Composite Digital Video

Review of last class session.

Sampling frequency is 4 times subcarrier 4FSC = 14.318181 MHz. Sample video at 10 bits per channel per sample. Professional video is 10 bits. D format runs 8 bits since couldn't get the desired bandwidth at that time it was developed. D-2 is composite digital.

10 bits of data sampled at a rate of 14.31818 MHz in parallel. Use shift register to convert to serial. Required clock rate is 10 * 14.31818 MHz = 143.1818 MHz. Data accepted in parallel by shift register, data outputted in serial by shift register. Diagram 1.

Shift registers used a lot in computer. For example, serial port, keyboard line.

D-1 Format – SDI – 601

SDI = serial digital interface. Also called 601, which comes from the ITU specification. This is component digital video. This is what you will find in today's professional studios. SDI is not composite digital video. Composite format was a bridge to try out digital video, and the industry learned a lot about digitizing and transferring in the process.

Don't call this format D-1. Call it SDI or 601.

Diagram 2 of a field with horizontal sync.

No setup in component. Why waste 7.5 IRE when we are never going to put chrominance on luminance? No subcarrier or burst. Not adding color onto luminance. Diagram 3.

Correct a problem: manufacturers must create PAL and NTSC versions. Make equipment so they could talk PAL and NTSC on the same machine. Chose **13.5 MHz** as a sampling rate to be compatible with both. This is for component video, and just for the luminance part.

R-Y, B-Y components are separate from Y (luminance). Broadcast industry will never put these signals together again like it did with composite video.

Learned important thing in the 50's when came up with color TV. Compressing RGB which would be 4.2 MHz each. R-Y bandwidth is 1.5 MHz, a huge savings. B-Y bandwidth is 0.5 MHz, another huge savings.

Realized this, so when decided to do component digital, decided to use R-Y and B-Y to save bandwidth. If sampled separately, would require 3 * 13.5 MHz = 40.5 MHz bandwidth. R-Y and B-Y each sampled at 2.1 MHz. This is half of the 4.2 MHz bandwidth of NTSC video. Can sample R-Y and B-Y each at 6.75 MHz (that's half the sampling rate of Y). To get half the bandwidth, sample at half the rate. Makes it cheaper and easier to relate the sampling frequencies by integral multiples (i.e., Y sampling rate is 2 times that of R-Y and B-Y sampling rates). Can use a frequency divider to get both frequencies from one source. Total bandwidth is 27.0 MHz (adding the bandwidths together).

Must put through shift register in order to convert parallel data to serial. Since still 10 bits per sample, run clock at 270 megabits per second (MHz). Why so much more than composite? Work with NTSC and PAL formats, added R-Y and B-Y separately instead of digitizing the interleaved NTSC color video signal.

How do we handle this SDI data rate in the studio? What sort of cabling do we need? Fiber optics had not yet been invented. Use standard 75 ohm coax. More limitations now since 270 MHz is much higher than 4.2 MHz. Does well as long as the cable run is not too long. Higher frequencies tend to ride on outside of conductor due to skin effect, so material has a higher effective resistance. This means that losses are greater. Cable is still 75 ohms *impedance* even though higher *DC resistance* at this frequency.

When get to the point where the signal is so weak as to be unusable. Digital video is usable as distance increases until the signal is suddenly unusable. This is the **cliff effect**. Old, corroded connectors, crimps will reduce the maximum distance, as will too many tees and barrels.

Digitize Y signal (no burst, no setup), sample at 13.5 MHz. Sample R-Y and B-Y at 6.75 MHz. Add together to get 27 MHz total bandwidth. Multiply by 10 bits per sample to get 270 MHz.

Use 5.5 MHz bandwidth in digital video. 2 times this is 11 MHz (Nyquist limit). 13.5 MHz is greater than that.

No longer have to deal with cross-chrominance, cross-luminance in this digital format. In digital component video, color is not represented with a phase relationship, unlike in NTSC.

Y = .30R + .59G + .11B – this is still part of digital.

R-Y and B-Y all come from Y (the above formula).

P_R and P_B, also C_R and C_B – these both mean R-Y and B-Y.

Notation for Sampling

Sampling frequency for component is 13.5 MHz, for composite is 14.318181 (4FSC).

Our notation says 4:2:2

the first number (luma) is 4 (representing 4 times subcarrier) 13.5 is considered 4 for these purposes

the second number (2) is for sampling chrominance – represents both R-Y and B-Y since they will always be sampled at the same rate

the third digit (2) is either 1) the same as the second digit, or 2) it is zero, indicating that Cb and Cr are subsampled 2:1 vertically

4:2:2 is what we use in the studio. This is SDI aka 601.

