Phil Crawley

15th February 2011, 16:30

101 Digital media: the size, the shape, and the frame rates

Since the 1970s digital imaging has been slowly making its way into television and film. Early digital video standards were not ideally suited to computer graphics and it's only in the last few years that those problems have been resolved.

- History – what did video look like?
- Computer Graphics Formats – MPEG and beyond
- Editing vs Transmission codecs
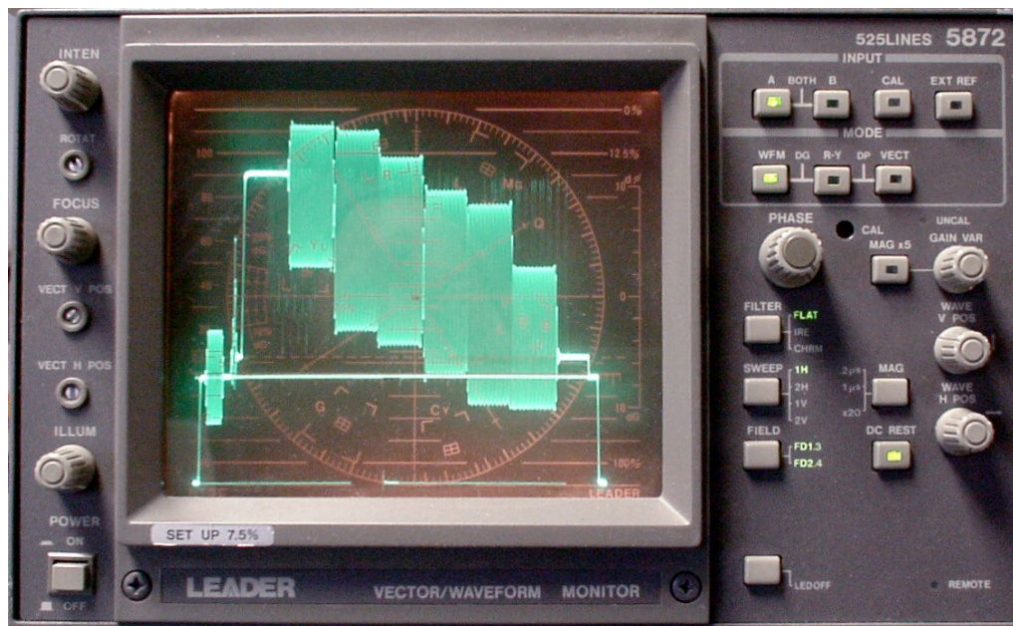- Contemporary MPEG4 codecs
- Frame rates – why 23.976 fps?
- Q&A

This is by its nature an introduction – you can't cover in a few hours what university training takes months over but it will give you a confidence in the basics to start investigating for yourself.

www.root6.com   +44 (0) 20 7437 6052

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.

2

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

www.root6.com   +44 (0) 20 7437 6052

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.

www.root6.com   +44 (0) 20 7437 6052

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.

Montréal 1976

www.root6.com   +44 (0) 20 7437 6052

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   +44 (0) 20 7437 6052

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   +44 (0) 20 7437 6052

Uncompressed video and codecs

- The data rate of uncompressed standard def video is 270Mbits^{-1}

- High def comes in at 1.48Gbits^{-1} and 3Gbits^{-1}

- These data rates are far too high to record on videotape or send over a network

- Using mathematical techniques the digital data that represents pixels – colour and luminance values – are transformed into a description that allows the pixels to be re-constituted and hence occupies much less space

- Depending on the application video can be compressed to 10% or less of its original size.

