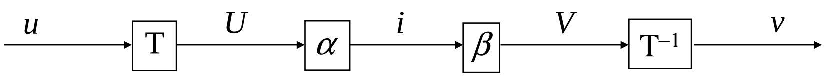
Transform Coding

- Principle of block-wise transform coding
- Properties of orthonormal transforms
- Discrete cosine transform (DCT)
- Bit allocation for transform coefficients
- Entropy coding of transform coefficients
- Typical coding artifacts
- Fast implementation of the DCT

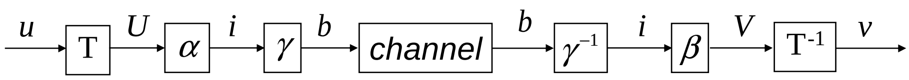
Transform Coding Principle

- Structure

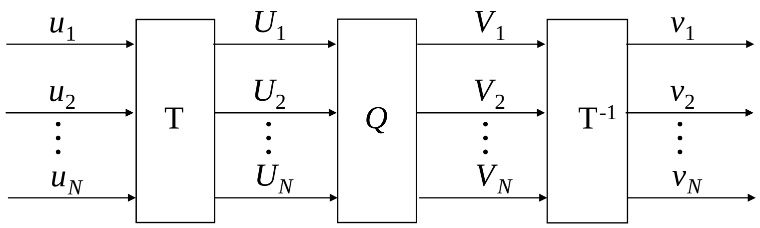
- Transform coder ($T\alpha$) / decoder (βT^{-1}) structure



- Insert entropy coding (γ) and transmission channel



Transform Coding and Quantization

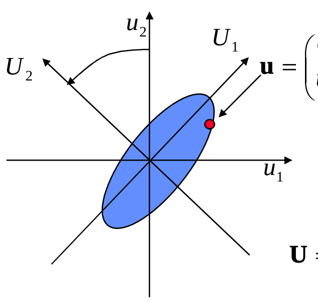


- Transformation of vector $\mathbf{u}=(u_1,u_2,\ldots,u_N)$ into $\mathbf{U}=(U_1,U_2,\ldots,U_N)$
- Quantizer Q may be applied to coefficients U_i
  - separately (scalar quantization: low complexity)
  - jointly (vector quantization, may require high complexity, exploiting of redundancy between U_i)
- Inverse transformation of vector $\mathbf{V}=(V_1,V_2,\ldots,V_N)$ into $\mathbf{v}=(v_1,v_2,\ldots,v_N)$

Why should we use a transform ?

Geometrical Interpretation

- A linear transform can decorrelate random variables
- An orthonormal transform is a rotation of the signal vector around the origin
- Parseval's Theorem holds for orthonormal transforms



$$\mathbf{u}=\begin{pmatrix}u_1\\u_2\end{pmatrix}=\begin{pmatrix}2\\1\end{pmatrix} \qquad \mathbf{T}=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\-1 & 1\end{pmatrix}$$

DC basis function

AC basis function

$$\mathbf{U}=\begin{pmatrix}U_1\\U_2\end{pmatrix}=\mathbf{Tu}=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\-1 & 1\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix}3\\-1\end{pmatrix}$$

Properties of Orthonormal Transforms

- Forward transform

$$\boxed{\mathbf{U} = \mathbf{T}\mathbf{u}}$$

N transform coefficients

input signal block of size N

Transform matrix
of size NxN

- Orthonormal transform property: inverse transform

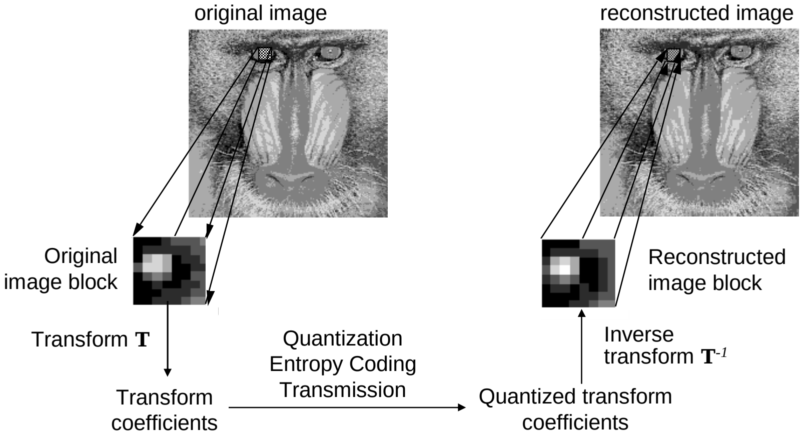
$$\boxed{\mathbf{u} = \mathbf{T}^{-1}\,\mathbf{U} = \mathbf{T}^{T}\,\mathbf{U}}$$

