

Image formats

David Bařina

March 22, 2013

Contents

1 Notions

2 Uncompressed formats

3 Lossless formats

4 Lossy formats

Introduction

- the need for compression: FullHD image with resolution $1920 \times 1080 \times 3 \approx 6 \text{ MB}$
- compression: none, lossless, lossy
- formats: BMP, TIFF, GIF, PNG, JPEG-LS, JPEG, JPEG-2000, . . .



Compression

- data (BMP, TIFF)
- data → RLE (BMP, TGA, TIFF)
- data → prediction → context EC (JPEG-LS)
- data → prediction → RLE → EC
- data → prediction → dictionary method (PNG)
- data → dictionary method (GIF)
- data → transform → RLE+EC (JPEG)
- data → transform → context EC (JPEG 2000)

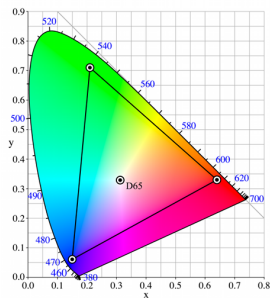
Color

- color
- primary colors of CRTs: R, G, B
- luminance
- ITU-R BT.601

$$Y = 0.299R + 0.587G + 0.114B$$

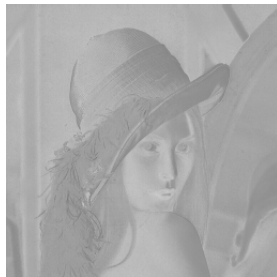
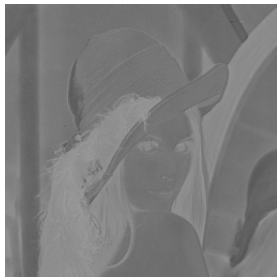
Color models and spaces

- model
- CIE xy chromaticity diagram
- space, gamut



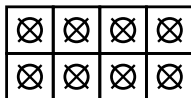
- RGB transform

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} +0.299 & +0.587 & +0.114 \\ -0.16875 & -0.33126 & +0.5 \\ +0.5 & -0.41869 & -0.08131 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

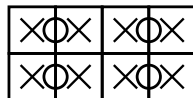


Chroma subsampling

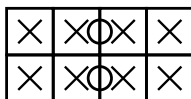
- Y vs. C_b, C_r
- unobservable by human
- block, macropixel
- $J:a:b$, 4:4:4, 4:2:2, 4:1:1, 4:2:0
- centroids



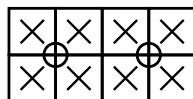
4:4:4



4:2:2



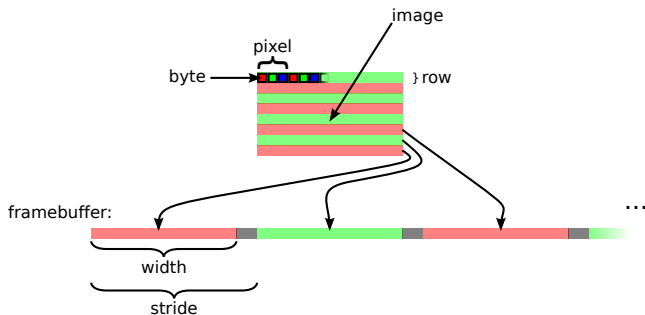
4:1:1



4:2:0

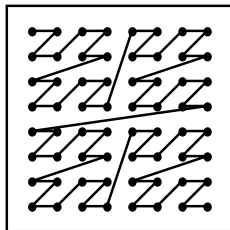
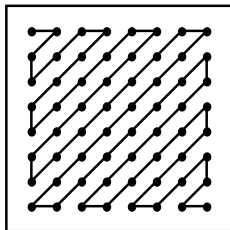
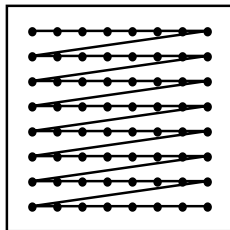
Pixelformat

