On Audio-Visual File Formats

Reto Kromer • AV Preservation by reto.ch

Akademie der bildenden Künste
Medienkunst:
Technologie und Erhaltungsstrategien III
Wien, 7.–10. Jänner 2020

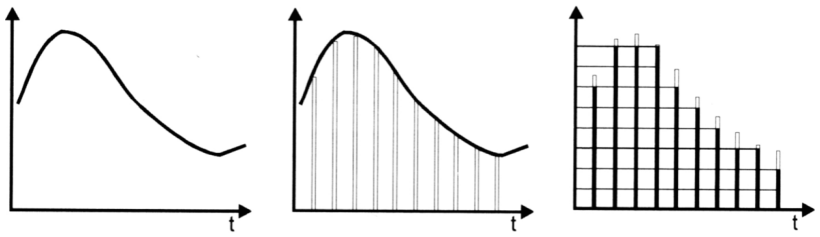
Summary

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

Digital Audio

Digital Audio

- sampling
- quantisation



digitisation = sampling + quantisation

Sampling

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

Quantisation

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

Digital Video

Digital Video

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

Resolution

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

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

Linear, Power, Logarithmic

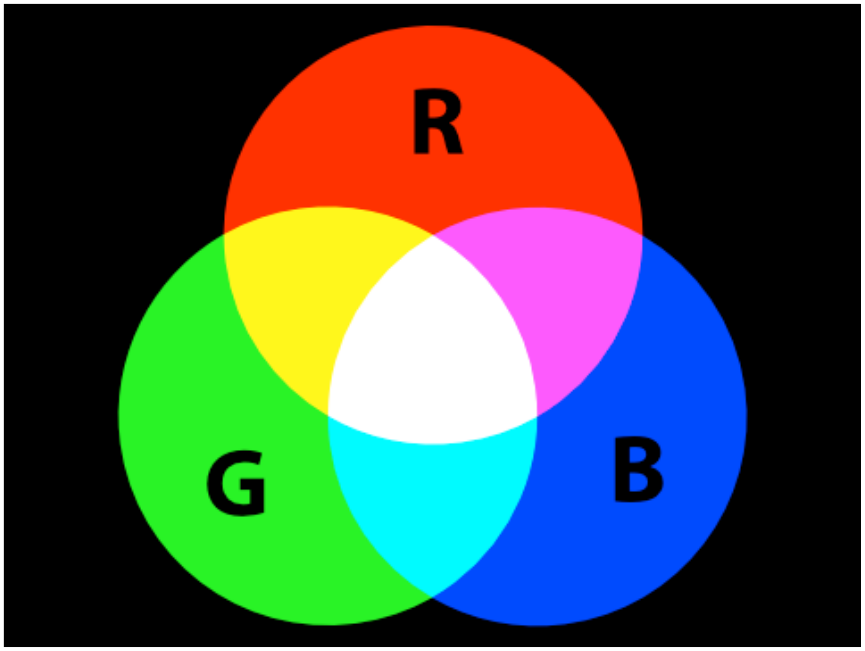
«medium grey»