- Linearity: $\mathbf{u}$ is represented as linear combination of "basis functions"

from: Girod

Transform Coding of Images

Exploit horizontal and vertical dependencies by processing blocks



from: Girod

Separable Orthonormal Transforms, I

- Problem: size of vectors $N*N$ (typical value of N: 8)
- An orthonormal transform is separable, if the transform of a signal block of size $N*N$-can be expressed by

$$\boxed{\mathbf{U} = \mathbf{T}\mathbf{u}\mathbf{T}^{-1}}$$

$N*N$ transform coefficients

Orthonormal transform matrix of size $N* N$

$N*N$ block of input signal

- The inverse transform is

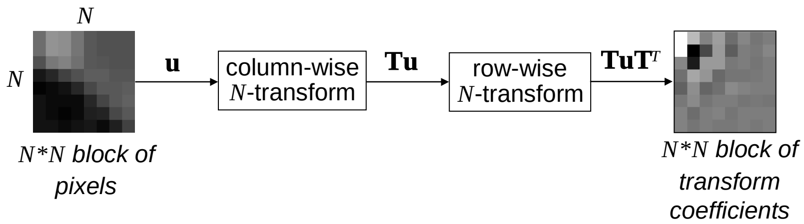
$$\boxed{\mathbf{u} = \mathbf{T}^{T}\,\mathbf{U}\,\mathbf{T}}$$

- Great practical importance: transform requires 2 matrix multiplications of size $N*N$ instead one multiplication of a vector of size $1*N^2$ with a matrix of size N^2*N^2
- Reduction of the complexity from $O(N^4)$ to $O(N^3)$

from: Girod

Separable Orthonormal Transforms, II

Separable 2-D transform is realized by two 1-D transforms
- along rows and
- columns of the signal block



from: Girod

Criteria for the Selection of a Particular Transform

- Decorrelation, energy concentration
  - KLT, DCT, …
  - Transform should provide energy compaction
- Visually pleasant basis functions
  - pseudo-random-noise, m-sequences, lapped transforms, …
  - Quantization errors make basis functions visible
- Low complexity of computation
  - Separability in 2-D
  - Simple quantization of transform coefficients

from: Girod

Karhunen Loève Transform (KLT)

- Decorrelate elements of vector $\mathbf{u}$

$$\mathbf{R_u} = \mathrm{E}\{\mathbf{uu}^T\}, \Rightarrow \quad \mathbf{U} = \mathbf{Tu}, \quad \Rightarrow$$

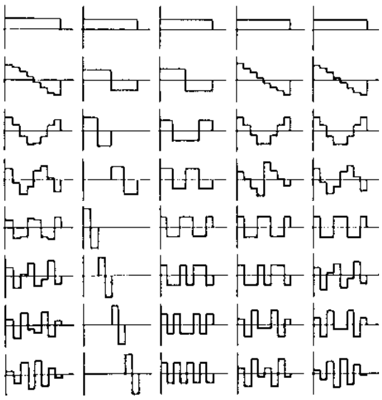
$$\mathbf{R_U} = \mathrm{E}\{\mathbf{UU}^T\} = \mathbf{T}\mathrm{E}\{\mathbf{uu}^T\}\mathbf{T}^T = \mathbf{TR_u T}^T = \mathrm{diag}\{\alpha_i\}$$

- Basis functions are eigenvectors of the covariance matrix of the input signal.
- KLT achieves optimum energy concentration.
- Disadvantages:
  - KLT dependent on signal statistics
  - KLT not separable for image blocks
  - Transform matrix cannot be factored into sparse matrices.