- The particular mathematical function used to achieve this is called a codec

- Different codecs have pros and cons depending on application (shoot, edit, TX etc)

www.root6.com   +44 (0) 20 7437 6052

Early computer-based graphics and video formats

With the exception of MPEG most computer video formats tend to;

- Have square pixels (because computer monitors do)

- Use RGB for their colour representation

- Use varying framerates (from 12fps up)

None of these lend those early computer video systems to television!

image: Wikipedia

The same can be said of computer still image formats – TIFF, Targa, BMP, etc – also;

- They may use CMYK colour space

- Graphics software may work in DPI rather than absolute resolutions

- Varying degrees and quality of anti-aliasing

How on Earth has any of this been reconciled?!

www.root6.com   +44 (0) 20 7437 6052

MPEG – Motion Picture Expert Group

- **MPEG-1**, 1993 – The basis for CDi and VideoCD (early DVD predecessors) – only 1.5Mbit/s and a quarter-screen resolution of 350x288 pixels. No interlaced video, 4x3 and stereo audio only. Long GOP

- **MPEG-2**, 1995 – Full standard def video, basis of DVD and DVB-T, C & S. Variable data rate and full 601 resolution using Y, Cr, Cb colour sampling. 16x9 or 4x3 video, up to 5.1 audio via AC3. 23.976, 25, and 29.97 FPS. Long GOP or i-frame (TX vs editing).

- **MPEG-4**, 1998 –The basis of most modern video encoding for acquisition, post and delivery. Multi-resolution, multi-framerate, multi-audio standards, editing or TX variants.

Domestic delivery -> TX -> Editing

www.root6.com   +44 (0) 20 7437 6052

Editing vs Transmission codecs

MPEG2 and all MPEG4 variants are based on a mathematical model of video called the **Discrete Cosine Transform** (DCT). Once applied to the video data this function then allows the codec to reduce the data rate of the video stream – you can transmit it and store more of it on disk.

P references preceeding I

B references preceeding I and succeeding P
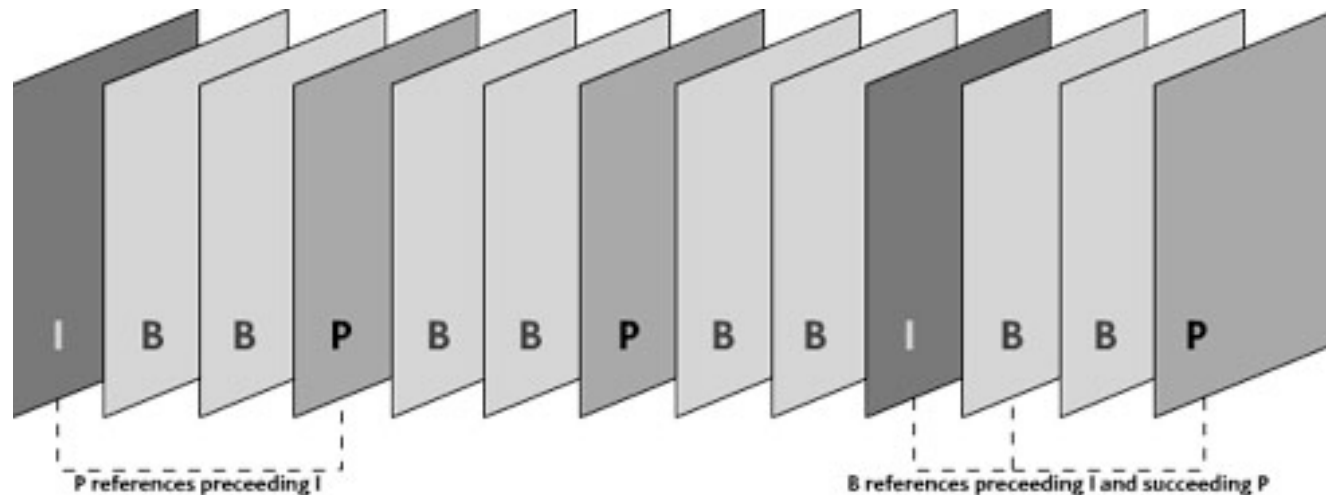
Once the data in the video frames has had the DCT function applied the codec can also define different types of video frames that go to make up the **Group of Pictures** (GOP).

www.root6.com   +44 (0) 20 7437 6052

Editing vs Transmission codecs cont.

- **I-frame:** An intra-frame, or I-frame, is a video frame which has been encoded without any reference to any other frame. A video file will always start with an I-frame and will have subsequent I-frames added at regular intervals. I-frames are also known as key-frames and are important for random access of video files such as rewind, fast-forward and seek operations. The downside to an I-frame is that they are the largest in terms of size as the whole video frame is encoded every time.

