Transform Coding

- Goal of Transform
  - to decorrelate the original signal
  - to compact a large fraction of the signal energy into a small set of transform components

The transform operation itself does not achieve any compression
The subsequent quantization of the transformed coefficients
and their efficient encoding result in compression

- The desired characteristic of the Transform
  - signal-independent basis function
    - Finding the basis functions of an optimal transform is computationally intensive task
    - The basis functions of a signal-dependent transform are required to transmit to the decoder
- Efficient bit allocation for quantizing the transform coefficients

The Transform

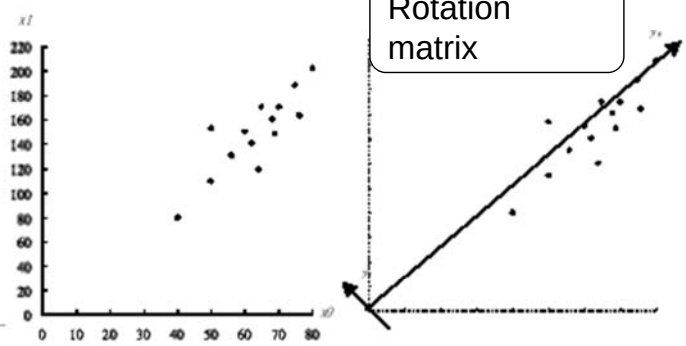
- Example
  - consider transform as coordinate axes rotations

X	(x_0,x_1)
X1	(65,170)
X2	(75,188)
X3	(60,150)
X4	(70,170)
X5	(56,130)
X6	(80,203)
X7	(68,160)
X8	(50,110)
X9	(40,80)
X10	(50,153)
X11	(69,148)
X12	(62,140)
X13	(76,164)
X14	(64,120)

$$y = Ax$$

$$\begin{bmatrix} y_0 \\ y_1 \end{bmatrix} = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix}$$

Rotation matrix



Y	(y_0,y_1)
Y1	(182,3)
Y2	(202,0)
Y3	(162,0)
Y4	(184,-2)
Y5	(141,-4)
Y6	(218,1)
Y7	(174,-4)
Y8	(121,-6)
Y9	(90,-7)
Y10	(161,10)
Y11	(163,-9)
Y12	(153,-6)
Y13	(181,-9)
Y14	(135,-15)

$$\sigma_{x_0}^2 = \frac{1}{14}\sum_{i=1}^{14}(xi_0 - \bar{x}_0)^2 = 117.31$$

$$\sigma_{x_1}^2 = 946.29$$

Q: $\sigma_{x_0}^2 + \sigma_{x_1}^2 = \sigma_{y_0}^2 + \sigma_{y_1}^2$

$$\sigma_{y_0}^2 = 1027.55$$

$$\sigma_{y_1}^2 = 36.05$$

Orthonormal Transform (1)

- Orthonormal transform
  - 1-D transform

$y = Ax$

$x = By \quad (B = A^{-1} = A^T)$

- e.g.

- $N = 2$

$$\begin{bmatrix} y_0 \\ y_1 \end{bmatrix} = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix}$$

Transform coefficients

$$y_n = \sum_{i=0}^{N-1} x_i a_{n,i}$$

$$\begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} a_{00} & a_{10} \\ a_{01} & a_{11} \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \end{bmatrix} = y_0 \begin{bmatrix} a_{00} \\ a_{01} \end{bmatrix} + y_1 \begin{bmatrix} a_{10} \\ a_{11} \end{bmatrix}$$

Basis vectors

- 2-D separable transform

$Y = AXA^T$

$X = BYB^T \quad (B = A^{-1} = A^T)$

- e.g.

$$\begin{bmatrix} x_{00} & x_{01} \\ x_{10} & x_{11} \end{bmatrix} = \begin{bmatrix} a_{00} & a_{10} \\ a_{01} & a_{11} \end{bmatrix} \begin{bmatrix} y_{00} & y_{01} \\ y_{10} & y_{11} \end{bmatrix} \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix}$$

$$= y_{00} \begin{bmatrix} a_{00}^2 & a_{00}a_{01} \\ a_{00}a_{01} & a_{01}^2 \end{bmatrix} + y_{01} \begin{bmatrix} a_{00}a_{10} & a_{00}a_{11} \\ a_{01}a_{10} & a_{01}a_{11} \end{bmatrix} + y_{10} \begin{bmatrix} a_{00}a_{10} & a_{01}a_{10} \\ a_{00}a_{11} & a_{01}a_{11} \end{bmatrix} + y_{11} \begin{bmatrix} a_{10}^2 & a_{10}a_{11} \\ a_{10}a_{11} & a_{11}^2 \end{bmatrix}$$

