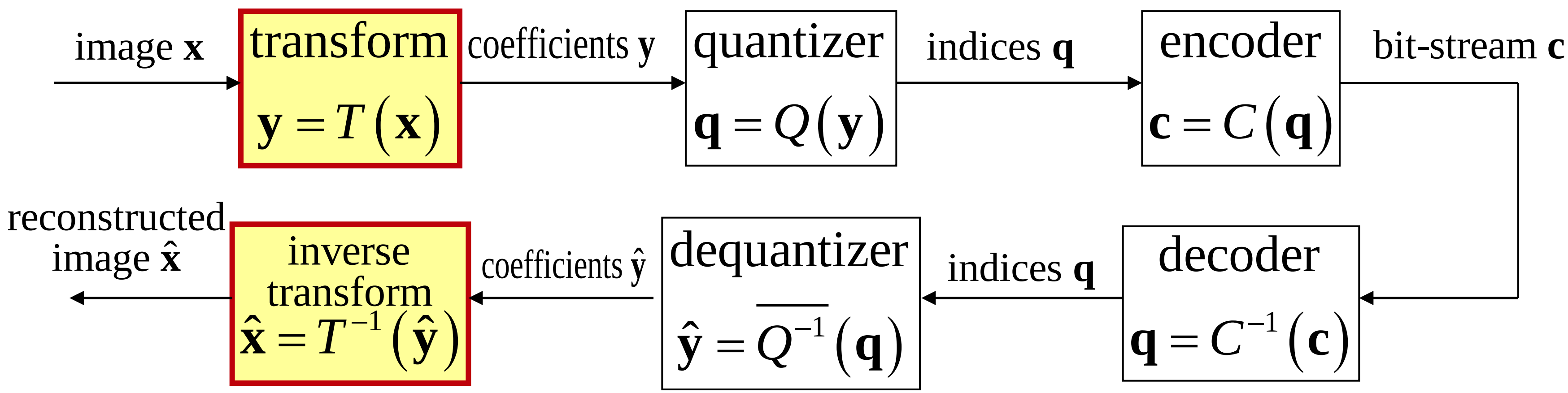
Image Compression

- Transform Coding
  - Typical codec structure
  - Comparison of transforms
  - Discrete cosine transform
  - Threshold coding
  - JPEG standard
- Subband Coding
  - Typical codec structure
  - JPEG 2000 standard

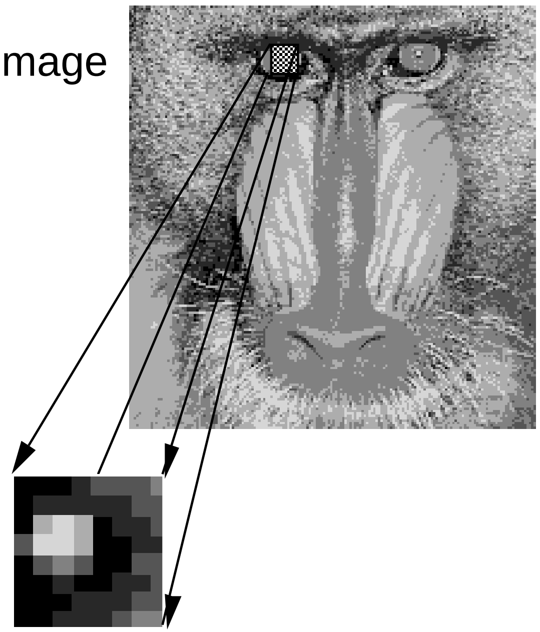
Transform Coding



- Transform *T(**x**)* usually invertible
- Quantization $Q(\mathbf{y})$ not invertible, introduces distortion
- Combination of encoder $C(\mathbf{q})$ and decoder $C^{-1}(\mathbf{c})$ lossless