- model + memory layout
- RGB24 vs. BGR24
- planar: $R_0 R_1 R_2 \dots G_0 G_1 G_2 \dots B_0 B_1 B_2 \dots$
- packed: $R_0 G_0 B_0 R_1 G_1 B_1 \dots$



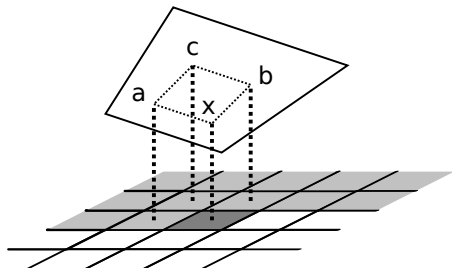
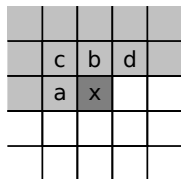
Multidimensional data

- linearization
- raster scan (everywhere), zig-zag (JPEG), z-curve (EZW, SPIHT), other (EBCOT)



Predictors

- prediction of symbol x , min. error
- predictor order, domain
- linear and non-linear predictors



$$p = a + b - c$$

Predictors

- trivial: a, b, c, d
- average, median
- p (fit plane)
- MED (JPEG-LS)

$$\hat{x} = \begin{cases} \min(a, b) & : c \geq \max(a, b) \\ \max(a, b) & : c \leq \min(a, b) \\ p & \end{cases}$$

- Paeth (PNG)

$$\hat{x} = \arg \min_{x \in a, b, c} |p - x|$$

Transforms

- sparse image representation
- lossy methods (quantization)
- often DCT, DWT
- blocks, blocking artifacts

DCT (discrete cosine transform)

$$g_{k,j}[n, m] = \lambda_k \lambda_j \frac{2}{N} \cos \left[\frac{k\pi}{N} \left(n + \frac{1}{2} \right) \right] \cos \left[\frac{j\pi}{N} \left(m + \frac{1}{2} \right) \right]$$

$$c[k, j] = \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} f[n, m] g_{k,j}[n, m]$$

$$f[n, m] = \sum_{k=0}^{N-1} \sum_{j=0}^{N-1} c[k, j] g_{k,j}[n, m]$$

$$0 \leq k, j < N \quad \lambda_k = \begin{cases} 1/\sqrt{2} & : k = 0 \\ 1 & \end{cases}$$

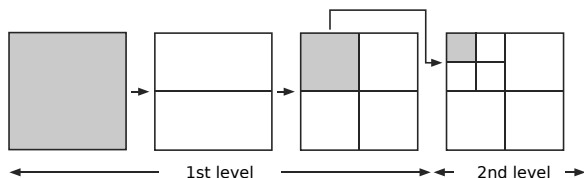
DCT (discrete cosine transform)



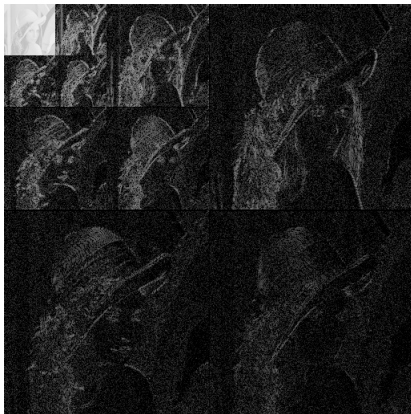
DWT (discrete wavelet transform)

