

On Audio-Visual File Formats

Reto Kromer • AV Preservation by reto.ch

Open-Source Tools and Resources for Audio-Visual Archives

Elías Querejeta Zine Eskola
Donostia (San Sebastián), Spain
19, 21, 26 and 28 May 2020

1

Summary

- digital audio and digital video
- container, codec, raw data
- different formats for different purposes
- audio-visual data transformations

2

Digital Audio

3

Digital Audio

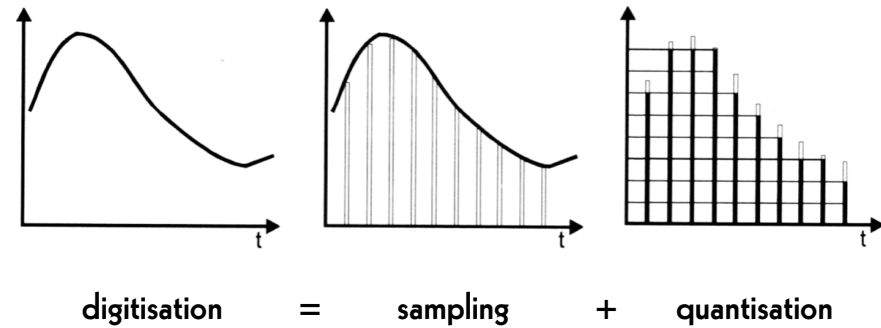
- sampling
- quantisation

4

Sampling

- 44.1 kHz
- 48 kHz
- 96 kHz
- 192 kHz
- 500 kHz

6



5

Quantisation

- 16 bit ($2^{16} = 65\,536$)
- 24 bit ($2^{24} = 16\,777\,216$)
- 32 bit ($2^{32} = 4\,294\,967\,296$)

7

Digital Video

8

Digital Video

- resolution
- bit depth
- linear, power, logarithmic
- colour model
- chroma subsampling
- illuminant

9

Resolution

- SD 480i / SD 576i
- HD 720p / HD 1080i
- 2K / HD 1080p
- 4K / UHD-1
- 8K / UHD-2

10

Bit Depth

- 8 bit ($2^8 = 256$)
- 10 bit ($2^{10} = 1\,024$)
- 12 bit ($2^{12} = 4\,096$)
- 16 bit ($2^{16} = 65\,536$)
- 24 bit ($2^{24} = 16\,777\,216$)

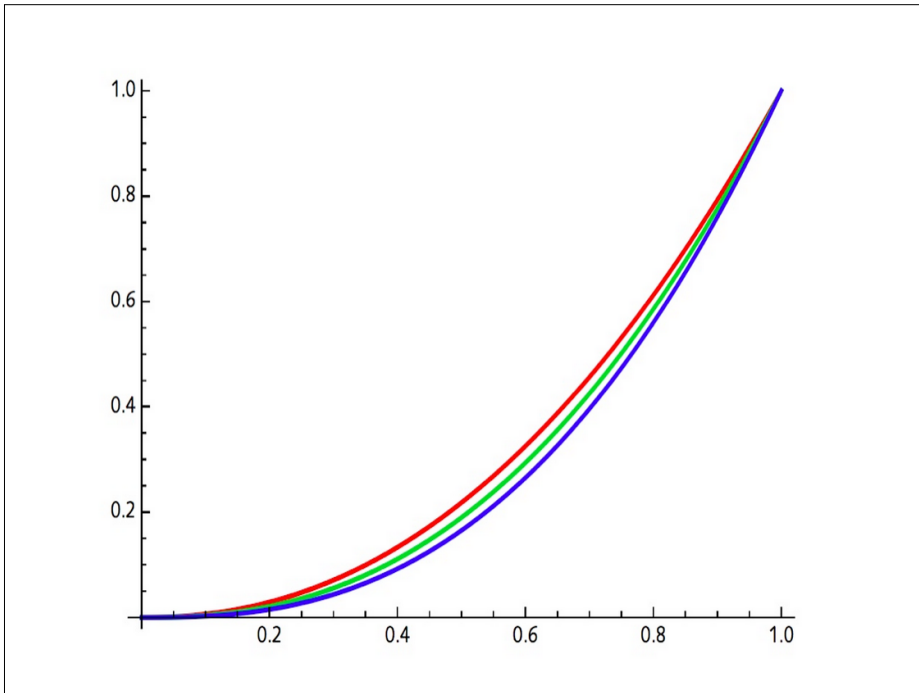
11

Linear, Power, Logarithmic

«medium grey»