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

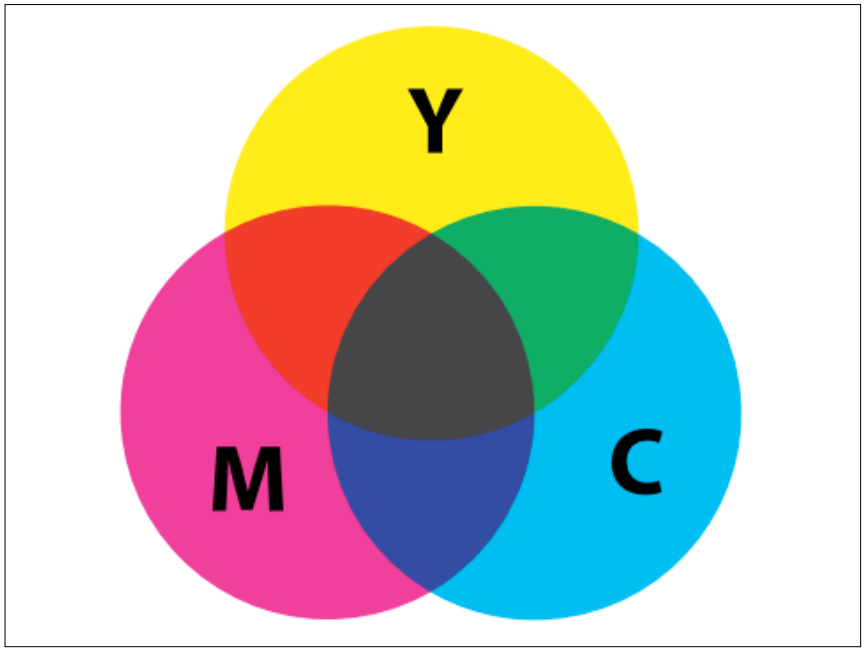
Colour Model

- 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′C_OC_G
- Y′P_BP_R

13



14



15

16

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
```


Compression

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

Uncompressed

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

Examples: TIFF, DPX, DNG, OpenEXR

Lossless Compression