$$a_{j+1} = (a_j * h) \downarrow 2$$

$$d_{j+1} = (a_j * g) \downarrow 2$$

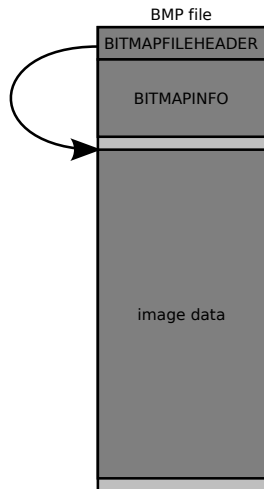


2-D DWT



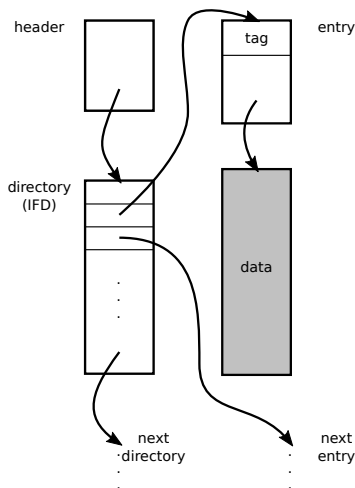
BMP

- BMP (DIB)
- uncompressed, RLE for palette
- stores: 1, 4, 8-bit indices into palette; 16, 24, 32 RGB (BGR order)
- bottom-top row order, 32-bit aligned
- Windows API, MSDN
- structures BITMAPFILEHEADER (header, data offset), BITMAPINFOHEADER (dimensions, compression)



TIFF

- TIFF (Tagged Image File Format), 1986, 1992, Adobe
- several pages, fragmented data, transparency, colorometry
- compression: RLE, LZW, JPEG
- structure: analogy to FS
- header, IFD, entry, tag, type+count, data
- types: byte (8-bit), ascii (characters), short (16), long (32), ...

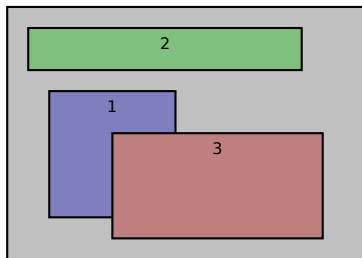


TIFF

- tags: Compression (none, PackBits, LZW, JPEG), ImageLength, ImageWidth, BitsPerSample, ColorMap, SamplesPerPixel
- fragmentation: stripes, tiles
- stores: bi-level (2), gray-scale (16, 256), palette (16, 256), RGB (16M), CMYK (2^{32}), YC_bC_r (2^{24})
- more tags: Artist (image creator), Copyright, DateTime, ImageDescription, Make (device manufacturer), Model (device model), Orientation, Software, ...
- PackBits: RLE variant, per bytes
 - ▶ read n
 - ▶ 0 to 127: copy next $n + 1$ bytes to output
 - ▶ -127 to -1: output next byte ($-n + 1$) times

GIF

- GIF (Graphics Interchange Format), CompuServe, 1987 (GIF87a), 1989 (GIF89a)
- LZW (patents, PNG)
- stores: frames (palette with indices up to 8-bits), animations, True Color (16M)
- transparent pixels (89a)
- logical screen, frames, background color



GIF

- global and local color palletes
- LZW: clear code, EOF
- compression per rows
- progressive transfer: stripes of 8 rows, 4 passes

row sequentially	pass of interleaving	row by interleaving
1	1	1
2	4	5
3	3	3
4	4	6
5	2	2
6	4	7
7	3	4
8	4	8

- Google for: true color gif

PNG

- PNG (Portable Network Graphics), replacement for GIF
- Deflate
- extensible, interleaving, alpha channel

pixel type	channels	bit-depths
palette	indices	1, 2, 4, 8
grayscale	Y	1, 2, 4, 8, 16
grayscale w. transparency	Y, A	8, 16
true color	R, G, B	8, 16
true color w. transparency	R, G, B, A	8, 16

- compression: predictor (filter) + Deflate

- prediction: per rows, per bytes, 5 predictors

predictor	predicted value
none	0
left	b
top	a
avg.	$\lfloor (a + b)/2 \rfloor$
Paeth	(above)

PNG

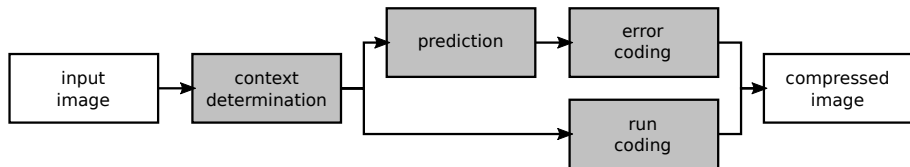
- interleaving: Adam7 (8×8 pixels, 7 passes, each pass double resolution)