$$= y_{00} \begin{bmatrix} a_{00} \\ a_{01} \end{bmatrix} \begin{bmatrix} a_{00} & a_{01} \end{bmatrix} + y_{01} \begin{bmatrix} a_{00} \\ a_{01} \end{bmatrix} \begin{bmatrix} a_{10} & a_{11} \end{bmatrix} + y_{10} \begin{bmatrix} a_{10} \\ a_{11} \end{bmatrix} \begin{bmatrix} a_{00} & a_{01} \end{bmatrix} + y_{11} \begin{bmatrix} a_{10} \\ a_{11} \end{bmatrix} \begin{bmatrix} a_{10} & a_{11} \end{bmatrix}$$

Orthonormal Transform (2)

$y = Ax$

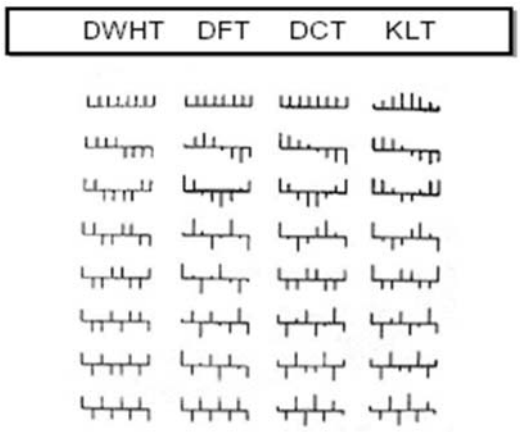
$$\begin{bmatrix} \\ y \\ \\ \end{bmatrix} = \begin{bmatrix} \leftarrow & b_0^T & \rightarrow \\ \leftarrow & b_1^T & \rightarrow \\ & & \\ \leftarrow & b_{N-1}^T & \rightarrow \end{bmatrix} \begin{bmatrix} \\ x \\ \\ \end{bmatrix}$$

$x = By = A^T y$

Basis vector $\boldsymbol{b}_i$

$$\begin{bmatrix} \\ x \\ \\ \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ b_0 & b_1 & & b_{N-1} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} \begin{bmatrix} \\ y \\ \\ \end{bmatrix}$$

- Basis vector
  - $N = 8$



Orthonormal Transform (3)

- Energy preserving $\quad \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \qquad \boldsymbol{x} = \boldsymbol{A}^T \boldsymbol{y}$

$$\sum_{i=0}^{N-1} y_i^2 = \boldsymbol{y}^T \boldsymbol{y} = (\boldsymbol{A}\boldsymbol{x})^T \boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^T \boldsymbol{A}^T \boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^T (\boldsymbol{A}^T \boldsymbol{A})\boldsymbol{x} = \boldsymbol{x}^T \boldsymbol{x} = \sum_{i=0}^{N-1} x_i^2$$

- Error introduced to x = Error introduced to y
  - ① forward transform $\quad \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$
  - ② quantization $\quad \hat{\boldsymbol{y}} = \boldsymbol{Q}[\boldsymbol{y}]$
  - ③ inverse transform $\quad \hat{\boldsymbol{x}} = \boldsymbol{A}^T \hat{\boldsymbol{y}}$

$$[A^T(y-\hat{y})]^T[A^T(y-\hat{y})]$$

$$\sum_{i=0}^{N-1}(x_i-\hat{x}_i)^2 = (x-\hat{x})^T(x-\hat{x}) = \|x-\hat{x}\|^2 = \left\|A^T y - A^T \hat{y}\right\|^2 = \|y-\hat{y}\|^2 = \sum_{i=0}^{N-1}(y_i-\hat{y}_i)^2$$

- The reconstruction error = error introduced by the set of quantizers

$$\sigma_r^2 = \frac{1}{N}\sum_{i=0}^{N-1}(x_i-\hat{x}_i)^2 = \frac{1}{N}\sum_{i=0}^{N-1}(y_i-\hat{y}_i)^2 = \frac{1}{N}\sum_{i=0}^{N-1}\sigma_{q_i}^2 = \sigma_q^2$$

Optimal Bit Allocation (1)

