Multimedia

§8 Image compression

Prof. Dr. Georg Umlauf

Contents

§8.1 Basics of image compression

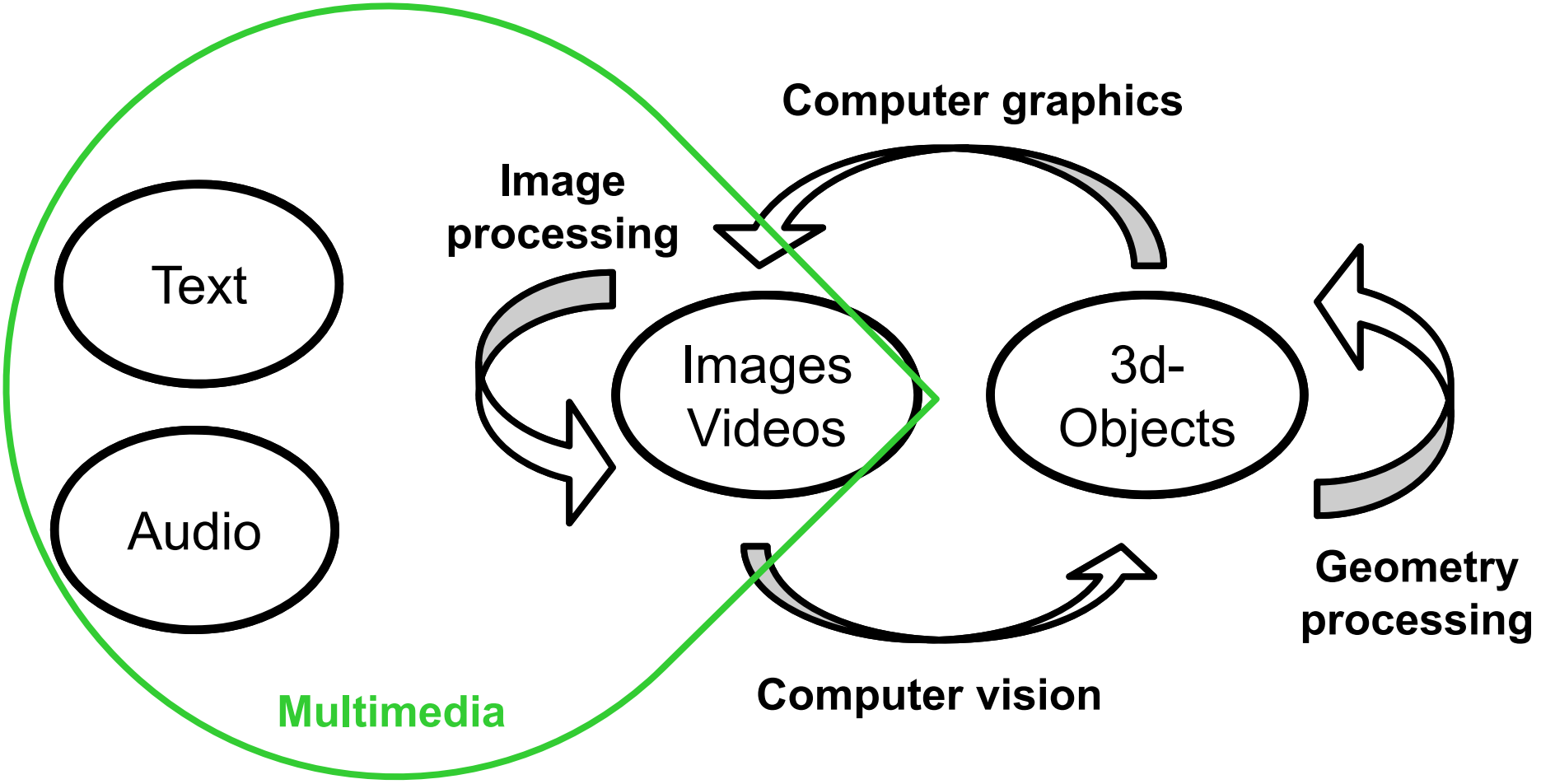


Image and graphics (1)

- **Image:** Structured in N rows of M **picture elements**, aka. **pixel**.
  - Abstractly: continuous function, that describes a rectangular range in the plane of the real world.
- **Image formats:**
  - Image recording format:
    - Spatial resolution ($N \times M$ Pixel)
    - Color depth (bits / pixel)
  - Image storage format:
    - Pixel matrix, color depth
    - Format of representation: RGB, reference to color tables, index
    - Compression
    - Meta-data (e.g. author, watermark, etc.)

Image and graphics (2)

