

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
1–4 and 9 June 2021

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
- compression

4

Sampling

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

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 and compression
- 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 ($2^{24} = 16\,777\,216$)

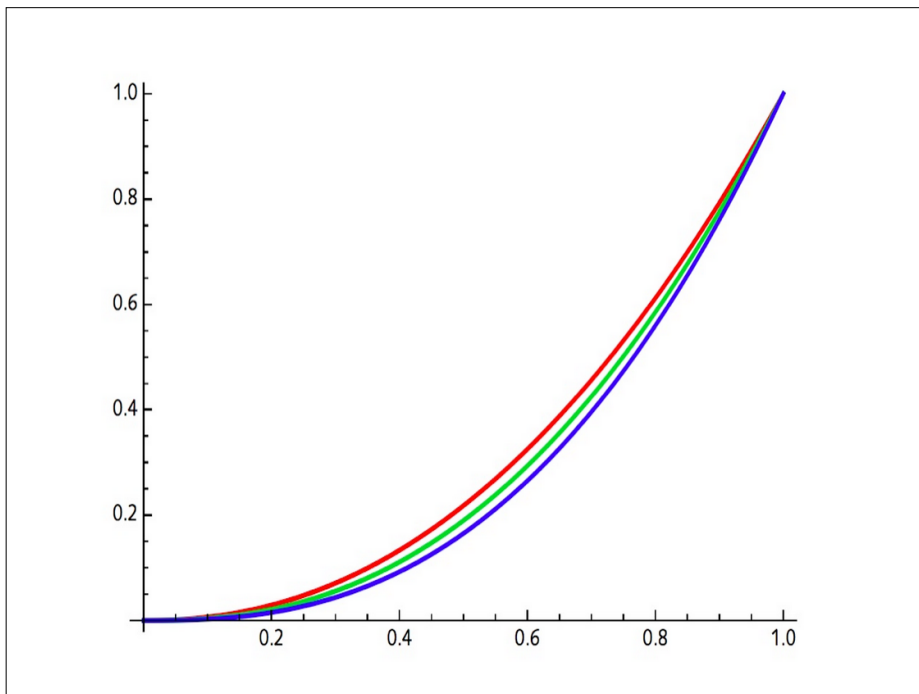
11

Linear, Power, Logarithmic

“medium grey”