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

Examples: JPEG 2000, FFV1

Lossy Compression

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

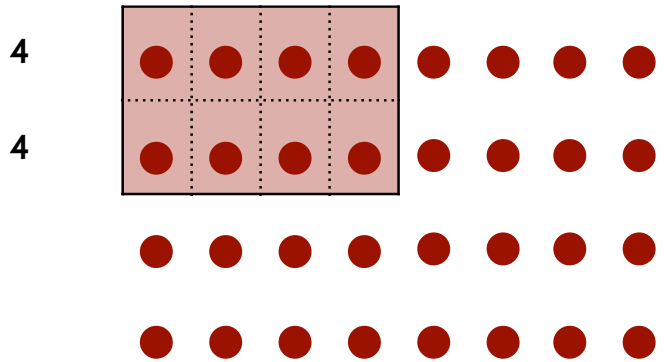
Examples (mezzanine): ProRes 422, ProRes 4444, DNxHD, DNxHR

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

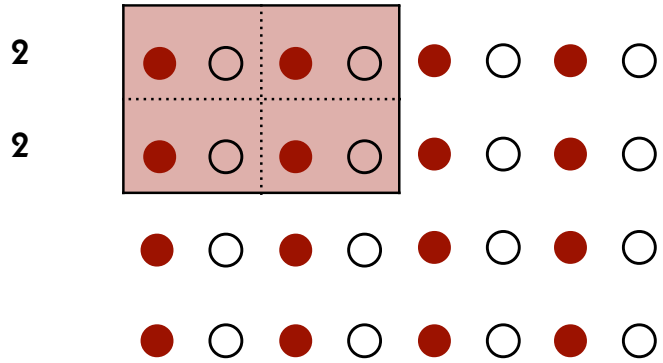
Chroma Subsampling

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

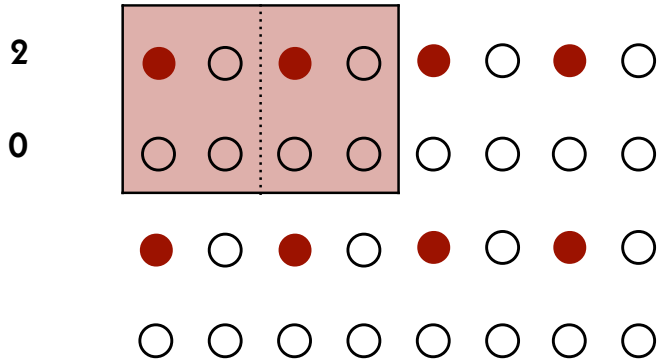
