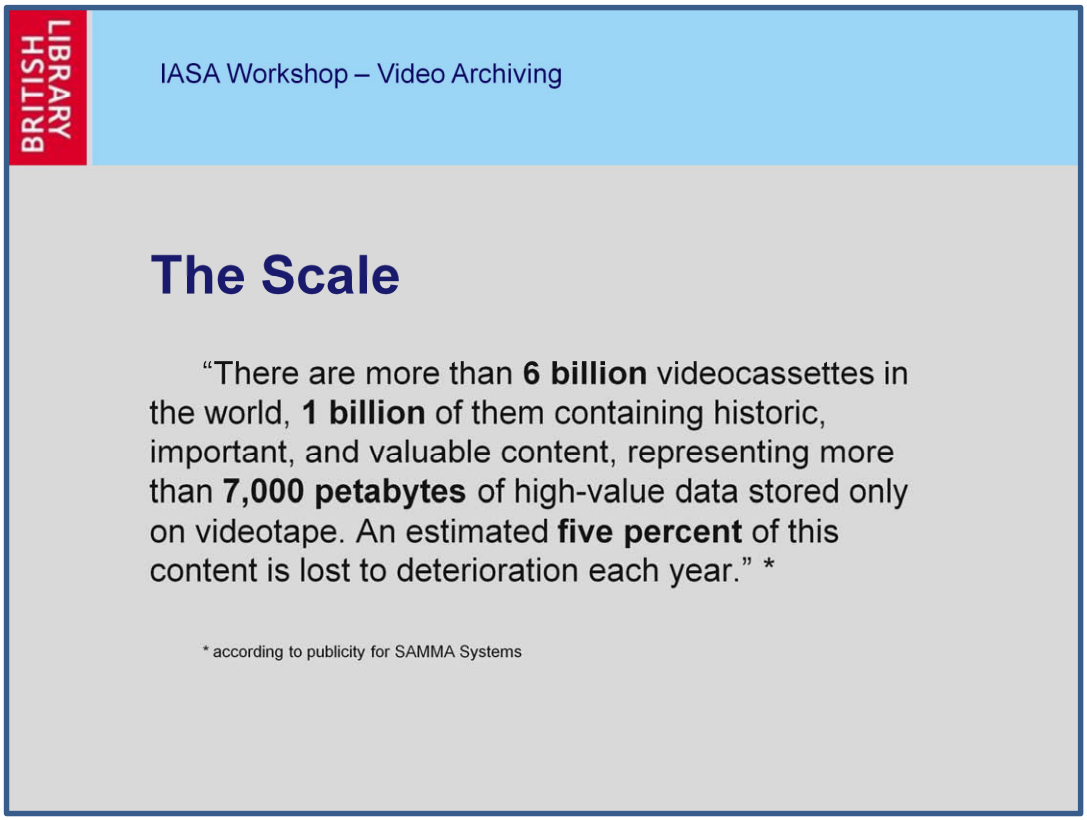
Video Archiving: One rule or many?

An introduction to technologies and formats

Andrew Pearson, British Library
Dominique Théron, BnF



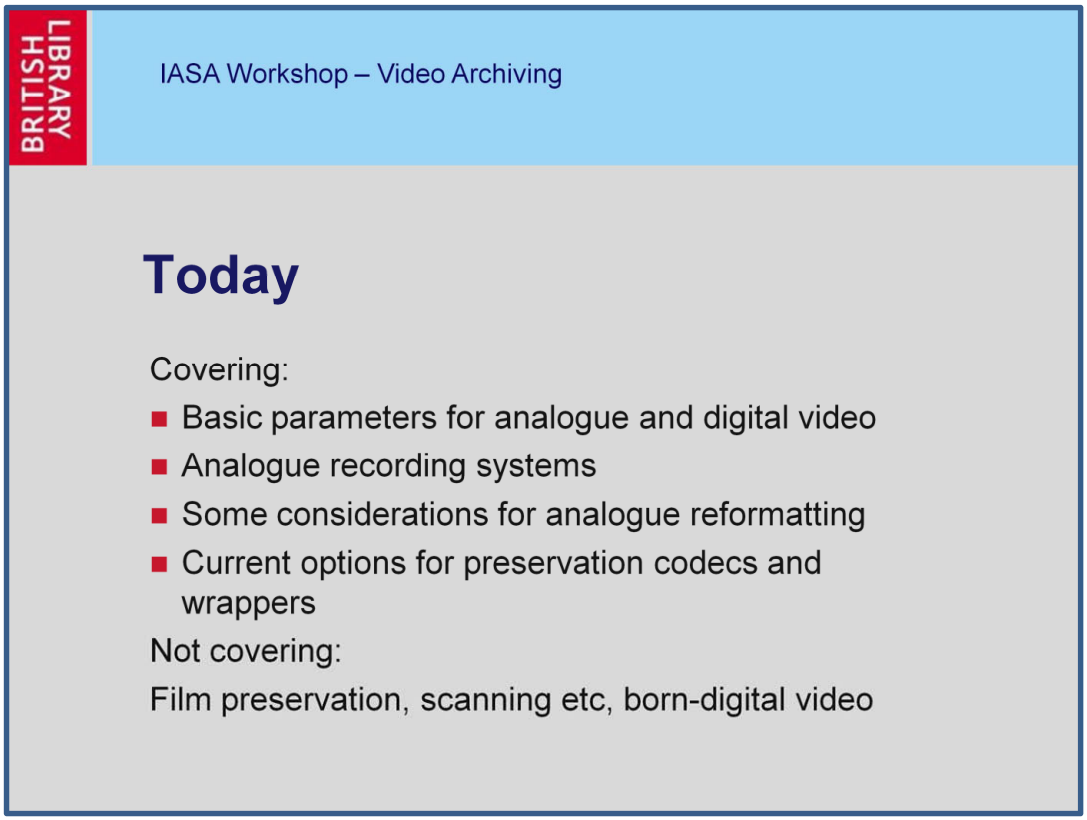
We're going to be dealing here mainly with reformatting of tape media

Figures may be based on worldwide sales of blank media, may or may not include rental video market.

Issues…

- Many recording tape formats
- Several worldwide TV standards, and new ones*
- Tape doesn't last forever
- Playback machines are becoming scarce
- Archival standards are still emerging*
- Digital video takes up large amounts of storage space*

* differs from the situation for audio



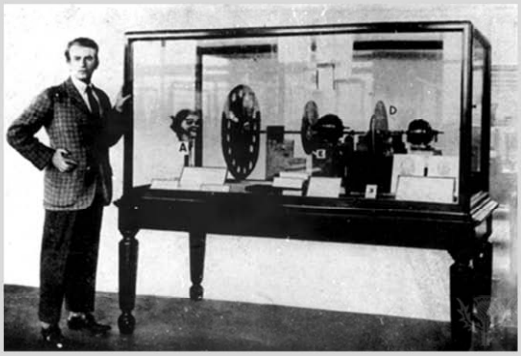
Not – born-digital, Film scanning

Introduction

Baird's mechanical TV

http://www.bairdtelevision.com/firstdemo.html

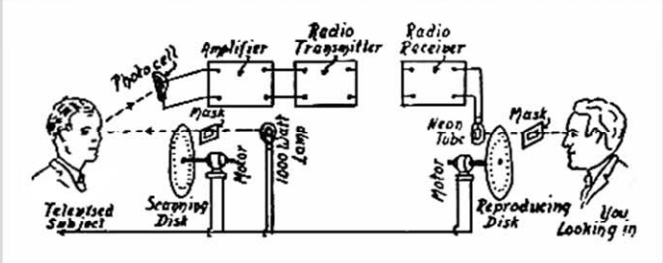
J L Baird and his Televisor



Science Museum/National Media Museum

Analogue sampling

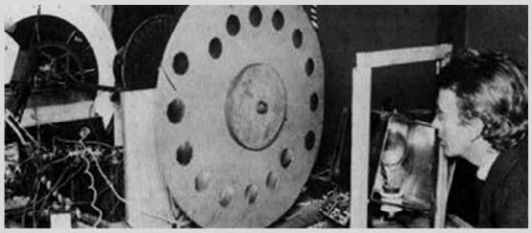
- Height
- Width
- Time



Baird's prototype

- Scanning wheel (Nipkow disk -1885) caters for height and width sampling
- Speed of rotation governs the number of pictures per second

Baird being 'televised'

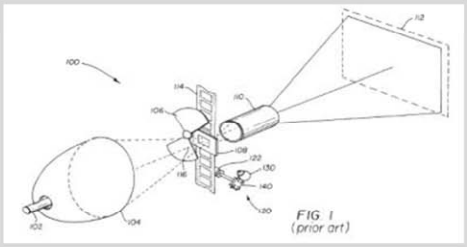


Hulton Getty

The time illusion

If there are more than about 16 images per second, the brain ‘sees’ a continuous stream of movement

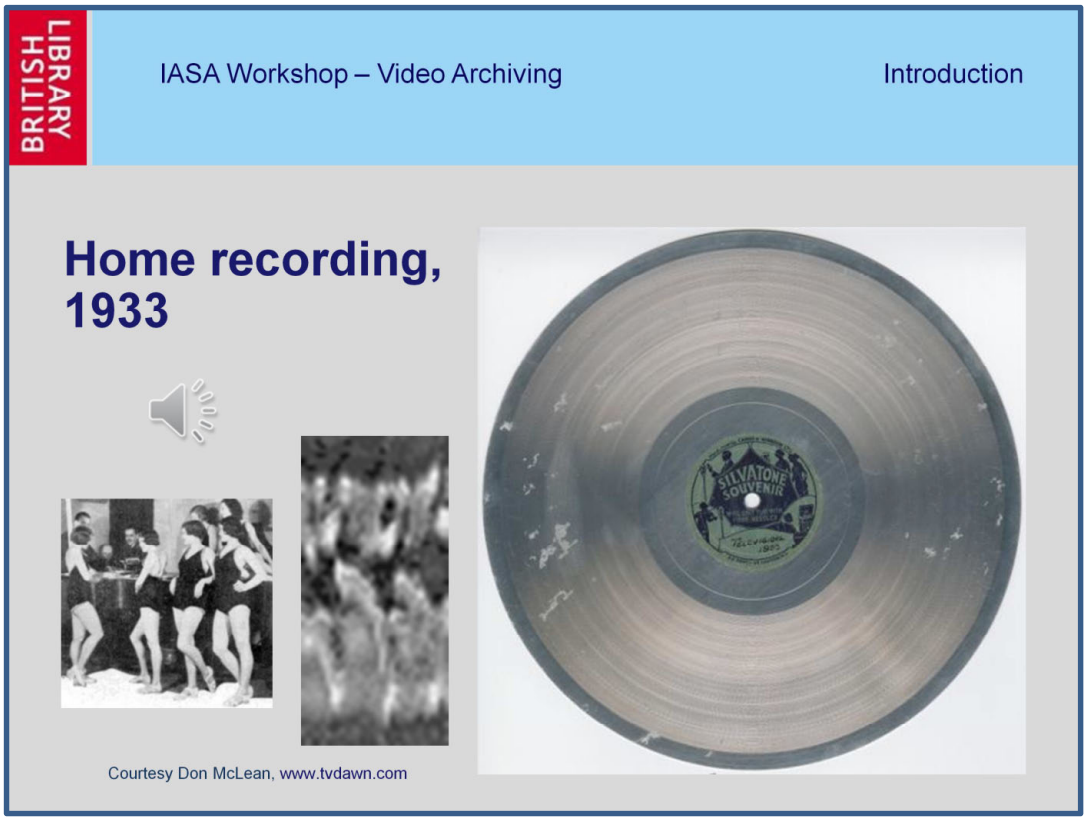
Cinema projector



Experimental mechanical recordings

- Baird
  1927-8

- Private recordings
  1932-5



Silvatone home recording apparatus – recorded TV from radio, somewhere in Ealing, west London.

Ealiest known broadcast TV recording, of earliest known TV Revue. 11:10pm on 21st April 1933.

