Run length coding

- Represent a sequence of values X_n by values and #s of repetitions

- (Value), $\underbrace{(\text{\# of repetitions})}_{\substack{\geq 1 \\ \leq 2^b}}$

Example: $X_n \in \{0, 1, 2\}$

$$\underbrace{0\ 0\ 0\ 0\ 0}_{0\,5}\ \underbrace{1\ 1}_{1\,2}\ \underbrace{0\ 0\ 0\ 0\ 0}_{0\,5}\ \underbrace{2}_{2\,1}\ \underbrace{0\ 0\ 0}_{0\,3}$$

Example:
Let $b = 3$

$$\underbrace{0\ 0\ 0\ 0\ 0\ 0\ 0\ 0}_{0\,8}\ \underbrace{0\ 0}_{0\,2}$$

$\#$ of code values $= M \cdot 2^b$
$= (\#\text{ of symbols})(\text{max run length})$

Alternative method

$$x_n \in \{0, 1\}$$

- Start with 0
- Run lengths of length $0 \to (2^b - 1)$

(run length) (run length)

Example $\quad b = 3$

$$1\ \underbrace{0\ 0\ 0\ 0\ 0\ 0\ 0}_{7}\ \underbrace{0\ 0\ 0}_{3}\ \underbrace{1\ 1\ 1}_{3}$$

$$0\ 1\ \ 7\ \ 0\ \ 3\ \ 3$$

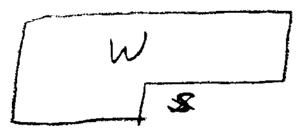
Comments

1) Run length coding reduces dependence

2) Requires subsequent Huffman coding for further compression

Predictive Coding (Lossless)

- Let $X_s \in \{0, 1\}$ binary, $s \in \mathbb{Z}^2$
- Let W be a causal window



- Compute

$$p(x_s \mid x_{s+r}\ r \in W)$$

- Predict x_s as

$$\hat{x}_s = \begin{cases} 1 & \text{if } p(1 \mid x_r\ r \in W) > p(0 \mid x_r\ r \in W) \\ 0 & \text{otherwise} \end{cases}$$

$\hat{X}_s$ is the minimum probability of error predictor

- If $X_s = \hat{X}_s \Rightarrow$ send 0
  $X_s \neq \hat{X}_s \Rightarrow$ send 1

- Run length code and Huffman code result.

How to pick $p(x_s \mid X_r \; r \in W)$?

Strategy 1) Typical predictor

a) Select of set of "typical" images

b) at each pixel of each image, form the vector $Z_s = (X_{s+r_1}, \ldots, X_{s+r_p})$ where $\{r_1, \ldots, r_p\} = W$