- linear: 18%
- power: 50%
- logarithmic: 50%

12

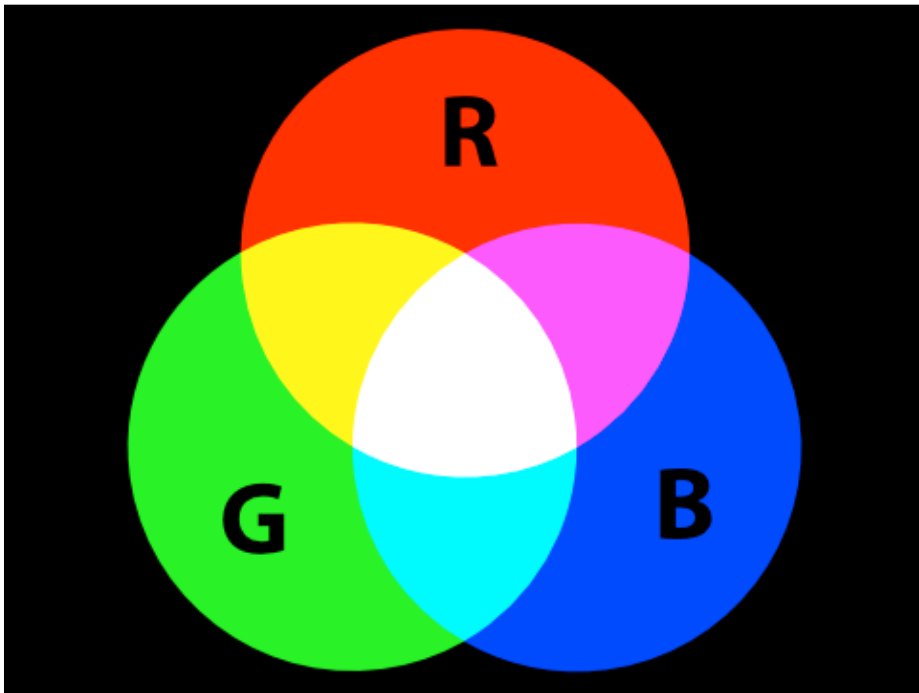


13

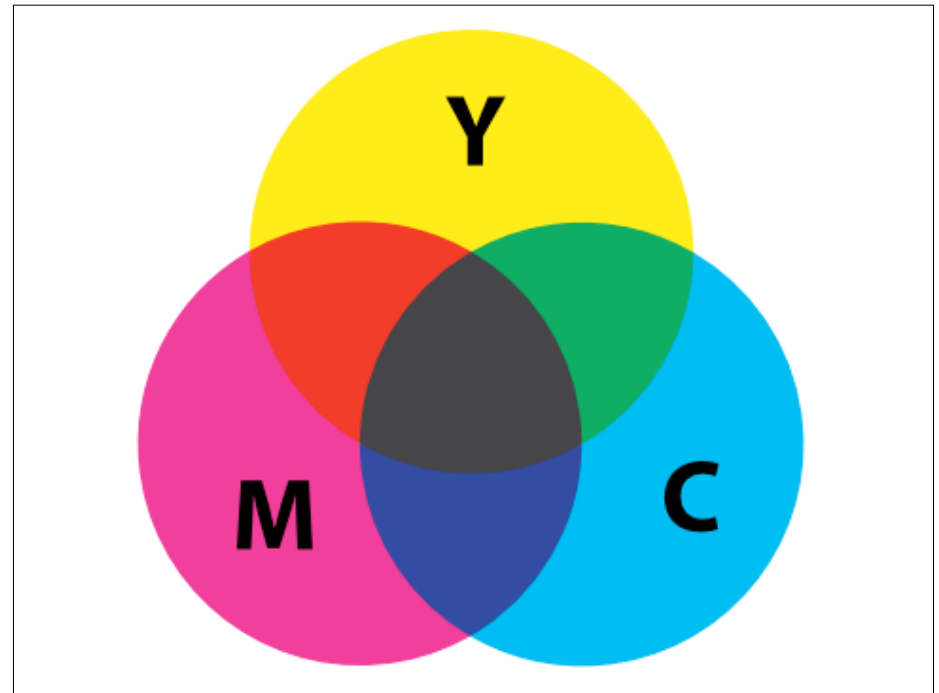
Colour Model

- XYZ, L*a*b*
- RGB / R'G'B' / CMY / C'M'Y'
- Y'IQ / Y'UV / Y'D_BD_R
- Y'C_BC_R / Y'CoC_G
- Y'P_BP_R

14



15



16



17



19

RGB24

00000000	11111111	00000000	00000000
00000000	00000000	11111111	00000000
00000000	00000000	00000000	11111111
00000000	11111111	11111111	11111111
11111111	00000000	11111111	11111111
11111111	11111111	00000000	11111111

18

Compression

- uncompressed
- lossless compression
- lossy compression
- chroma subsampling
- born compressed

20

Uncompressed

- + data simpler to process
- + software runs faster
- bigger files
- slower writing, transmission and reading

Examples: TIFF, DPX, DNG, OpenEXR

21

Lossless Compression

- + smaller files
- + faster writing, transmission and reading
- data processing complexer
- software runs slower