from: Girod

Comparison of Various Transforms, I

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



(1) (2) (3) (4) (5)

(1) Karhunen Loève transform (1948/1960)

(2) Haar transform (1910)

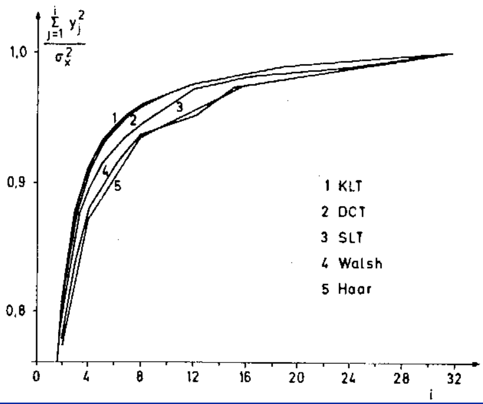
(3) Walsh-Hadamard transform (1923)

(4) Slant transform (Enomoto, Shibata, 1971)

(5) Discrete CosineTransform (DCT) (Ahmet, Natarajan, Rao, 1974)

Comparison of Various Transforms, II

- Energy concentration measured for typical natural images, block size 1x32 (Lohscheller)
- KLT is optimum
- DCT performs only slightly worse than KLT

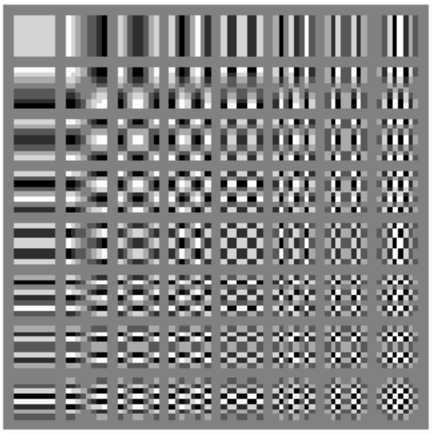


DCT

- Type II-DCT of blocksize $M \times M$ is defined by transform matrix A containing elements

$$
\begin{aligned}
&a_{ik} = \alpha_i \cdot \cos \frac{\pi(2k+1)i}{2M} \\
&i,k = 0 \ldots (M-1) \\
&\text{with} \\
&\quad \alpha_0 = \sqrt{\frac{1}{M}} \\
&\quad \alpha_i = \sqrt{\frac{2}{M}} \qquad i \neq 0
\end{aligned}
$$

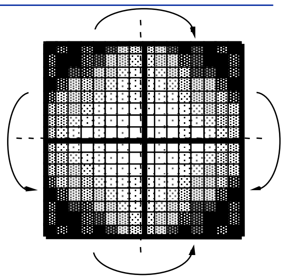
- 2D basis functions of the DCT:



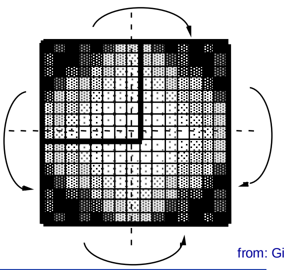
Discrete Cosine Transform and Discrete Fourier Transform

- 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
- DFT of larger symmetric block -> “DCT“ due to circular topology of the DFT and Gibbs phenomena.
- Remedy: reflect image at block boundaries

edge folded



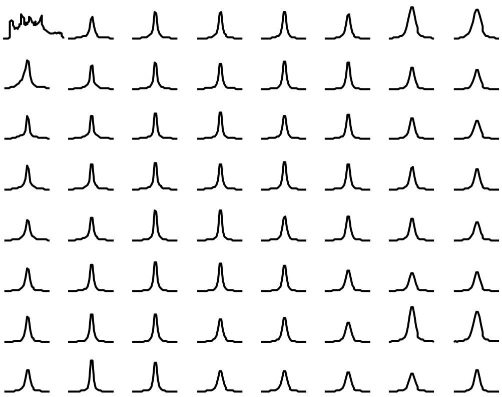
pixel folded



from: Girod

Histograms of DCT Coefficients:

- Image: Lena, 256x256 pixel
- DCT: 8x8 pixels
- DCT coefficients are approximately distributed like Laplacian pdf