- **P-frame:** A predictive inter-frame, or P-frame uses previous I or P-frames as a reference when encoding. This means a P-frame will analyze a previous I or P-frame for any static elements which do not change between frames. Any areas which do not change are not encoded therefore a P-frame only stores video which registers movement making them much smaller than I-frames. The downside to P-frames is that they are sensitive to transmission errors because of their dependency on earlier frames.

- **B-frame:** A bi-predictive inter frame, or B-frame makes reference to both a preceding reference frame as well as a future reference frame. Using B-frames improves the prediction and ultimately the quality of decoded video but it also increases the processing requirements and latency.

www.root6.com   +44 (0) 20 7437 6052

Editing vs Transmission codecs cont.

The requirements of editing and transmission differ somewhat;

- Editing requires immediate access to each video frame and should not have to build a complete frame by looking at the frames that surround it, so an **I-Frame** only (or 'short-GOP' codec) is used.

- Transmission would rather leverage the additional image quality available to a long-GOP system and so a 12-frame (typical) GOP is used.

Rule of five;

- Uncompressed standard definition video ~ 250Mbits/sec

- I-Frame editing codec, MPEG2 ~ 50Mbits/sec

- Long GOP transmission codec, MPEG2 ~ 10Mbits/sec

- Statistical Multiplexed DVB stream to the home ~ 2Mbits/sec

www.root6.com   +44 (0) 20 7437 6052

Contemporary standards – HD and MPEG4

HD video makes things considerably easier – the Rec 709 standard defines;

- 1920x1080 resolution at 16x9 now has square pixels!

- Wider variety of frame rates (including 24 PsF)

MPEG4 is the basis for most acquisition, editing and transmission codecs currently.

- MPEG4 improves the performance of MPEG2 by using variable sized macroblocks.

- MPEG4-part 10 (aka 'H.264' or AVC) further improves performance by allowing macroblocks to the referenced across I-Frame boundaries.

www.root6.com   +44 (0) 20 7437 6052

Qualitative comparison of MPEG2, MPEG4 & H.264

Video reams at the same resolution, encoded three times.

www.root6.com   +44 (0) 20 7437 6052
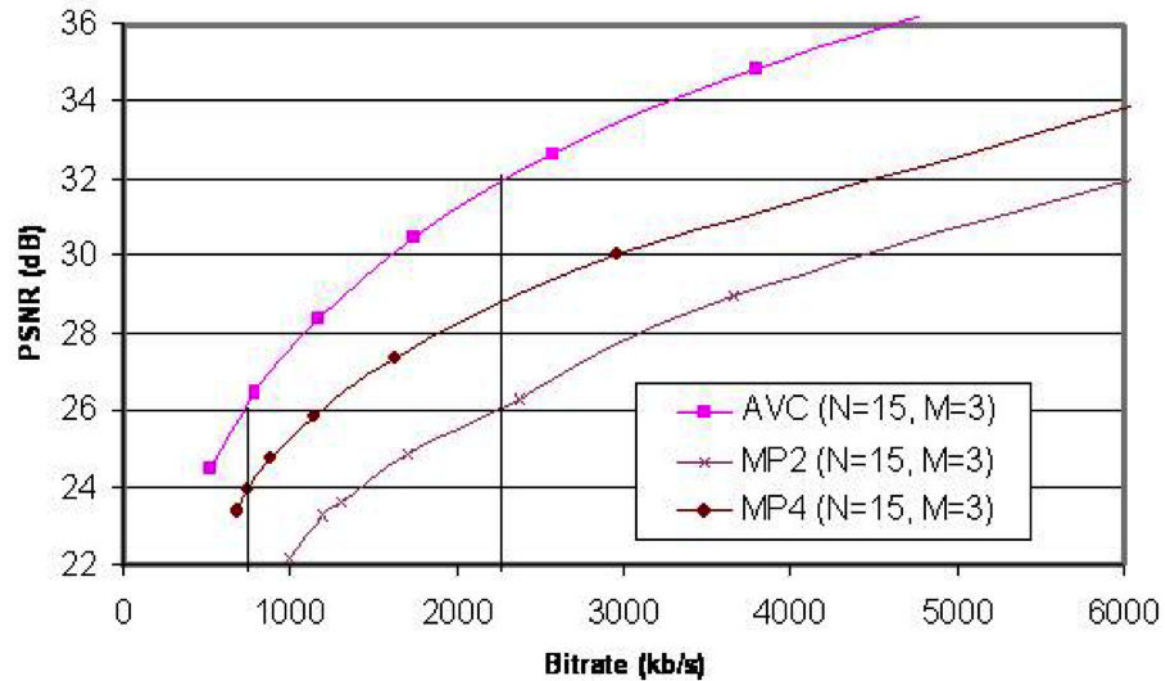
Contemporary editing & acquisition formats

