Lecture 14:
Predictive and Transform Coding

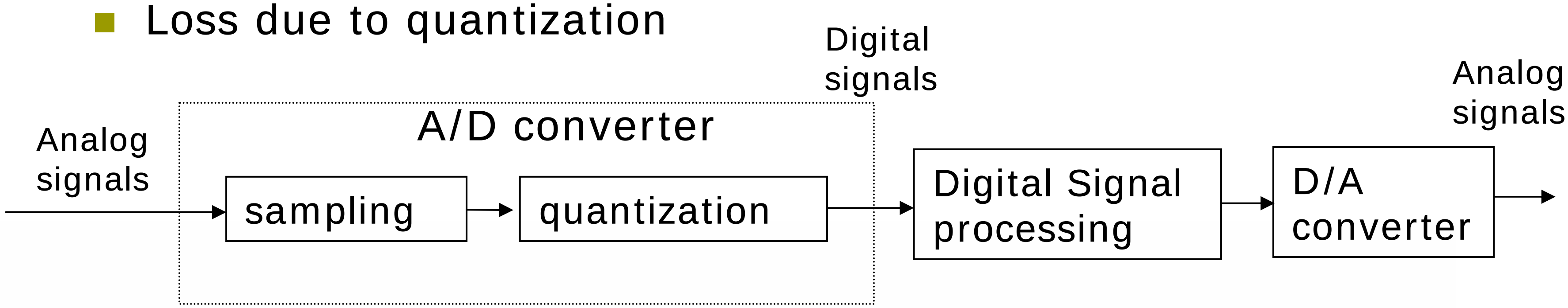
Thinh Nguyen
Oregon State University

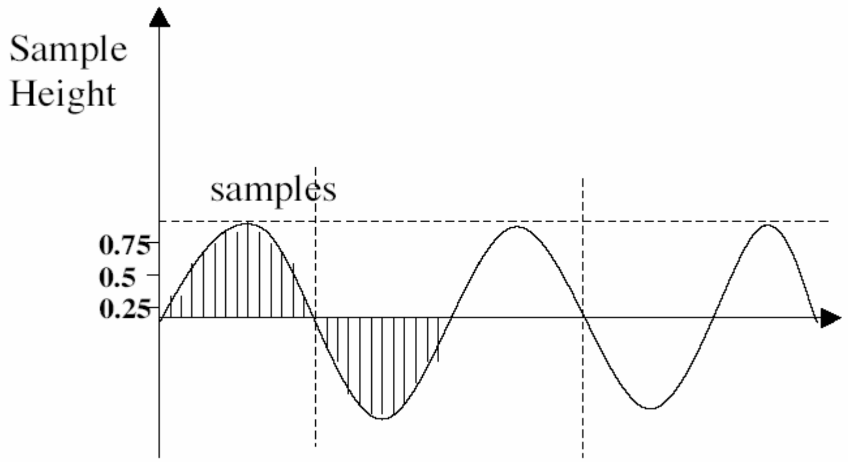
Outline

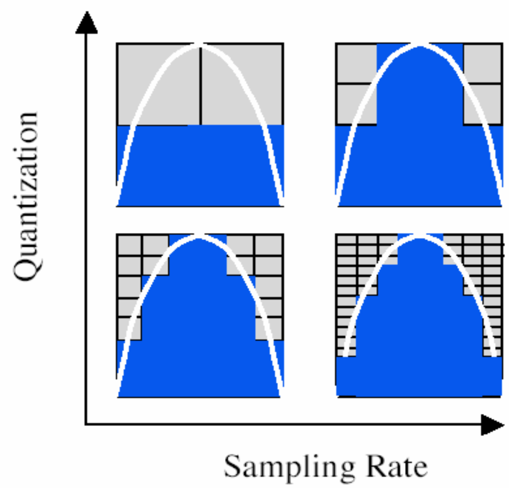
- PCM (Pulse Code Modulation)
- DPCM (Differential Pulse Code Modulation)
- Transform coding
- JPEG

Digital Signal Representation

- Loss from A/D conversion
  - Aliasing (worse than quantization loss) due to sampling
  - Loss due to quantization







Signal representation

- For perfect re-construction, sampling rate (Nyquist's frequency) needs to be twice the maximum frequency of the signal.
- However, in practice, loss still occurs due to quantization.
- Finer quantization leads to less error at the expense of increased number of bits to represent signals.

Audio sampling

- Human hearing frequency range: 20 Hz to 20 Khz.
- Voice:50Hz to 2 KHz

- What is the sampling rate to avoid aliasing? (worse than losing information)

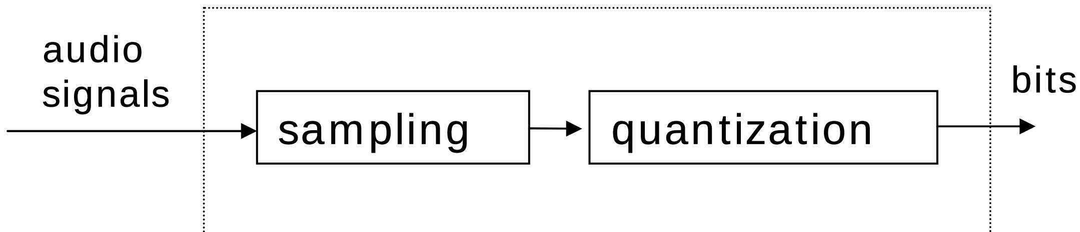
Audio CD : 44100Hz

Audio quantization

- Sample precision – resolution of signal
- Quantization depends on the number of bits used.
- Voice quality: 8 bit quantization, 8000 Hz sampling rate. (64 kbps)
- CD quality: 16 bit quantization, 44100Hz (705.6 kbps for mono, 1.411 Mbps for stereo).

