Research Institute for Future Media Computing
未来媒体技术与研究所

Institute of Computer Vision
计算机视觉研究所

图像压缩标准
Image Compression Standards

知识点回顾

◆ Distortion Measures

– *mean square error* (MSE,均方差) σ^2,

$$\sigma^2 = \frac{1}{N}\sum_{n=1}^{N}(x_n - y_n)^2$$

where x_n, y_n, and N are the input data sequence, reconstructed data sequence, and length of the data sequence respectively.

– *signal to noise ratio* (SNR,信号噪声比), in decibel units (dB),

$$SNR = 10\log_{10}\frac{\sigma_x^2}{\sigma_d^2}$$

σ_d^2 is the MSE.

均方 $$\sigma_x^2 = \frac{x_1^2 + x_2^2 + \cdots + x_N^2}{N}$$

σ_x^2 is the average square value of the original data sequence

– *peak signal to noise ratio* (PSNR, 峰值信噪比),

$$PSNR = 10\log_{10}\frac{x_{peak}^2}{\sigma_d^2}$$

1D DCT and IDCT with 8 numbers

- Consider a data sequence with 8 numbers
- 1D DCT

$$F(u) = \frac{C(u)}{2} \sum_{i=0}^{7} \cos \frac{(2i+1)u\pi}{16} f(i)$$

- 1D IDCT

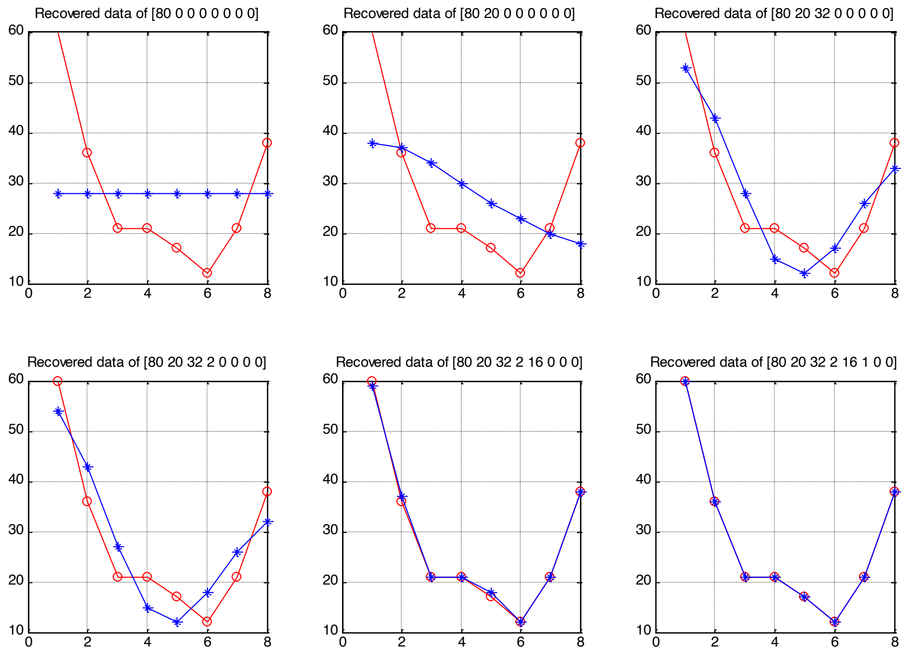
$$\tilde{f}(i) = \sum_{u=0}^{7} \frac{C(u)}{2} \cos \frac{(2i+1)u\pi}{16} F(u)$$

- The constants C(u) are determined by

$$C(\xi) = \begin{cases} \frac{\sqrt{2}}{2} & if \quad \xi = 0, \\ 1 & otherwise. \end{cases}$$

Example of Compression Scheme for 1D DCT

- Original data X: { 60 36 21 21 17 12 21 38 }
- Transform coding T: 1D DCT
- Transformed vector Y: 80 20 32 2 16 1 0 0



