

On Audio-Visual File Formats

Reto Kromer • AV Preservation by reto.ch

Using FFmpeg in a film archive
FIAF, on-line, 30 Mars 2023

1

Digital Audio

3

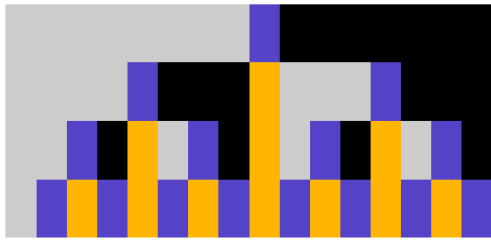
Summary

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

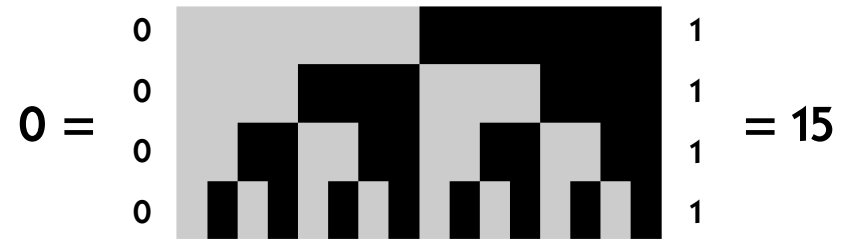
2

Frank Gray
(1887–1969)

4



5



6

Digital Audio

- sampling
- quantisation
- compression

8

March 17, 1953

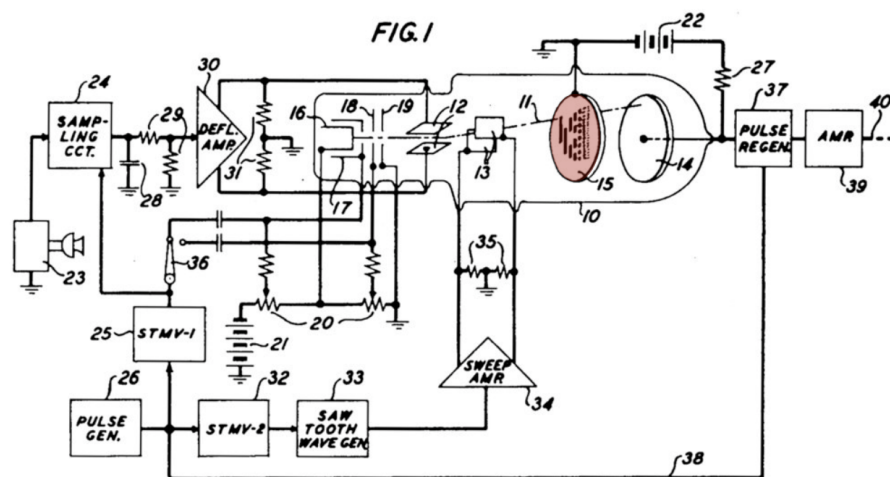
F. GRAY

2,632,058

PULSE CODE COMMUNICATION

Filed Nov. 13, 1947

4 Sheets-Sheet 1



7

Sampling

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

9

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

13

Resolution

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

14

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

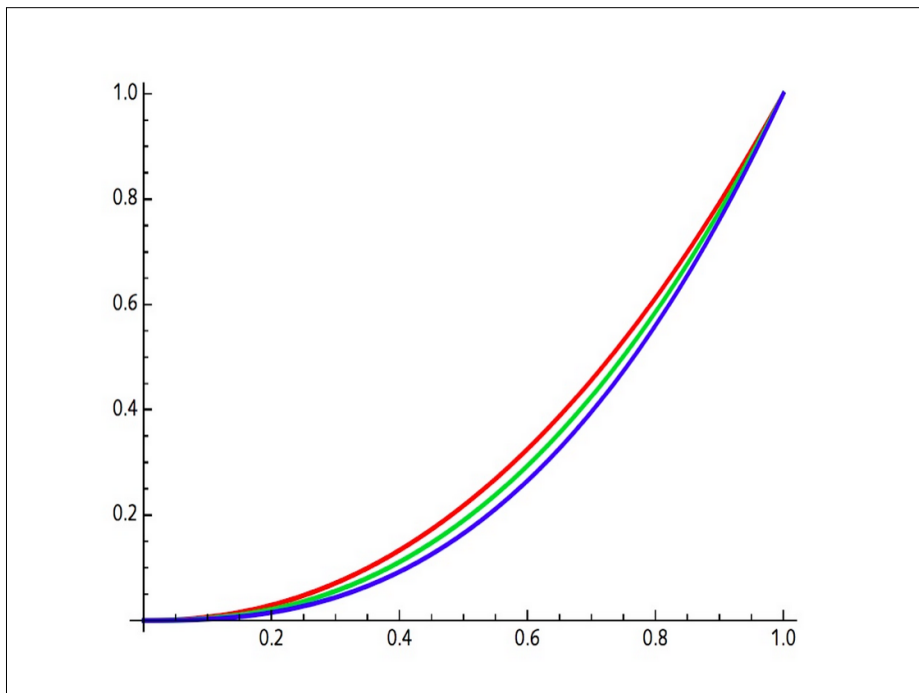
15

Linear, Power, Logarithmic

“medium grey”