Distribution of the DCT Coefficients

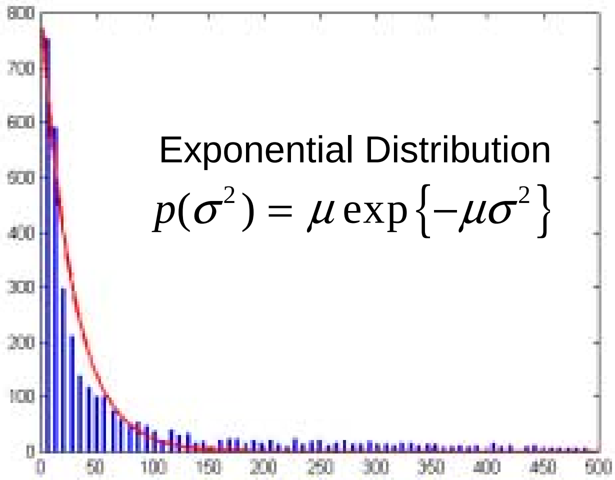
- Central Limit Theorem requests DCT coefficients to be Gaussian distributed
- Model variance of Gaussian DCT coefficients distribution as random variable (Lam & Goodmann‘2000)

$$p(u \mid \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{u^2}{2\sigma^2} \right\}$$

- Using conditional probability

$$p(u) = \int_0^\infty p(u \mid \sigma^2) \cdot p(\sigma^2) \cdot \mathrm{d}\sigma^2$$

Distribution of the Variance



Distribution of the DCT Coefficients

$$
\begin{aligned}
p(u) &= \int_0^\infty p(u \mid \sigma^2) \cdot p(\sigma^2) \cdot \mathrm{d}\sigma^2 \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{u^2}{2\sigma^2}\right\} \cdot \mu \exp\{-\mu\sigma^2\} \cdot \mathrm{d}\sigma^2 \\
&= \sqrt{\frac{2}{\pi}}\mu \int_0^\infty \exp\left\{-\frac{u^2}{2\sigma^2} - \mu\sigma^2\right\} \cdot \mathrm{d}\sigma \\
&= \left(\sqrt{\frac{2}{\pi}}\mu\right)\left(\frac{1}{2}\sqrt{\frac{\pi}{\mu}}\right) \exp\left\{-2\sqrt{\frac{\mu u}{2}}\right\} \\
&= \frac{\sqrt{2\mu}}{2} \exp\{-\sqrt{2\mu}\,|u|\} \qquad \text{Laplacian Distribution}
\end{aligned}
$$

Bit Allocation for Transform Coefficients I

- *Problem*: divide bit-rate R among N transform coefficients such that resulting distortion D is minimized.

$$\underbrace{D(R)}_{\text{Average distortion}} = \frac{1}{N}\sum_{i=1}^{N} \underbrace{D_i}_{\text{Distortion contributed by coefficient } i}(\underbrace{R_i}_{\text{Rate for coefficient } i}), \quad \text{s.t.} \quad \frac{1}{N}\sum_{i=1}^{N} R_i \le \underbrace{R}_{\text{Average rate}}$$

- *Approach*: minimize Lagrangian cost function

$$\frac{d}{dR_i}\sum_{i=1}^{N} D_i(R_i) + \lambda \sum_{i=1}^{N} R_i = \frac{dD_i(R_i)}{dR_i} + \lambda \overset{!}{=} 0$$

- *Solution*: Pareto condition $\boxed{\frac{dD_i(R_i)}{dR_i} = -\lambda}$

- Move bits from coefficient with small distortion reduction per bit to coefficient with larger distortion reduction per bit