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

12

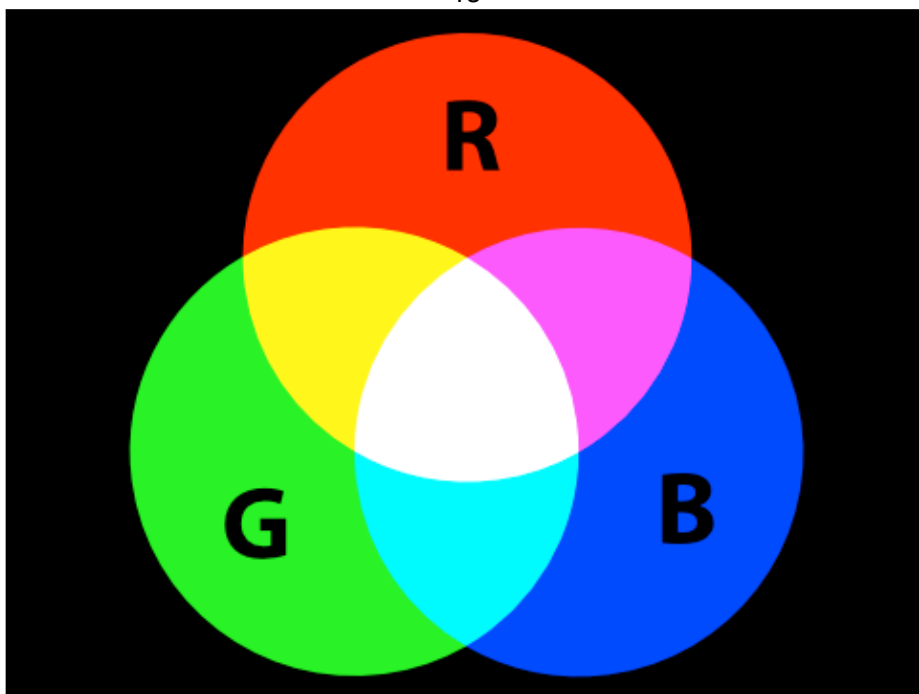


13

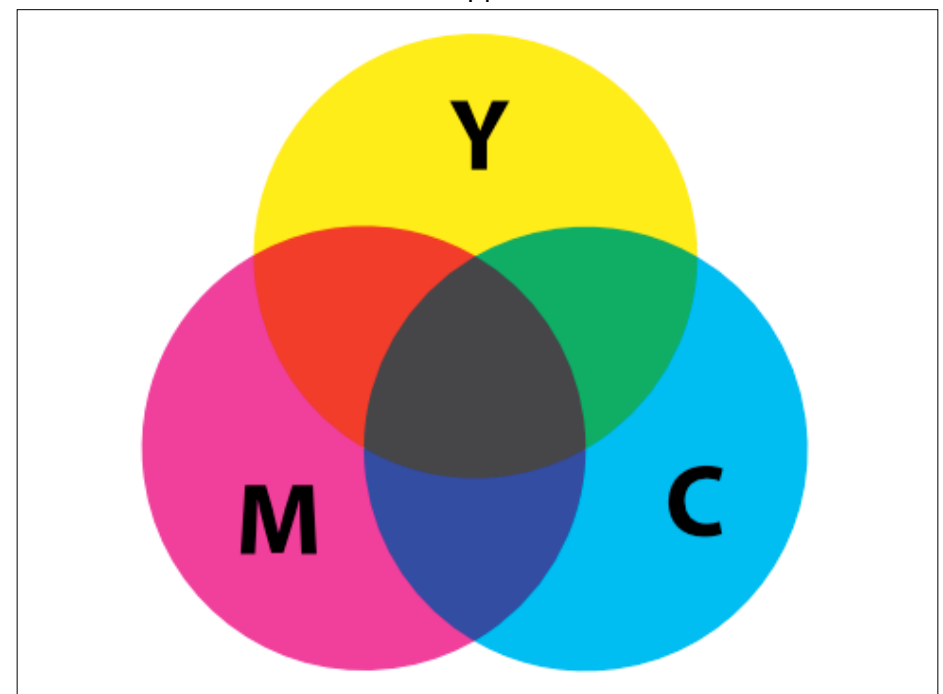
Colour Model

- XYZ, $L^*a^*b^*$
- RGB / $R'G'B'$ / CMY / $C'M'Y'$
- $Y'IQ$ / $Y'UV$ / $Y'D_BD_R$
- $Y'C_{BCR}$ / $Y'CoCG$
- $Y'P_{BP_R}$

14



15



16

$$\begin{pmatrix} R' \\ G' \\ B' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1.396523 \\ 1 & -0.342793 & -0.711348 \\ 1 & 1.765078 & 0 \end{pmatrix} \begin{pmatrix} Y' \\ C_B \\ C_R \end{pmatrix}$$

$$\begin{pmatrix} Y' \\ C_B \\ C_R \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.168074 & -0.329965 & 0.498039 \\ 0.498039 & -0.417947 & -0.080992 \end{pmatrix} \begin{pmatrix} R' \\ G' \\ B' \end{pmatrix}$$

