

Metropolitan Police Video Training

Phil Crawley @IsItBroke on Twitter

Video Measurement - Principals (4 days)

Course Aim

The aim of the course is to provide grounding in video fundamentals, compression and picture quality analysis for staff who work with video processing, detection and vision systems.

Course Structure

- Day 1 TV Fundamentals; Scanning and Sampling, Colour Systems, Analogue Composite Coding
- Day 2 TV Fundamentals; Digital Component Coding, Conversion of Film to Television, High Definition
- Day 3 Compression; DCT principles, Intra-Frame vs Inter-Frame Encoding, blocks and macroblocks etc.
- Video Tape Recording; Magnetic recording principles, rotary recording, Simple VT maintenance.
- Day 4 Television Measurements; Analogue, Digital
- Picture Quality Analysis; Analogue picture impairments, digital picture impairments, compression
- TV Displays; CRT, LCD and Plasma displays. Problems with LCD and Plasma, Projection systems.

www.root6.com



History - what did video look like?



- Until the 1970s video was entirely analogue and didn't have pixels.
- Everything from the monochrome video level to the colour content and all of the synchronising information was encoded onto an analogue signal.
- Cameras, videotape, telecine, colour-correction, and captioning was all done in the analogue domain with not a hint of digital imaging.

This all changed in the mid 70's with three requirements; Timebase correction, synchronisation and standards conversion.

www.root6.com



Timebase Correction

- 1" and ¾" Umatic VTRs have an inherently unstable timebase the stability of the video signal is insufficient to mix with camera sources etc.
- The only way to stabilise and lock to the station reference is to turn the analogue off-tape signal into a digital representation, write it into a store and read it out with the super-stable station genlock.
- In the early seventies being able to store eight lines of video required a large unit that typically sat underneath the VTR.





Synchronisation

- In a television studio all the cameras and VTRs are locked to a common reference which allows those sources to be seamlessly mixed with each other.
- In a studio centre this is easy to achieve with an outside broadcast contribution or a camera in a helicopter not so much!
- Montreal 1976 first Olympics that had live coverage from helicopters as well as other remote cameras. This was achieved with early-model synchronisers from Quantel – the DFS 3000.
- A frame-store synchroniser operates much like a TBC but it has a whole video-frame (625 lines, a 25th of a second) of storage.





www.root6.com



Standards Conversion

Different territories around the world use different standards for their television;

- PAL common in Europe, 625 lines per frame, 25 interlaced frames per second
- NTSC Common in the Americas, 525 lines per frame, 30 interlaced frames per second

It turns out you need eight frames of storage to do good quality standards conversion.

Digital Video Effects, Painting systems, Slow Motion machines.

In subsequent years all of these devices started to be used in television production and post but it wasn't until the introduction of D1 VTRs in the late eighties that it became common place to interconnect equipment digitally (using the seminal rec-601 system) rather than via their analogue i/o.

Until the 601 standard different manufacturer's equipment operated internally at whatever raster the designer had landed on.

www.root6.com



Digital Video - but isn't video an analogue signal?

There is a clear advantage in representing video now as an analogue signal travelling down a wore, rather as a digital bitstream;

- It is can be stored in computer-based editing and special-effect systems
- It can be copied, manipulated and sent long distances without the degradation of analogue video
- It can be compressed for delivery over cable, the air, or network delivery

Analogue to Digital conversion

The resolution of the converter indicates the number of discrete values it can produce over the range of analogue values. The values are usually stored electronically in binary form, so the resolution is usually expressed in bits. In consequence, the number of discrete values available, or "levels", is a power of two. For example, an ADC with a resolution of 8 bits can encode an analogue input to one in 256 different levels, since $2^8 = 256$.