- Quantization error introduced by the *i*-th quantizer

Example: an uniform quantizer with 2^R level $\Rightarrow \Delta = \dfrac{2X_{\max}}{2^R}$

$$\sigma_q^2 = \frac{1}{\Delta}\int_{-\Delta/2}^{\Delta/2} q^2 dq = \frac{\Delta^2}{12} = \frac{1}{12}\left(\frac{2X_{\max}}{2^R}\right)^2 = \frac{1}{3}\frac{X_{\max}^2}{2^{2R}}$$

$$\sigma_x^2 = \int_{-X_{\max}}^{X_{\max}} x^2 \frac{1}{2X_{\max}} = \frac{X_{\max}^2}{3}$$

$$\therefore \sigma_q^2 = 2^{-2R}\sigma_x^2$$

$$\Rightarrow \quad \sigma_{q_i}^2 = \alpha_i 2^{-2R_i}\sigma_{y_i}^2 \quad \left(i.e.\ R_i = \frac{1}{2}\log_2\alpha_i + \frac{1}{2}\log_2\frac{\sigma_{y_i}^2}{\sigma_{q_i}^2}\right)$$

where R_i is # bits used by the i-th coefficient y_i and α_i is a factor that depends on the input distribution and quantizer

- Average reconstruction error

$$\sigma_r^2 = \sigma_q^2 = \frac{1}{N}\sum_{i=0}^{N-1}\sigma_{q_i}^2 = \frac{1}{N}\sum_{i=0}^{N-1}\alpha_i 2^{-2R_i}\sigma_{y_i}^2$$

Optimal Bit Allocation (2)

To minimize $\sigma_r^2 = \sigma_q^2 = \frac{1}{N}\sum_{i=0}^{N-1}\sigma_{q_i}^2 = \frac{1}{N}\sum_{i=0}^{N-1}\alpha_i 2^{-2R_i}\sigma_{y_i}^2$ (1)

with the constraint of $\frac{1}{N}\sum_{i=0}^{N-1}R_i = R$ (2)

Sol: Assume α_i is constant α for all i and using the Lagrange multiplier method

Let $J = \frac{\alpha}{N}\sum_{i=0}^{N-1}2^{-2R_i}\sigma_{y_i}^2 - \lambda(R - \frac{1}{N}\sum_{i=0}^{N-1}R_i)$ and set $\nabla J = 0$

(i.e. $(\frac{\partial J}{\partial R_0}, \frac{\partial J}{\partial R_1}, \ldots, \frac{\partial J}{\partial R_{N-1}}) = (0,0,\ldots 0)$

$\Rightarrow \quad R_i = \frac{1}{2}\log_2(2\alpha \ln 2\sigma_{y_i}^2) - \frac{1}{2}\log_2\lambda$ (3)

(3) 代入 (2)

$\Rightarrow \quad \frac{1}{N}\log_2(\prod_{i=0}^{N-1}(2\alpha \ln 2\sigma_{y_i}^2) = 2R + \log_2\lambda \quad \Rightarrow \quad \lambda = \prod_{i=0}^{N-1}(2\alpha \ln 2\sigma_{y_i}^2)^{1/N}2^{-2R}$ (4)

Optimal Bit Allocation (3)

(4) 代入 (3)

$\Rightarrow \quad R_i = R + \frac{1}{2}\log_2\frac{\sigma_{y_i}^2}{\prod_{i=0}^{N-1}(\sigma_{y_i}^2)^{1/N}}$ (5)

(5) 代入 (1)

$\Rightarrow \because \sigma_{q_i}^2 = \alpha 2^{-2R_i}\sigma_{y_i}^2 = \alpha 2^{-2R}(\prod_{i=0}^{N-1}\sigma_{y_i}^2)^{1/N}$

$\therefore \sigma_r^2 = \frac{1}{N}\sum_{i=0}^{N-1}\alpha 2^{-2R_i}\sigma_{y_i}^2 = \alpha 2^{-2R}(\prod_{i=0}^{N-1}\sigma_{y_i}^2)^{1/N}$

* The optimal bit allocation also implies identical error variances introduced in each y_i coefficient quantization

- Transform coding gain

$$G_{TC} = \frac{\sigma_{r,PCM}^2}{\sigma_{r,TC}^2} = \frac{\frac{1}{N}\sum_{i=0}^{N-1}\alpha 2^{-2R}\sigma_{y_i}^2}{\alpha 2^{-2R}\left(\prod_{i=0}^{N-1}\sigma_{y_i}^2\right)^{\frac{1}{N}}} = \frac{\frac{1}{N}\sum_{i=0}^{N-1}\sigma_{y_i}^2}{\left(\prod_{i=0}^{N-1}\sigma_{y_i}^2\right)^{\frac{1}{N}}}$$

