CS150 FPGA Board Digital Video in a Nutshell

Many video standards have been developed over the years since television was invented, and with the advent of digital video, even more formats and standards have evolved. But for a number of reasons one specific digital video standard is of interest to us with respect to the CS150 FPGA board: the 4:2:2 format section of the ITU-R BT.601/656 specification. A couple of these reasons are:

- It is a very common standard that encompasses many of the currently popular video formats: normal 525-line NTSC and 625-line PAL composite video, wide-screen NTSC and PAL, HDTV, as well as component video formats used in many broadcast studios and video production centers;

- The Analog Devices video encoder and decoder chips used on the CS150 board are specifically designed to be compatible with this standard.

The purpose of this standard is to define a digital coding scheme for these commonly used video formats. It applies to both *component video*, where the analog source and destination signals are separate red, green and blue signals, and *composite video*, where a single analog signal is a composite of three values derived from the red, green and blue signals. These composite values are the luminance (intensity) and two color difference values. Conversion between the component values and the composite values is a straightforward mathematical process.

The video encoder and decoder chips used on the CS150 board are designed to be compatible with composite video signals, as well as component video signals. However, as configured on the CS150 board, these chips are used only for converting analog composite video to digital data and then back to analog composite video. The AD7185 video decoder accepts analog composite video and converts it to 8- or 10-bit digital video, whereas the AD7194 video encoder accepts 8- or 10-bit digit video and converts it to analog composite video.

Analog Video

Two very common types of video that most of us observe almost daily are television broadcasts and computer video output. In both cases, what we are basically watching is full color, vertically and horizontally scanned moving pictures. If you look at the signals contained in the cable going from your computer to your computer monitor (VGA, XGA, etc.) you would find five signals: three color signals and two sync signals. On the other hand, If you look at the cable going to a composite video (television) monitor, you would find only one signal. How does the technology accomplish this?

We can reproduce color, in an additive sense, by superimposing appropriate amounts of red, green and blue colors. In addition, it is quite straightforward to produce a picture on a cathode ray tube by scanning an electron beam in horizontal and vertical directions. By combining these two seemingly separate operations, we can produce a full color picture on the tube, and if we repeat this operation many times a second the picture becomes a moving picture. We basically need five signals to do this: three analog signals representing the three colors red, green and blue, and two digital signals that command the electron beam to start the scan of a line (horizontal sync) and start a new screen (vertical sync). This group of five signals is at the heart of what is referred to as “component video” which is very commonly used in professional video applications, high-end amateur applications, and in computer video.

However, with regular television we have only one wire! So, over the years video engineers have developed some very ingenious ways of getting all the required information on one wire. Television started as black and white, and so there was only one analog-type signal, representing the overall intensity, generally called *luminance.* When color television was invented, the government commanded that the signal format be compatible with all the black and white televisions in existence, so the luminance signal was retained and the color information as added by using some very interesting techniques that included generating color difference signals. The result is a set of three analog signals, called Y, U, and V that can all coexist on a single wire. The Y signal represents the luminance value, and the U and V contain the color information. Additional tricks, carried over from black and white technology, allowed the horizontal and vertical sync signals to also end up on the single composite video wire.

Digital Video

The purpose of the AD7185 video decoder on the CS150 board is to convert the analog composite video signal containing the Y, U, and V color values to digital form and extract the horizontal and vertical sync signals, and then produce a continuous stream of 8- or 10-bit words which can be processed, stored, and transmitted. Conversely, the purpose of the AD7194 video encoder is to take a properly formatted digital stream of 8- or 10-bit words and convert it to an NTSC compatible analog composite video signal. The specific form of the digital video data is the so-called 4:2:2 format of the ITU-BT.601 specification.

When the Y, U, and V signals are in digital form, they are labeled Y, Cb, Cr.

The picture being viewed on a video screen is comprised of pixels, lines and two interlaced fields. The following defines the values for these items for U.S. standard NTSC video:

- Samples (pixels) per line 858
- Lines per frame 525
- Frames per second 29.97
- Resulting samples (pixels) per second 13.5 MHz

- Active (viewable) samples per line 720
- Active (viewable) line per frame 487

For the purpose of sending this information as digital video:

- Words per sample 2
- Resulting words per second 27 MHz

The digital video information is transmitted synchronously with a 27 MHz clock. For every four clocks the active video information is sent in the following order: Cr, Y, Cb, Y. This means that the sampling rate for the luminance is 13.5 MHz but for the color signals the sampling rate is only 6.75 MHz. This is allowable because human visual perception is less sensitive to color than it is to intensity.

Horizontal Format

The overall format of one complete line of the digital video data stream includes the following sections. The horizontal scan is considered to start with the beginning of the EAV section.

- EAV - End of Active Video (timing reference signal) 2 samples (4 words)
- Horizontal blanking 134 samples (268 words)
- SAV - Start of Active Video (timing reference signal) 2 samples (4 words)
- Active video 720 samples (1440 words)

EAV Timing Reference Signal:

The EAV Timing Reference consists of four words in the following format:

P9	P8	P7	P6	P5	P4	P3	P2
1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
1	F	V	1(H)	E3	E2	E1	E0

where:

F = Field select (defines which vertical scan during interlace scanning)
V = Vertical blanking
H = 1 indicating EAV
E3 = V xor H
E2 = F xor H
E1 = F xor V
E0 = F xor V xor H

Horizontal Blanking:

The horizontal blanking section consists of a repeating pattern: 1000 0000 0001 0000 ....

SAV Timing Reference Signal:

The SAV Timing Reference consists of four words in the following format:

P9	P8	P7	P6	P5	P4	P3	P2
1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
1	F	V	0(H)	E3	E2	E1	E0

where:

F = Field select (defines which vertical scan during interlace scanning)
V = Vertical blanking
H = 0 indicating SAV
E3 = V xor H
E2 = F xor H
E1 = F xor V
E0 = F xor V xor H

Active Video:

The active video section consists of the Y, Cb, Cr data in the following sequence. The Y, Cb and Cr values are the scaled, offset, digitized versions of Y, U and V.

- Cr
- Y
- Cb
- Y

Vertical Format

The details of the vertical format are shown in the complete screen diagrammed below. The scan is of the interlace type which means that the odd lines are scanned first, followed by the even lines. The symbols "F" and "V" refer to the values contained in bits P8 and P7 of the EAV and SAV timing reference words.