1	6	4	6	2	6	4	6
7	7	7	7	7	7	7	7
5	6	5	6	5	6	5	6
7	7	7	7	7	7	7	7
3	6	4	6	3	6	4	6
7	7	7	7	7	7	7	7
5	6	5	6	5	6	5	6
7	7	7	7	7	7	7	7

- structure: blocks (type, size) – critical and ancillary
- block: size, name, data, CRC
- name: four characters (critical, registered, –, do not copy)
- critical: IHDR (header), PLTE (palette), IDAT (image data), IEND (footer)

JPEG-LS

- replacement for JPEG (lossless mode), not JPEG modification
- based on LOCO-I (LOW COMplexity LOSSless COMpression for Images)
- predictor + context EC (Rice), also RLE variant
- compression on rows
- generic method, stores e.g. RGB, palette images
- file format based on JIF (further)



JPEG-LS

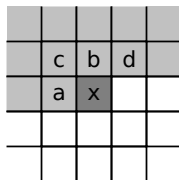
- gradient $D = (D_1, D_2, D_3)$ (smoothness, edge)

$$D_1 = R_d - R_b$$

$$D_2 = R_b - R_c$$

$$D_3 = R_c - R_a$$

- chooses compression mode (zero gradient)
- quantize gradient: $D_i \rightarrow Q_i \in \langle -4, 4 \rangle$
- $9^3 = 729$, merge, 365 contexts

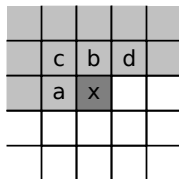


JPEG-LS

- MED predictor
- value on edge, or plane fitting

$$P_x = \begin{cases} \min(R_a, R_b) & : R_c \geq \max(R_a, R_b) \\ \max(R_a, R_b) & : R_c \leq \min(R_a, R_b) \\ R_a + R_b - R_c & \end{cases}$$

- subsequent context dependent correction of prediction
- near-lossless method: quantization of prediction error



JPEG-LS

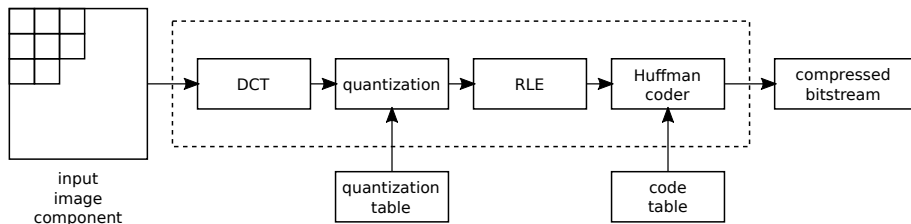
- codes error of prediction
- limited Golomb-Rice code (max. length)
- parameter of Rice code dependent on context
- negative values mapping

$$E_M = \begin{cases} 2 \cdot |E| & : E \geq 0 \\ 2 \cdot |E| - 1 & : E < 0 \end{cases}$$

JPEG

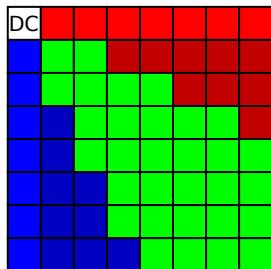
- JPEG committee (Joint Photographic Experts Group), 1992
- standard ISO/IEC 10918-1, recommendation CCITT/ITU-T T.81
- incorrectly file format (correctly JFIF or Exif)
- lossy and lossless compression
- sequential and progressive transmission
- defines JIF
- baseline process: mandatory, 8-bits, Huffman code with 2 AC + 2 DC tables, sequential
- extended process: 8 + 12-bits, 4 AC + 4 DC

