Quiz

MPEG compression and digital formats.

Review of Quiz from Last Time

What is bit rate for 4:4:4?

Rate of 13.5 MHz * 3 (Y, R-Y, B-Y). Chrominance is fully-sampled, as much as Y. 405 megabits per second.

What is binary representation of 3FFh?

Ten ones.

Reserved upper and lower part of numbering system for reserved values. 3FF and 0 are two of the reserved values.

What is EAV?

End of active video.

What information is between SAV and EAV?

The video payload (active video).

What information is between EAV and SAV?

Ancillary data. Audio too.

What is the sample rate of composite digital video?

This rate is notated as 4Fsc, four times the subcarrier frequency. 4 * 3579545 = 14.318181. Why multiply by 4? Nyquist theorem says must have a sample rate of at least two times the highest frequency.

What is the bit rate for composite digital video?

143 megabits per second.

Is there any room to embed audio in D2 composite digital video? Where?

Must digitize the entire NTSC video signal in order to maintain the color subcarrier reference; can't do this precisely without digitizing it. As long as we have to capture burst, we might as well capture the entire sync.

What's the biggest problem when we carry NTSC video into the digital domain? **Interleaving of chroma and luminance** (cross-color and cross-luminance). Gets encoded into the digital format. We really don't want to keep this but it is a consequence of digitizing the entire NTSC video.

Digitize up to 5 or 5.5 MHz baseband, not just the 4.2 MHz.

What is the relationship between sync and subcarrier? **SCH** – subcarrier to horizontal. Must preserve this relationship as well, else we won't have correct color.

HANC allows for how many bits (X-bit) of data? What is the value of X?

Value of X is 10. That's how many bits we capture per sample.

VANC allows for how many bits (X-bit) of data? What is the value of X?

Value of X is 8. Goes back to the days of D-2, made this compatible with D-2.

Lecture

Finish MPEG.

Notation

Number:number:number:number.

First number means luminance horizontal sampling rate (4 Fsc). The value of this number is typically 4. Originally, luma f_S as multiple of 3 3/8 MHz (one-fourth of 13.5 MHz).

PAL uses 4.2 MHz (approximately) as its subcarrier. The people that make professional television equipment found it too expensive to make separate machines for US and European markets, so chose to make them the same. So, had to find a suitable standard sampling rate; came up with 13.5 MHz, which works for US and European systems.

Second number is the Cb and Cr horizontal factor (relative to the first digit). By using 2, we compress video by including less chroma data.

Originally, third number was Cr and second number was Cb, but always sampled red and blue at the same rate, so no need to have separate numbers. This scheme failed to anticipate vertical subsampling.

Third number will be either same as second digit or zero. Zero indicates Cb and Cr are subsampled 2:1 vertically. Nonzero indicates there is no vertical subsampling.

Fourth number optional; if present, same as luma digit, indicating alpha (key) component. Keys are not meant to be choppy, so want to use a large number to get good detail on a key. The keying information sent on a separate wire from the 4:2:2 since else we'd have to add another 13.5 MHz * 10 onto the bandwidth of 601.

See diagram notes for diagrams of the following.

Using 4:4:4

Four:four:four samples every pixel.

Using 4:2:2

four:two:two samples chrominance on every other pixel, samples luminance every time. No vertical subsampling. Horizontal chroma resolution cut in half. Reduces bandwidth of a video signal by one-third with little to no visual difference.

Using 4:2:0

Four:two:zero is emitted MPEG. Two chrominance samples for every 4 luminance samples, odd lines only. Chrominance halved in both directions.

Using 4:1:1

Four:one:one If look at the Internet, there is a big error on a huge number of people's web pages. Will see picture like diagram 2. Is actually 4:2:0 since the chroma is just shifted over. It does not follow the definition of 4:1:1 (it violates the rule of 4:1:1) since 4:1:1 is not vertically subsampled, but what is shown is vertically subsampled. See real 4:1:1 in diagram. Chrominance sampled every 4 pixels horizontally and every pixel vertically (within the vertical column that is being sampled horizontally).

Using 4:1:0

Four:one:zero. Possible (some codecs support it), but not widely used. Half the vertical and one quarter the horizontal color resolutions, with only one eighth of the bandwidth of the maximum color resolutions used. Uncompressed video in this format with 8-bit quantization uses 10 bytes for every macropixel (4 x 2 pixels).

Using 3:1:1

Used by Sony in their HDCam High Definition recorders. In the horizontal dimension, luma is sampled horizontally at three quarters of the full HD sampling rate – 1440 samples per row instead of 1920. Chroma is sampled at 480 samples per row, a third of the luma sampling rate. In the vertical dimension, both luma and chroma are sampled at the full HD sampling rate (1080 samples vertically).