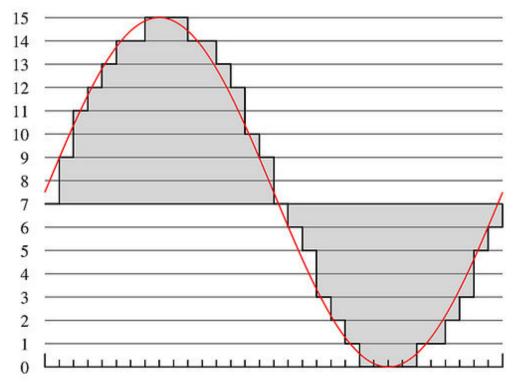
- In broadcast television digital video started out at 8-bit
- Now we use 10-bit for television and 12 or more bits for Film work

www.root6.com

rete

technology for the creative

Digital Video - sampling structure



Two things have to be considered;

- Bit depth; 10-bits gives 1024 descrete levels which seems to be enough 'dynamic range' to faithfully capture video
- Sampling frequency; The Nyquist theory says you need to measure more than twice as often as the highest frequency being dealt with.
- 5.5Mhz standard definition video would therefore need to be sampled at more than 11Mhz
- SD video is actually sampled at 13.5Mhz
- Because we have to deal with the luminance channel AND the chrominance channel we have a sampling clock of 27Mhz
- Because the system is ten bit the SDi data rate is 270Mbits/sec

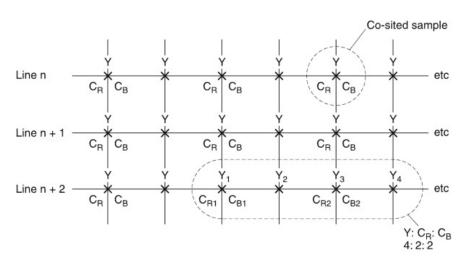
www.root6.com

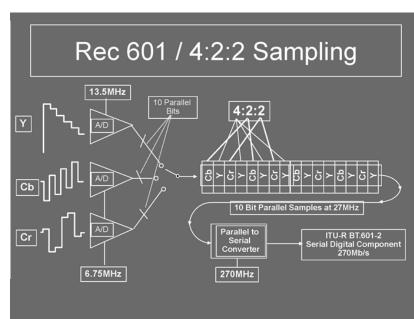
+44 (0) 20 7437 6052

7



Digital Video - sampling structure





- For every four luminance samples we have two chroma
- Alternate luma samples are **co-sited** with Cr and Cb
- The other luma samples are on their own
- This is why we refer to this style of digital video as 4:2:2

www.root6.com



CCIR rec 601 - Standard Definition Digital Video

Originally the 1982 standard defined;

- 4 x 3 aspect ratio
- 720 pixels x 576 lines enough pixels for 5.5Mhz video
- Y Cb Cr luminance/colour encoding at 4:2:2 data rate half res colour difference.

There are several things worth noting;

- 4 x 3 display with 720 x 576 gives **non-square pixels** (almost square, but not quite)
- When 16 x 9 came along pixels got very non-square, same 720 x 576 resolution.
- Colour space & sampling structure unlike graphics formats

Remember – at this point Photoshop was still pre v.1 and 601 served the needs of TV images.

www.root6.com



SDi – a standard that has grown with time

Whereas CCIR Rec 601 defines the sampling structure of standard definition digital video the SDi standard allows multiple data rates;

Standard	Name	Bitrates	Example Video Formats		
SMPTE 259M	SD-SDI	270 Mbit/s, 360 Mbit/s, 143 Mbit/s, and 177 Mbit/s	480i, 576i		
SMPTE 344M	ED-SDI	540 Mbit/s	480p, 576p		
SMPTE 292M	HD-SDI	1.485 Gbit/s, and 1.485/1.001 Gbit/s	720p, 1080i		
SMPTE 372M	Dual Link HD-SDI	2.970 Gbit/s, and 2.970/1.001 Gbit/s	1080p		
SMPTE 424M	3G-SDI	2.970 Gbit/s, and 2.970/1.001 Gbit/s	1080p		

image: Wikipedia

We'll talk a bit more about the HD variants later on.