Block-Based Transform Coding

original image



original image block

Transform A

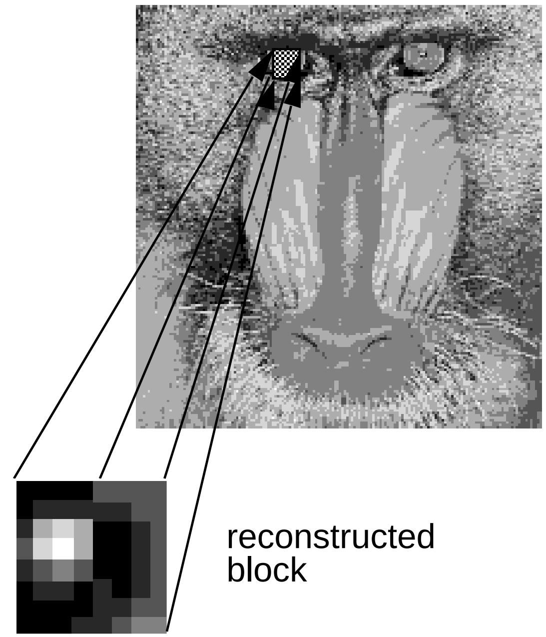
Transform coefficients

Quantization, entropy coding & storage or transmission

Quantized transform coefficients

Inverse transform A^{-1}

reconstructed block

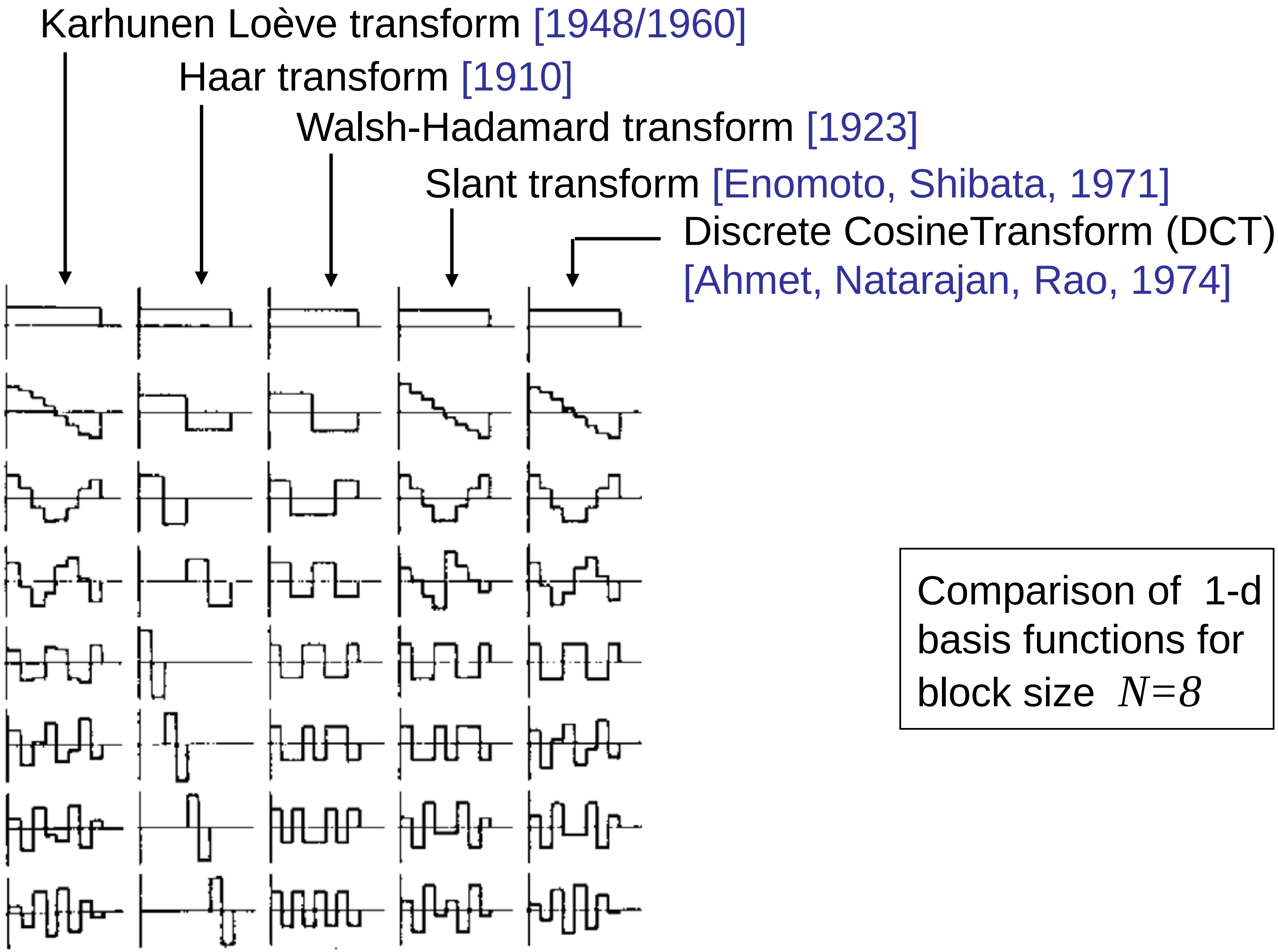


reconstructed image

Orthonormal Transforms

- Energy conservation
- Rotation of the coordinate system around the origin
- Unevenly distributed energy among coefficients
- KLT decorrelates transform coefficients
- KLT achieves optimum energy concentration
- Disadvantages of KLT:
  - depends on signal statistics
  - usually not separable for image blocks

