On Audio-Visual File Formats

Reto Kromer • AV Preservation by reto.ch

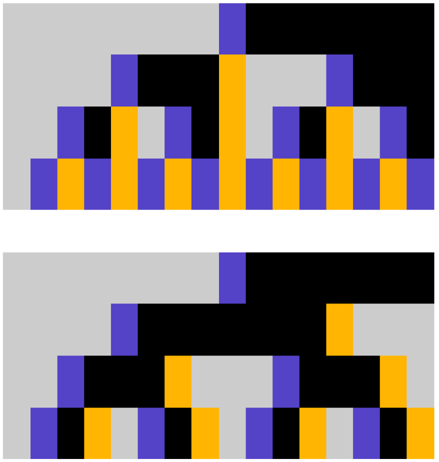
Using FFmpeg in a film archive
FIAF, on-line, 24 November 2022

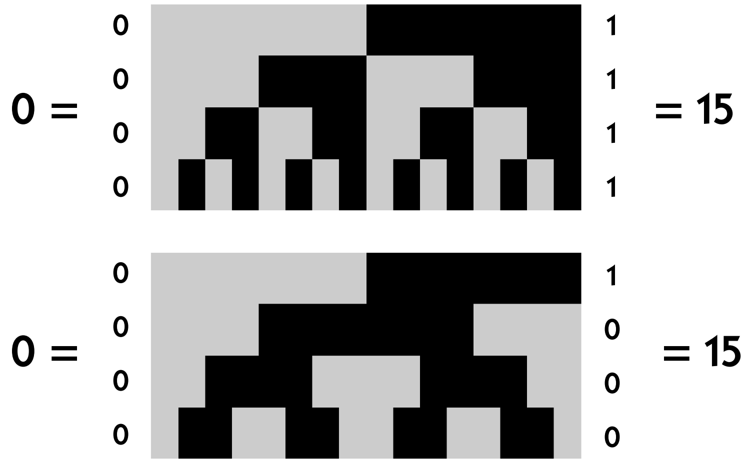
Summary

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

Digital Audio

Frank Gray (1887–1969)

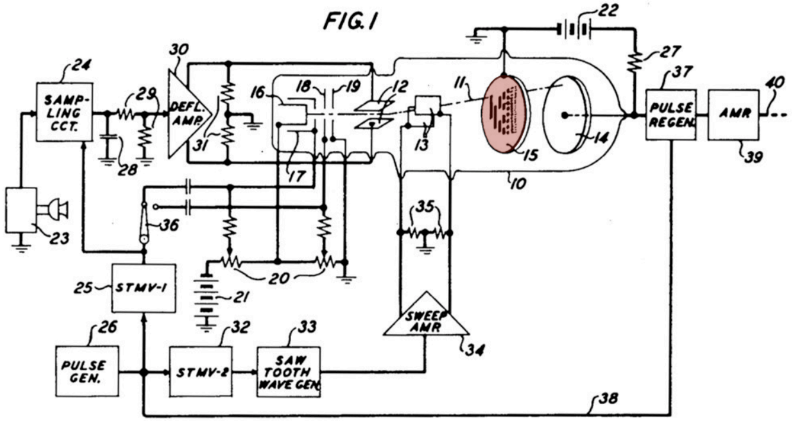




March 17, 1953 F. GRAY 2,632,058

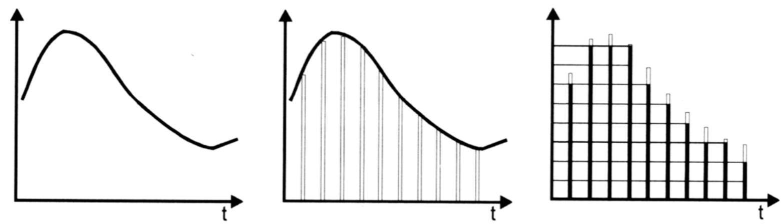
PULSE CODE COMMUNICATION

Filed Nov. 13, 1947 4 Sheets-Sheet 1



Digital Audio

- sampling
- quantisation
- compression



digitisation = sampling + quantisation

9

Sampling

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

10

Quantisation

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

11

Digital Video

12

Digital Video

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

Resolution

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

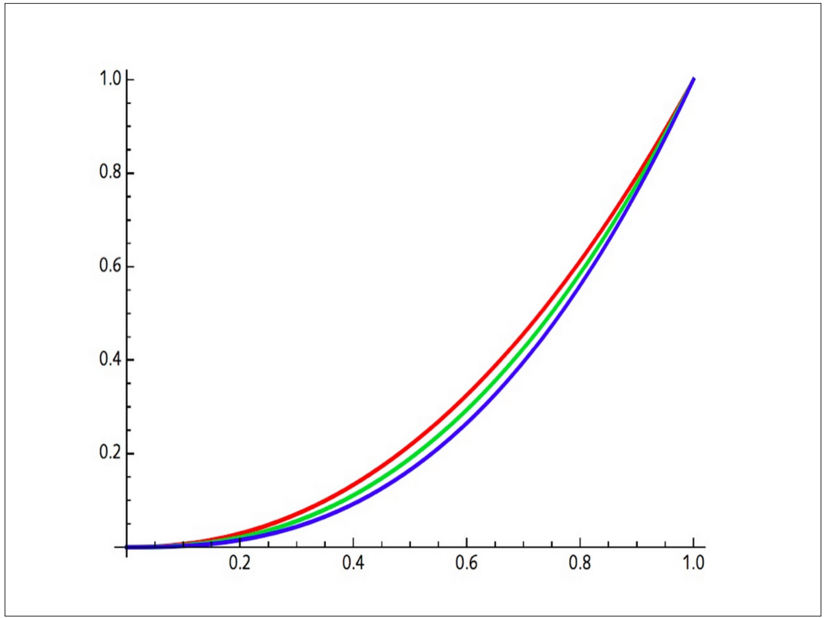
Bit Depth

- 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)

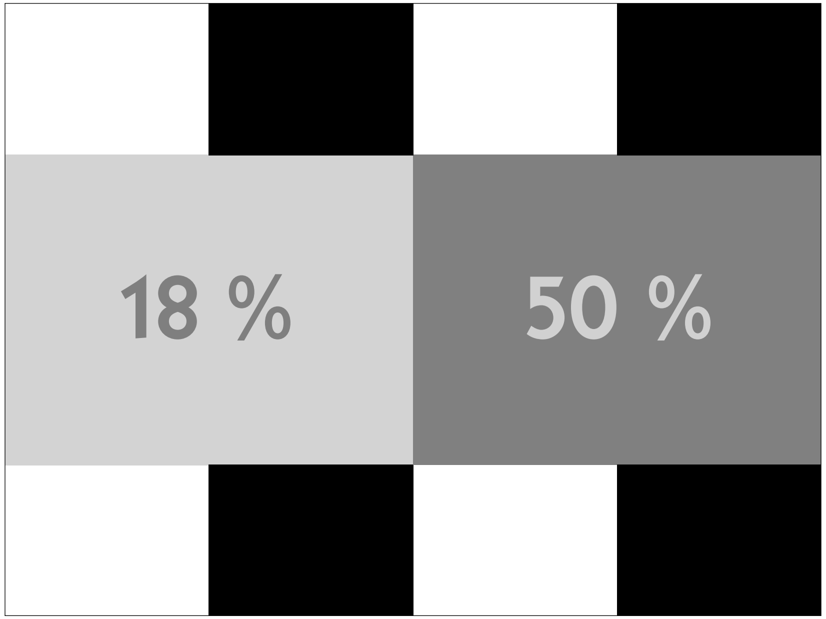
Linear, Power, Logarithmic

"medium grey"