c) Compute histogram $h(i,j)$

$$h(i,j) = (\# \text{ of } X_s = i \text{ and } Z_s = j)$$

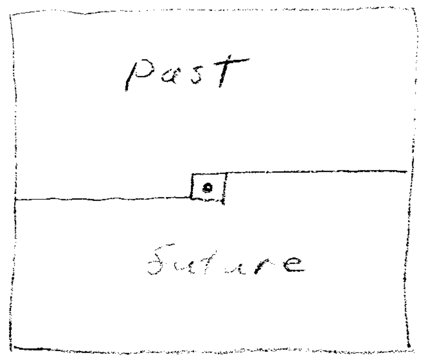
j is vector valued

$$j = (j_1, \ldots, j_p) \in 2^p$$

d) Estimate $p(x_s \mid z_s)$ as

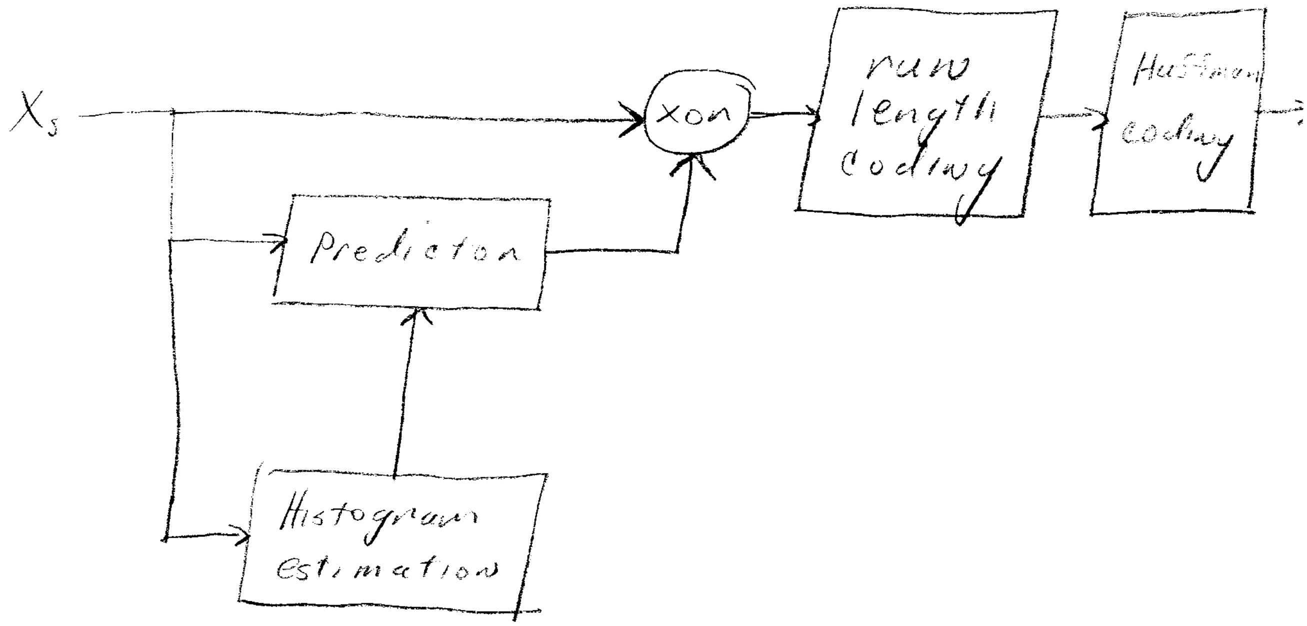
$$\hat{p}(x_s \mid z_s) = \frac{h(i,j)}{\sum_{i=0}^{1} h(i,j)}$$

