CMPT 365 Multimedia Systems

Lossy Compression

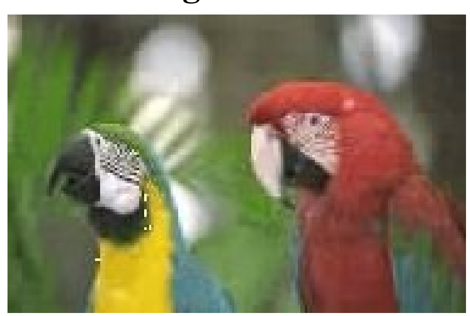
Fall 2023

Lossless vs Lossy Compression

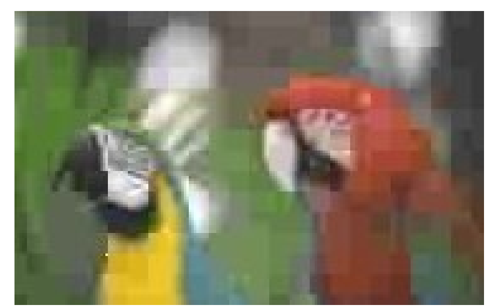
- If the compression and decompression processes induce no information loss, then the compression scheme is **lossless**; otherwise, it is **lossy**.
- Why is lossy compression possible ?

Original

Compression Ratio: 7.7



Compression Ratio: 12.3



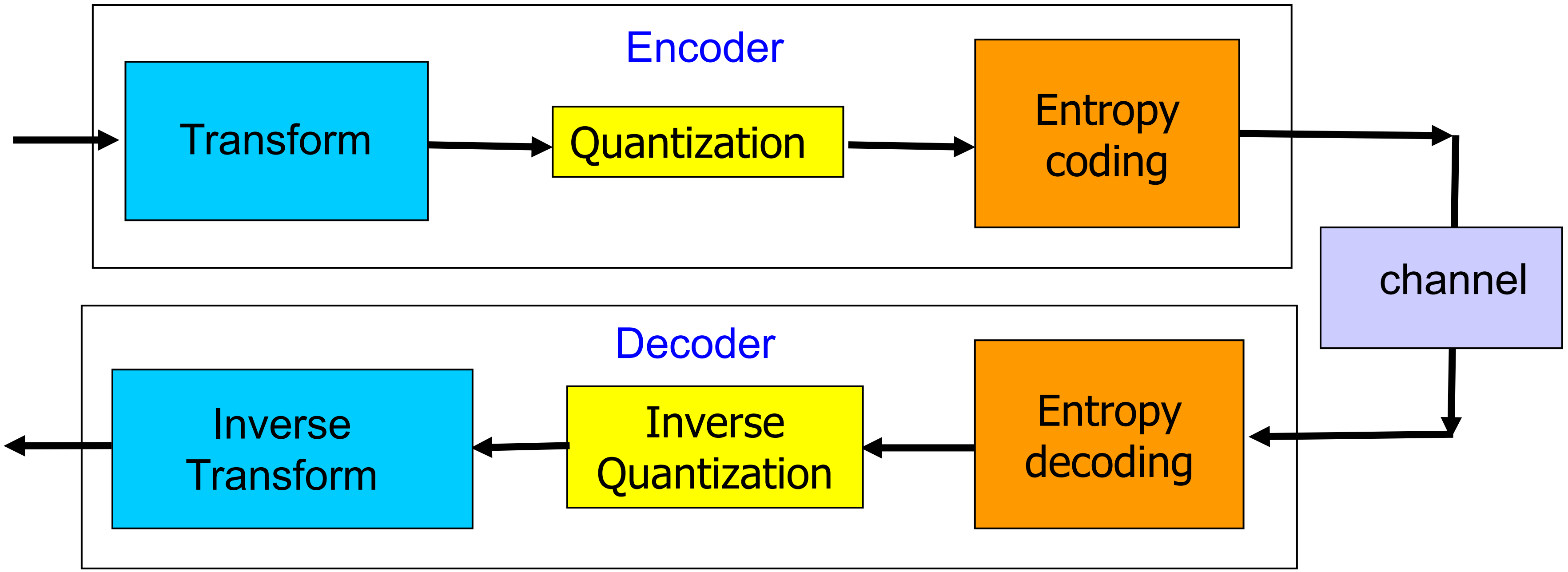
Compression Ratio: 33.9

Outline

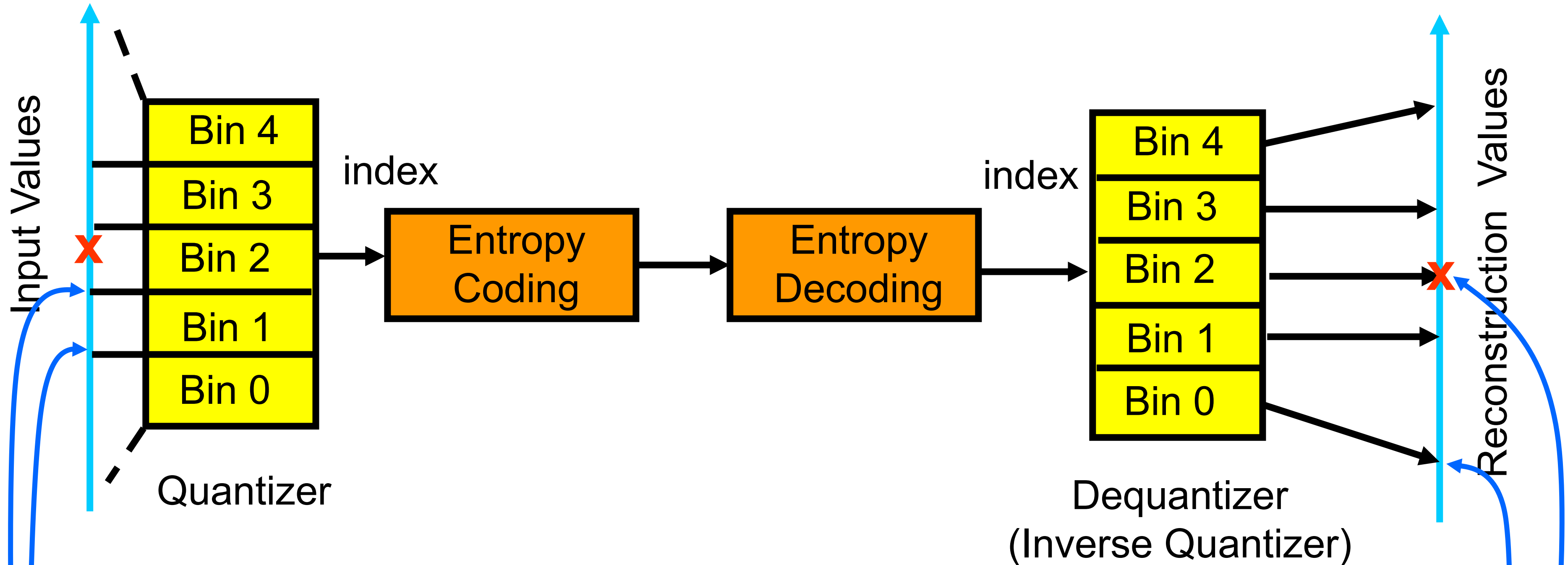
- Quantization
  - Uniform
  - Non-uniform
- Transform coding
  - DCT

Quantization

- The process of representing a large (possibly infinite) set of values with a much smaller set.
  - Example: A/D conversion
- An efficient tool for lossy compression
- Review …



Review: Basic Idea



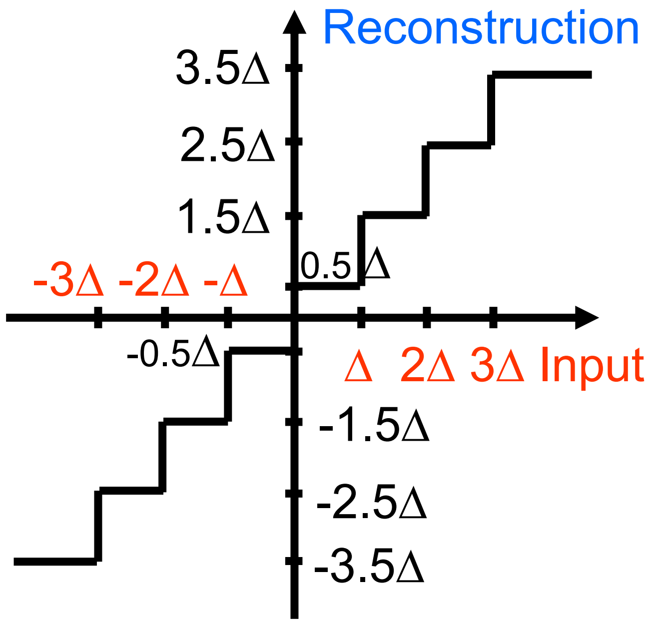
- Quantization is a function that maps an input interval to one integer
- Can reduce the bits required to represent the source.
- Reconstructed result is generally not the original input
- Terminologies:
  - Decision boundaries b_i: bin boundaries
  - Reconstruction levels y_i: output value of each bin by the dequantizer.