Examples: JPEG 2000, FFV1

22

Lossy Compression

- optimised for image acquisition and/or postproduction
- optimised for access

Examples (mezzanine): ProRes 422, ProRes 4444, DNxHD, DNxHR

Examples (access): H.264 (AVC), H.265 (HEVC), AV1

23

Chroma Subsampling

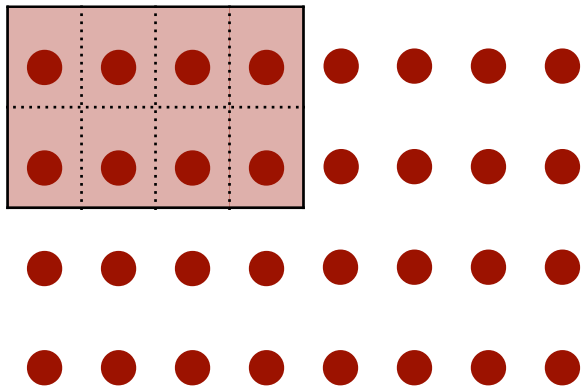
- 4:4:4
- 4:2:2
- 4:2:0 / 4:1:1

24

4:4:4

4

4

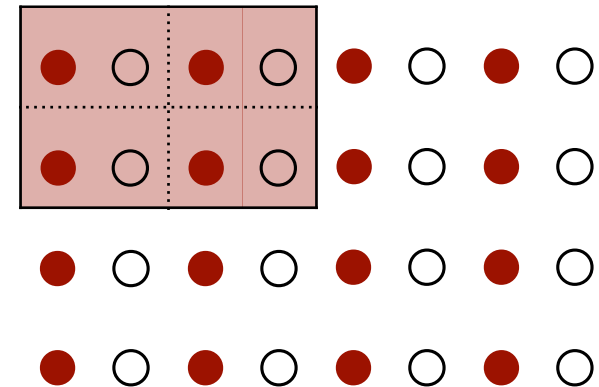