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



17

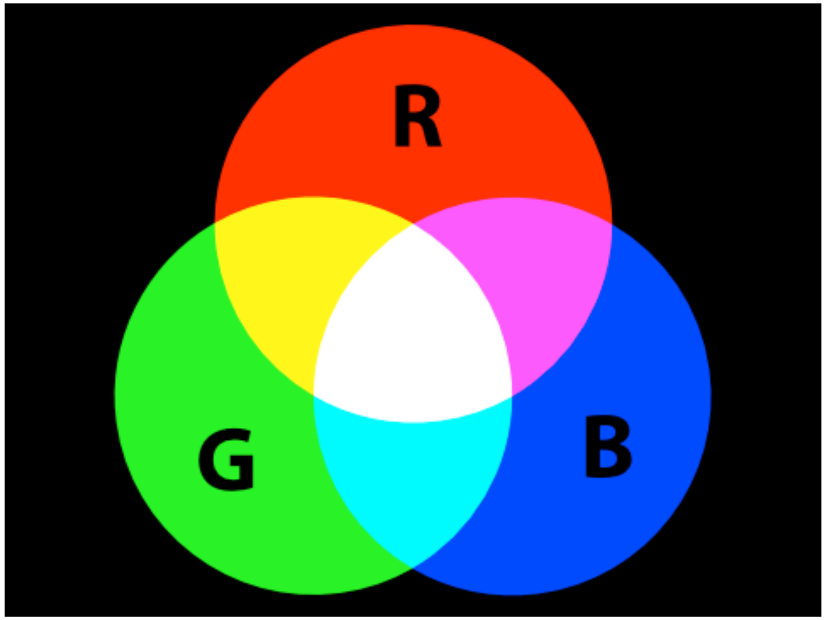


18

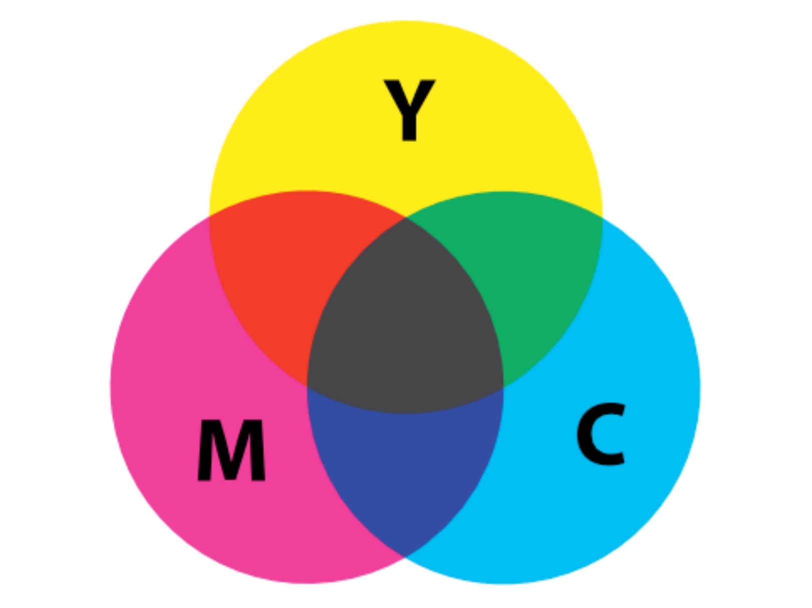
Colour Model

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

19



20



21

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

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

22

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

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

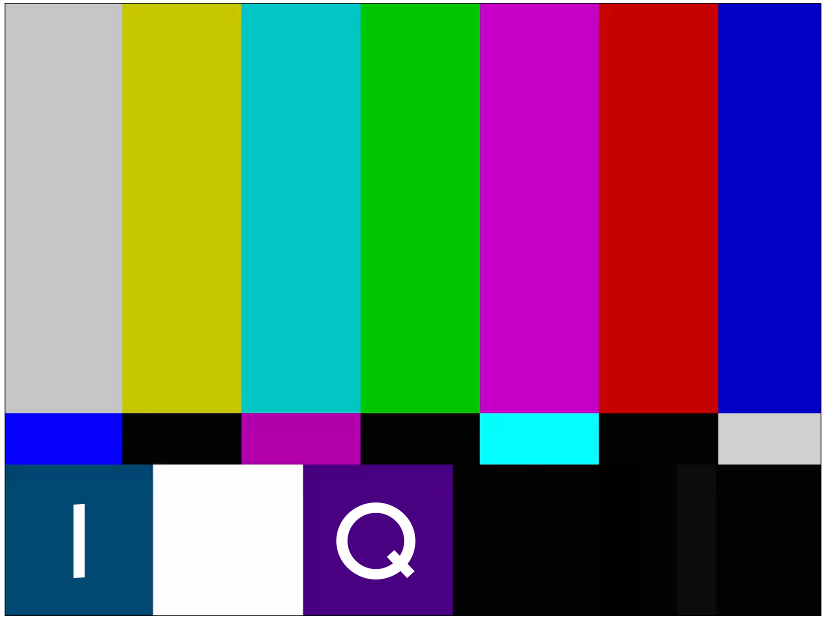
23

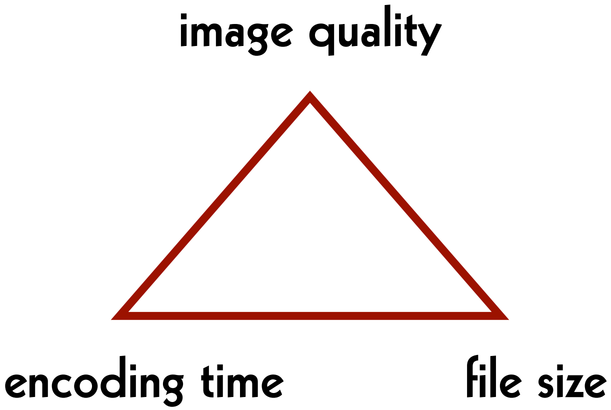
24

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

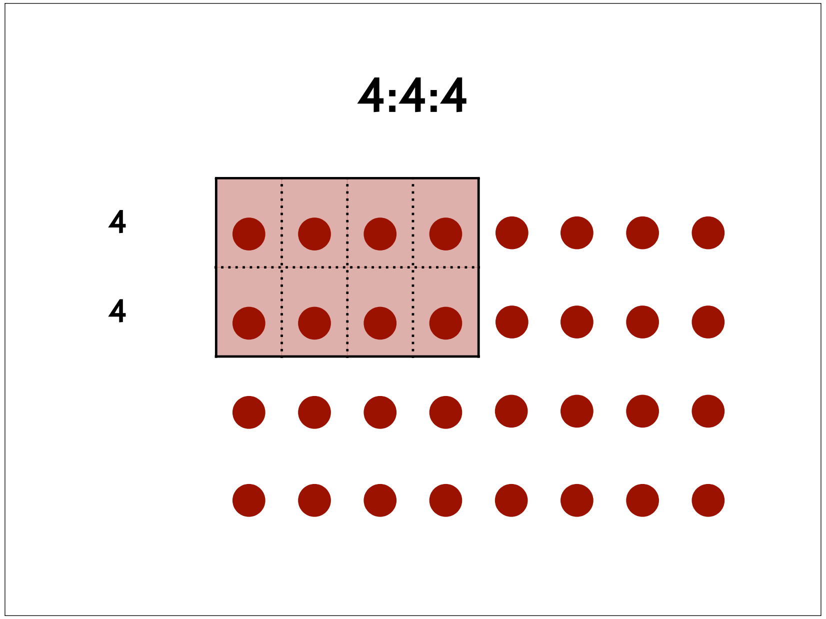
- optimised for image acquisition and/or postproduction
- optimised for access