Overall Sampling Lesson

Can't recover what you don't have.

Catch up with this.

Digital Bit Rates

Current television: 30 fps * 729 & 480 * 1.5 * 8 = 124 megabits per second

HDTV 9:16: 30 fps * 1920 * 1152 * 1.5 * 8 = 796 megabits per second. Here using 1152 instead of 1080; this is fudged. So actually 747 megabits per second.

This motivates compression.

The 1.5 comes from chrominance and luminance sampling (4:2:0), where chrominance is sampled at one-fourth the rate of luminance. X plus ¼ X plus ¼ X = 1.5 X.

Why drop frame video?

In NTSC, went to 29.97 fps because of color interleaving.

No interleaving in digital, so why do 59.94 in HD? Backward compatibility. In case running a B-roll of analog. Would get a vertical roll at the boundary when viewers watch; last thing we want.

Digital Video Line Timing

Video Standard	Total Samples	Active Samples	Blanked Samples
NTSC 525/60	858 (63.56 us)	720	138
PAL 625/50	864	720	144

Active region of digital video longer than that of analog video. Why? Can fit more in, but why? Because we can. Today's components are better, and we don't need the timing we were restricted to in the CRT days.

Overscanned portions of the line in the digital data allow for slight errors in any digital processing without losing any of the sampled active analog video.

Typical errors include sampling offsets, filter pre-load, or pipeline delay mismatch.

Since overscanned samples extend into the front and back porch regions of the analog video, they are known as front overscan (FOS) and back overscan (BOS).

Digital Overscan Amounts

601 specifies the number of samples from the EAV to the 50% point of horizontal sync tip. This positions the digital data with respect to the analog line and defines the number of FOS and BOS samples.

Video Standard	**Digital end to analog sync**	**FOS samples**	**BOS samples**
NTSC 525/60	16 (1.19 us)	4	6
PAL 625/50	12	8	10

Sampling and Serial Rates

This is in the studio. All of these are at standard definition, even the 16:9 formats.

Format	Parallel Clock Rate (MHz)	Serial Data Rate (Mbps)
4Fsc – NTSC	14.3	143
4Fsc – PAL	17.7	177
4:2:2 4:3	27.0	270
16:9	13.5	270
16:9	18	360

Can sample wide screen at 13.5 or 18, but 18 will give us the same resolution on a wide screen as we have on 4:3 on standard definition.

Why keep standard definition now that we can go HD?

Broadcasters can run more channels and get more income from more sponsors at the same time.

Ancillary Data

Placed between EAV and SAV data words.

In component digital domain, space for over 55 megabits of ancillary data. This is probably for the entire picture/frame.

Diagram of timing reference signal (TRS). HANC space is 288 words: 4 words for EAV, 4 words for SAV, leaving 280 words for data.

Luminance Quantizing

Name	Voltage Level (mV)	10-bit value in hex
Excluded	766.3	3FF
Peak	700.0	3AC
Black	0.0	040
Excluded	-51.1	000

Having EAV/SAV being all ones then all zeros differentiates from the case where video is too hot, where there would be multiple values of all ones or other consecutive high value reserved values.

Chrominance Quantizing

Name	Voltage Level mV	10-bit value in hex
Excluded	399.2	3FF
Max positive	350.0	3C0
Black	0.0	200
Max negative	-350.0	040
Excluded	-400.0	000

Words 1440, 1441, 1442, and 1443 used to transmit end of active line (EAV) TRS message. Words 1712, 1713, 1714, 1715 are used to transmit the start of active line (SAV) TRS message.

TRS is timing reference signal.

XYZ has 3 flags.

Flag	Values
F	0 if the line is in the odd field, 1 if in the even field.
V	0 if line is in active region, 1 if in VBI.
H	0for SAV symbols, 1 for all EAV symbols.

Video Timing Reference Codes

Why are 1 and 0 always zeros? If go to 8-bit, remove them. First word is always 3FF, second word is always 0, third word is always 0.

For 4:2:2 component video, active video has samples from 0 to 719. EAV is samples 720, 721. 722 through 855 are HANC, 856 and 857 are SAV. H = 0 for SAV symbols, H=1 for all EAV symbols. V = 0 if line in active region, 1 if in VBI.

Data Bit Number	Fourth Word (XYZ)
9 (MSB)	1
8	F
7	V
6	H
5	P3
4	P2
3	P1
2	P0
1	0
0	0