Uniform Quantizer

- All bins have the same size except possibly for the two outer intervals:
  - bi and yi are spaced evenly
  - The spacing of bi and yi are both **Δ (step size)**

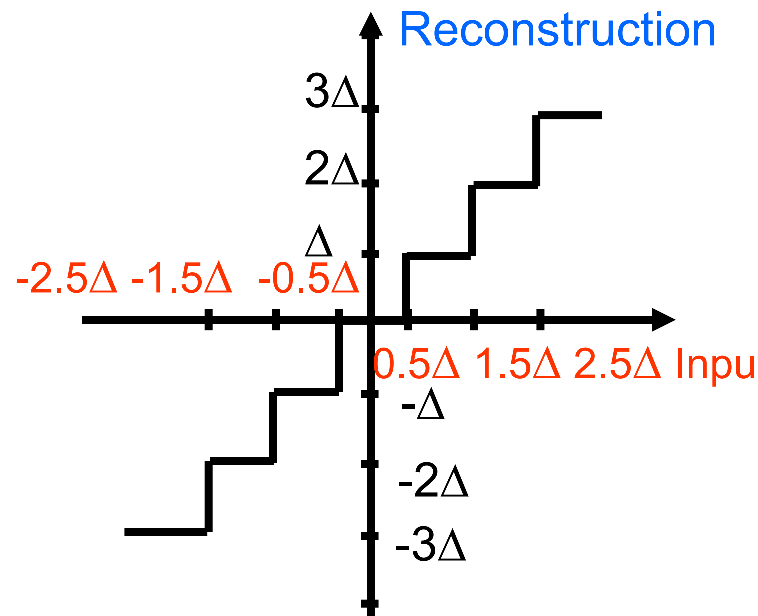
$$y_i = \frac{1}{2}\left(b_{i-1} + b_i\right)$$ for inner intervals.

Uniform Midrise Quantizer



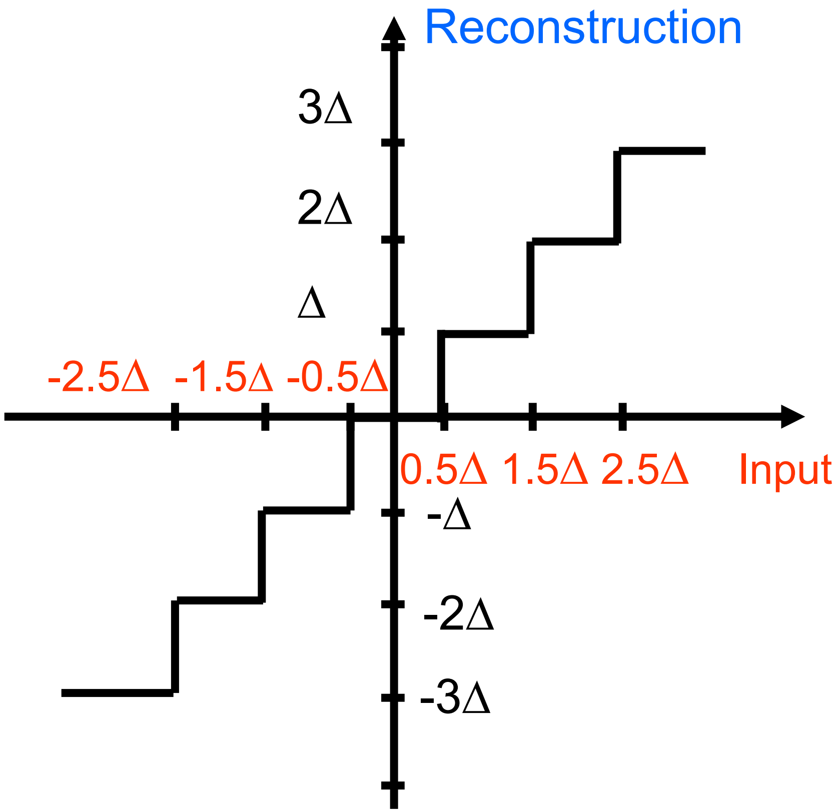
Even number of reconstruction levels
0 is not a reconstruction level

Uniform Midtread Quantizer



Odd number of reconstruction levels
0 is a reconstruction level

Midtread Quantizer



- Quantization mapping:
  Output is an index

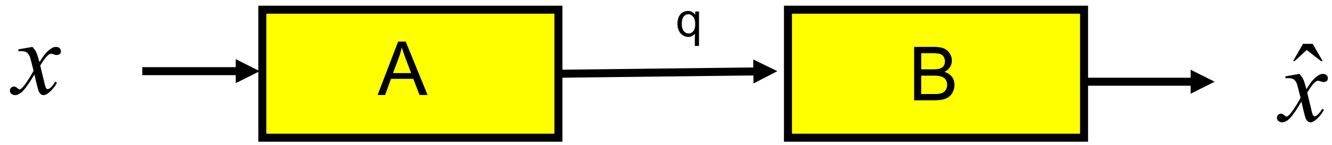
$$q = A(x) = sign(x)\left\lfloor \frac{|x|}{\Delta} + 0.5 \right\rfloor$$

- Example:
  x = -1.8Δ, q = -2.

- De-quantization mapping:

$$\hat{x} = B(q) = q\Delta$$

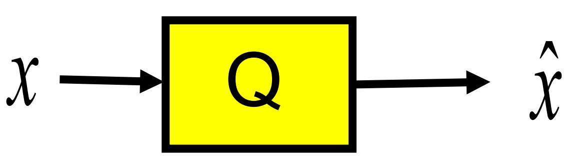
Model of Quantization



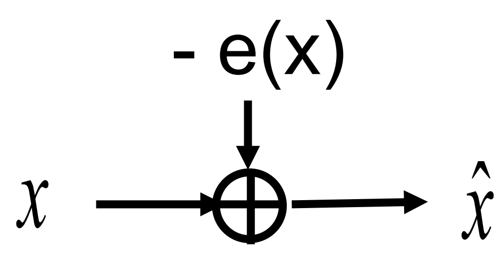
- Quantization: $q = A(x)$
- Inverse Quantization: $\hat{x} = B(q) = B(A(x)) = Q(x)$

B(x) is not exactly the inverse function of A(x), because $\hat{x} \neq x$

- Quantization error: $e(x) = x - \hat{x}$
- Combining quantizer and de-quantizer:

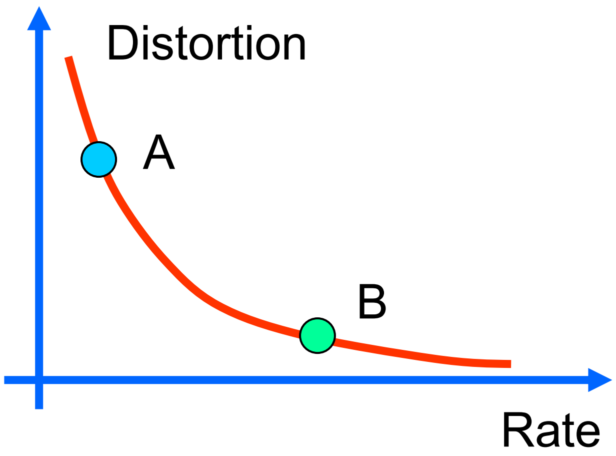


or



Rate-Distortion Tradeoff

- Things to be determined:
  - Number of bins
  - Bin boundaries
  - Reconstruction levels



- A tradeoff between rate and distortion:
  - To reduce the size of the encoded bits, we need to reduce the number of bins
  - Less bins ➔ More reconstruction errors

Measure of Distortion

- Quantization error: $e(x) = x - \hat{x}$
- Mean Squared Error (MSE) for Quantization
  - Average quantization error of all input values
  - Need to know the probability distribution of the input

- Number of bins: M
- Decision boundaries: b_i, i = 0, …, M
- Reconstruction Levels: y_i, i = 1, …, M
- Reconstruction: $\hat{x} = y_i \quad \text{iff } b_{i-1} < x \le b_i$

- MSE: $$MSE_q = \int_{-\infty}^{\infty} (x - \hat{x})^2 f(x)dx = \sum_{i=1}^{M} \int_{b_{i-1}}^{b_i} (x - y_i)^2 f(x)dx$$
  - Same as the variance of e(x) if μ = E{e(x)} = 0 (zero mean).
  - Definition of Variance: $$\sigma_e^2 = \int_{-\infty}^{\infty} (e - \mu_e)^2 f(e)de$$

Rate-Distortion Optimization

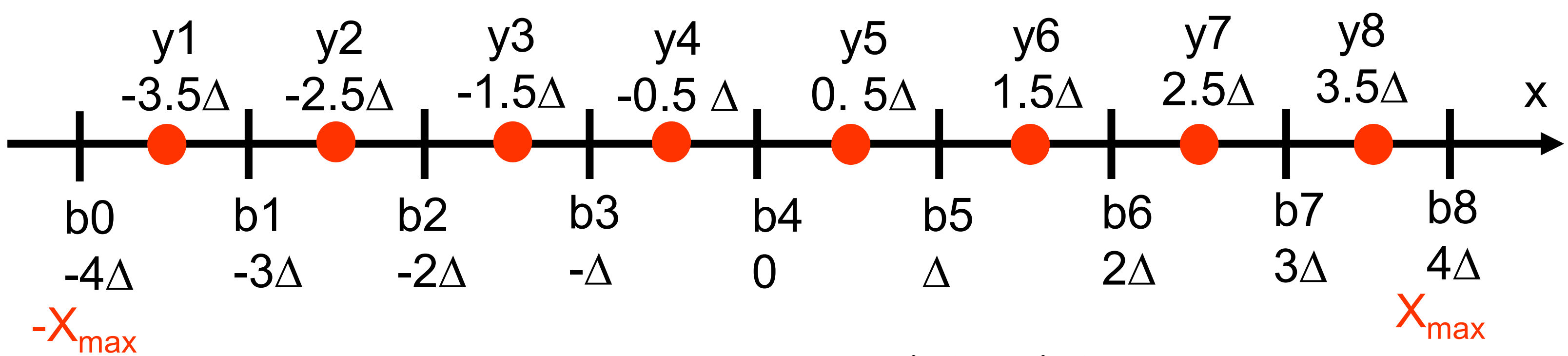
- Two Scenarios:
  - Given M, find b_i and y_i that minimize the MSE.
  - Given a distortion constraint D, find M, b_i and y_i such that the MSE $\leq$ D.