BRITISH LIBRARY

Video systems

Pixellated image

12 x 9 grid

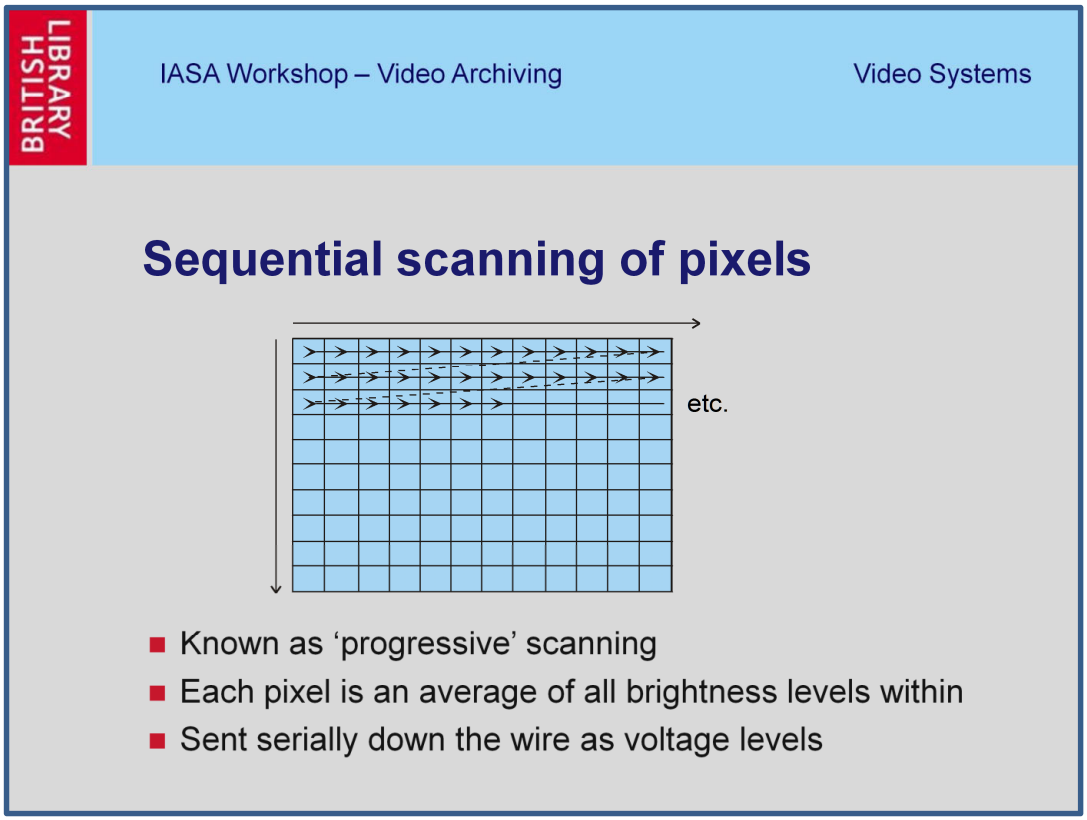
60 x 45 grid

768 x 576 pixels

- 'Standard definition' TV – 4:3 aspect ratio

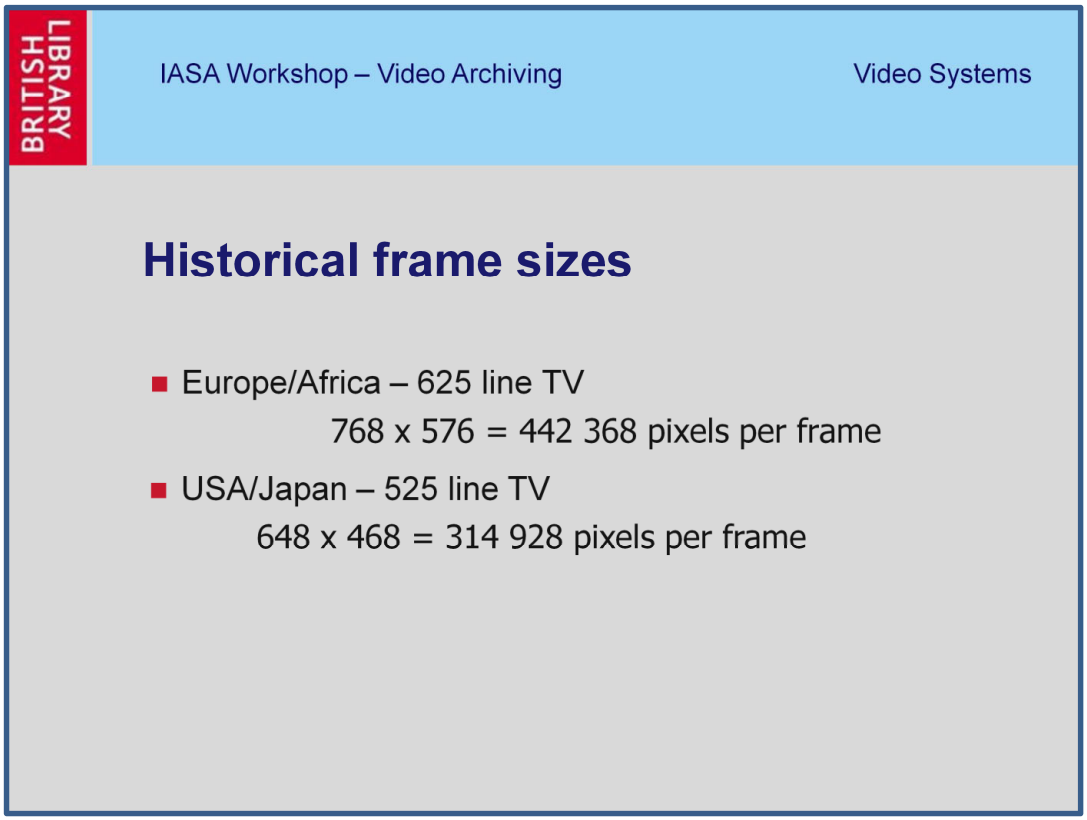
768 x 576 indistinguishable from HD from the back of the room!

Pixels = picture cells or picture elements

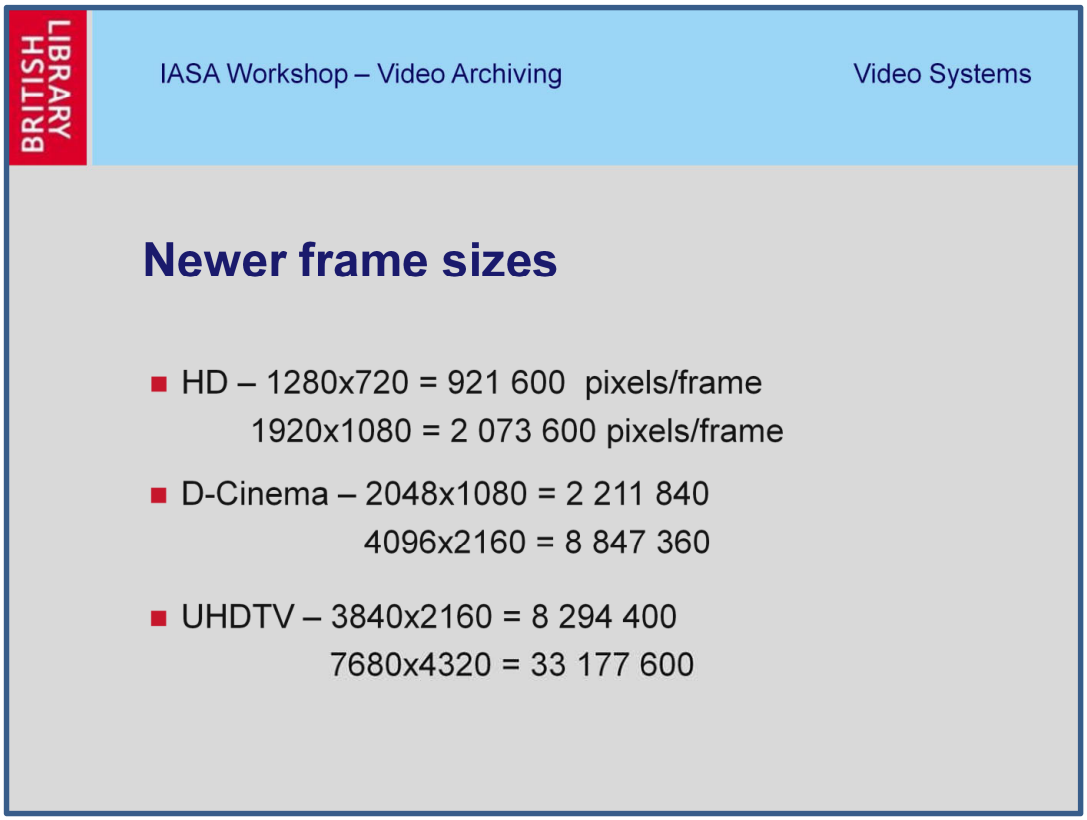


All systems scan in a sequence from top left to bottom right, though not always quite like this.

Note that Rows of pixels are discrete and defined from each other, but in an analogue system, horizontal pixels do not actually exist – it's a continuous 'stream' signal, similar to analogue audio.

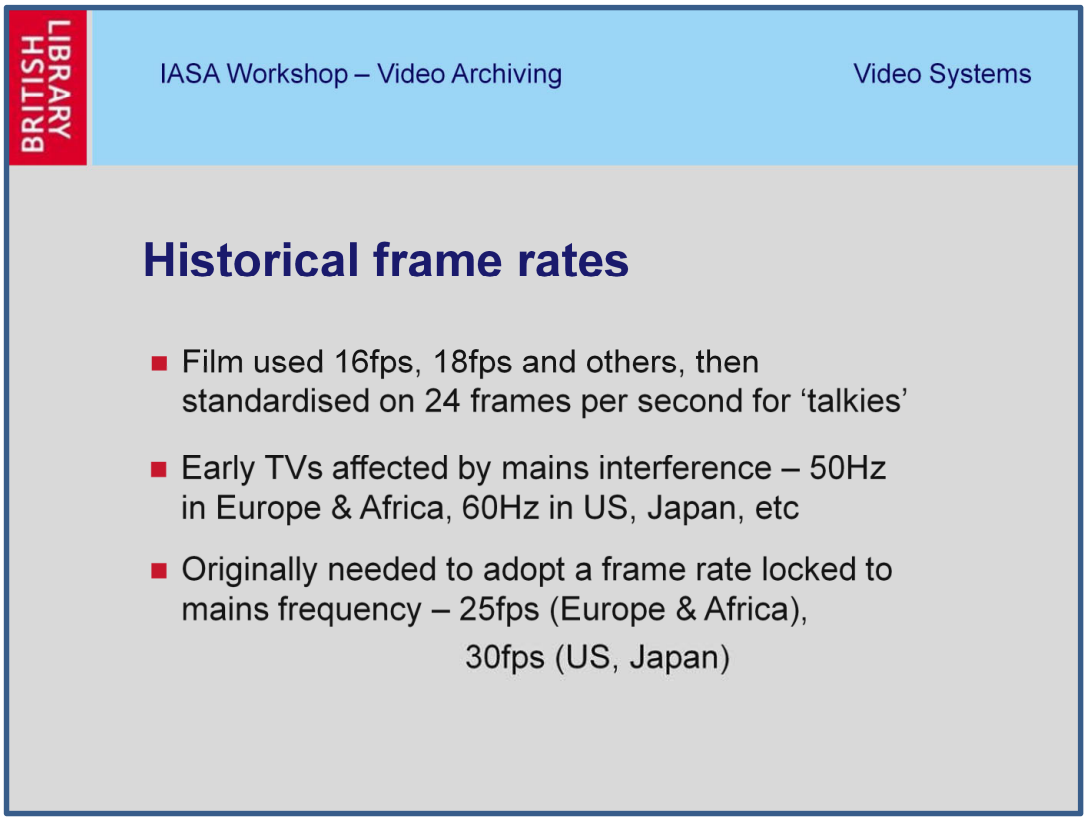


768 figure not quite right for digital video, but more later

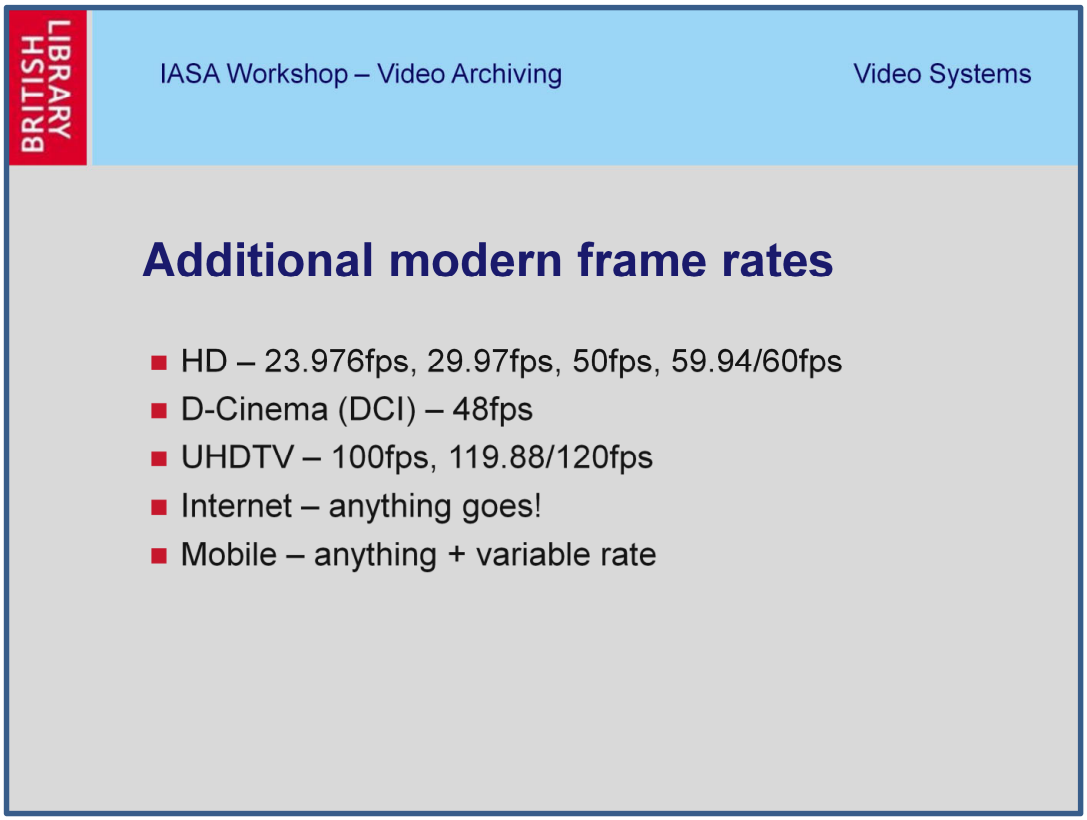


Mentioned at start – newer systems coming along all the time.

Always need to preserve frame size as is.



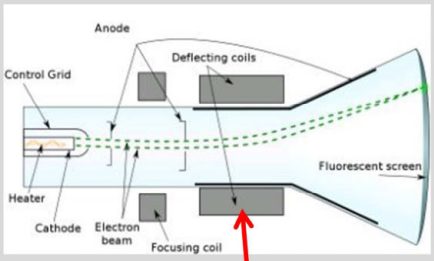
Mains interference = 'hum' on audio, light and dark bands across image on video



This slide includes rates in previous frame rate slide.

Always need to preserve frame rate as is.

Display



Wikipedia Commons

Scanning coils

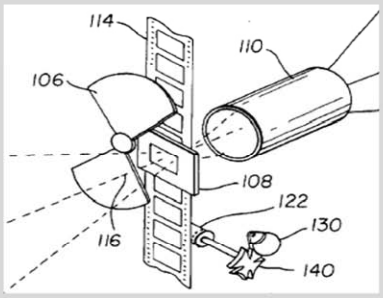
50cm Telefunken Cathode Ray Tube, 1938

http://members.chello.nl/~h.dijkstra19/page3.html - The Cathode Ray Tube Site

Until recently, displays were cathode ray tubes, with very particular characteristics

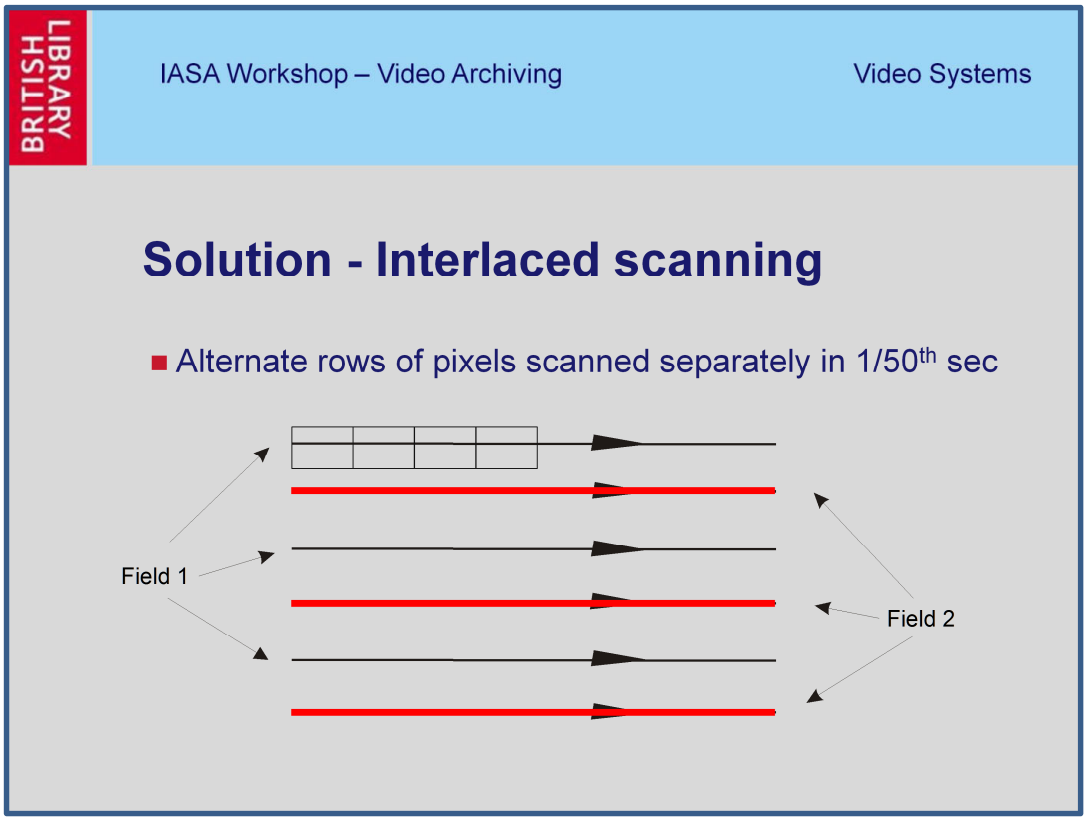
Problem!