, where $\sigma_{y_i}^2$ is the variance of the ith coefficient y_i.

$$10\log_{10}G_{TC} = 10\log_{10}\frac{\sigma_x^2/\sigma_{r,TC}^2}{\sigma_x^2/\sigma_{r,PCM}^2} = 10\log_{10}\frac{\sigma_x^2}{\sigma_{r,TC}^2} - 10\log_{10}\frac{\sigma_x^2}{\sigma_{r,PCM}^2} \quad \Rightarrow \quad \boxed{SNR_{TC}(db) = SNR_{PCM}(db) + 10\log G_{TC}}$$

Optimal Transform (1)

- Karhunen-Loeve Transform(KLT)
  - the optimal transform in an energy-packing sense

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$$

$$\boldsymbol{R}_{yy} = \boldsymbol{E}\left[\boldsymbol{y}\boldsymbol{y}^T\right] = \boldsymbol{E}\left[(\boldsymbol{A}\boldsymbol{x})(\boldsymbol{A}\boldsymbol{x})^T\right] = \boldsymbol{A}\cdot\boldsymbol{E}\left[\boldsymbol{x}\boldsymbol{x}^T\right]\cdot\boldsymbol{A}^T = \boldsymbol{A}\cdot\boldsymbol{R}_{xx}\cdot\boldsymbol{A}^T$$

Correlation matrix of the transformed coefficient

Correlation matrix of the input signal

$\because \boldsymbol{R}_{xx}$ is a real symmetric matrix

$\therefore$ its eigenvalues λ_i are real and there exactly N eigenvectors $\boldsymbol{l}_i$ which are orthogonal and can be chosen to be orthormal

$$\boldsymbol{R}_{xx}\cdot\boldsymbol{l}_i = \lambda_i\cdot\boldsymbol{l}_i, \qquad i = 0,\ldots,N-1$$

  - KLT is defined by the fact that its basis vectors are the eigenvectors of $\boldsymbol{R}_{xx}$

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{l}_0^T \\ \vdots \\ \boldsymbol{l}_{N-1}^T \end{bmatrix}$$

$$\boldsymbol{A}^T = \begin{bmatrix} \boldsymbol{l}_0 & \boldsymbol{l}_1 & \cdots & \boldsymbol{l}_{N-1} \end{bmatrix}$$

Optimal Transform (2)

- KLT diagonalize $\boldsymbol{R}_{xx}$

$$\boldsymbol{R}_{yy} = \boldsymbol{A}\cdot\boldsymbol{R}_{xx}\cdot\boldsymbol{A}^T = \begin{bmatrix} \boldsymbol{l}_0^T \\ \vdots \\ \boldsymbol{l}_{N-1}^T \end{bmatrix}\begin{bmatrix} \boldsymbol{R}_{xx}\boldsymbol{l}_0 & \cdots & \boldsymbol{R}_{xx}\boldsymbol{l}_{N-1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{l}_0^T \\ \vdots \\ \boldsymbol{l}_{N-1}^T \end{bmatrix}\begin{bmatrix} \lambda_0\boldsymbol{l}_0 & \cdots & \lambda_{N-1}\boldsymbol{l}_{N-1} \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_0 & 0 & \cdots & 0 \\ 0 & \lambda_1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_{N-1} \end{bmatrix} = \Lambda$$

  - the transformed coefficients are uncorrelated and can quantized independently
  - e.g
    - if N = 2 and $\boldsymbol{R}_{xx} = \begin{bmatrix} R_{xx}(0) & R_{xx}(1) \\ R_{xx}(1) & R_{xx}(0) \end{bmatrix}$
    - then

$$\lambda_0 = R_{xx}(0) + R_{xx}(1) \Rightarrow l_0 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \lambda_1 = R_{xx}(0) - R_{xx}(1) \Rightarrow l_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\therefore \boldsymbol{A} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Suboptimum Transforms (1)

