18-796

Multimedia Communications:
Coding, Systems, and Networking

*Prof. Tsuhan Chen*
tsuhan@ece.cmu.edu

# JPEG

# JPEG

- Joint Photographic Experts Group
  - ISO/IEC JTC1 SC29 WG1
  - Formed in 1986 by ISO and CCITT (ITU-T)
  - Became International Standard (IS) in 1991
  - Compression ratio 10 to 50; 0.5 to 2 bpp. At 1 bpp, one 256×256 image takes only 2 sec at 33.6 kbits/s
  - Digital Compression and Coding of Continuous-Tone Still Images (grayscale or color)
    - ISO/IEC IS 10918-1 (ITU-T T.81): Requirements and guidelines
    - ISO/IEC IS 10918-2 (ITU-T T.83): Compliance testing
    - ISO/IEC IS 10918-3 (ITU-T T.84): Extensions



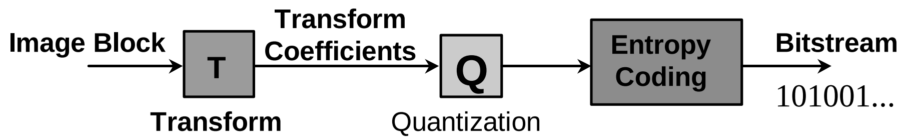
# Picture Formats

- Up to 65535 lines and 65535 pels/line
- 8 or 12 bits precision
- Color-space independent
  - Up to 255 color components
  - Each component can be subsampled
  - Interleaving
  - To save bits: YUV is better than RGB
- Typical picture sizes

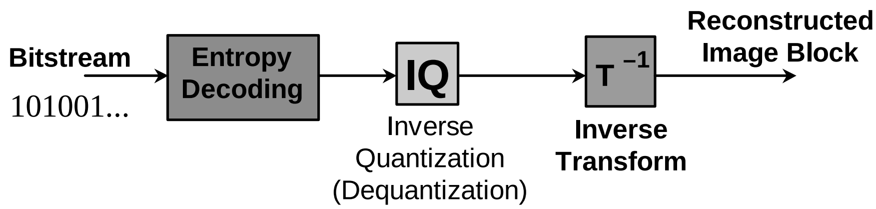
| | CGA ~SIF | VGA ~CCIR601 | SVGA ~HDTV |
|---|---|---|---|
| Pels/line | 320 | 640 | ~1280 |
| Lines | 240 | 480 | ~960 |

# Basic Framework

## Encoder




## Decoder






# Linear Transform

- De-correlate pixels
- Allow the most efficient representation
  - Energy concentration
  - Removal or heavy quantization of some coefficients
- Allow perceptually weighted quantization
- Easy for entropy coding
- Karhunen-Loeve Transform (KLT) is optimal
- Discrete Cosine Transform (DCT)
  - Close to KLT for typical images
  - Widely used in JPEG, H.26x, MPEG

# 2D Discrete Cosine Transform

$$\begin{bmatrix} Y_{mn} \end{bmatrix} = \begin{bmatrix} C_{mn} \end{bmatrix}^T \begin{bmatrix} X_{mn} \end{bmatrix} \begin{bmatrix} C_{mn} \end{bmatrix}$$

Transform Coefficients — Image Block

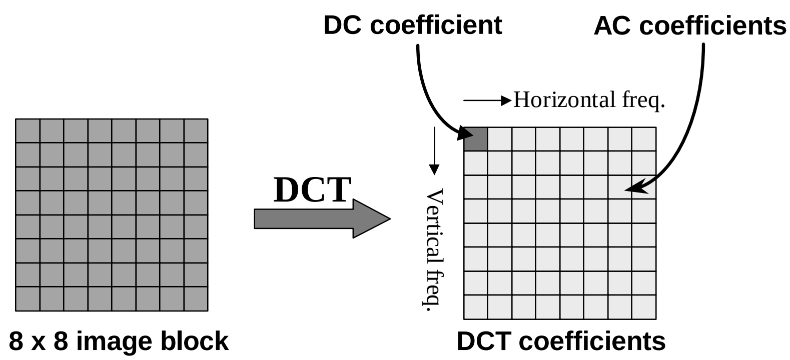
- For 8×8 blocks

$$C_{mn} = k_n \cos\left[\frac{(2m+1)np}{16}\right] \text{ where } k_n = \begin{cases} 1/(2\sqrt{2}) & \text{when } n = 0 \\ 1/2 & \text{otherwise} \end{cases}$$