- Original CRT display couldn't hold onto image
- 'Persistence of vision' means that the retina can only hold an image if it flashes faster than 40 times per second – 24, 25 or 30 fps produces annoying flicker
- Solved in cinemas by rotating shutter – each frame shown twice



Try taking a still picture off a CRT!

See here for example and details: http://www.drhdmi.eu/dictionary/refresh-rate.html



Lots of problems for modern systems:

Temporal order of fields is important, but differs in different systems.

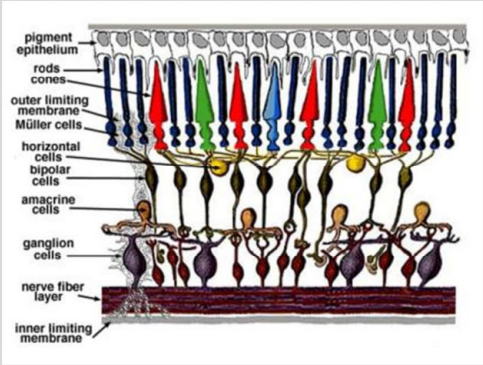
Interlace is a big nuisance for digital systems. Fields must be compressed separately, and need to be ‘doubled’ up to progressive for display on modern flat screens.

Interlace order can be hard to detect without a traditional CRT display.

Advantage of an updated image every 50th sec (Europe), rather than every frame like film – smoother motion

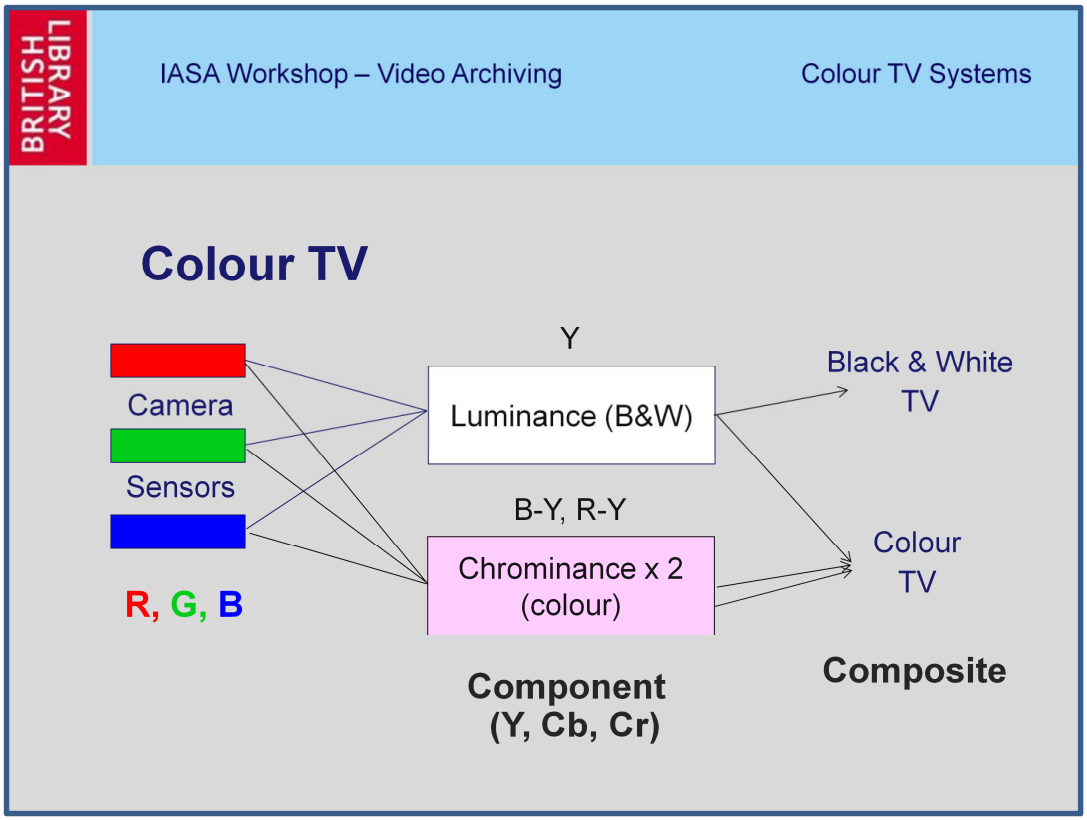
Colour TV systems

Human retina

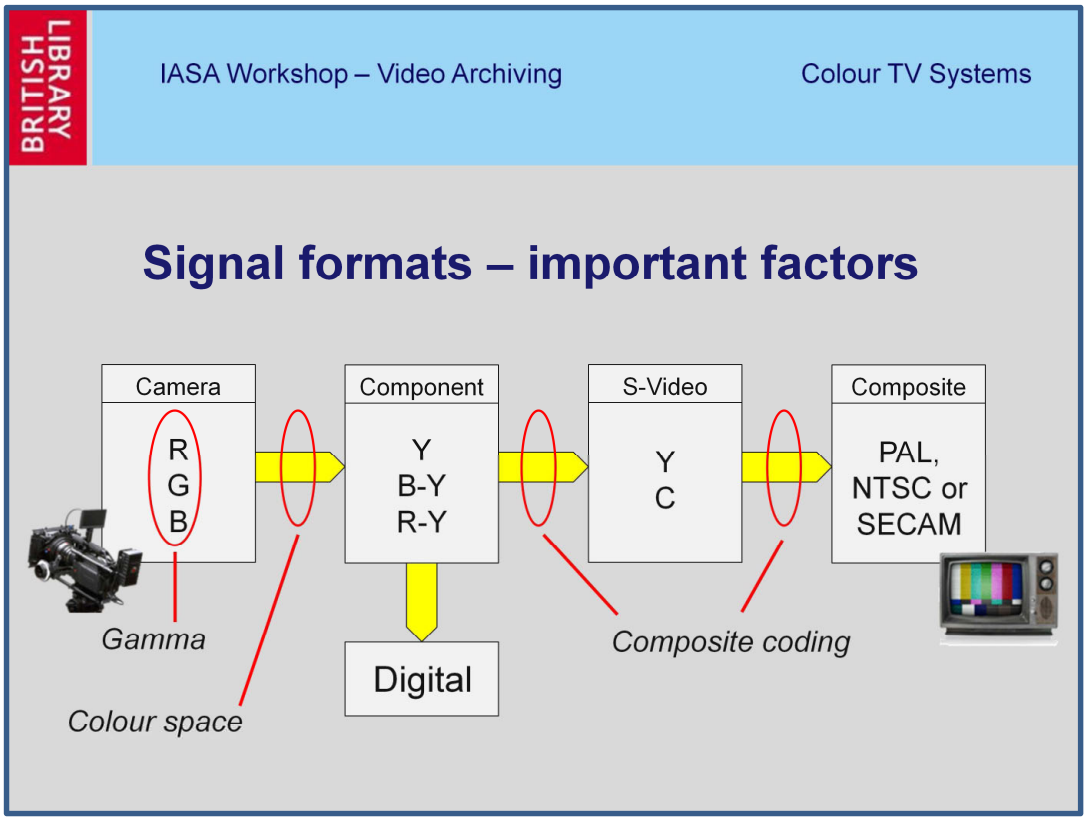


- 90 million rods for 'black and white' vision
- 4.5 million cones for colour vision

Rods and cones not spread evenly over retina, however.



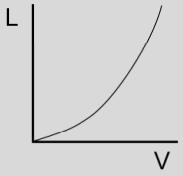
System of luminance and chrominance made new colour system compatible with existing TVs



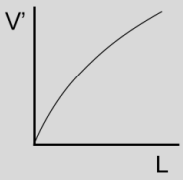
Signal chain from camera to screen – analogue signal progressively degrades as more processing is applied

Gamma correction

- Original CRT screens had non-linear light to voltage 'transfer characteristic'
- Reciprocal correction needed to be added by camera (or source), removed by display
- Added advantage of reducing perceptual noise in dark areas
- Different values for γ have been used historically – 1.8, 2.2, 2.4
- Differing methods of handling it in files/wrappers



Display Function $L = V^{\gamma}$



Camera Function $V' = L^{1/\gamma}$

Needs to be electronically removed by flat-screen displays which don't naturally have CRT gamma characteristic

Colour Space

- 3 variables always needed to convey colour – R,G,B; Y, B-Y, R-Y; X, Y, Z, etc
- However, sensors and displays operate in R, G, B
- Precise characteristics of R, G, B need to be known, and can vary, e.g:

Color space	White point (D_{65})		Primary color					
	x_W	y_W	x_R	y_R	x_G	y_G	x_B	y_B
625 line	0.3127	0.3290	0.640	0.330	0.290	0.600	0.150	0.060
525 line	0.3127	0.3290	0.630	0.340	0.310	0.595	0.155	0.070

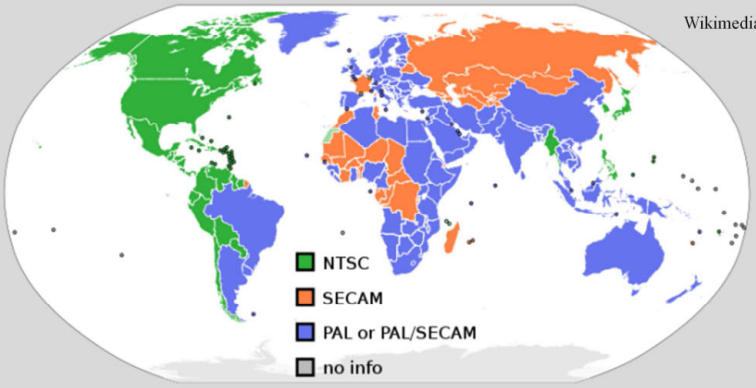
- Any conversion between systems should be precise and REVERSIBLE (i.e. no rounding with 8/10/12 bit binary maths)
  e.g: *ITU Rec. 601 - Y' = 0.299R' + 0.587G' + 0.114B' - needs rounding*

Composite encoding – SD only

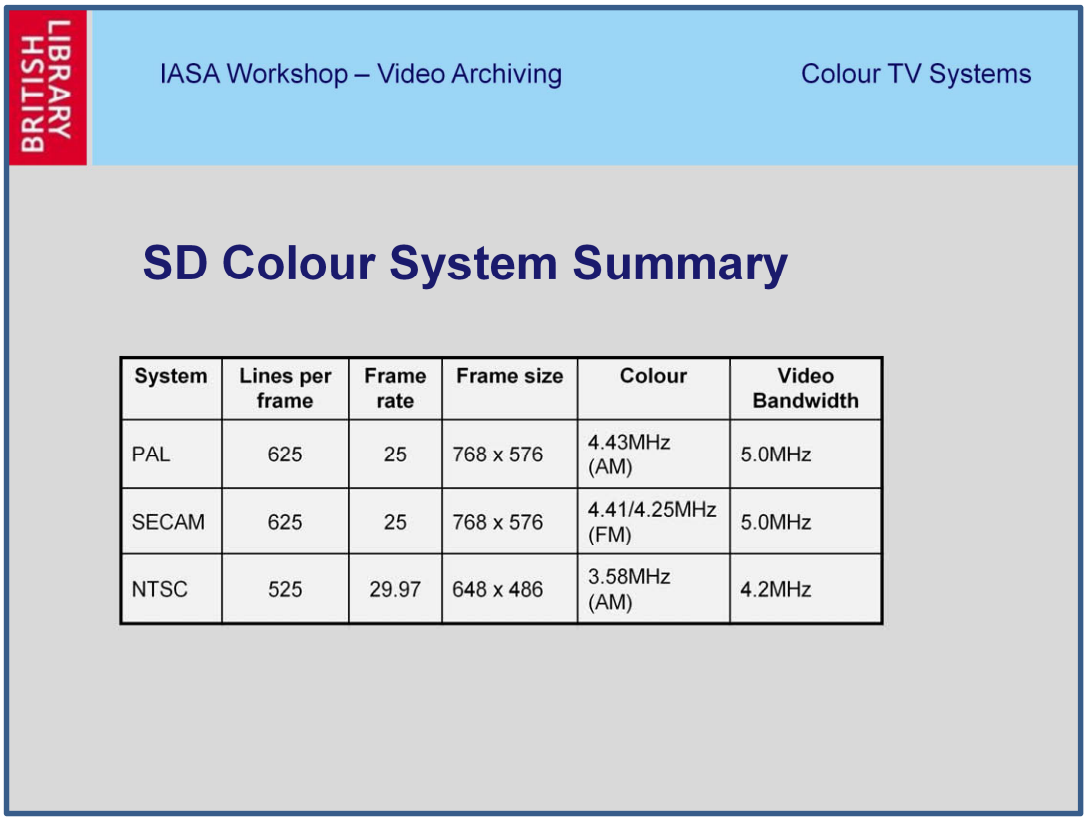
- B-Y and R-Y are low-pass filtered – removing detail not perceived – and then weighted to produce similar level
- Now known as U and V (PAL) or I and Q (NTSC)
- Amplitude modulate a colour subcarrier at 4.43MHz (PAL) or 3.58MHz (NTSC) or frequency modulate 4.41 and 4.25MHz alternately (SECAM)
- Encoded Chroma is added to Luma to produce Composite, or kept separate in the case of S-Video
- Separation on decoding is imperfect, resulting in dot crawl, cross-colour, PAL footprint, etc

Worldwide standards

Wikimedia Commons



http://www.videouniversity.com/articles/world-wide-tv-standards



System name, strictly speaking only applies to Composite video. I.e. Proper name for digital NTSC is 525-line or 525/29.97.

Compare video bandwidth with that for audio – 20Hz-20kHz.

The extra/missing lines

- 768 x 576 frame size is called '625 line'
- ~ 50 scanning lines do not contain picture information
- 'Vertical Interval' lines are reserved for technical information –
  - Synchronising signals
  - Vertical interval test signals (VITS)
  - Vertical Interval Timecode (VITC)
  - Teletext
  - Closed Captions
  - Active Format Description

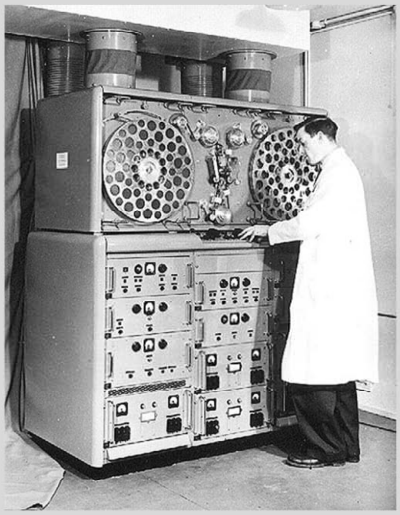
'In-raster' data

Some or all of this 'ancillary data' needs to be preserved, in some way. In D10 (IMX) it can be stored it as video – 32 extra lines - 608 lines for PAL, 512 lines for NTSC

Video recording

Linear recording

- High bandwidth video signal very difficult to record on tape as in audio
- Needs very high head-to-tape speed



BBC 'VERA'
Fixed heads, 200 ips tape speed

http://www.vtoldboys.com/vera.htm#btm

See here for film of Vera pictures: http://www.vtoldboys.com/vera02b.htm

Moving heads

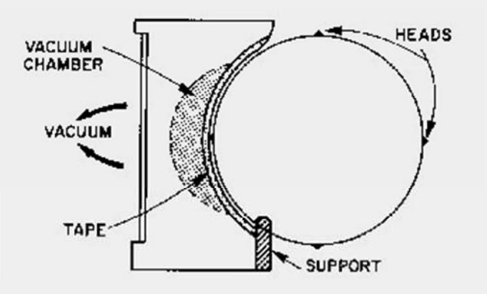
- Instead of moving tape past the heads, move the heads past the tape!
- Ampex introduced the first machine in 1956
- First broadcast use was 30th Nov. 1956

Ampex VR1000A

Wikimedia/Science Museum

Just out of sight on left of picture – 2 whole bays of electronics for the transport shown.