17

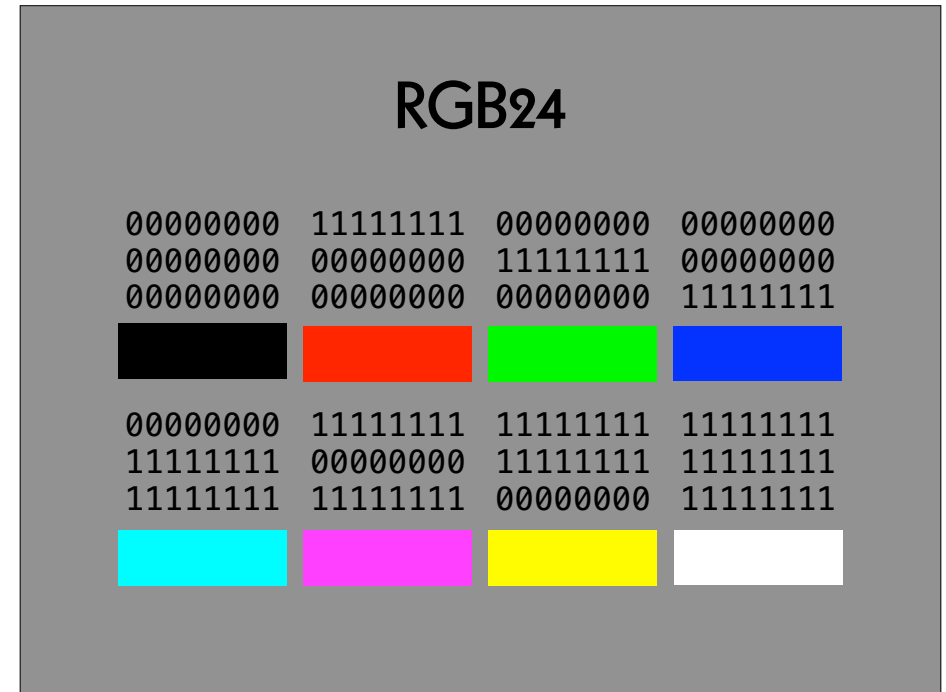


19

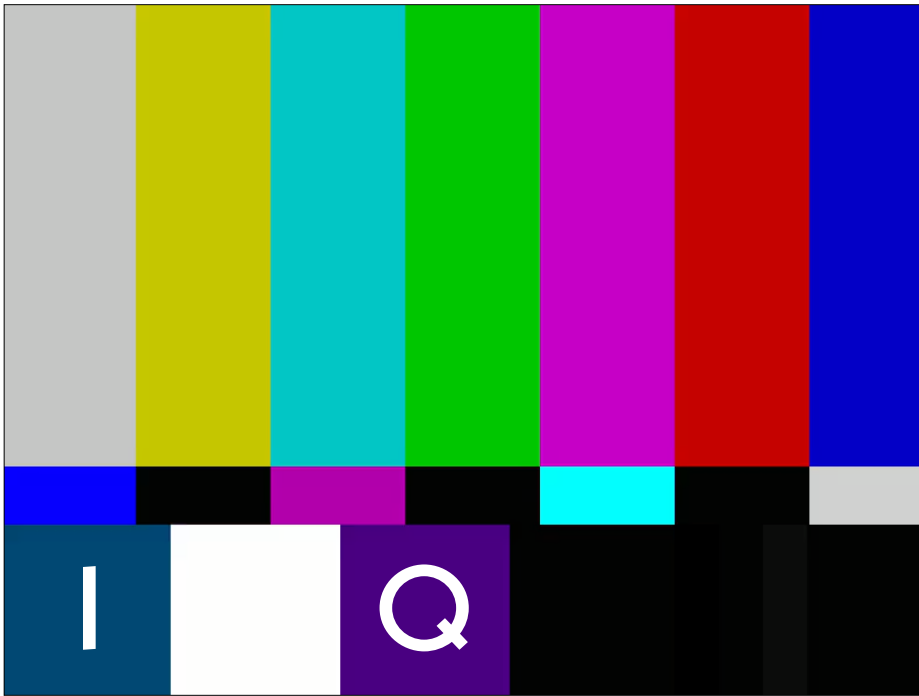
$$\begin{pmatrix} R' \\ G' \\ B' \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} Y' \\ C_O \\ C_G \end{pmatrix}$$

$$\begin{pmatrix} Y' \\ C_O \\ C_G \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \end{pmatrix} \begin{pmatrix} R' \\ G' \\ B' \end{pmatrix}$$