- Inverse DCT

18-796/Spring 1999/Chen

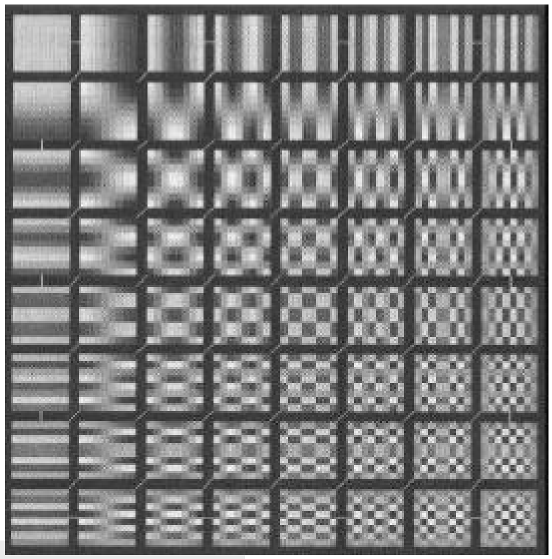
# DC and AC Coefficients




(Zero-shift to [-128,127])

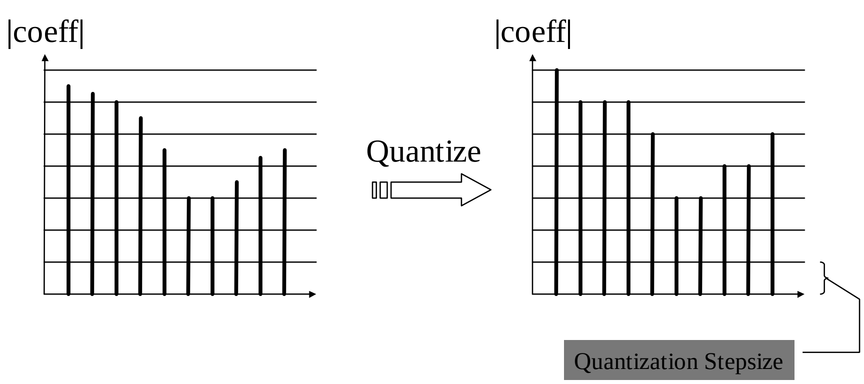
18-796/Spring 1999/Chen

# DCT Basis Functions



64 8×8 blocks

# Quantization



# Quantization (cont.)

- 8×8 quantization table Q[u,v]
  - High-freq coefficients can be quantized more
  - Color components can be quantized more
  - q-factor (in some implementation)
    - A scale factor applied to a fixed Q
- Q: $$FQ_{uv} = round\left(\frac{c_{uv}}{Q[u,v]}\right)$$
- IQ: $$\tilde{c}_{uv} = FQ_{uv} \times Q[u,v]$$



# Example Quantization Tables

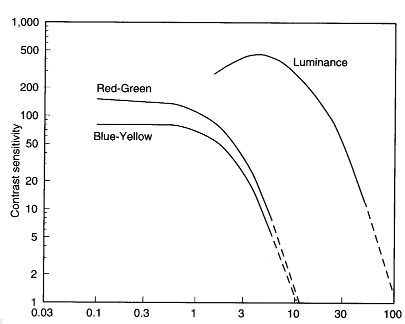
- Luminance

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 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 | 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 |

- Chrominance

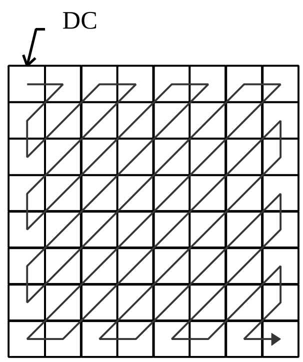
| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 17 | 18 | 24 | 47 | 99 | 99 | 99 | 99 |
| 18 | 21 | 26 | 66 | 99 | 99 | 99 | 99 |
| 24 | 26 | 56 | 99 | 99 | 99 | 99 | 99 |
| 47 | 66 | 99 | 99 | 99 | 99 | 99 | 99 |
| 99 | 99 | 99 | 99 | 99 | 99 | 99 | 99 |
| 99 | 99 | 99 | 99 | 99 | 99 | 99 | 99 |
| 99 | 99 | 99 | 99 | 99 | 99 | 99 | 99 |
| 99 | 99 | 99 | 99 | 99 | 99 | 99 | 99 |

