18-899 Special Topics in Signal Processing

Multimedia Communications:
Coding, Systems, and Networking

Prof. Tsuhan Chen
tsuhan@ece.cmu.edu

Lecture 3

JPEG

JPEG

- Joint Photographic Experts Group
  - ISO/IEC JTC1 SC29 WG1
  - Formed in 1986 by ISO and CCITT (ITU-T)
  - Digital Compression and Coding of Continuous-Tone Still Images (grayscale or color)
  - Compression ratio 10 to 50; 0.5 to 2 bpp
  - At 1 bpp, one 256×256 image takes 2 sec at 33.6 kbits/s

JPEG (cont.)

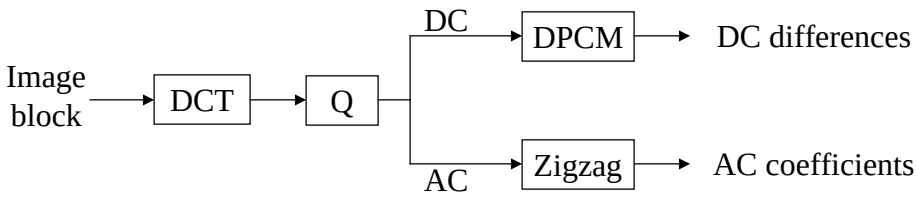
- 1988: Agreed on DCT
- 1988-1991: Refinements
  - Including lossless mode
- 1991: became International Standard (IS)
  - ISO/IEC IS 10918-1 (ITU-T T.81): Requirements and guidelines
  - ISO/IEC IS 10918-2 (ITU-T T.83): Compliance testing
  - ISO/IEC IS 10918-3 (ITU-T T.84): Extensions

Picture Formats

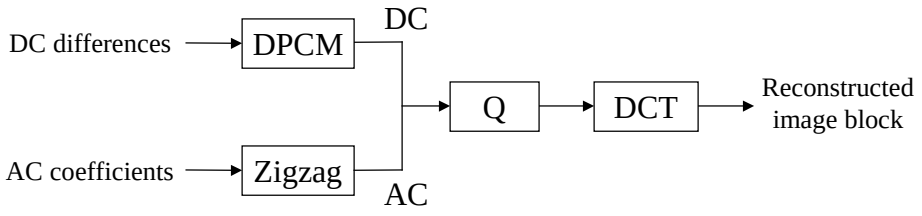
- Up to 65535 lines and 65535 pels/line
- 8 or 12 bits precision
- Color-space independent
  - Up to 255 color components
  - Each component can be subsampled
  - Interleaving
  - To save bits: YUV is better than RGB
- Typical picture sizes

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

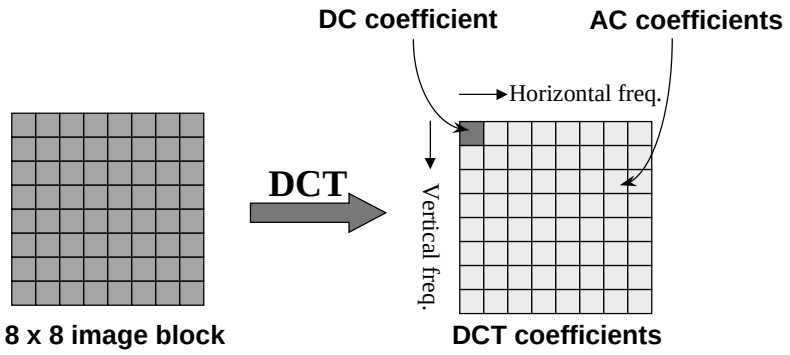
JPEG Encoder



JPEG Decoder

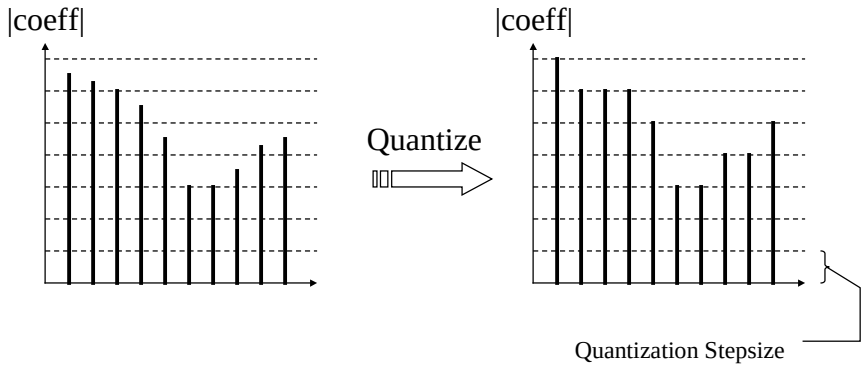


8 × 8 DCT



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

Quantization



Quantization (cont.)