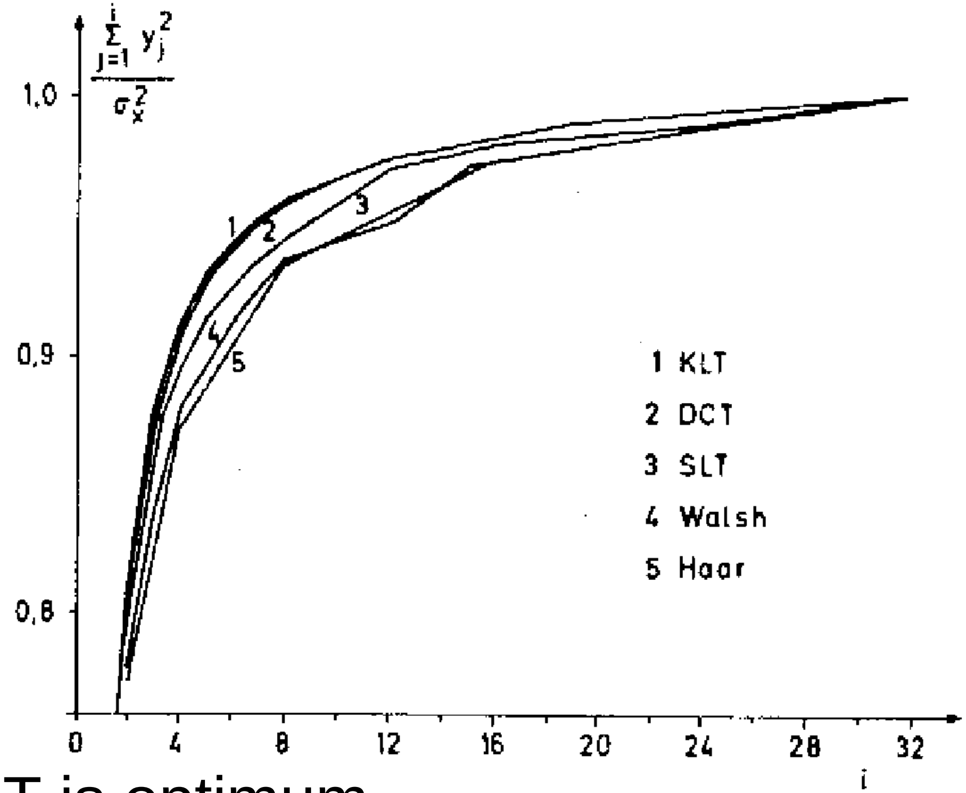
Comparison of Various Transforms



Comparison of 1-d basis functions for block size $N=8$

Comparison of Various Transforms

- Energy concentration measured for typical natural images, block size 1x32

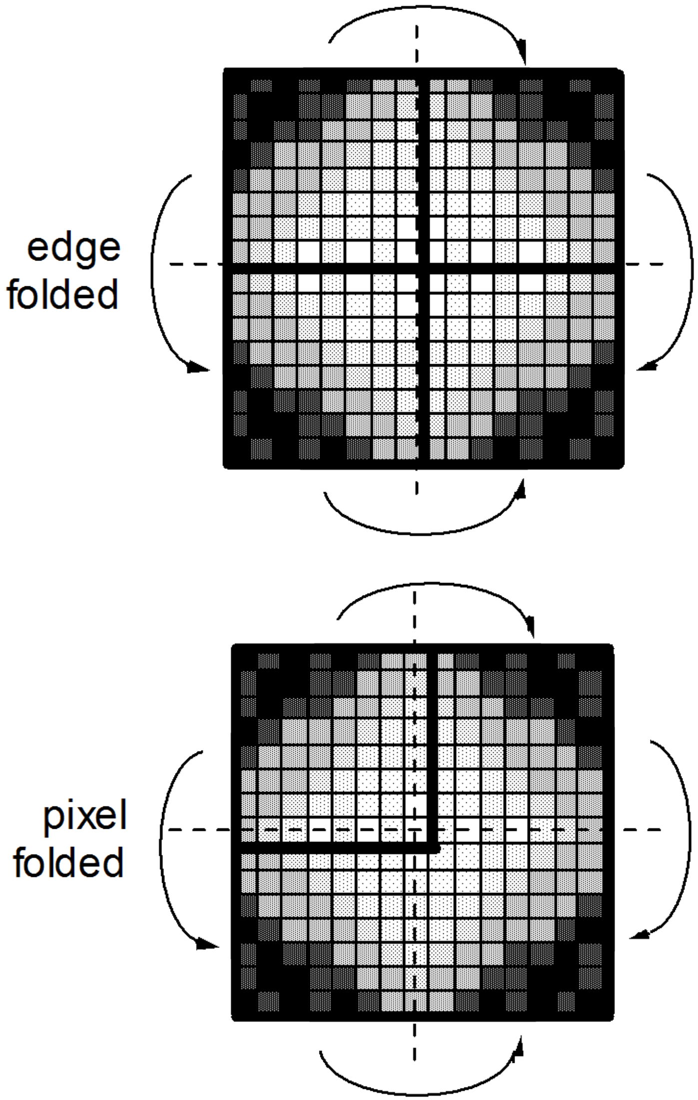


[Lohscheller]

- KLT is optimum
- DCT performs only slightly worse than KLT

Discrete Cosine Transform

- Transform coding of images using the Discrete Fourier Transform (DFT):