25

4:2:2

2

2

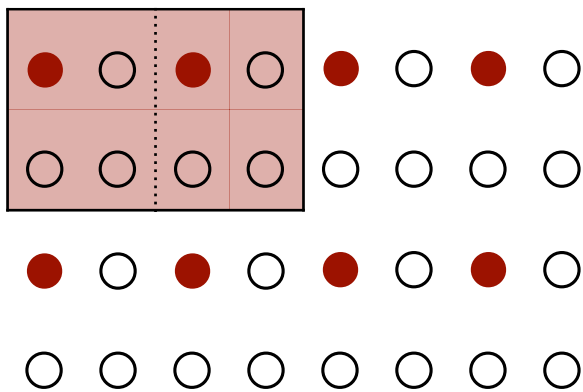


26

4:2:0

2

0

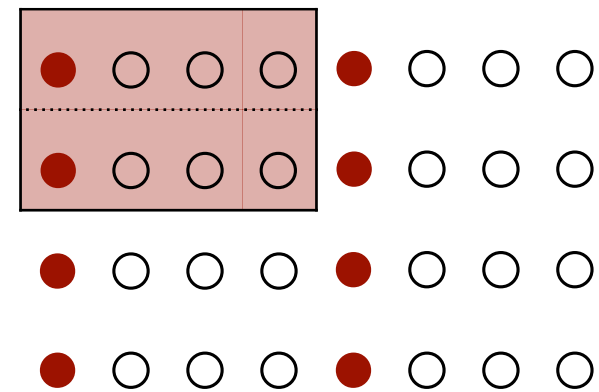


27

4:1:1

1

1



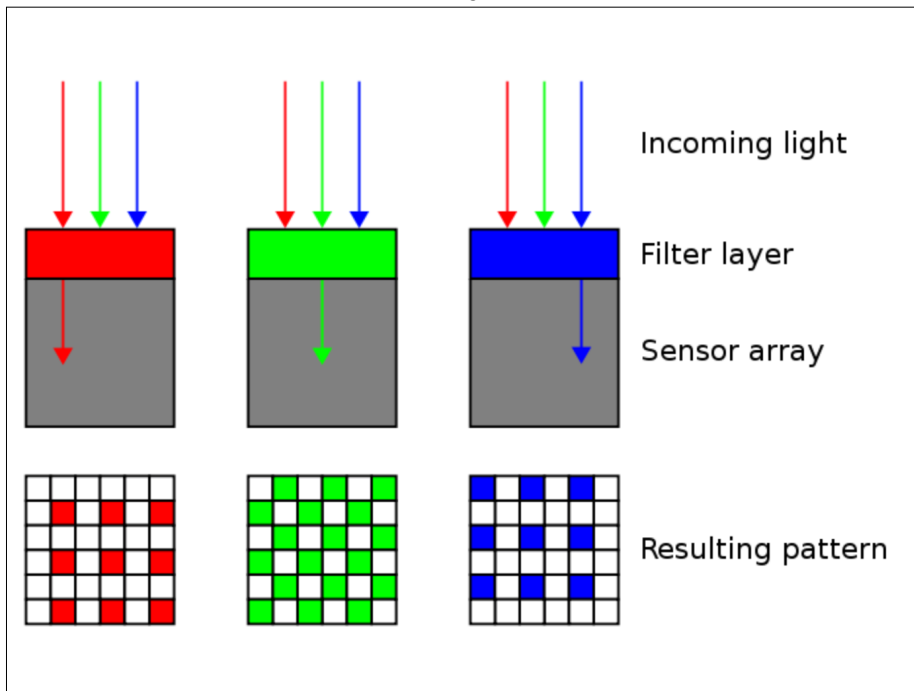
28

Born Compressed

- optimised for both image acquisition and postproduction

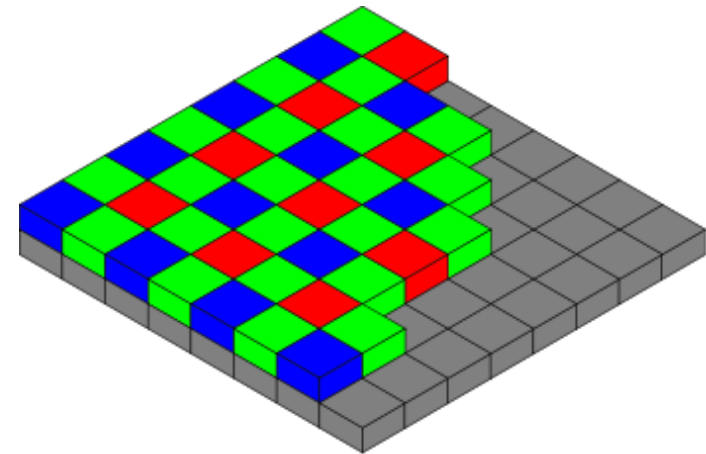
Examples: CineForm RAW, ProRes RAW

29

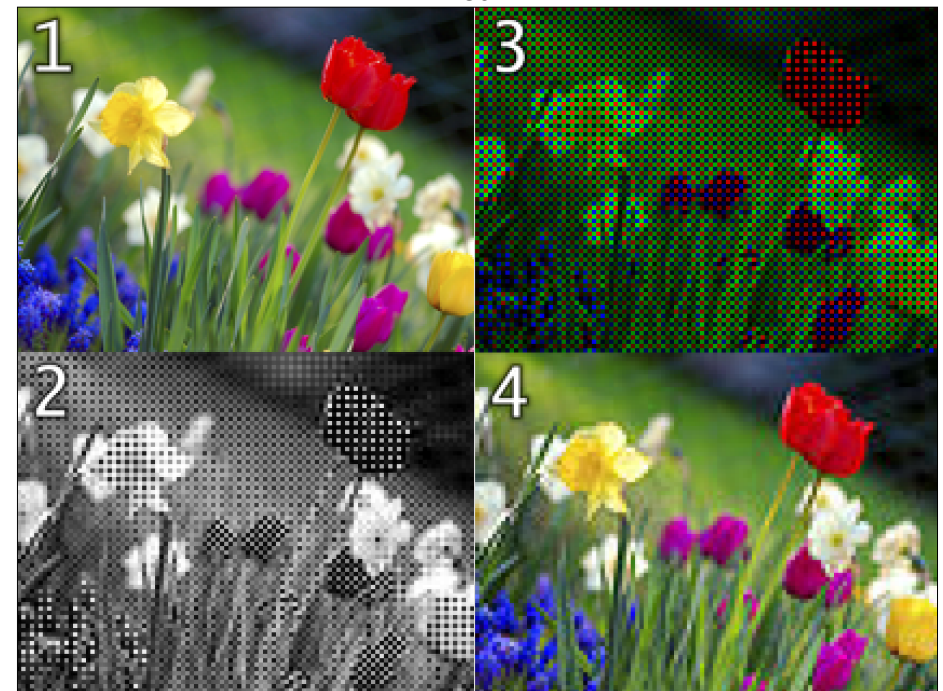


