Digital Video Overview

The following provides a very brief overview of video signal fundamentals. For more detailed information, the reader is referred to the following books:

"*Video Demystified: a Handbook for the Digital Engineer*", by Keith Jack, Newnes Publishers, 2001.

"*Digital Video and HDTV: Algorithms and Interfaces*", by Charles Poynton, Morgan Kaufmann Publishers, 2003.

The NTSC/PAL Composite Signal

Figure 1 shows an illustration of the structure of the common National Television System Committee (NTSC) analog composite baseband video signal. PAL timings vary slightly (as well as number of lines, colorburst phase, etc..), but the overall video signal structure is very similar.

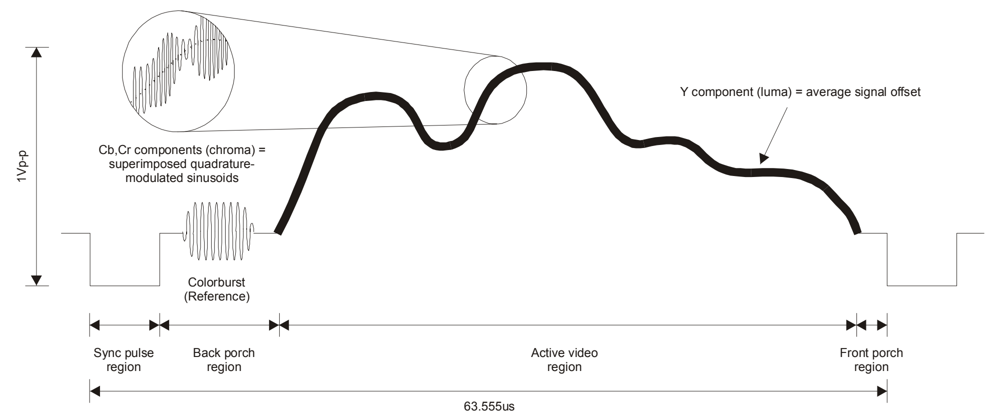

Figure 1: A typical single line composite NTSC video signal

There are four regions of interest: the sync pulse, the back porch, the active video region, and the front porch.

Sync pulse: this is the lowest level of the video signal used to synchronize a receiver to the incoming signal on a line-by-line basis.

Back porch: the back porch is set at the black level and contains a sinusoidal waveform called the colorburst. The colorburst serves as a color subcarrier reference with respect to which the color information of the active region is decoded.

Active video: the active video region contains the information for three different components, the luma (Y) as well as two chroma components (Cr, Cb). The luma represents the intensity information of the image at every point along the line. In the active region, the luma component is represented as the average value of the signal (this preserves compatibility with the old black & white TV standard). To add color information (while preserving b&w backwards compatibility), two color components are added to the luma signal using quadrature-modulation. The resulting superimposed color signal component is characterized by a sinusoidal wave whose phase with respect to the colorburst reference represents the hue of the color, and whose amplitude represents the saturation of that color.

Front porch: the front porch is set at the black level.

Image Encoding

An image is converted into an NTSC video signal by a video camera. The encoding process breaks down the information contained in the full image into a sequence of single-line video signals. The lines are encoded in a raster scan pattern as shown in Figure 2. The NTSC standard calls for scanning of 525 horizontal lines in a video frame at 30 frames per second. This just exceeds the threshold of human retinal persistence where image flicker is no longer perceived during playback.

