

how it works

PERCEIVED TELEVISION-IMAGE QUALITY DEPENDS ON SOURCE-MATERIAL CHARACTERISTICS, ON HOW THAT MATERIAL GETS TRANSPORTED TO THE RECEPTION SYSTEM, AND ON THE CAPABILITIES OF THE CIRCUITS IN THAT RECEPTION SYSTEM.

A crash course in color conversion

By Brian Dipert, Technical Editor

ON ONE END OF THE VIDEO CHAIN is the camera used to capture the images onto film, magnetic tape, or other media. On the other end is the television

used to display those images (Figure 1). In between, numerous image transformations take place, both to reduce required transmission bandwidth and to reduce required storage capacity. These transformations take advantage of the human vision system's greater sensitivity to luminance (black, white, and in-between shades of gray) versus color and to some colors in the visible spectrum versus others. They can, however, create perceptible image degradations (see sidebar "Clearer views").

IMAGE CAPTURE

Early attempts at color television included one proposal, briefly approved by the FCC, that placed a tricolor filter wheel in front of a black-and-white

set. The wheel rotated at 20 revolutions per second, synchronized with a TV station that sent a sequential RGB pattern that varied every 60th of a second. The approach delivered highly saturated colors and did a good job of reproducing still images. But consumers could not directly view the transmitted signals on the black-and-white set, a backward-compatibility issue for the installed base of televisions whose owners might not want to pay for a then-expensive color wheel. Color-wheel-equipped TVs weren't backward-compatible with black-and-white broadcasts, either. Frame refreshes every 20th of a second were also too slow to acceptably reproduce objects in fast motion, and filter-wheel synchronization with multiple broadcast signals from multiple stations was problematic.

Instead, the NTSC decided to shoehorn color information into the 4.2-MHz black-and-white video signal, with sound located at 4.5 MHz. Studies of human light sensitivity by the CIE (Commission International de L'Eclairage) in 1931 showed that our visual system is most sensitive to green (G) light (specifically yellow-green), followed by red (R), and

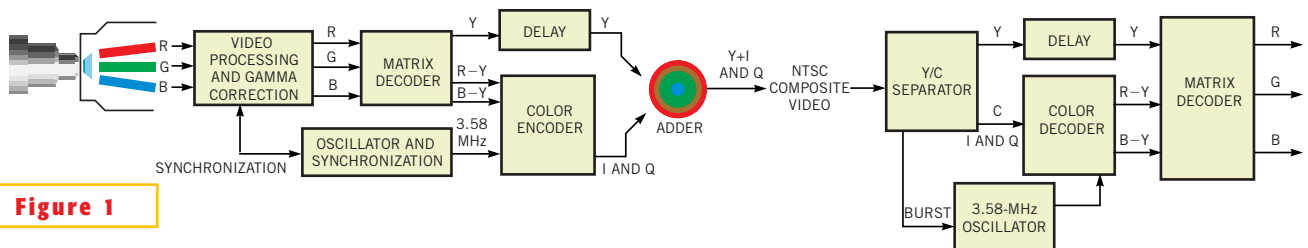


Figure 1

The conversion from RGB to composite video and back to RGB again for display is complex and beset with prospects for image degradation (courtesy *Widescreen Review* magazine).

then blue (B). Therefore, the luminance (Y) signal derives from the following proportional equation: $Y=0.3R+0.59G+0.11B$. Luminance alone, displayed on a black-and-white television, assured the desired backward compatibility. Along with the R-Y and B-Y signals, you now had three equations and three unknowns and could therefore reconstruct the original RGB information.

The R-Y and B-Y signals undergo additional alterations deriving from the CIE data, both to prevent amplitude overexcursion and to squeeze them into the available broadcast signal bandwidth. The color difference signal U derives from the equation $U=0.493(B-Y)$, and its companion $V=0.877(R-Y)$. Closely related I and Q derive from the equations $I=0.736(R-Y)-0.268(B-Y)$, and $Q=0.478(R-Y)+0.413(B-Y)$. The bandwidth of I is limited to 1.3 MHz, whereas Q extends to only 0.6 MHz. U and V are both bandwidth-limited to 1.3 MHz. Because most videocameras of the era couldn't capture information above approximately 2.8 MHz, the spectrum space between 2.8 and 4.2 MHz was available. Therefore, the NTSC modulated both color-difference signals on a common 3.58-MHz carrier signal (the midpoint between approximately 2.8 and 4.2 MHz), 90° out of phase with each other, and added the resultant chrominance (C) signal to a time-delayed version of Y (Figure 2a).