Original digital video cabling

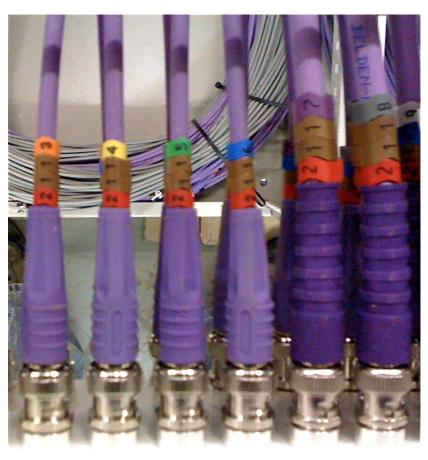


- 1988-vintage Sony DVR-1000 D1 VTR
- Digital video carried on a 25-pin connector
- Signals could travel a few metres down very expensive cables
- Not really a scalable format for studios/post/transmission
- All eight or ten bits are carried on separate conductors
- This is why it was referred to as *parallel digital video*



Digital video - the current state of things

- Digital video is now carried over the same style of coax cable with BNC connectors
- SD video can use the skinny coax
- HD video needs the fatter coax for any distance
- Purple is the common colour for digital video!
- All eight or ten bits of data are fed down the same cable
- This is referred to as serial digital video
- Other things can be added to the video
- When four, eight or twelve channels of digital audio are included then it is referred to as serial digital interface; SDi



www.root6.com



Monitoring in SDi



- With analogue video we're looking at the state of the signal coming down the cable; voltage = video level
- With SDi the quality of the bits on the cable don't represent the video, they must be decoded by a digital TV test set.
- This display from the Tektronix WVR-series shows the numbers recovered and what the waveform would look like.



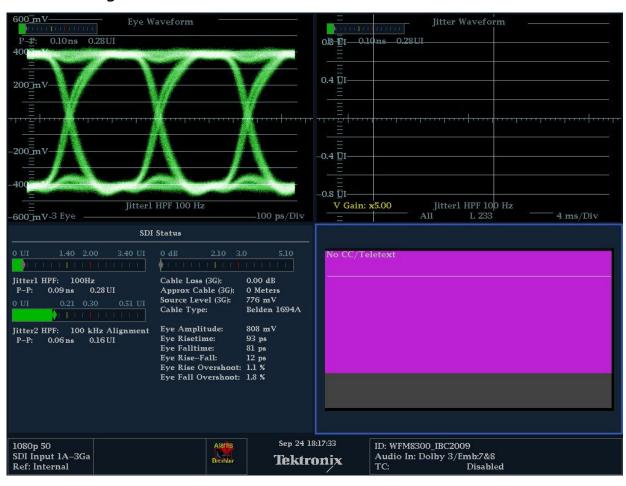
Monitoring in SDi - timing



- Unlike analogue video where the timing pulses are real voltage changes, in SDi
- No timing pulses are stored, only active picture
- There are Timing Recovery Signal data packets (TRS is the abbreviation)
- Again, the test set must decode the digital data to identify the timing data.



SDI data signal & cable transmission characteristics



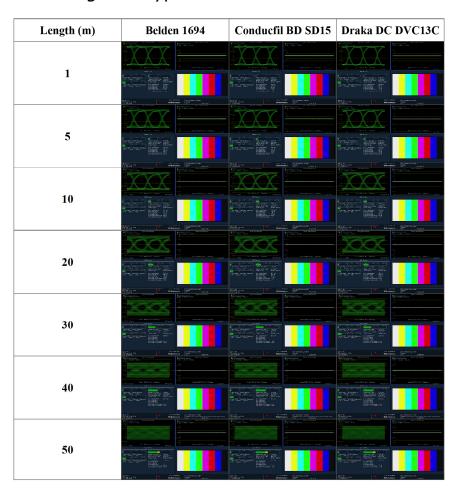
Physical layer measurements

- Eye Pattern
- Jitter measurements
- Cable loss

15



Cable lengths v type



This shows how the signal quality degrades over distance for three different cable types

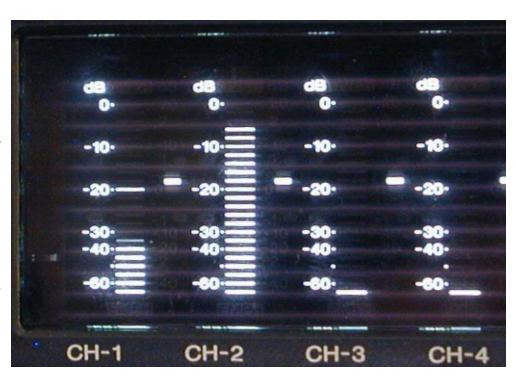
- This is the highest-res 3G video
- Standard def 270mBit/sec video goes much further over good quality coax
- Notice how bad the eye-pattern is at 50m but the machine is still recovering the video

www.root6.com



Embedded Audio.