Eye Sensitivity



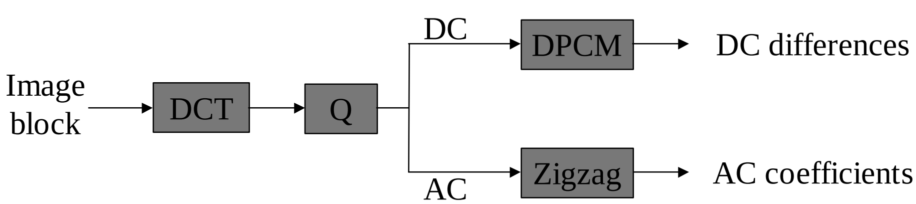
# Zigzag Scan

- Convert 2-D coefficients block to 1-D coefficients
- To generate long runs of zeros

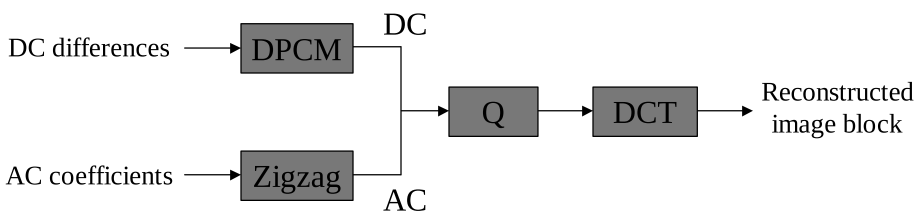





# JPEG Encoder



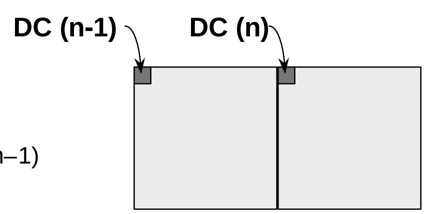
# JPEG Decoder






# DC Coding

- DC Prediction




Diff(n) = DC(n) – DC(n–1)

- Each Diff(n) is coded as
  - Size: in VLC, indicating the size of the following VLI
  - Amplitude: in VLI (variable length integer)

| Size | Amplitude |
|---|---|
| 0 | 0 |
| 1 | -1,1 |
| 2 | -3,-2,2,3 |
| 3 | -7,-6,-5,-4,4,5,6,7 |

⋮

# AC Coding

- Use VLC to code Run-Size
  - Run
    - The number of zeros before a nonzero coefficient
  - Size
    - The size of the following VLI
- Amplitude
  - The amplitude of the nonzero AC coefficient
  - coded in VLI
- EOB: end-of-block
- ZRL: zero-run-length, a run of 16 zeros



# Entropy

- Entropy
  - Uncertainty of a signal source X
  - Bits needed to resolve uncertainty

$$x(n) \in \{a_1, a_2, \cdots, a_K\}$$

$$\text{Probability}: \quad p_1, p_2, \cdots, p_K \qquad \sum_k p_k = 1$$

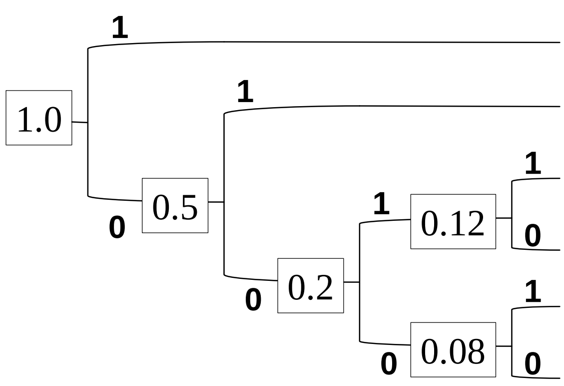
$$\text{Entropy}: \quad H(X) = -\sum_{k=1}^{K} p_k \log p_k \quad \text{(bits/symbol)}$$