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

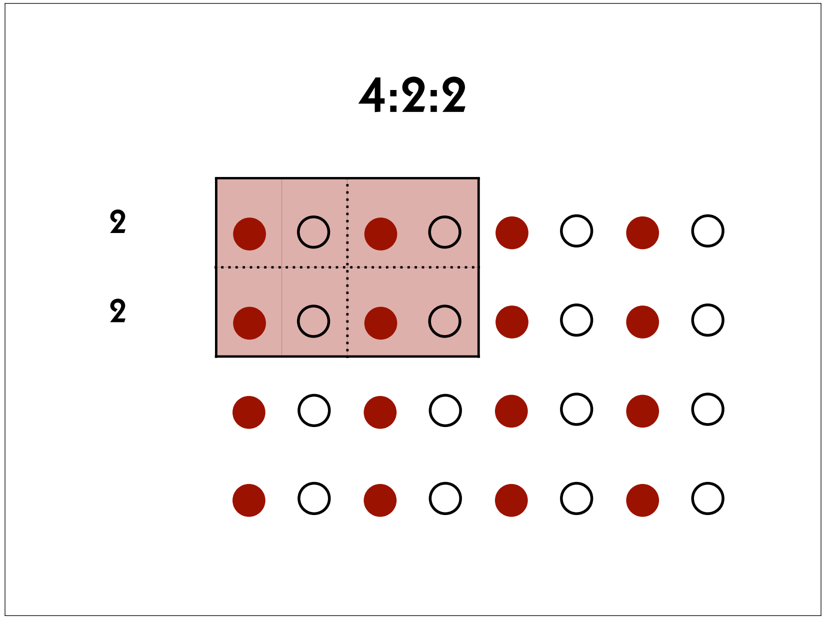
Examples (access): H.264 (AVC), H.265 (HEVC), H.266 (VVC); AV1

Chroma Subsampling

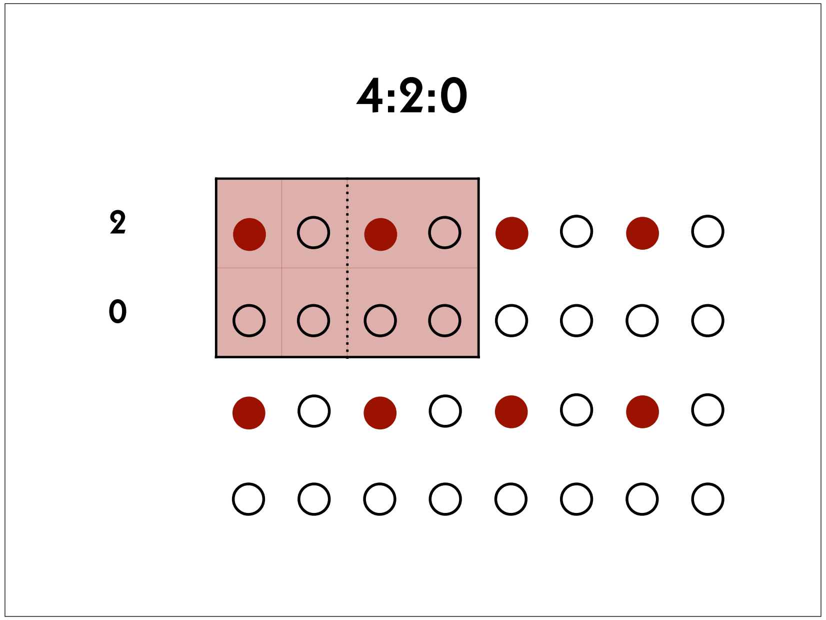
- 4:4:4
- 4:2:2
- 4:2:0 / 4:1:1



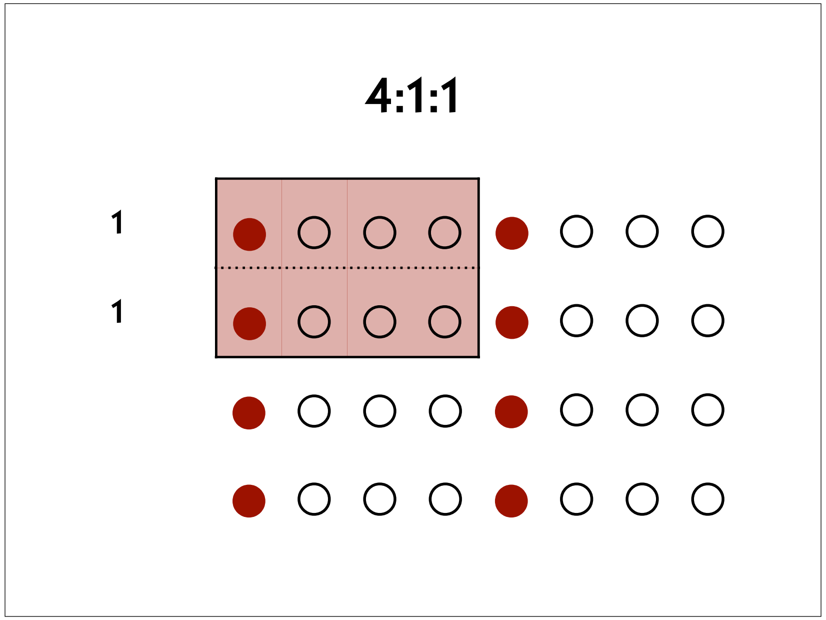
33



34



35



36

Born Compressed

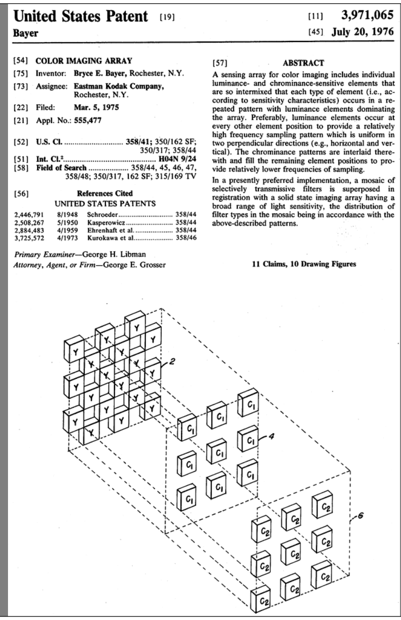
- optimised for both image acquisition and postproduction

Examples: CineForm RAW, ProRes RAW, Blackmagic RAW

Uncomfortable Truths

- sensors are colour blind
- Bayer sensors do not generate full RGB

Bryce E. Bayer (1929–2012)



United States Patent [19] — [11] **3,971,065**

Bayer — [45] **July 20, 1976**

[54] **COLOR IMAGING ARRAY**

[75] Inventor: **Bryce E. Bayer**, Rochester, N.Y.

[73] Assignee: **Eastman Kodak Company**, Rochester, N.Y.