  - For stationary image statistics, the energy concentration properties of the DFT converge against those of the KLT for large block sizes.
  - **Problem** of blockwise DFT coding: blocking effects due to circular topology of the DFT and Gibbs phenomena.
  - **Remedy:** reflect image at block boundaries, DFT of larger symmetric block := DCT



Discrete Cosine Transform

- Type II DCT of blocksize NxN is defined by transform matrix A with elements

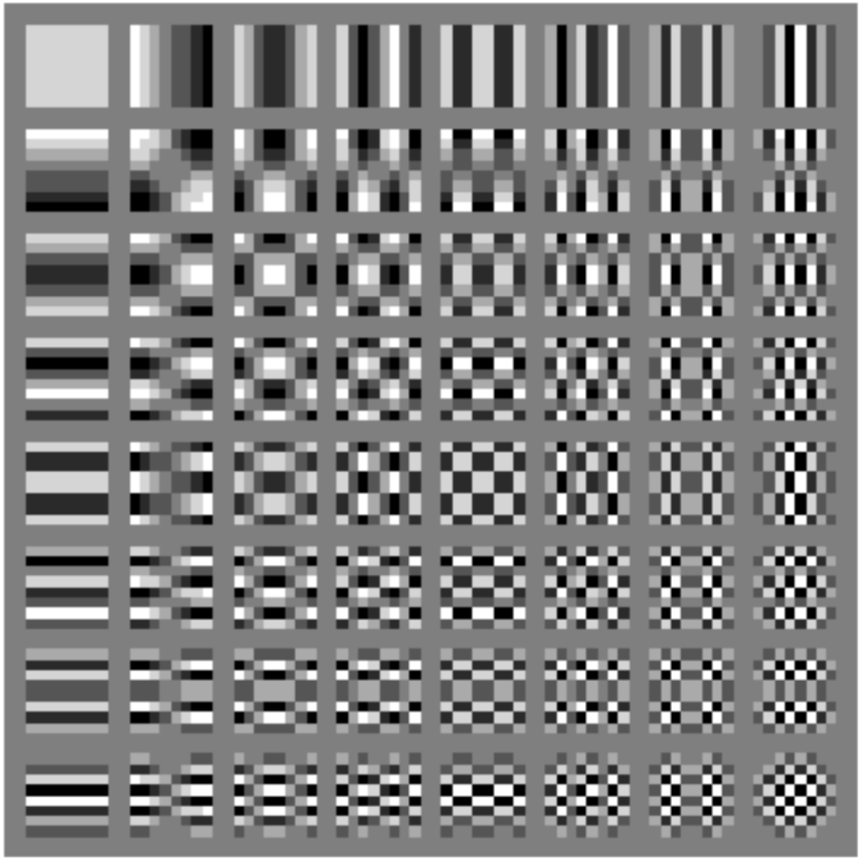
$$a_{ik} = \alpha_i \cos \frac{\pi(2k+1)i}{2N}$$