- 8×8 quantization table Q[u,v]
  - Higher-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

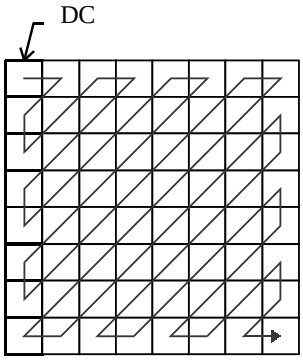
16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

- 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

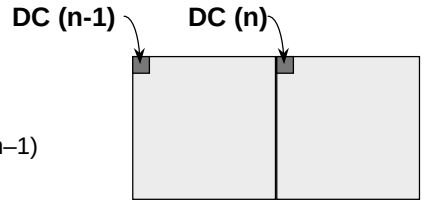
Zigzag Scan

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



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

- Using 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

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

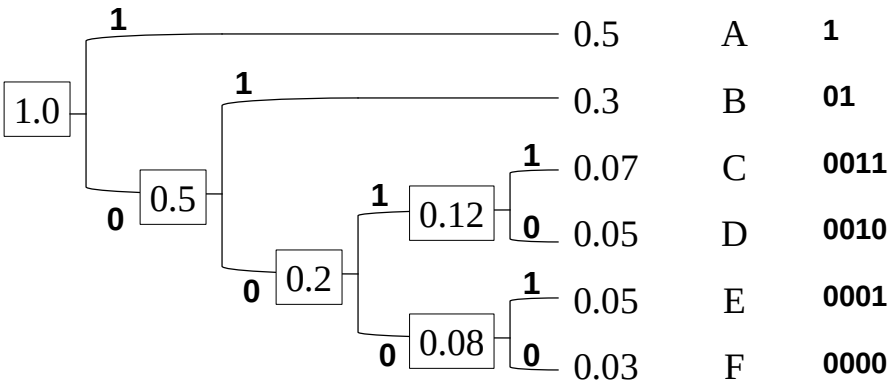
$$\text{Probability}: \quad p_1, p_2, \cdots, p_K \qquad \sum_k p_k = 1$$

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

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

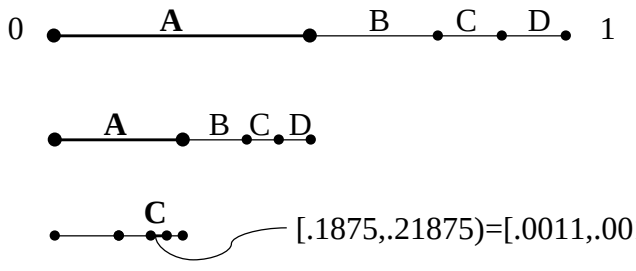
Huffman Coding



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

Arithmetic Coding

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



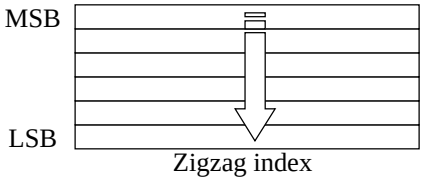
AAC… ⇨ 00110...

Modes of Operation

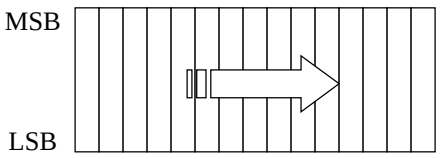
- Sequential DCT-based
  - A subset of sequential mode is the **Baseline Mode**
    - 8 bits/pel
    - Huffman coding: 2 DC and 2 AC tables
  - Extended
    - 8 or 12 bits/pel
    - Huffman coding: 4 DC and 4 AC tables

Modes of Operation (cont.)

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

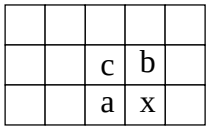


  - Spectral selection: Low-freq first, high-freq later



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

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

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