31

Bayer



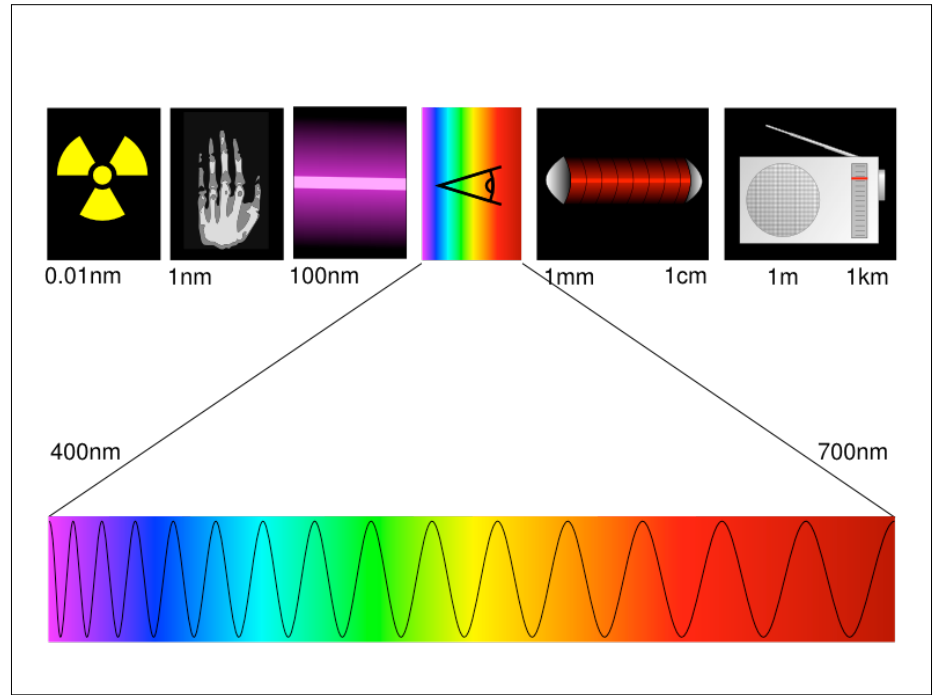
30



32

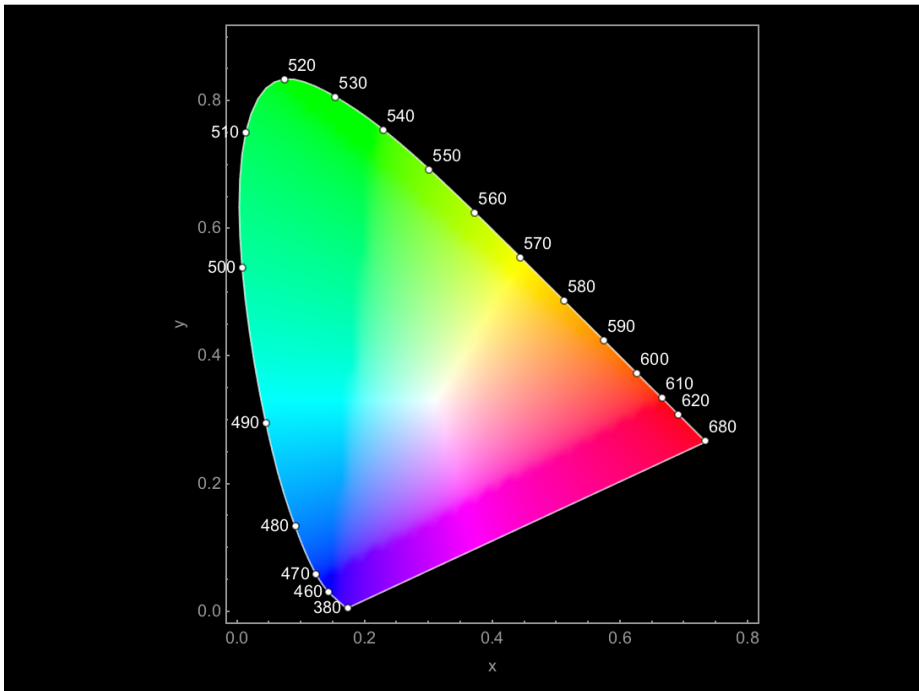
Illuminant

- D50
- D55
- D65
- D75

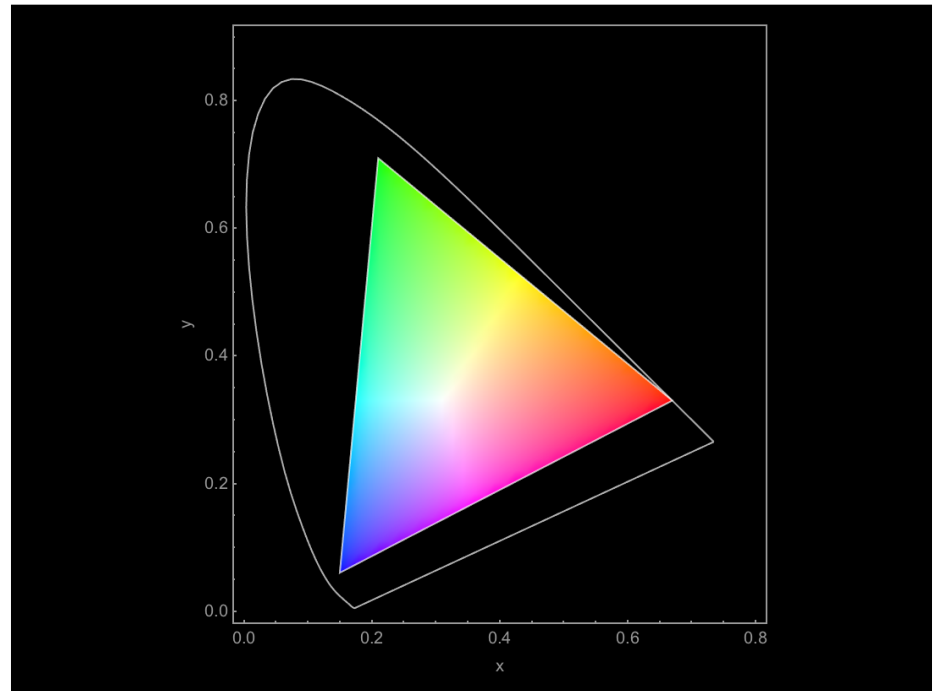


33

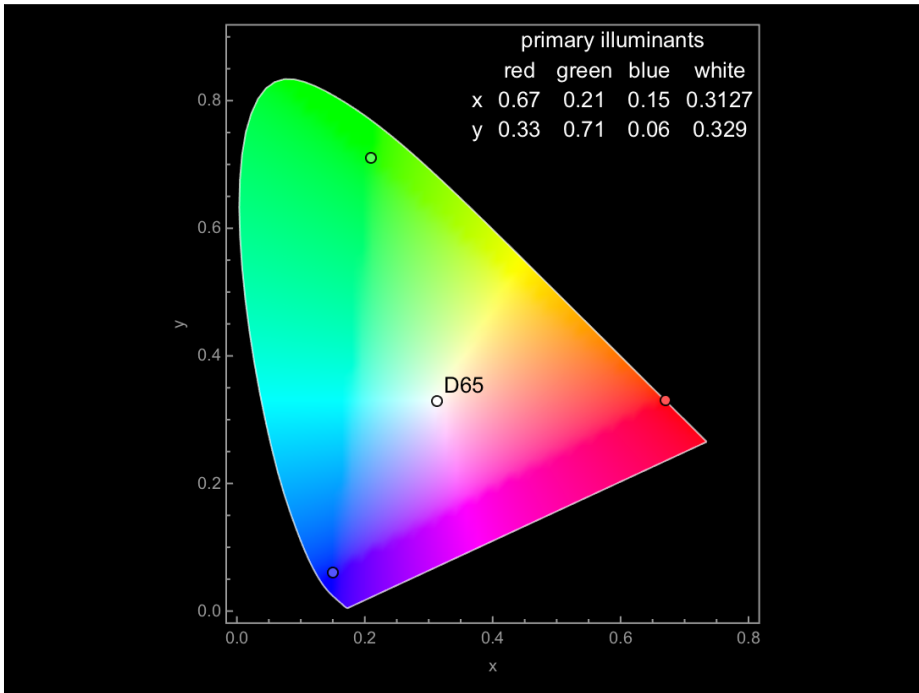
34



35



36



37

File Structure

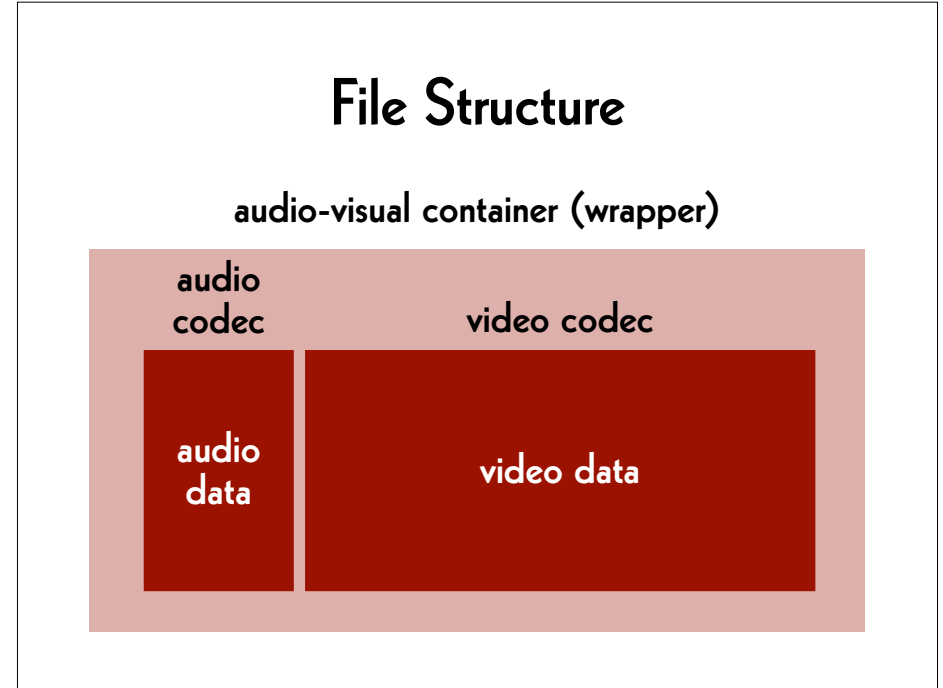
38

```

0111010100101010100010110101011110
0100110101010101010100001011101010
0111010100101010100010110101011110
0001110101010101010100001011101010
011010101001001010100010111010101111
0010101010101010000101110101010000
0111010100101010100010110101011110
0101010101010101000010111010100110
1001011101010010101010001011010101
1110010101010101010000101110101010
0111010100101010100010110101011110
0101010101010101001101010100000001
00101000101010101001010101010101