4:4:4



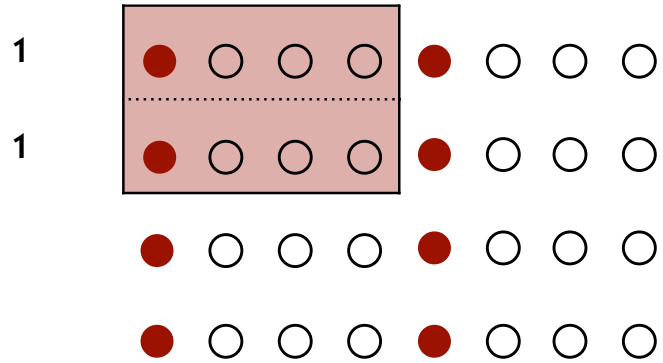
4:2:2



4:2:0



4:1:1

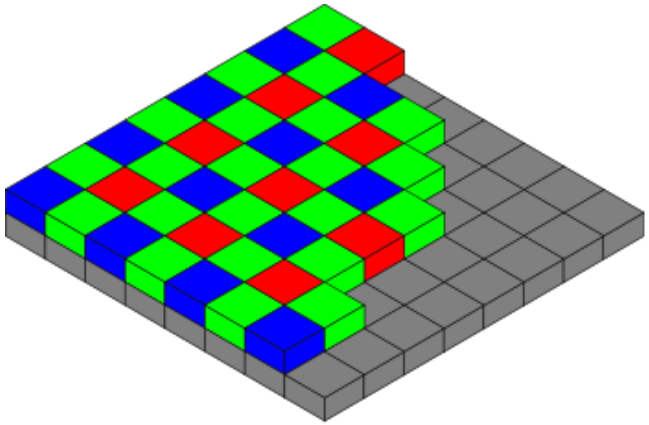


Born Compressed

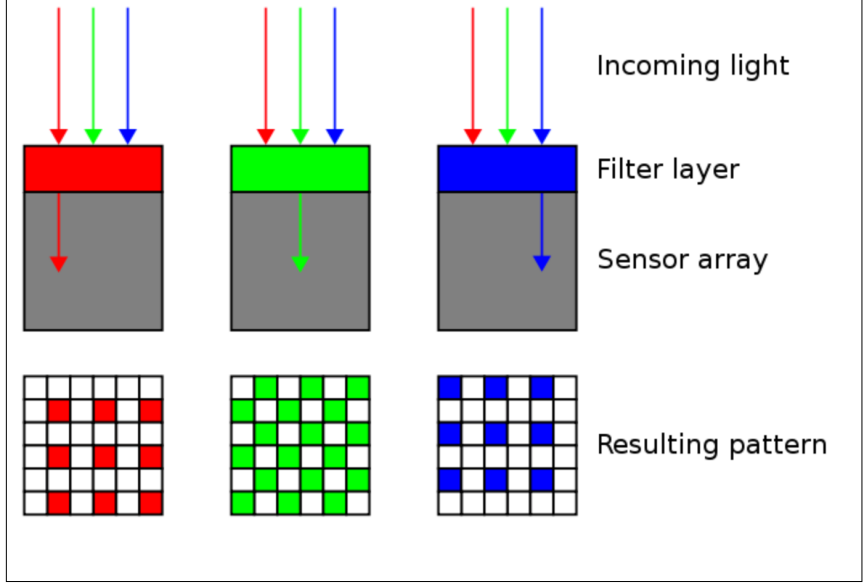
- optimised for both image acquisition and postproduction

Examples: CineForm RAW, ProRes RAW

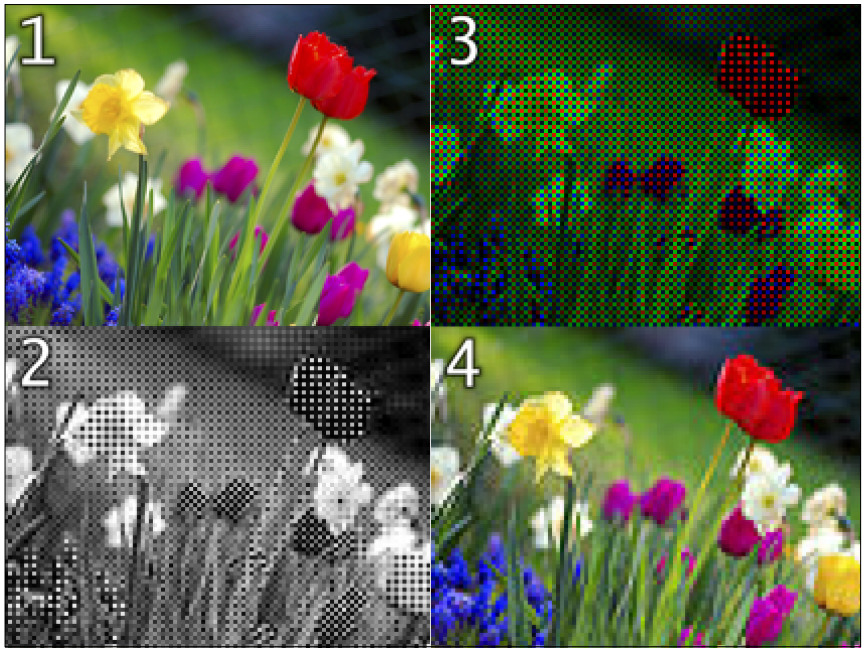
Bayer



29



30

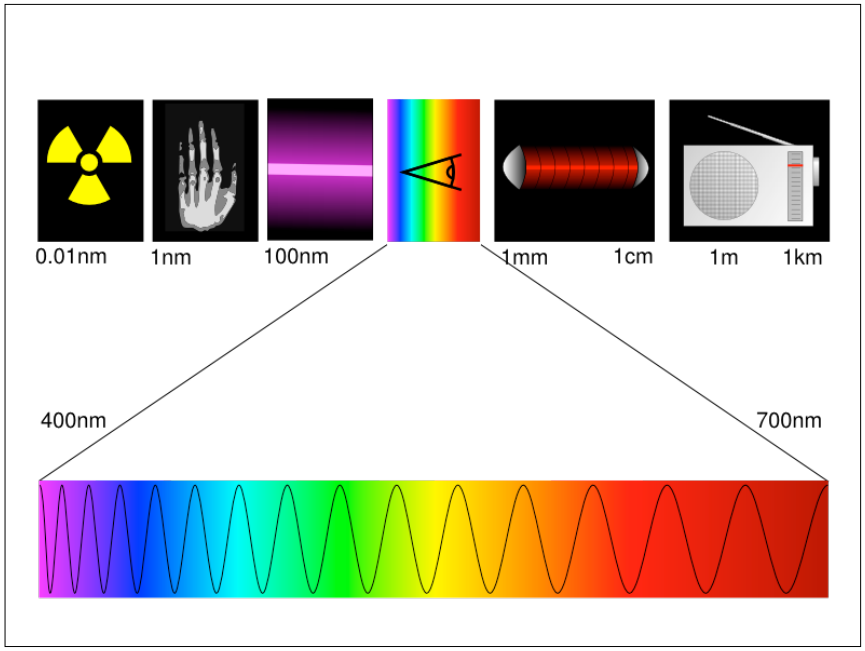


31

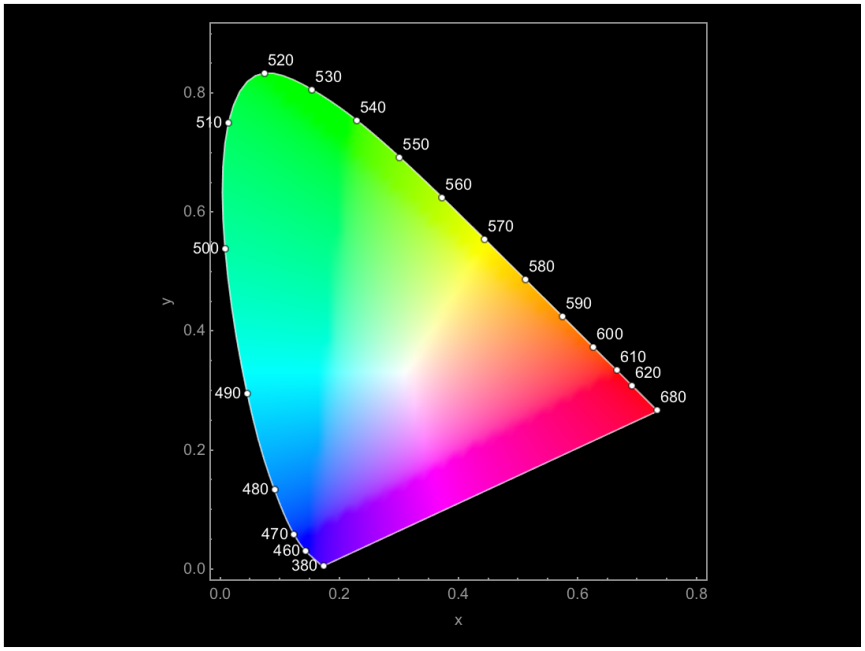
Illuminant

- D50
- D55
- D65