# Entropy Coding

- Huffman coding
  - Variable length coding (VLC)
  - Short codewords for frequent symbols
  - For JPEG: the encoder can transmit the tables
- Arithmetic coding
  - Non-integer length coding
  - Probability distribution can be derived in real time
  - Usually more efficient than Huffman coding
    - For JPEG test images, 10% or more better



# Huffman Coding



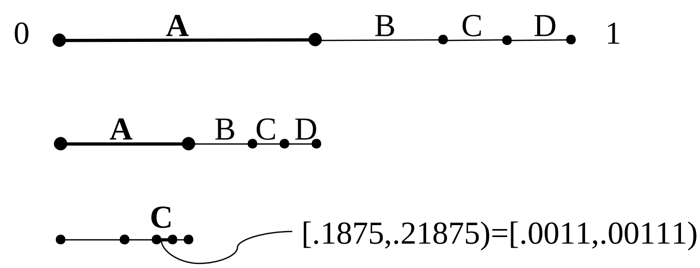

| Prob. | Symbols | |
|---|---|---|
| 0.5 | A | **1** |
| 0.3 | B | **01** |
| 0.07 | C | **0011** |
| 0.05 | D | **0010** |
| 0.05 | E | **0001** |
| 0.03 | F | **0000** |

Uniquely decodable, e.g., **0100101010101001100101 00…**

18-796/Spring 1999/Chen

# Arithmetic Coding

$p_A = 0.5 \quad p_B = 0.25 \quad p_C = 0.125 \quad p_D = 0.125$




AAC… ⇒ 00110...

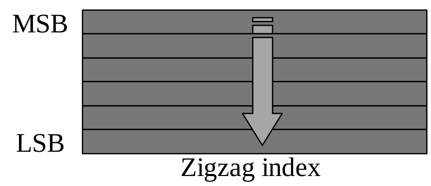


# Modes of Operation

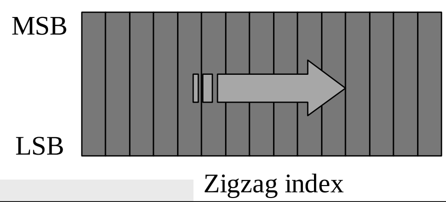
- Sequential DCT-based
  - A subset of sequential mode is the **Baseline Mode**
    - 8 bits/pel
    - Huffman coding: 2 DC and 2 AC tables
  - Extended
    - 8 or 12 bits/pel
    - Huffman coding: 4 DC and 4 AC tables

## Modes of Operation (cont.)

- Progressive DCT-based
  - Successive approximation: MSB first, LSB later



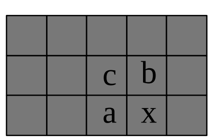

  - Spectral selection: Low-freq first, high-freq later






## Modes of Operation (cont.)

- Sequential lossless
  - Compression factor only 2 to 3




| | Prediction |
|---|---|
| 0 | No |
| 1 | a |
| 2 | b |
| 3 | c |
| 4 | a+b-c |
| 5 | a-(b-c)/2 |
| 6 | b-(a-c)/2 |
| 7 | (a+b)/2 |

- Hierarchical
  - Pyramid coding

18-796/Spring 1999/Chen

# JPEG Picture Quality

- Quality vs. bit rates

| | CGA | VGA | SVGA |
|---|---|---|---|
| 0.5 bits/pel | poor | fair | good |
| 1.0 bits/pel | fair | good | excellent |
| 2.0 bits/pel | good | excellent | excellent+ |

- Perceptually lossless at 1.5-2 bpp



# Extension to Video

- Motion JPEG (M-JPEG)
- Compared to MPEG, M-JPEG has
  - No error propagation
  - Random access
  - Low complexity
  - But the compression ratio is low

# JPEG 2000

- Image Coding System (JTC 1.29.14, ISO 15444)
- Goals
  - – Low bit-rate compression
    - e.g., below 0.25 bpp for highly detailed gray-level images
  - – Lossless and lossy compression in a single bitstream
  - – Large images
    - More than 64K by 64K
  - – Single decompression architecture
  - – Transmission in noisy environments
  - – Computer generated imagery
  - – Compound documents: bi-level and gray-scale