$$\text{for } \; i,k = 0,\ldots,N-1$$

$$\text{with} \quad \alpha_0 = \sqrt{\frac{1}{N}}$$

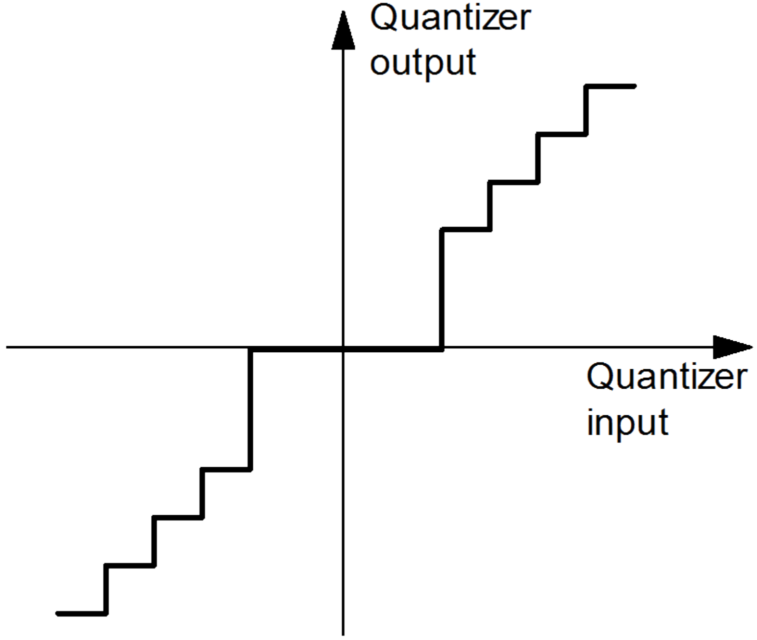
$$\alpha_i = \sqrt{\frac{2}{N}} \quad \forall i \neq 0$$

- 2D basis functions of the DCT:



Threshold Coding

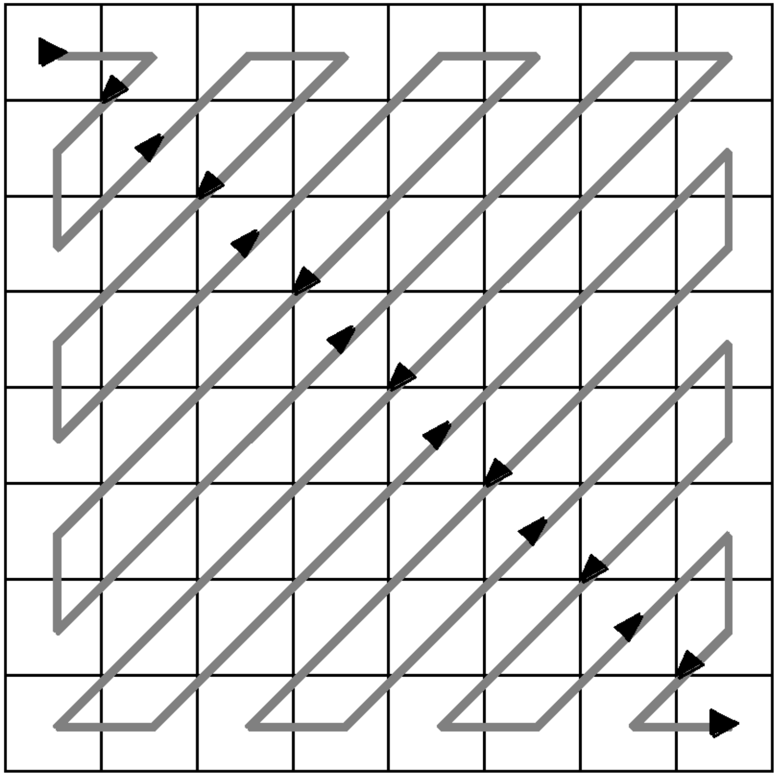
- Uniform deadzone quantizer: transform coefficients that fall below a threshold are discarded