18



20



21

Uncompressed

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

Examples: TIFF, DPX, DNG, OpenEXR

23

Compression

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

22

Lossless Compression

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

Examples: JPEG 2000, FFV1

24

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), H.266 (VVC); AV1

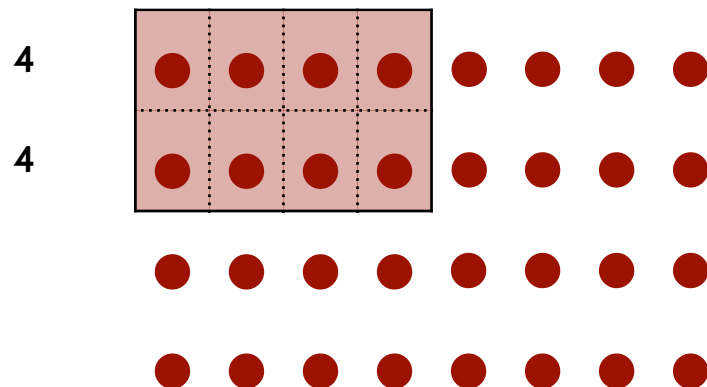
25

Chroma Subsampling

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

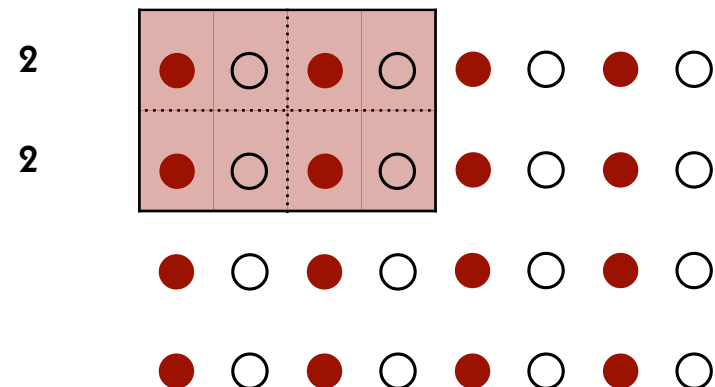
26

4:4:4



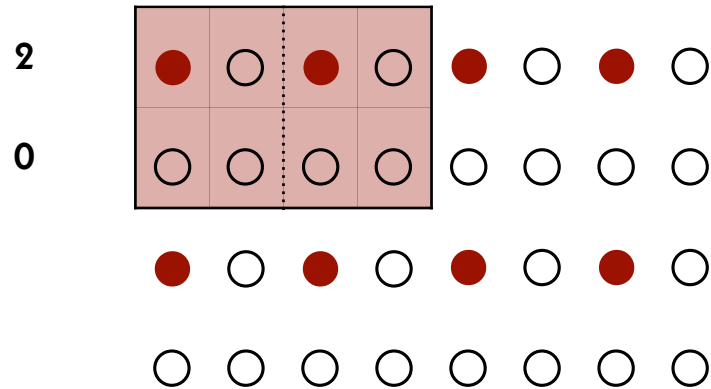
27

4:2:2



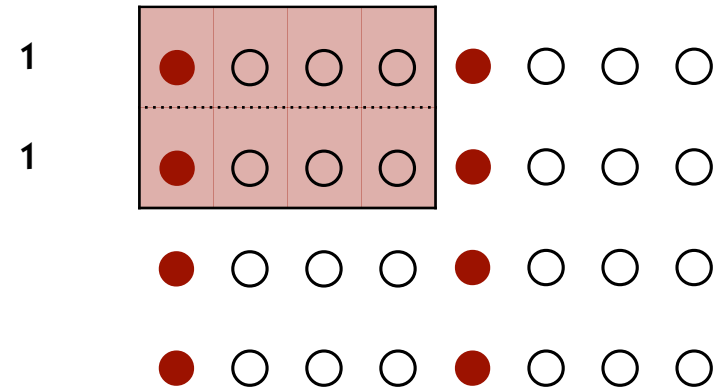
28

4:2:0



29

4:1:1



30

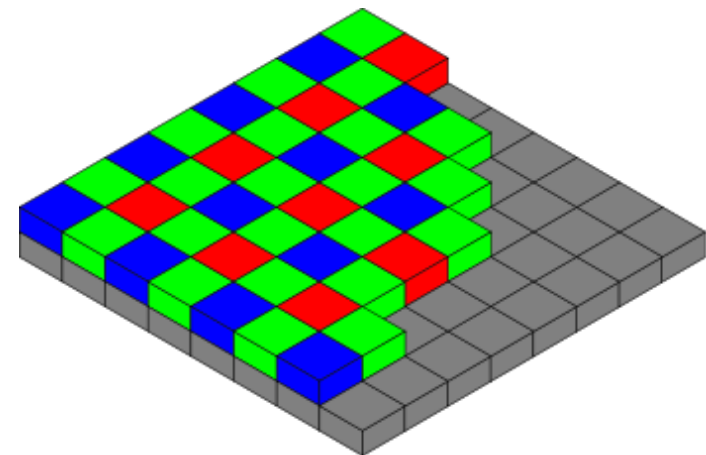
Born Compressed

- optimised for both image acquisition and postproduction

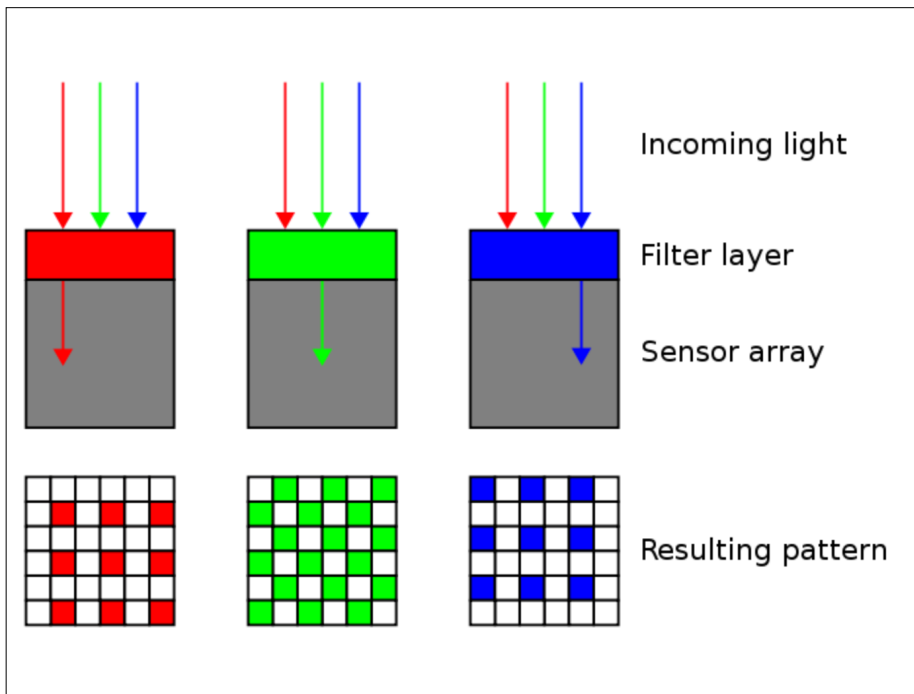
Examples: CineForm RAW, ProRes RAW, Blackmagic RAW