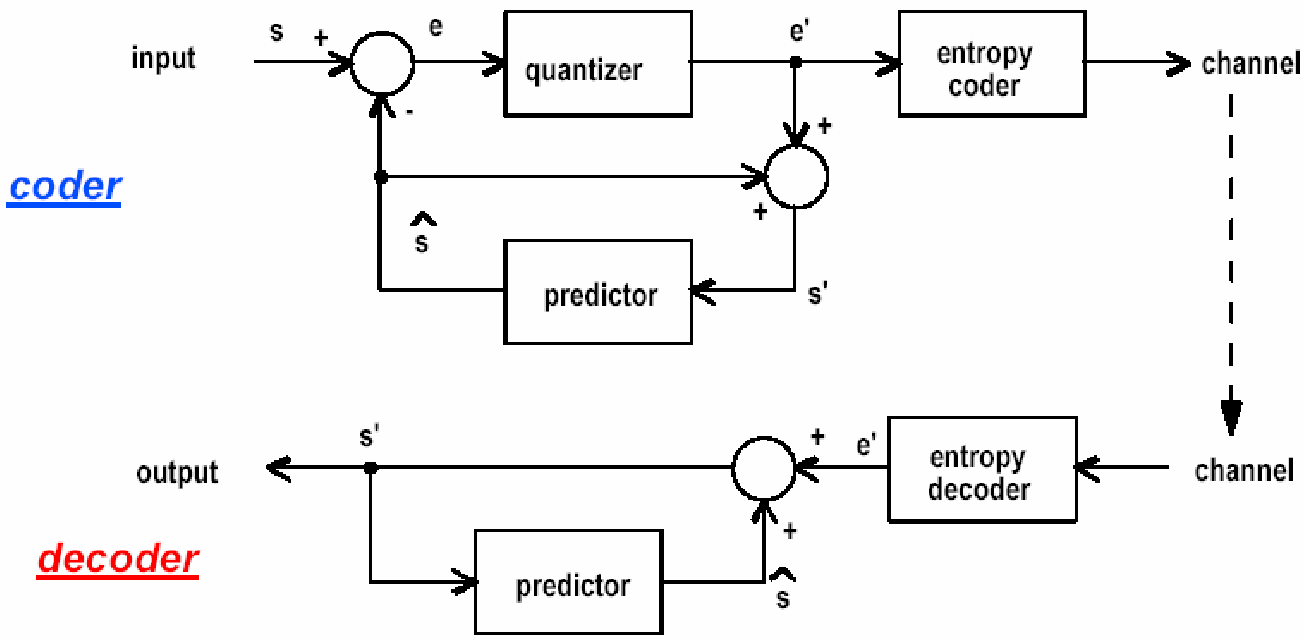
Pulse Code Modulation (PCM)

- The 2 step process of sampling and quantization is known as *Pulse Code Modulation.*
- Used in speech and CD recording.



No compression, unlike MP3

DPCM (Differential Pulse Code Modulation)



Prediction error

$$e = s - \hat{s}$$

Reconstruction

$$s' = e' + \hat{s}$$

Reconstruction error = quantization error

$$s' - s = e' - e = q$$

DPCM (Differential Pulse Code Modulation)

- Simple example:
- Code the following value sequence:
  - 1.4 1.75 2.05 2.5 2.4
  - Quantization step: 0.2
  - Predictor: current value = previous quantized value + quantized error.

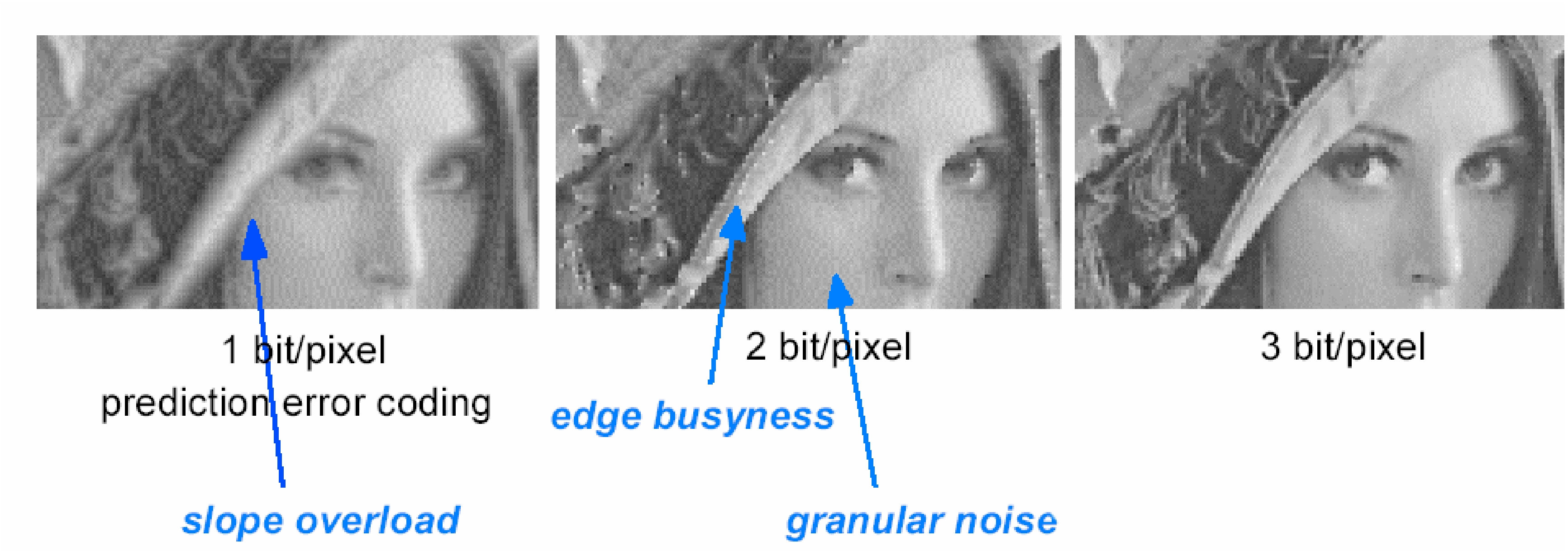
  - Error = 1.75-1.4 = 0.35 => .4
  - Prediction value = 1.4 + 0.4 = 1.8

  - Error = 2.05 - 1.8 = => 0.25 => 0.2
  - Prediction value = 1.8 + .2 = 2.0

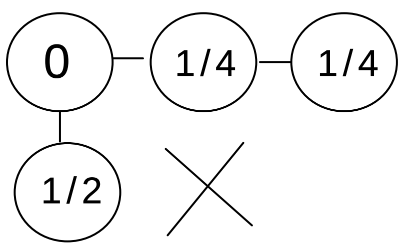
  - Error = 2.5 – 2.0 = 0.5=> 0.4
  - Prediction value = 2.0 + .4 = 2.4

Send 1.4, 0.4, 0.2, 0.4, ...