- D75

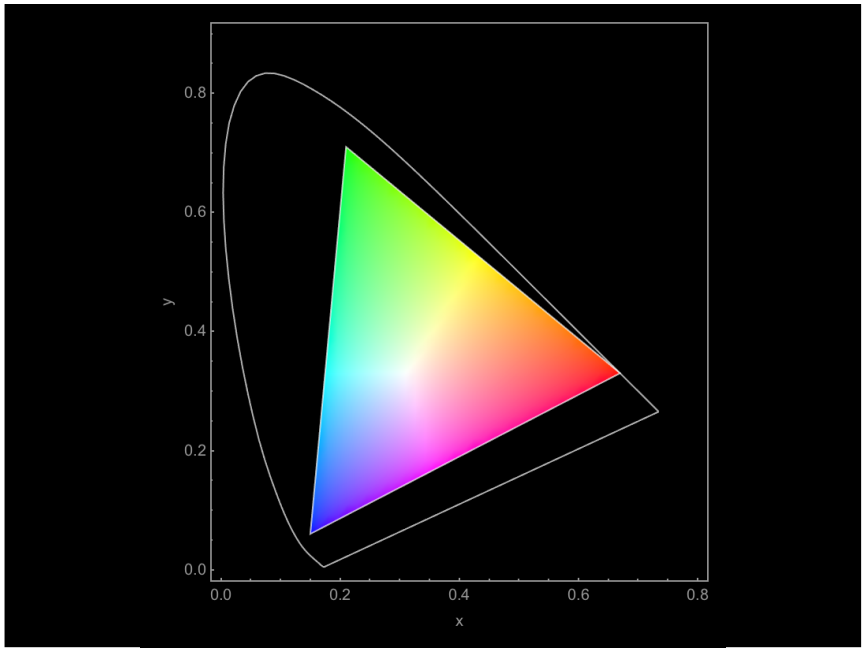
32


0.01nm
1nm
100nm
1mm
1cm
1m
1km
400nm
700nm

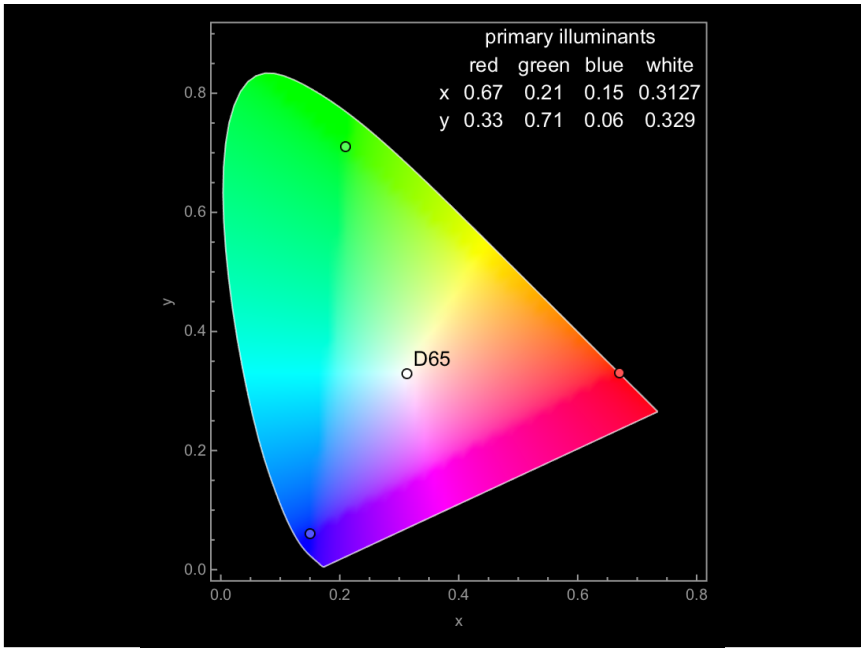
33


520
530
540
550
560
570
580
590
600
610
620
680
510
500
490
480
470
460
380
0.0
0.2
0.4
0.6
0.8
x
y

34


0.0
0.2
0.4
0.6
0.8
x
y

35


primary illuminants
red green blue white
x 0.67 0.21 0.15 0.3127
y 0.33 0.71 0.06 0.329
D65
0.0
0.2
0.4
0.6
0.8
x
y

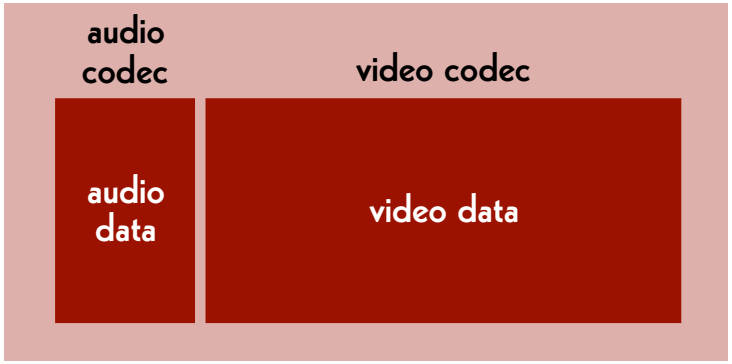
36

File Structure