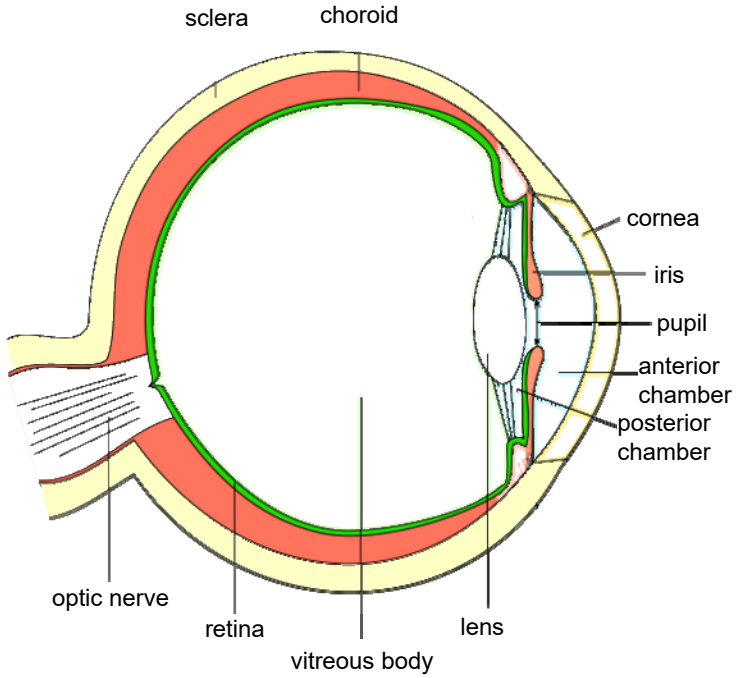
- **Graphics:** Structured in **graphical primitives** (point, line, rectangle, ellipse, Text) and its attributes (line width, color, etc.).
  - declarative description of image content
- Usage of graphical primitives of a graphical description language
  - Positioning („go to point x,y“)
  - Drawing (point, line, form)
  - Definition of graphical properties (color, patterns, line style)
- Examples
  - Computer Graphics Metafile (CGM)
  - Proprietary Formats (e.g. Corel Draw, Adobe Illustrator, Freehand)
  - Postscript/pdf
  - Web-based vector graphics: Flash, vrml

Contents

§8.2 The eye

Perception of light in two steps

1. Perception of the stimulus by two types of receptors on the retina:
   - Rods: Black-and-white vision, also of low intensities (ca. 120 Mio),
   - Cones: Color perception (ca. 6.5 Mio).
2. Processing of the stimulus in multiple steps:
   - Contrast enhancement at the outlet of the retina,
   - Interpretation in the visual cortex of the brain.

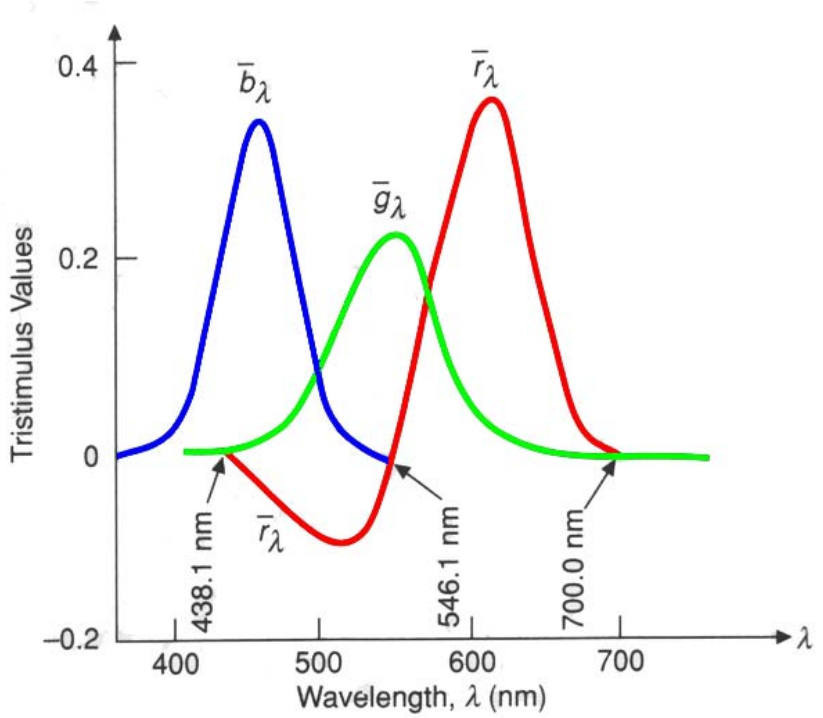


Source: Wikipedia

Color perception

- Cones for color perception:
  - blue-sensitive: 4% at 430 nm
  - green-sensitive: 32% at 530 nm
  - red-sensitive: 64% at 560 nm

➡ Every color can be generate as a (weighted) combination of three primary colors.



Color functions of the human eye to combine the complete human spectrum.

Source: Foley et al.

§8.2 The eye

- **Objective color characteristic:**
  - **Dominant wave length:** The wave length in the spectrum, with the maximum radiation power.
  - **Excitation purity:** Ratio of dominant wave length to the white content, e.g. white/gray levels → 0% purity.
  - **Luminance:** Radiation energy [energy/surface area]

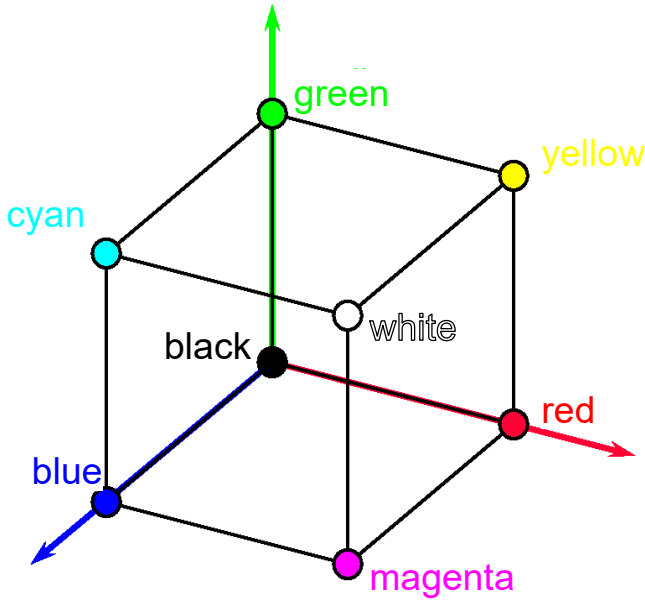
- **Subjective color characteristic:**
  - **Hue/tone:** Differentiates between different color patterns and pure colors (e.g. red, yellow, green, blue, etc.)
  - **Saturation:** Distance of a color from a grey level of the same intensity, e.g. rose is less saturated than red.
  - **Lightness/brightness:** Intensity of the perceived total energy flow
    - **Lightness:** Light intensity of a reflecting object.
    - **Brightness:** Light intensity of a light emitting object.