Both the HD and SD serial interfaces provide for 16 channels of embedded audio. The two interfaces use different audio encapsulation methods - SD uses the SMPTE 272M standard, whereas HD uses the SMPTE 299M standard. In either case, an SDI signal may contain up to audio channels sixteen (8 pairs) embedded 48 kHz, 24-bit audio channels along with the video. Typically, 48 kHz, 24-bit (20-bit in SD, but extendable to 24 bit) PCM audio is stored, in a manner directly compatible with the AES3 digital audio interface. These are placed in the (horizontal) blanking periods, when the SDI signal carries nothing useful.



Digibeta VTR audio meters

www.root6.com





The Tektronix WVR & WFM-series rasteriser test-sets include full audio monitoring including multi-channel and Dolby encoding (data) over audio.

This display shows a 5.1 (six channel) audio with a derived stereo mix (where channels seven and eight would be in a Dolby E stream).

As well as the bar-graphs (like the audio meters we saw before) there is a phase diagram display that plots two channels against each other and allows you to see how 'wide' a stereo feed is or if it's a mono feed (or if it's out of phase with itself).



Dolby E and advanced audio parameters

Multiple-channel audio (5.1 and now 7.1 – six and eight channels respectively) is becoming more common and may be delivered on discrete channels (of a DA88 or videotape) or, more commonly, encoded using the Dolby-E system onto two channels of a VTR running in 'data mode'.



The following are important;

- Loudness the broadcasters are getting very hot on perceived loudness as a common complaint is that 'the commercials are louder'. The emerging standard is that loudness is measured on a similar scale to PPM for audio level – this is what the Tektronix measures, and in this display shows it over the duration of the show.
- Dialnorm a Dolby parameter than gets encoded with the audio – the Tek shows the presence and value.
- **Dynamic Range** Another Dolby parameter that gets carried with the data stream.

www.root6.com



Conversion of Film to Television

Telecine; A machine for transferring images on film (both positive prints and negative camera rushes) to video for recording on videotape. The Telecine machine (AKA "TK") has to deal with the following;

- Wrong framerate film is 24fps, TV is 25 or 29.97fps
- Interlace film is progressive, TV is interlaced
- Resolution film captures much more detail than even HD television can
- Dynamic range the difference between the smallest difference in film and TV means the TK has to deal with more subtlety than video can handle.



www.root6.com



Telecine - framerates; film -> PAL TV

- You can shoot film at 25fps commonly used in high-end TV production
- You can run the TK 4% fast so that the 24fps film is now running at a TV-friendly 25fps
 - This can be problematic as the 120 min film now lasts 115 mins and the soundtrack is pitch-shifted up by nearly a semitone
 - If capturing into Avid this is called "Method 1" the Avid makes allowances
- You can digitally 'make-up' an extra frame on average every second
 - o This can lead to an uncomfortable cadence to fast camera pans etc.
 - Avid refer to this as "Method 2" with the advantage of being able to keep the same digital audio captured on set without having to re-sample, 1 sec time = 1 sec!

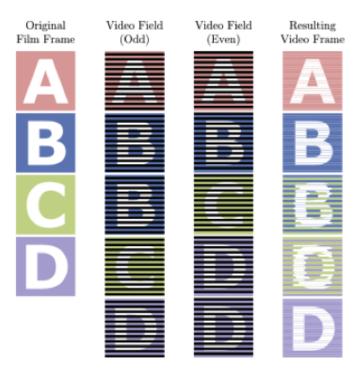
Interlaced – in the past TK machines had to go to great lengths to make interlaced video using twin-lenses or jump-scan CRT patches. Nowadays the film data is written into a digital store and read-out in two passes to turn the film frame into two TV fields.