```

39



40

Audio-Visual Container

- MP4
- MOV
- AVI
- MXF
- Matroska (.mkv)
- Flash

41

Single Images

- folder
- TAR
- ZIP
- MXF
- Matroska (.mkv)
- CinemaDNG

42

Audio Codec

- WAVE
- BWF
- AAC
- MP3
- FLAC

43

Video Codec (Master)

- | images | streams |
|---------------|----------------|
| • TIFF | • 8 bit raw |
| • DPX | • 10 bit raw |
| • JPEG 2000 | • HuffYUV |
| • OpenEXR | • FFV1 |
| • DNG | |

44

Video Codec (Mezzanine)

- ProRes 422, ProRes 4444, ProRes RAW
- DNxHD, DNxHR
- CineForm RAW

45

Video Codec (Access)

- H.264, H.265 (HEVC), AV1

46

**Data is anything
but «raw».**

47

Audio Data

- pcm_s16le
- pcm_s24le
- pcm_s32le

48

Video Data

- rgb48le
- rgb24
- rgb72le
- bayer_bggr16le
- bayer_bggr24le
- yuv444p16le
- yuv422p10le
- uyvy422
- yuv420p
- yuv444p24le

49

File Formats

51

What is inside my DPX?

- log neg encoding
- log RGB encoding or quasi-log encoding
- gamma encoding or power function encoding
- scene-linear encoding

50

Principles

- **The archive must be able to handle the file formats it holds.**
- open source
- simple to use and well documented
- widely used by the community

52

Different Purposes

archive master format:

→ for preservation

mezzanine format:

→ for professional use in post-production

dissemination formats:

→ for widely spreading and easy access

53

Archive Master (Today)

film

- folder, TIFF, 2K, RGB, 4:4:4, 16 bit
- MXF, DPX, 2K, R'G'B', 4:4:4, 10 bit

video

- AVI, «raw», HD, Y'CbCr, 4:2:2, 10 bit
- Matroska, FFV1, HD, Y'CbCr, 4:2:2, 10 bit

audio

- BWF, 96 kHz, 24 bit
- FLAC, 96 kHz, 24 bit

55

Elena Rossi-Snook:

Archiving without access
isn't preservation,
it's hoarding.

54

Mezzanine (Today)

video

- ProRes 4444, 2K
- DNxHR, 2K
- ProRes 422 HQ, HD
- DNxHD 175x, HD

audio

- BWF, 48 kHz, 24 bit
- WAVE, 48 kHz, 24 bit

56

Dissemination (Today)

MP4

Video

- H.264, SD, yuv420p, «lossy»
- H.264, HD, yuv420p, «lossy»

Sound

- AAC, 44.1 kHz, 16 bit
- AAC, 48 kHz, 16 bit

57

Archive Master and Mezzanine

film

- Matroska, FFV1, 2K, «RGB», 4:4:4, 16 bit

video

- Matroska, FFV1, HD, Y'CbCr, 4:2:2, 10 bit

audio

- Matroska, FLAC, 96 kHz, 24 bit

58

Access

WebM (a subset of Matroska)

Video

- «H.265», HD, yuv420p
- AV1, HD, yuv420p

Sound

- FLAC, 48 kHz, 16 bit

59

Reading

Reto Kromer: **Matroska and FFV1: One File Format for Film and Video Archiving?**, in «Journal of Film Preservation», n. 96 (April 2017), FIAF, Brussels, Belgium, p. 41–45

→ https://retokromer.ch/publications/JFP_96.html

60

Pros & Cons

61

container:

- folder
- TAR
- ZIP
- MXF
- Matroska

codec:

- TIFF
- DPX
- JPEG 2000
- FFV1
- OpenEXR
- CineForm RAW
- ProRes RAW

62

	avantages	disavantages
TIFF DPX OpenEXR	data easier to process	bigger files
JPEG 2000 FFV1	smaller files	data complexer to process

63

Transformations

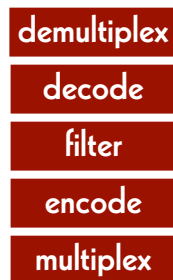
64

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.140251 \\ 1 & -0.393931 & -0.580809 \\ 1 & 2.028398 & 0 \end{bmatrix} \cdot \begin{bmatrix} Y'_{601} \\ U \\ V \end{bmatrix}$$