DPCM - Image



Linear predictor



DPCM - Image

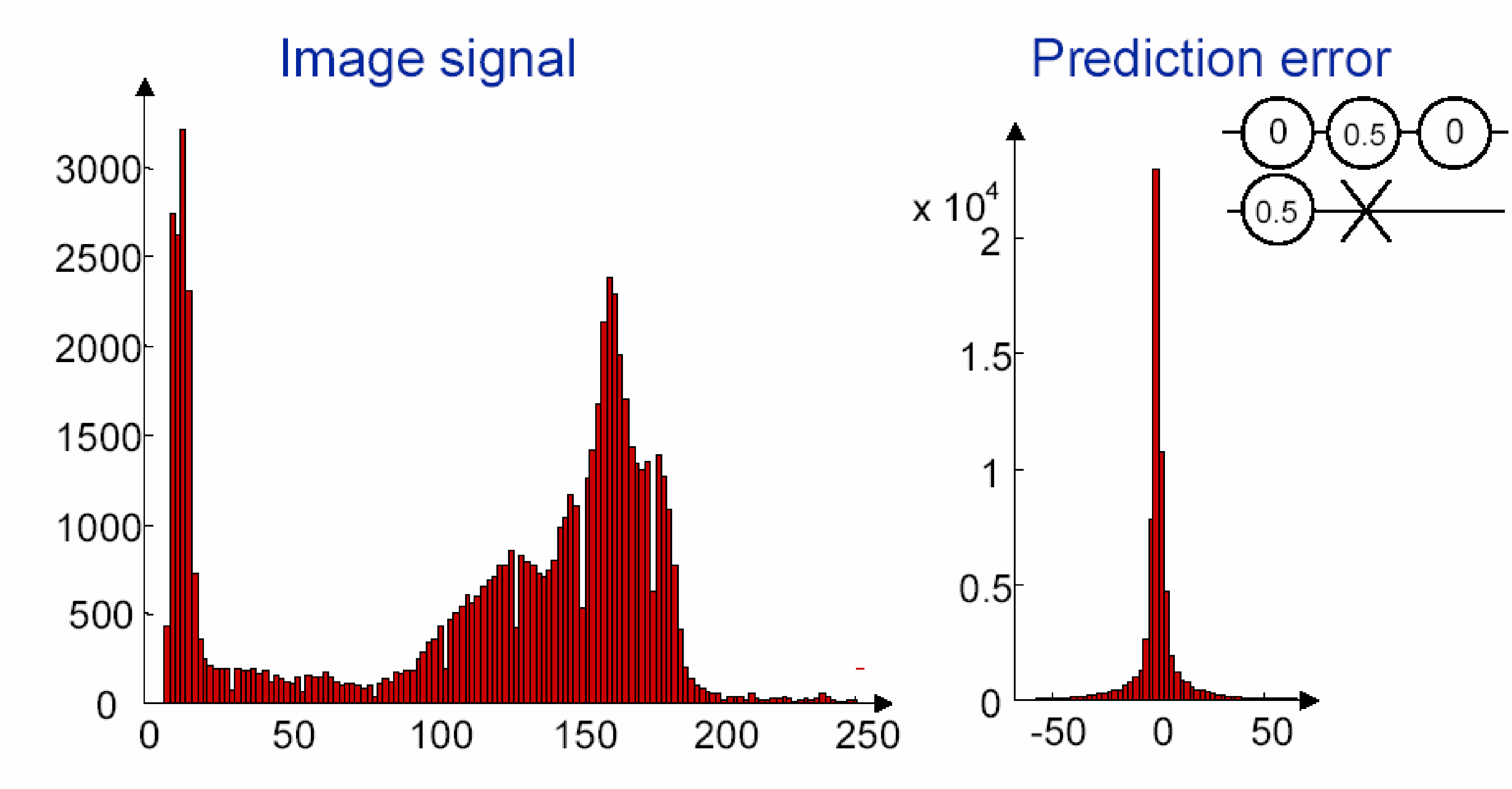


Image signal
3000
2500
2000
1500
1000
500
0
0
50
100
150
200
250
Prediction error
x 10^4
2
1.5
1
0.5
0
-50
0
50
0
0.5
0
0.5

Transmission errors in a DPCM system

- For a linear DPCM decoder, the transmission error response is superimposed to the reconstructed signal.

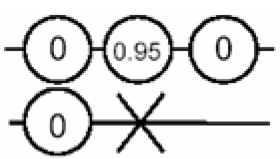
- For variable length coding, e.g. Huffman, loss of synchronization due to a single bit error can lead to many prediction error samples.

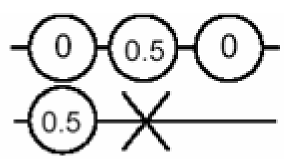
Transmission errors in a DPCM system

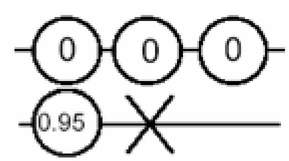
Example: Lena, 3 *bpp* (fixed code word length)

Error rate: $p=10^{-3}$

Match the following predictors to each images



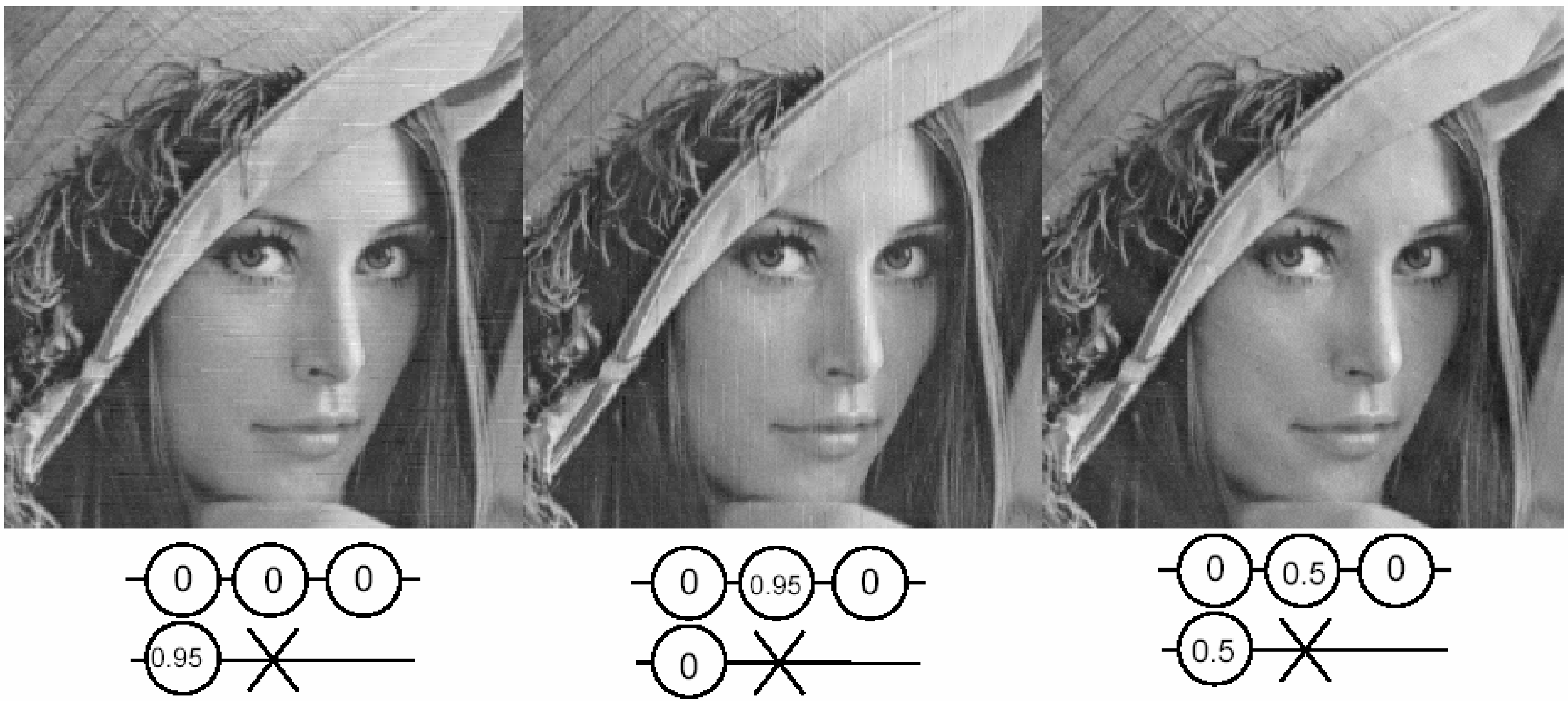




Transmission errors in a DPCM system

Example: Lena, 3 *bpp* (fixed code word length)

Error rate: $p=10^{-3}$



DPCM-Video

- Interframe coding exploits similarity of temporal successive pictures.