- Discrete Cosine Transform (DCT)

$$C(k,l)=\begin{cases}\sqrt{\frac{1}{N}}\cos\frac{(2l+1)k\pi}{2N}, & k=0\\ \sqrt{\frac{2}{N}}\cos\frac{(2l+1)k\pi}{2N}, & k=1,\ldots,N-1\end{cases}\quad ,l=0,1,\ldots,N-1$$

$$\begin{cases}F_k=\sqrt{\frac{2}{N}}\,c_k\sum_{l=0}^{N-1}f_l\cos\left[k\pi\left((2l+1)/2N\right)\right],\ 0\le k\le N\text{-}1\\ f_l=\sqrt{\frac{2}{N}}\sum_{k=0}^{N-1}c_kF_k\cos\left[k\pi((2l+1)/2N)\right],\ 0\le l\le N\text{-}1\\ \text{where } c_k=\frac{1}{\sqrt{2}} \text{ if } k\text{=0, and } c_k\text{=1, otherwise.}\end{cases}$$

i.e. $F=Cf$ & $f=C^TF$

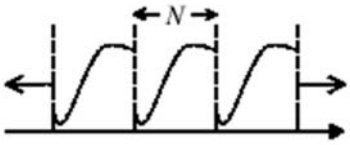
– for signals with high correlation $\boldsymbol{\rho}\rightarrow 1$ $\quad(\rho=\frac{E[x_nx_{n+1}]}{E[x_n^2]})$

- the energy compaction ability of DCT is very close to that of KLT

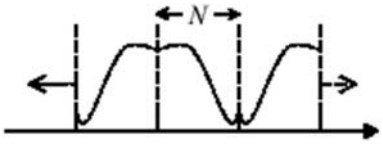
– DCT has a higher compression efficiency than DFT does

- DCT avoids the generation of spurious spectral components

Basis matrices (8x8) for DCT



DFT periodicity



DCT periodicity

Suboptimum Transforms (2)

- Discrete Fourier Transform (DFT)

$$F(k,l)=\frac{1}{\sqrt{N}}\exp(-\frac{2\pi ikl}{N})=\frac{1}{\sqrt{N}}[\cos(\frac{2\pi kl}{N})-i\sin(\frac{2\pi kl}{N})],\quad k,l=0,1,\ldots,N-1$$

- Discrete Hartley Transform (DHT)

$$F(k,l)=\frac{1}{\sqrt{N}}\mathrm{cas}(\frac{2\pi kl}{N})=\frac{1}{\sqrt{N}}[\cos(\frac{2\pi kl}{N})+\sin(\frac{2\pi kl}{N})],\quad k,l=0,1,\ldots,N-1$$

- Discrete Sine Transform (DST)

$$S(k,l)=\sqrt{\frac{2}{N+1}}\sin(\frac{(k+1)(l+1)\pi}{N+1}),\quad k,l=0,1,\ldots,N-1$$

– DST performs close to the optimum KLT when correlation ρis small

Suboptimum Transforms (3)

- Discrete Hadamard Transform
  - Hadamard matrix of order N : an N×N matrix H_N , $H_N H_N^T = N I_N$

$$H_{2N} = \begin{bmatrix} H_N & H_N \\ H_N & -H_N \end{bmatrix}$$

$$H_1 = [1] \quad H_2 = \begin{bmatrix} H_1 & H_1 \\ H_1 & -H_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$H_4 = \begin{bmatrix} H_2 & H_2 \\ H_2 & -H_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

0
3/2
1/2
2/2

$$H_8 = \begin{bmatrix} H_4 & H_4 \\ H_4 & -H_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix}$$

Sequency
0
7/2
3/2
4/2
1/2
6/2
2/2
5/2

- Discrete Walsh-Hadamard Tranform
  - normalization $\Rightarrow H_N H_N^T = I_N$
  - row re-ordering in increasing sequency order
    - sequency = (number of sign changes) / 2

$$H_4 = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

$$H_8 = \frac{1}{\sqrt{8}}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix}$$

Transform Coding Strategies (1)

- Zonal sampling
  - retain only those transform coefficients located in a predefined zone in the transform block
  - each retained coefficient is then quantized and encoded with a fixed-length codeword

or

  - allocate bits to the individual coefficients within the predefined zone so as to minimize the total quantization error
  - e.g.
    Bit allocation map for an 8×8 transform
  - Drawback
    - the coefficients outside the zone may contain significant energy
      $\Rightarrow$ large reconstruction errors

Zonal mask

8	7	5	3	1	1	0	0
7	5	3	2	1	0	0	0
4	3	2	1	1	0	0	0
3	3	2	1	1	0	0	0
2	1	1	1	0	0	0	0
1	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Transform Coding Strategies (2)

- Threshold coding
  - coefficients with magnitude above the threshold are quantized and encoded
  - e.g.
    - quantization table for an 8×8 block

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

  - Ordering for threshold coding
    - Zigzag scan ordering

0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63