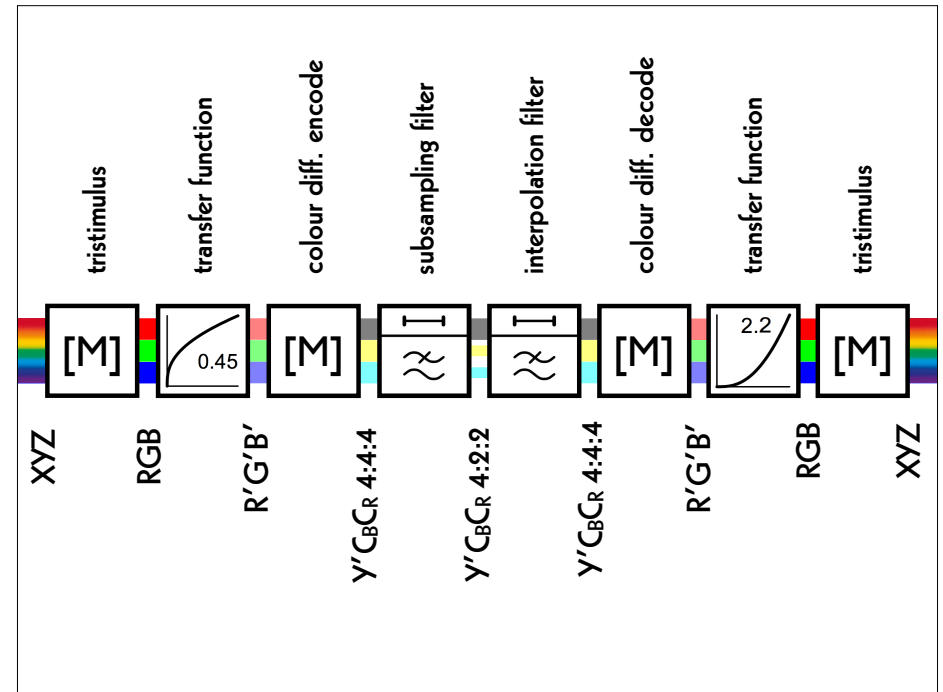
$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} 1 & 0.956295 & 0.621025 \\ 1 & -0.272558 & -0.646709 \\ 1 & -1.104744 & 1.701157 \end{bmatrix} \cdot \begin{bmatrix} Y'_{601} \\ I \\ Q \end{bmatrix}$$

65

Data Transformations

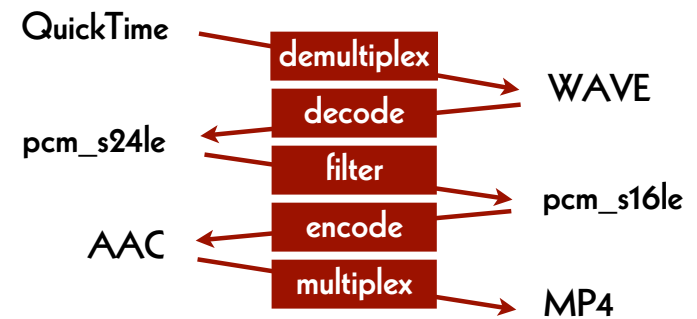


67



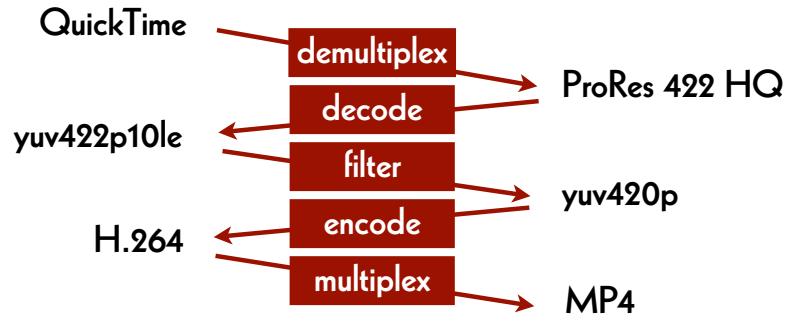
66

Audio Exemple



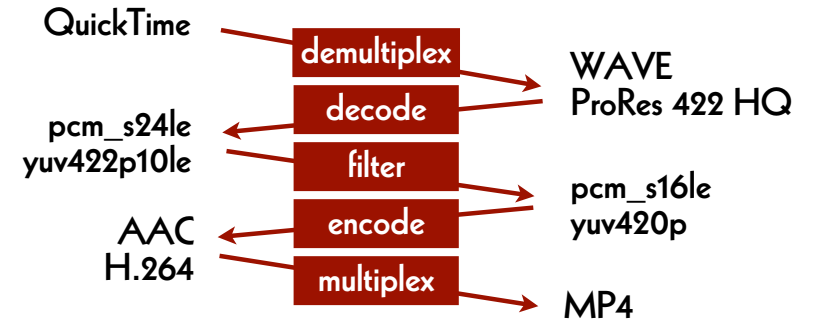
68

Video Exemple



69

Audio-Visual Exemple



70

Acknowledgements

- Swiss Federal Institute of Technology
- Massachusetts Institute of Technology
- Kinemathek Lichtspiel, Bern

- Charles Poynton
- Dave Rice & Misty De Meo
- Agathe Jarczyk & David Pfluger

71

AV Preservation by reto.ch

zone industrielle Le Trési 3
1028 Préverenges
Switzerland

Web: reto.ch
Twitter: @retoch
Email: info@reto.ch



72