# Criteria of JPEG 2000

- Superior low bit-rate performance
- Continuous-tone and bi-level compression
- Lossless and lossy compression
- Progressive transmission by pixel accuracy and resolution
- Fixed-rate, fixed-size, limited workspace memory
- Random bitstream access and processing
- Robustness to bit errors
- Open architecture
- Sequential build-up capability (real time coding)
- Backwards compatibility with JPEG
- Content-based description
- Protective image security
- Interface with MPEG-4
- Side channel spatial information (transparency)

# Applications

- Low bandwidth dissemination of imagery
- Medical imagery lossless/lossy compression
- Pre-press imagery
- Client/server applications (World Wide Web)
- Electronic photography
- Photo and art digital libraries
- Security
- Facsimile
- Laser print rendering
- Scanner and digital copier memory buffers



| | | | | | | | | Overall System | | | | | | | | | | Decoder Spec | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 5.1 | 5.2 | 5.3 | 5.4 | 5.5 | 5.6 | 5.7 | 5.8 | 5.9 | 5.10 | 5.11 | 5.12 | 5.13 | 5.14 | 5.15 | 5.16 | 5.17 | 5.18 |
| | Application Profile | Image Type | Uncompressed | Lossless Compression | Visually Lossless Compression | Visually Lossy Compression | Progressive Spatial | Progressive Quality | Security | Error Resilience | Complexity Scaleability | Strip Processing | Sensor Specific Compression Flexibility | Information Embedding | Repetitive Encoding/Decoding | Object-Based Functionality | MPEG4 VTC Compatability | Backward Compatability Client Side Ease of Transcoding | ROI Decoding |
| 6.1 | Internet | M(1,3) O(4+) | | | M | M | M | M | O | O | | | | M | | M | O | M(Baseline) O(non-baseline) | O |
| 6.2 | Color Fascimile | M | | O | M | | | | | O | | M | | | M | | | O | |
| 6.3 | Printing | M | | | M | | | | | | | | | | | | | O | |
| 6.4 | Scanning (Consumer, pre-press) | M | O | M | M | | | | M | | | | | | | | | | |
| 6.5 | Digital Photography | M | | | M | O | O | O | | O | | | | M | | | O | O | |
| 6.6 | Remote Sensing | M(1,3) O(4+) | | M | M | M | M | O | | O | | M | M | O | | O | | O | O |
| 6.7 | Mobile | M(1,3) O(4+) | | | M | M | O | M | O | M | O | | | | | O | O | O | O |
| 6.8 | Medical | M(1,3) O(4+) | | M | M | M | M | M | O | O | | M | M | M | O | O | | M(Baseline) O(non-baseline) | O |
| 6.9 | Digital Library | M(1,3) O(4+) | | | M | M | M | M | O | O | | | | M | | M | | M(Baseline) O(non-baseline) | O |
| 6.1 | E-Commerce | M(1,3) O(4+) | | | M | M | M | M | M | O | | | | M | | M | | M(Baseline) O(non-baseline) | O |
| | **Legend:** M - Mandatory O - Optional | | | | | | | | | | | | | | | | | | |

# Schedule

- Call for contributions: Mar 97
- Submission of architectural frameworks: June 97
- Presentation of architectural frameworks: July 97, Japan
- Submission of algorithms: September 97
- Presentation of algorithms: November 97, Australia
- Working Draft (WD): Mar 98
- Committee Draft (CD): Jul 98
- DIS: Mar 2000
- IS: Nov 2000

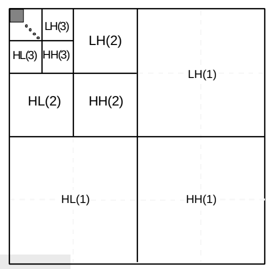
24 proposals, 35 test images



# Sample Test Image

## Popular Wavelet Codecs

- Embeded zerotree wavelet (EZW) [Shapiro'93]
- Layer zero coding (LZC) [Taubman and Zakhor'94]
- Set partition in hierarchical tress (SPHIT) [Said and Pearlman'96]
- Multi-thresholds wavelet codec (MTWC) [Kuo'97]






## References

- JPEG
  - William B. Pennebaker, Joan L. Mitchell, *JPEG: Still Image Data Compression Standard*, Van Nostrand Reinhold, New York, NY, 1993
- Arithmetic coding
  - Cleary, Witten, and Neal, "Arithmetic coding for data compression, *Comm. of ACM*, June 1987