2D DCT and 2D IDCT

- Consider an image block with 8 ×8 numbers
- 2D DCT

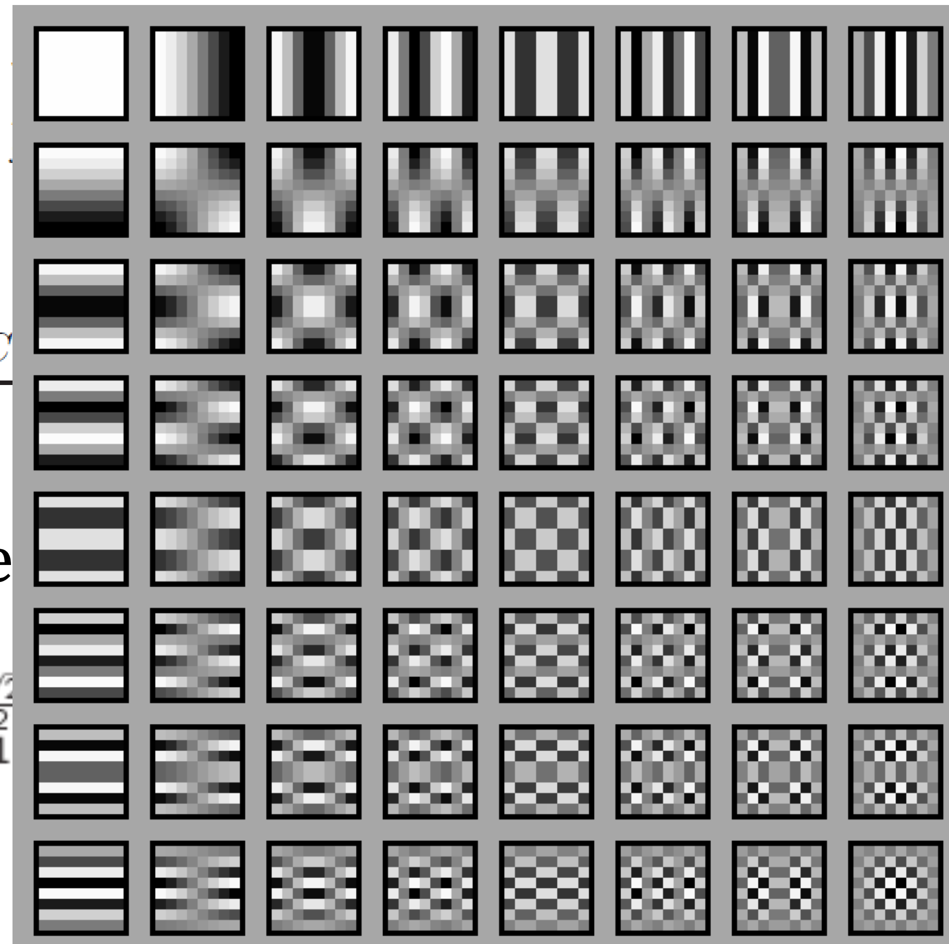
$$F(u,v) = \frac{C(u)\,C(v)}{4} \sum_{i=0}^{7}$$

- 2D IDCT

$$\tilde{f}(i,j) = \sum_{u=0}^{7} \sum_{v=0}^{7} \frac{C(u)\,C}{4}$$

- The constants C(u) are

$$C(\xi) = \begin{cases} \frac{\sqrt{2}}{2} \\ 1 \end{cases}$$



Outline of Lecture 09

- Image compression standard JPEG
- Block diagram for JPEG encode
  - Transform RGB color model to YUV or YIQ
  - DCT coding for image block
  - Uniform scalar quantization
  - ZigZag scan for quantized DCT coefficients
    - DPCM for DC coefficients
    - RLC for AC coefficients
  - Entropy coding
  - JPEG file generation

Image Compression Standard JPEG

- JPEG is an image compression standard that was developed by the "Joint Photographic Experts Group" （联合图像专家组）
- JPEG was formally accepted as an international standard in 1992
- JPEG combines several lossless and lossy compression techniques
  - Transform coding
  - Quantization
  - DPCM （差分脉冲编码调制）
  - Run-length coding
  - Entropy coding

Image Compression Standard JPEG

- The effectiveness of the coding method in JPEG relies on 3 major observations:

Observation 1: Useful image contents change relatively slowly across the image, i.e., it is unus
times in a small area, for