Outline

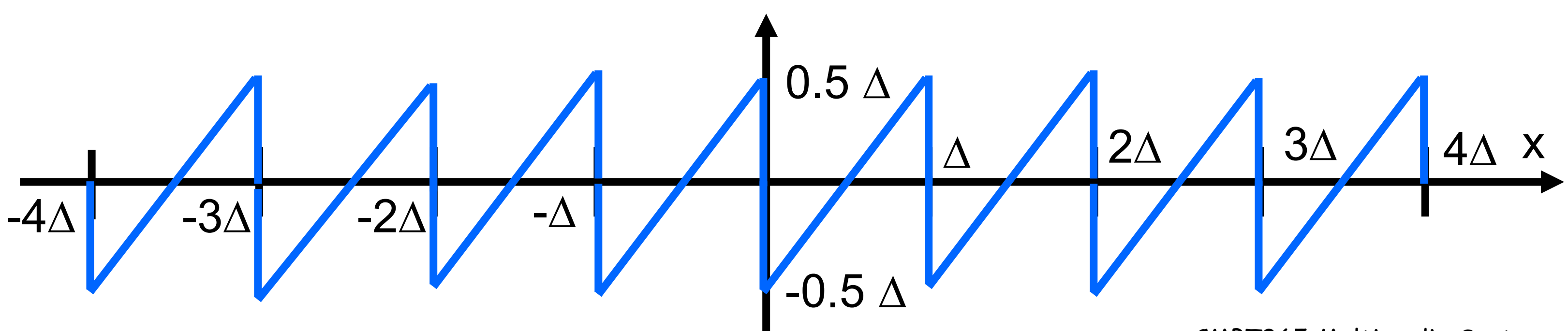
- Quantization
  - Uniform
  - Non-uniform
- Transform coding
  - DCT

Uniform Quantization of a Uniformly Distributed Source

- Input X: uniformly distributed in $[-X_{max}, X_{max}]$: $f(x)= 1 / (2X_{max})$
- Number of bins: M (even for midrise quantizer)
- Step size is easy to get: $\Delta = 2X_{max} / M$.
- $b_i = (i - M/2)\,\Delta$



- ➔ e(x) is uniformly distributed in $[-\Delta/2, \Delta/2]$.



Uniform Quantization of a Uniformly Distributed Source

- MSE

$$MSE_q = \int_{-\infty}^{\infty} (x - \hat{x})^2 f(x)dx = \sum_{i=1}^{M} \int_{b_{i-1}}^{b_i} (x - y_i)^2 f(x)dx$$

$$= M \frac{1}{2X_{\max}} \int_0^{\Delta} \left(x - \frac{\Delta}{2} \right)^2 dx = \frac{M}{2X_{\max}} \frac{1}{12} \Delta^3 = \frac{1}{12} \Delta^2$$

- M increases, Δ decreases, MSE decreases
- Variance of a random variable uniformly distributed in [- Δ/2, Δ/2]:

$$\sigma^2{}_q = \int_{-\Delta/2}^{\Delta/2} (x - 0)^2 \frac{1}{\Delta} dx = \frac{1}{12} \Delta^2$$

- Optimization: Find M such that MSE ≤ D

$$\frac{1}{12} \Delta^2 \le D \Rightarrow \frac{1}{12} \left(\frac{2X_{\max}}{M} \right)^2 \le D \Rightarrow M \ge X_{\max} \sqrt{\frac{1}{3D}}$$

Signal to Noise Ratio (SNR)

- Variance is a measure of signal energy
- Let $M = 2^n$
- Each bin index is represented by n bits

$$SNR(dB) = 10\log_{10}\frac{Signal\ Energy}{Noise\ Energy} = 10\log_{10}\frac{1/12(2X_{\max})^2}{1/12\Delta^2}$$

$$= 10\log_{10}\frac{(2X_{\max})^2}{(2X_{\max}/M)^2} = 10\log_{10}M^2 = 10\log_{10}2^{2n} = (20\log_{10}2)n$$

$$\approx 6.02n \ \ dB$$

- If $n \rightarrow n+1$, Δ is halved, noise variance reduces to 1/4, and SNR increases by 6 dB.

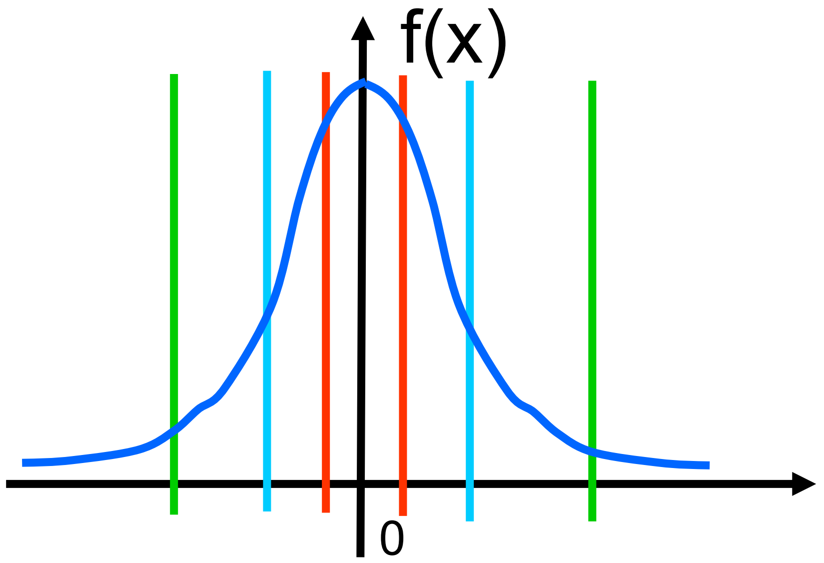
Outline

- Quantization
  - Uniform
  - Non-uniform
- Transform coding
  - DCT

Non-uniform Quantization

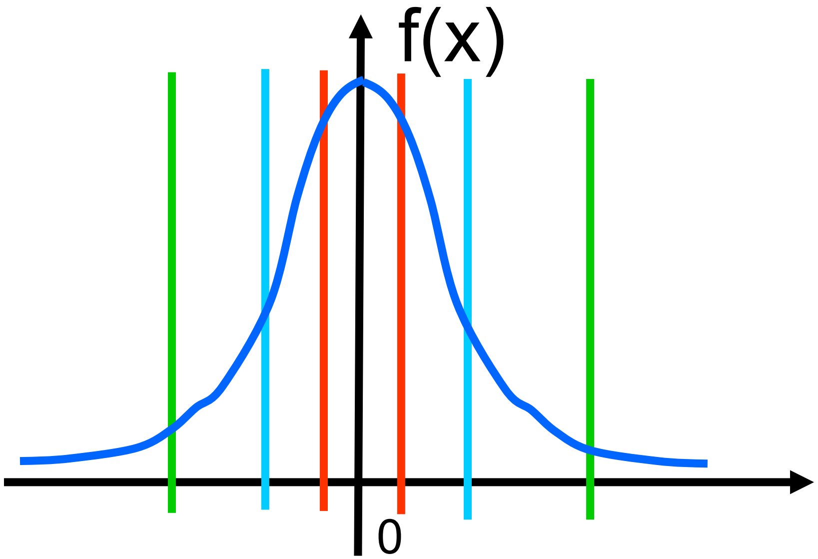
- Uniform quantizer is not optimal if source is not uniformly distributed
- For given M, to reduce MSE, we want narrow bin when f(x) is high and wide bin when f(x) is low

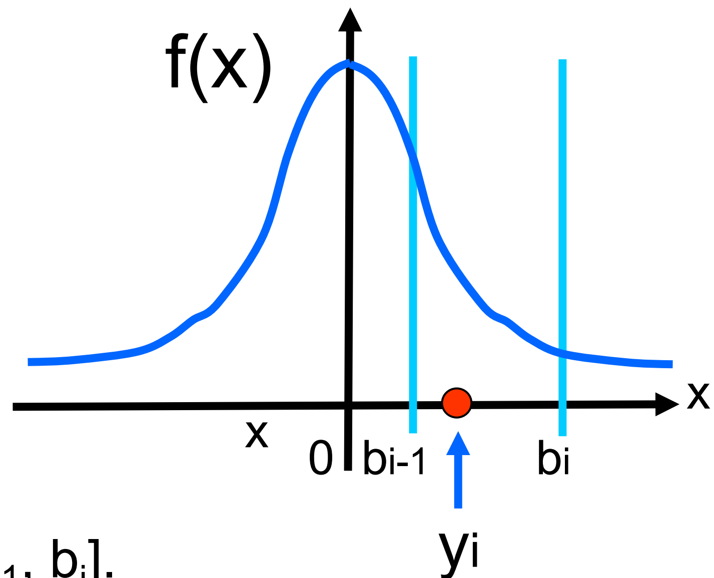
$$\sigma_q^2 = \int_{-\infty}^{\infty} (x - \hat{x})^2 f(x)dx = \sum_{k=1}^{M} \int_{b_{k-1}}^{b_k} (x - y_k)^2 f(x)dx$$