www.root6.com



Telecine - framerates; film -> NTSC TV

- There is no film framerate that gets near NTSC's 29.97fps!
- Imagine for a moment that we need to turn 24fps into 30fps; we use the 3:2 pulldown
- However! NTSC is 29.97 so the 24fps film needs to be run at 23.976fps for the pulldown to work
- The tiny speed-change from 24 to 23.976fps doesn't cause any trouble with the sound.
- NTSC video that has come from film has burnt-in trouble; some frames have cuts between fields 1 and 2 which makes subsequent editing hard.
- If that interlaced video is then de-interlaced for progressive use some frames will appear as flash-frames on playback.
- Software can undo the 3:2 pulldown.



www.root6.com



Telecine - dynamic range

Since film has a much wider dynamic range than video – 15 f-stops (equiv. to 90dBs) verses video's 12 f-stops (70'ish dBs); a "colourist" is needed to adjust those pictures with bigger latitude to fit into the smaller space that video offers. At this point in the process the richer colour and dynamic range means that this is the best time to do colour correction and artistic grading.

Telecine – legacy for digital film production

Because most digitally shot films will either end up on TV (shown on air, DVD, or BluRay) OR will be edited on a TV-type editing machine (Avid, Final Cut etc) AND digital audio needs to be maintained then modern digital film cameras run at 23.976fps (just like old telecines when transferring to NTSC).

In 2012 this is the legacy of the choice of colour subcarrier frequency for NTSC back in 1950!



Telecine - framerates; film -> NTSC TV, Drop Frame Timecode

NTSC video is 29.97fps but timecode runs true (i.e. 30 frames/sec, no fractional frames) – this means that real-time and timecode would run around 3.5 seconds wrong every hour.

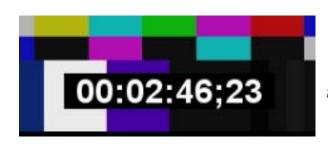
To correct this, drop frame SMPTE timecode was invented. What's actually being dropped are some of the timecode "labels".

In order to make an hour of timecode match an hour on the clock, DFTC drops;

- frame numbers 0 and 1 of the first second of every minute,
- except when the number of minutes is divisible by ten

That is, drop frame TC drops 2 frames every minute, except every tenth minute, achieving $30 \times 0.999 = 29.97$ frame/s.

The same is applied to digital film cameras that shoot 24fps (they really shoot 23.976fps)



the semicolon indicates DFTC

www.root6.com



Digital Intermediate

DI is a phrase that people started to use in the early noughties and refers to

- Scanning / Telecine as the input to the process rather than a creative part
- Maintaining a film in a digital form for all parts of the post production
- Effects, grading, mastering all in the file-domain

All this takes film from being a very physical/chemical business to being something a lot like television production.

Intermediate originally referred to the fact that film was still shot and delivered on celluloid;

- High Resolution Telecine to scan the rushes (at either 2k or 4k resolution)
- Shared storage so that grading/editing/effects/mastering can access the rushes
- Film laser recorder to write the finished film back to a distribution master

Now it is usual for a film to be shot on a digital camera and delivered on a 500gig drive as a DCI master (with the *volume key* on a USB stick).

www.root6.com



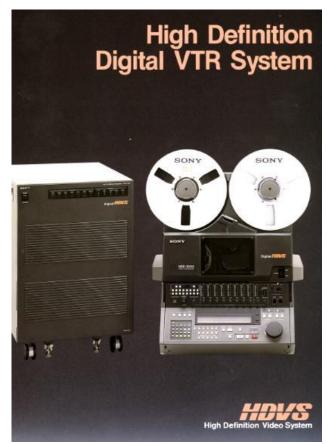
High Definition

There have been efforts to push up the resolution of TV since the 1970s

- 1984 Sony HDVS 100 camera and VTR 1" analogue RGB 1250-line
- 1988 Sony HDD-1000 1" Digital uncompressed VTR
- 1997 Sony HDCam system ½" cassette compressed 1440x1080 7:1
- 2003 Sony HDCamSR mildly compressed RGB 1920x1080, some models 1080P



Picture taken in Television Centre, Studio 1 c.1989



www.root6.com

+44 (0) 20 7437 6052



High Definition

There are currently many formats!