Strategy 2) Adaptive Predictor

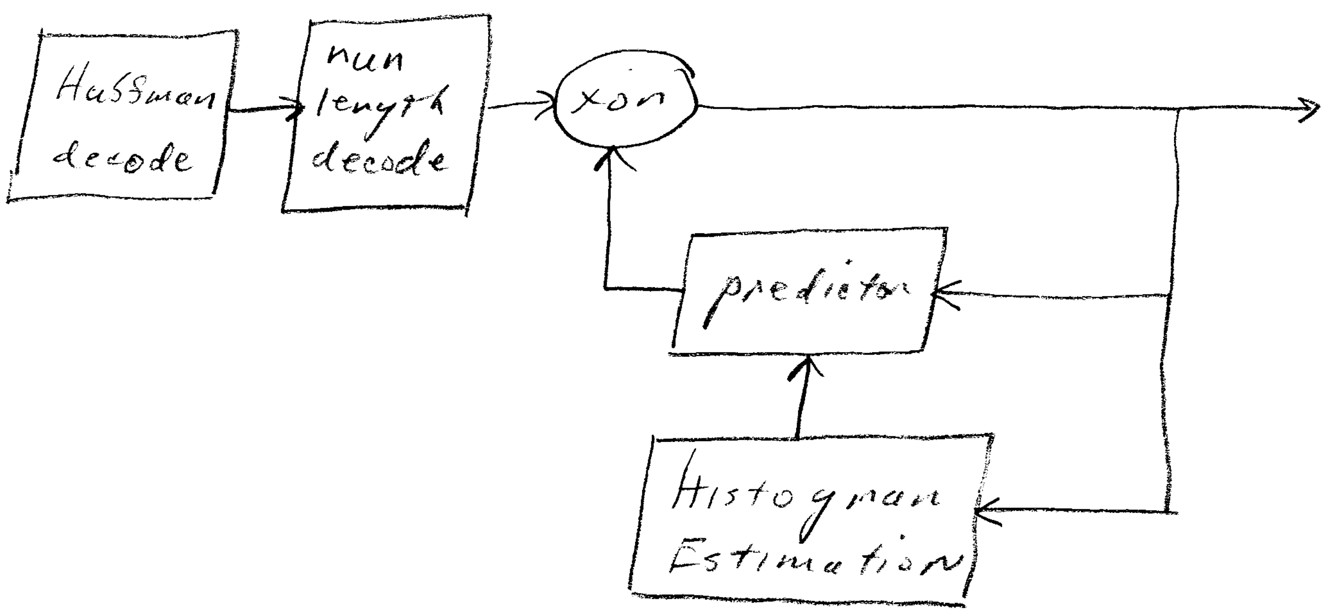


- Use all past points to estimate $\hat{p}(x_s|z_s)$
- These points are also available at decoder

Lossless Predictive Encoder



Decoder



Lossy Coding

Rate-distortion theory
- For each level of distortion there is a optimal bit rate

Define

$X_n \in \{0, \ldots, M-1\} \leftarrow$ source symbols

$X \leftarrow$ input signal $X_0, \ldots, X_{N-1}$

$N \leftarrow$ number of symbols

$Y \leftarrow$ Coded image; binary sequence of length K

$Z \leftarrow$ decoded signal $Z_0, \ldots, Z_{N-1}$

$$\frac{K}{N \log_2 M} = \text{compression ratio}$$

$$\frac{K}{N} = \text{bit rate} \left(\frac{\text{bits}}{\text{pixel}} \right)$$

$d(x, z)$ = distortion measure

$= \|x - z\|^2$

$$I(x, z) \triangleq \int -p(x) \log_2 p(x)\, dx - \int -p(x, y) \log_2 p(x|y)\, dx\, dy$$

= mutual information

<u>Distortion</u>

$$D = \frac{1}{N} E\left[d(x, z)\right] = \frac{1}{N} E\left[\|x - z\|^2\right]$$

<u>Rate</u>

$$R(\delta) = \lim_{N \to \infty} \frac{1}{N} \inf_{\text{all } z\text{'s}} \left\{ I(x, z) : D \le \delta \right\}$$

Comment:

Usually $D(\delta)$ and $R(\delta)$ are <u>both</u> functions of a free parameter δ.

Theorems (loosely stated)

1) $R(\delta)$ is a monotone decreasing function of δ

2) There exists some δ such that $R(\delta) = 0$

3) $R(\delta)$ is a convex function

4) Shannon's source coding theorem: For any $R' > R(\delta)$ $D' > D(\delta)$, and N sufficiently large, there exists a code which achieves

$$\frac{K}{N} < R' \qquad E\left[d(X, Z)\right] < D'$$

- In other words, we can achieve a bit rate arbitrarily close to $R(\delta)$ at distortion $D(\delta)$

- Proof works by selecting codewords, Z, randomly with distribution of X

Examples 1)

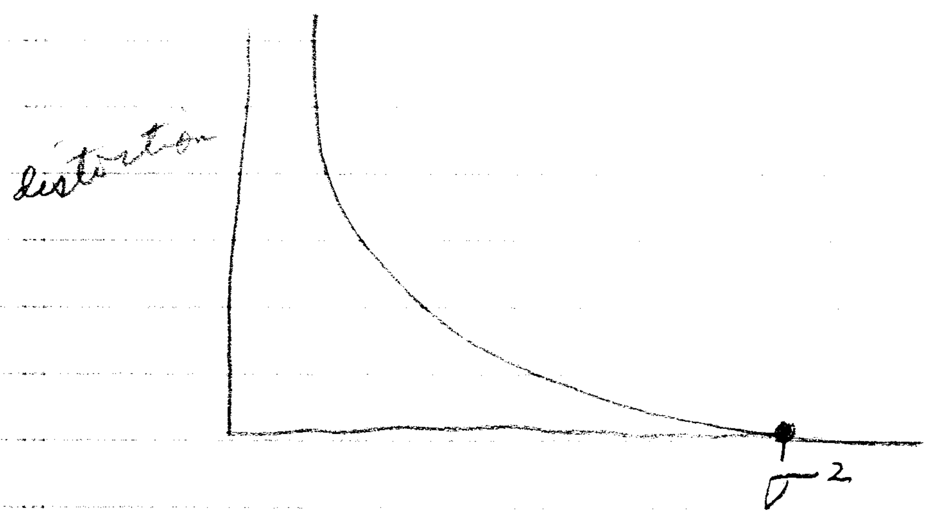
$$X \sim N(0, \Sigma) \quad RV$$

$$d(x, z) = |x - z|^2$$

Can't be shown that

$$R(\delta) = \begin{cases} \frac{1}{2} \log \frac{\sigma^2}{\delta} & \text{for } \delta \leq \sigma^2 \\ 0 & \text{for } \delta > \sigma^2 \end{cases}$$