Contents

RGB color model

- Uses the basic colors red R, green G and blue B for **additive color generation** (e.g. computer or TV screens).
  - Description of a color by a triple (R, G, B) of weights $R, G, B \in [0,1]$.

    $(0,0,0)$ = black $(1,1,1)$ = white
    $(1,0,0)$ = red $(0,1,1)$ = cyan
    $(0,1,0)$ = green $(1,0,1)$ =magenta
    $(0,0,1)$ = blue $(1,1,0)$ = yellow

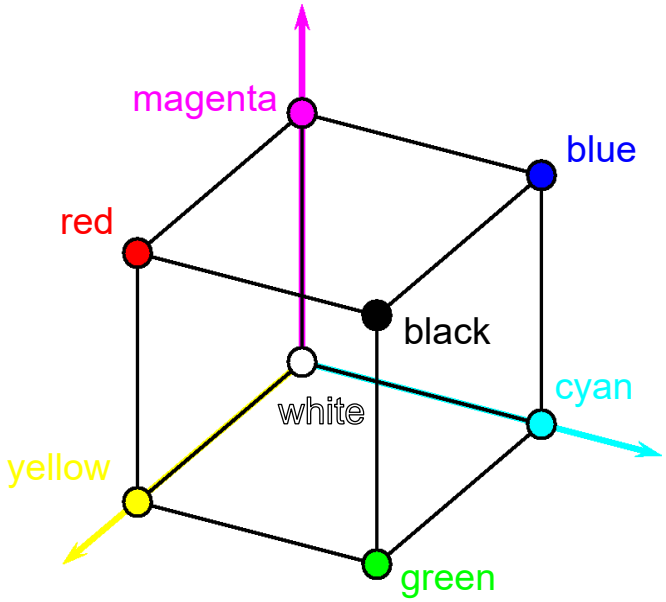


  - In the computer e.g. 8 bit per basic color, i.e. $0 \leq R, G, B \leq 255$.
- Set of all specified colors can be visualized as a 3d cube (color gamut).
  - It does not cover all perceivable colors.

CMY(K) color model

- Uses the basic colors cyan C, magenta M and yellow Y for **subtractive color generation** (e.g. printing).
- Relation to the RGB model:

$$(\mathrm{C}\ \mathrm{M}\ \mathrm{Y}) = (1\ 1\ 1) - (\mathrm{R}\ \mathrm{G}\ \mathrm{B}).$$



- Extension with the color black K (black as ***k***ey color) for printing:

$$\mathrm{K} = \min(\mathrm{C}, \mathrm{M}, \mathrm{Y}), \quad \mathrm{C}' = \mathrm{C} - \mathrm{K}, \quad \mathrm{M}' = \mathrm{M} - \mathrm{K}, \quad \mathrm{Y}' = \mathrm{Y} - \mathrm{K}.$$

YUV color model (1)

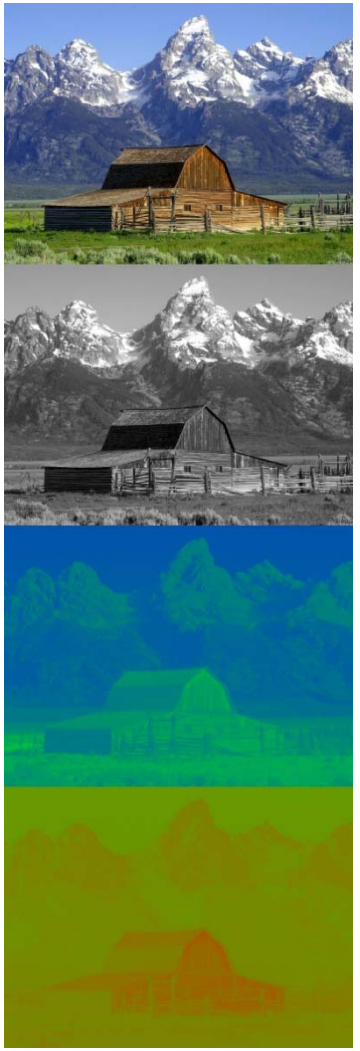
- Used for analog TV-transmissions to enhance the transmission efficiency and downwards compatibility for black-white-TVs:
  - **Luminance signal** Y

    $$\mathrm{Y} = 0.299\,\mathrm{R} + 0.587\,\mathrm{G} + 0.114\,\mathrm{B}.$$

    Used in black-white-TVs.
  - **Chrominance signals** U and V (for analog PAL)

    $$\mathrm{U} = 0.493\,(\mathrm{B} - \mathrm{Y}),$$

    $$\mathrm{V} = 0.877\,(\mathrm{R} - \mathrm{Y}).$$
  - Analog: YIQ for analog NTSC,
    YPbPr for analog video transmission,
    YCbCr for digital PAL/NTSC.