[22] Filed: **Mar. 5, 1975**

[21] Appl. No.: **555,477**

[52] **U.S. Cl.** .... **358/41**; 350/162 SF; 350/317; 358/44

[51] **Int. Cl.²** **H04N 9/24**

[58] **Field of Search** .... 358/44, 45, 46, 47, 358/48; 350/317, 162 SF; 315/169 TV

[56] **References Cited**

UNITED STATES PATENTS

2,446,791	8/1948	Schroeder	358/44
2,508,267	5/1950	Kasperowicz	358/44
2,884,483	4/1959	Ehrenhaft et al.	358/44
3,725,572	4/1973	Kurokawa et al.	358/46

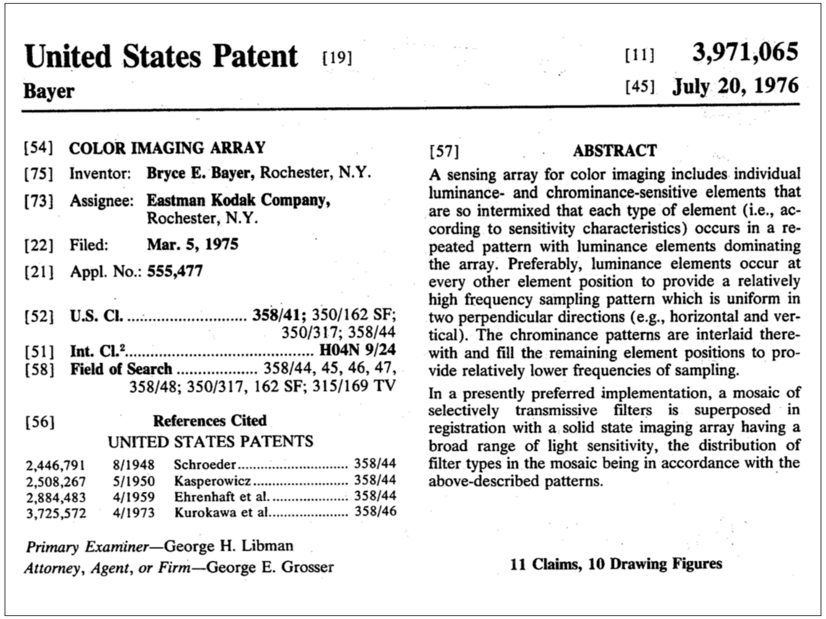
Primary Examiner—George H. Libman
Attorney, Agent, or Firm—George E. Grosser

[57] **ABSTRACT**

A sensing array for color imaging includes individual luminance- and chrominance-sensitive elements that are so intermixed that each type of element (i.e., according to sensitivity characteristics) occurs in a repeated pattern with luminance elements dominating the array. Preferably, luminance elements occur at every other element position to provide a relatively high frequency sampling pattern which is uniform in two perpendicular directions (e.g., horizontal and vertical). The chrominance patterns are interlaid therewith and fill the remaining element positions to provide relatively lower frequencies of sampling.

In a presently preferred implementation, a mosaic of selectively transmissive filters is superposed in registration with a solid state imaging array having a broad range of light sensitivity, the distribution of filter types in the mosaic being in accordance with the above-described patterns.

11 Claims, 10 Drawing Figures



United States Patent [19]

Bayer

[11] **3,971,065**

[45] **July 20, 1976**

[54] **COLOR IMAGING ARRAY**

[75] Inventor: **Bryce E. Bayer,** Rochester, N.Y.

[73] Assignee: **Eastman Kodak Company,** Rochester, N.Y.

[22] Filed: **Mar. 5, 1975**

[21] Appl. No.: **555,477**

[52] **U.S. Cl.** .... **358/41;** 350/162 SF; 350/317; 358/44

[51] **Int. Cl.²**.. **H04N 9/24**

[58] **Field of Search** .... 358/44, 45, 46, 47, 358/48; 350/317, 162 SF; 315/169 TV

[56] **References Cited**

UNITED STATES PATENTS

2,446,791	8/1948	Schroeder	358/44
2,508,267	5/1950	Kasperowicz	358/44
2,884,483	4/1959	Ehrenhaft et al.	358/44
3,725,572	4/1973	Kurokawa et al.	358/46

Primary Examiner—George H. Libman
Attorney, Agent, or Firm—George E. Grosser

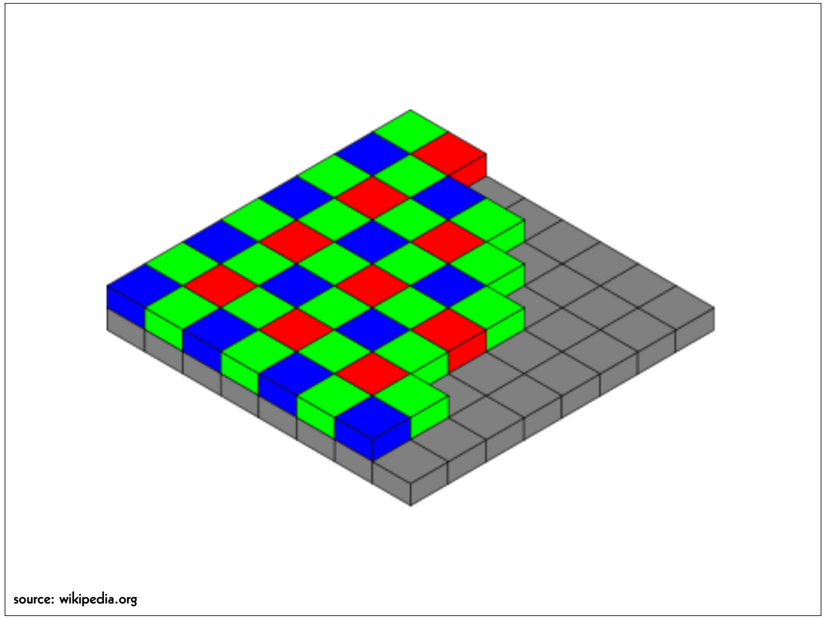
[57] **ABSTRACT**

A sensing array for color imaging includes individual luminance- and chrominance-sensitive elements that are so intermixed that each type of element (i.e., according to sensitivity characteristics) occurs in a repeated pattern with luminance elements dominating the array. Preferably, luminance elements occur at every other element position to provide a relatively high frequency sampling pattern which is uniform in two perpendicular directions (e.g., horizontal and vertical). The chrominance patterns are interlaid therewith and fill the remaining element positions to provide relatively lower frequencies of sampling.

In a presently preferred implementation, a mosaic of selectively transmissive filters is superposed in registration with a solid state imaging array having a broad range of light sensitivity, the distribution of filter types in the mosaic being in accordance with the above-described patterns.