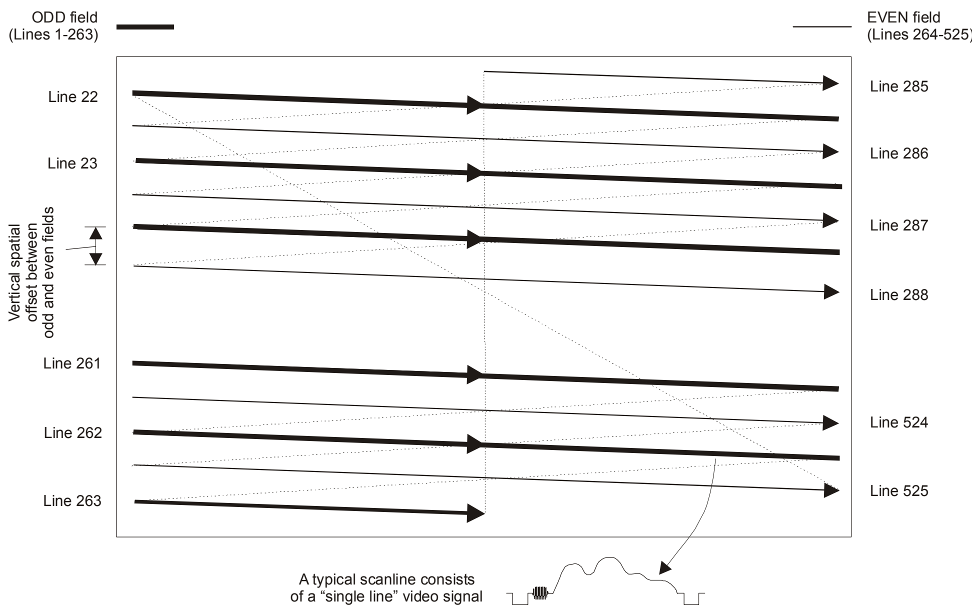

Figure 2: Video image raster scan pattern

Early TV pioneers decided that it would be possible to double the image refresh rate to 60 Hz by introducing the video interlacing technique. This trick again uses the retinal persistence of the human eye. Instead of scanning all the 525 lines of the image from top to bottom, the image is generated by scanning the 262.5 odd lines first, then followed by scanning the 262.5 even lines. Because these two half-images, or fields, are actually vertically offset in space (and displayed as such), the eye is again fooled into seeing a static 525 line picture in space.

Odd video field: is composed of the 262.5 odd lines of the video image.

Even video field: is composed of the 262.5 even lines of the video image.

Video frame: the full 525 image with odd and even field interlaced together.

For PAL, there are 625 lines per frame, and 312.5 lines per field, and 50 Hz field rate (25 Hz frame rate).

Freezing an Image

One of the side-effects of the video interlace technique is that the two fields ***are actually acquired at different instants in time***. For the NTSC standard, the delay between acquiring the ODD and EVEN fields is 1/60 sec. What that means is that the ODD and EVEN half images that belong to the same video frame are slightly temporally mismatched. When displayed on a TV screen, our eyes are easily fooled by the rapidly changing image sequence. However, when freezing a full frame of video, the mismatch becomes obvious when an image of a fast moving object is taken. This is illustrated in Figure 3.

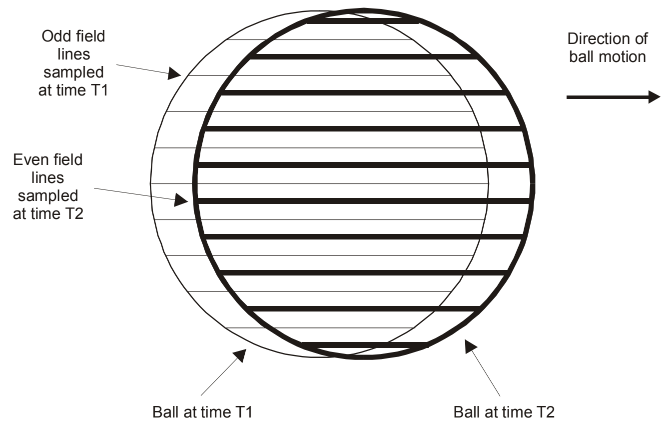

Figure 3: Spatio-temporal mismatch of the odd and even fields in a frozen video frame

There are some de-interlacing techniques available to try and compensate for the motion artifacts by identifying the regions of large motion and interpolating. This usually requires a fair bit of computational resources. The most common approach is to only freeze ONE of the two fields. The tradeoff is that only half of the vertical resolution is available, but at a fixed instant in time. If required, the other field can simply be interpolated from the acquired one through simple vertical interpolation techniques.

The uCFG allows the digitization of a single field of video with full control over which field is acquired (ODD, or EVEN). This provides the option to acquire a full-frame, high resolution image of static scenes. To do this, the ODD field is first captured and downloaded, followed by the EVEN field capture and download. The two are interlaced together and a full frame of 525 lines is obtained.

ITU-R BT656 Digital Video Stream Format

Digitizing a composite analog video signal requires and analog-to-digital converter (ADC). Some front-end analog signal preprocessing is also required to clamp the video signal and make sure it is positioned within bounds of the ADC. A video decoder is a specialized IC that performs all these steps and outputs video samples in a standard digital video format.

The ITU-R BT601 standard describes the fundamentals of the video digitization process. The ITU-R BT656 standard further specifies an 8 bit digital video parallel interface. Both standards apply to 4:2:2 sampled digital video. The 4:2:2 terminology simply means that during the video sampling, the luma is digitized at every pixel position (13.5 MHz sample rate) and the chroma components are sub-sampled by a factor 2 (6.75 MHz each). This sub-sampling is possible since the human eye is less sensitive to color spatial resolution then to intensity spatial resolution. The direct benefit is that a 4:2:2 data stream is 2/3 the overall data size as compared to a non-decimated image.