Source: Wikipedia

PGF-YUV color model (2)

RBG

R: ~30%

G: ~40%

B: ~30%

YUV

Y: ~90%

U: ~5%

V: ~5%

% of the total energy

Source: C. Stamm

YCbCr color model

- Used in the jpg format.
  - Luminance signal Y

$$Y = 0.299\,R + 0.587\,G + 0.114\,B.$$

  - **Chrominance blue/red** Cb and Cr

$$Cb = -0.1687\,R - 0.3313\,G + 0.5\,B,$$

$$Cr = 0.5\,R - 0.4187\,B - 0.0813\,B.$$

  - Advantages:
    - Better decorrelation of the chrominance components than for YUV

      and
    - smaller variance of the chrominance components.

Source: Wikipedia

YCgCo color model

- Used in the mpg-format.
  - Luminance signal Y

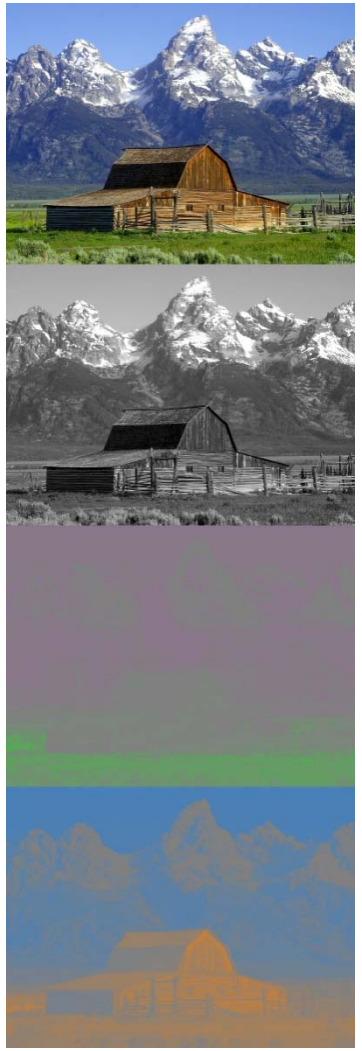
$$\mathrm{Y} = \frac{1}{4}\mathrm{R} + \frac{1}{2}\mathrm{G} + \frac{1}{4}\mathrm{B}.$$

  - **Chrominance green/orange** Cg and Co

$$\mathrm{Cg} = -\frac{1}{4}\mathrm{R} + \frac{1}{2}\mathrm{G} - \frac{1}{4}\mathrm{B},$$

$$\mathrm{Co} = \frac{1}{2}\mathrm{R} - \frac{1}{2}\mathrm{B}.$$

  - Advantages:
    - Simpler color space transformations and
    - better decorrelation of the chrominance components than for any other Yxx color model.

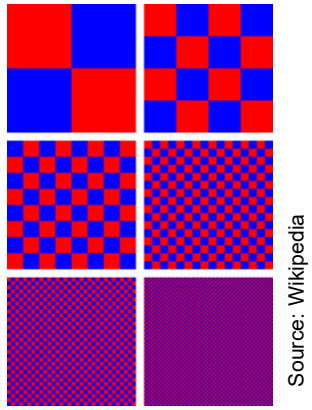


Source: Wikipedia

Contents

Properties of raster images (1)

- Maximum image size
- Resolution
- Color depth: bits per pixel (bpp)
  - **Dithering:** Generate impression of larger color depth using patterns.
  - *True Color*: as of 24 bit color depth
- Color storage / color table
  - Pseudo-colors: with color table
    - Adaptation to displays with low color depth
    - Usage of standardized color tables (Windows, Mac, Netscape – „browser-safe colors“)
  - **Alpha-Channel** for transparency



Dithering example

Properties of raster images (2)

- **Compression:** lossless/lossy
- **Suitable for streaming:** for low transfer rates progressive image build-up
- **Suitable for animation:** allows for coding of single images and image sequences (animations)
- **Container-Format:** supports multiple image code formats

Examples

- Properties of various image formats (in standard used)

Format	Color-depth	Alpha-channel	Compression	Streaming	Animation	Container
BMP	1/4/8/24	–	no rsp. lossless	–	–	–
GIF	8	+	lossless	–	+	–
PNG	24/48	+	lossless	+	–	–
TIFF	32	+	any	–	–	+
JPG	24/32	–	lossy	+	–	+

BMP-Format

(Windows **bitmap**)

- Pros:
  - Windows standard
  - Very simple
- Cons:
  - Compression: no compression or RLE
    - ➡ Large file sizes
  - Maximum image size: 32.767×32.767 Pixel
- Structure:

BITMAPFILEHEADER	BITMAPINFO	Image data
• Size of file	• Image size • Color depth: 1, 4, 8, 24 • optional color table • Optional compression: RLE	• Pixel values (RGB-triples) row-wise

GIF-Format (Graphic Interchange Format)

- Pros:
  - Lossless compression using LZW
  - 1-bit-transparency
  - Animation possible
- Cons:
  - Maximum 256 true-color-values in color table
  - Not stream-able (only row-interlacing)
  - Until 2004 license fee
- Structure:

HEADER	LOCAL-SCREEN-DESCRIPTOR	GLOBAL-COLOR-TABLE	
• Version	• Image size • Color table size • Background color	• Up to 256 true-color-values	• IMAGE-DESCRIPTOR-BLOCK (position, size, color table) • LOCAL-COLOR-MAP • GRAPHICS-CONTROL-EXTENSION-BLOCK (transparency, animation control) • RASTER-DATA-BLOCK (LZW-Codes, data length, indices for color table)

PNG-Format (Portable Networks Graphics)