```
01110101001010101000101101010111101
01001101010101010101000010111010101
01110101001010101000101101010111101
00011101010101010101000010111010101
01101010100101010100010110101011111
00101010101010100001011101010100001
01110101001010101000101101010111101
01010101010101010000101110101001101
10010111010100101010100010110101011
11100101010101010100001011101010101
01110101001010101000101101010111101
01010101010101010011010101000000011
00101000101010101010010101010101011
```


File Structure

audio-visual container (wrapper)



Audio-Visual Container

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

Single Images

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

Audio Codec

- WAVE
- BWF
- FLAC
- AAC
- MP3

Video Codec (Master)

images

- TIFF
- DPX
- JPEG 2000
- OpenEXR
- DNG

streams

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

Video Codec (Mezzanine)

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

Video Codec (Access)

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

Data is anything
but «raw».

Audio Data

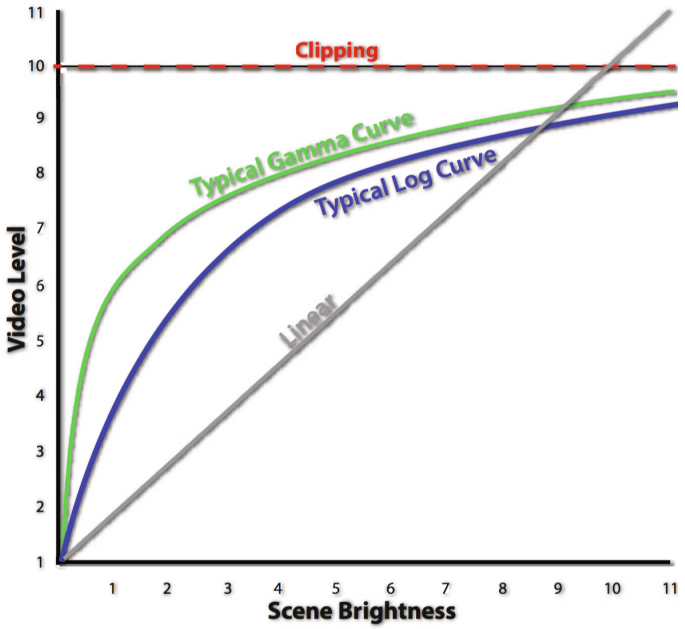
- pcm_s16le
- pcm_s24le
- pcm_s32le

Video Data

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

What is inside my DPX?

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



File Formats

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

Different Purposes

archive master format:
➔ for preservation

mezzanine format:
➔ for professional use in post-production

dissemination formats:
➔ for widely spreading and easy access

Elena Rossi-Snook:

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

