The vertical blanking interval contains several blanked lines that synchronize the picture vertically. In addition, it contains a number of blanked lines that have been pressed into service for a number of uses, some of which include test and alignment signals, both internal and public; time signals; source identification signals; and closed captioning. When closed captioning was adopted, it was implemented on Line 21 of both fields, which moved the start of active video to Line 22, reducing the transmitted active line count to 483.

In addition to the vertical blanking interval, there is also a horizontal blanking interval between scanned lines, during which interval the raster is blanked while, in a CRT, the electron beam "flies back" from the end of the line, at the right-hand side of the raster, to the left-hand side of the raster to begin the next scanning line. During this period, horizontal synchronization signals are placed, as in Figure 2, which depicts portions of two lines of a flat field video signal.



Enlarge

Figure 2 Horizontal Blanking Interval [2]

Figure 2 shows the horizontal synchronization signals. We see that NTSC, as used in the United States, has "setup", that is, black level is "set up" 7.5 percent above blanking level. When a scan line reaches blanking, it stays at that level for a period, called the "front porch of sync", then it falls to sync level, the lowest voltage level the signal reaches. After the sync period, the signal rises back to blanking level, forming the "back porch of sync", on which appears a burst of color subcarrier, depicted in Figure 2 as the burst envelope. This burst is 9 cycles of sine wave at the color subcarrier frequency. After the back porch of sync, the signal starts scanning the next line. It should be noted here that for many years, color burst has occupied the back porch of sync of all NTSC signals, and color scanning frequencies have been used exclusively, even when monochrome programming was being transmitted.

NTSC transmitters employ negative modulation. That is, the video signal is inverted so that the peak of sync is also the transmitter's peak power output. Video modulates the carrier in the downward direction: the brighter the instantaneous video signal, the lower the transmitted power output level. In this way, the strongest signal is emitted during sync periods, so that a television receiver may synchronize on even a very weak signal.

In order to better understand signal levels, we must understand IRE units. The IRE scale is a linear scale used to measure the relative amplitude of the components of a television signal. The IRE scale is marked from -50 to +120 units, with blanking level being at 0 IRE and reference white level at 100 IRE units. Synchronization signals are below blanking level, and have negative IRE values. Horizontal sync level is at -40 IRE, setup level is 7.5 IRE, and the burst envelope reaches from -20 to 20 IRE. The IRE scale is relative; for composite NTSC signals, the full 140 IRE unit signal's amplitude is 1.0 volt peak-to-peak[3].

Figure 2 shows the relationships between IRE units, transmitter carrier amplitude, and composite video voltage.

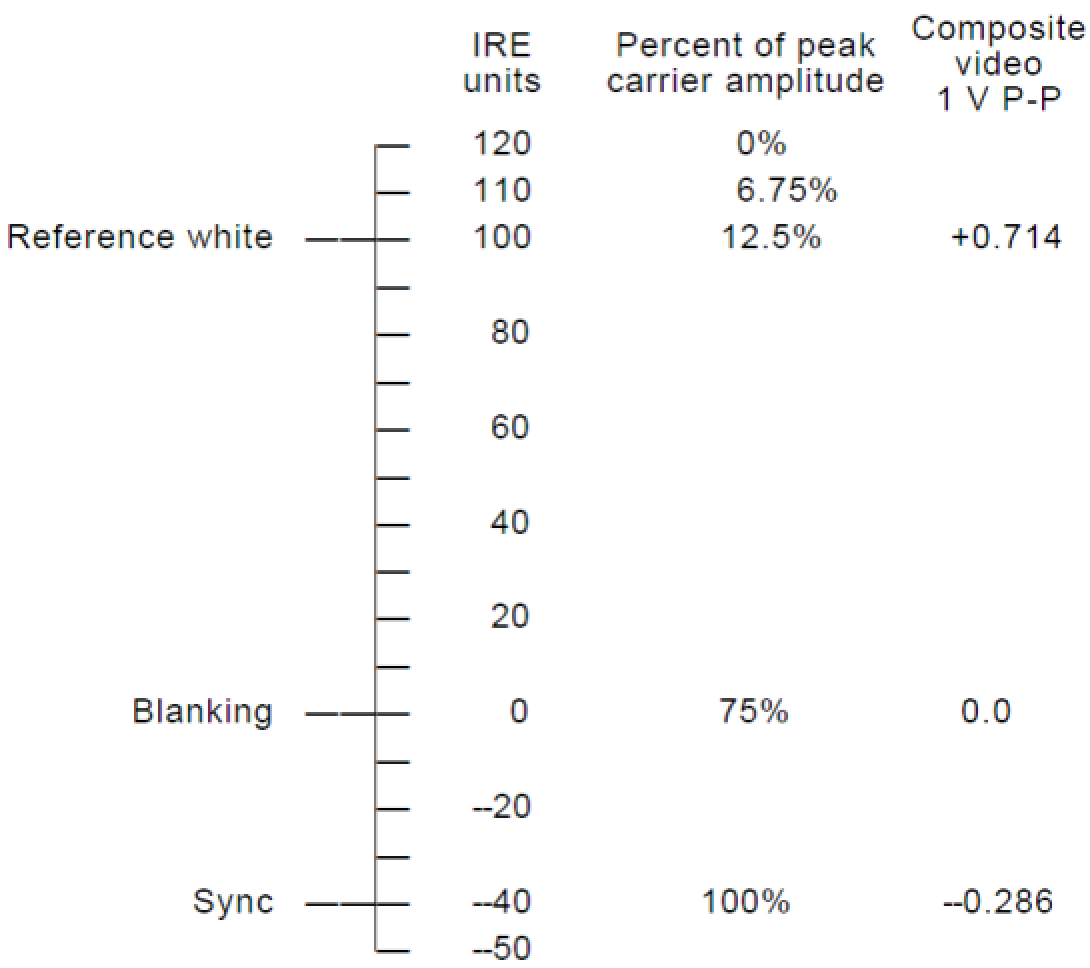


Figure 3, NTSC Video Signal Amplitudes[4]