- Advancement of GIF
- Draft for standardization of the World Wide Web Consortium (W3C)
- Pros:
  - Lossless compression using LZ77 of the **pixel differences**
  - 8/16-bit-transparency possible
  - Color table or up to 48 bit color depth per pixel
  - Stream-able using 2d-interlacing (Adam7)
  - Non-proprietary
  - 10-30% better compression rate than gif.
- Cons:
  - No animations
  - No random access to image data (e.g. sub-images)

TIF-Format

(Tagged Image File Format)

- Used in the printing industry, because is supports the CMYK-color model.
- Pros:
  - Transparency possible
  - Various tilings of the image in multiple resolutions for previews
  - Compression: optional LZW, RLE, jpg
  - Different color models
- Cons:
  - Not stream-able (although tiles)
  - Maximal 4GB
  - Very complex format

JPG-Format (1)

(Joint Photographic Experts Group)

- Developed 1991-1993 by the Joint Photographic Experts Group (ISO-10918).
- Compression method, rather than a proper file format.
- Pros:
  - High compression rates.
  - Supported by almost all systems.
  - Supported by all image processing software.
- Cons:
  - Bad compression for sharp edges in the image.
    - Unsuitable for line drawings with few colors and sharp edges.
  - Lossy compression
    - Loss of sharpness and color caused by the compression.
  - Block-artifacts

JPG-Compression (1)

Original image

Adaptation to the human perception

- Change color space
- Low pass filter & Chroma-subsampling

Filtering of areas with large contrast variations

- Blocking & DCT
- Quantization

Coding

- Lossless compression

Compressed image

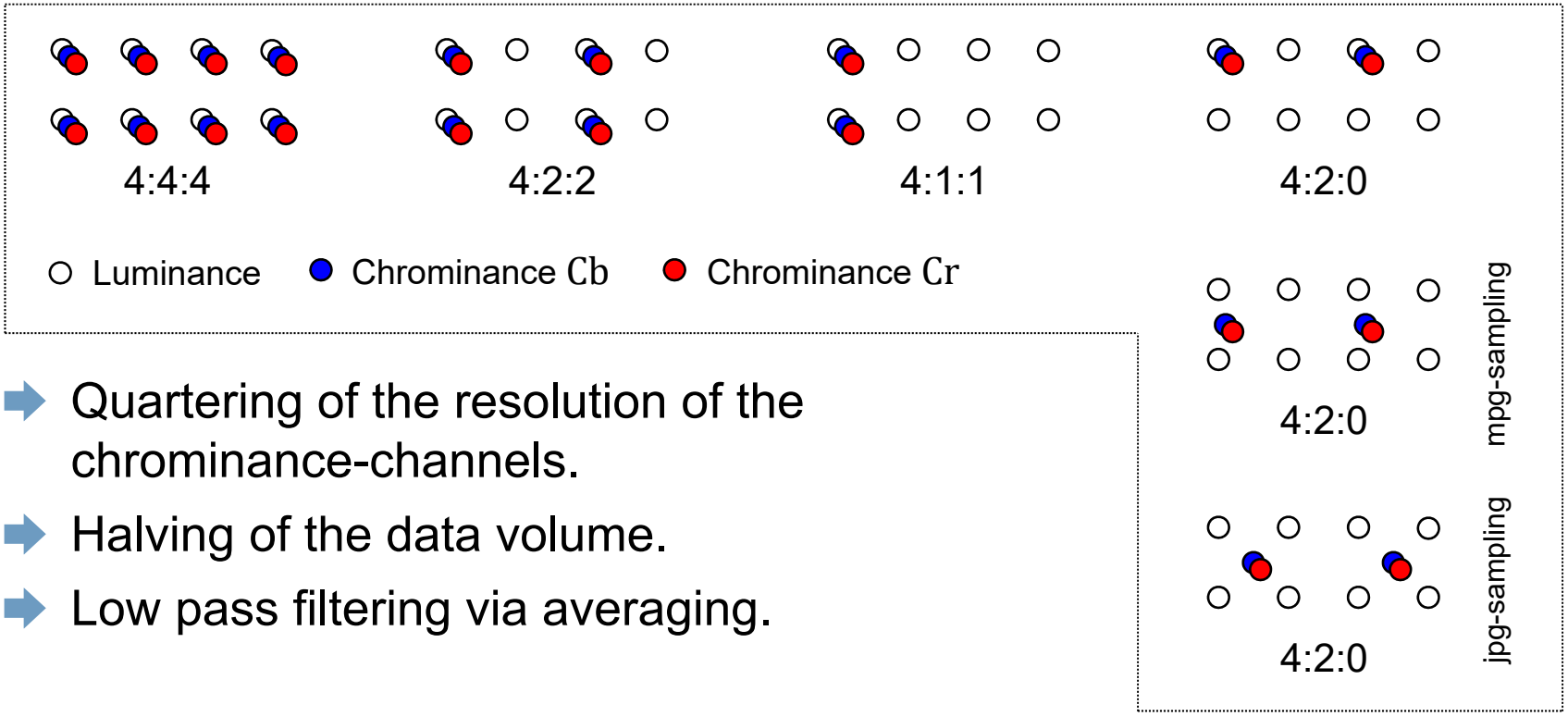
JPG-Compression (2)

1. **Change of color space**
   - Conversion from RGB color model to YCbCr color model, see §8/13.
     - Lossy because of rounding errors (and poss. sub-sampling).

2. **Low pass filter & chroma-subsampling**
   - Luminance information is perceived by humans with a higher resolution than color information.
   - ➡ Sub-sampling of chrominances Cb and Cr with the ratio 4:2:0.
   - ➡ Lossy compression stage!

JPG-Compression (3)

- **Sampling ratio A:B:C**: In a (2×A)-array of values B values are sampled in the first row and C values in the second.