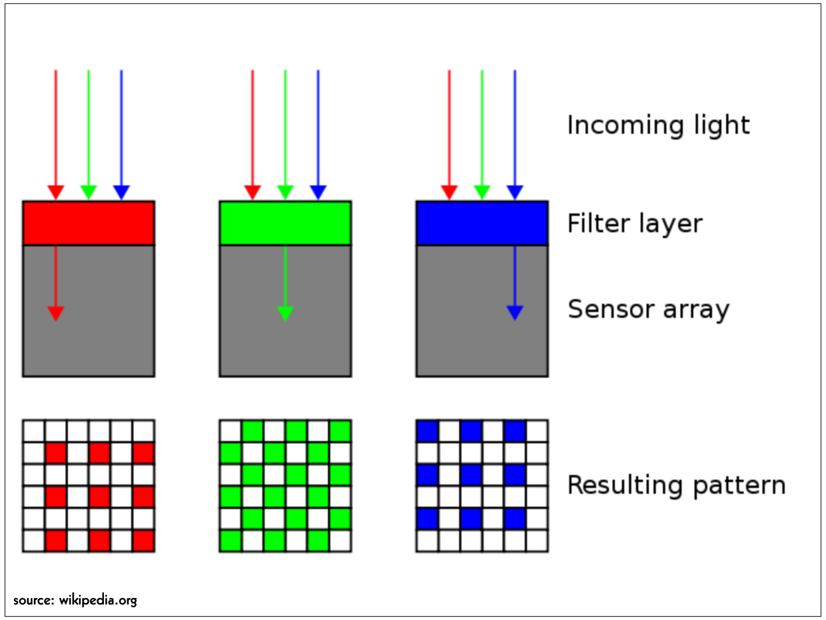
11 Claims, 10 Drawing Figures

41



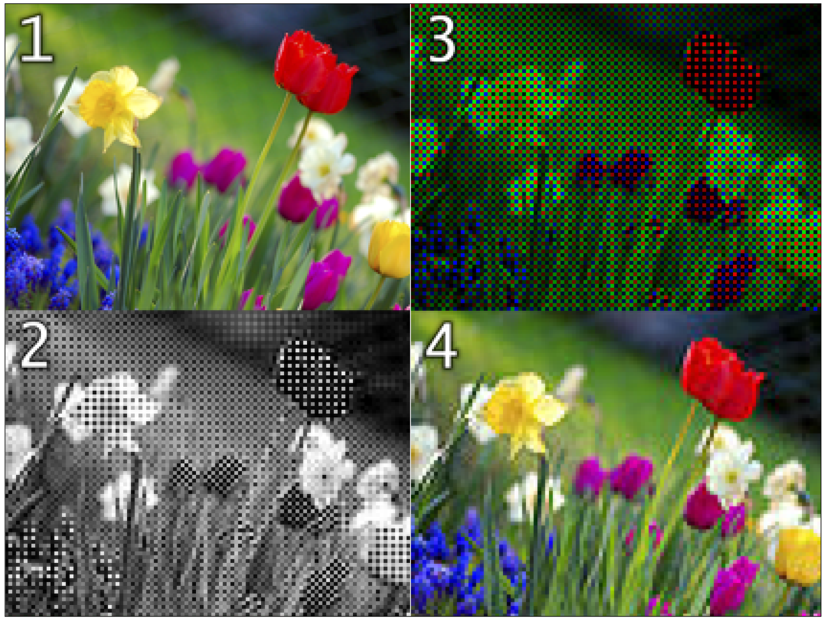
source: wikipedia.org

42



source: wikipedia.org

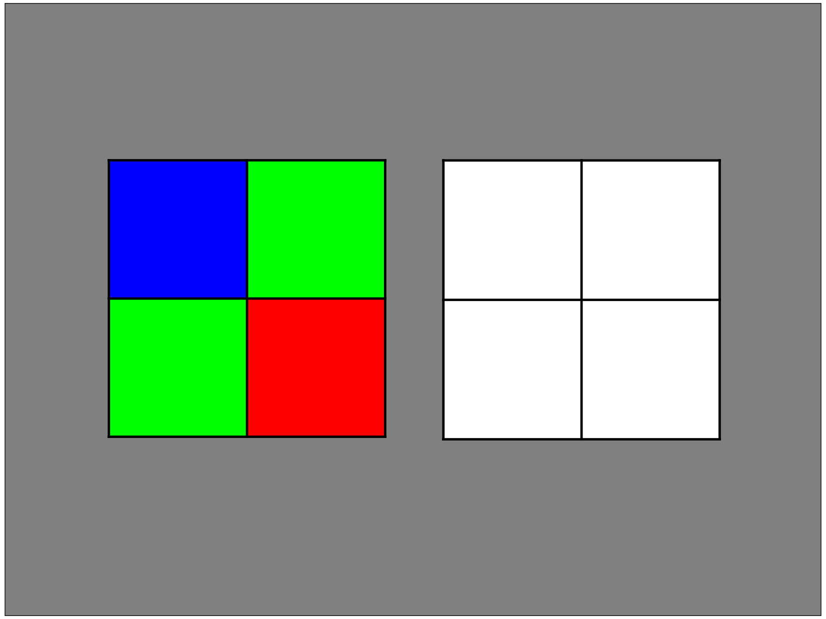
43



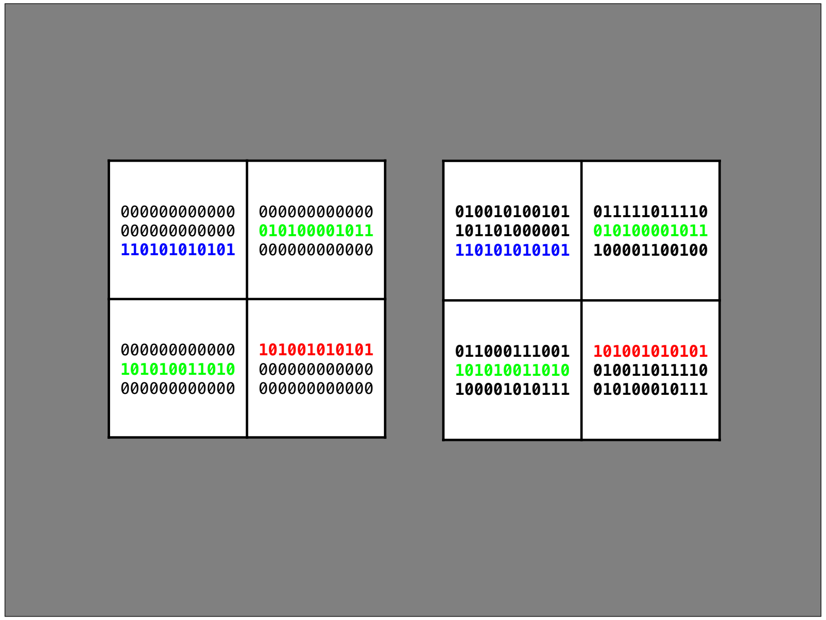
44

```
01110101001010101000101101010111110
01001101010101010101000010111101010
01110101001010101000101101010111110
00011101010101010101000010111101010
01101010100101010100010110101011111
00101010101010100001011101010100000
01110101001010101000101101010111110
01010101010101010000101110101001110
10010111010100101010100010110101010
11100101010101010100001011101010100
01110101001010101000101101010111110
01010101010101010011010101000000001
00101000101010101010010101010101010
```