31

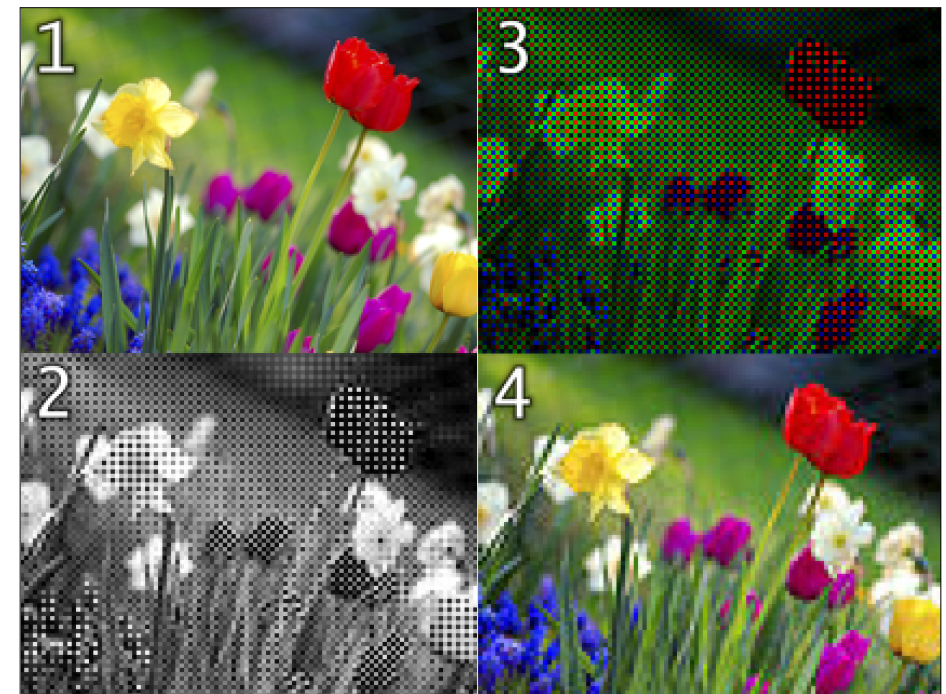
Bayer



32



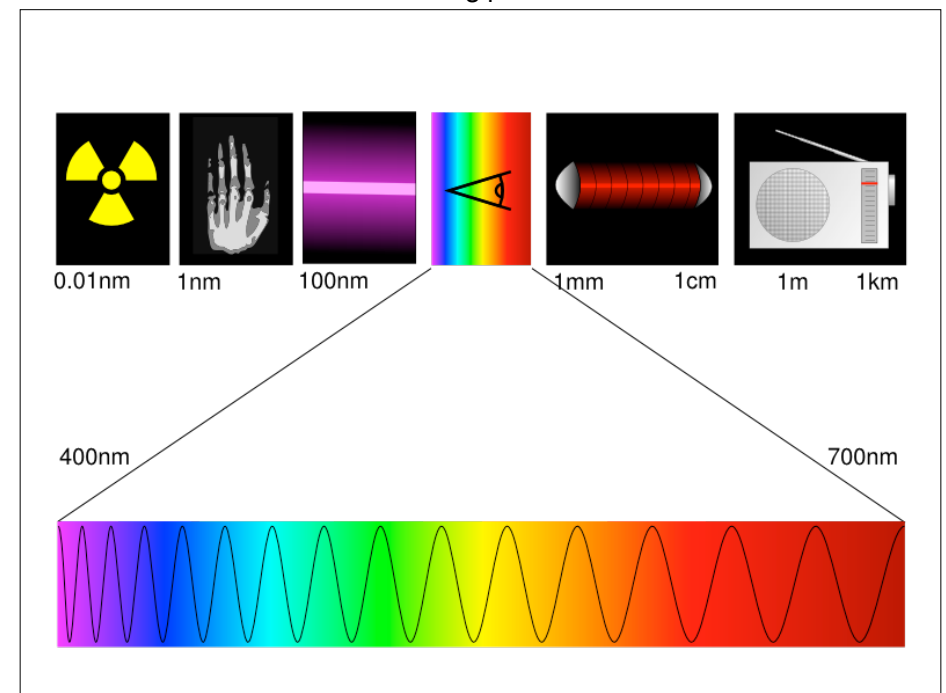
33



34

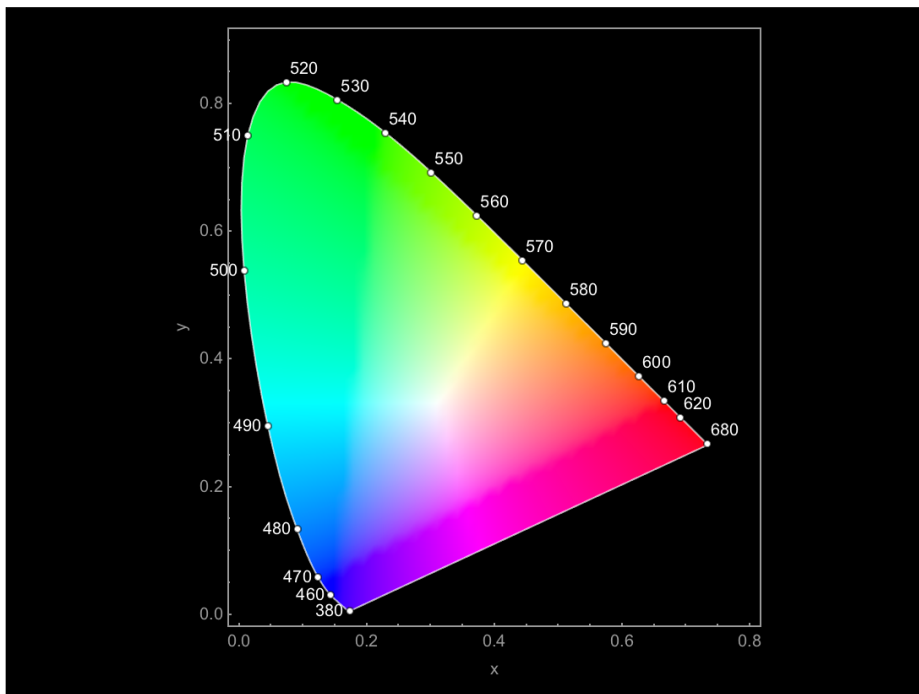
Illuminant

- D50
- D55
- D65
- D75

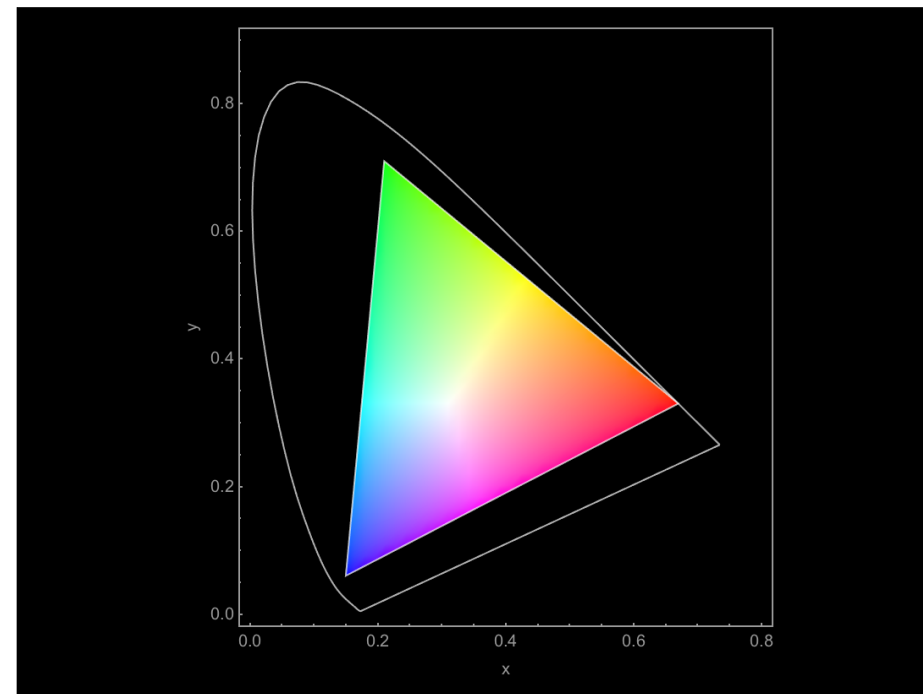


36

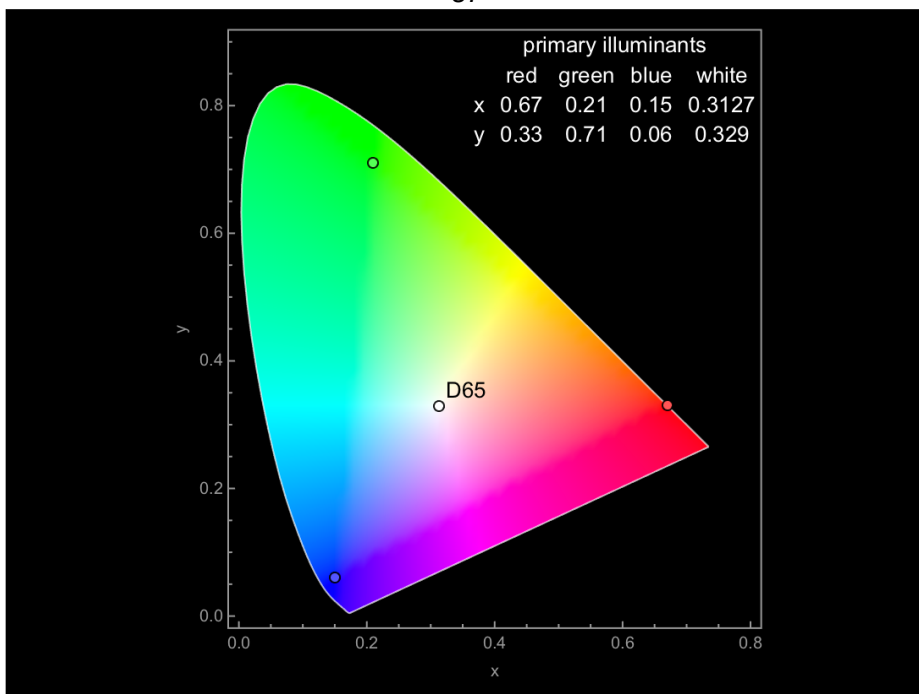
35



37



38



39

File Structure

40

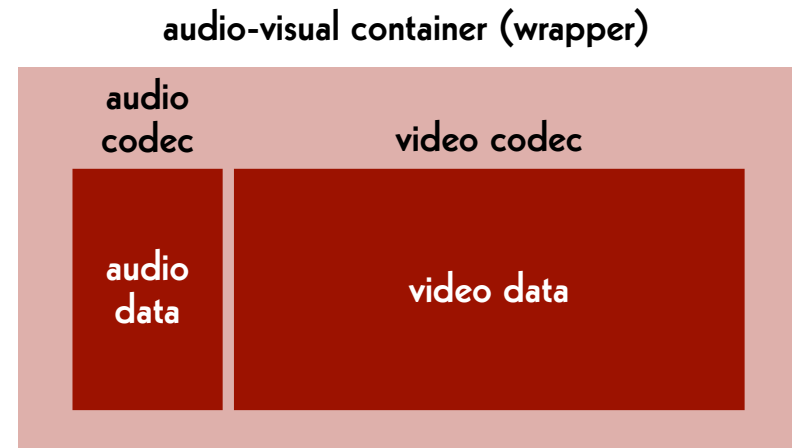
```

011101010010101010100010110101011110
01001101010101010101010100001011101010
011101010010101010100010110101011110
00011101010101010101010100001011101010
011010101001010101010001011010101111
001010101010101010000101110101010000
011101010010101010100010110101011110
010101010101010101000010111010100110
100101110101001010101010001011010101
1110010101010101010000101110101010
011101010010101010100010110101011110
0101010101010101001101010100000001
0010100010101010101001010101010101

```

41

File Structure



42

Audio-Visual Container

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