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

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

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

Archive Master and Mezzanine

film

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

video

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

audio

- Matroska, FLAC, 96 kHz, 24 bit

Access

WebM (a subset of Matroska)

Video

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

Sound

- FLAC, 48 kHz, 16 bit

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

Pros & Cons

container:

- folder
- TAR
- ZIP
- MXF
- Matroska

codec:

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

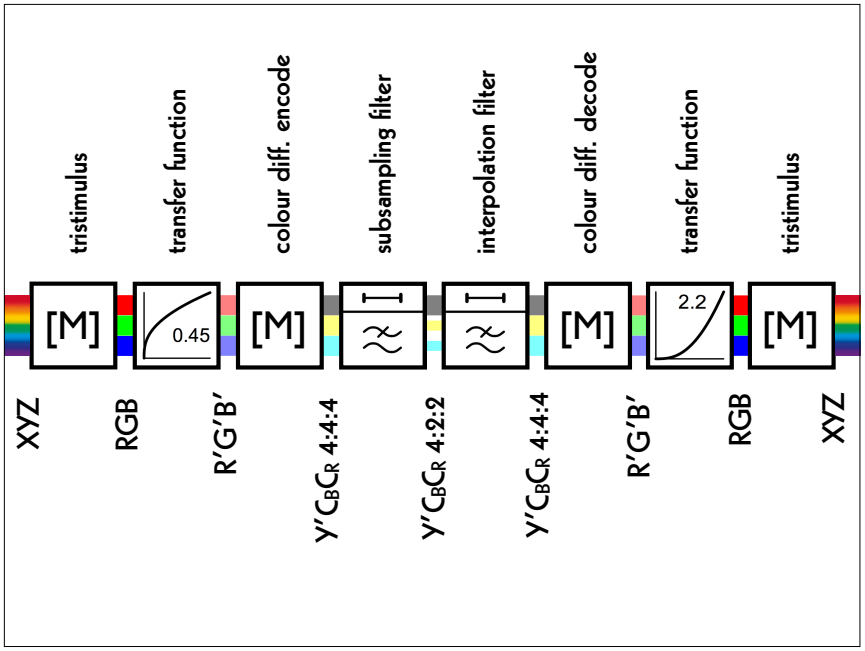
	avantages	disavantages
TIFF DPX OpenEXR	data easier to process	bigger files
JPEG 2000 FFV1	smaller files	data complexer to process

Transformations

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

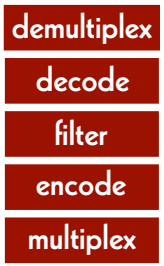
$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} 1 & 0.956295 & 0.621025 \\ 1 & -0.272558 & -0.646709 \\ 1 & -1.104744 & 1.701157 \end{bmatrix} \cdot \begin{bmatrix} Y'_{601} \\ I \\ Q \end{bmatrix}$$

65



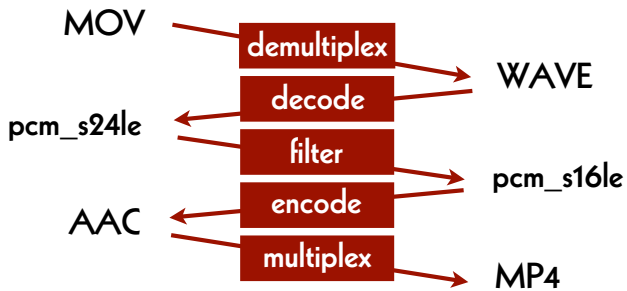
66

Data Transformations



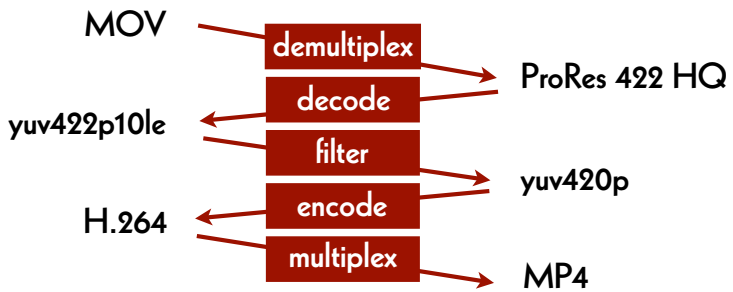
67

Audio Exemple

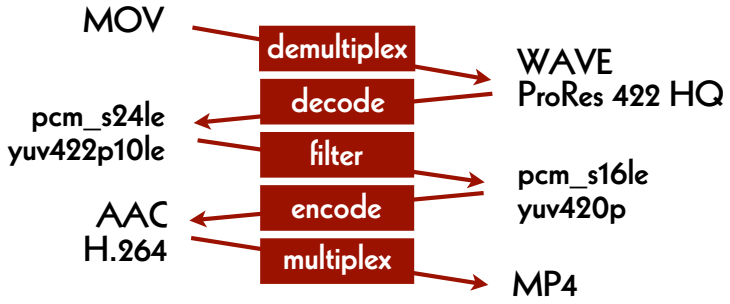


68

Video Exemple



Audio-Visual Exemple



Acknowledgements

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

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

AV Preservation by reto.ch

chemin du Suchet 5
1024 Écublens
Switzerland

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