Figure 4 illustrates the sampling process. As shown, every one of the 720 pixel positions sampled has a Y, Cb, Cr value, except that every 2nd Cb and Cr components are dropped (decimated). The resulting BT656 4:2:2 data stream (gray box) is a simple concatenation of these 8-bit luma and chroma samples. The aggregate BT656 data stream rate is of 13.5 MHz + 2 x 6.75 MHz = 27 MHz.

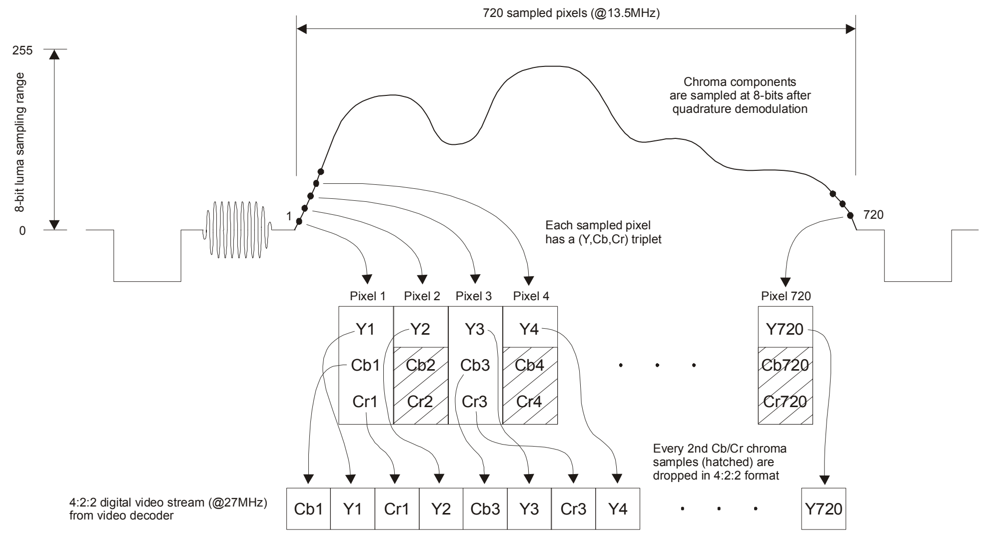

Figure 4: ITU-R BT656 4:2:2 digital video stream

Color Space Conversions

So far, we have explored the fundamentals of the analog composite video standard as well as digital video encoding techniques. As described previously, the color information contained in a composite video signal is decoded into the Y, Cb, Cr color space as per the ITU-R BT656 standard. This color space is very convenient for many types of image processing where operations are mainly performed on the black & white video component (in that case, the Y luma component can be used directly). To display color images on a PC, however, the more familiar RGB color space is used. In this color space, each pixel is described by an 8-bit value of its red, green and blue components. In order to perform a conversion from the Y, Cb, Cr color space to the RGB color space, a simple transformation can be applied:

$$R = 1.164 \times (Y-16) + 1.596 \times (Cr-128)$$
$$G = 1.164 \times (Y-16) - 0.392 \times (Cb-128) - 0.813 \times (Cr-128)$$
$$B = 1.164 \times (Y-16) + 2.017 \times (Cb-128)$$

(R, G, B are gamma-corrected and ranging between [0-255])

Equation 1: (Y, Cb, Cr) to (R, G, B) color space transformation equations

Liability Disclaimer

Digital Creation Labs Incorporated gives no warranty either expressed or implied and specifically disclaims all other warranties, including warranties for merchantability and fitness. In no event shall the liability of Digital Creation Labs Incorporated exceed the buyer's purchase price nor shall Digital Creation Labs Incorporated be liable for any indirect or consequential damages.

Life Support Policy

Digital Creation Labs Incorporated's products are not intended or authorized for use in any critical medical, life support, or life-saving devices or applications or any other system whose failure to perform correctly could result in life support failure.

FCC, CE, IC Certification

Digital Creation Labs Incorporated's products are provided as OEM unfinished products to be integrated into a customer's end product design. The customer is therefore responsible to test for, and obtain all the required EMI/EMC certifications of their final product wherever applicable.

Trademarks

Microcontroller Frame Grabber, uCFG and uCFGEVAL are all Trademarks of Digital Creation Labs Incorporated.