2 inch Quad tape



- Four heads (hence ‘quadruplex’)
- Rotate at 14,400 rpm on compressed air bearings
- Video tracks recorded across the tape – curved around heads

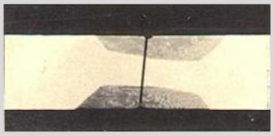
Each head records around 16 lines, so image is divided into bands, which are sometimes visible if machine needs adjustment

Shibaden - early semi-domestic machine

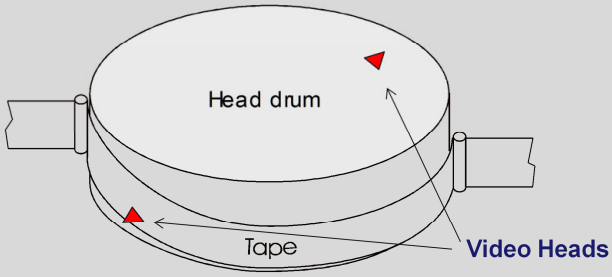
- EIAJ ½" system
- Much shallower head angle
- 'Helical' scan

Helical scan uses narrower tape, so less. Longer path means segmenting not necessary. Late 60s

Helical scan principle

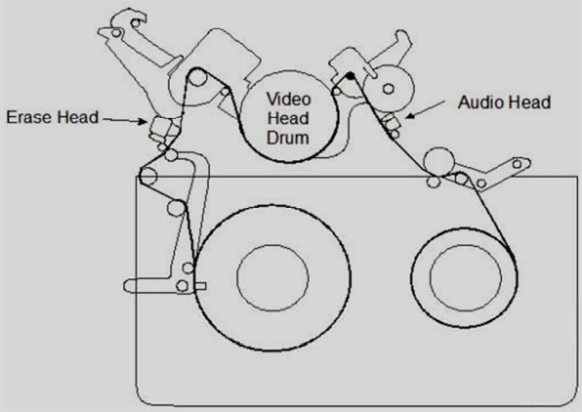


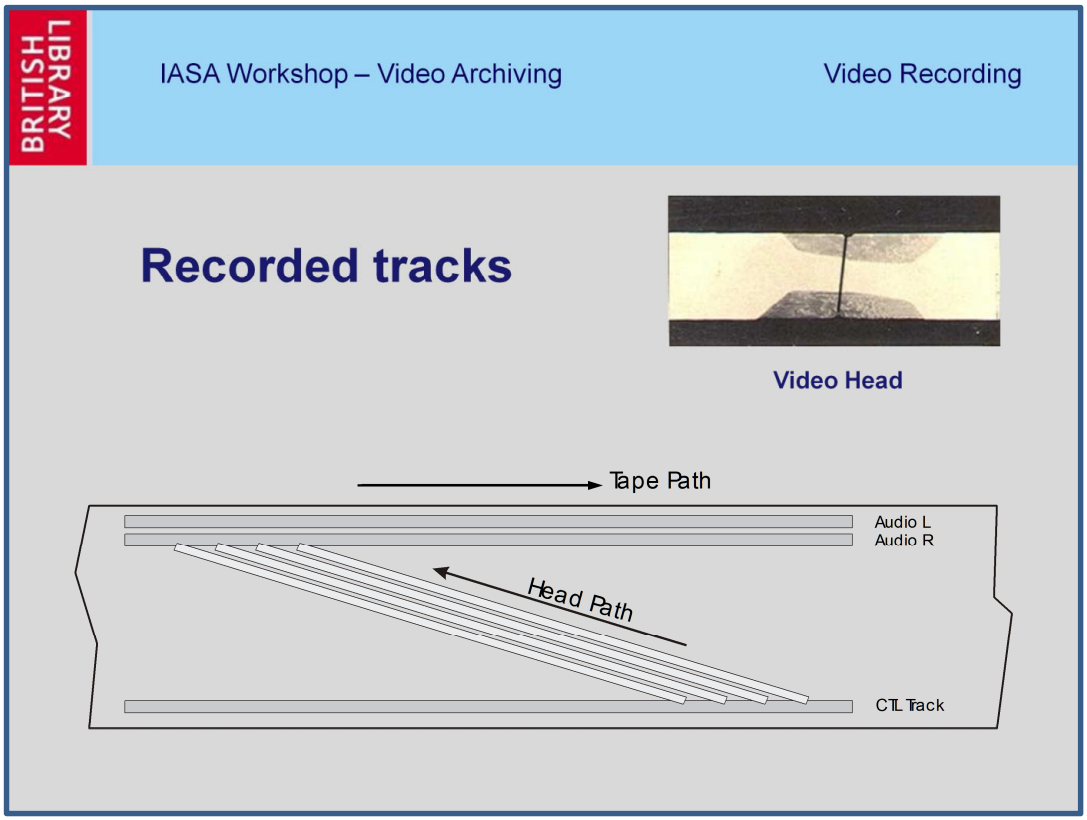
Video Head



- Used in all machines since quadruplex
- Two heads on opposite sides of drum – always one in contact with tape

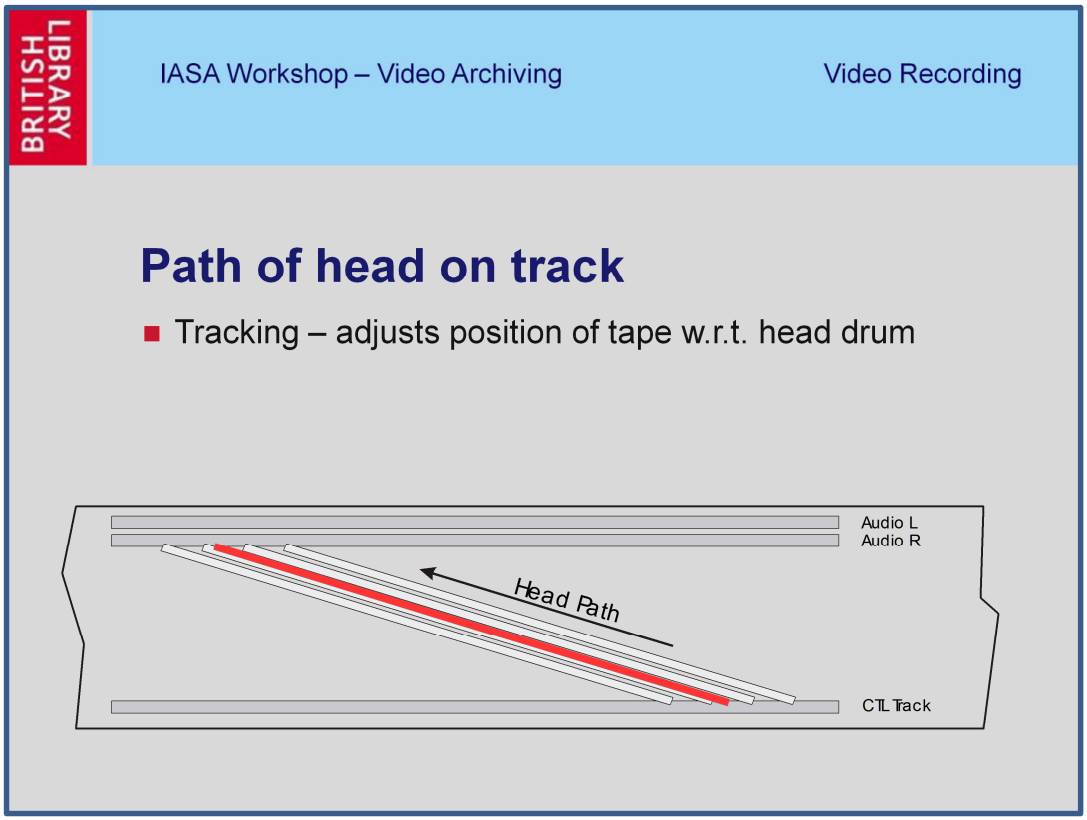
Typical tape path - VHS



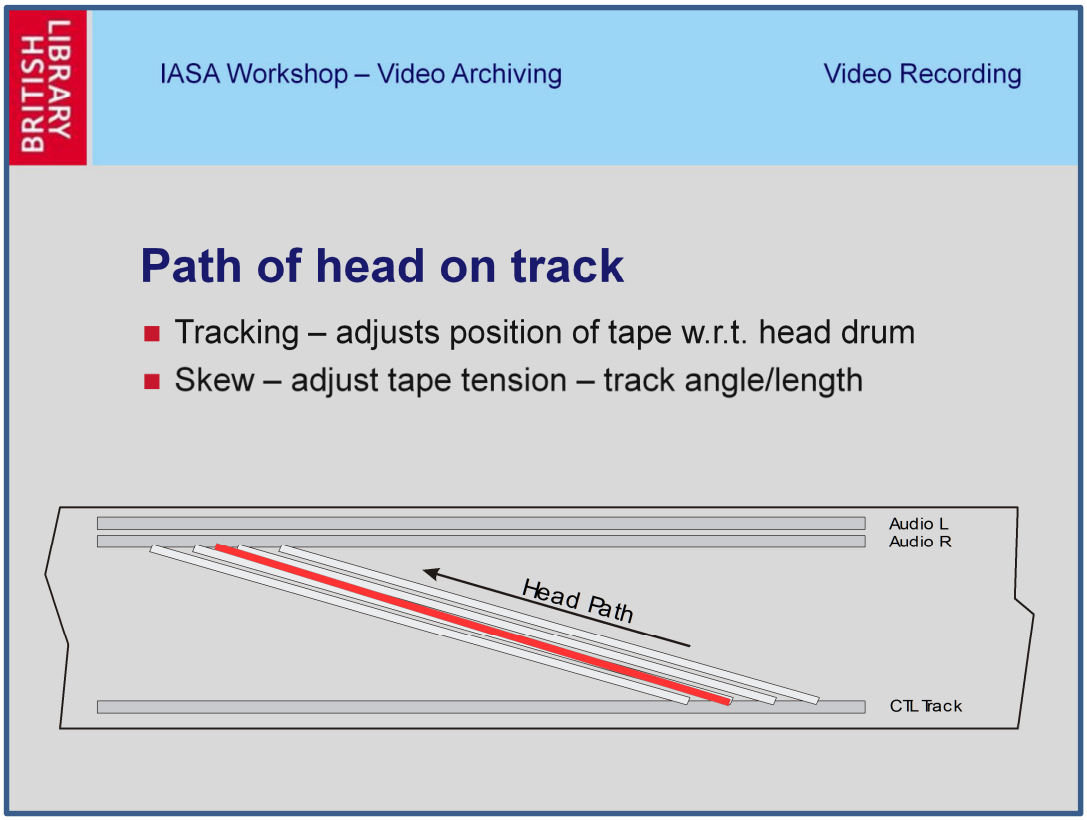


Eg. VHS. Actual track layouts differ between formats – different numbers of audio tracks, maybe also a TC track (see later). Maybe even a longitudinal track 'underneath' the video track.

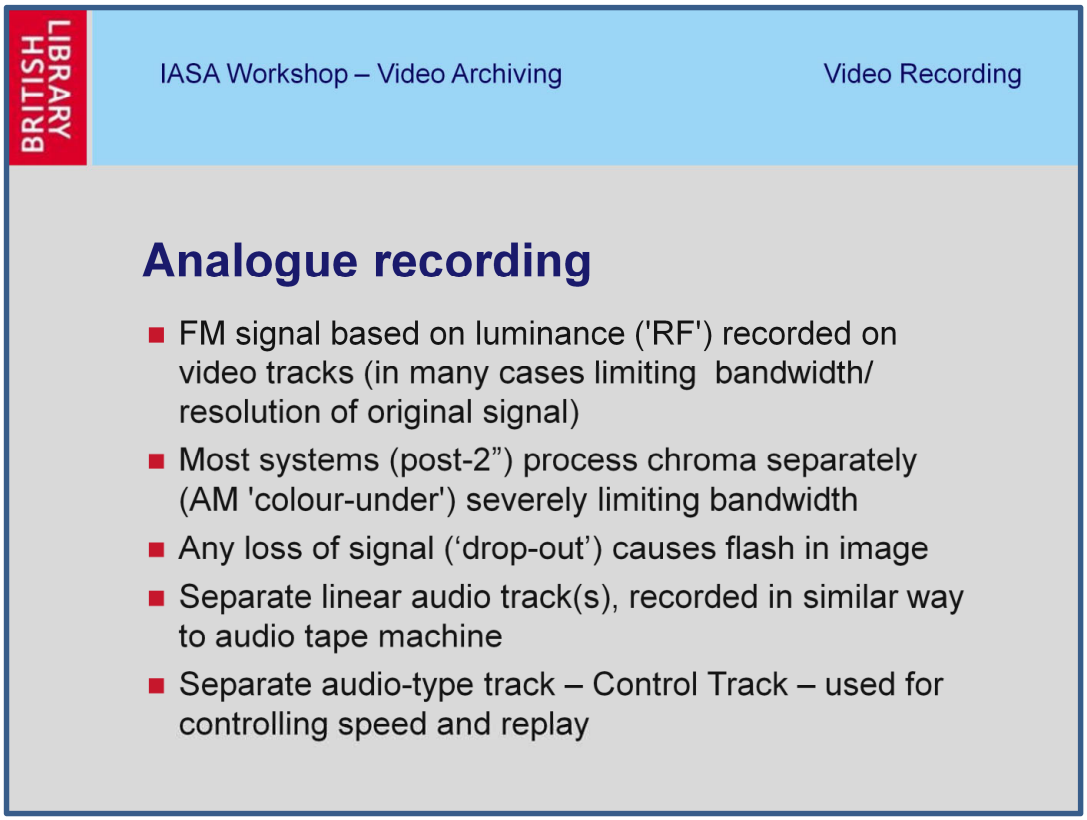
Control track vital for locking speed and phase of servos for capstan and head drum



Servos control capstan and head drum rotation to position rotating head precisely on-track.