- Quartering of the resolution of the chrominance-channels.
- Halving of the data volume.
- Low pass filtering via averaging.

JPG-Compression (4)

3. Blocking & DCT

a) Blocking

- Each color channel is subdivided into 8×8 blocks.
  - Poss. the image is enlarged and suitably extended (ringing-artifacts).

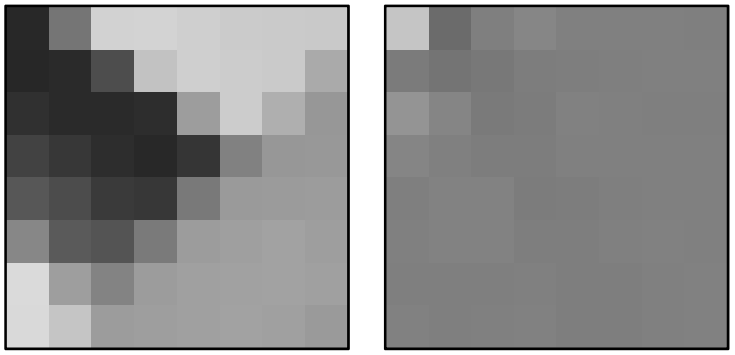
b) Discrete Cosine Transformation

- Apply the 2d-DCT-II to the 8×8 blocks.

➡ The resulting values represent the geometric structure of the image content.

  - In the extreme a single-colored area is represented by a single value.

➡ Lossy because to rounding errors.

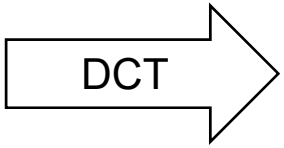


Original | After DCT

JPG-Compression (5)

139	144	149	153	155	155	155	155
144	151	153	156	159	156	156	156
150	155	160	163	158	156	156	159
159	161	162	160	160	159	159	155
159	160	161	162	162	155	155	157
161	161	161	161	160	157	157	157
162	162	161	163	162	157	157	157
162	162	161	161	163	158	158	158

Original image (luminance)



235	-1	-12	-5	2	-2	-3	1
-22	-17	-6	-3	-3	0	0	-1
-11	-10	-2	1	0	-1	-1	0
-7	-2	0	1	1	0	0	0
-1	-1	0	1	0	-1	1	1
2	0	2	0	-1	1	1	-1
-1	0	0	-1	0	2	1	-1
-3	2	-4	-2	2	1	-1	0

Luminance coefficients
after DCT (rounded)

JPG-Compression (6)

4. Quantization

- The width of the quantization intervals varies for each DCT-coefficient according to perception-based quantization.
- The uniform quantization mapping is changed to

$$Q(x_{ij}) = \left\lfloor \frac{x_{ij}}{\Delta_{ij}} + \frac{1}{2} \cdot \operatorname{sign}(x_{ij}) \right\rfloor,$$

  where Δ_{ij} is the quantization value for the (i,j)-coefficients.
- The matrix (Δ_{ij}) is the so-called **quantization table**.

➡ Lossy compression step!

JPG-Compression (7)

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Quantization table for luminance

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

Quantization table for both chrominances

JPG-Compression (8)

235	-1	-12	-5	2	-2	-3	1
-22	-17	-6	-3	-3	0	0	-1
-11	-10	-2	1	0	-1	-1	0
-7	-2	0	1	1	0	0	0
-1	-1	0	1	0	-1	1	1
2	0	2	0	-1	1	1	-1
-1	0	0	-1	0	2	1	-1
-3	2	-4	-2	2	1	-1	0

Luminance coefficients
after DCT (rounded)