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

16

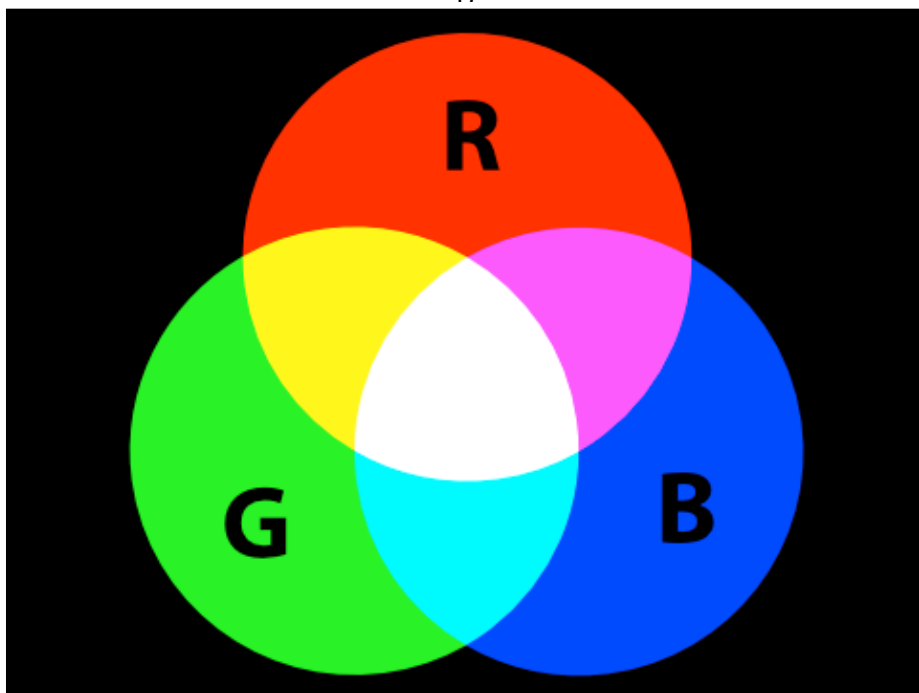


17

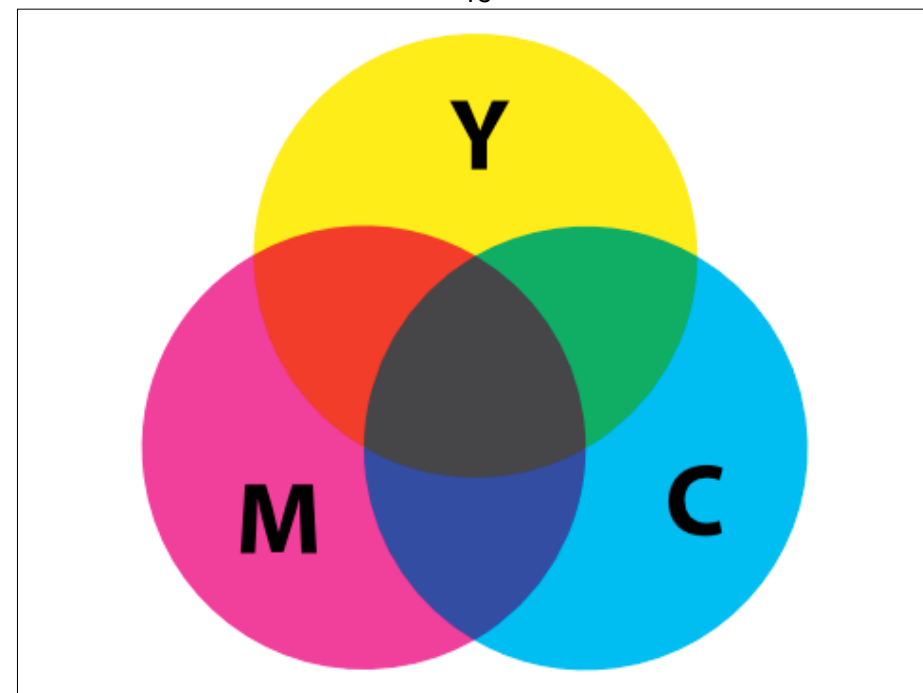
Colour Model

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

18



19



20

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

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

21

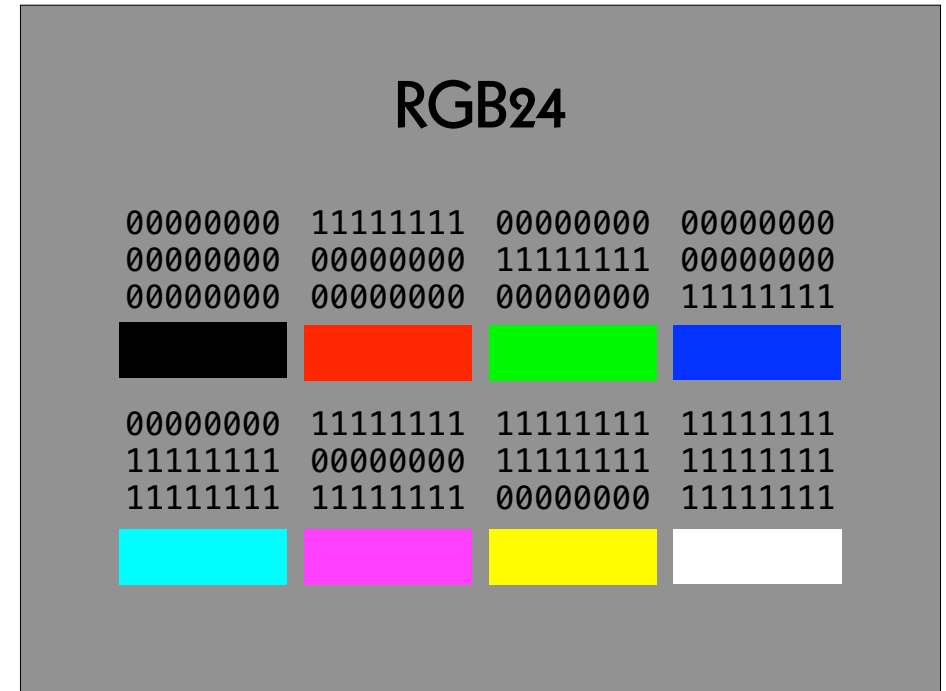


23

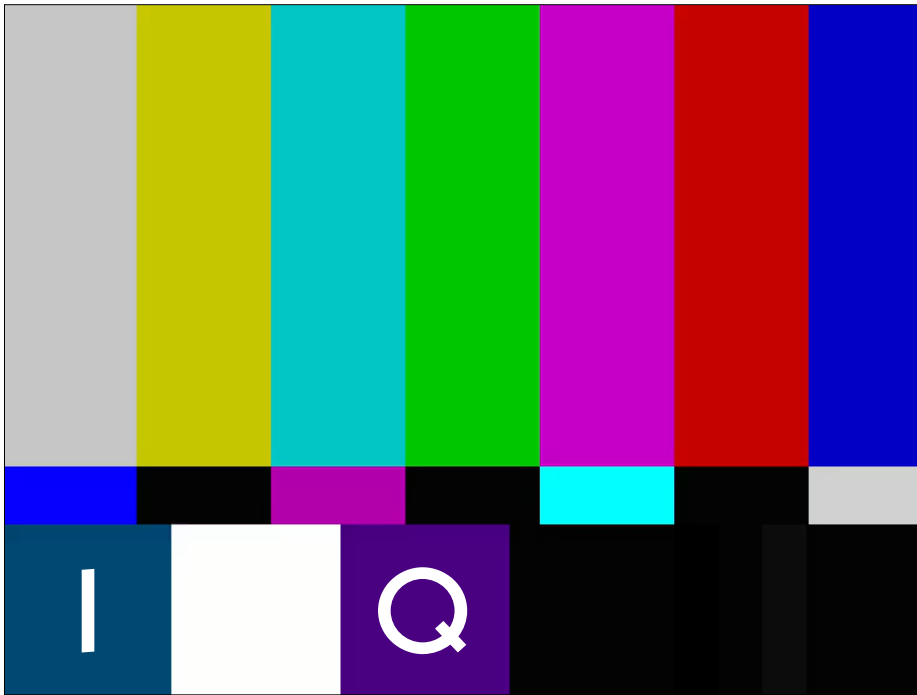
$$\begin{pmatrix} R' \\ G' \\ B' \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} Y' \\ C_O \\ C_G \end{pmatrix}$$