43

Single Images

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

44

Audio Codec

- WAVE
- BWF
- AAC
- MP3
- FLAC

45

Video Codec (Master)

images

- TIFF
- DPX
- JPEG 2000
- OpenEXR
- DNG

streams

- 8 bit raw
- 10 bit raw
- HuffYUV
- FFV1

46

Video Codec (Mezzanine)

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

47

Video Codec (Access)

- H.264 (AVC), H.265 (HEVC), H.266 (VVC)
- AV1

48

Data is anything
but “raw”.

49

Audio Data

- pcm_s16le
- pcm_s24le
- pcm_s32le

50

Video Data

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

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

52

File Formats

53

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

54

Different Purposes

archive master format:

→ for preservation and archiving

mezzanine format:

→ for professional use in post-production

dissemination formats:

→ for widely spreading and easy access

55

Elena Rossi-Snook:

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

56

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

audio

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

57

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

58

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

59

Archive Master and Mezzanine

film

- Matroska, FFV1, 2K, R'G'B' 4:4:4, 12–16 bit

video

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

audio

- Matroska, FLAC, 96 kHz, 24 bit

60

Access

WebM (a subset of Matroska)

Video

- H.265, "HD", yuv420p
- H.266, "HD", yuv420p
- AV1, "HD", yuv420p

Sound

- FLAC, 48 kHz, 16 bit

61

container:

- folder
- TAR
- ZIP
- MXF
- Matroska

codec:

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

63

Pros & Cons

62

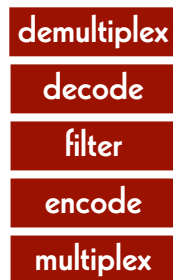
	advantages	disadvantages
TIFF DPX OpenEXR	data easier to process	bigger files
JPEG 2000 FFV1	smaller files	data complexer to process