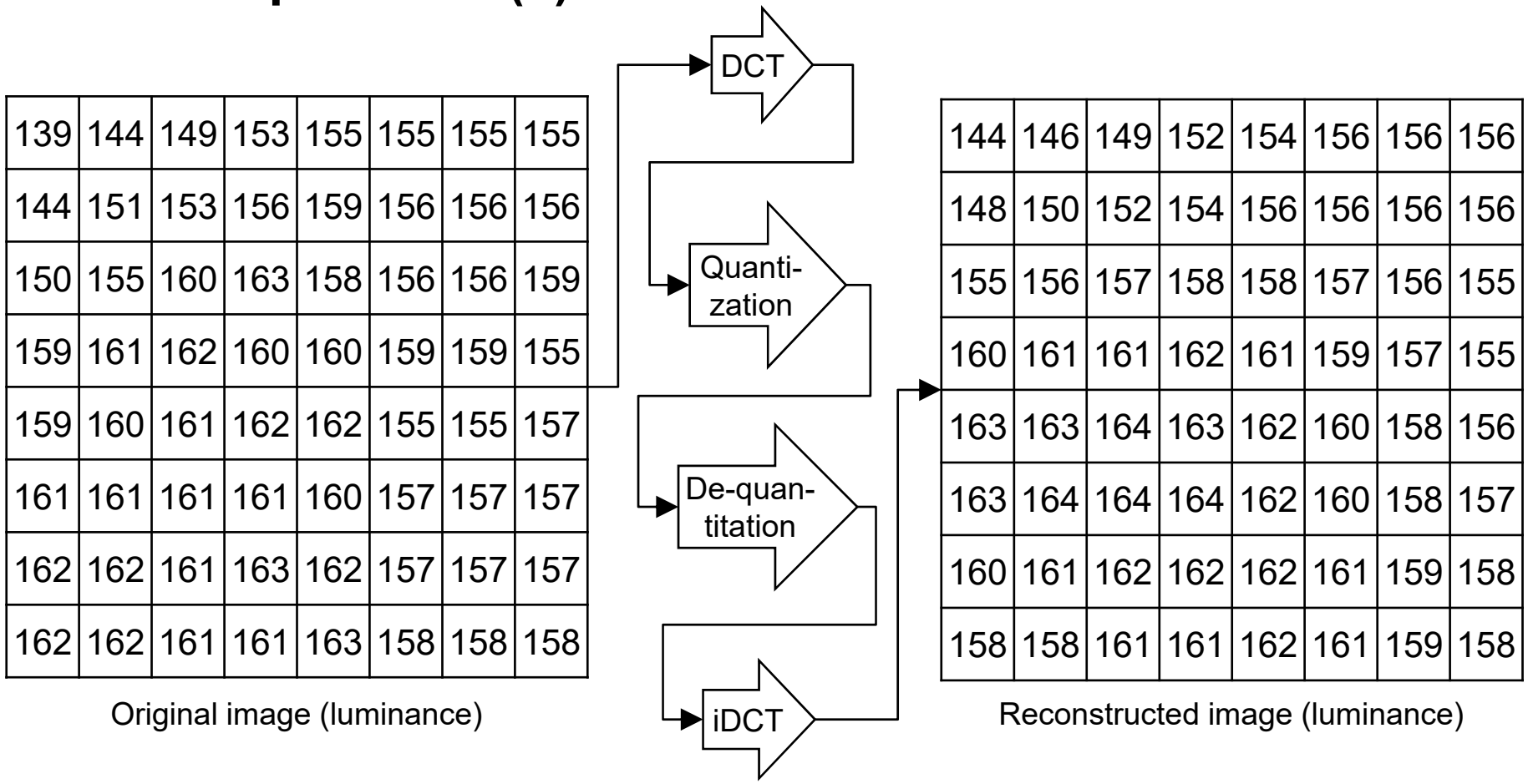
16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Quantization

15	0	-1	0	0	0	0	0
-2	-1	0	0	0	0	0	0
-1	-1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Quantized values (normalized)

JPG-Compression (9)



139	144	149	153	155	155	155	155
144	151	153	156	159	156	156	156
150	155	160	163	158	156	156	159
159	161	162	160	160	159	159	155
159	160	161	162	162	155	155	157
161	161	161	161	160	157	157	157
162	162	161	163	162	157	157	157
162	162	161	161	163	158	158	158

Original image (luminance)

144	146	149	152	154	156	156	156
148	150	152	154	156	156	156	156
155	156	157	158	158	157	156	155
160	161	161	162	161	159	157	155
163	163	164	163	162	160	158	156
163	164	164	164	162	160	158	157
160	161	162	162	162	161	159	158
158	158	161	161	162	161	159	158

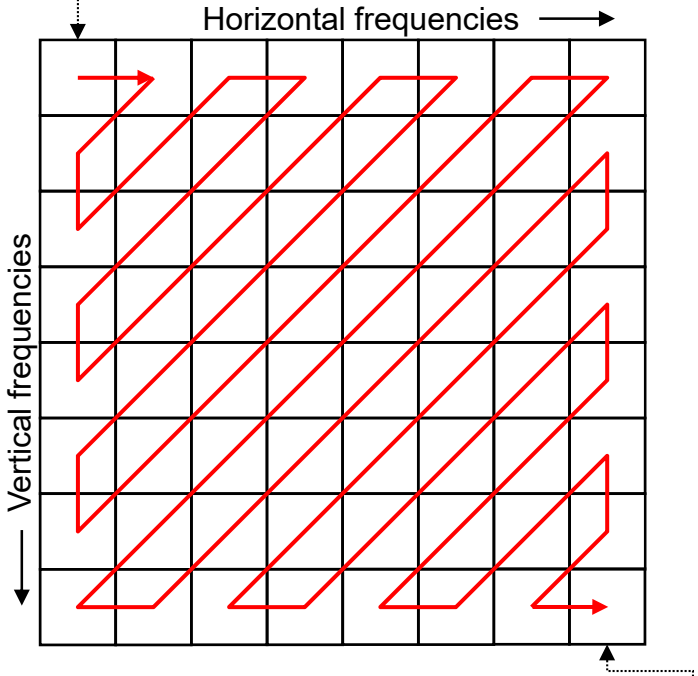
Reconstructed image (luminance)

JPG-Compression (10)

5. **Lossless compression**
   a) **Resorting of the coefficients**
   b) **Coding of the DC-components**
      - Huffman coding of the differences of the DC-components
   c) **Coding of the AC-components**
      i. RLE and subsequent
      ii. Huffman or arithmetic coding

➡ Lossless compression step!

DC-component (background):
- Is coded as difference to the DC-component of the previous block.



AC-component with maximal frequency

JPG Example

83.261 Byte, 2,6:1

15.135 Byte, 15:1

9.553 Byte, 23:1

4.787 Byte, 46:1

1.523 Byte, 144:1

Source: Wikipedia

JPG Variants

- Quantization: Choice of different quantization tables
- Entropy coding: Choice of Huffman-coding or arithmetic coding
- Coding depth: 8 or 12 bit / Pixel
- Image build-up
  - Sequential: Transmit the coefficients of each block subsequently.
  - Progressive: Transmit first the first coefficients of all blocks, then the second coefficients of all blocks, etc.