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

22



24

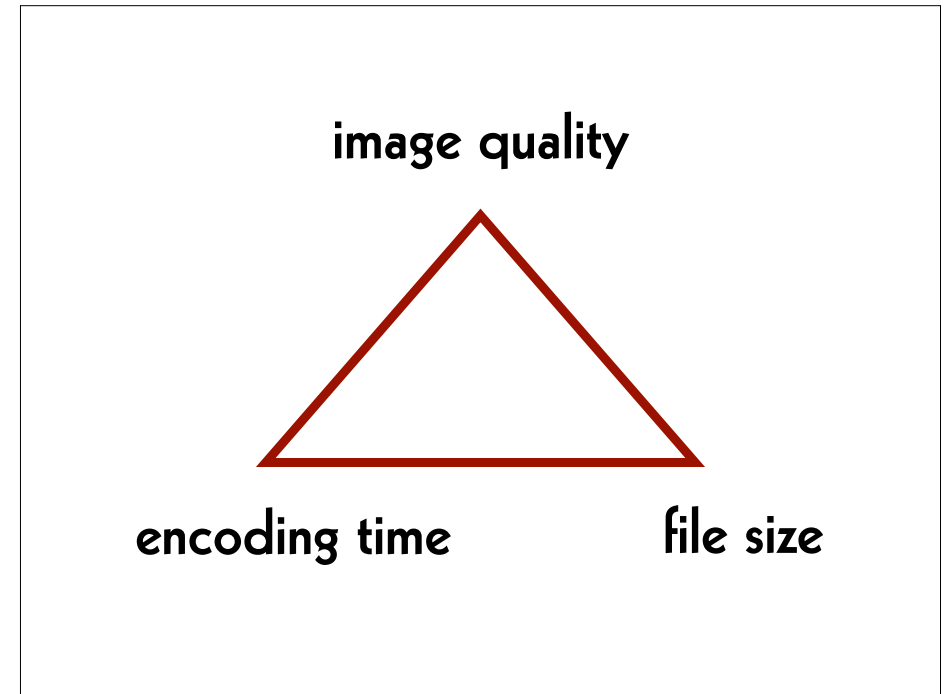


25

Compression

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

27



26

Uncompressed

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

Examples: TIFF, DPX, DNG, OpenEXR

28

Lossless Compression

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

Examples: JPEG 2000, FFV1

29

Chroma Subsampling

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

31

Lossy Compression

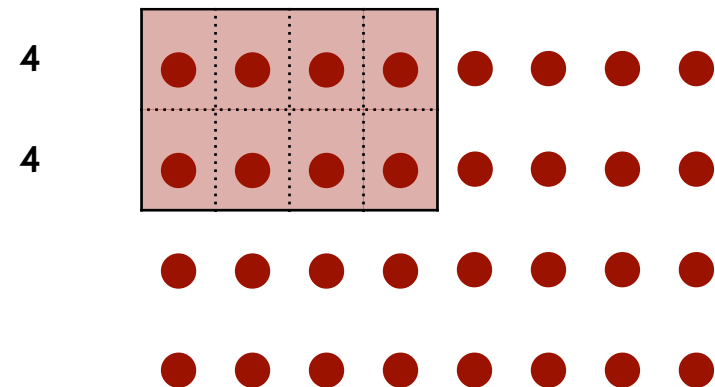
- optimised for image acquisition and/or postproduction
- optimised for access

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

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

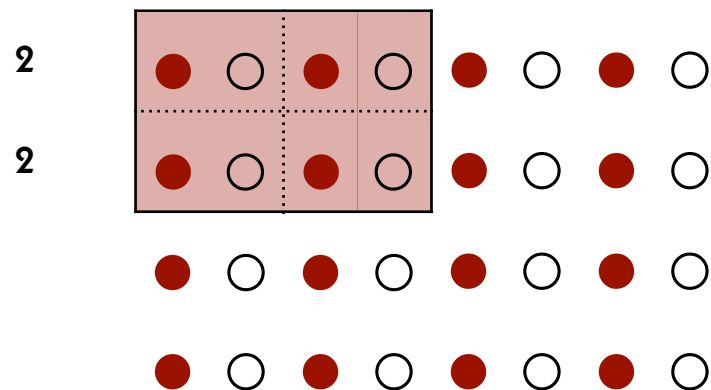
30

4:4:4



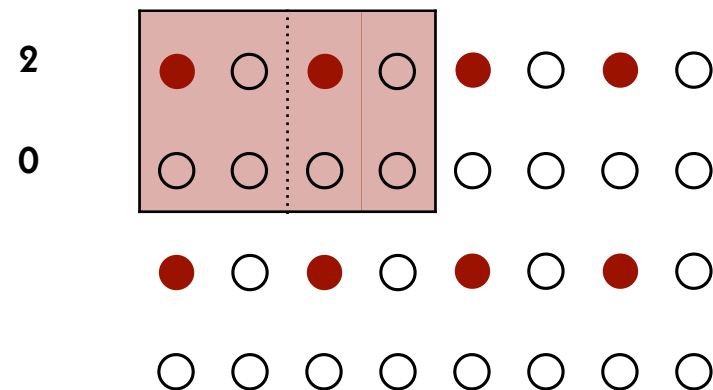
32

4:2:2



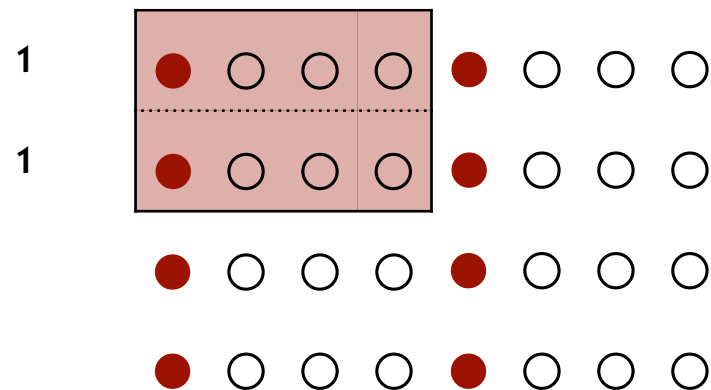
33

4:2:0



34

4:1:1



35

Born Compressed

- optimised for both image acquisition and postproduction

Examples: CineForm RAW, ProRes RAW, Blackmagic RAW

36

Uncomfortable Truths

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

37

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

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

39

Bryce E. Bayer (1929–2012)

38

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

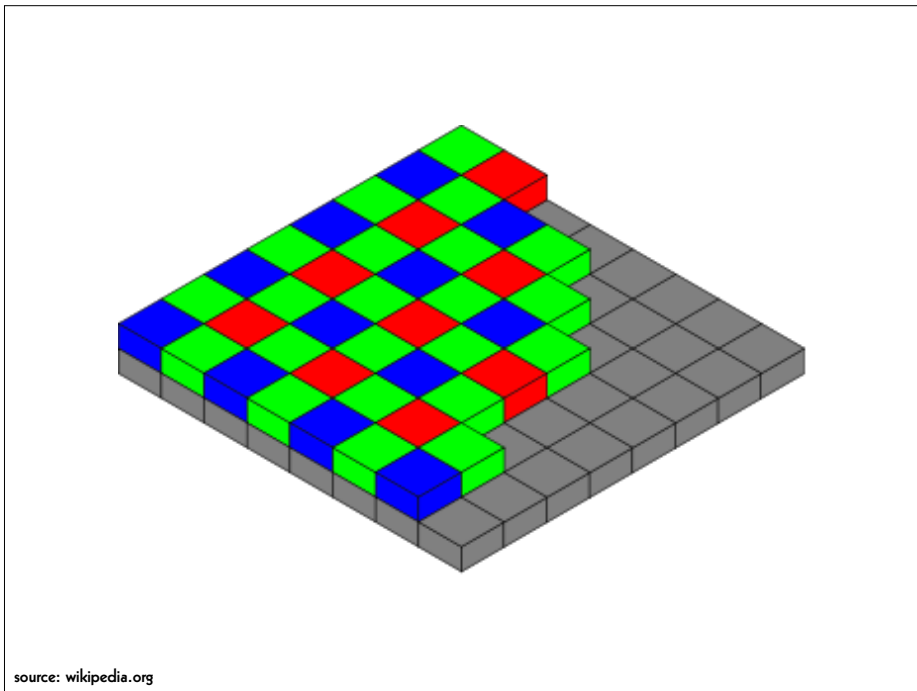
ABSTRACT

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.