Lloyd-Max Quantizer

$$\sigma_q^2 = \int_{-\infty}^{\infty} (x - \hat{x})^2 f(x)dx = \sum_{k=1}^{M} \int_{b_{k-1}}^{b_k} (x - y_k)^2 f(x)dx$$





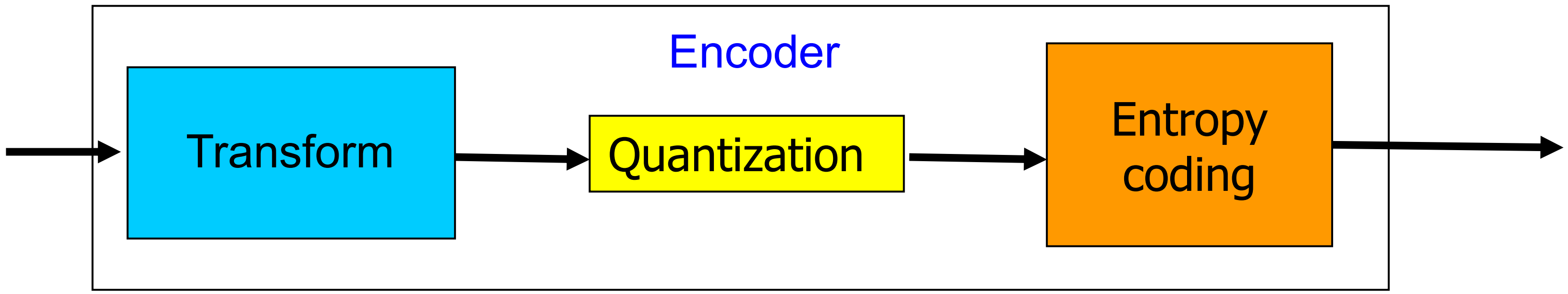
y_i is the centroid of interval $[b_{i-1}, b_i]$.

Outline

- Quantization
  - Uniform quantization
  - Non-uniform quantization
- Transform coding
  - Discrete Cosine Transform (DCT)

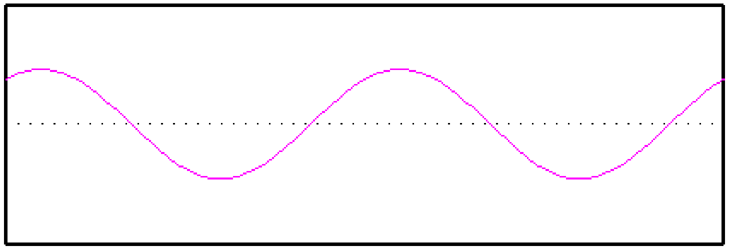
Why Transform Coding ?

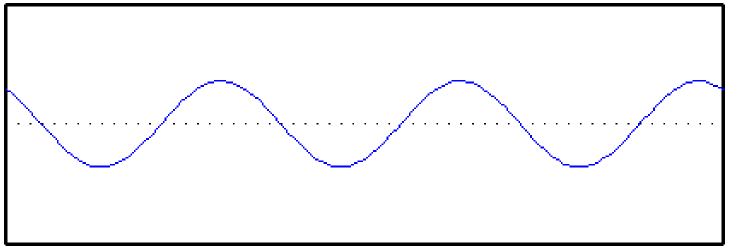
- Transform
  - From one domain/space to another space
  - Time -> Frequency
  - Spatial/Pixel -> Frequency
- Purpose of transform
  - Remove correlation between input samples
  - Transform most energy of an input block into a few coefficients
  - Small coefficients can be discarded by quantization without too much impact to reconstruction quality

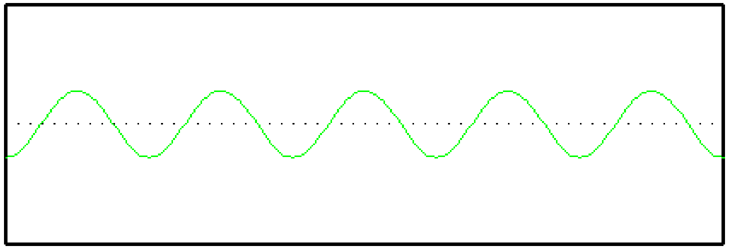


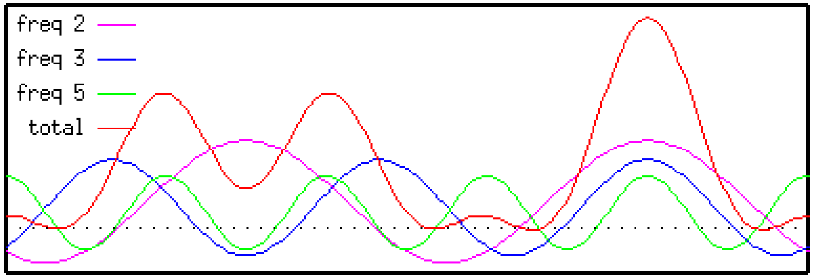
1-D Example

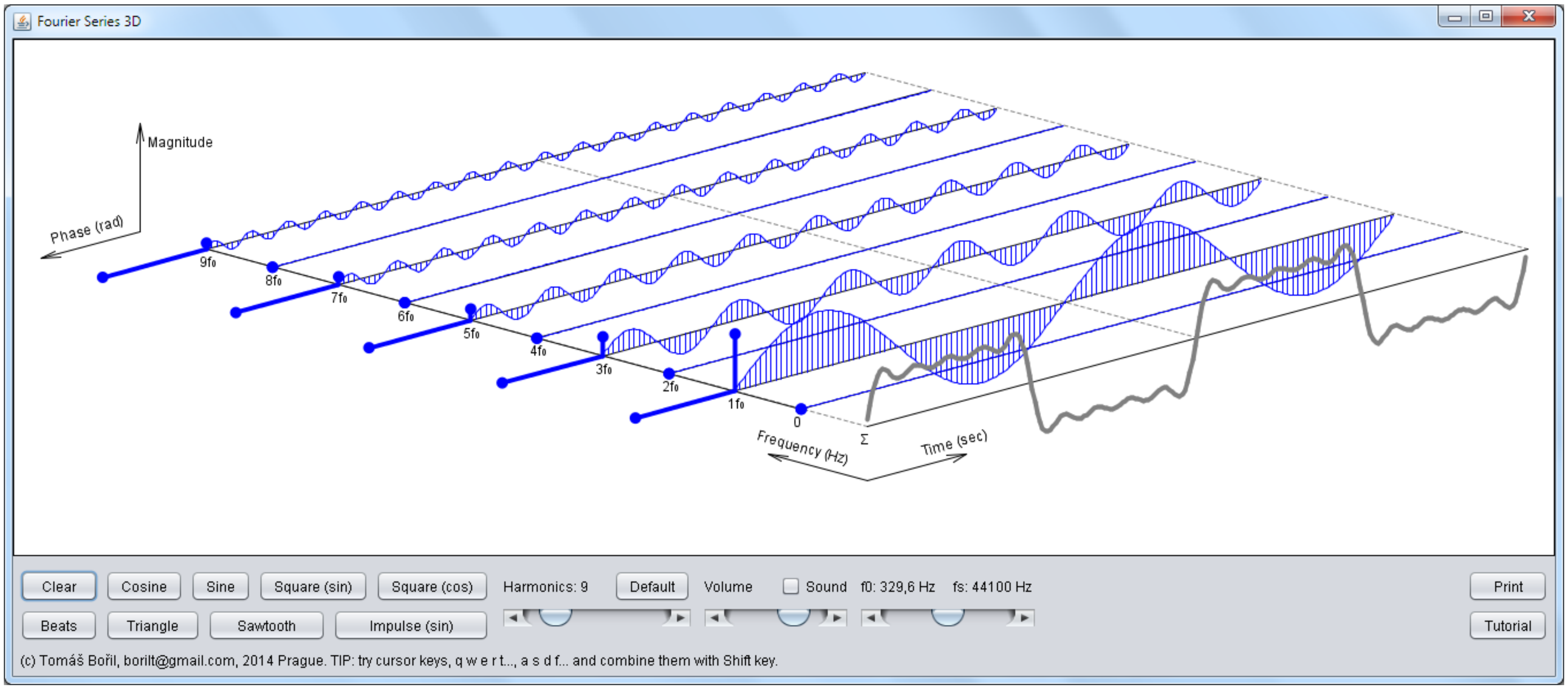
- Fourier Transform










Fourier Series 3D
Magnitude
Phase (rad)
9f0
8f0
7f0
6f0
5f0
4f0
3f0
2f0
1f0
0
Σ
Frequency (Hz)
Time (sec)
Clear
Cosine
Sine
Square (sin)
Square (cos)
Harmonics: 9
Default
Volume
Sound
f0: 329,6 Hz
fs: 44100 Hz
Print
Beats
Triangle
Sawtooth
Impulse (sin)
Tutorial
(c) Tomáš Bořil, borilt@gmail.com, 2014 Prague. TIP: try cursor keys, q w e r t..., a s d f... and combine them with Shift key.