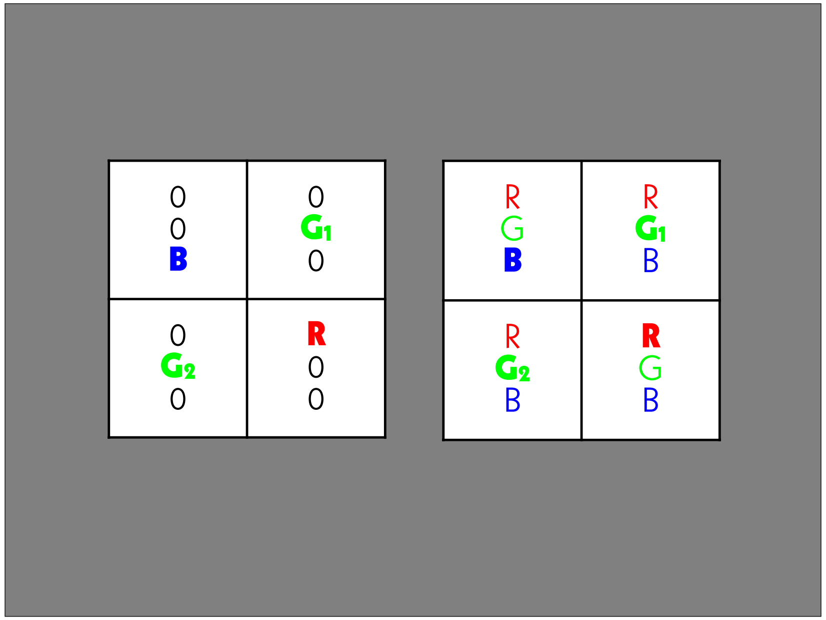
45



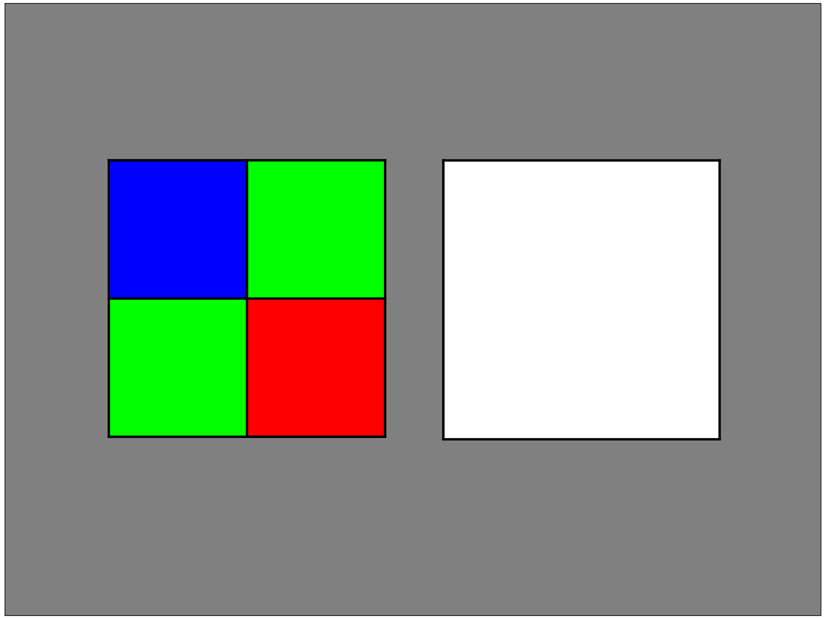
46



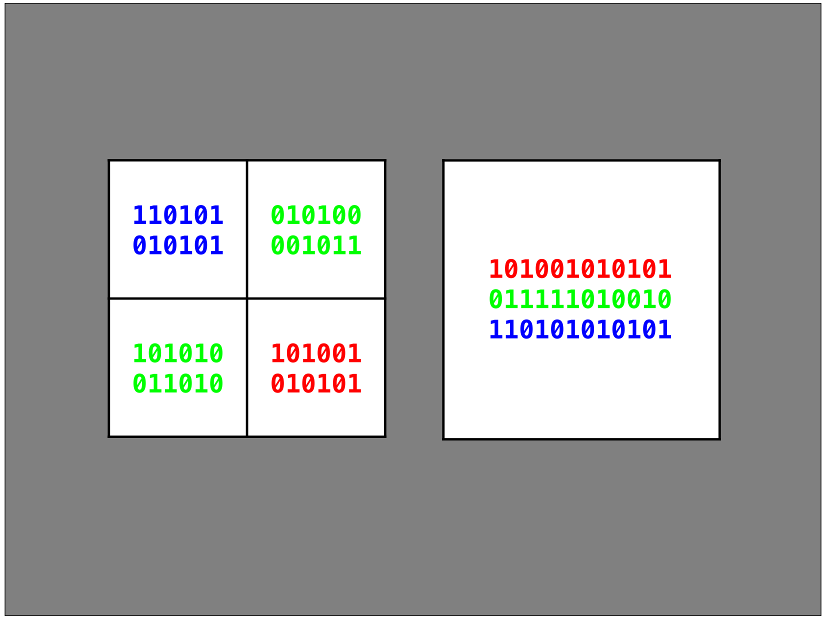
47



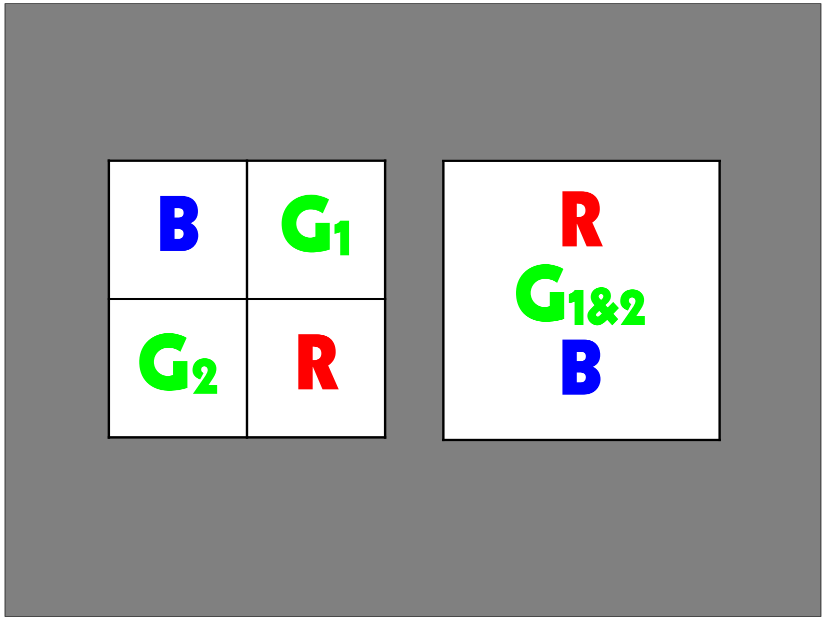
48



49



50



51

Ways to use Bayer-type data

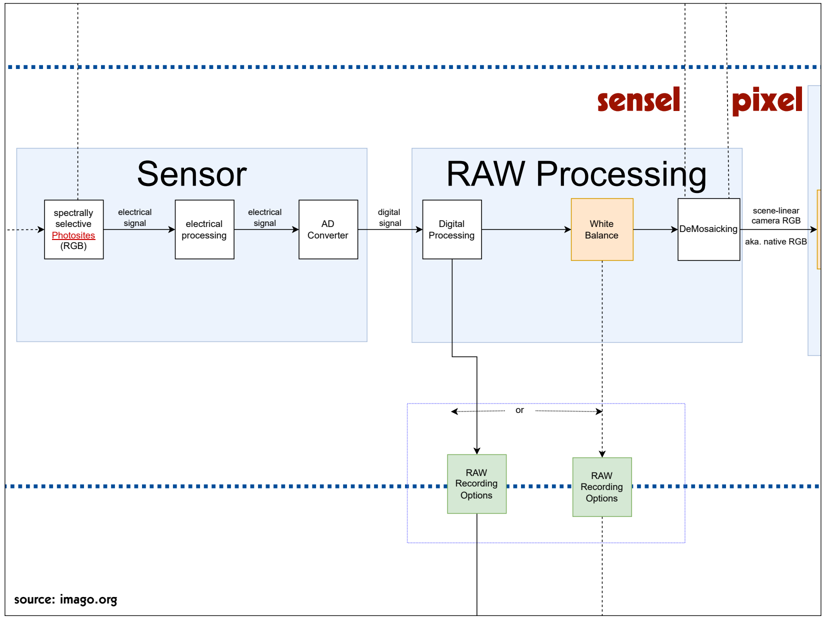
digital blow-up to RGB

- 3 times the amount of the generated data
- the file has the full sensor resolution
- only ⅓ of the data are real

digital reduction to RGB

- ¾ the amount of the generated data
- the file has ½ of the sensor resolution
- all data are real