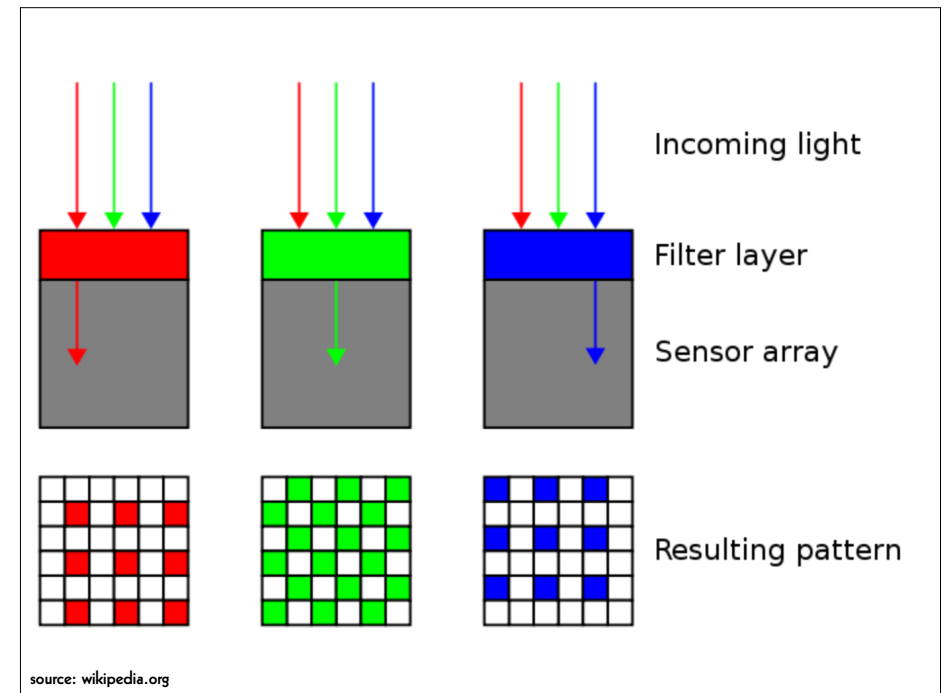
In a presently preferred implementation, a mosaic of selectively transmissive filters is superposed in registration with a solid state imaging array having a broad range of light sensitivity, the distribution of filter types in the mosaic being in accordance with the above-described patterns.