1-D Example

- https://www.youtube.com/watch?v=1kFeBtsevz0
- https://www.youtube.com/watch?v=t8ZE1MIkg2s
  - Sine wave/sound/piano

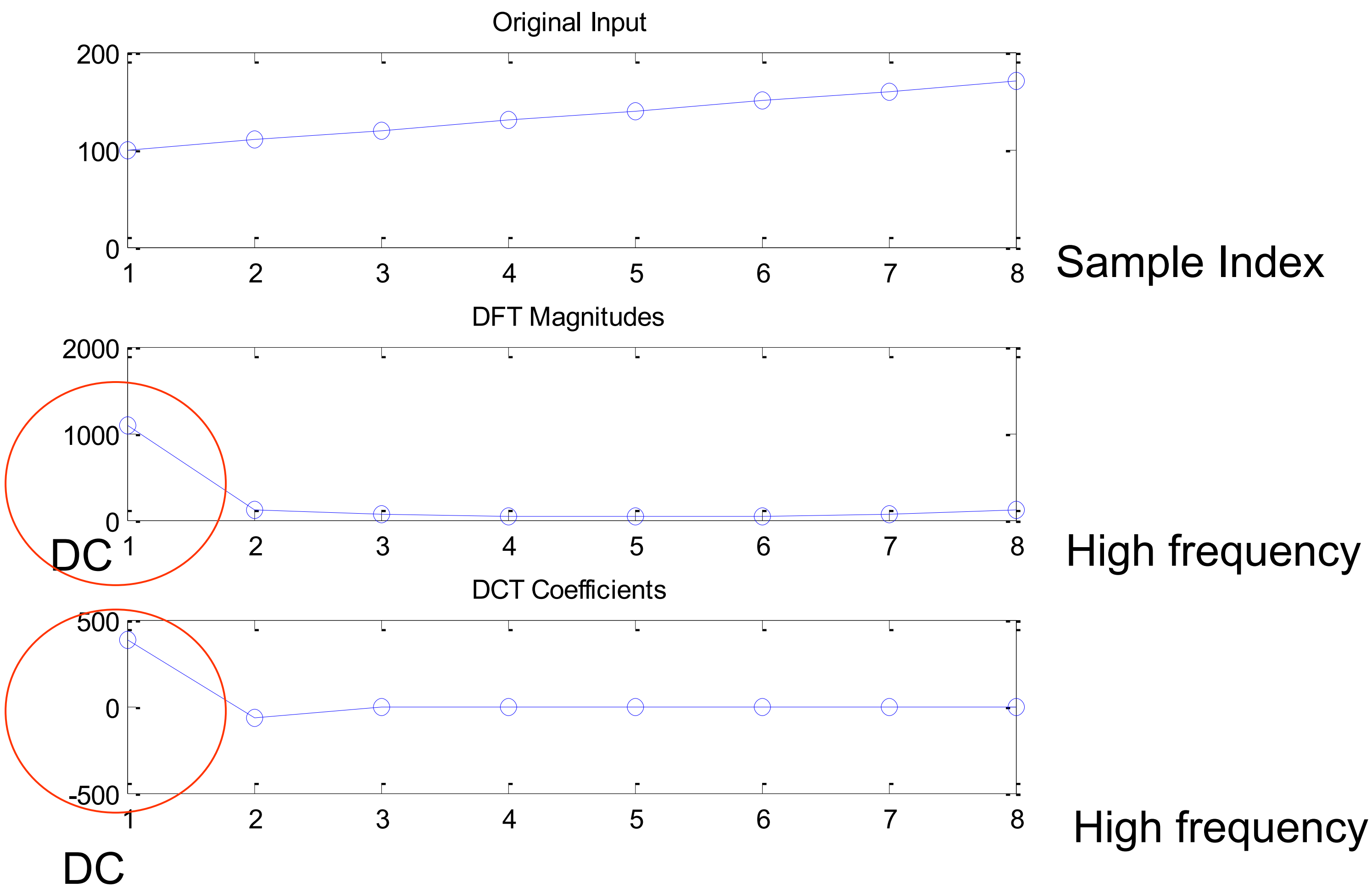
- www.sagebrush.com/mousing.htm
  - An electronic instrument that allows direct control of pitch and amplitude

Rationale behind Transform

- If **Y** is the result of a linear transform **T** of the input vector **X** in such a way that the components of **Y** are much less correlated, then **Y** can be coded more efficiently than **X**.
- If most information is accurately described by the first few components of a transformed vector, then the remaining components can be coarsely quantized, or even set to zero, with little signal distortion.

1-D Example

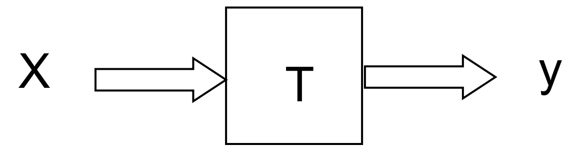
- Smooth signals have strong DC (direct current, or zero frequency) and low frequency components, and weak high frequency components



Matrix Representation of Transform

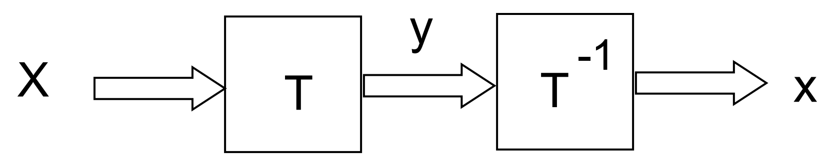
- Linear transform is an N x N matrix:

$$\mathbf{y}_{N\times 1} = \mathbf{T}_{N\times N}\mathbf{x}_{N\times 1}$$



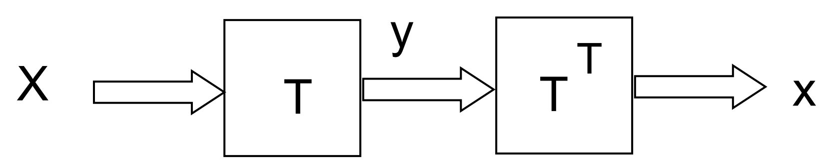
- Inverse Transform:

$$\mathbf{x} = \mathbf{T}^{-1}\mathbf{y}$$



- Unitary Transform (aka orthonormal):

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



- For unitary transform: rows/cols have unit norm and are orthogonal to each others

$$\mathbf{T}\mathbf{T}^{T} = \mathbf{I} \;\Rightarrow\; \mathbf{t}_i\mathbf{t}_j^{T} = \delta_{ij} = \begin{cases} 1, & \mathrm{i} = \mathrm{j} \\ 0, & \mathrm{i} \neq \mathrm{j} \end{cases}$$

Discrete Cosine Transform (DCT)

- DCT - close to optimal (known as KL Transform) but much simpler and faster

- Definition:

$$c_{i,j} = a\cos\left(\frac{(2j+1)i\pi}{2N}\right), \quad \mathrm{i,j} = 0, \ldots, \mathrm{N}-1.$$

$$a = \sqrt{1/N} \quad \text{for i} = 0,$$

$$a = \sqrt{2/N} \quad \text{for i} = 1, \ldots, \mathrm{N}-1.$$

- Matlab function:
  - `dct(eye(N));`

DCT

- Definition:

$$C_{i,j} = a \cos\left(\frac{(2j+1)\, i\, \pi}{2N}\right), \quad \mathrm{i, j = 0, ..., N-1.}$$

$$a = \sqrt{1/N} \quad \text{for i} = 0,$$

$$a = \sqrt{2/N} \quad \text{for i} = 1, \ldots, \mathrm{N-1}.$$

- N = 2 (Haar Transform):