- much of the information i
redundancy".
- Use DCT transform

Image Compression Standard JPEG

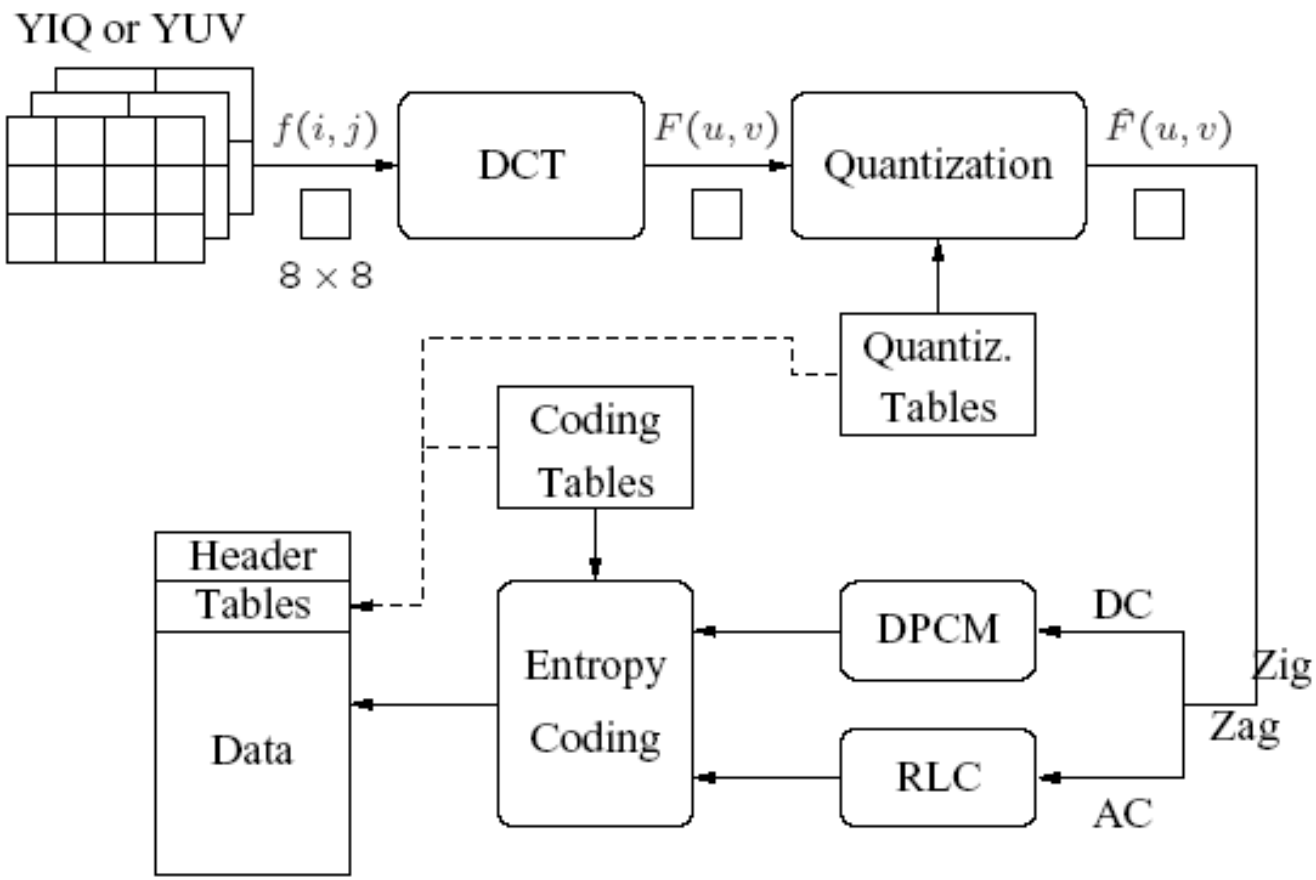
◆ **Observation 2**: Psychophysical（心理物理学）experiments suggest that humans are much less likely to notice the loss of very high spatial frequency components than the loss of lower frequency components.

- The spatial redundancy can be reduced by largely reducing the high spatial frequency contents.

◆ **Observation 3**: Visual acuity (敏锐度) is much greater for gray (“black and white”) than for color.