$$D(\delta) = \min(\sigma^2, \delta)$$



How

Example 2)

$$X \in \mathbb{R}^p$$
$$X \sim N(0, B)$$

$$B = E^t \Lambda E$$

where E is an orthonormal matrix of eigenvectors and Λ is a diagonal matrix of eigenvalues

Then let

$$\tilde{X} = E X$$

$$\tilde{B} = \mathbb{E}[\tilde{X} \tilde{X}^t]$$
$$= \mathbb{E}[E X X^t E^t]$$
$$= E B E^t = E E^t \Lambda E^t E$$
$$= \Lambda = \mathrm{diag}\{\sigma_1, \ldots, \sigma_p\}$$

$\Rightarrow \tilde{X}$ are indepent

Since E is orthonormal

$$E[\|X - Z\|^2] = E[\|\tilde{X} - \tilde{Z}\|^2]$$

$\Rightarrow$ Rate distortion of X and $\tilde{X}$ are the same

Can be show that

$$R(\delta) = \sum_{i=1}^{p} \max\left(0, \frac{1}{2} \log_2 \frac{\sigma_i^2}{\delta}\right)$$

$$D(\delta) = \sum_{i=1}^{p} \min\left(\sigma_i^2, \delta\right)$$

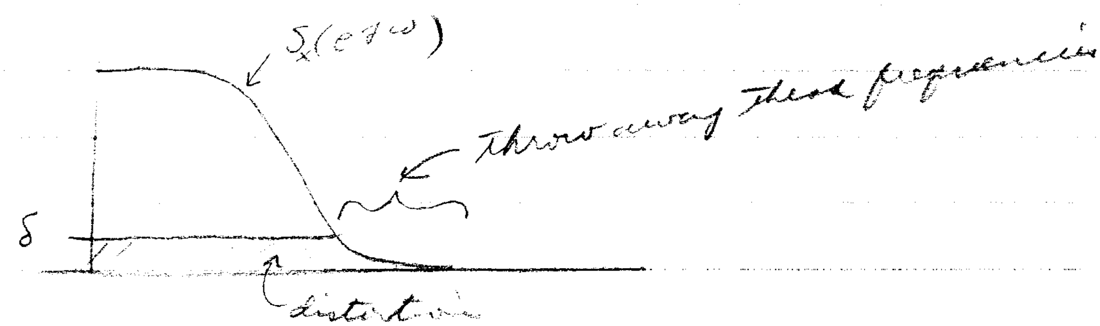
$\Rightarrow$ Throw away smallest eigenvalues

Example 3)

MSE Coding of DT Gaussian random process with power spectrum $S_x(e^{j\omega})$

$$R(\delta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \max\left(0,\ \frac{1}{2}\log_2 \frac{S_x(e^{j\omega})}{\delta}\right) d\omega$$

$$D(\delta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \min\left(S_x(e^{j\omega}),\ \delta\right) d\omega$$



General Coder Structure

$X \rightarrow$ [decorrelating transformation] $\rightarrow$ [quantization "bit allocation"] $\rightarrow$ [entropy coding] $\rightarrow$

Typical Transformations

Block Transforms
1) KLT
2) DCT (JPEG)

Other Transforms
1) QMF (subband)
2) Wavelet

Typical Quantizers

1) Uniform scalar
2) Nonuniform scalar
3) Vector quantizer
4) More bits for low spatial frequencies $\Rightarrow$ not MSE distortion

Block coding using Karhunen-Loeve (KL) transform

1) Break image into $N \times N$ blocks

2) Form vectors $X_n \in \mathbb{R}^{N^2}$ for each block.