- Avid DNX HD – I-frame only (i.e. editing) **codec**.

- Quicktime – Apple's wrapper format, especially **ProRes codec**.

- MXF –a 'universal' wrapper format that can encapsulate different codecs

- MPEG2 – the **codec** used in HDV, XDCam

- DV – I-Frame only **codec**, initially domestic cameras but extends up to HD (100Mbits^{-1})

You have to distinguish between *codecs* (that mathematical function that changes raw pixels into a file-description of how to re-create the pixels; compression) and *container* formats (AKA wrapper).

www.root6.com   +44 (0) 20 7437 6052

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-AVI	DV	8	4:2:0	720x576	25i, 25p	25 (CBR)	PCM 4 ch/16 bit/48 kHz
				4:1:1	720x480	29.97i, 29.97p		
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
					720x480	29.97i, 29.97p, 23.98p		
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 (CBR)	PCM 8 ch/16 bit/48 kHz, or 4 ch/24 bit/48 kHz
					1280x720	59.94p, 50p, 23.98p (pulldown)		
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.

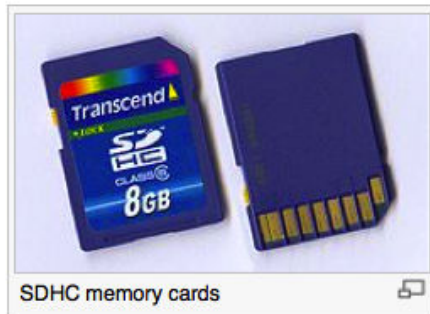www.root6.com   +44 (0) 20 7437 6052

image: Wikipedia

Professional Disc

Just for illustration;

- Blue Ray recordable disk

- SxS Memory cards
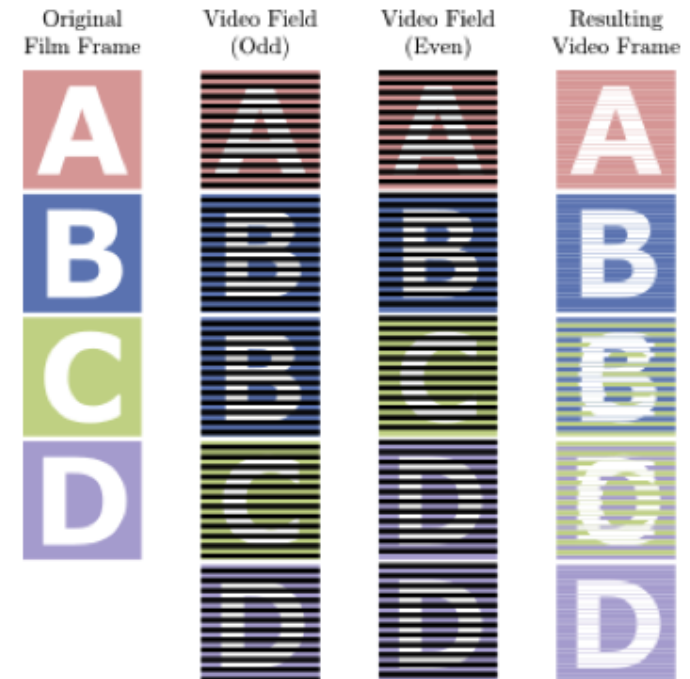
- SD memory cards

SxS memory card

Recording formats as used in the Sony XDCam product line.