JPEG

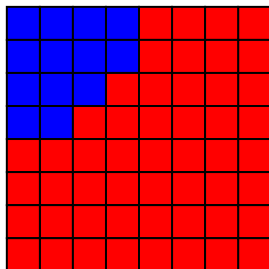


- $YCbCr$ input
- chroma subsampling (4:2:2 a 4:2:0)
- blocks 8×8 , artifacts
- DCT
- coefficients: 1 DC + 63 AC
- yet fully invertible

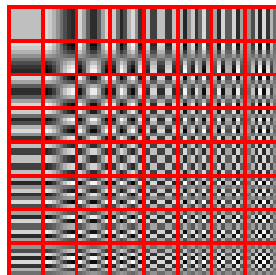
JPEG



(a) Edge direction



(b) Frequency



(c) Bases

- quantization, quantization table (JPEG parametr)

$$S_{v,u}^q = \text{round}(S_{v,u}/Q_{v,u})$$

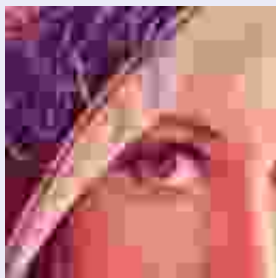
JPEG

16	11	10	16	124	140	151	161
12	12	14	19	126	158	160	155
14	13	16	24	140	157	169	156
14	17	22	29	151	187	180	162
18	22	37	56	168	109	103	177
24	35	55	64	181	104	113	192
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	199

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

JPEG

blocking artifacts



- zig-zag linearization, more zeros at the end, EOB
- DC differentially
- AC using RLE of leading zeros
- followed by arithmetic or Huffman coding (5–10%)
- Huffman coding: tables for DC and AC coefficients

JPEG

- 15+ zeros using ZRL
- AC/DC code = Huffman code (tables) + extra bits
- Huffman code cannot consist of only 1-bits
- code tables indexing
 - ▶ DC: category
 - ▶ AC: category + zero run length

category	range
0	0
1	-1, 1
2	-3, -2, 2, 3
3	-7... -4, 4... 7
4	-15... -8, 8... 15
5	-31... -16, 16... 31
...	
10	-1023... -512, 512... 1023
11	-2047... -1024, 1024... 2047

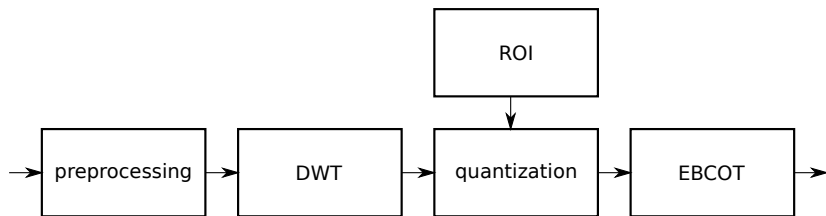
JPEG

JIF, JFIF, Exif, SPIFF

- bitstream, big endian
- consists of segments, max. 65 535 B
- segment preceded by marker $0\text{xff} + 0\text{x01}$ až 0xfe
- markers: APP0 (JFIF header), APP1 (Exif), SOF0 (image in baseline process), SOS (compressed data), RSTm (restart points), DHT (Huffman tables), DQT (quantization tables)
- JIF extended to JFIF and Exif (incompatible), JFIF+Exif
- thumbnails, metadata (device model)
- Exif: TIFF (thumbnail in JFIF)
- SPIFF not used

JPEG 2000

- successor of JPEG
- many improvements, better quality at the same compression ratio, more computationally demanding
- lossy and lossless compression
- officially ISO/IEC 15444-1, ITU-T T.800
- based on DWT