$$\mathbf{C}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

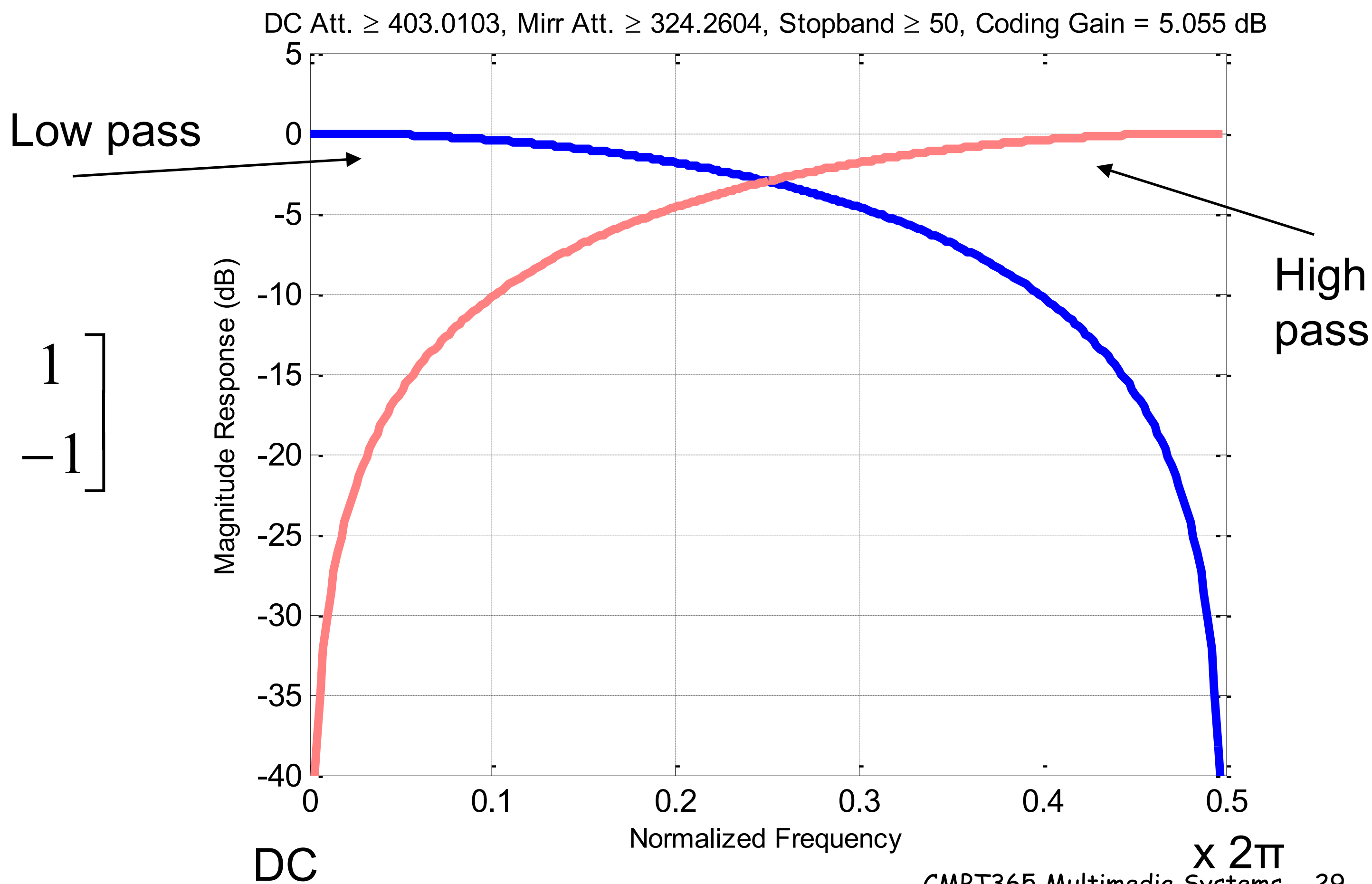
$$\begin{bmatrix} y_0 \\ y_1 \end{bmatrix} = \mathbf{C}_2 \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} x_0 + x_1 \\ x_0 - x_1 \end{bmatrix}$$

- y_0 captures the mean of x_0 and x_1 (low-pass)
  - $x_0 = x_1 = 1$ ➔ y_0 = sqrt(2) (DC), $y_1 = 0$
- y1 captures the difference of x0 and x1 (high-pass)
  - $x_0 = 1$, $x_1 = -1$ ➔ $y_0 = 0$ (DC), y_1 = sqrt(2).

DCT

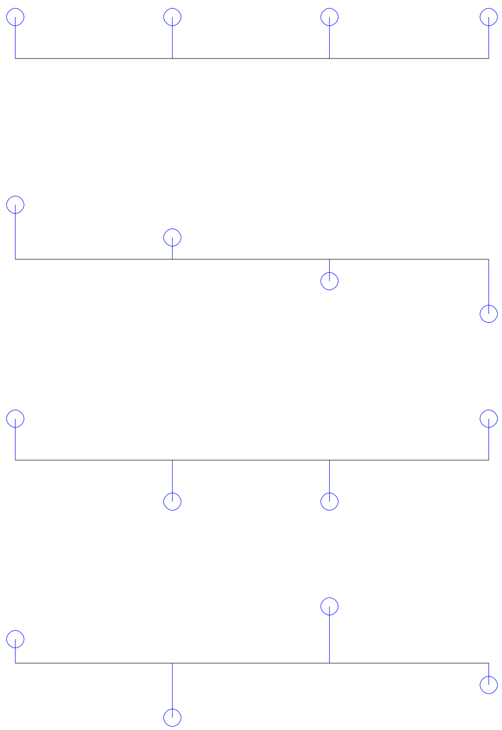
- Magnitude Frequency Responses of 2-point DCT:
  - Can be obtained by **freqz( )** in Matlab.

$$\mathbf{C}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$



4-point DCT

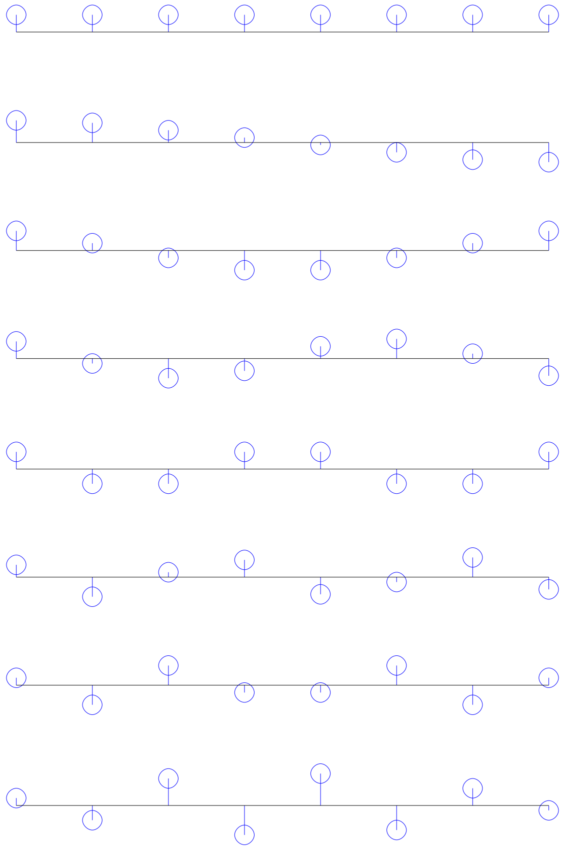
0.5000	0.5000	0.5000	0.5000
0.6533	0.2706	-0.2706	-0.6533
0.5000	-0.5000	-0.5000	0.5000
0.2706	-0.6533	0.6533	-0.2706



$$C = \sqrt{\frac{2}{(4)}} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \cos\frac{\pi}{2(4)} & \cos\frac{3\pi}{2(4)} & \cos\frac{5\pi}{2(4)} & \cos\frac{7\pi}{2(4)} \\ \cos\frac{2\pi}{2(4)} & \cos\frac{6\pi}{2(4)} & \cos\frac{10\pi}{2(4)} & \cos\frac{14\pi}{2(4)} \\ \cos\frac{3\pi}{2(4)} & \cos\frac{9\pi}{2(4)} & \cos\frac{15\pi}{2(4)} & \cos\frac{21\pi}{2(4)} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}}\cos\frac{\pi}{8} & \frac{1}{\sqrt{2}}\cos\frac{3\pi}{8} & \frac{1}{\sqrt{2}}\cos\frac{5\pi}{8} & \frac{1}{\sqrt{2}}\cos\frac{7\pi}{8} \\ \frac{1}{\sqrt{2}}\cos\frac{\pi}{4} & \frac{1}{\sqrt{2}}\cos\frac{3\pi}{4} & \frac{1}{\sqrt{2}}\cos\frac{5\pi}{4} & \frac{1}{\sqrt{2}}\cos\frac{7\pi}{4} \\ \frac{1}{\sqrt{2}}\cos\frac{3\pi}{8} & \frac{1}{\sqrt{2}}\cos\frac{9\pi}{8} & \frac{1}{\sqrt{2}}\cos\frac{15\pi}{8} & \frac{1}{\sqrt{2}}\cos\frac{21\pi}{8} \end{bmatrix}$$

8-point DCT



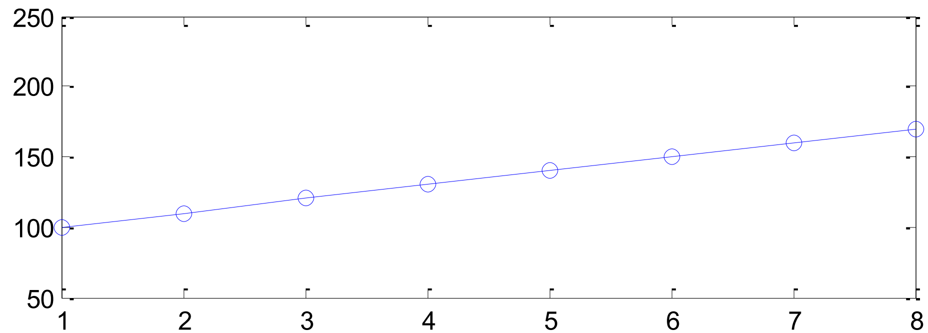
$$T = \begin{bmatrix} .3536 & .3536 & .3536 & .3536 & .3536 & .3536 & .3536 & .3536 \\ .4904 & .4157 & .2778 & .0975 & -.0975 & -.2778 & -.4157 & -.4904 \\ .4619 & .1913 & -.1913 & -.4619 & -.4619 & -.1913 & .1913 & .4619 \\ .4157 & -.0975 & -.4904 & -.2778 & .2778 & .4904 & .0975 & -.4157 \\ .3536 & -.3536 & -.3536 & .3536 & .3536 & -.3536 & -.3536 & .3536 \\ .2778 & -.4904 & .0975 & .4157 & -.4157 & -.0975 & .4904 & -.2778 \\ .1913 & -.4619 & .4619 & -.1913 & -.1913 & .4619 & -.4619 & .1913 \\ .0975 & -.2778 & .4157 & -.4904 & .4904 & -.4157 & .2778 & -.0975 \end{bmatrix}$$