11 Claims, 10 Drawing Figures

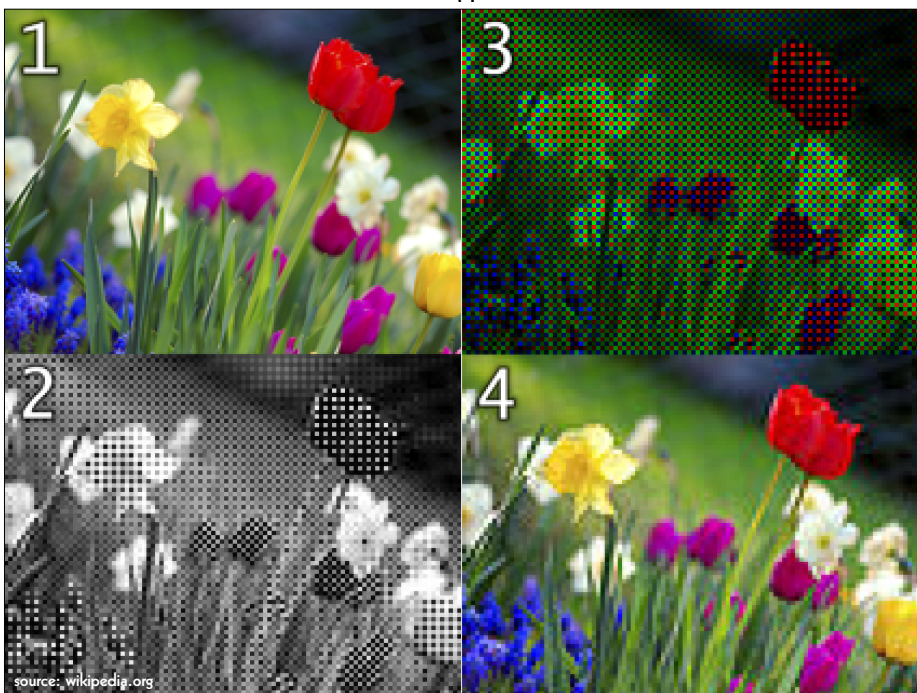
40



41



42



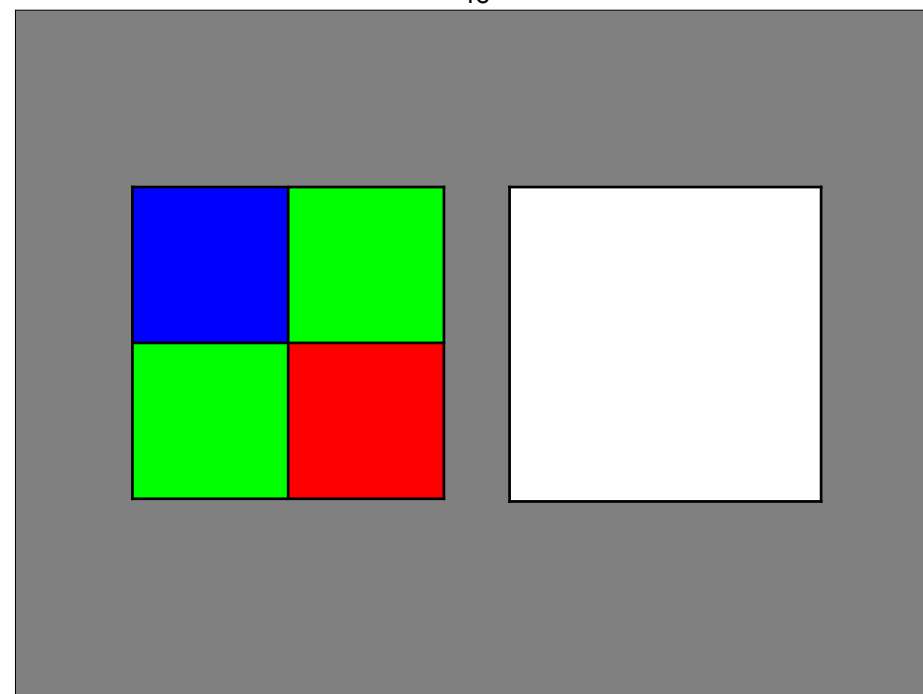
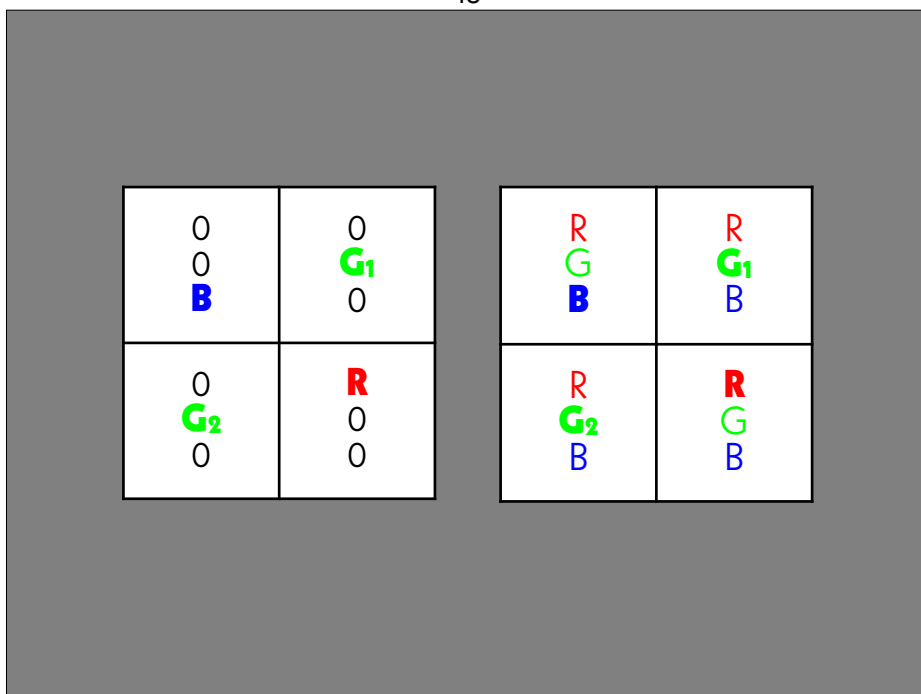
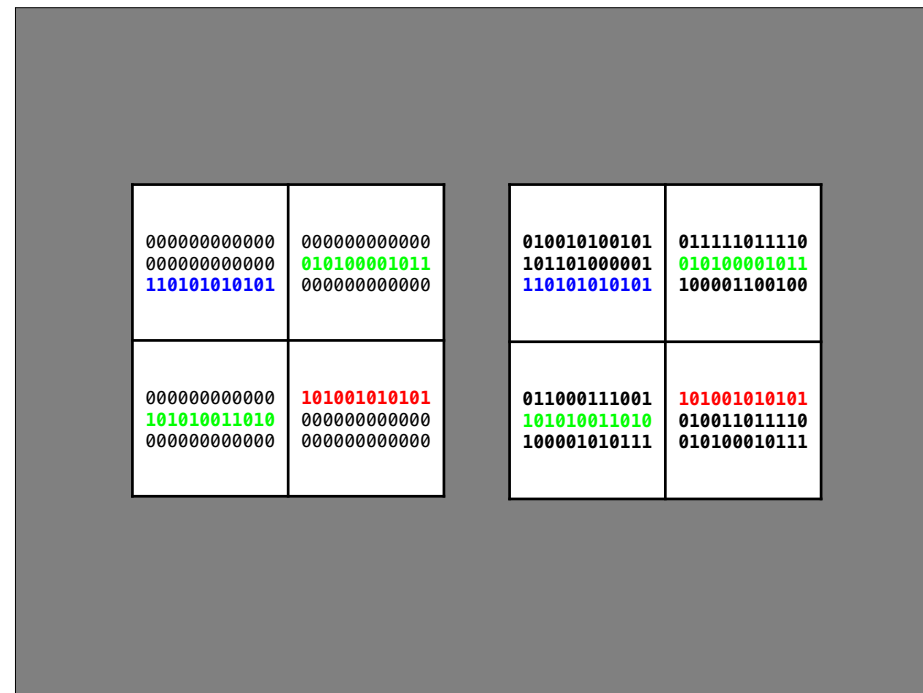
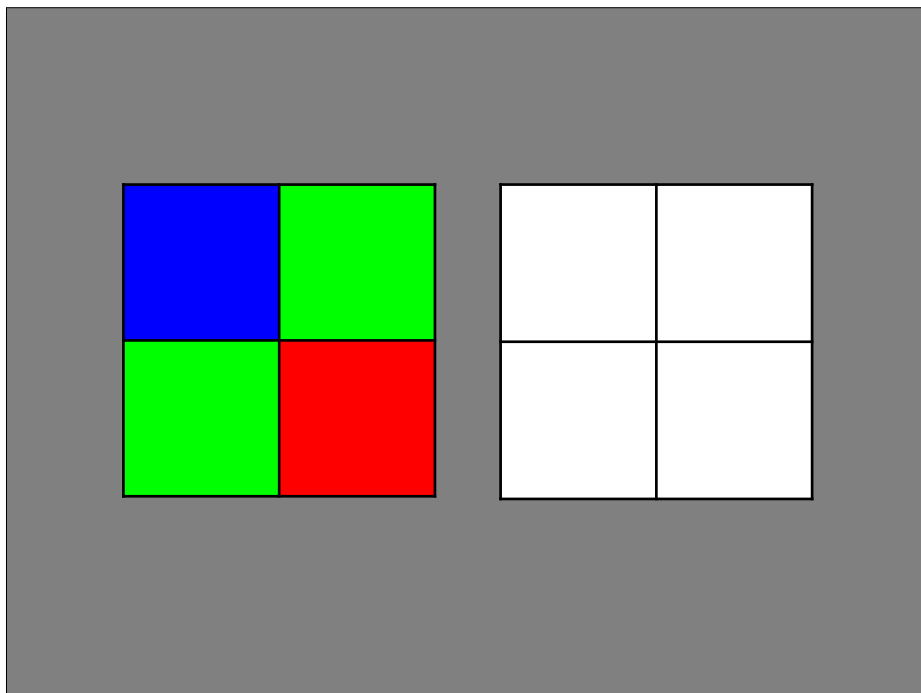
43