All of these – codec, media, edit system govern how usable your edit workflow will be.

SDHC memory cards

www.root6.com   +44 (0) 20 7437 6052

23.976 frames per second – what's that all about?!

- NTSC video is actually 29.97fps rather than 30fps

- When transferring 24fps film to video there are two methods

- For PAL (25fps) the film is just played 4% fast

- For NTSC the 3:2 pulldown is used.

- Since the aim is to produce 29.97 video (rather than 30fps) an NTSC telecine has to run internally at 23.976fps rather than 24fps

- So if material is to be shot 'filmic' but will end up on DVD or television in NTSC mode then 23.976 is a must.

- This limitation remains to this day with 23.976 being the preferred mastering format for HDCamSR.
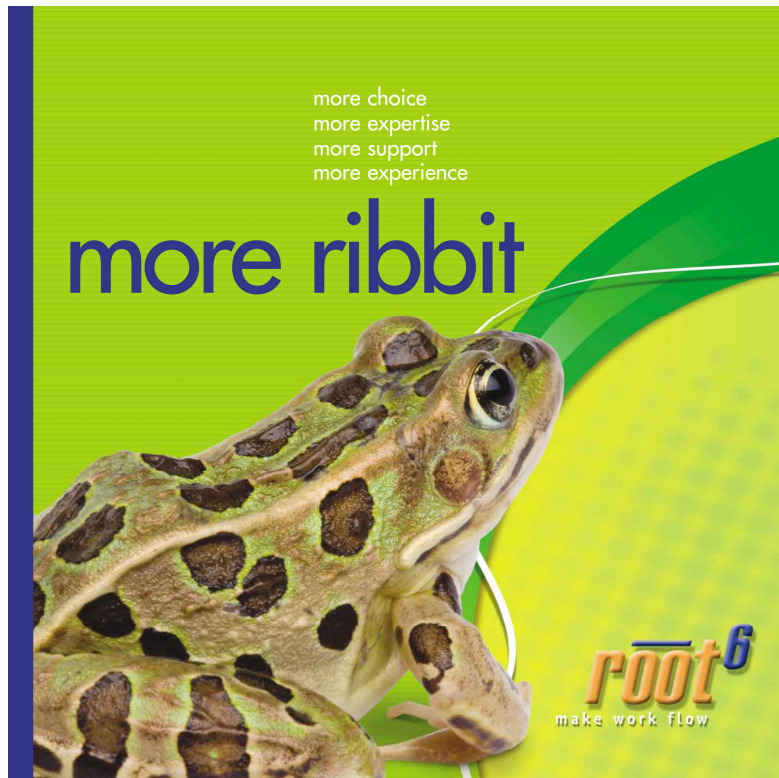

Original Film Frame | Video Field (Odd) | Video Field (Even) | Resulting Video Frame

www.root6.com   +44 (0) 20 7437 6052

In Conclusion

Things were a lot simpler in the days of SD!

- Fixed frame size

- Only two frame rates – PAL 25fps & NTSC 29.97fps

- Videotape – several formats but all usable in an SDi environment

- Uncompressed or only a handful of compression codecs.

However, with careful consideration of shooting format & media, edit workflow and delivery spec a file-based workflow will bring immense benefits. Always best to ask someone who has used your workflow already – there are countless examples of people who have assumed "..it'll just work" (like it would ten years ago) and discovered to their expense that it's more involved.

www.root6.com   +44 (0) 20 7437 6052

root6

technology for the creative

more choice
more expertise
more support
more experience

more ribbit

root6

make work flow

Please see www.root6.com for details of all of root6's training packages.

If you get **your badge scanned** at the root6 booth then you will be eligible for a discount when attending the full version of this taster session.

TCP/IP for Broadcast Engineers

Audio 101

Video 101

Television QC101

http://www.root6.com/blog
download today's notes

www.root6.com   +44 (0) 20 7437 6052