- Sony HDCam ½" 4:2:2 videotape
 - HDCamSR ½" 4:4:4 videotape
 - EXCam Solid state SxS memory cards
 - XDCam Proprietary BluRay disks see next two slides for details!
 - HDV 8mm videotape
- Panasonic P2 4xSD cards
 - D5 ½" videotape
 - DVCProHD 8mm videotape
- Various Digital SLRs shooting on Compact Flash etc Canon 5D mk.2
- RedONE digital film camera
- · Alexa digital film camera
- D21 digital film camera these all shoot to various data storage devices.

www.root6.com

XDCAM formats

Format name	Container	Video coding	Bit depth	Color sampling	Frame size	Frame rate and scanning type	Video bit rate, Mbit/s	Audio coding	
DVCAM	MXF,	DV		4:2:0	720x576	25i, 25p	OF (ODD)	PCM 4 ch/16 bit/48 kHz	
	DV-AVI	DV	8	4:1:1	720x480	29.97i, 29.97p	25 (CBR)		
MPEG IMX	MXF	MPEG-2 422P@ML	8	4:2:2	720x576	25i, 25p	30 (CBR), 40 (CBR), 50(CBR)	PCM 8 ch/16 bit/48 kHz, or 4 ch/24 bit/48 kHz	
	IVIAF				720x480	29.97i, 29.97p, 23.98p	30 (CBH), 40 (CBH), 50(CBH)		
MPEG HD	MXF, MP4	MPEG-2 MP@H14/HL	. 8	4:2:0	1920x1080	29.97i, 25i, 29.97p, 25p, 23.98p	35 (VBR)	PCM 4 ch/16 bit/48 kHz	
					1440x1080	29.97i, 29.97p, 23.98p (pulldown), 25p	18 (VBR), 25 (CBR), 35 (VBR)		
					1280x720	59.94p, 29.97p, 23.98p, 50p, 25p	25 (CBR), 35 (VBR), 19 (CBR) ^[1]		
MPEG HD422	MXF	MPEG-2 422P@HL	8	4:2:2	1920x1080	29.97i, 25i, 29.97p, 25p, 23.98p	50 (CDD)	PCM 8 ch/16 bit/48 kHz, or 4 ch/24 bit/48 kHz	
					1280x720	59.94p, 50p, 23.98p (pulldown)	50 (CBR)		
Proxy AV	?	MPEG-4 Part-2 (ASP)	8	?	CIF (50i - 352x288)	?	1.5	A-Law 4 ch/8 bit/8 kHz	

image: Wikipedia

- various codecs (DV, MPEG2, MPEG4)
- data rates (18Mbits⁻¹ to 50Mbits⁻¹)
- resolution/colour sampling (352x288 -> 1920x1080)
- wrappers (MXF, AVI, Quicktime)

As used in the Sony XDCam product line.

www.root6.com

+44 (0) 20 7437 6052

28





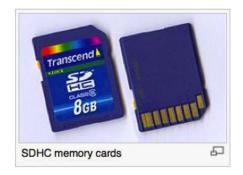
image: Wikipedia

Just for illustration;



- Blue Ray recordable disk
- SxS Memory cards
- SD memory cards

Recording formats as used in the Sony XDCam product line.



www.root6.com



High definition Sampling Standards (ITU 709)

First proposed in 1990(!) rec.709 part 2 defines what has become our HD TV standard;

- 1920x1080, 1280x720 720 has nearly disappeared
- 24, 25, 30, 50 or 60fps interlaced, progressive or PsF (more later!)
- New colour primaries defined
- Standardise on 6500 for white point (NTSC uses 5500k remember)
- 8 or 10 bit 'studio swing' levels (i.e. 16 235 for luminance)
- 16x9 aspect
- Square pixels

Quite a lot has changed since rec.601 (1981) which defines all the parameters for SD video.