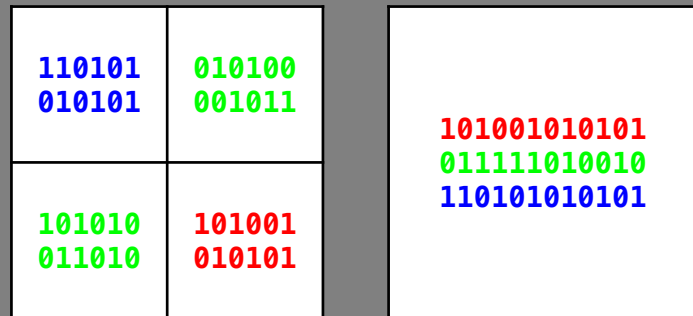
```

0111010100101010100010110101011110
0100110101010101010100001011101010
0111010100101010100010110101011110
0001110101010101010100001011101010
011010101010101010001011010101111
0010101010101010100001011101010000
0111010100101010100010110101011110
0101010101010101000010111010100110
1001011101010010101010001011010101
1110010101010101010000101110101010
0111010100101010100010110101011110
0101010101010101001101010100000001
0010100010101010101001010101010101

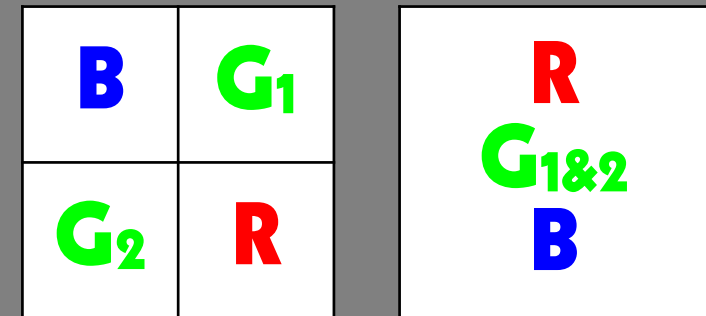
```

44





49



50

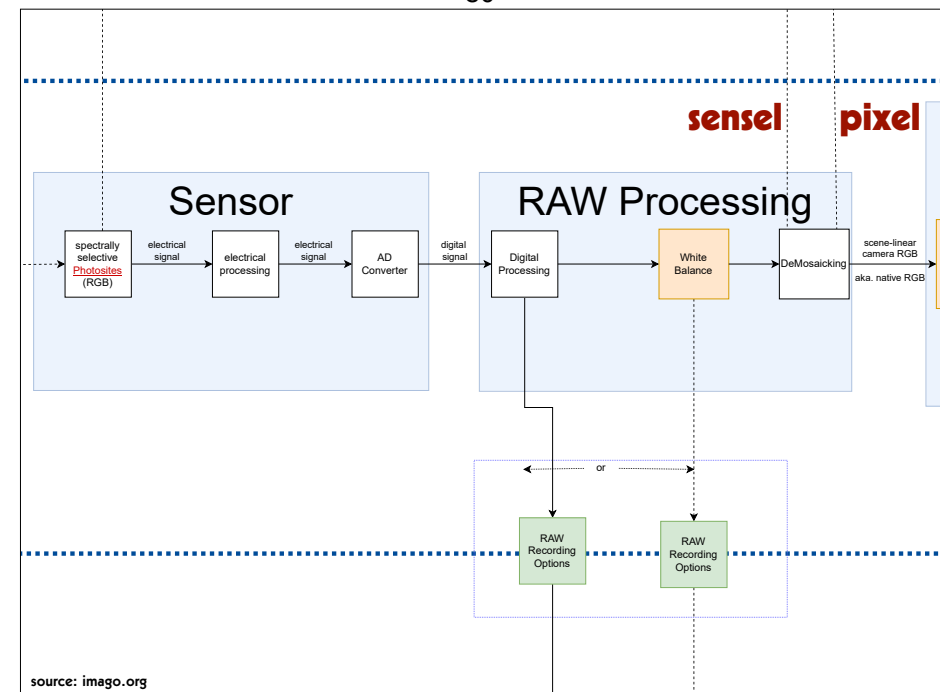
Ways to use Bayer-type data

digital blow-up to RGB

- 3 times the amount of the generated data
- the file has the full sensor resolution
- only $\frac{1}{3}$ of the data are real

digital reduction to RGB

- $\frac{3}{4}$ the amount of the generated data
- the file has $\frac{1}{2}$ of the sensor resolution
- all data are real



52

51

Ways to store Bayer-type data

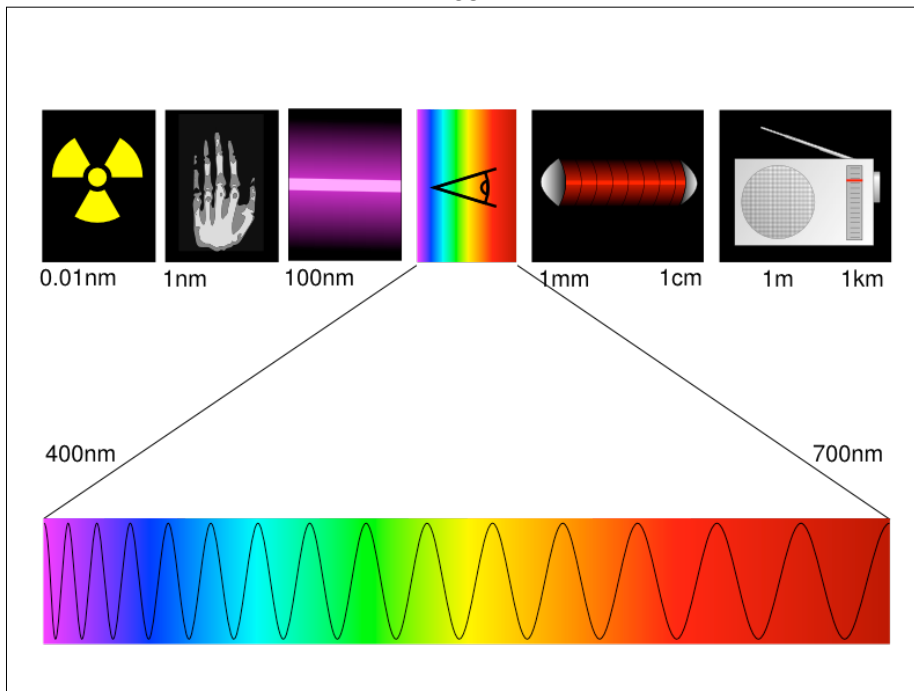
- pixel values generated by one de-mosaicking algorithm (digital blow-up)
- pixel values generated by mixing two green sensel values into one (digital reduction)
- raw sensel values