JPEG 2000

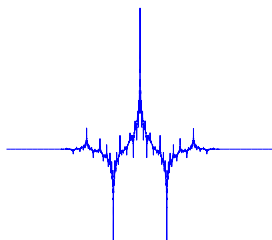
- color model transform using RCT, ICT

$$Y_r = \left\lfloor \frac{R + 2G + B}{4} \right\rfloor$$

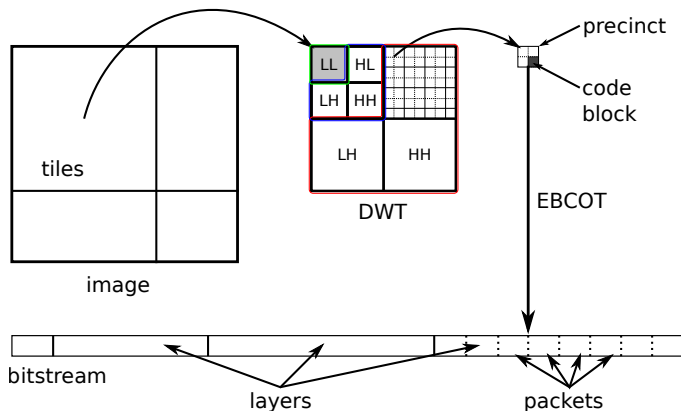
$$C_b = B - G$$

$$C_r = R - G$$

- division into tiles (blocking artifact)
- DWT (CDF 9/7, 5/3)

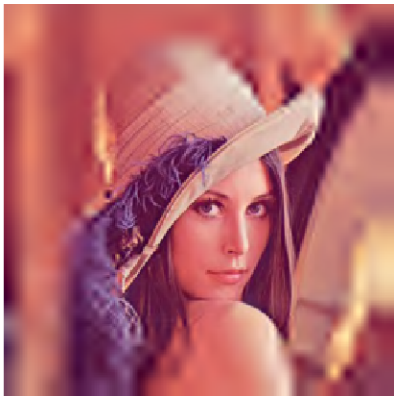
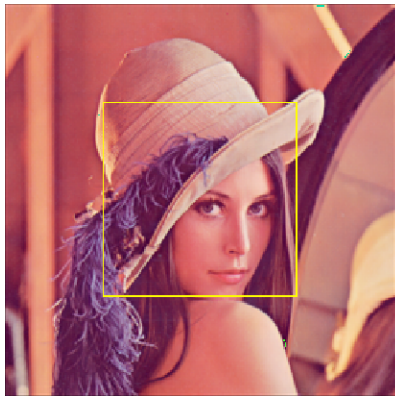


JPEG 2000



- decomposition level, resolution, subbands (LL, HL, LH, HH)
- quantization, ROI (Maxshift)
- precinct, code-block, EBCOT
- layer (quality), packet, bitstream

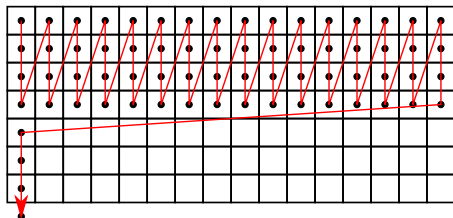
JPEG 2000



JPEG 2000

EBCOT

- profile-0: code blocks of size of 32×32 or 64×64
- bit-planes from MSB to LSB
- bits per stripes (4 bits tall)



- three passes: significance propagation, magnitude refinement, cleanup
- context AC, 8-bits context

D	V	D
H	X	H
D	V	D

JPEG 2000

- packets: component, tile, resolution, precinct, layer
- 4 axes of progressive transmission: quality, color component, resolution, position
- order can be changed
- bitstream compatible to JIF, but other markers
- JP2 file format
- metadata in XML (e.g. XMP)

JPEG 2000

Comparison with JPEG (108:1, ImageMagick)