- Important interframe coding methods
  - Adaptive intra-interframe coding
  - Conditional replenishment
  - Motion-compensated prediction

[Second Edition.]

PATENT SPECIFICATION

Convention Date (United States): April 25, 1929.

341,811

Application Date (in United Kingdom): April 25, 1930. No. 12,805/30.

Complete Specification Accepted: Jan. 22, 1931.

COMPLETE SPECIFICATION.

Improvements relating to Electric Picture Transmission Systems.

We, THE BRITISH THOMSON-HOUSTON COMPANY LIMITED, a British Company, having its registered office at Crown House, Aldwych, London, W.C. 2, (Assignees of RAY DAVIS KELL, of 111, Sanders Avenue, Scotia, County of Schenectady, State of New York, United States of America, a citizen of the United States of America), do hereby declare the nature of this invention and in what manner the same is to be performed, to be particularly described and ascertained in and by the following fineness of detail is limited only by the speed of the action to be transmitted.

The invention will be better understood from the following description when considered in connection with the accompanying drawings in which Fig. 1 illustrates a picture transmitting apparatus wherein the invention has been embodied; and Figs. 2 to 5 illustrate various details of an apparatus which may be utilised to receive the difference between the successive images of a picture or moving object.

"It has been customary in the past to transmit successive complete images of the transmitted picture."

[...]

"In accordance with this invention, this difficulty is avoided by transmitting only the difference between successive images of the object."

Transform Coding

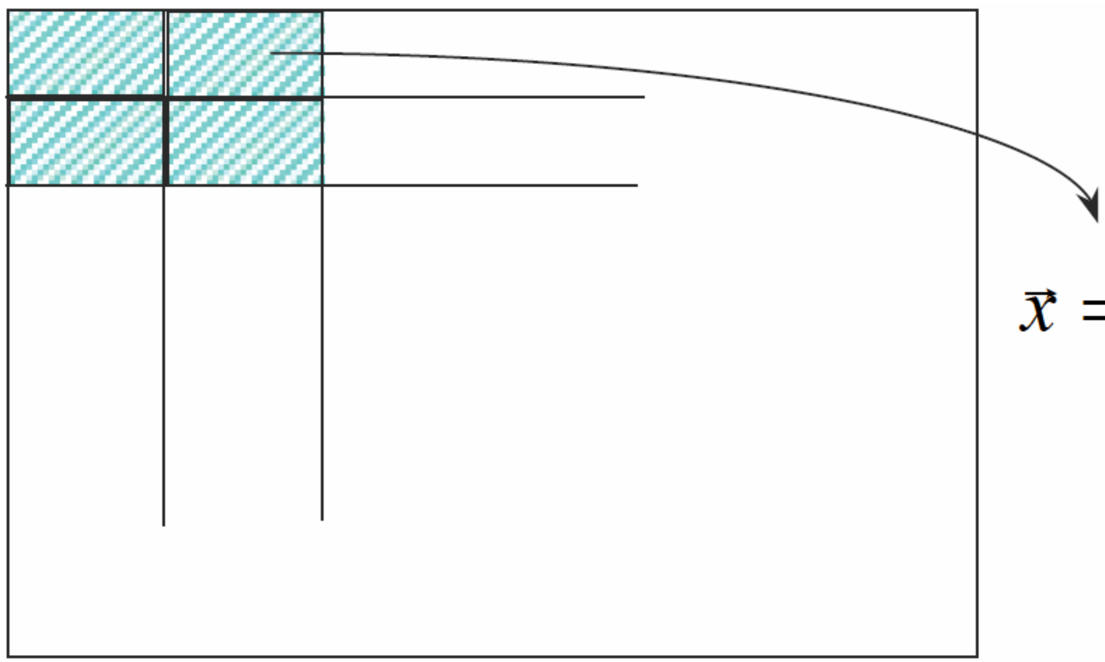
- Why transform Coding?
  - Purpose of transformation is to convert the data into a form where compression is easier.

- Transformation yields energy compaction (skewed probability distribution)
  - Facilitates reduction of irrelevant information.

- The transform coefficients can now be quantized according to their statistical properties.
  - This transformation will reduce the correlation between the pixels.)

How Transform Coders Work

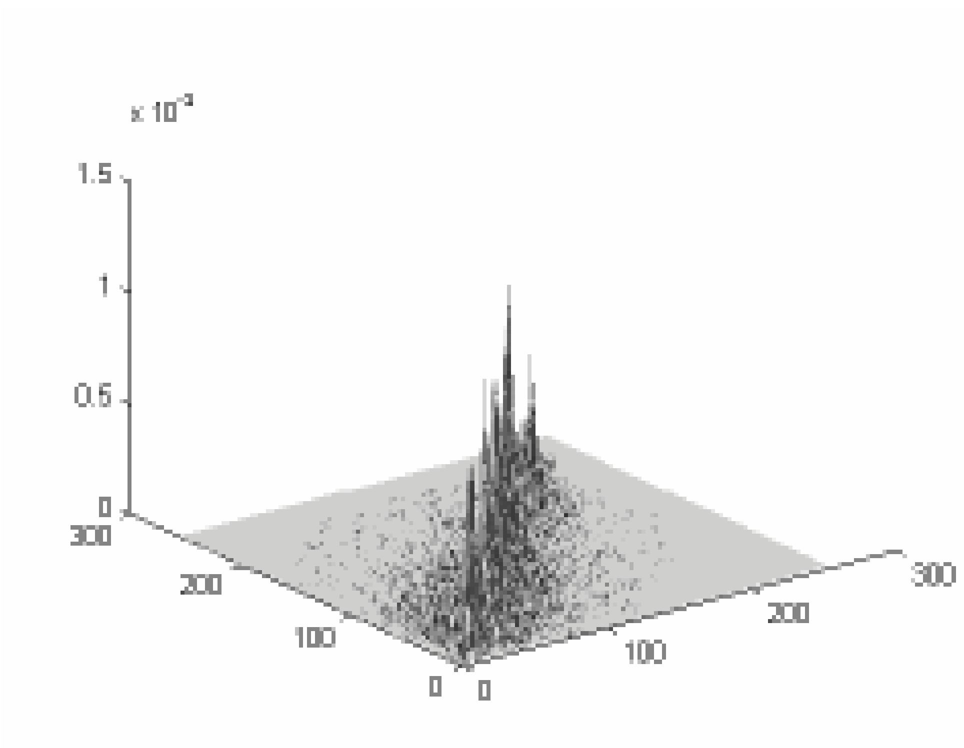
- Divide the image into 2x1 blocks
  - Typical transforms are 8x8 or 16x16