3) Compute KL transform

$$C = E[X_n X_n^t]$$

$$C = E^t \Lambda E$$

E - orthonormal

$$\Lambda = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_{N^2}^2)$$

where $\sigma_1^2 > \sigma_2^2 \ldots > \sigma_{N^2}^2$

4) $$\tilde{X}_n = E X_n$$

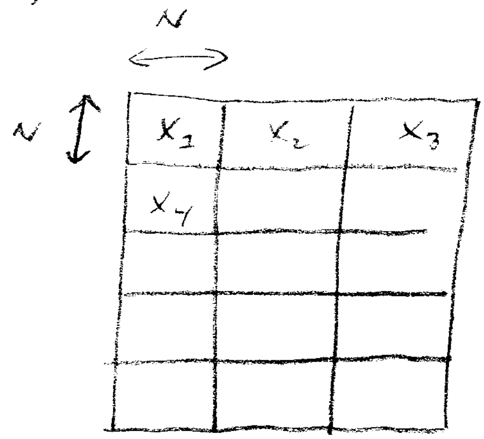
5) Apply (uniform) scalar quantization

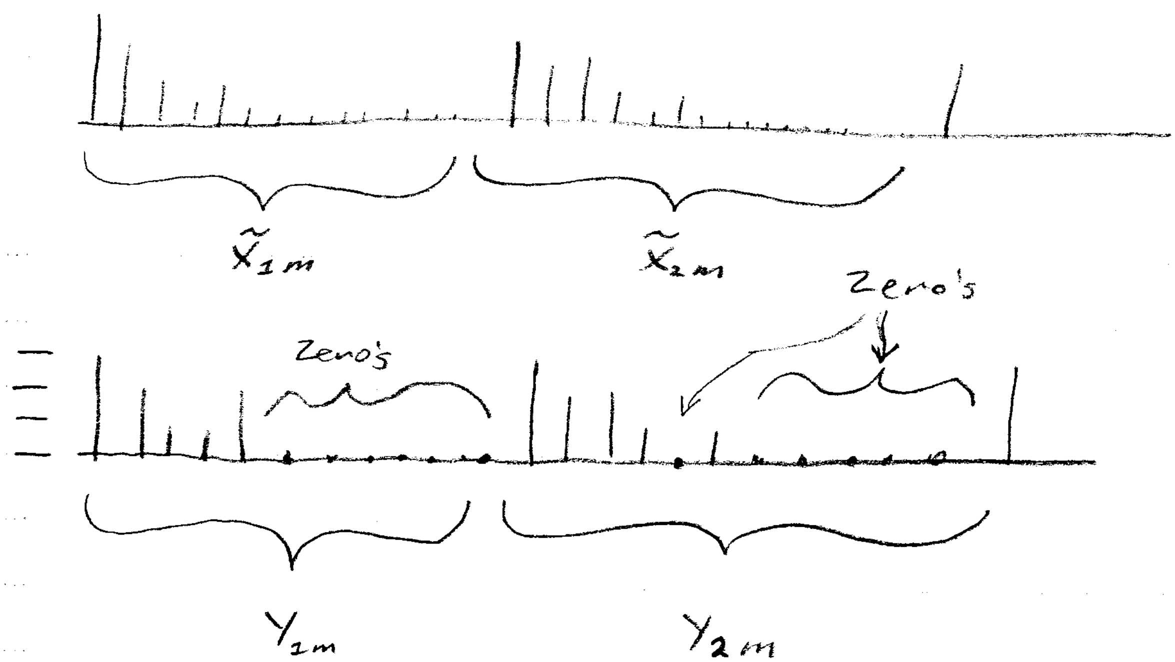
$$Y_{nm} = Q(\tilde{X}_{nm})$$

$1 \leq m \leq N^2$

6) Run length and Entropy code Y_{nm}

Example





Comments:

1) Transform matrix E must be "provided" to decoder.
2) Runs of zeros compress well
3) MMSE – not minimum visual error
4) Non uniform quantizer may improve rate/distortion performance.