- Positions of non-zero transform coefficients are transmitted in addition to their amplitude values

Threshold Coding

- Efficient encoding of the position of non-zero transform coefficients: zig-zag-scan + run-level-coding



ordering of the transform coefficients by zig-zag-scan

Threshold Coding

198	202	194	179	180	184	196	168
187	196	192	181	182	185	189	174
188	185	193	179	188	188	187	170
184	188	182	187	183	186	195	174
194	193	189	187	180	183	181	185
193	195	193	192	170	189	187	181
181	185	183	180	175	184	185	176
195	185	177	178	170	179	195	175

Original 8x8 block

DCT

1480	26.0	9.5	8.9	-26.4	15.1	-8.1	0.3
11.0	8.3	-8.2	3.8	-8.4	-6.0	-2.8	10.6
-5.5	4.5	9.0	5.3	-8.0	4.0	-5.1	4.9
10.7	9.8	4.9	-8.3	-2.1	-1.9	2.8	-8.1
1.6	1.4	8.2	4.3	3.4	4.1	-7.9	1.0
-4.5	-5.0	-6.4	4.1	-4.4	1.8	-3.2	2.1
5.9	5.8	2.4	2.8	-2.0	5.9	3.2	1.1
-3.0	2.5	-1.0	0.7	4.1	-6.1	6.0	5.7

Transformed 8x8 block

Q

185	3	1	1	-3	2	-1	0
1	1	-1	0	-1	0	0	1
0	0	1	0	1	0	0	0
1	1	0	-1	0	0	0	-1
0	0	1	0	0	0	-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

Zig-zag scan

Run-level coding

Mean of Block: 185

(0,3) (0,1) (1,1) (0,1) (0,1) (0,1) (0,-1) (1,1)
(1,1) (0,1) (1,-3) (0,2) (0,-1) (6,1) (0,-1) (0,-1)
(1,-1) (14,1) (9,-1) (0,-1) EOB

Transmission

Mean of Block: 185

(0,3) (0,1) (1,1) (0,1) (0,1) (0,1) (0,-1) (1,1)
(1,1) (0,1) (1,-3) (0,2) (0,-1) (6,1) (0,-1) (0,-1)
(1,-1) (14,1) (9,-1) (0,-1) EOB

Run-level decoding

185	3	1	1	-3	2	-1	0
1	1	-1	0	-1	0	0	1
0	0	1	0	1	0	0	0
1	1	0	-1	0	0	0	-1
0	0	1	0	0	0	-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

Inverse zig-zag scan

Scaling and inverse DCT

Reconstructed 8x8 block

192	201	195	184	177	184	193	174
189	191	195	182	182	187	190	171
188	185	190	181	185	187	189	171
189	188	185	183	183	182	190	175
191	192	186	189	179	182	188	178
190	191	189	190	177	186	184	179
189	188	185	184	175	186	187	179
189	188	178	176	173	183	193	180