$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Joint Probability Distribution

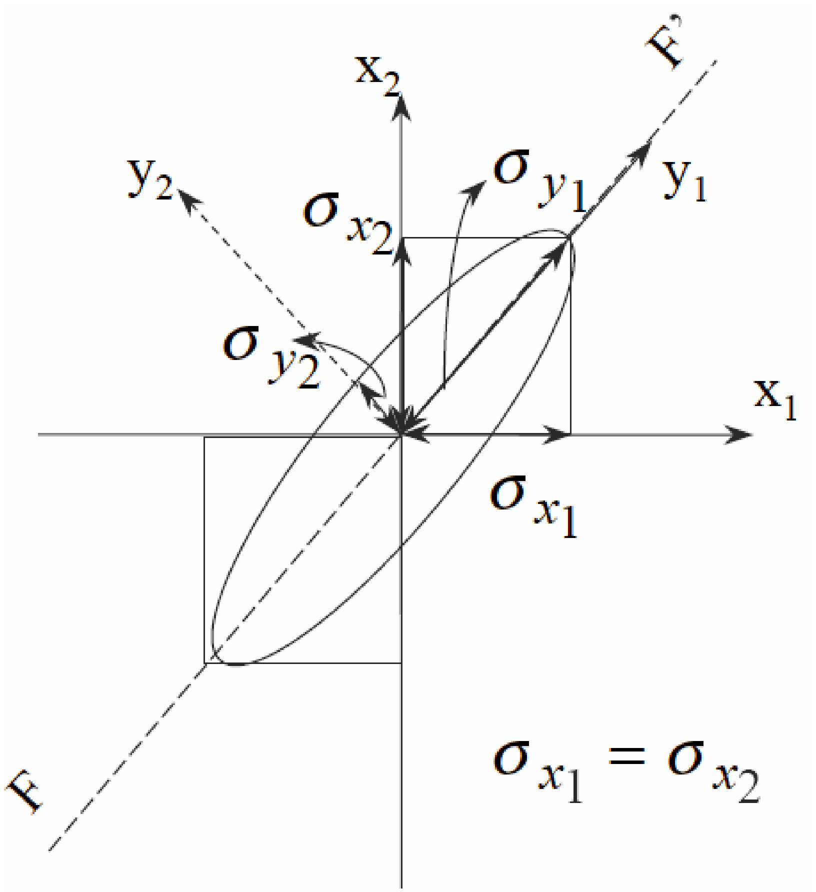
- Observe the Joint Probability Distribution or the Joint Histogram.



Correlated Pixels

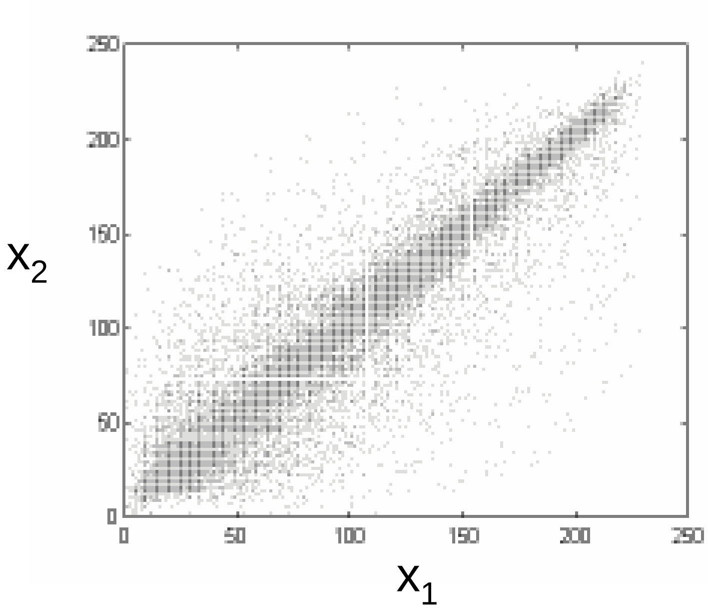
- Since adjacent pixels x_1 and x_2 are highly correlated the joint probability $p(x_1, x_2)$ is concentrated around the line FF' and variances are equal because they are the samples of the same image.

$$\sigma_{x_1} = \sigma_{x_2}$$



Pixel Correlation Map

- X_{ave1} =100, X_{var1} =2774
- X_{ave2} =100, X_{var2} =2761



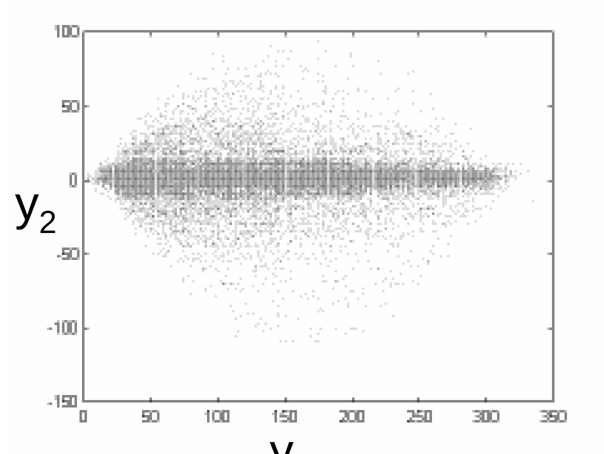
Rotated Pixels

- Rotated 45 degrees

  $$Y = \mathbf{A} X$$

- $y_{ave1} = 141, \; y_{var1} = 5365$
- $y_{ave2} = 0.13, \; y_{var2} = 170$

- Energy is packed to y_1. Can apply entropy coder on y_2.



y_2

y_1

- Rotation matrix

$$A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}_{\theta = 45^{O}} = \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ -\sqrt{2} & \sqrt{2} \end{bmatrix}$$