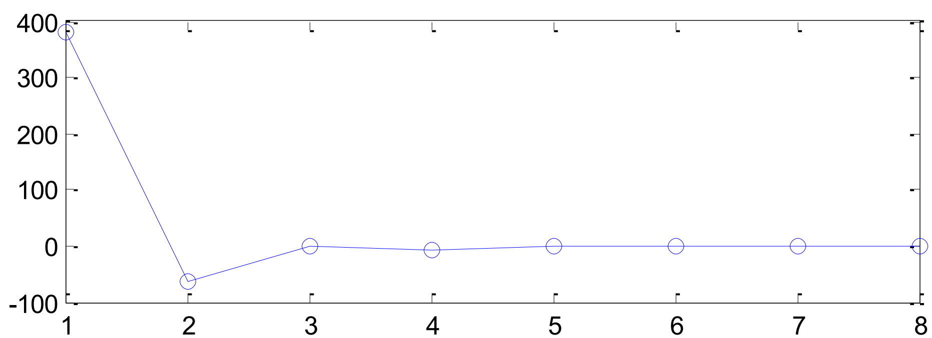
Example

- x = [100 110 120 130 140 150 160 170]^T;
- 8-point DCT:

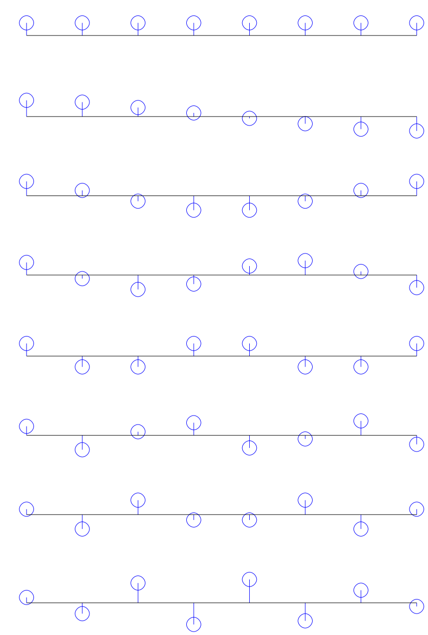
  [381.8377, -64.4232, 0.0, -6.7345, 0.0, -2.0090, 0.0, -0.5070]

  Most energy are in the first 2 coefficients.





$$T = \begin{bmatrix} .3536 & .3536 & .3536 & .3536 & .3536 & .3536 & .3536 & .3536 \\ .4904 & .4157 & .2778 & .0975 & -.0975 & -.2778 & -.4157 & -.4904 \\ .4619 & .1913 & -.1913 & -.4619 & -.4619 & -.1913 & .1913 & .4619 \\ .4157 & -.0975 & -.4904 & -.2778 & .2778 & .4904 & .0975 & -.4157 \\ .3536 & -.3536 & -.3536 & .3536 & .3536 & -.3536 & -.3536 & .3536 \\ .2778 & -.4904 & .0975 & .4157 & -.4157 & -.0975 & .4904 & -.2778 \\ .1913 & -.4619 & .4619 & -.1913 & -.1913 & .4619 & -.4619 & .1913 \\ .0975 & -.2778 & .4157 & -.4904 & .4904 & -.4157 & .2778 & -.0975 \end{bmatrix}$$



Interpretation of Transform

- Forward transform y = Tx (x is N x 1 vector)
  - Let t_i be the i-th row of T
  - ➔ $y_i = t_i x = \langle t_i^T, x \rangle$ (Inner product)
  - y_i measures the similarity between x and t_i
  - Higher similarity ➔ larger transform coefficient

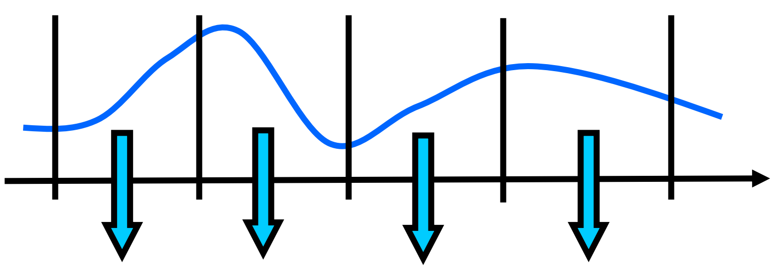
- Inverse transform:

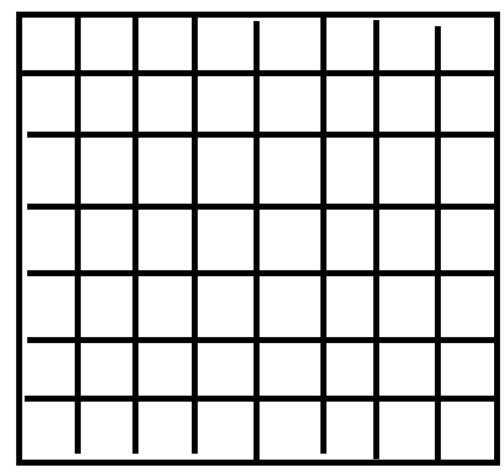
$$\mathbf{x} = \mathbf{T}^T \mathbf{y} = \begin{bmatrix} \mathbf{t}_0^T & \mathbf{t}_1^T & \dots & \mathbf{t}_{N-1}^T \end{bmatrix} \mathbf{y} = \sum_{i=0}^{N-1} \mathbf{t}_i^T y_i$$

- x is the weighted combination of t_i.
  - Rows of T are called basis vectors.

2D Block Transform

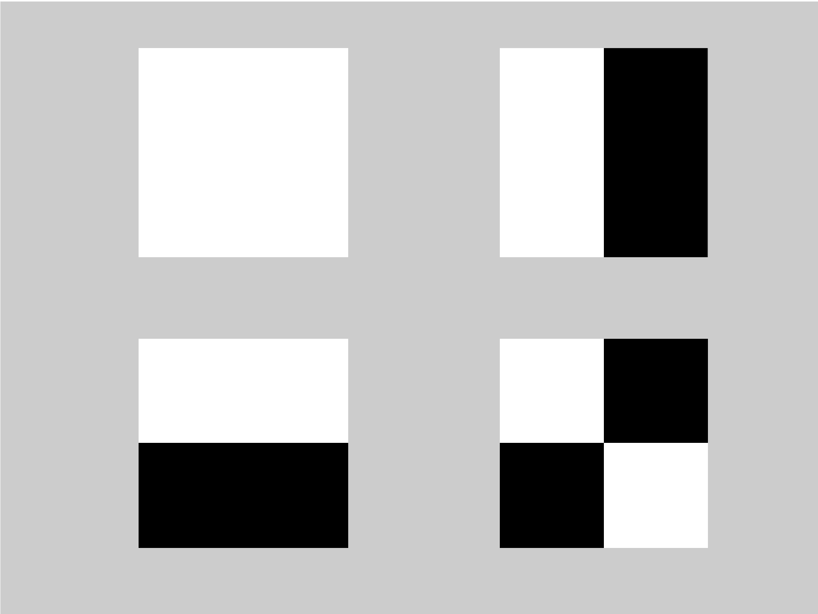
- Divide input data into blocks (2D)
- Encode each block separately (sometimes with information from neighboring blocks)
- Examples:
  - Most DCT-based image/video coding standards