4:1:1 is consumer-grade video

4:1:0

4:2:0

4:4:4 is pristine.

4:0:0 is black and white. Nobody does this. Even if there is a black and white sequence, people will be sampling at something higher.

There is better than 4:4:4 since could just sample at a higher rate, like 8:8:8. When Hollywood films something with digital video they will go 8:8:8.

Originally, the second digit specified the horizontal subsampling of C_B and the third digit specified the horizontal subsampling of C_R. That scheme failed to anticipate vertical subsampling, and in any event, all practical systems have the same subsampling ratios for both C_B and C_R.

Third digit now has two possibilities. If third digit is same as second digit, there is no vertical subsampling. If the third digit is zero, this indicates 2:1 vertical subsampling of both C_B and C_R.

If a fourth digit is present, it must be identical to the first digit, and indicates the presence of a fourth signal channel containing transparency (key, or alpha) information, sampled identically to luma (first digit). Must have same number of samples to ensure the keyed picture doesn't flicker. Rarely see a fourth digit.

Digital Television Standards

DV is based on CCIR 601, but is 4:1:1. Not easily edited; see shadows of colors during editing, see big gaps of color, since chrominance only sampled 1 time for every 4 luminance.

High-quality DVPRO and Digital-S is 4:2:2 chrominance subsampling. This is Panasonic's system of digitizing. This provides better editing.

MPEG-2 (DVD, HDTV) is 4:2:0 subsampling (1 pixel per 2x2 square).

H.261 – we don't use it – it is 4.2.0 and small format, used for teleconferencing (don't have to know it).

4:2:2

4:2:2 – see diagram 4.

For every 4 luminance samples, take 2 chrominance samples from odd lines and 2 from even lines.

Chrominance planes are just as tall and half as wide.

Diagram 5. Chrominance and luminance, luminance only, chrominance and luminance, luminance only.

This results in full vertical color for a column but only half horizontal color.

JPEG does this.

4:2:0

2 chrominance samples for every 4 luminance samples, odd lines only. Diagram 6.

This is more balanced for chrominance since the sampling is the same horizontally and vertically.

Chrominance is halved in both directions; this is the object of this standard.

4:1:1

One chrominance for every 4 luminance for both odd and even lines. Sample chrominance between the lines. See diagram 7. Average the chroma among the four pixels in a 2x2 square.

Hard to edit because we have to wait a long time (relatively) to get chrominance information.

There's a few different methods to do this. One way is to cheat and move the color sample spot into the first line but between the two pixels on left and right. Diagram 8. This is even cheaper since don't have to have a line store.

Overall Sampling Lesson

Can't recover what you don't have (if you didn't record it you have lost it forever).

Conversion between representations requires estimation of missing samples.

Interpolation causes errors:

- **spatially** – at the edges
- **temporally** – when moving

Spatial is one single picture, what happens on one single picture. Temporal is for more than one frame; between/among frames.

Common Digital Video Sizes

Name	*Geometry*	*Digital Standard*
CCIR-601	720x480	4:2:2, 4:2:0
SIF	360x240	4:2:0
CIF	360x288	4:2:0
4:3 HDTV	1440x1152	4:2:2, 4:2:0
9:16 HDTV	1920x1152	4:2:2, 4:2:0
4CIF, 16CIF, QCIF		

Steve will look into this since the 1152 doesn't seem to make sense.

Digital Bit Rates

How much to stream digital video without compression.

Current television: 30 fps * 720 (the 4 of the 4:3 aspect ratio, including sync (else would be 640)) * 480 (active lines) * 1.5 (luminance + compressed chrominance) * 8 (minimum number of bits) = 124 megabits per second.

9:16 HDTV: 30 fps * 1920 * 1152 * 1.5* 8 = 796 megabits per second.

This is a large bit rate. This motivates compression.

ITU-R BT.601.Rec.601

This is the full name of what we refer to as 601. This is the European name.

4:2:2 component digital video sampled at 13.5 MHz.

601 video.

SDI.

Since the 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). How to adjust for overscan? It is very subjective/qualitative and not quantitative. You just eyeball it to make sure there's a little bit.

In digital, it is more critical on how much overscan/underscan.

BT.601 specifies the number of samples from the end of the digital active video (**EAV**) to the 50% point of the horizontal sync tip. The samples are at 13.5 MHz. EAV = end of the active video.