Reconstruct the Image

- Recover X from Y.
- Since Y = AX, so

  $X = A^{-1}Y$

- Reverse matrix

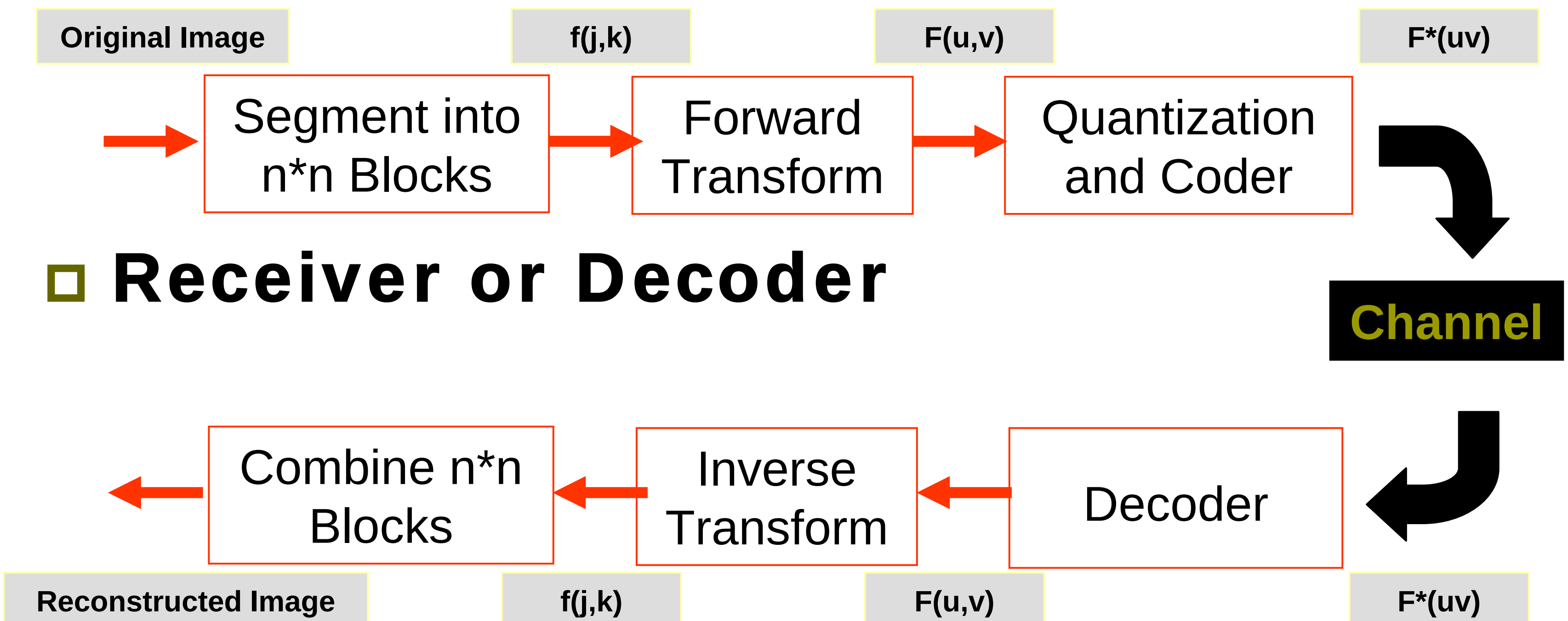
$$A^{-1} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}_{\theta=45^{O}} = \begin{bmatrix} \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & \sqrt{2} \end{bmatrix}$$

Idea of Transform Coding

- Transform the input pixels $Y_0, Y_1, Y_2, \ldots, Y_{N-1}$ into coefficients $Y_0, Y_1, \ldots, Y_{N-1}$ (real values)
  - The coefficients are have the property that most of them are near zero.
  - Most of the "energy" is compacted into a few coefficients.
- Scalar quantize the coefficient
  - This is bit allocation.
  - Important coefficients should have more quantization levels.
- Entropy encode the quantization symbols

Transform Coding Block Diagram

- **Transmitter or Encoder**
- **Receiver or Decoder**



Discrete Cosine Transform (DCT)

- For conventional image data having reasonably high inter-element correlation.

- Avoids the generation of the spurious spectral components which is a problem with DFT and has a fast implementation which avoids complex algebra.

$$X_k = \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}\left(n + \frac{1}{2}\right)k\right]$$

Basic Cosine Transform

- The basic Cosine Transform is:

$$F(u,v) = \frac{4C(u)C(v)}{n^2} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} f(j,k) \cos\left[\frac{(2j+1)u\pi}{2n}\right] \cos\left[\frac{(2k+1)v\pi}{2n}\right]$$

- where:

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

Inverse Cosine Transform

- The inverse cosine transform is:

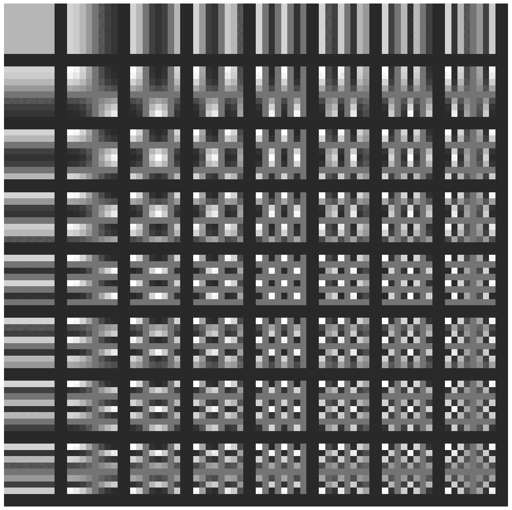
$$f(j,k) = \sum_{u=0}^{n-1}\sum_{v=0}^{n-1} C(u)C(v)F(u,v)\cos\left[\frac{(2j+1)u\pi}{2n}\right]\cos\left[\frac{(2k+1)v\pi}{2n}\right]$$

- where:

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

Fast Cosine Transform

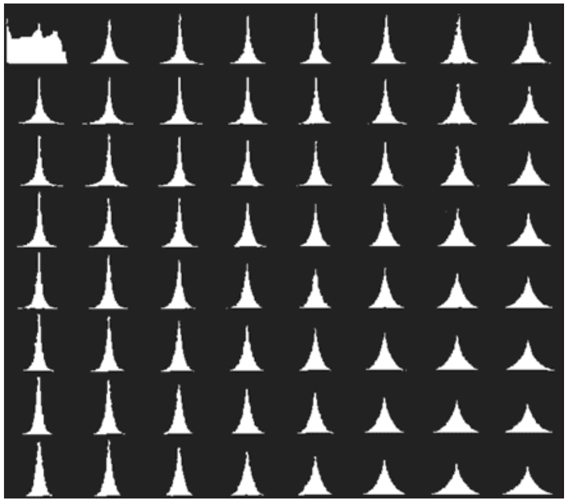
- 2D basis functions of the DCT:
  - Apply 1D horizontal DCT then 1D vertical DCT



- Fast algorithm for scaled 8-DCT
  - 5 multiplications
  - 29 additions.

Amplitude distribution of the DCT coefficients

- Histograms for 8x8 DCT coefficient amplitudes measured for natural images
  - DC coefficient is typically uniformly distributed.
  - The distribution of the AC coefficients have a Laplacian distribution with zero-mean.



Importance of Coefficients

- The DC coefficient is the most important.
- The AC coefficients become less important as they are farther from the DC coefficient.
- Example bit allocation table:

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

Karhunen Loève Transform (KLT)

- Karhunen Loève Transform (KLT) yields decorrelated transform coefficients.
- Basis functions are eigenvectors of the covariance matrix of the input signal.
- KLT achieves <u>optimum</u> energy concentration.
- Disadvantages:
  - KLT dependent on signal statistics
  - KLT not separable for image blocks
  - Transform matrix cannot be factored into sparse matrices

Other Transforms

- Wavelets
- DFT
- Remember BWT?