Typical DCT Coding Artifacts

DCT coding with increasingly coarse quantization, block size 8x8

quantizer stepsize
for AC coefficients: 25

quantizer stepsize
for AC coefficients: 100

quantizer stepsize
for AC coefficients: 200

JPEG Standard

- JPEG: "Joint Photographic Experts Group"
- Formally: ISO/IEC JTC1/SC29/WG1

International Organization for Standardization

International Electrotechnical Commission

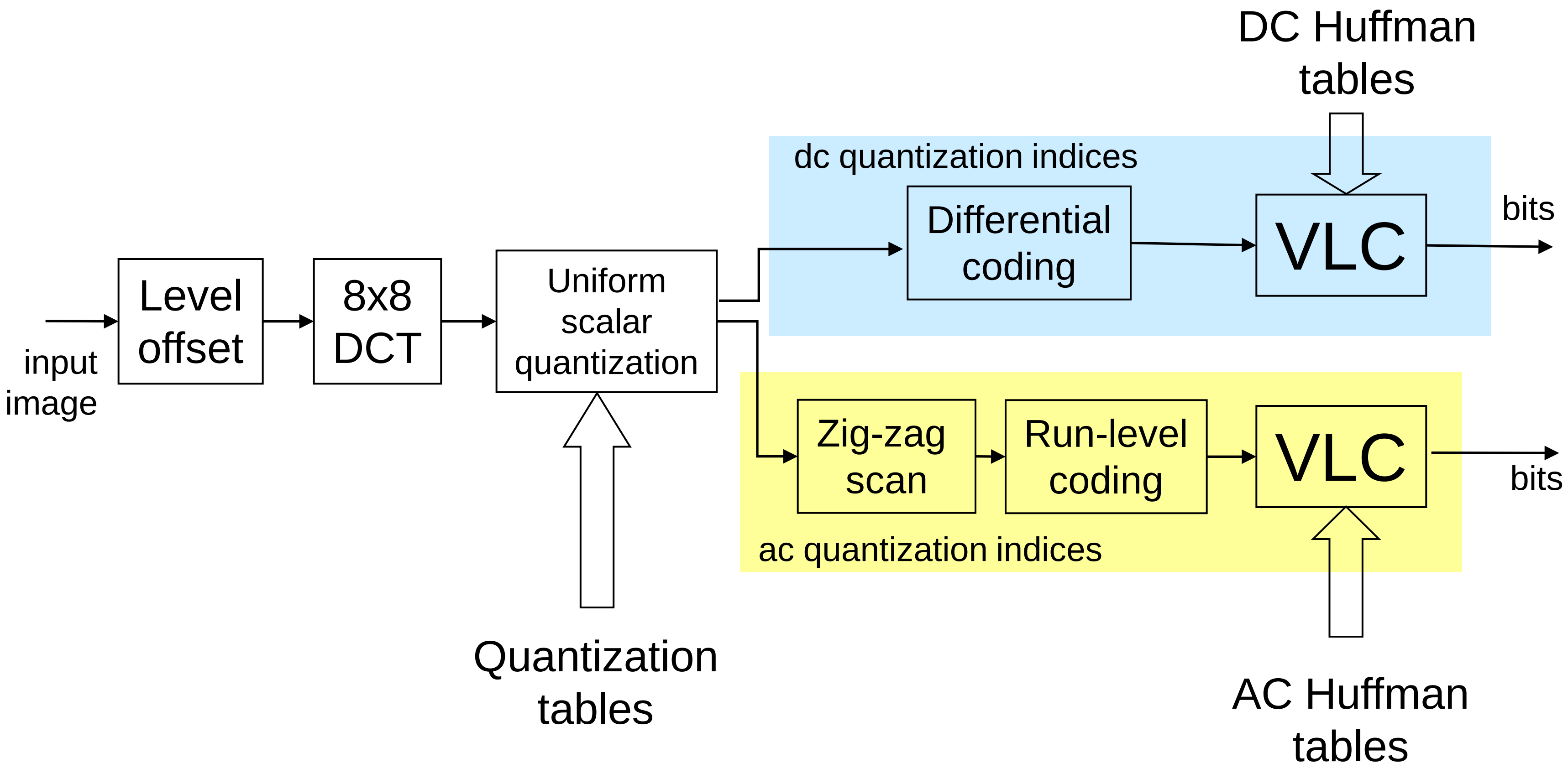
Joint ISO/IEC Technical Committee (Information Technology)

Subcommittee 29 (Coding of Audio, Picture, Multimedia and Hypermedia Information)