Rec 709 - New colour primaries defined

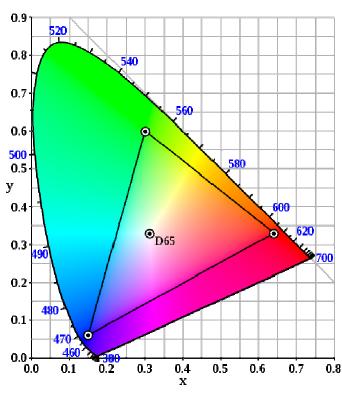
Newer display technologies can display a greater range ('gamut').

Color enges	White	Primaries						
Color space	xw	yw	XR	УR	XG	yg	XB	ув
ITU-R BT.709	0.3127	0.3290	0.64	0.33	0.30	0.60	0.15	0.06

The upshot of all this is that the RGB -> YUV transform has now changed;

$$Y = 0.213R + 0.715G + 0.072B$$

 $Cb = 0.539(B-Y) + 350mV$
 $Cr = 0.635(R-Y) + 350mV$



Digital devices that can receive an SD/HD-SDi bitstream have to be able to switch in the appropriate matrix. If that isn't the case then you'd notice a green cast on pictures if you switched between standards (going from HD to SD) or a magenta error going the other way. In the case of a monitor you'd have to re-calibrate the white point to D65.



Rec 709 - 8 / 10 bit levels

- Like 601 HD television allows for eight or ten bit levels; HDCam is 8 bit, everything else is 10.
- Only really an issue when importing graphics or video sourced from non-TV equipment
- Shows itself by crushing the blacks or bleaching-out the whites.

Import Settings - Current Image OMFI/AAF Shot Log Audio Image Size Adjustment Image sized for current format Crop/Pad for DV scan line difference Do not resize smaller images Resize image to fit format raster File Pixel to Video Mapping Alpha Channel Computer RGB (0-255) Invert on import (white = opaque) Computer RGB, dither image colors Do not invert (black = opaque) 601 SD or 709 HD (16-235) Ignore Frame Import Duration Dilate Fill 80 seconds Duration: Autodetect Sequentially-Numbered Files OK Cancel

Avid import tool

www.root6.com



Rec 709, aspect ratio and square pixels

- 709 only defines 16x9 widescreen displays
- 1920x1080 on a 16x9 monitor = square pixels
- In SD pixels are never square!

This shows the problem of multiple aspect ratios in television. In the US they didn't really have much 16x9 SD television and so their aspect ratio conversion issues are really only between SD and HD.

In Europe we had SD 16x9 since the mid-90s.



16:9 to 4:3 Letterbox

- Full 16:9 source width mapped to 4:3 monitor width
- Full source height compressed to 75% of monitor height
- Horizontal black bands added to top and bottom
- No picture lost



16:9 to 14:9 Letterbox

- Central 87.5% of 16:9 source width stretched to 4:3 monitor width
- Full source height compressed to 87.5% of monitor height
- Horizontal black bands added to top and bottom
- Minimal picture lost



16:9 to 4:3 Full Screen

- Central 75% of 16:9 source width stretched to 4:3 monitor width
- Full source height mapped to monitor height
- No horizontal black bands required
- Significant picture lost



4:3 to 16:9 Pillarbox

- Full 4:3 source width compressed to central 75% of 16:9 monitor width
- Full source height mapped to monitor height
- Vertical black bands added to left and right
- No picture lost



4:3 to 14:9 Pillarbox

- Full 4:3 source width compressed to central 87.5% of 16:9 monitor width
- Central 87.5% of source height stretched to monitor height
- Vertical black bands added to left and right
- Minimal picture lost



4:3 to 16:9 Full Screen

- Full 4:3 source width mapped to 16:9 monitor width
- Central 75% of source height stretched to monitor height
- No vertical black bands required
- Significant picture lost

+44 (0) 20 7437 6052

www.root6.com



Tri-Syncs in High Definition Television.