JPG extensions

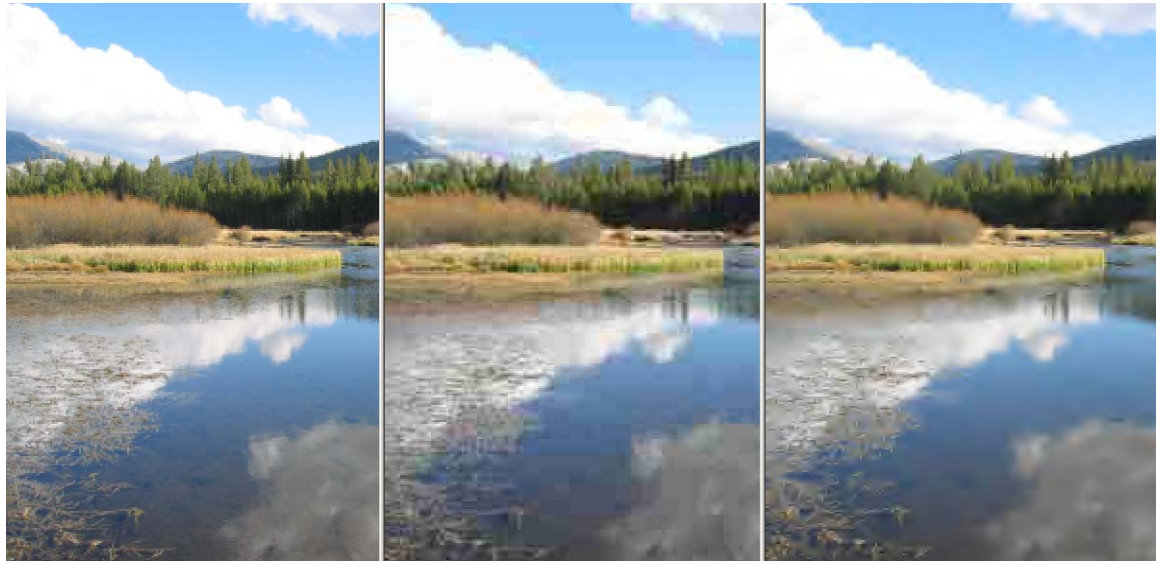
- JPG stereoscopic:
  - One image of double width, that contains two half-images of half the total width.
  - The left half-image is for the right eye and vice versa.

- JPG multi-picture format (.mpo):
  - JPEG-based format for multi-view images.
  - It contains two or more JPEG files appended together.

- JPEG-LS
  - Method for lossless compression based on prediction coding (median adaptive prediction) and Golomb-codes.

JPG 2000

- DWT instead of DCT ➡ less block artifacts
- Particular image regions of interest (ROI) can be compress with higher quality.
- Up to 256 color channels.
- Transparency
- Larger computational costs
- Quality depends on the implementation

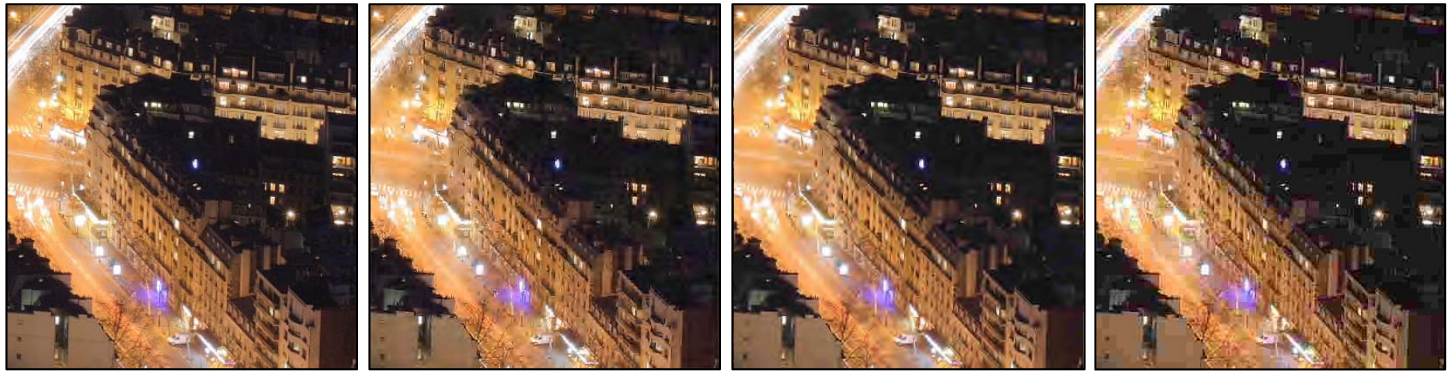


Original jpg jpg2000

Source: Wikipedia

JPG XR (eXtended Range) (lossy/lossless) (.jxr, *.hdp, .wdp*)

- Similar to JPG2000, patented by Microsoft.
- DCT-variant without round-off errors (PCT).
- 4×4 core transform in a two-level hierarchy within 16×16 macroblock.
- Before PCT optional overlap pre-filtering (POT) of 4×4 blocks offset by 2 Pixels in both directions.
- Prediction of coefficients across transform blocks applied to DC coefficients and to an additional row or column of AC coefficients.
- Color space transformation without round-off errors.



Original, png | JPG XR, 9.16KB | JPG2000, 9.16KB | JPG, 9.14KB

Source: Wikipedia

Goals

- How are the RGB-, CMY(K)-, Yxx-color models defined and how do they differ?
- What is color depth? What is „true color“?
- Which compression methods are used in bmp, gif, png?
- Which formats are suitable formats for streaming or animations?
- What are the main building blocks of jpg-coding?
- Why yields jpg-compression better compression rates than other compression methods?
- What are the typical artifacts of jpg-compression?