64

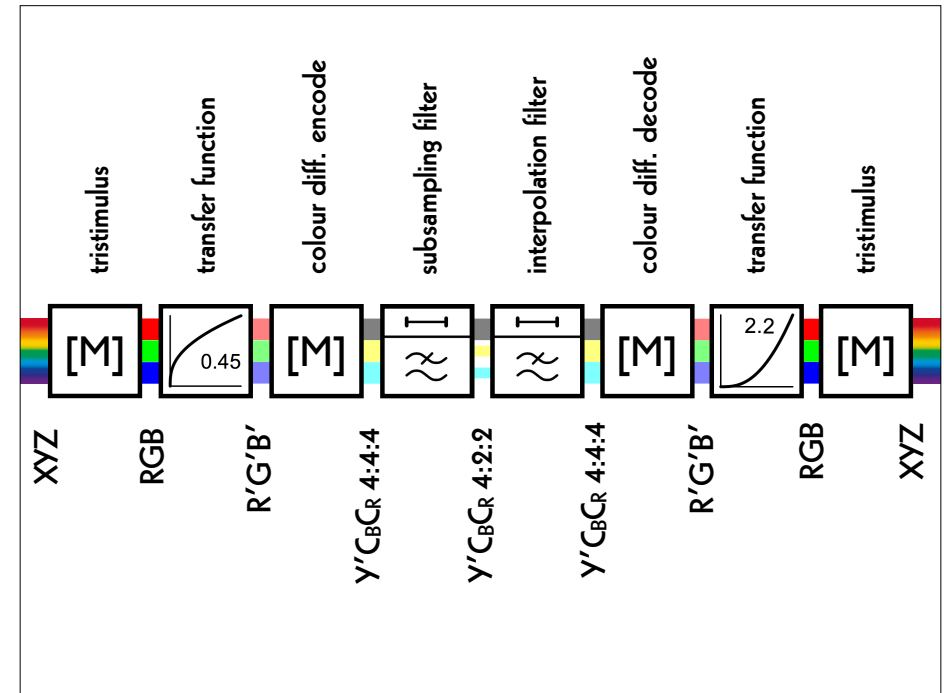
Transformations

65

Data Transformations

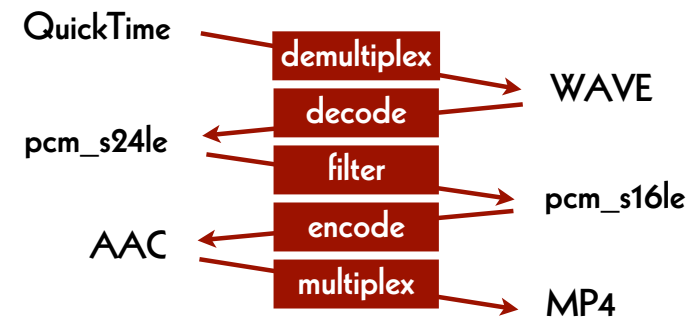


67



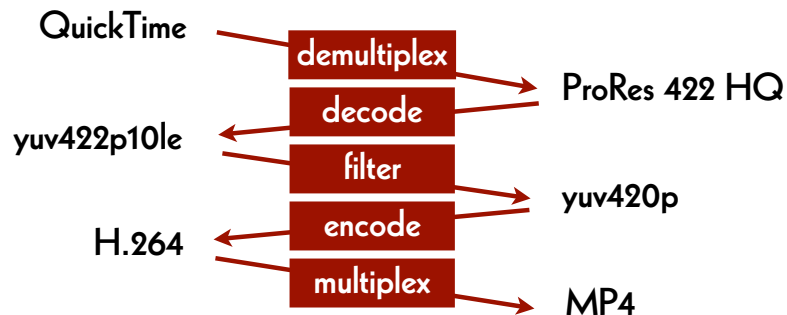
66

Audio Example



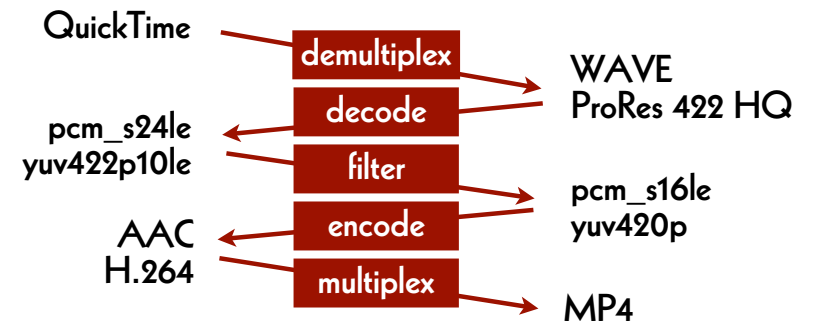
68

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