U-matic has manual skew due to long tape path. Other machines have mechanical tensioning systems or reel servos to control tension

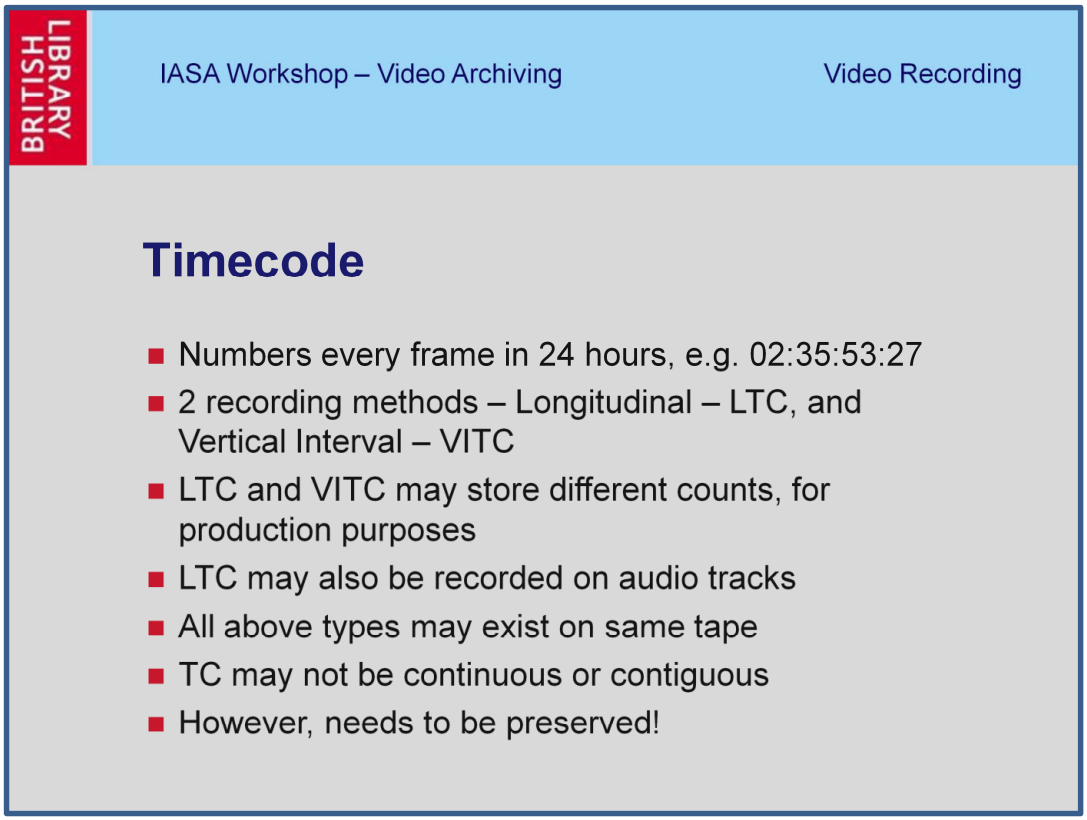


Different recording systems have differing capability to record the full bandwidth of Y. More domestic systems record less, resulting in softer image horizontally, and generally utilising additional artificial sharpening circuits.

Q. Should 'sharpness control' be on or off for preservation?

Timing errors

- Like any mechanical process, VTRs suffer 'wow and flutter'
- Affects recording and replay, multiplies with copying
- Needs to be corrected on digitising
- Can be done with A to D converter or with TBC – built-in or external
- TBC may also perform drop-out compensation – is this a good thing?
- TBC may add another A-D-A loop into path



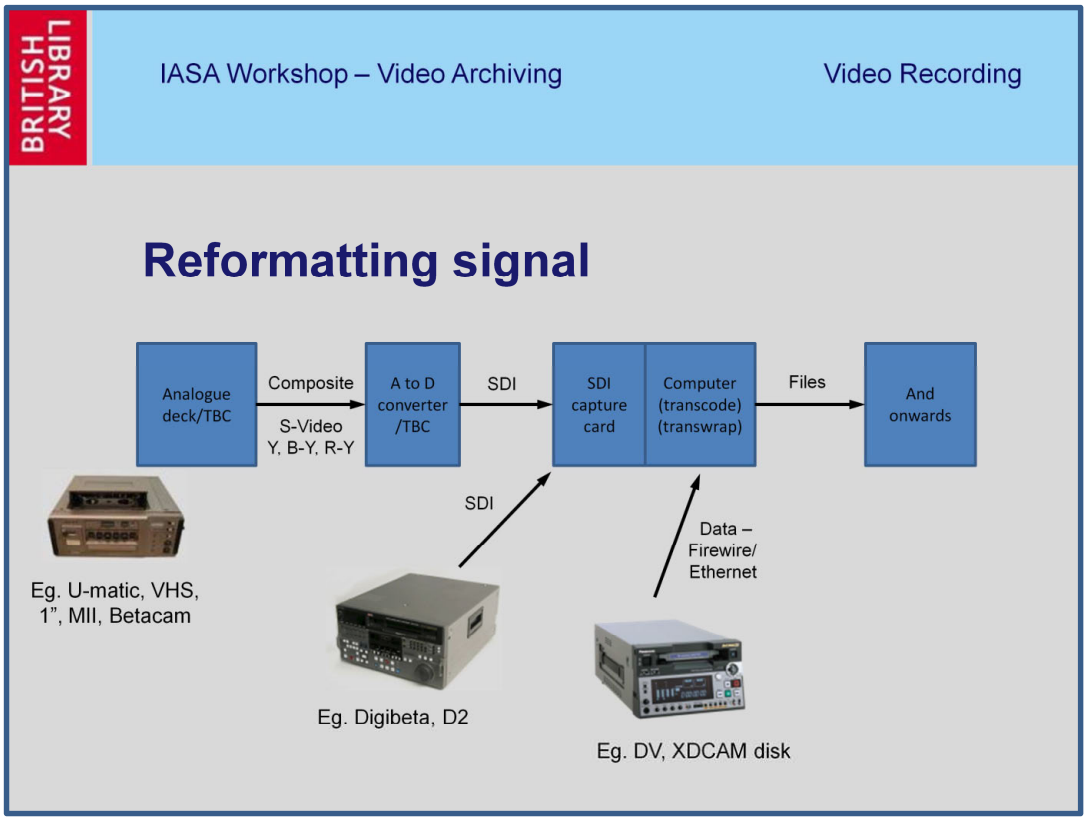
LTC is basically a modulated audio signal, so can be recorded on audio tracks for formats without TC (eg. VHS). Maybe an additional timecode on formats which do have LTC.

Only come wrappers allow for timecode, either as a start time or a track (to allow for TC jumps)

Digital recording – since around 1990

- Uses same helical scan principle
- Video and audio are digitised, then data is recorded
- Digitised video may be composite or component
- No separate longitudinal audio tracks are needed
- Timebase correction built-in for playback
- Small drop-outs can be concealed – error correction
  - large dropouts can result in severe degradation
- Compression used on most formats – may/may not be supported by corresponding file-based system

Portapack image for nostalgic reasons – not a viable archive player! Capture cards usually have SDI or HDSDI inputs, with embedded audio or AES



Analogue involves levels (colour bar alignment?), recording settings,

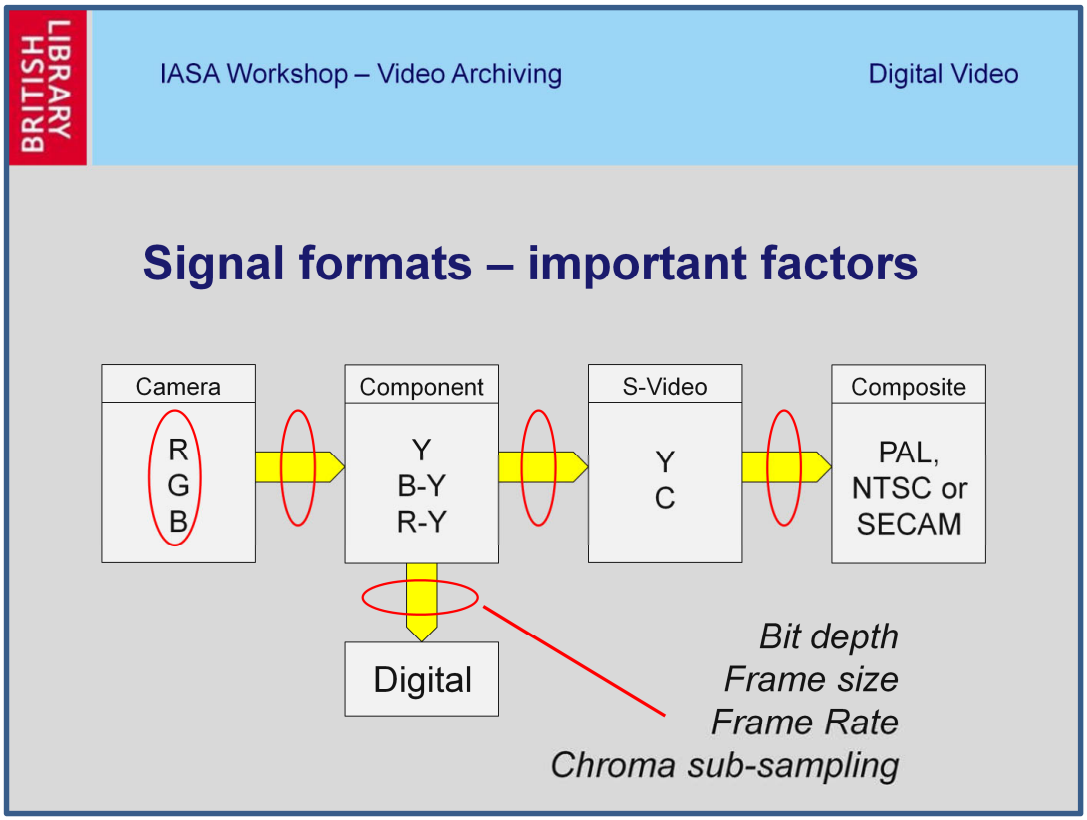
Digital straightforward, but may need data converter for earlier composite digital decks. Should involve recording something of error state

Digital data route preserves compression used on tape and metadata from stream – error reporting (concealment), record date, camera settings, etc, etc

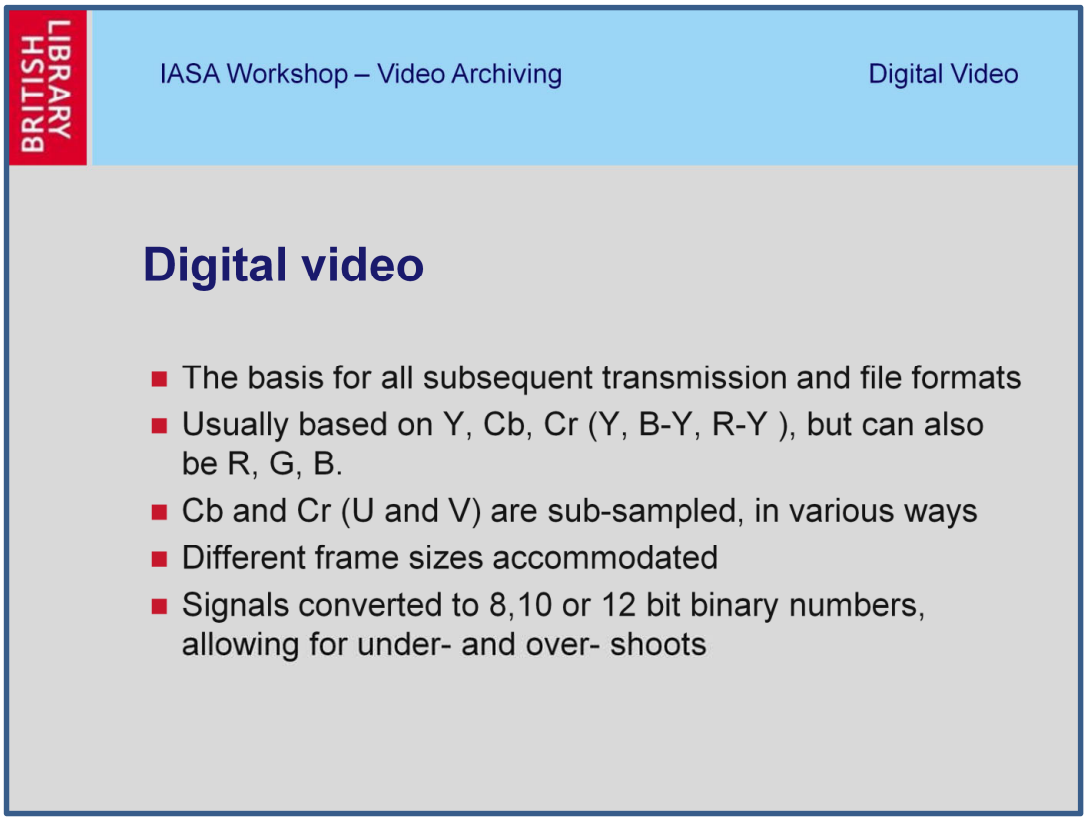
c.f. born digital: Last week – took in 3956 pop promo files from record company (10 year's work). 2000+ stills and other random files excluded - remaining 987 files were in 258 distinct digital formats!

VCR museum: http://www.totalrewind.org/mainhall.htm

Digital video

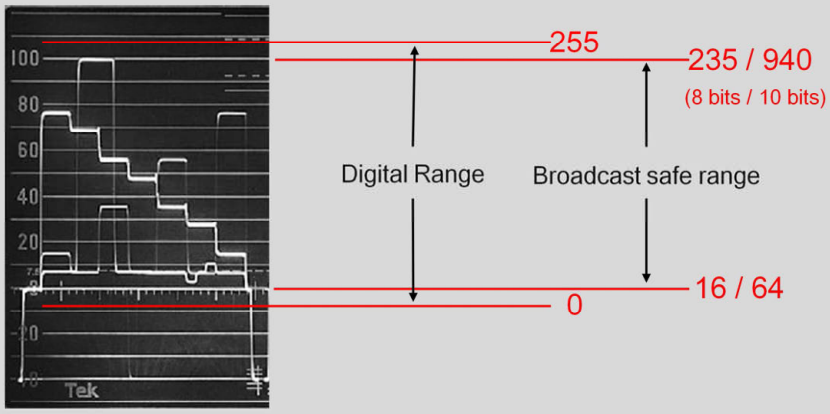


Gamma corrected signals should have prime – ‘– added as in Y’, but left off here for simplicity