DECODE AND DISPLAY

Voilà. You have the NTSC composite-video signal. For broadcast purposes, transmitters RF-modulate it on their specific channel-carrier signals. Laserdiscs also store video in analog composite form, but they aren't handicapped by the 4.2-MHz bandwidth restriction and can therefore at least theoretically deliver higher resolution images. To decode and display the composite-video image, you reverse the steps of the encoding process, and end up with RGB information that goes to the television tube, albeit with a more limited bandwidth, and, therefore, more limited resolution, than the original. However, one complication has caused a host of problems for composite video over the years.

Improvements in videocamera technology have increased the resolution of captured images and, along with it, the bandwidth of the luminance signal. This situation resulted in an intermingling of luminance and chrominance between 2.8 and 4.2 MHz. Fortunately, the spectral energy of Y and C and aliases of these signals cluster around specific frequencies related to the color subcarrier frequency (FSC) and horizontal-scan rate (FH) (Figure 2b). For still images, Y and C interleave where they overlap, fitting together and looking something like the teeth of two combs. However, in response to object movement within images, the frequency clusters

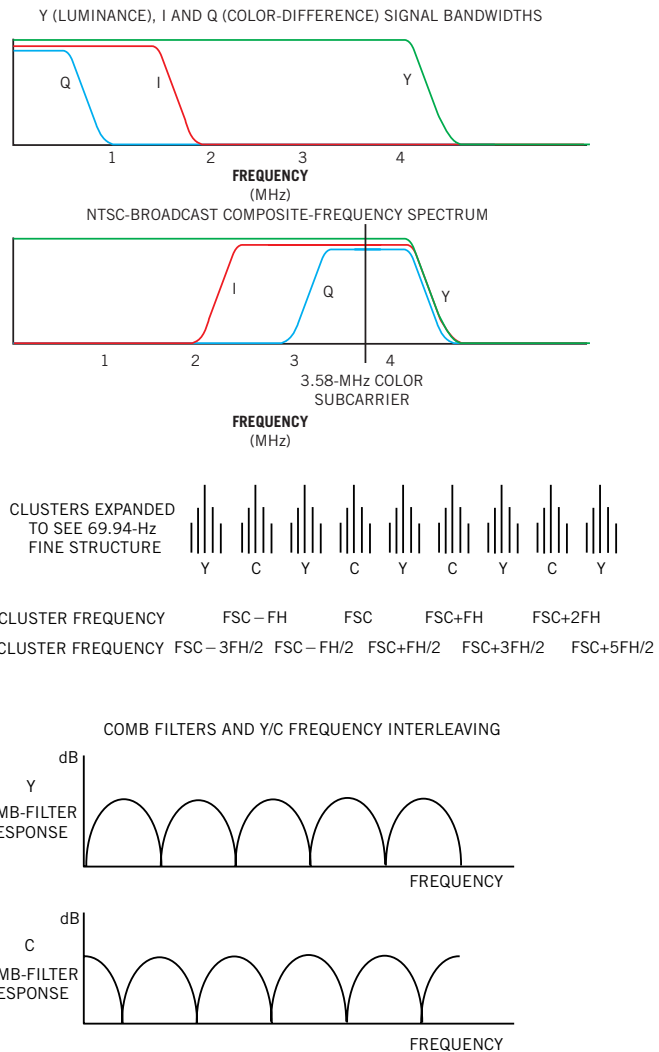


Figure 2

Composite-video (a) images with no motion contain interleaved patterns of luminance and chrominance and their aliases (b). A comb filter, intended to separate chrominance and luminance, has a similar interleaved pattern (c) (courtesy *Widescreen Review* magazine).

spread out, “smear,” and become difficult or even impossible to cleanly separate.

The simplest and cheapest “brute-force” approach to signal division employs a lowpass filter at approximately 2.5 MHz to derive luminance and a bandpass filter centered at 3.58 MHz and extending down to 2.3 MHz or so to derive chromi-

CLEARER VIEWS

This short article provides only a simplified explanation of the color-video-conversion process and of composite-video filtering. To further explore these topics, check out *Imaging Science Theater 2000*, published by *Widescreen Review* magazine (www.widescreenreview.com) or any of the additional information sources listed in **Reference 2**.

Also, please note that the terms Cr and Cb, and Pr and Pb, have two common meanings. In this article, Cr and Cb, and Pr and Pb, respectively, denote the digital and analog video representations of R-Y and B-Y. However, they can also refer to, respectively, the BT.601 and BT.709 color-space versions of R-Y and B-Y.