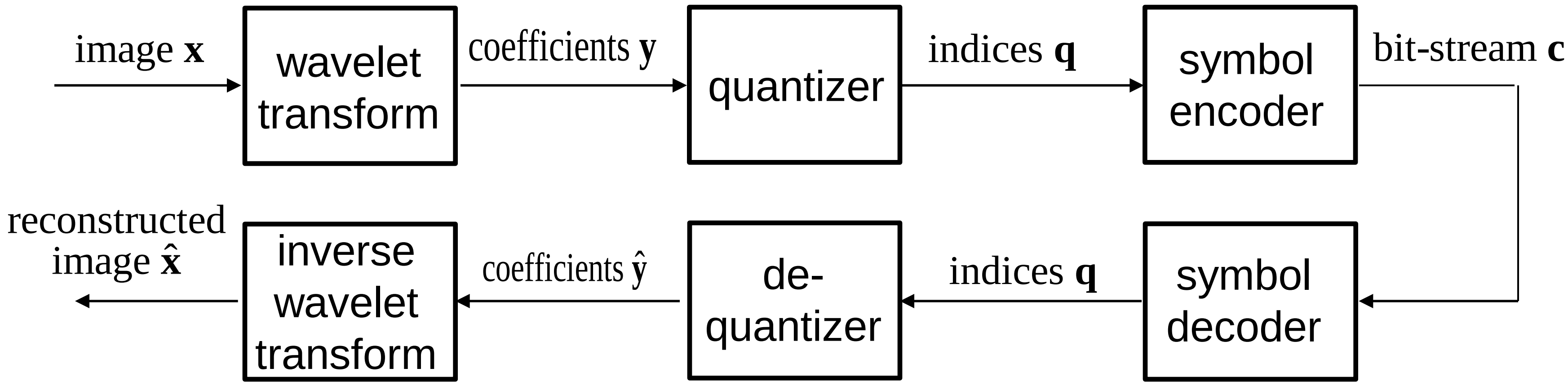
Working Group 1 (JBIG, JPEG)

- Work commenced in mid-1980's
- Draft international standard 1991
- Widely used for image exchange, WWW, and digital photography
- Motion-JPEG is de facto standard for digital video editing

Basic JPEG Algorithm

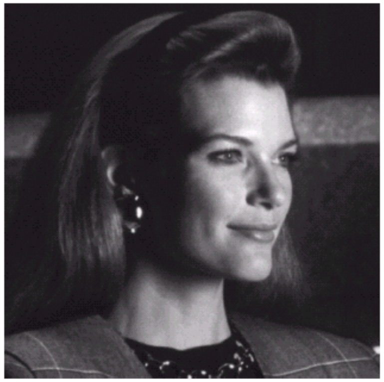


Subband Coding



- Usually biorthogonal wavelets
- Multiple levels of wavelet decomposition
- Sophisticated symbol encoder/decoder
- Usually generate embedded bit-streams

Wavelet Compression Results



Original
512x512
8bpp

0.074 bpp

0.048 bpp

Error
images

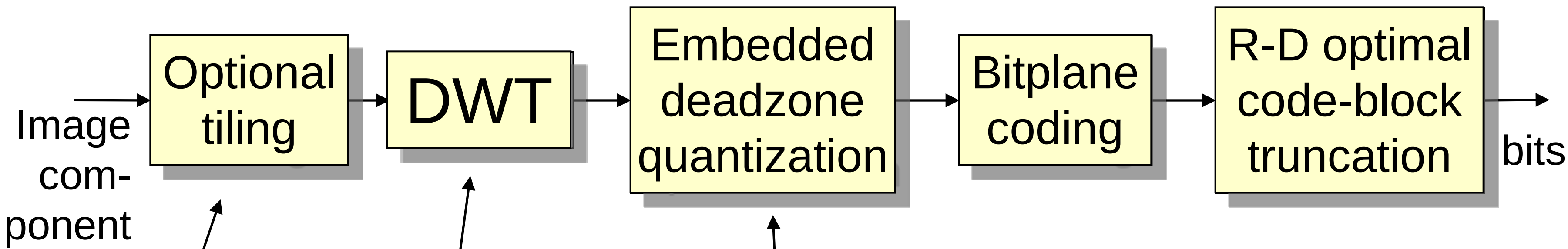
enlarged

[Gonzalez, Woods, 2001]

JPEG 2000

- Joint Photographic Experts Group ISO/IEC JTC1/SC29/WG1
- Still image compression standard
- Features
  - Improved compression efficiency (vs. JPEG)
  - Highly scalable embedded data streams
  - Progressive lossy to lossless compression within a single data stream
  - Arbitrarily crop images in the compressed domain
  - Selectively enhance quality of spatial “regions of interest”
  - Support for very large images
- JPEG 2000 Part I (minimum compliant decoder) international standard since December 2000.

JPEG 2000 Compression



Subdivision of images into rectangles for independent coding

Reversible 5/3 or floating-point 9/7 Daubechies wavelet.

Uniform scalar quantization, adjustable deadzone

Embedded block coding with optimal truncation (EBCOT)

Comparison JPEG vs. JPEG 2000

Lenna, 256x256 RGB
Baseline JPEG: 4572 bytes

Lenna, 256x256 RGB
JPEG 2000: 4572 bytes