Block Coding Using Discrete Cosine Transform

DCT

$k, n \in \{0, \ldots, N-1\}$

$$F(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} f(n) \; C(k) \cos \frac{\pi(2n+1)k}{2N}$$

where

$$C(k) = \begin{cases} 1 & k=0 \\ \sqrt{2} & k=1, \ldots, N-1 \end{cases}$$

$$f(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} F(k) \; C(k) \cos\left(\frac{\pi(2n+1)k}{2N}\right)$$

Define:

$$T_{kn} = \frac{1}{\sqrt{N}} \; C(k) \cos\left(\frac{\pi(2n+1)k}{2N}\right)$$

1) T is an orthonormal transform

$$T^{-1} = T$$

2) Real valued

3) Not the real part of a DFT

Relationship to DFT

$$F(k) = \frac{c(k)}{\sqrt{N}} \sum_{n=0}^{N-1} f(n) \cos\left(\frac{\pi(2n+1)k}{2N}\right)$$

$$= \frac{c(k)}{\sqrt{N}} \sum_{n=0}^{N-1} f(n)\, \mathrm{Re}\left\{ e^{-j\frac{2\pi n k}{2N}}\, e^{-j\frac{\pi k}{2N}} \right\}$$

$$= \frac{c(k)}{\sqrt{N}}\, \mathrm{Re}\left\{ e^{-j\frac{\pi k}{2N}} \sum_{n=0}^{N-1} f(n)\, e^{-j\frac{2\pi n k}{2N}} \right\}$$

Define

$$f_p(n) = \begin{cases} f(n) & 0 \le n \le N-1 \\ 0 & N \le n \le 2N-1 \end{cases}$$

$0 \le k \le N-1$

$$F(k) = \frac{c(k)}{\sqrt{N}}\, \mathrm{Re}\left\{ e^{-j\frac{\pi k}{2N}} \underbrace{\sum_{n=0}^{2N-1} f_p(n)\, e^{-j\frac{2\pi n k}{2N}}} \right\}$$

$$F_p(k) = \underline{\underline{DFT}}\{f_p(n)\}$$

$$= \frac{c(k)}{\sqrt{N}}\, \mathrm{Re}\left\{ e^{-j\frac{\pi k}{2N}}\, F_p(k) \right\}$$

$$= \frac{c(k)}{2\sqrt{N}} \left(F_p(k)\, e^{-j\frac{\pi k}{2N}} + F_p(k)\, e^{+j\frac{\pi k}{2N}} \right)$$

$$= \frac{c(k)\, e^{-j\frac{\pi k}{2N}}}{2\sqrt{N}} \left(F_p(k) + F_p(k)\, e^{+j\frac{2\pi k}{2N}} \right)$$

$$= \frac{c(k)\, e^{-j\frac{\pi k}{2N}}}{2\sqrt{N}} \left(F_p(k) + \Big(F_p(k)\, \underbrace{e^{-j\frac{2\pi k}{2N}}}_{\text{delay by 1 sample}}\Big)^* \right)$$

Notice that for $N=4$

$F_p(k) \Leftrightarrow$	a	b	c	d	0	0	0	0
$F_p(k)e^{-j2\pi\frac{k}{2N}} \Leftrightarrow$	0	a	b	c	d	0	0	0
$(F_p(k)e^{-j2\pi\frac{k}{2N}})^* \Leftrightarrow$	0	0	0	0	d	c	b	a

$F_p(k) + (F_p(k)e^{-j2\pi\frac{k}{2N}})^* \Leftrightarrow$ a b c d b c b a

Define

$$f_r(n) = (f(0), \ldots, f(N-1), f(N-1), \ldots, f(0))$$

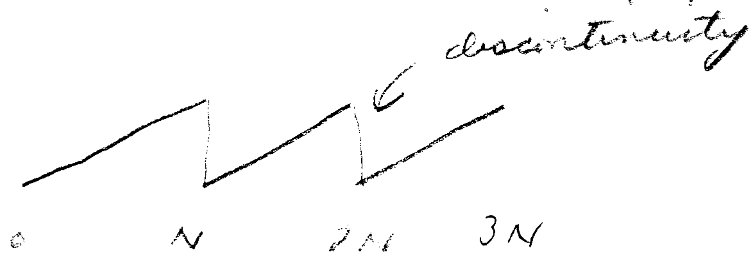
$$F_r(k) = DFT\{f_r(n)\}$$

$$\underset{\substack{\uparrow \\ \text{DCT}}}{F(k)} = \underbrace{\frac{c(k)\, e^{-j\frac{\pi k}{2N}}}{2\sqrt{N}}\, F_r(k)}_{\text{real valued}}$$