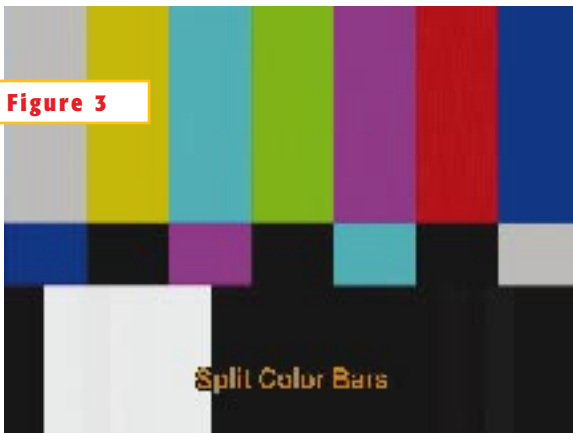
nance. Several problems exist with this method. First, it limits the bandwidth of the luminance signal, throwing away fine-horizontal-resolution detail. Color leaking into the luminance signal causes cross-luminance artifacts, which manifest themselves as stationary or moving black-and-white dot patterns at abrupt color transitions (Figure 3a). Luminance leaking into the filtered color signal, conversely, causes cross-color effects: rainbow patterns in picture areas with fine detail (Figures 3b, 3c, and 3d).

The next step up in complexity is the 1-H (one-horizontal-delay) comb filter (Figure 2c). It subtracts the current scan line's information from a time-delayed version of the previous scan line to derive the current C and adds the two lines together to form the current Y. More complex 2-H filters are less common, as their added cost doesn't result in proportionally greater performance. Because the comb filter doesn't lowpass-filter luminance, it doesn't degrade horizontal resolution (though it can degrade vertical resolution), and it does a better job of suppressing cross-color artifacts, though it still

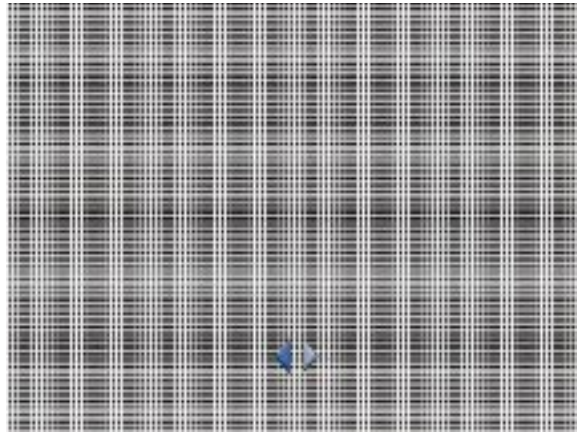
struggles with fine-detail diagonal lines. However, when the color information is radically different from one scan line to the next, such as at an object edge, cross-luminance artifacts remain.

Today's most common composite-video-filter type, 2-D adaptive filters focus their efforts on "dot crawl." These filters examine three consecutive scan lines at a time. Say, for example, the first two lines are the same color, and the third line is different. The 2-D adaptive filter sends the first two lines off to a 1-H filter and, next time, focuses on the second two lines. Two-dimensional adaptive filters suppress cross-luminance artifacts at simple horizontal edges, although they can't help in the rare cases in which three consecutive scan lines contain three different colors. They also don't suppress cross-luminance at horizontal or diagonal edges, and they only marginally improve cross-color suppression over simpler filter types.

Three-dimensional motion-adaptive filters encompass multiple frames' scan lines. If they detect no changes from one frame to the next, they subtract and add identical scan lines of both frames to



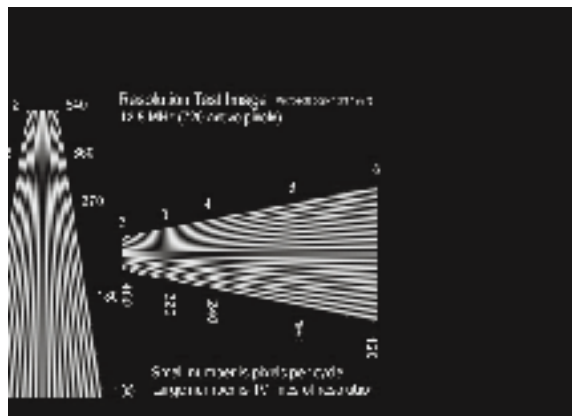
(a)