53

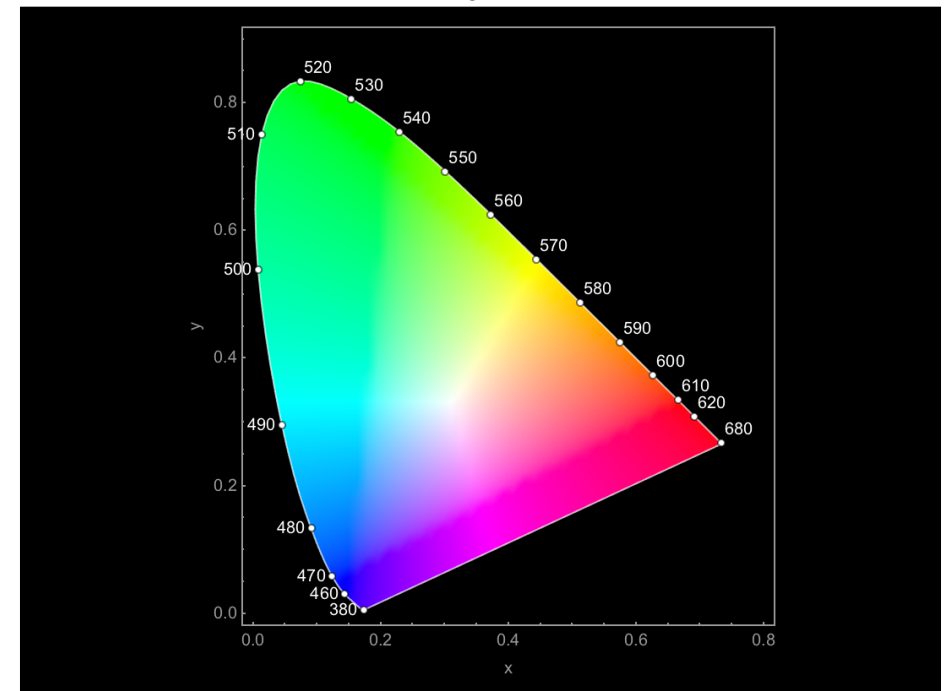
Standard Illuminant

- D50
- D55
- D65
- D75

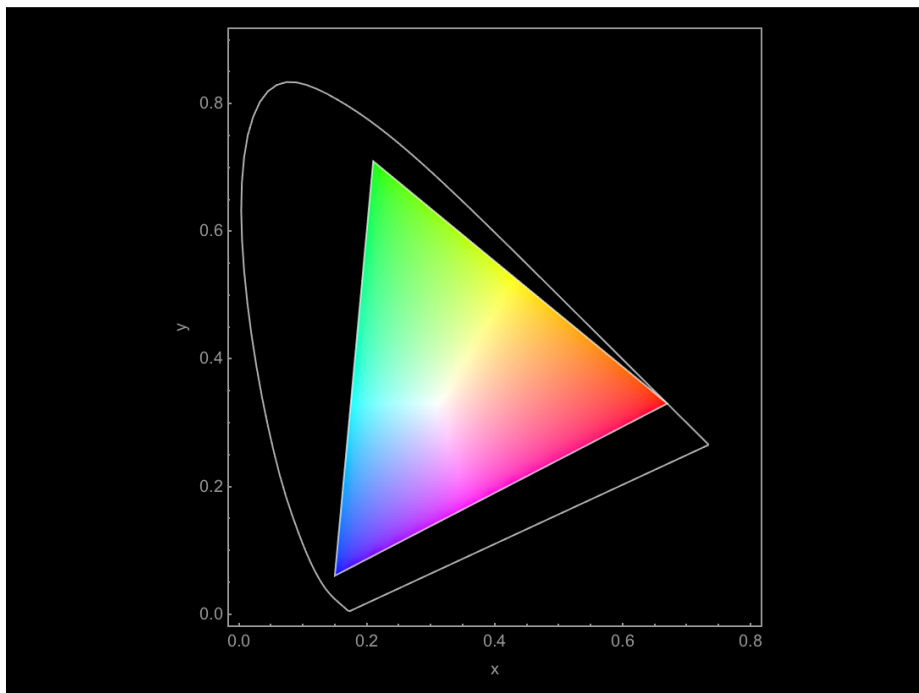
54



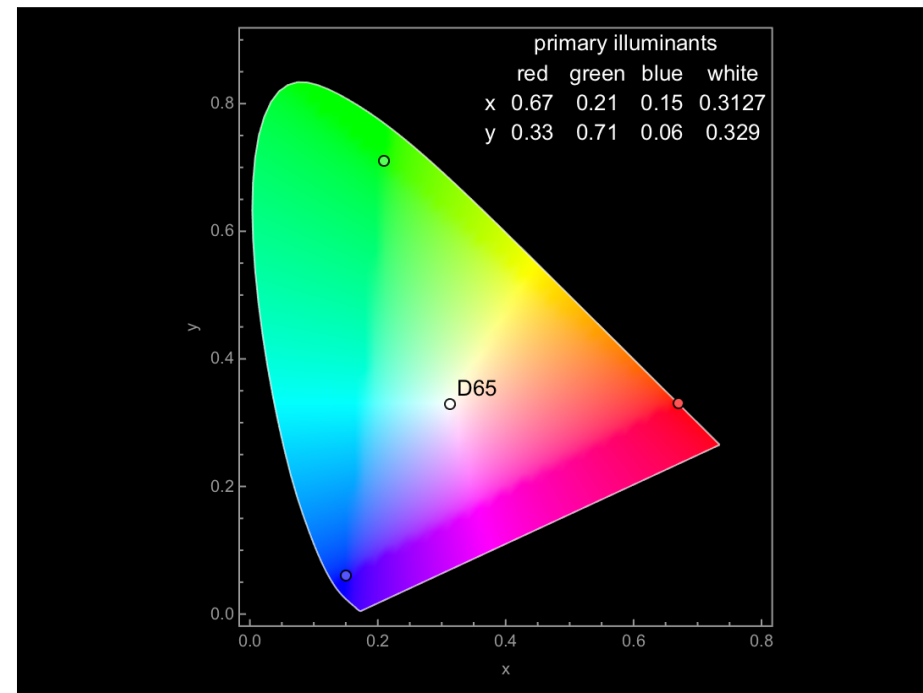
55



56



57



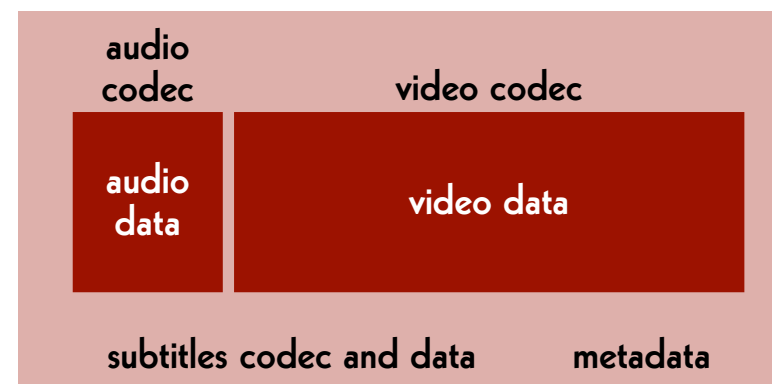
58

File Structure

59

File Structure

audio-visual container (wrapper)



60

Audio-Visual Container

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

61

Single Images

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

62

Audio Codec

- WAVE
- BWF
- AAC
- MP3
- FLAC

63

Video Codec (Master)

images

- TIFF
- DPX
- JPEG 2000
- OpenEXR
- DNG

streams

- Y'C_BC_R 8 bit
- Y'C_BC_R 10 bit
- HuffYUV
- FFV1

64

Video Codec (Mezzanine)

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

65

Video Codec (Access)

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

66

RAW data are cooked.

67

Audio Data

- pcm_s16le
- pcm_s24le
- pcm_s32le

68

Video Data

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

69

File Formats

71

What is inside my DPX?