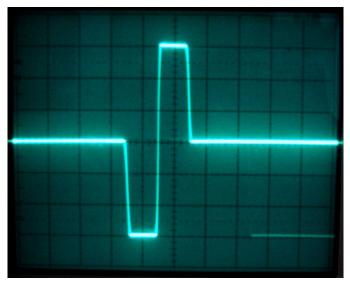
- Becoming more significant in this HD-world
- Unlike HD-SDi video tri-syncs can be distributed much like normal black & burst
- Same cables, jackfields, and even distribution amplifiers and routers can be used
- Equipment that uses tri-syncs will often need reference input switched so small composite routers and multi-o/p SPGs are a good idea.



www.root6.com



Tri-Syncs in High Definition Television.



The horizontal timing reference point for a standard bilevel sync signal is defined as the 50% point of the falling edge of the horizontal sync pulse. Looking for the sync pulse has always been one of the "trickiest" of tasks for the display signal processor. It requires careful biasing of the sync processing circuitry so that the sync pulse is made as distinguishable as possible from the other voltage levels within the video signal. In order to ascertain this point precisely, it is necessary to clamp both black-level and sync-tip level and derive a value half way between. If the signal is in any way distorted, this will give H timing inaccuracy. With a relatively pedestrian

standard definition signal (line rate in the province of 15 Khz) this isn't such a big deal, but with 1080 lines per frame this becomes more significant and with a tri-level sync signal, the timing reference point is the rising edge of the sync signal as it passes through the black-level. This point is much easier to determine accurately, and can be implemented relatively easily. It is also more immune to signal distortion.

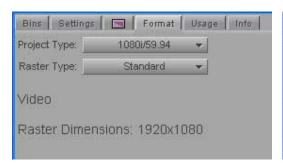


Full Raster / Thin Raster

Several format have lower-resolution in the horizontal direction than rec-709 allows

- Sony HDV
- Sony HDCam

They both use 1440x1080 (so non-square pixels). Other 'thin raster' systems use 960x1280 Capturing these into Avid or FinalCut and then mixing with 'full raster' material on the same timeline can be problematic.





So even if the timeline is 1080i/59.94 (in this example) you can capture thin-raster material which Avid converts on import.

www.root6.com



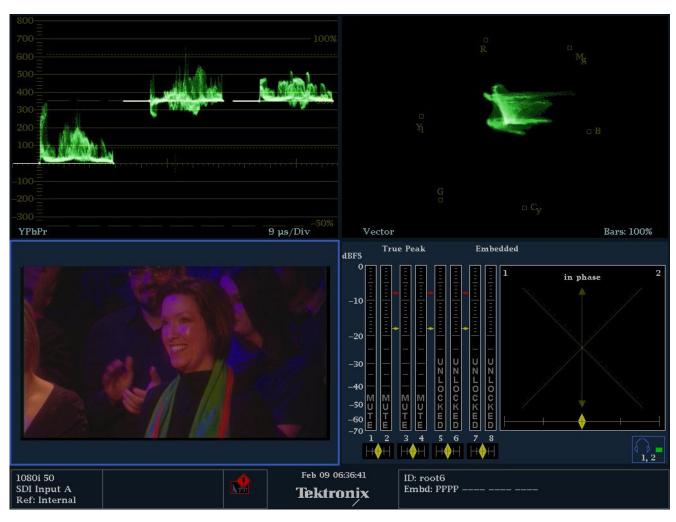
High Def 25Psf video

In the 1990s Sony realised that progressive video was the future but existing HD equipment (typically the BVM-D series monitors) couldn't lock to such a slow framerate (24/25/29.97 as opposed to 48/50/59.98 fields). The answer for progressively-sourced pictures was the Psf standard which makes progressive frames look like interlaced video.

- There is no difference between a Psf signal and an interlaced signal from a technical standpoint
- Sending 1080 pictures via Psf doesn't degrade them in the slightest in fact if you're laying off 1080 to HDCamSR it is recording Psf!

This shows the output from a Symphony NitrisDX BOB. The footage had come from a Sony EX3 cameras recorded at 35mBits 1080/25P onto Memory Stick and imported straight into a progressive timeline. The Avid plays back Psf which the Tek shows as 1080i (for the reasons discussed above). Laying this off to HDCamSR (a 5500 deck) gives a 25Psf recording on tape.





www.root6.com

+44 (0) 20 7437 6052



Show & tell with the Tektronix 7120 HD television test set



www.root6.com