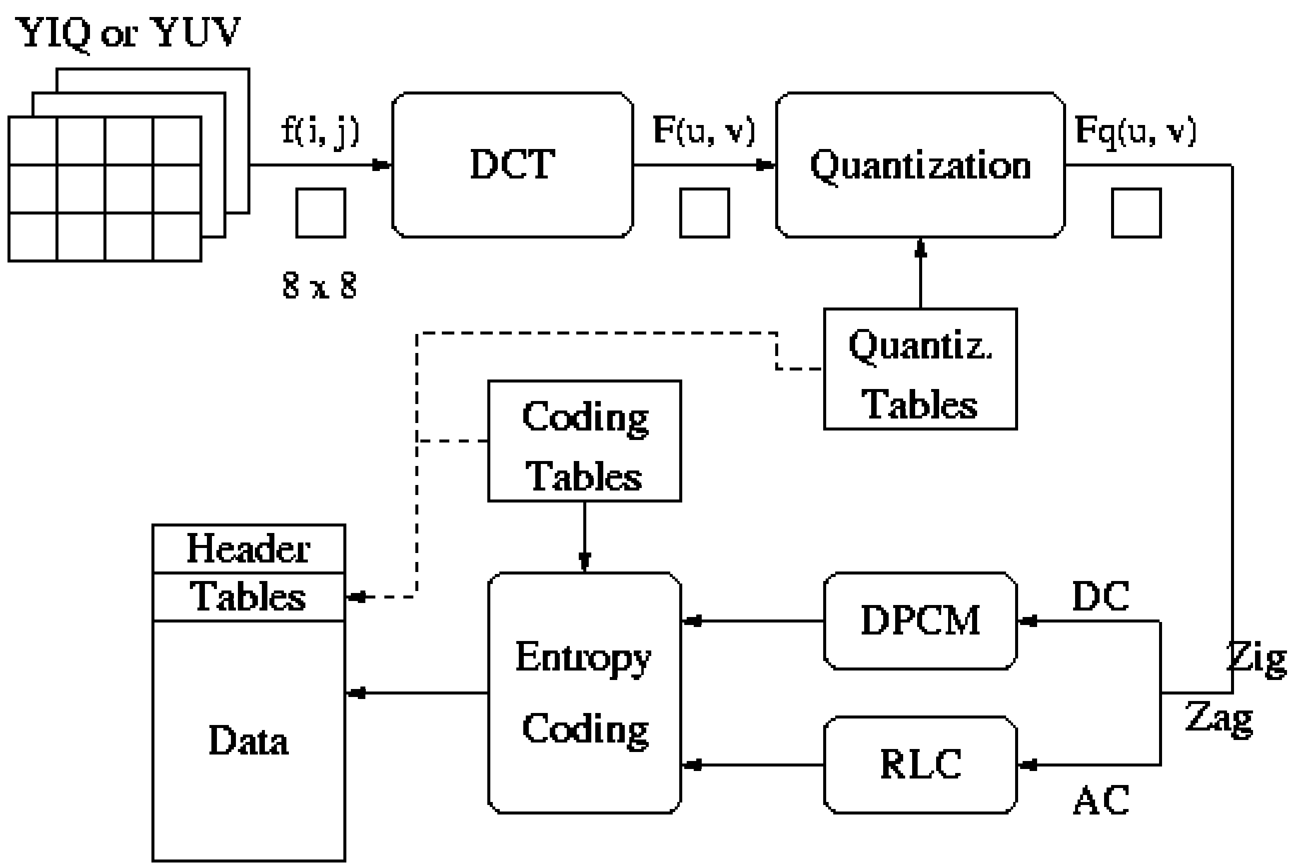
JPEG

- JPEG stands for Joint Photographic Expert Group
- A standard image compression method is needed to enable interoperability of equipment from different manufacturer
- It is the first international digital image compression standard for continuous-tone images (grayscale or color)

JPEG

- “very good” or “excellent” compression rate, reconstructed image quality, transmission rate
- be applicable to practically any kind of continuous-tone digital source image
- good complexity
- have the following modes of operations:
  - sequential encoding
  - progressive encoding
  - lossless encoding
  - hierarchical encoding

JPEG Encoder Block Diagram



JPEG Main operations

- **Quantization**
  - Different quantization level for different DC and AC coefficients
- **Zig-zag Scan**
  - Why? -- to group low frequency coefficients in top of vector.
  - Maps 8 x 8 to a 1 x 64 vector
- **Differential Pulse Code Modulation (DPCM) on DC component**
  - DC component is large and varied, but often close to previous value.
  - Encode the difference from previous 8 x 8 blocks -- DPCM
- **Run Length Encode (RLE) on AC components**
  - 1 x 64 vector has lots of zeros in it
  - Keeps *skip* and *value*, where *skip* is the number of zeros and *value* is the next non-zero component.
  - Send (0,0) as end-of-block sentinel value.
- **Entropy Coding (Huffman, arithmetic, …)**
  - Differential DC and run-length AC values

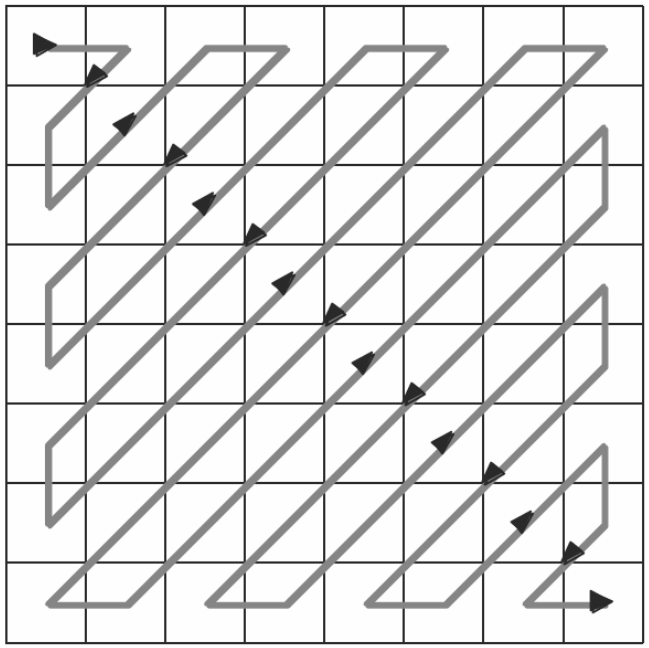
DCT Quantization Table (Luminance)

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	33	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