52



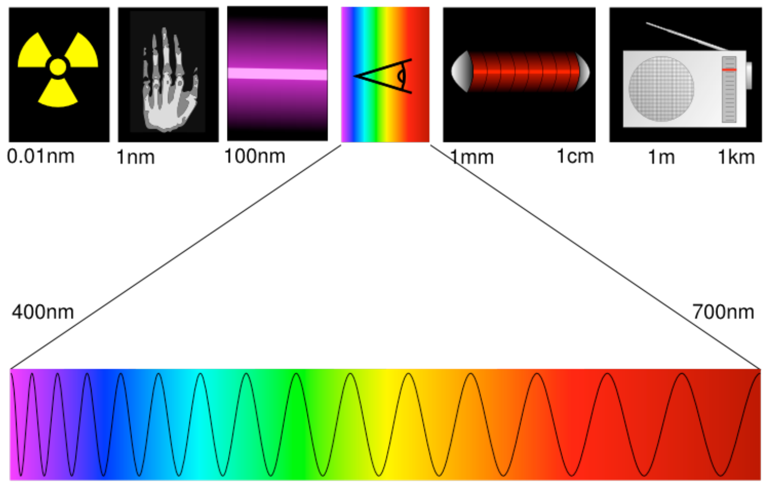
source: imago.org

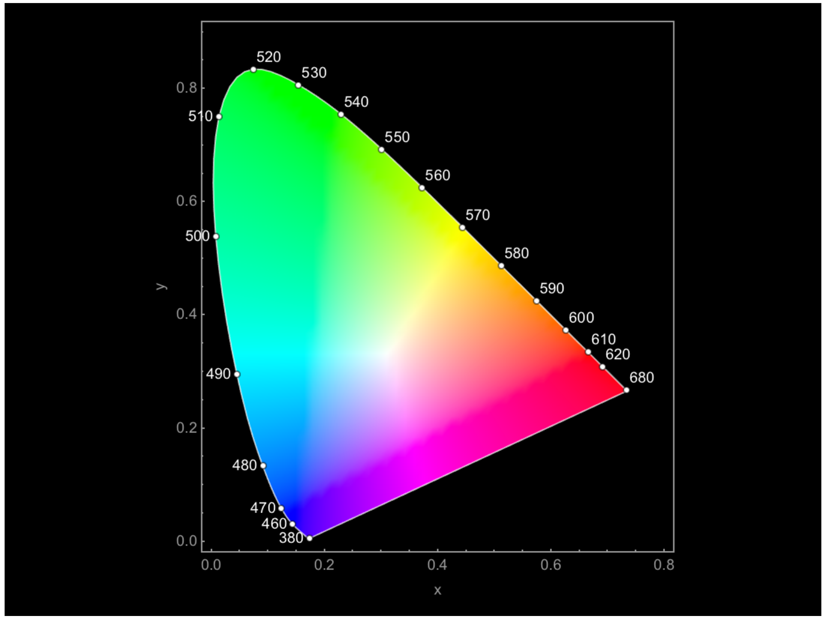
Ways to store Bayer-type data

- pixel values generated by one de-mosaicking algorithm (digital blow-up)
- pixel values generated by mixing two green sensel values into one (digital reduction)
- raw sensel values

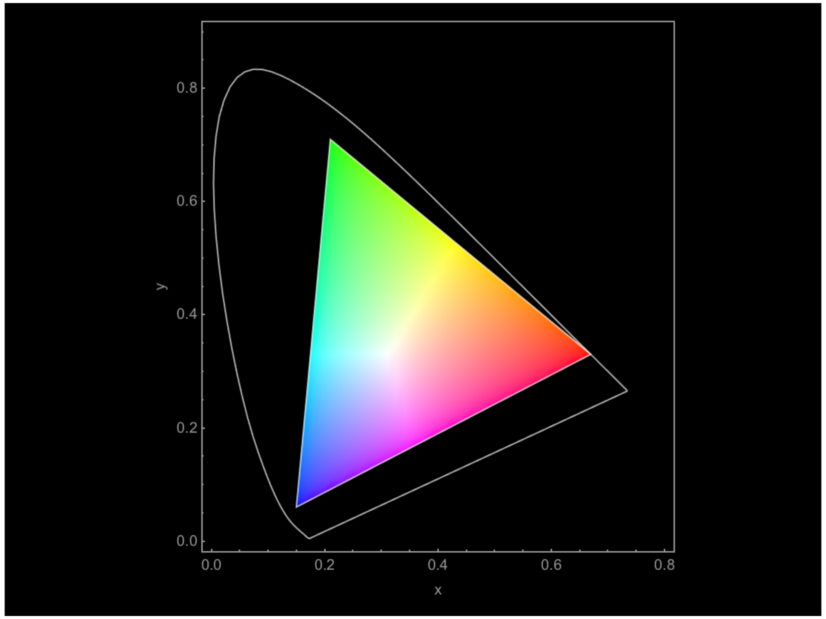
Standard Illuminant

- D50
- D55
- D65
- D75

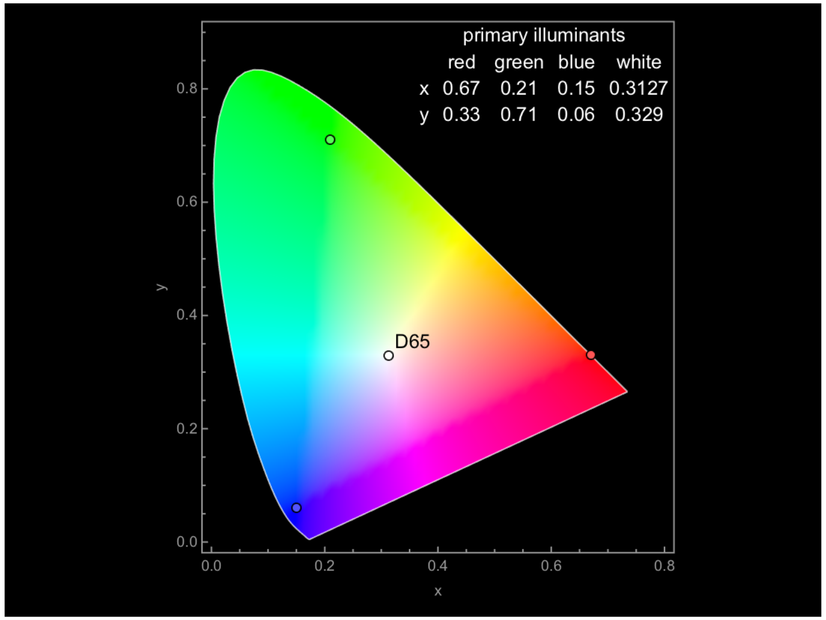




57



58

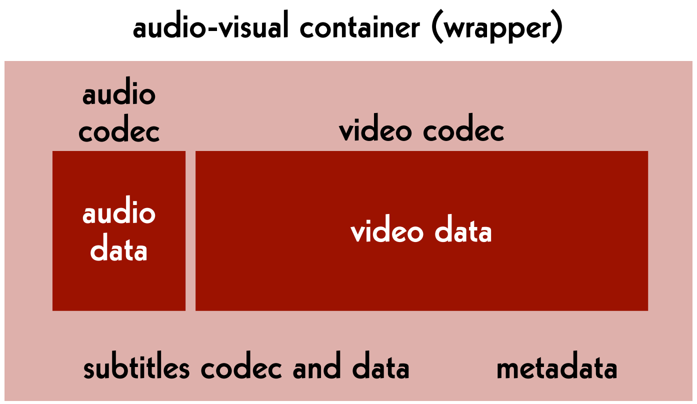


59

File Structure

60

File Structure



61

Audio-Visual Container

- MP4
- QuickTime (.mov)
- AVI
- Flash
- MXF
- Matroska (.mkv)

62

Audio-Visual Container

- MP4
- QuickTime (.mov)
- AVI
- Flash
- MXF
- Matroska (.mkv)

63

Single Images

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

64

Audio Codec

- WAVE
- BWF

- AAC
- MP3

- FLAC

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
- Blackmagic RAW

Video Codec (Access)

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

- AV1

RAW data are cooked.