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

70

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

72

Different Purposes

archive master format:

→ for preservation

mezzanine format:

→ for professional use in post-production

dissemination formats:

→ for widely spreading and easy access

73

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

75

Elena Rossi-Snook:

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

74

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

76

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

77

Access

video

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

audio

- FLAC, 48 kHz, 16 bit
- FLAC, 96 kHz, 16 bit

79

Archive Master and Mezzanine

film

- Matroska, FFV1, 2K, RGB, 16 bit
- Matroska, FFV1, 2K, R'G'B', 12 bit

video

- Matroska, FFV1, "HD", Y'C_BC_R 4:2:2, 10 bit
- Matroska, FFV1, "HD", Y'C_BC_R 4:4:4, 12 bit

audio

- Matroska, FLAC, 96 kHz, 24 bit
- Matroska, FLAC, 192 kHz, 24 bit

78

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

80

Pros & Cons

81

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

83

container:

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

video codec:

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

82

The Missing Piece of Software

RAWcooked (CLI)

→ mediaarea.net/RAWcooked

84

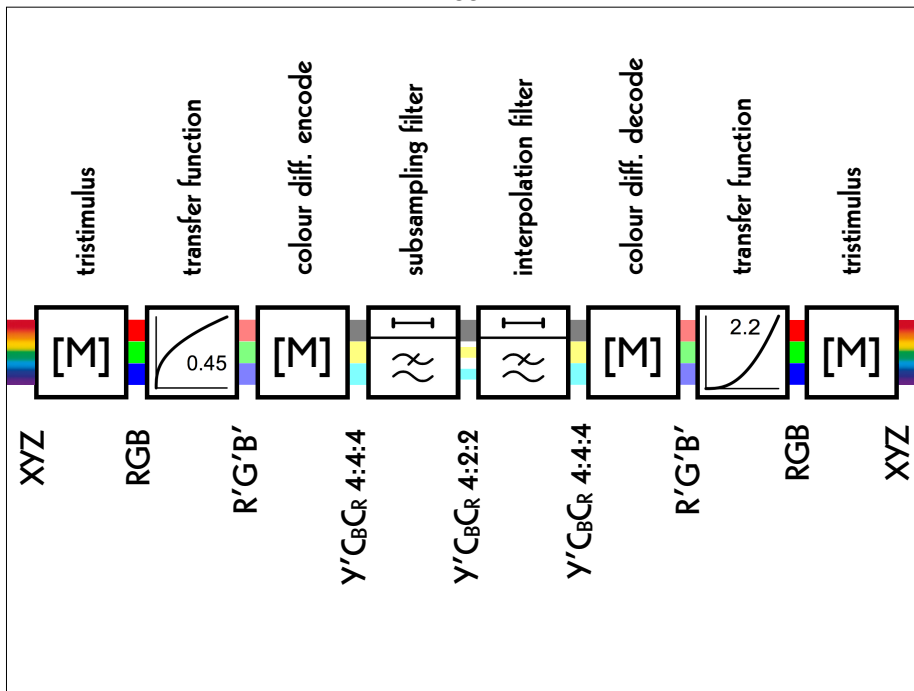
RAWcooked

- 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

85

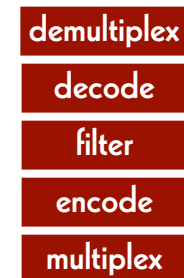
Transformations

86



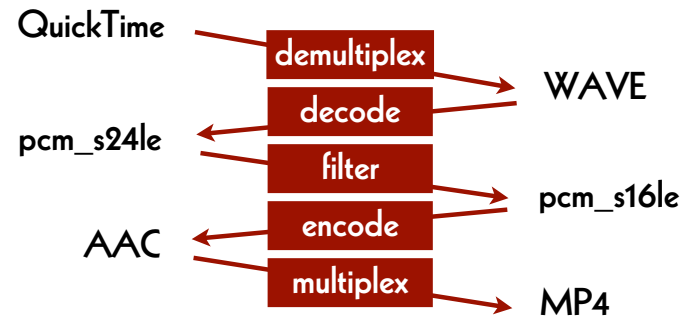
87

Data Transformations



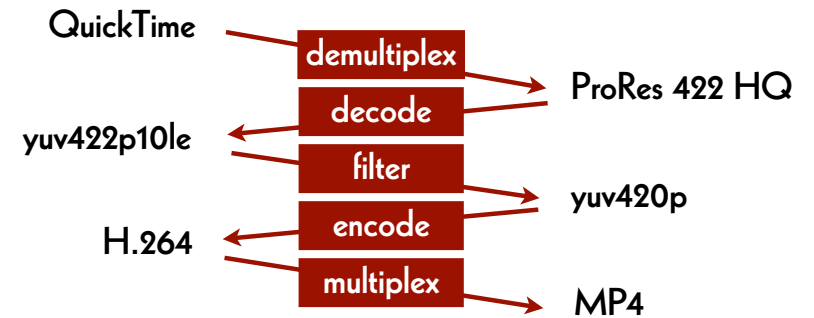
88

Audio Exemple



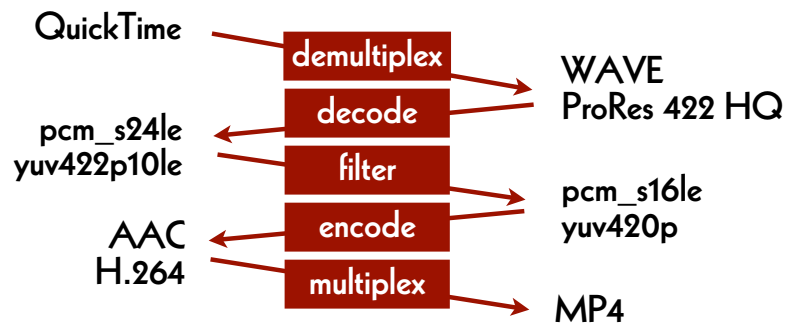
89

Video Exemple



90

Audio-Visual Exemple



91

Data Maintenance

92

Plan the Next Migration

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

93

File Naming (Example)

- title_codec.container
- title_codec_container_algorithm.txt
- film_H264.mp4
- film_H264_mp4_md5.txt

94

Checksums

cryptographic

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

non-cryptographic

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

95

Longterm

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

96

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)

97

read | script | write

script to modify

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

99

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

98

#1: ProRes-born Content

from:

- ProRes stored in a QuickTime (.mov) container

to:

- ProRes stored in a Matroska (.mkv) container

100

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

101

#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

103

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.

102

Container and Codec

- read file from source LTO
- demultiplex file
- decode file
 - Y'CbCr, 4:2:2, 8 bit, uyvy422
- encode file
- multiplex file
- write file to destination LTO

104

Container and Codec

- read file from source LTO
- demultiplex file
- decode file
 - $Y'CbCr$, 4:2:2, 10 bit, yuv422p10le
- encode file
- multiplex file
- write file to destination LTO

105

#3: Filename

from:

- Title_YUV422.mkv

to:

- Title_YCbCr422_9d5084b5b0a08d5022b39e0e75241d12.mkv

106

Always remember:

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

107

Coda

108

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

109

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

110

Acknowledgements (2)

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

111

AV Preservation by
reto.ch

Sandrainstrasse 3
3007 Bern
Switzerland

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



112