This positions the digital data with respect to the analog line and defines the number of FOS and BOS

Digital Overscan Amounts

Standard	***Digital end to analog sync***	***FOS samples***	***BOS samples***
NTSC 525/60	16 (1.19 microseconds)	4 (0.30 microseconds)	5 (0.37 microseconds)
PAL 625/50	12	8	10

The active region of digital video is longer than that of analog video (by not alot).

The 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.

Great thing about D-1 format is that ituses the same 13.5 MHz sampling frequency for both 525- and 625-line television signals.

This does not mean that a signal that has been recorded on one standard may be played back on the other, but that the hardware is identical for both.

Today's monitors by law are required to lock up to any format.

These samples are taken at a frequency of 13.5 MHz for the Y channel and 6.75 for R-Y and B-Y. When multiplexed the resultant sample frequency is 27 MHz.

In the 4:2:2 sampling standard, the blanking intervals are not sampled. Instead, end of active video (**EAV**) and start of active video (**SAV**) data words are inserted in the data stream as markers.

Have EAV, front porch, sync pulse, back porch, SAV. Diagram 9. We throw all of this stuff away. Why sample stuff we know is always going to be the same? 10.9 microseconds out of 63.5 microseconds. As opposed to composite video, where encoded everything in the video signal.

We still have this space, so what do we do with it? Embedded audio, embedded data,ancillary data. Includes audio, closed captioning, SAP, program information. **Ancillary data** is placed between the EAV and SAV data words.

In fact, in the component digital domain, there is spare space for over 55 megabits of ancillary data, enough room to carry digital audio, time code, error detection, and handling information with plenty of room to spare for future use.

Sampling and Serial Rates

Format	*Parallel Clock Rate*	*Serial Data Rate*
4FSC – NTSC	14.3 MHz	143 Mbps
4FSC – PAL	17.7 MHz	177 Mbps
4:2:2	27.0 MHz	270 Mbps
16:9	13.5 MHz	270 Mbps
some 16:9	18 MHz	360 Mbps

Why use the 18 MHz option? 16:9 not being treated fairly for its definition as compared with 4:3. If have a longer scan but use the same frequency, must sacrifice definition. Boosting parallel clock rate is a way to make 16:9 look like 4:3 as far as detail goes.

Should have this stuff in your notes.

FILL ME IN

Coaxial cable between serial transmitter and serial receiver.

At the transmit end, the parallel data is shift registered using a 10 times parallel rate clock. The resulting 1 bit wide serial data is passed through a data scrambler. A known algorithm is applied to ensure that there are a sufficient number of transitions to recover the clock signal at the receiving end and to limit any DC content on the signal. We ran into this in our time code readers Can't have a run of the same number (a long string of 1s or 0s) without the receiving end getting out of sync. Clocks drift, so we need transitions regularly to reset the clock in the receiver.

The data is then passed through a non return to zero inverted (NRZI) encoder, which takes a transition in the signal and changes it to a 0 and takes a non transition in the signal and changes it to a 1. Uses lots of XOR gates. Creates a lot of changes, so the receiving clock can sync to it.

This further limits any DC content in the signal. The signal is then sent down the coax.

Block diagram D10.

Asynchronous is sending signal without a clock. We did it in our time coder by using a phase locked loop.

The square wave that we created by transitioning between ones and zeroes is affected by the losses in the cable and is attenuated (via skin effect) and distorted. In cables, especially high frequencies (transitions in square wave), wave reflects and causes distortion. Take note that the higher the data rate, the shorter the cable length the signal can pass through.

The 270 MHz (component digital) signals will not pass through as much cable as a signal at 143 Mbits/sec.

The signal starts with a peak-to-peak voltage of 800 mV +/- 10%. At the end of a 220-m (600 ft) cable run using a 270 Mb/sec signal, the peak-to-peak level will be attenuated down to only about 30 mV (a 143-Mbit/sec signal after 300 m is also 30 mV). This is a drop of 28.5 dB (20 log (V1/V2)). This is a big loss. Not only will the amplitude be reduced, but also the square shaped pulses will have become more like a sine wave. Lose high frequencies easier than low frequencies: skin effect, plus reflections.

Period of 270 MHz is 1/frequency. 3.7 nanoseconds. Our scopes can handle 20 microseconds time base so how are we going to read this?

Will continue this Tuesday.

Homework

Read 12.19 to 12.34 in textbook. Be prepared for a **quiz on Tuesday** on what we read. Skip the math/calculations, just read for content.