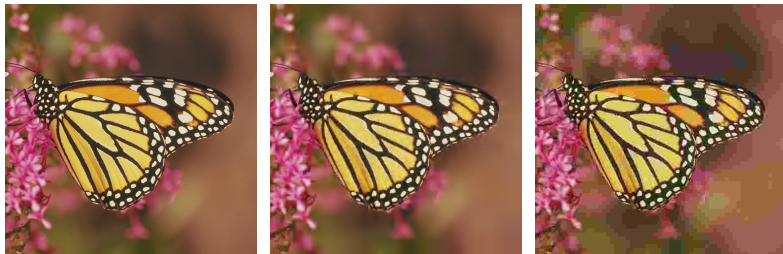


Figure: original, JPEG 2000, JPEG

JPEG 2000

Error resilience

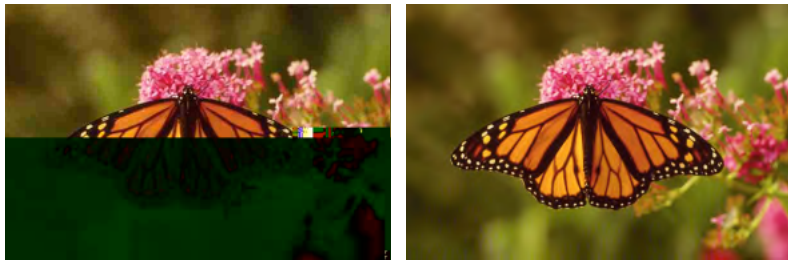
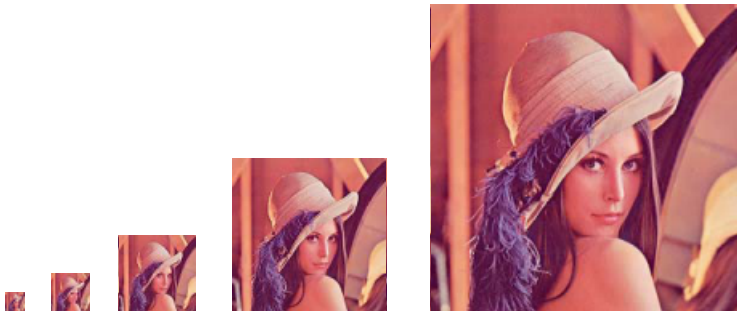


Figure: JPEG, JPEG 2000

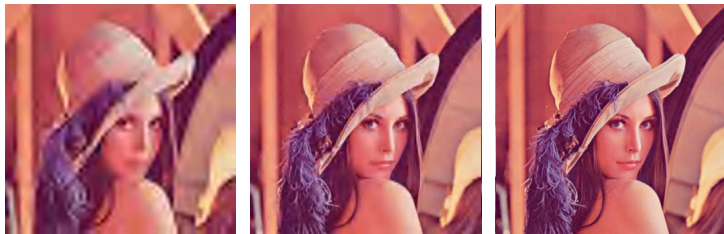
JPEG 2000

Progressive transmission (resolution)



JPEG 2000

Progressive transmission (quality)



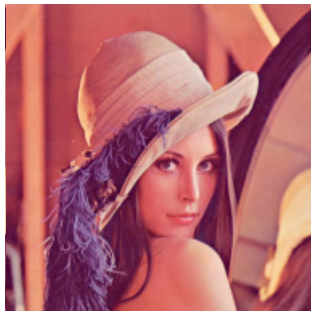
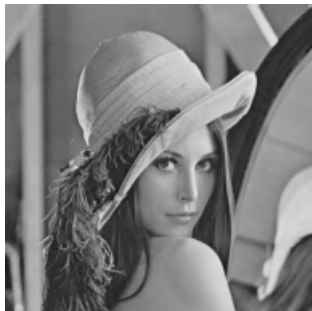
JPEG 2000

Progressive transmission (position)



JPEG 2000

Progressive transmission (components)



Summary

- colors, predictors
- BMP: uncompressed, RLE for palette images
- TIFF: directory structure, tags, LZW, RLE, JPEG
- GIF: LZW, 256 colors, 16M colors
- PNG: Deflate, predictors
- JPEG-LS: context, predictor, Golomb-Rice codes
- JPEG: $YCbCr$, DCT, RLE, Huffman code, JIF, segments, markers, JFIF, Exif
- JPEG 2000: DWT, EBCOT