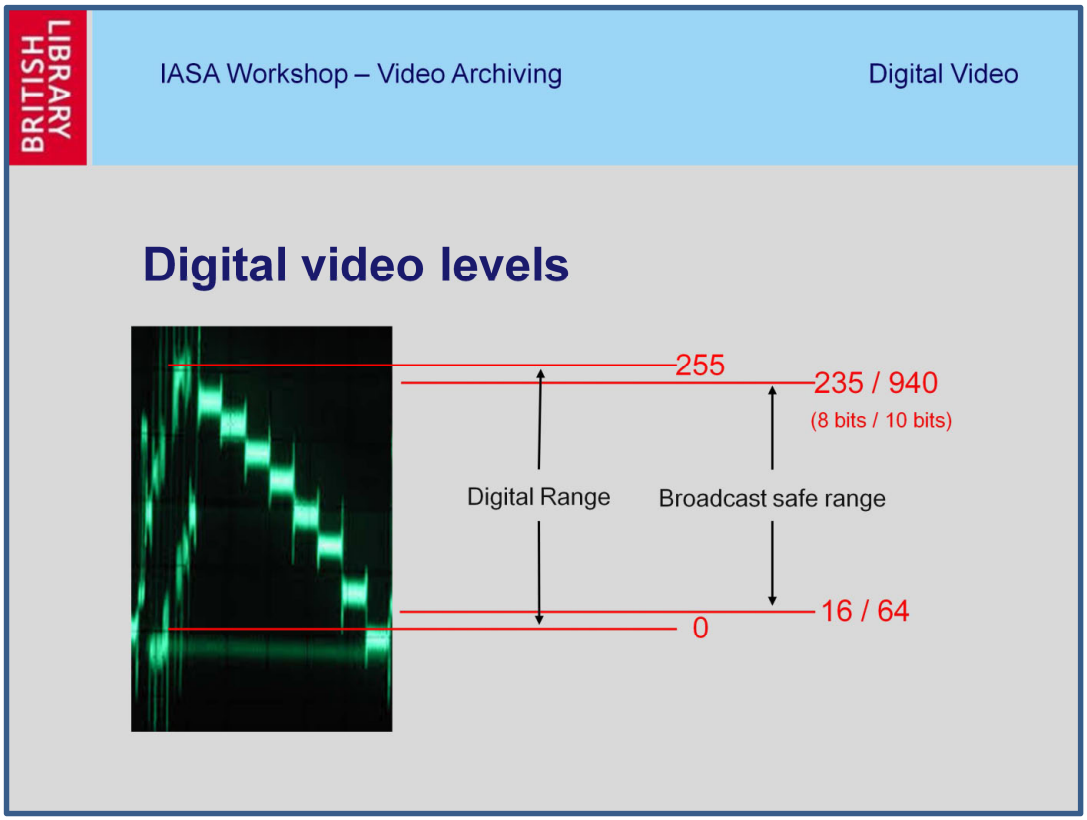
See later for sub-sampling

Digital video levels

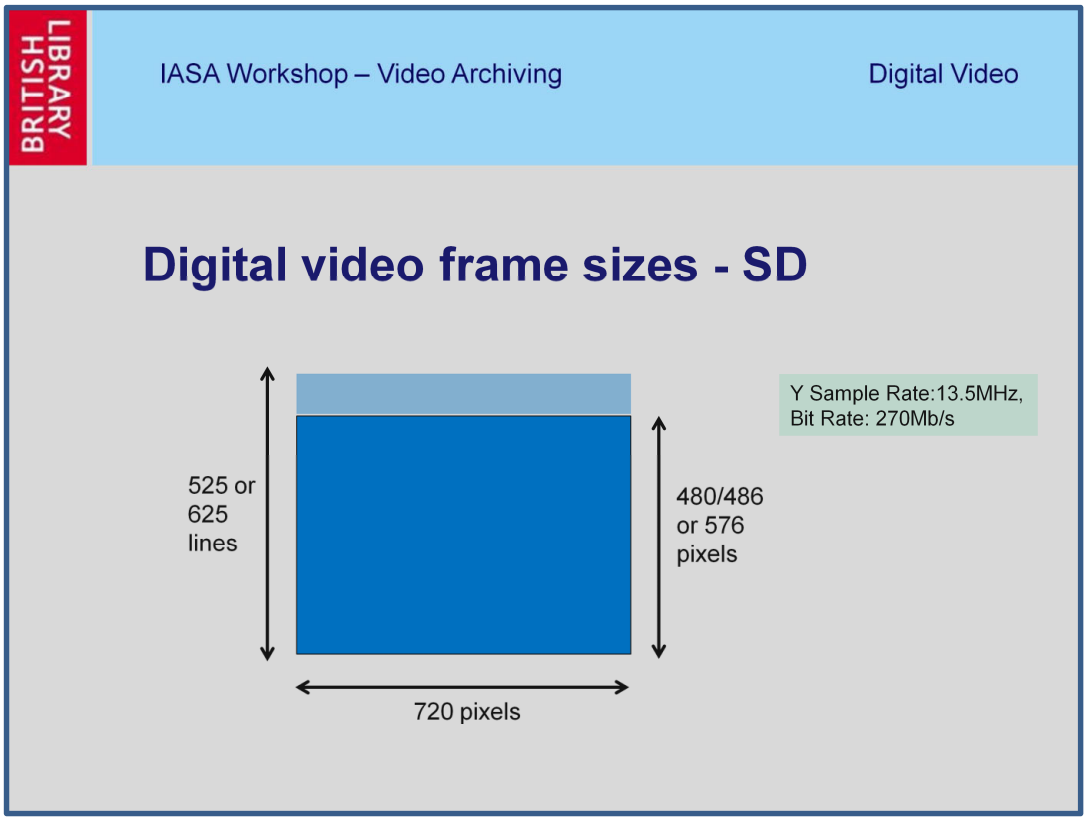


Video always standardised as 1v overall, including syncs – shown here at start and end of each line (left and right edge of image).

See here for an image and description of the picture corresponding to the test signal above: https://en.wikipedia.org/wiki/SMPTE_color_bars

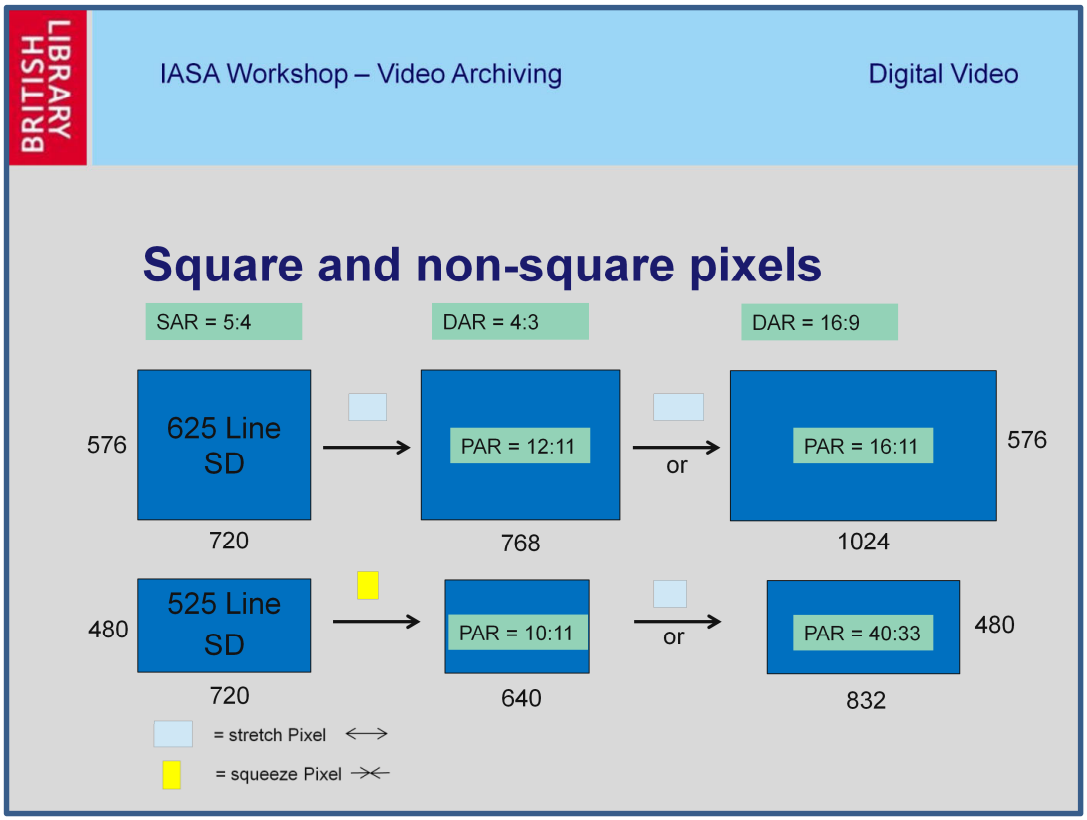


What happens to overshoots or over-exposed/misaligned levels? Need to adjust video levels to ensure no clipping and loss of information.



Frame is sampled at 720 pixels horizontally for PAL and NTSC. Difference in frame rates means this results in similar total number of pixels for both, and similar sampling rate.

720 pixels actually overscans width of image – only 704 have active picture. Allows for mis-timing - shifting image left or right.

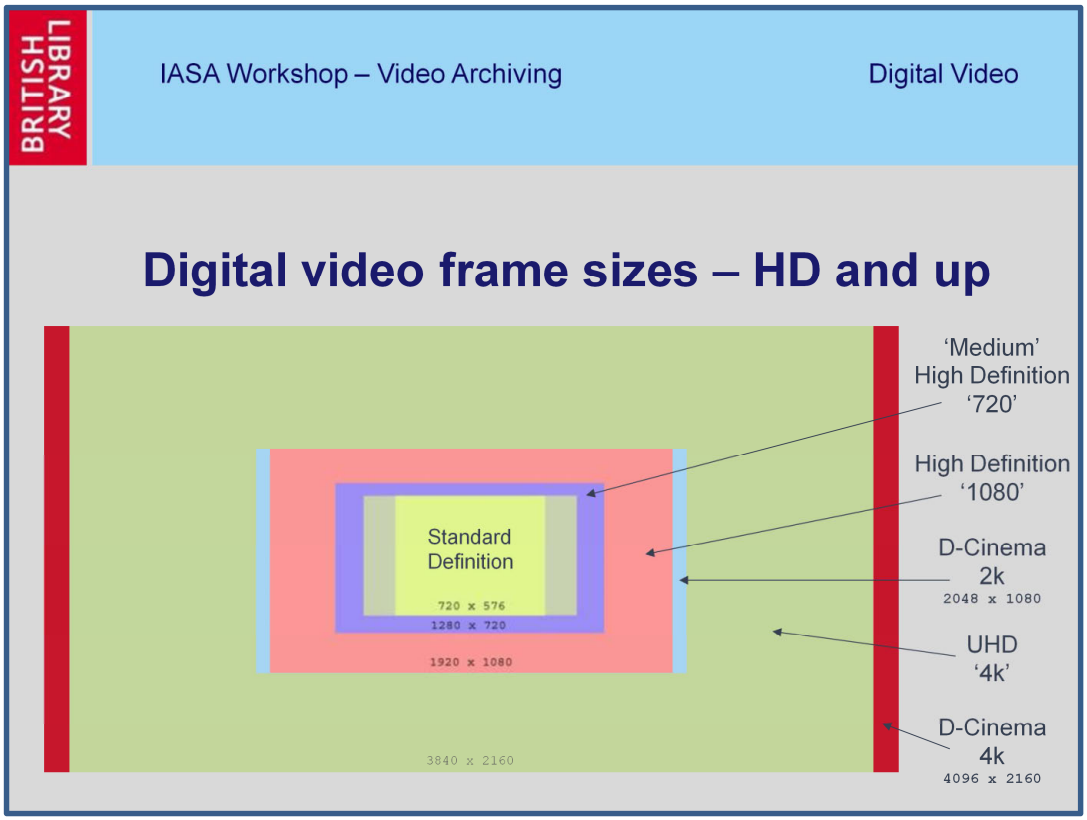


SAR = Stored Aspect Ratio, DAR = Display Aspect Ratio, PAR = Pixel Aspect Ratio

Slightly more complex when you take into account the fact that 704 not 720 pixels = 4:3! Above PAR takes this into account.

SAR also written as – NTSC4:3 - 1:0.909; PAL4:3 - 1:1.09 (typo in TC06)

SAR × PAR = DAR

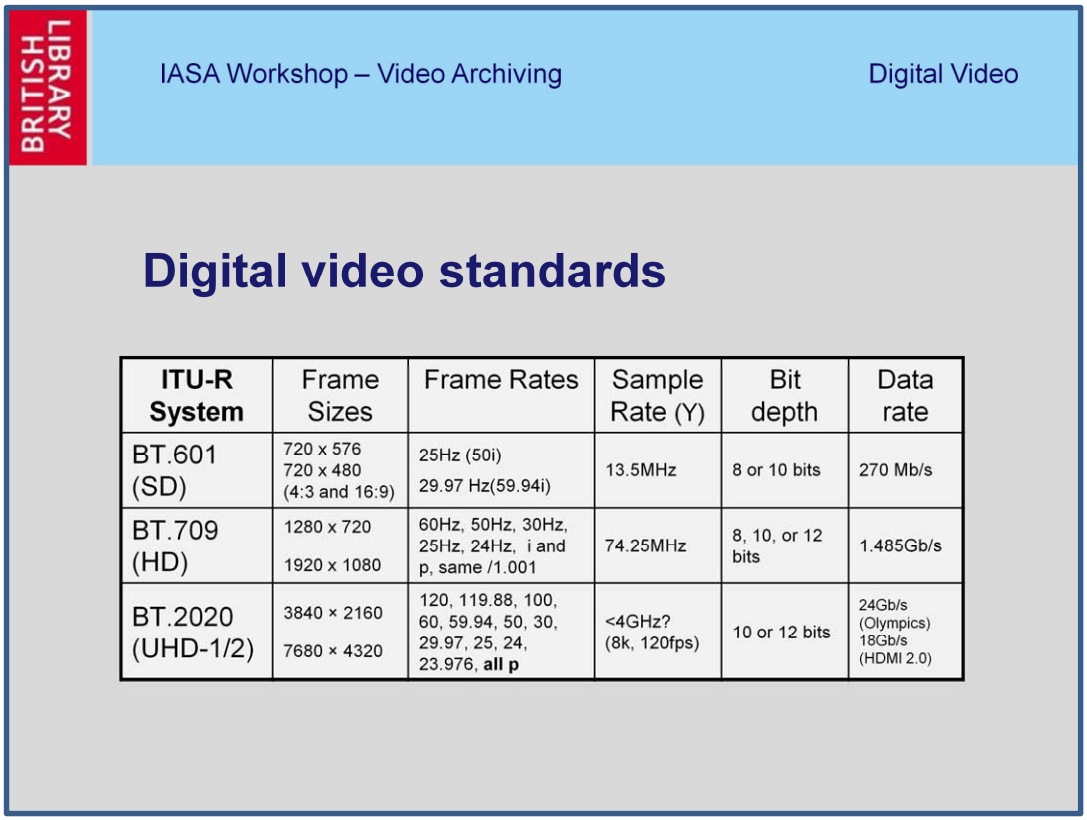


Plus UHD 8k!!

UHD 8K is approx 80 x SD pixels.

At 100fps, 320 x SD PAL data!!!

Please - no 3D!



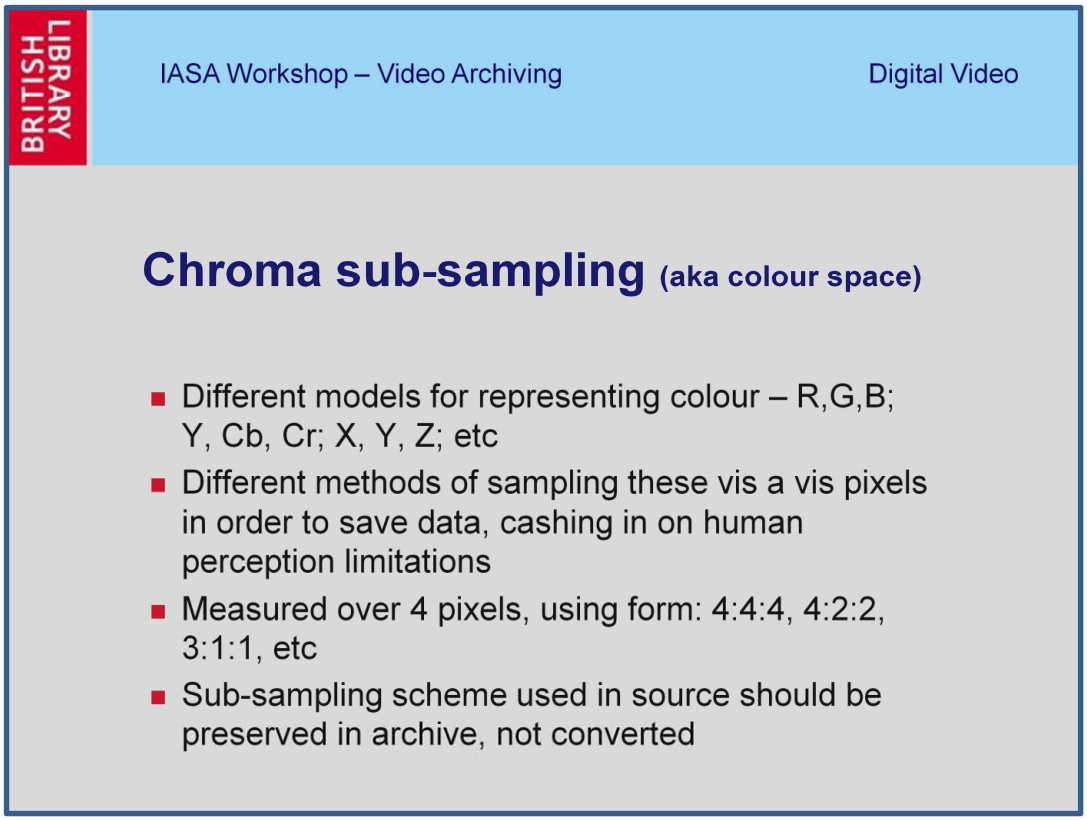
Standards built on standards – many SMPTE standards also involved. Avoiding standards alphabet soup here!