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, 16 bit
- MXF, DPX, 2K, R'G'B', 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

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

Archive Master and Mezzanine

film
- Matroska, FFV1, 2K, R'G'B', 16 bit

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

audio
- Matroska, FLAC, 96 kHz, 24 bit

Access

video

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

audio

- 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

→ retokromer.ch/publications/JFP_96.html

Pros & Cons

container:

- folder
- TAR
- ZIP
- MXF
- Matroska
- AXF

video codec:

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

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

The Missing Piece of Software

RAWcooked (CLI)

→ mediaarea.net/RAWcooked

RAWcooked

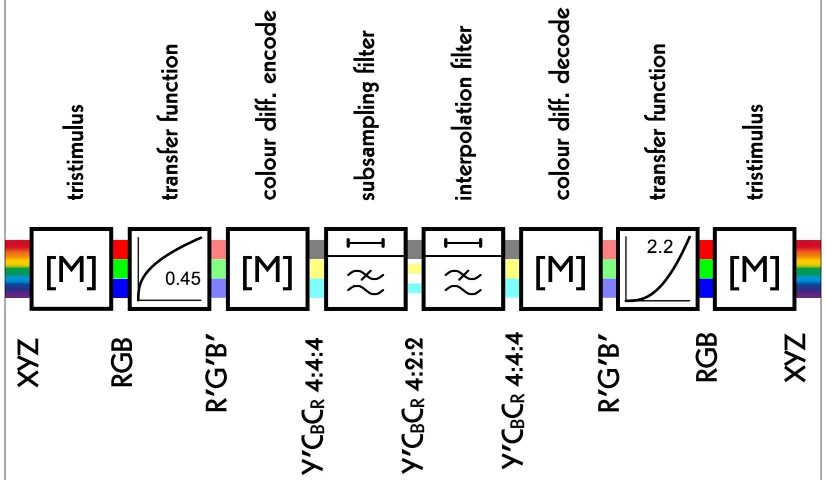
- encoding into Matroska (.mkv) using FFV1 video codec and FLAC audio codec
- all metadata preserved
- decoding with bit-by-bit reversibility
- possibility to embed sidecar files (e.g. MD5, LUT, XML)
- compatibility with media players

Transformations

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

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

89



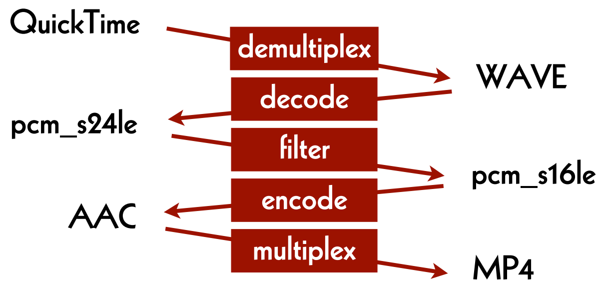
90

Data Transformations

- demultiplex
- decode
- filter
- encode
- multiplex

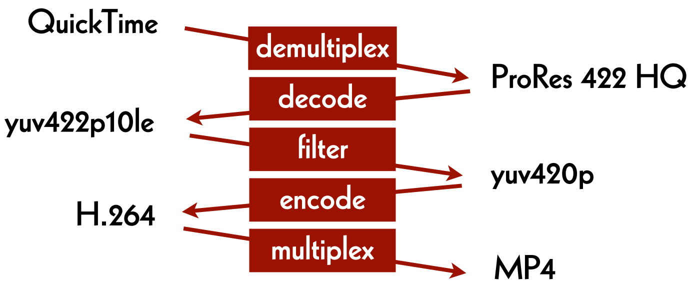
91

Audio Exemple



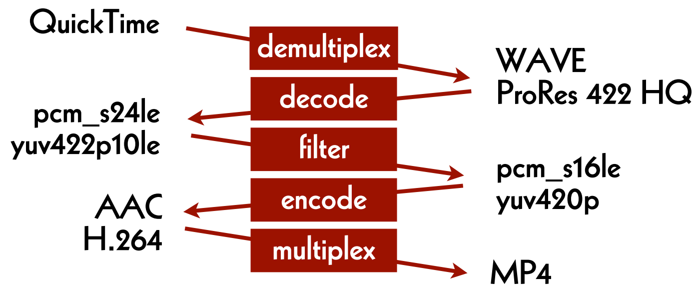
92

Video Exemple



93

Audio-Visual Exemple



94

Data Maintenance

95

Plan the Next Migration

- file naming
- barcodes
- checksums
- write the full index onto the cartridge
- technical metadata
- code to retrieve the files

96

File Naming (Example)

- title_codec.container
- title_codec_container_algorithm.txt

- film_H264.mp4
- film_H264_mp4_md5.txt

Checksums

cryptographic

- MD5
- SHA-1
- SHA-256
- SHA-512

non-cryptographic

- CRC-32
- xxHash 32
- xxHash 64
- xxHash 128

Longterm

- storage of the cartridges
- three copies …
- … in geographically distant locations
- data integrity check
- data migration
- availability of LTO desks

Data Migrations

2014

- our internal archive from LTO-4 to LTO-6 (5.7 PB)

2014–2021

- two dozen migrations for clients

2021

- our internal archive from LTO-6 to LTO-8 (25.2 PB)

Reading

Reto Kromer: **On the Bright Side of Data Migrations**, in «IASA Journal», n. 49 (December 2018), IASA, p. 18–22

→ retokromer.ch/publications/IASA_49.html

read | script | write

script to modify

- container
- codec
- both container and codec
- metadata
- filename

#1: ProRes-born Content

from:

- ProRes stored in a QuickTime (.mov) container

to:

- ProRes stored in a Matroska (.mkv) container

Update the Container

→ read file from source LTO

→ demultiplex file

- ProRes 422, 10 bit [yuv422p10le]
- ProRes 4444, 10 bit [yuv444p10le or yuva444p10le] or 12 bit [yuv444p12le]

→ multiplex file

→ write file to destination LTO

SMPTE RDD 36:2015

SMPTE REGISTERED DISCLOSURE DOCUMENT

Apple ProRes Bitstream Syntax and Decoding Process

Page 1 of 39 pages

The attached document is a Registered Disclosure Document prepared by the sponsor identified below. It has been examined by the appropriate SMPTE Technology Committee and is believed to contain adequate information to satisfy the objectives defined in the Scope, and to be technically consistent.

This document is NOT a Standard, Recommended Practice or Engineering Guideline, and does NOT imply a finding or representation of the Society.

Every attempt has been made to ensure that the information contained in this document is accurate. Errors in this document should be reported to the proponent identified below, with a copy to eng@smpte.org.

105

#2: Video

from:

- AVI / 8-bit and 10-bit uncompressed
- MOV / 8-bit and 10-bit uncompressed
- MP4 / 8-bit and 10-bit uncompressed

to:

- Matroska / FFV1

106

Container and Codec

→ read file from source LTO
→ demultiplex file
→ decode file
- $Y'C_BC_R$, 4:2:2, 8 bit, «raw» [uyvy422]

→ encode file
→ multiplex file
→ write file to destination LTO

107

Container and Codec

→ read file from source LTO
→ demultiplex file
→ decode file
- $Y'C_BC_R$, 4:2:2, 10 bit, «raw» [yuv422p10le]

→ encode file
→ multiplex file
→ write file to destination LTO

108

#3: Filename

from:

- Title_YUV422.mkv

to:

- Title_YCbCr422_9d5084b5b0a08d5022b3 9e0e75241d12.mkv

Always remember:

To do nothing is **never** an option!

Coda

Live in the real world!

There is only one efficient way:

- keep the analogue source elements as long as possible
- more prevention:
  → better insulation
  → more efficient air conditioning
- less handling of the source elements
- make digital masters and access copies

Acknowledgements (1)

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

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

Acknowledgements (2)

- Tommy Aschenbach

- Claudio Weidmann
- Jim Lindner
- Carl Eugen Hoyos
- Peter Bubestinger-Steindl
- Jérôme Martinez
- Michael Niedermayer

until 31st January 2023

AV Preservation by

reto.ch

zone industrielle Le Trési 3
1028 Préverenges
Switzerland

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

from 1st March 2023

AV Preservation by

reto.ch

Sandrainstrasse 3
3007 Bern
Switzerland

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