DFT of reflected signal

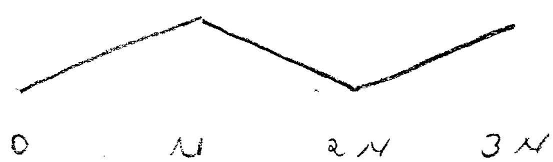
Intuition:

- DFT is transform of periodic extension of signal



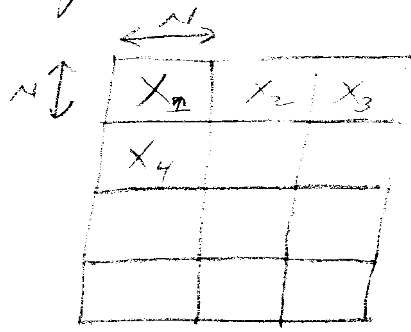
- Discontinuities create high frequency "noise" that must be coded

- DCT is a smoother extension of signal
$\Rightarrow$ less high frequency energy to code.



2-D DCT

Apply DCT to rows and columns



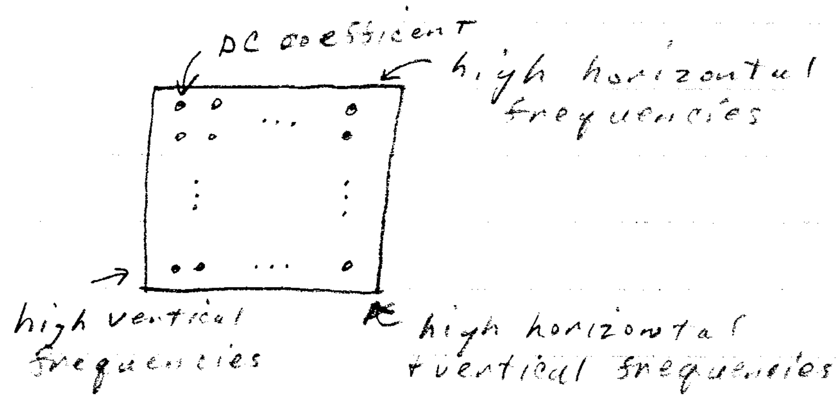
X_n is an $N \times N$ <u>matrix</u>

$T X_n \leftarrow$ DCT applied to each <u>column</u> of X_n

$X_n T^t \leftarrow$ DCT applied to each <u>row</u> of X_n

$\tilde{X}_n = T X_n T^t \leftarrow$ 2-D DCT of X_n

$\tilde{X}_n$



Threshold matrix

$\Delta \leftarrow N \times N$ threshold matrix

$Q(\cdot) \leftarrow$ uniform quantizer with step size of 1

$$Y_{nij} = Q(X_{nij} / \Delta_{ij})$$

larger $\Delta_{mn} \Rightarrow$ larger quantization step size
$\Rightarrow$ more distortion at frequencies corresponding to X_{mn}

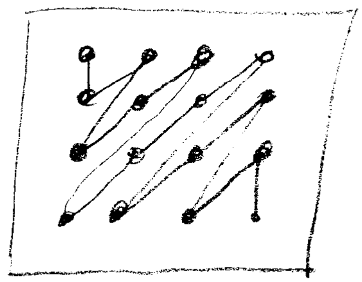
Objective: Choose Δ_{ij} to be well matched to human visual system

higher frequencies $\Rightarrow$ larger Δ_{mn}

- JPEG threshold matrix is <u>important</u>.

Diagonal Scan Patterns

4×4 DCT

$Y_{nij} \Rightarrow$  $\Rightarrow Z_{ni}$

- Smallest coefficients are scanned last.
- Most high frequency coeffiscients are zero.
- Runs of zero compress effectively

Summary of Basic Block Transform Coding Algorithm (JPEG)

1) Break image into blocks of size $N \times N$

2) Compute $N \times N$ DCT of each block

3) Divide by threshold matrix $t_{m,n}$ and quantize with $Q(\cdot)$ to form $Y_{n,i,j}$

4) For each n, perform diagonal scan to form $Z_{n,i}$

5) For each n, run length code $Z_{n,i}$, entropy code result.