Bit Allocation for Transform Coefficients II

- *Assumption*: high rate approximations are valid

$$D_i(R_i) \approx a\sigma_i^2 2^{-2R_i}, \quad \rightarrow \quad \frac{dD_i(R_i)}{dR_i} \approx -2a \ln 2\sigma_i^2 2^{-2R_i} = -\lambda$$

$$R_i \approx \log_2 \sigma_i + \log_2 \sqrt{\frac{2a\ln 2}{\lambda}} \rightarrow R_i \approx \log_2 \sigma_i - \log_2 \tilde{\sigma} + R = \log_2 \frac{\sigma_i}{\tilde{\sigma}} + R$$

$$R = \frac{1}{N}\sum_{i=1}^{N} R_i = \underbrace{\frac{1}{N}\sum_{i=1}^{N}\log_2 \sigma_i}_{\log_2 \tilde{\sigma}} + \log_2 \sqrt{\frac{2a\ln 2}{\lambda}} = \log_2 \underbrace{\left(\prod_{i=1}^{N}\sigma_i\right)^{\frac{1}{N}}}_{\tilde{\sigma}} + \log_2 \sqrt{\frac{2a\ln 2}{\lambda}}$$

- Operational Distortion Rate function for transform coding:

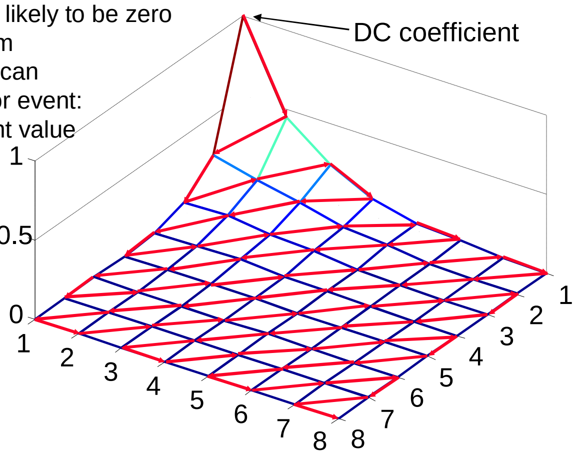
$$D(R) = \frac{1}{N}\sum_{i=1}^{N} D_i(R_i) \approx \frac{a}{N}\sum_{i=1}^{N}\sigma_i^2 2^{-2R_i} = \frac{a}{N}\sum_{i=1}^{N}\sigma_i^2 2^{-2\log_2\frac{\sigma_i}{\tilde{\sigma}} - 2R} = \boxed{a\tilde{\sigma}^2 2^{-2R}}$$

Geometric mean: $\tilde{\sigma} = \left(\prod_{i=1}^{N}\sigma_i\right)^{\frac{1}{N}}$

Entropy Coding of Transform Coefficients

- Previous derivation assumes: $R_i = \frac{1}{2}\max[(\log_2 \frac{\sigma_i^2}{D}), 0]\,bit$
- AC coefficients are very likely to be zero
- Ordering of the transform coefficients by zig-zag scan
- Design Huffman code for event: # of zeros and coefficient value
- Arithmetic code maybe simpler

Probability that coefficient is not zero when quantizing with: $V_i = 100 * \mathrm{round}(U_i / 100)$



Entropy Coding of Transform Coefficients II

Encoder

DCT $\rightarrow$ α $\rightarrow$

Original 8x8 block

201	195	188	193	169	157	196	190
193	188	187	201	195	193	213	193
184	192	180	195	182	151	199	193
176	172	179	179	152	148	198	183
196	195	169	171	159	185	218	175
214	213	205	170	173	185	206	150
207	205	207	184	180	167	173	160
198	203	205	186	196	149	159	163

DCT:

1480	49	33	-15	-14	33	-38	20
10	-52	11	-12	16	17	-13	-12
19	32	-22	-10	22	-20	9	8
16	10	17	27	-31	12	6	-5
-30	-6	13	-12	8	4	-3	-3
-25	16	6	-24	9	3	3	3
-2	17	4	-6	0	-4	-9	8
1	-2	6	0	7	-5	-8	-7

α:

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 $\rightarrow$ run-level coding $\rightarrow$ entropy coding & transmission

```
(185 3 1 0 1 1 1 -1 0 1 0 1 1 0 -3
 2 -1 0 0 0 0 0 0 1 -1 -1 0 -1 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0
 0 0 0 0 0 -1 -1  EOB)
```

```
Mean of block 185
(0,3) (0,1) (1,1) (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)
```

from: Girod

Entropy Coding of Transform Coefficients III

Decoder

transmission →

Mean of block: 185
(0,3) (0,1) (1,1) (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 0 1 1 1 -1 0 1 0 1 1 0 -3 2 -1 0 0 0 0 0 0 1 -1 -1 0 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 -1 -1 EOB)