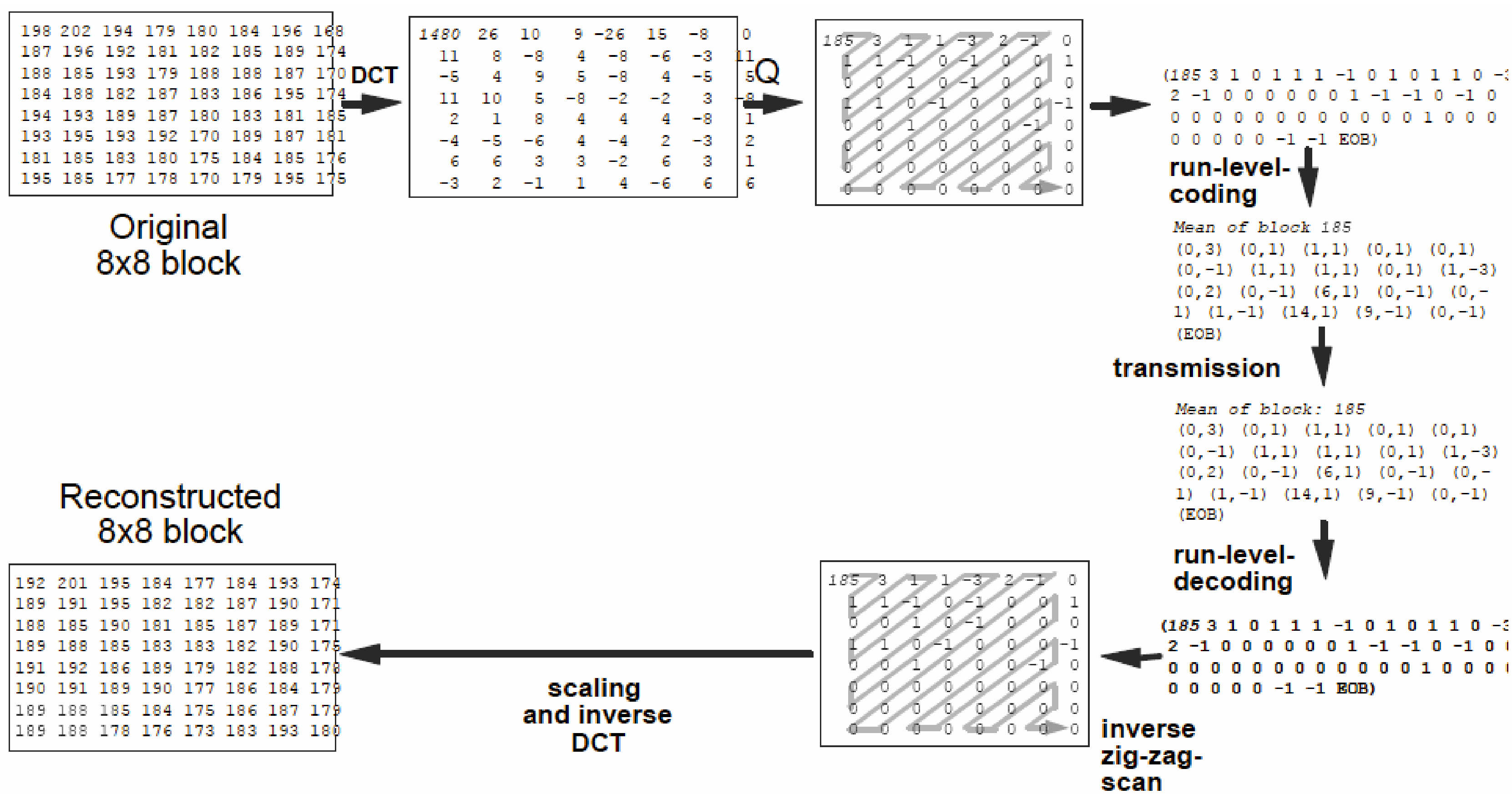
$F'[u, v] = \text{round}(F[u, v] / q[u, v])$.

DCT Zig-Zag-Scan

- The variances of the DCT transform coefficients are decreasing in a zig-zag manner approximately.
- zig-zag-scan + run-level-coding



Example

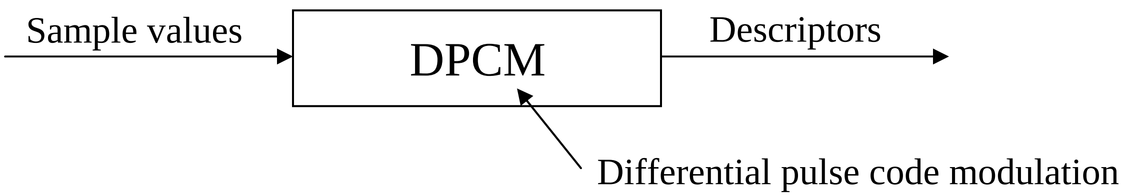


DCT
Q
Original
8x8 block
run-level-
coding
Mean of block 185
(0,3) (0,1) (1,1) (0,1) (0,1)
(0,-1) (1,1) (1,1) (0,1) (1,-3)
(0,2) (0,-1) (6,1) (0,-1) (0,-
1) (1,-1) (14,1) (9,-1) (0,-1)
(EOB)
transmission
Mean of block: 185
(0,3) (0,1) (1,1) (0,1) (0,1)
(0,-1) (1,1) (1,1) (0,1) (1,-3)
(0,2) (0,-1) (6,1) (0,-1) (0,-
1) (1,-1) (14,1) (9,-1) (0,-1)
(EOB)
run-level-
decoding
inverse
zig-zag-
scan
Reconstructed
8x8 block
scaling
and inverse
DCT

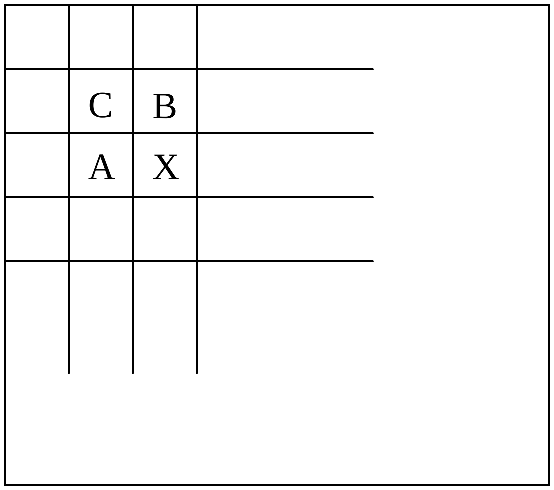
JPEG Progressive Model

- Why progressive model?
  - Quick transmission
  - Image built up in a coarse-to-fine passes
- First stage: encode a rough but recognizable version of the image
- Later stage(s): the image refined by successive scans till get the final image
- Two ways to do this:
  - Spectral selection – send DC, AC coefficients separately
  - Successive approximation – send the most significant bits first and then the least significant bits

JPEG Lossless Model



Predictors for lossless coding	
selection value	prediction strategy
0	no prediction
1	A
2	B
3	C
4	A+B-C
5	A+(B-C)/2
6	B+(A-C)/2
7	(A+B)/2



JPEG Hierarchical Model

- Hierarchical model is an alternative of progressive model (pyramid)
- Steps:
  - filter and down-sample the original images by the desired number of multiplies of 2 in each dimension
  - Encode the reduced-size image using one of the above coding model
  - Use the up-sampled image as a prediction of the origin at this resolution, encode the difference
  - Repeat till the full resolution image has been encode

JPEG-2000

- JPEG 2000 is the upcoming standard for Still Pictures
- Major change from the current JPEG is that wavelets will replace DCT as the means of transform coding.
- Among many things it will address:
  - Low bit-rate compression performance,
  - Lossless and lossy compression in a single codestream,
  - Transmission in noisy environment where bit-error is high,
  - Application to both gray/color images and bi-level (text) imagery, natural imagery and computer generated imagery,
  - Interface with MPEG-4,
  - Content-based description.