601 also called D1 after format which introduced it. Carried on 'SDI'

2020 based on (NHK) Super Hi-Vision. Main change with BT 2020 is big increase in colour gamut.

2020: 75.8%, DCI: 53.6%, Adobe RGB: 52.1%, and Rec. 709 colour space covers 35.9% (601 is similar).

Necessitates >10bits. Additional bits also needed for new 'HDR' TVs and production.



Last thing

Not really correct to refer to sub-sampling as ‘colour space’, but hey ho…

4:4:4 and 4:2:2

Four pixels:

+	R	R	R	R	= 4 samples
◯	G	G	G	G	= 4 samples
X	B	B	B	B	= 4 samples

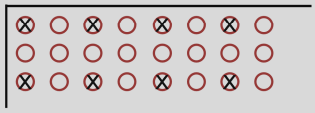
Four pixels:

◯	Y	Y	Y	Y	= 4 samples
X	B-Y		B-Y		= 2 samples
	R-Y		R-Y		= 2 samples

- 4:4:4 – generally denotes R,G,B colour space, no sub-sampling, i.e. no compression
- 4:2:2 – most common, Y, Cb, Cr colour space, chroma sub-sampled at half Y, results in 3:2 compression

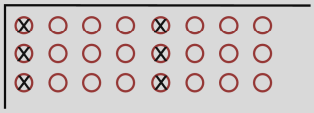
Other sub-sampling schemes

4:2:0



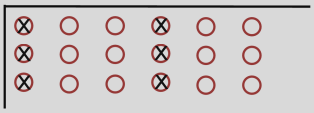
- 4:2:0 - used in PAL DV, DVDs, DTV, etc

4:1:1



- 4:1:1 – used in NTSC DV, DVCPRO, etc

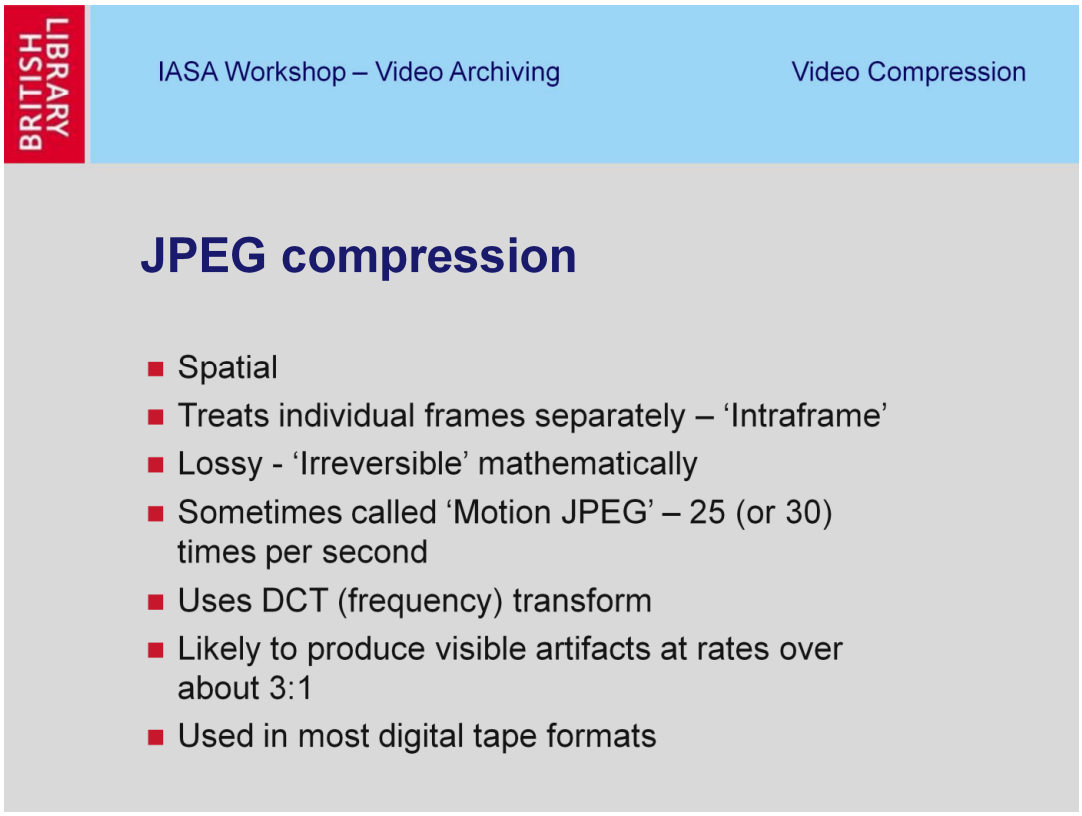
3:1:1



- 3:1:1 – Used in HDCAM and some tapeless codecs

DVCPRO HD even uses 2.6666:1:1! (1280 x 1080) to represent 1920 x 1080 HD

Video compression



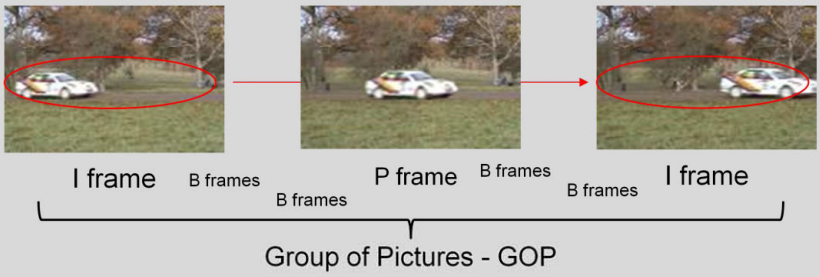
Amount of compression / picture quality is closely related to source material content

JPEG artifacts

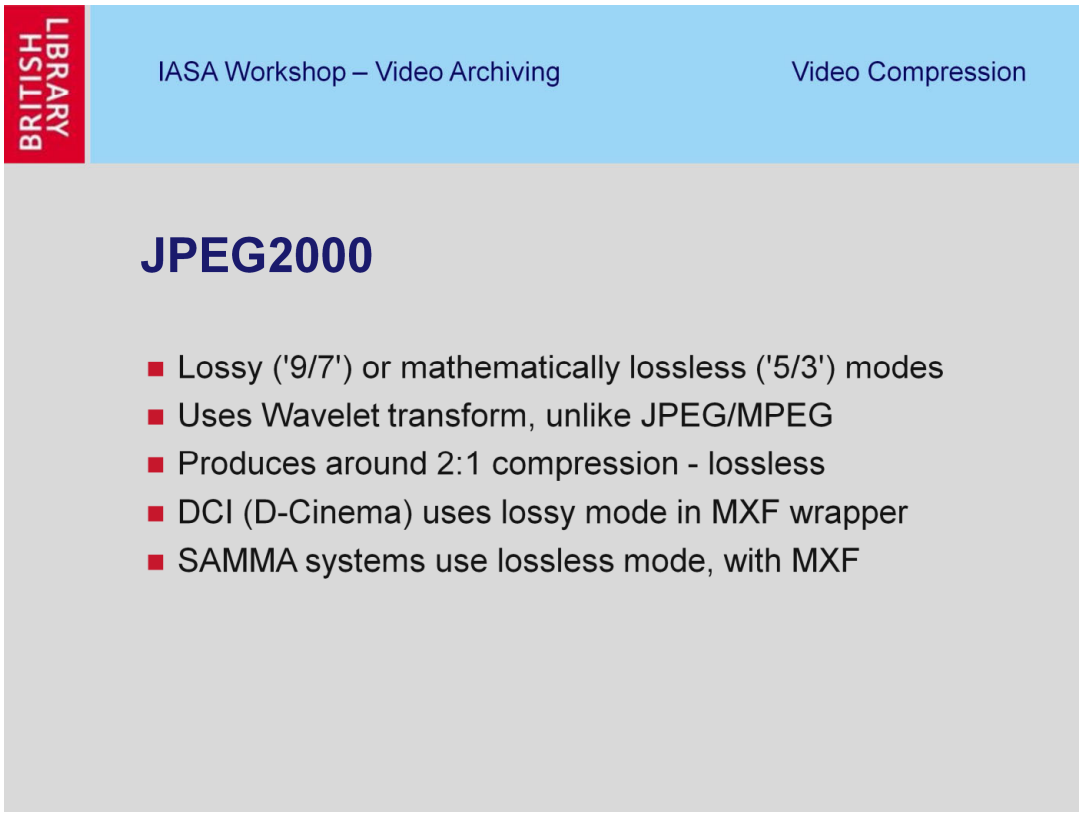
'Butterflies'

MPEG compression

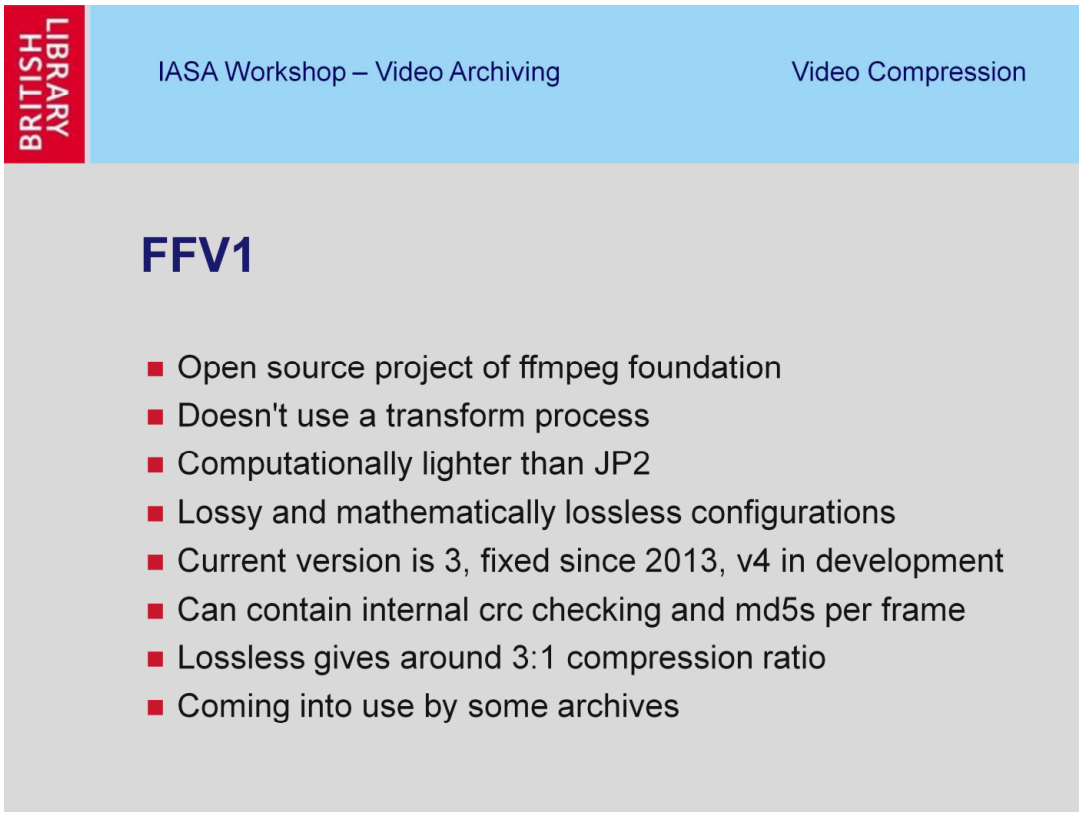
- Lossy – spatial and temporal
- Uses JPEG intra-frame compression first – I frame
- Looks for similarities between frames – 'Interframe'



Only used for archive in Intra mode (no GOP). But archiving should not be using lossy compression at all!



Quite CPU-intensive to process. Very few tools available. Some compatibility issues. Need RCT option for lossless RGB.

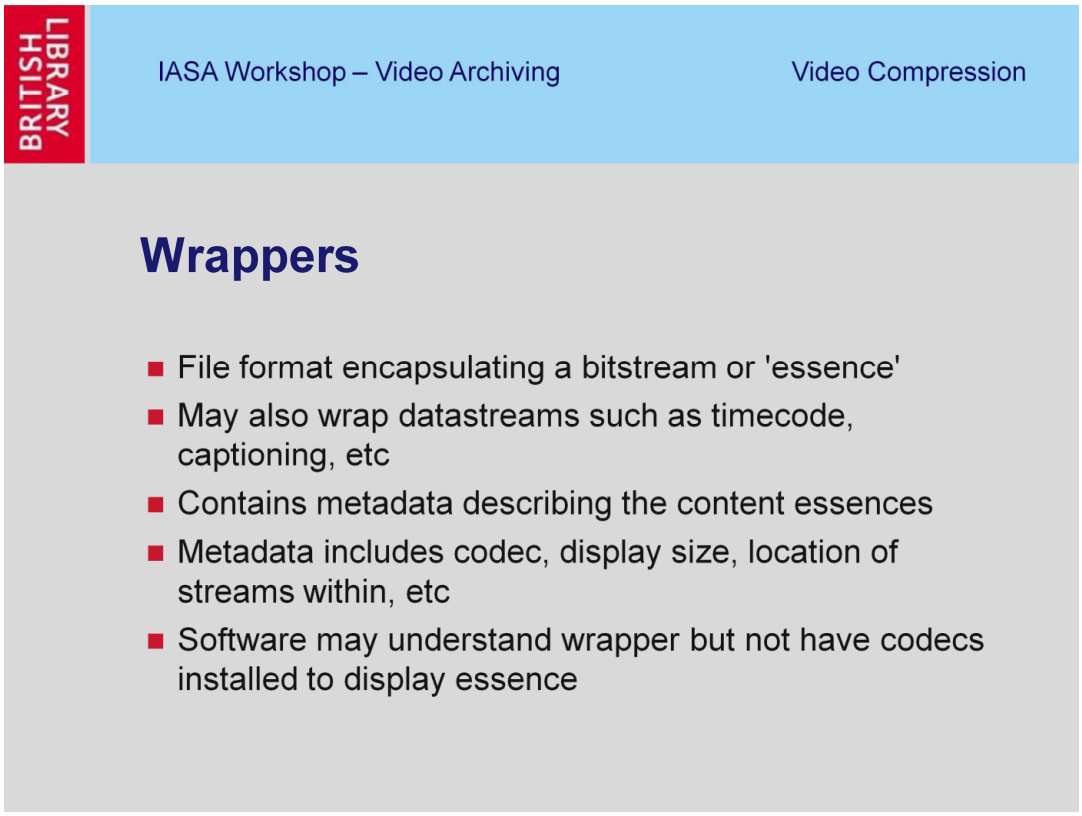


2007 MSU test showed, of 16 lossless codecs tested (not inc. JP2), FFV1 had best balance of processing speed and compression ratio (using old version). Also should mention BBC Dirac codec – has lossless mode.