→ Inverse zig-zag-scan →

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

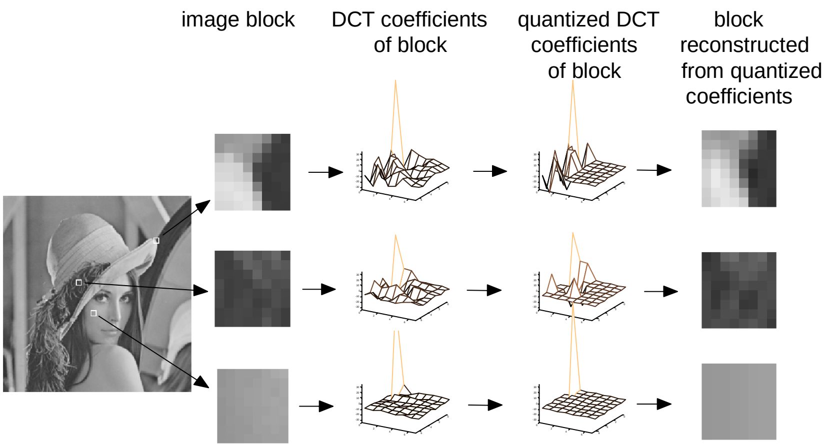
→ Scaling and inverse DCT: β →

196	193	187	192	179	176	196	189
198	188	182	198	196	192	208	200
185	189	191	197	174	159	184	189
167	181	182	177	154	153	187	189
201	199	178	165	163	185	206	179
220	217	193	176	165	179	197	170
194	198	195	193	169	156	180	179
210	196	192	209	185	149	157	160

Reconstructed 8x8 block

from: Girod

Detail in a Block vs. DCT Coefficients Transmitted



from: Girod

Typical DCT Coding Artifacts

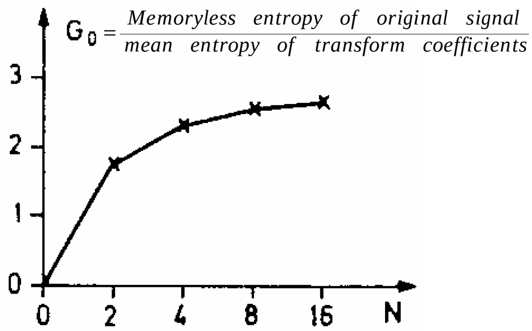
DCT coding with increasingly coarse quantization, block size 8x8

quantizer step size for AC coefficients: 25 | quantizer step size for AC coefficients: 100 | quantizer step size for AC coefficients: 200

from: Girod

Influence of DCT Block Size

- Efficiency as a function of block size NxN, measured for 8 bit Quantization in the original domain and equivalent quantization in the transform domain.



- Block size 8x8 is a good compromise between coding efficiency and complexity

from: Girod

Fast DCT Algorithm I

DCT matrix factored into sparse matrices (Arai, Agui, and Nakajima; 1988):

$$\boxed{\begin{aligned} \underline{y} &= \underline{M} \cdot \underline{x} \\ &= \underline{S} \cdot \underline{P} \cdot \underline{M}_1 \cdot \underline{M}_2 \cdot \underline{M}_3 \cdot \underline{M}_4 \cdot \underline{M}_5 \cdot \underline{M}_6 \cdot \underline{x} \end{aligned}}$$