(b)



(c)



(d)

Color bars (a, courtesy Ovation Software) are a good way to observe cross-luminance patterns; crosshatch grids (b, courtesy Unapix Entertainment) and even the salt-and-pepper hair color of well-known politicians (c, photo by Bruce Fox, University Relations, Michigan State University) bring out cross-color rainbow effects. Images in motion also help test for proper deinterlacing (d, courtesy Microsoft).

derive Y and C, thereby nearly eliminating cross-luminance and cross-color. When they do detect frame-to-frame differences, suggesting object motion or object color change, they back off to the previously discussed intraframe 2-D adaptive function. Because they require multiframe buffer memory, 3-D motion-adaptive filters are more expensive than simpler filter types. If they incorrectly interpret an absence of frame-to-frame differences, the quality results can be even worse than if you used a simpler, cheaper filter. Similarly, if they sense frame-to-frame differences that haven't actually occurred, they produce results no better than those you'd obtain with a 2-D adaptive filter. Only when they guess right does the enhanced quality offset their higher implementation cost.

ANALOG ALTERNATIVES, DIGITAL DERIVATIONS

My hands-on project appearing on page 83 of this issue of *EDN* covers, among other deinterlacers, Sage's (www.sage.com) FLI2200 digital video processor. The chip's motion-adaptive cross-chroma suppression feature is impressive. On both test patterns and real-life video material, it did an effective though incomplete job of suppressing rain-

bow-pattern artifacts that remained after the composite video passed through the 2-D-adaptive filter in the evaluation board's FLI2000 video decoder. In the process, the FLI2200 didn't seem to inject any artifacts of its own into the images. The DVD player, progressive-scan television, and other deinterlacers tested also included 2-D adaptive filters but not the additional color processing of the FLI2200 and therefore provided good data points for comparison.

Pragmatically, though, the *best* option for dealing with composite video is to *not* deal with it but instead to always hold the video in a form that separates chrominance and luminance. VHS recorders, for example, separate chroma and luma, and DVD-video players take the additional step of splitting Y, Cr (R-Y, known as Pr in the analog domain), and Cb (B-Y, known as Pb in the analog domain). As more and more video equipment incorporates S-Video outputs and, at the high end, component-video outputs, and as television broadcasts enter the digital era, you'll theoretically need to worry less about cross-luminance and cross-color suppression. These steps, however, are insufficient to ensure that the quality of images on your television matches the

quality of images the camera originally captured.

To completely eliminate cross-luminance and cross-color effects, the video material must never experience a composite-video conversion. Realize that when you record the latest episode of the *X-Files* onto VHS tape, you're capturing a composite signal that travels through the recorder's chroma/luma-splitting input filter. And, when you watch that episode through a player's composite-video output, the player is recombining luma and chroma, only for the television's filter to again separate it. Connecting together the recorder's S-Video output and the TV's S-Video input, if they exist, will eliminate this final redundant and further degrading conversion.

Prerecorded VHS movies you rent from the video store aren't necessarily cross-luma- and cross-color-free, either, especially if they're old films, and, therefore, the only master tape is a composite-video version. Remember, too, that S-Video's C (that is, I+Q or U+V) signal still comprises reduced-bandwidth versions of R-Y and B-Y.

The move to the digital domain exemplified by DVD and DTV is also limited

by the inherent quality of the source material. And even pristine RGB or 4:4:4 component video still goes through bandwidth- and therefore resolution-limiting steps to meet digital storage- and transmission-capacity requirements. For example, DVDs store video in the 4:2:0 format. This convention indicates that not only are there half as many Cr and Cb samples as Y samples on each scan line, there are half as many scan lines of Cr and Cb data as there are Y lines (**Reference 2**). Oversampling during playback reconstructs an approximation of the original 4:4:4 material. But it's only an estimate. Interpolation cannot reconstruct previously lost resolution.

Fortunately, for most gradual-transition video material, subsampled component video is virtually indistinguishable from the original RGB. You wouldn't want to use the 4:2:0 video format, though, to transmit or archive fine-detail computer graphics. □

You can reach
Technical Editor
Brian Dipert at
1-916-454-5242,
fax 1-916-454-5101,
e-mail bdipert@pacbell.net.



REFERENCES

1. Dipert, Brian, "Video quality, a hands-on view," *EDN*, June 7, 2001, pg 83.
2. Dipert, Brian, "Compression puts images on a diet," *EDN*, June 18, 1998, pg 71.