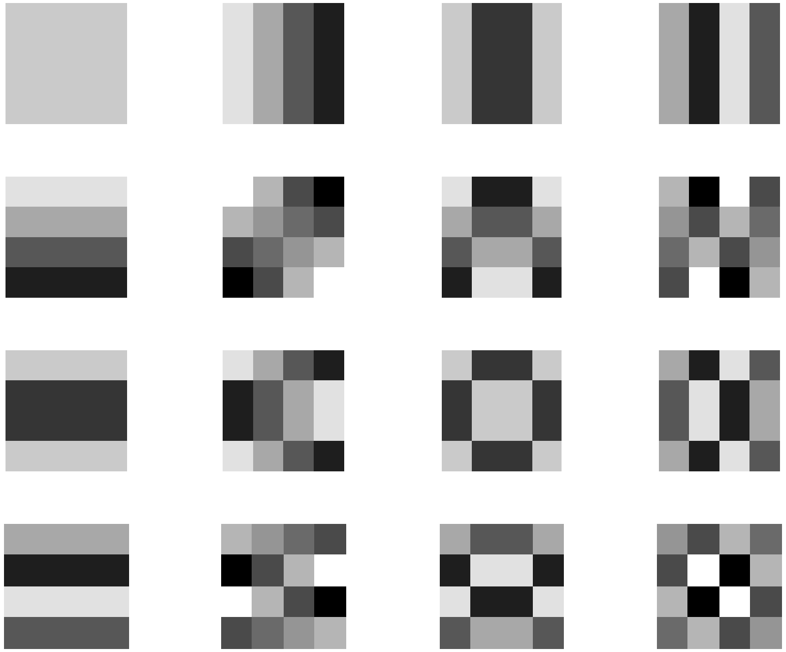


2-D DCT Basis

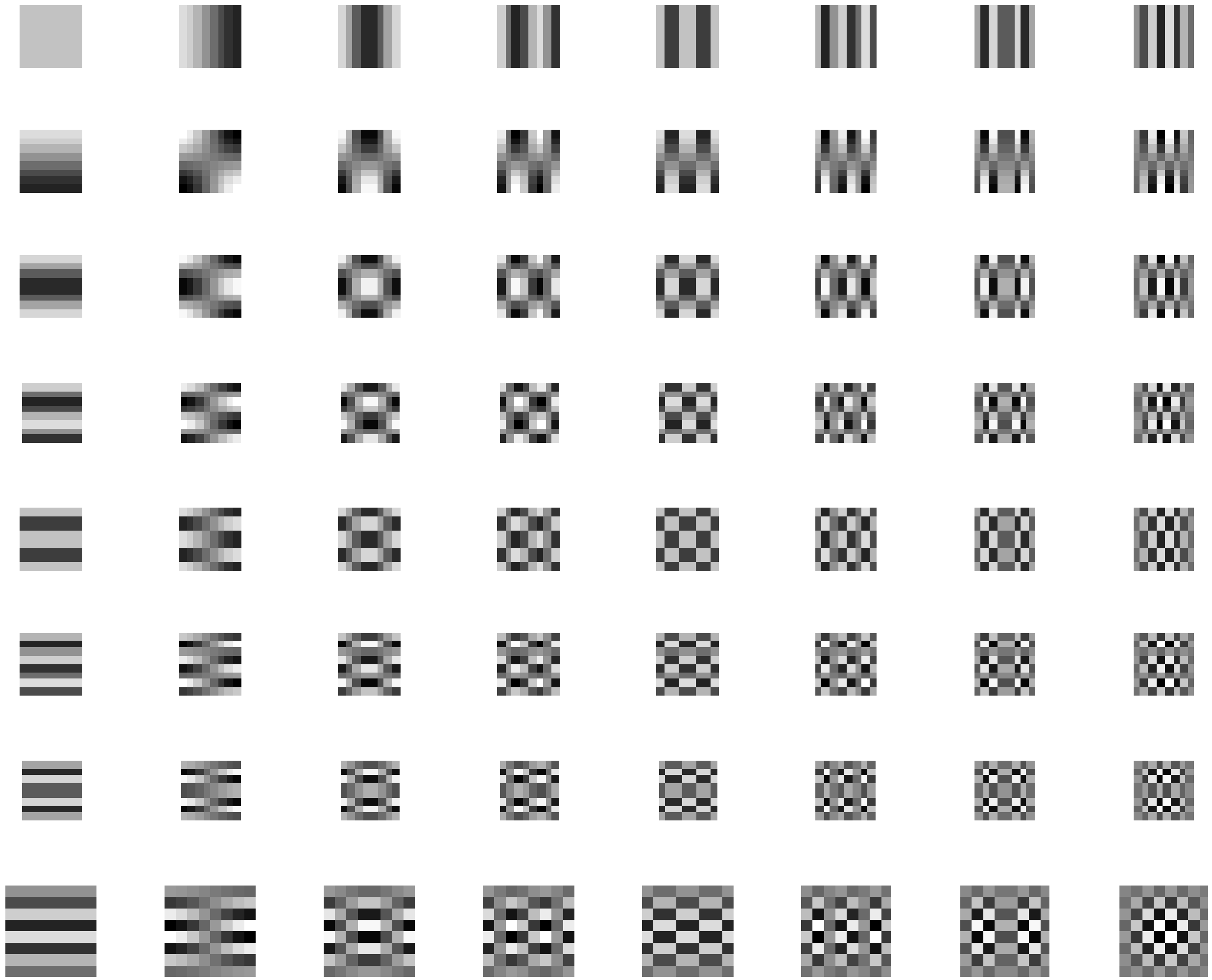
For 2-point DCT



For 4-point DCT

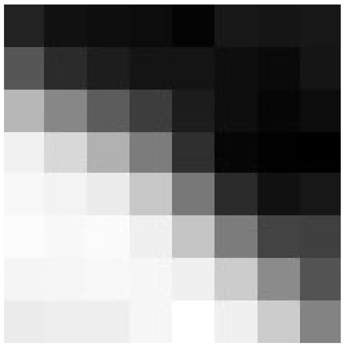


2-D DCT Basis Matrices: 8-point DCT



2-D 8-point DCT Example

- Original Data:



89	78	76	75	70	82	81	82
122	95	86	80	80	76	74	81
184	153	126	106	85	76	71	75
221	205	180	146	97	71	68	67
225	222	217	194	144	95	78	82
228	225	227	220	193	146	110	108
223	224	225	224	220	197	156	120
217	219	219	224	230	220	197	151

- 2-D DCT Coefficients (after rounding to integers):

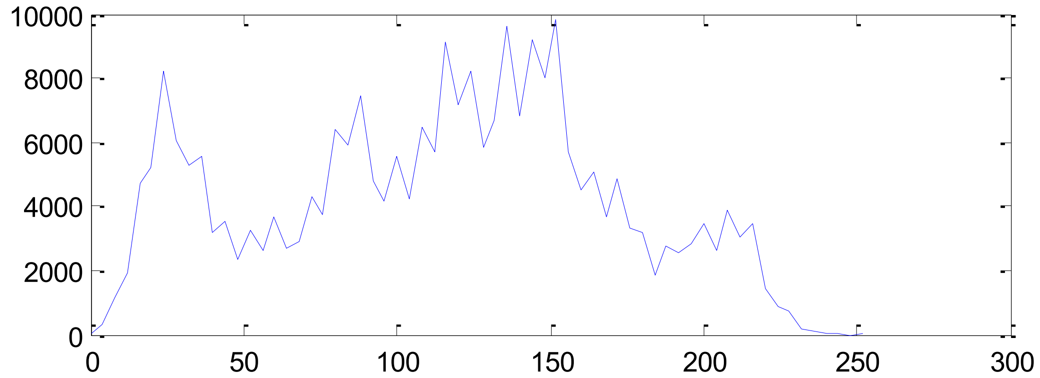
1155	259	-23	6	11	7	3	0
-377	-50	85	-10	10	4	7	-3
-4	-158	-24	42	-15	1	0	1
-2	3	-34	-19	9	-5	4	-1
1	9	6	-15	-10	6	-5	-1
3	13	3	6	-9	2	0	-3
8	-2	4	-1	3	-1	0	-2
2	0	-3	2	-2	0	0	-1

Most energy is in the upper-left corner

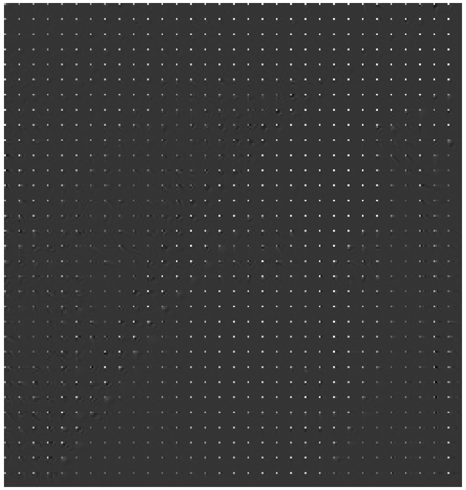
2-D Example

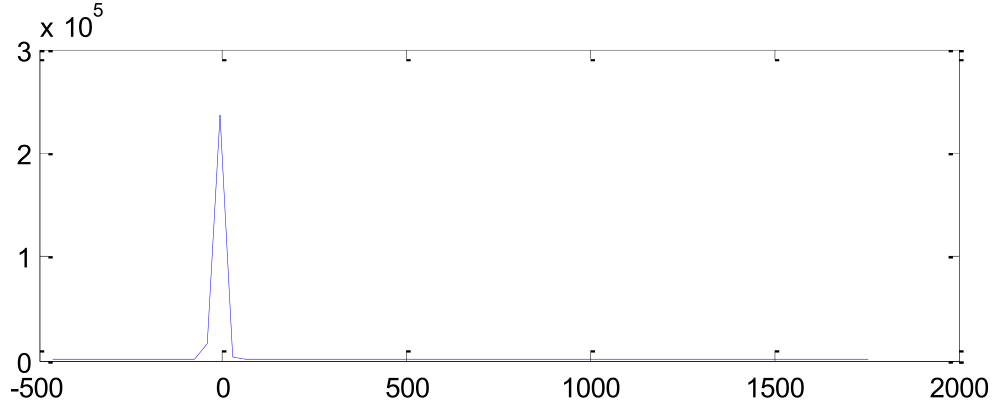
- Apply transform to each 8x8 block
- Histograms of source and DCT coefficients

Original Image



2-D DCT Coefficients. Min= -465.37, max= 1789.00



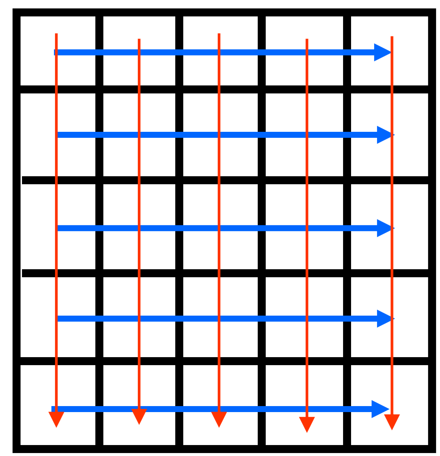


- Most transform coefficients are around 0.
- Desired for compression

From 1D to 2D: Separable Transform

- X: N x N input block
- T: N x N 1D transform
- A = TX: Apply T to each column of X
- B=XT^T: Apply T to each row of X
- 2-D Separable Transform:
  - Apply T to each row
  - Then apply T to each column

$$\mathbf{Y} = \mathbf{T}\mathbf{X}\mathbf{T}^T$$



- Inverse Transform:

$$\mathbf{X} = \mathbf{T}^T\mathbf{Y}\mathbf{T}$$

- Not all 2D transforms are separable, but DCT is !

Further Exploration

- **Textbook 8.1-8.5**