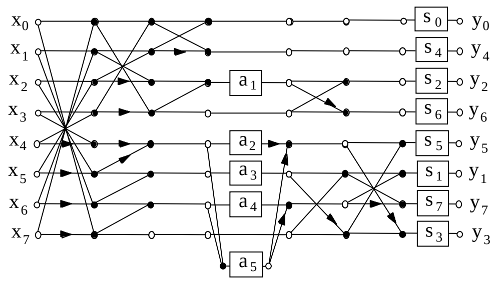
$$\underline{S} = \begin{bmatrix} s_0 & & & & & & & \\ & s_1 & & & & & & \\ & & s_2 & & & 0 & & \\ & & & s_3 & & & & \\ & & & & s_4 & & & \\ & & & & & s_5 & & \\ & 0 & & & & & s_6 & \\ & & & & & & & s_7 \end{bmatrix} \quad \underline{P} = \begin{bmatrix} 1 & & & & & & & \\ & & & & & 1 & & \\ & & 1 & & & & & \\ & & & & & & & 1 \\ & 1 & & & & & & \\ & & & & 1 & & & \\ & & & 1 & & & & \\ & & & & & & 1 & \end{bmatrix} \quad \underline{M}_1 = \begin{bmatrix} 1 & & & & & & & \\ & 1 & & 0 & & & & \\ & & 1 & & & & & \\ & & & 1 & & & & \\ & & & & 1 & & & 1 \\ & 0 & & & & 1 & 1 & \\ & & & & & 1 & -1 & \\ & & & & -1 & & & 1 \end{bmatrix} \quad \underline{M}_2 = \begin{bmatrix} 1 & & & & & & & \\ & 1 & & & 0 & & & \\ & & 1 & 1 & & & & \\ & & -1 & 1 & & & & \\ & & & & 1 & & & \\ & & & & & 1 & & 1 \\ & 0 & & & & & 1 & \\ & & & & & -1 & & 1 \end{bmatrix}$$

$$\underline{M}_3 = \begin{bmatrix} 1 & & & & & & & \\ & 1 & & & 0 & & & \\ & & C_4 & & & & & \\ & & & 1 & & & & \\ & & & & -C_2 & & -C_6 & \\ & 0 & & & & C_4 & & \\ & & & & -C_6 & & C_2 & \\ & & & & & & & 1 \end{bmatrix} \quad \underline{M}_4 = \begin{bmatrix} 1 & 1 & & & & & & \\ 1 & -1 & & & 0 & & & \\ & & 1 & 1 & & & & \\ & & & 1 & & & & \\ & & & & 1 & & & \\ & 0 & & & & 1 & & \\ & & & & & & 1 & \\ & & & & & & & 1 \end{bmatrix} \quad \underline{M}_5 = \begin{bmatrix} 1 & & & 1 & & & & \\ & 1 & 1 & & 0 & & & \\ & 1 & -1 & & & & & \\ 1 & & & -1 & & & & \\ & & & & -1 & -1 & & \\ & 0 & & & & 1 & 1 & \\ & & & & & & 1 & 1 \\ & & & & & & & 1 \end{bmatrix} \quad \underline{M}_6 = \begin{bmatrix} 1 & & 0 & & & & & 1 \\ & 1 & & & & & 1 & \\ & & 1 & & & 1 & & \\ 0 & & & 1 & 1 & 0 & & \\ & & & 1 & -1 & & & \\ & & 1 & & & -1 & & \\ & 1 & 0 & & & & -1 & \\ 1 & & & & & & & -1 \end{bmatrix}$$

from: Girod

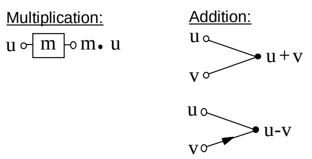
Fast DCT Algorithm II

Signal flow graph for fast (scaled) 8-DCT according to Arai, Agui, Nakajima:



scaling

only 5 + 8 Multiplications (direct matrix multiplication: 64 multiplications)



$a_1 = C_4$

$a_2 = C_2 - C_6$

$a_3 = C_4$

$a_4 = C_6 + C_2$

$a_5 = C_6$

$s_0 = \frac{1}{2 \cdot \sqrt{2}}$

$s_k = \frac{1}{4 \cdot C_k}; k = 1 \ldots 7$

$C_k = cos(ks/16)$

from: Girod

Transform Coding: Summary

- Orthonormal transform: rotation of coordinate system in signal space
- Purpose of transform: decorrelation, energy concentration
- KLT is optimum, but signal dependent and, hence, without a fast algorithm
- DCT shows reduced blocking artifacts compared to DFT
- Bit allocation proportional to logarithm of variance
- Threshold coding + zig-zag scan + 8x8 block size is widely used today (e.g. JPEG, MPEG-1/2/4, ITU-T H.261/2/3)
- Fast algorithm for scaled 8-DCT: 5 multiplications, 29 additions