PREFORMA, PREservation FORMATs for culture information/e-archives, EU project started on January 1st, 2014, - 'implementing good quality standardised file formats for preserving data content in the long term. The main objective is to give memory institutions full control of the process of the conformity tests of files to be ingested into archives'. Moving image format based on FFV1/MKV. http://www.preforma-project.eu/

IETF (Internet Engineering Taskforce) just started CELLAR project: Codec Encoding for LossLess Archiving and Realtime transmission to codify standards and documentation, and formalise specifications. inc Matroska wrapper

Content Integrity ../.. Fixity



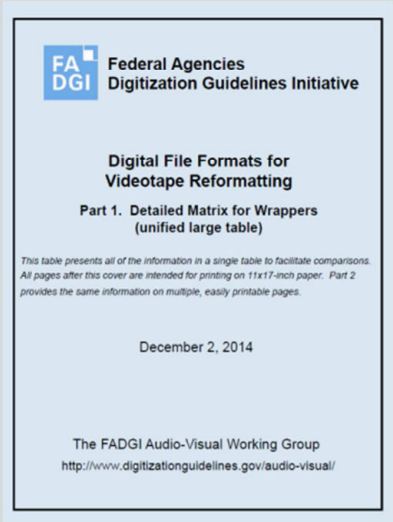
Wrapper allows inclusion of additional data in file without altering essence encoding

Point 4 – some essence streams (such as MPEG2) can contain metadata which conflicts with wrapper metadata (as can actual video encoded in essence stream – eg. AFD)

Standards for archiving - discussion

Requirements for reformatting

- Use original TV standard (525, 625, etc),
- Use SD/HD/etc, as original, p or i
- 4:2:2 is appropriate, although compressed 3:2
- Use 10-bit uncompressed for video (?)
- Use best available signal path (component, s-video, etc)
- Capture all audio tracks, at 48kHz/16bit, 96/24?
- Preserve ancillary data, either digitised or converted to metadata



FADGI Federal Agencies Digitization Guidelines Initiative

Digital File Formats for Videotape Reformatting

Part 1. Detailed Matrix for Wrappers (unified large table)

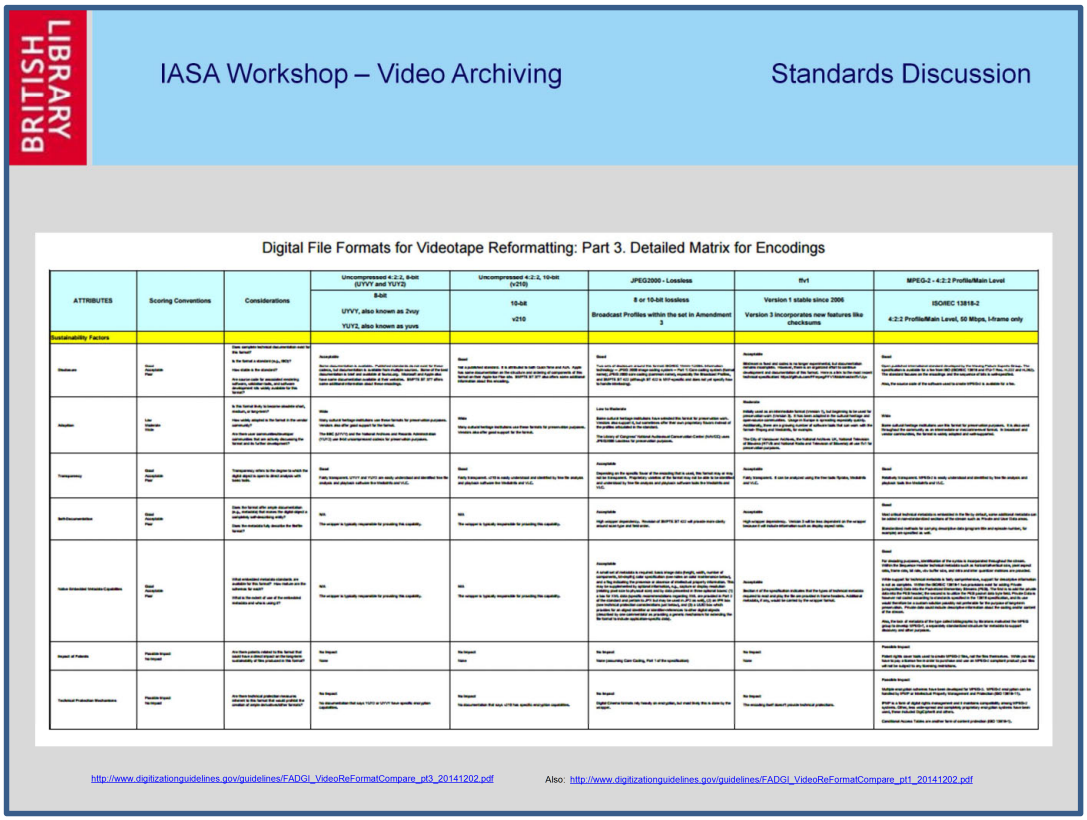
This table presents all of the information in a single table to facilitate comparisons. All pages after this cover are intended for printing on 11x17-inch paper. Part 2 provides the same information on multiple, easily printable pages.

December 2, 2014

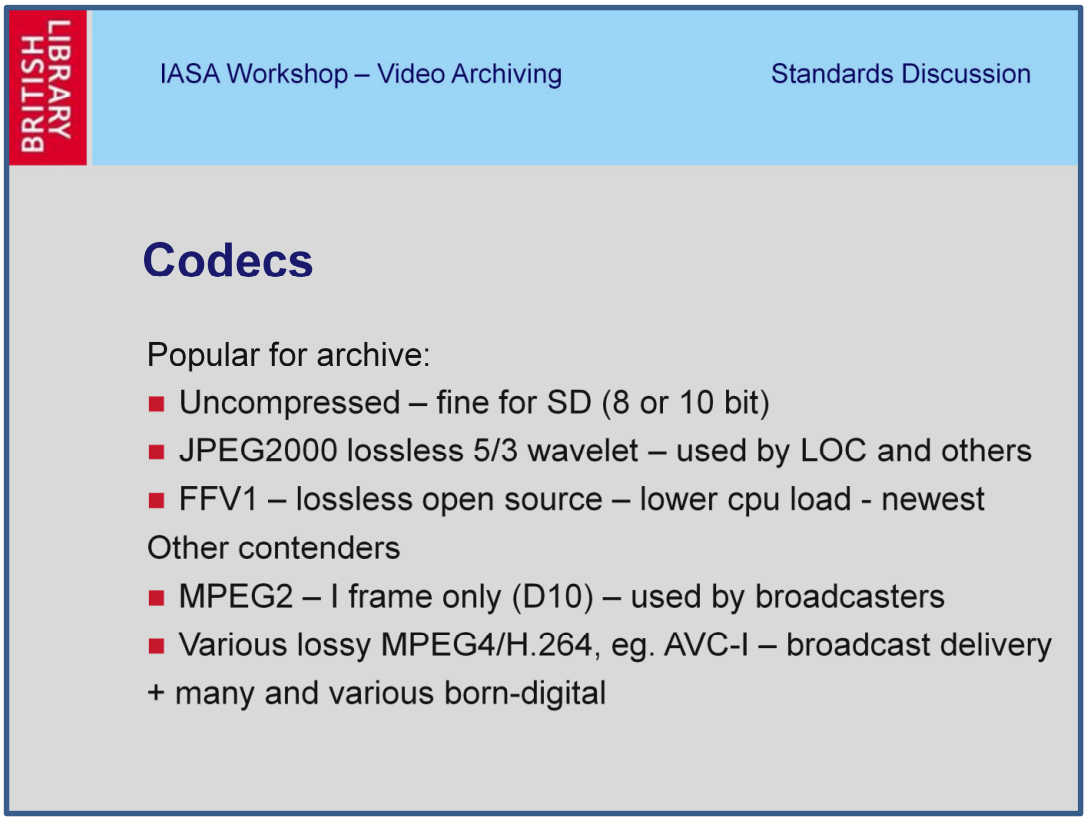
The FADGI Audio-Visual Working Group
http://www.digitizationguidelines.gov/audio-visual/

http://www.digitizationguidelines.gov/guidelines/video_reformatting_compare.html

Check the FADGI website for most recent documents



Note – summary table above contains slightly different information to the full discussion document on the websites mentioned



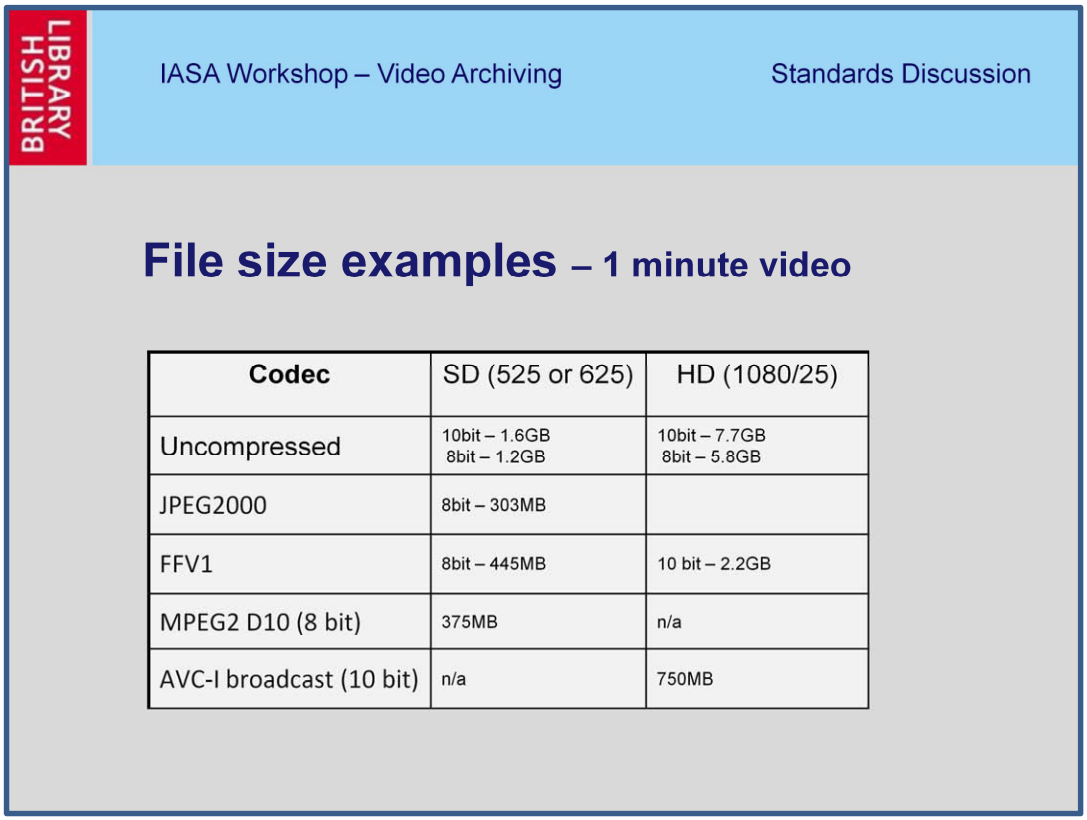
Maybe these are 'safe haven' interim formats, until one emerges as a clear winner in the stakes – Disclosure, Adoption, Transparency, Self-documentation, External dependencies, Technical protection mechanisms (i.e. no DRM) (LOC Sustainability).

Measure of compression is bpp – bits per pixel. E.g. uncompressed 422 8bit is 16bpp, youtube might be 0.26bpp.

Not good to transcode between lossy codecs

MPEG2 should be main profile/main level, 422.

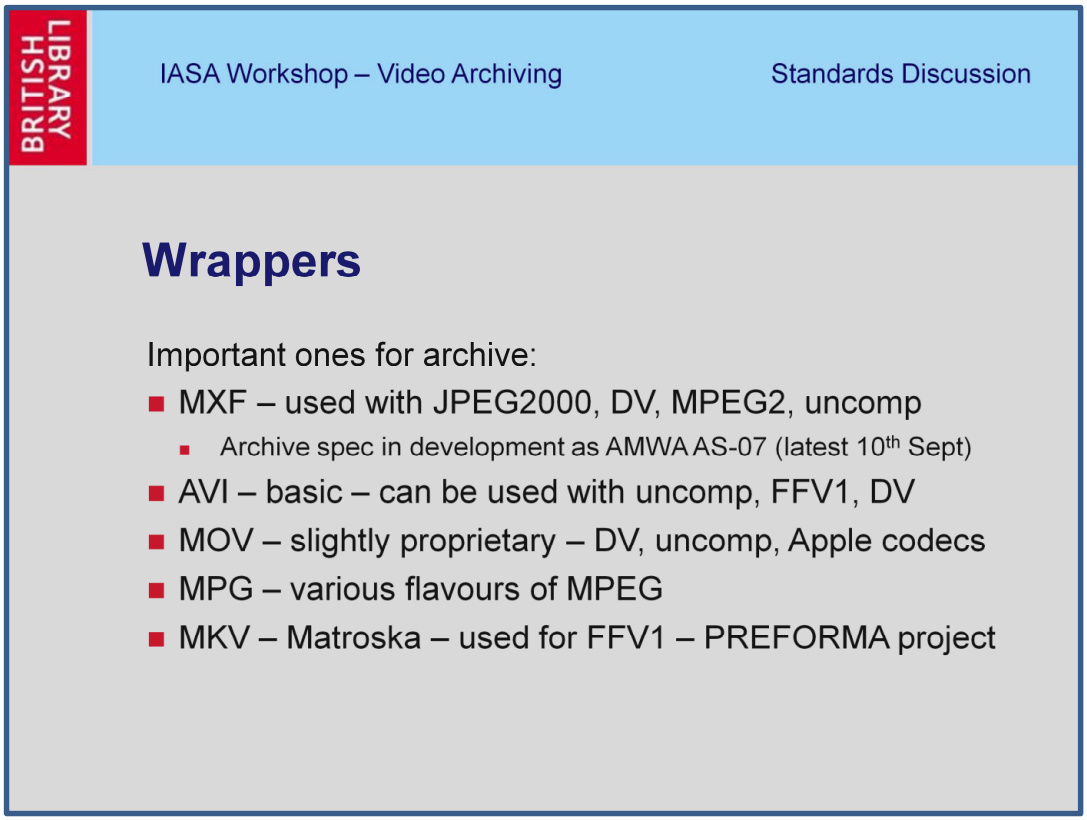
Q. Is it acceptable to use lower quality codec for certain videotape formats?



c.f. Audio (stereo 24/96): 1 min = 34MB

FFV1 may be useful for replacing Avid and Apple post-production codecs, but is about 2x size

AVC-I achieves good compression considering 10bit and HD, compared to MPEG2.



At this stage, these may turn out to be interim formats.

Point 1: MXF is Operational Pattern 1a – Video and Audio essence in same wrapper, only one package (clip)

Point 2: AVI has limited scope for metadata, but this has been extended by including EXIF fields within it

Points 1 and 3: Both MXF and QT can also refer to items outside the wrapper

Talking points?

- Interim/preservation codecs and wrappers
- Content integrity checking – internal/external
- Dealing with auxilliary files – pdfs, scans, text, etc
- Formats for metadata – technical, preservation, etc
- Metrics – measurement and recording of analogue and digital equipment and settings before and during reformatting