The NTSC luminance signal, known as the "Y" signal, has a bandwidth of 4.2 MHz for on-air transmission, while the color difference components that are modulated onto the color subcarrier are known as the "I" (In Phase) and "Q" (Quadrature) signals. As the names imply, I and Q signals modulate the color subcarrier in quadrature, or offset from one another by 90 degrees. The bandwidths of the I and Q signals are much lower than that of the Y signal, and I and Q are also asymmetrical in bandwidth. This is a good fit with the human visual system, which perceives luminance at much higher resolution than it does chrominance. The fact that the Q signal's bandwidth is lower than the I signal's bandwidth also takes advantage of the human visual system's characteristics, as the visual system's sensitivity to the orange-to-blue color range, contained in the I signal, is greater than its sensitivity to the purple-to-green color range, contained in the Q signal. When compared to U and V signals, as used in PAL, I is offset 33 degrees counterclockwise from V, the axis through the red-to-green range; and Q is offset 33 degrees counterclockwise from U, the axis through the blue-to-yellow range, on a color plot.

NTSC Audio

Until 1984, NTSC television audio was monophonic. The aural carrier, located 4.5 MHz above the visual carrier (See Figure 1), is frequency-modulated. The maximum carrier deviation for monophonic audio is ±25 kHz, which gives monophonic FM radio, with its maximum carrier deviation of ±75 kHz, a 10 dB signal-to-noise advantage. Both systems employ 75 microsecond pre-emphasis to the incoming audio signals.

Multichannel television sound was added to NTSC, using the BTSC system, in 1984. BTSC stereo uses a scheme similar to that used in FM stereo, in which the sum of the left and right stereo channels is transmitted on the monophonic channel, while the left minus right (L-R) information is modulated onto a double-sideband suppressed-carrier subcarrier, and a stereo pilot is added at one-half the subcarrier frequency. In the case of the BTSC system, the stereo pilot is at the horizontal scan frequency, approximately 15,734 Hz, and the L-R subcarrier is at twice the horizontal frequency, about 31, 468 Hz. Both signals are locked to horizontal sync. In the case of the BTSC system, the signal-to-noise ratio of the L-R subcarrier is enhanced by doubling the amplitude of the L-R signal, and employing noise reduction. The pilot and subcarrier, and the other additional components of the BTSC signal occupy an additional ±50 kHz of carrier deviation, raising the total carrier deviation to ±73 kHz. (This is not an arithmetical error; some interleaving occurs, so the total deviation is not the exact arithmetical sum of the individual deviations.) Because of these enhancements, there is essentially no signal-to-noise penalty for operating in stereo versus mono, unlike the case for FM radio.

The BTSC MTS system also provides for a frequency-modulated Second Audio Program (SAP) subcarrier located at 5_{fH}, also locked to sync, and also employing BTSC noise reduction companding; and for a small nonpublic Professional Channel located at 6.5_{fH}, that has been used for such applications as cuing and IFB. For more information about BTSC MTS, see Hoffner, Randy, *Multichannel Television Sound Broadcasting in the United States*, presented at the 82nd Convention of the Audio Engineering Society, London, March, 1987, and published in the *Journal of the Audio Engineering Society*, Vol. 35, No. 9, September, 1987, and available at www.aes.org .

NTSC, like PAL, and unlike the video systems that followed it, is both a production and a transmission system. It has historically been used to produce video as well as to transmit video, with the recognition that when it is used for production, it does not have the same bandwidth and resolution constraints that are required for its transmission.

The other two analog television broadcast systems used in the world, are PAL (Phase Alternating Line) and SECAM (Sequential Avec Memoire). PAL is an NTSC variant, while SECAM is a system that uses frequency-modulated color subcarriers, unlike the amplitude-modulated subcarriers of NTSC and PAL.

At this time, all full-power television broadcast stations in the U.S. are broadcasting ATSC digital signals, having shut off their analog signals. However, NTSC translators and repeaters may still be operated, and cable systems are required to provide analog signals to their subscribers for an indefinite period of time. Add to this the enormous volume of television programming produced in NTSC since the late 1940's, and still in daily use, and NTSC is still very much with us.

[1]I am grateful to television technology pioneer Bernie Lechner, one of the luminaries at RCA Laboratories, for this table, from a private communication.

[2] Figure 2 is taken from SMPTE 170M-1999, *Composite Analog NTSC Signal - NTSC for Studio Applications*, Society of Motion Picture and Television Engineers, White Plains, NY, 1999, www.smpte.org, p. 9.

[3]Information extracted from Ibid, *Annex B (informative), IRE units, p. 17*.
[4]Ibid, *Annex B (informative), IRE units, p. 17*.

All the analog television transmission systems of the world use ________ scanning.

- Progressive
- Interlaced
- Quincunx
- Vertical

The NTSC broadcast system uses ______________ to transmit video.

- Double-sideband suppressed-carrier amplitude modulation
- Frequency modulation
- Phase modulation
- Vestigial sideband

The NTSC broadcast system has _________ total scan lines.

- 525
- 404
- 576
- 1080

Ancillary signals may be carried in the NTSC _________.

- Fukinuki holes
- Horizontal blanking interval
- Vertical blanking interval
- Vestigial sideband