- chroma subsampling (4:2:0) is used in JPEG.

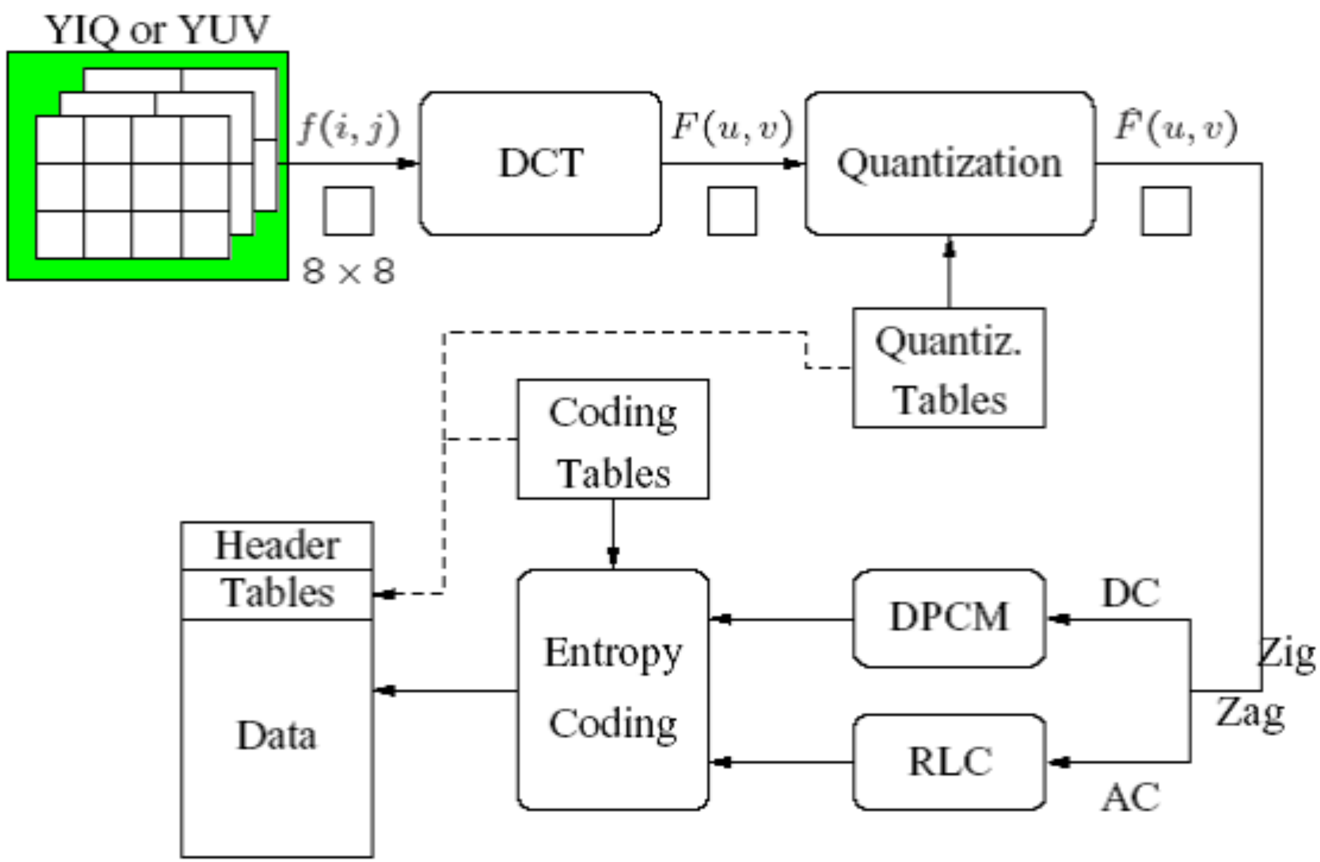
Block Diagram for JPEG Encode



Block Diagram for JPEG Encode

- Main Steps in JPEG Image Compression
  - Transform RGB to YIQ or YUV and subsample color.
  - DCT on image blocks.
  - Quantization.
  - Zig-zag ordering and run-length encoding.
  - Entropy coding.

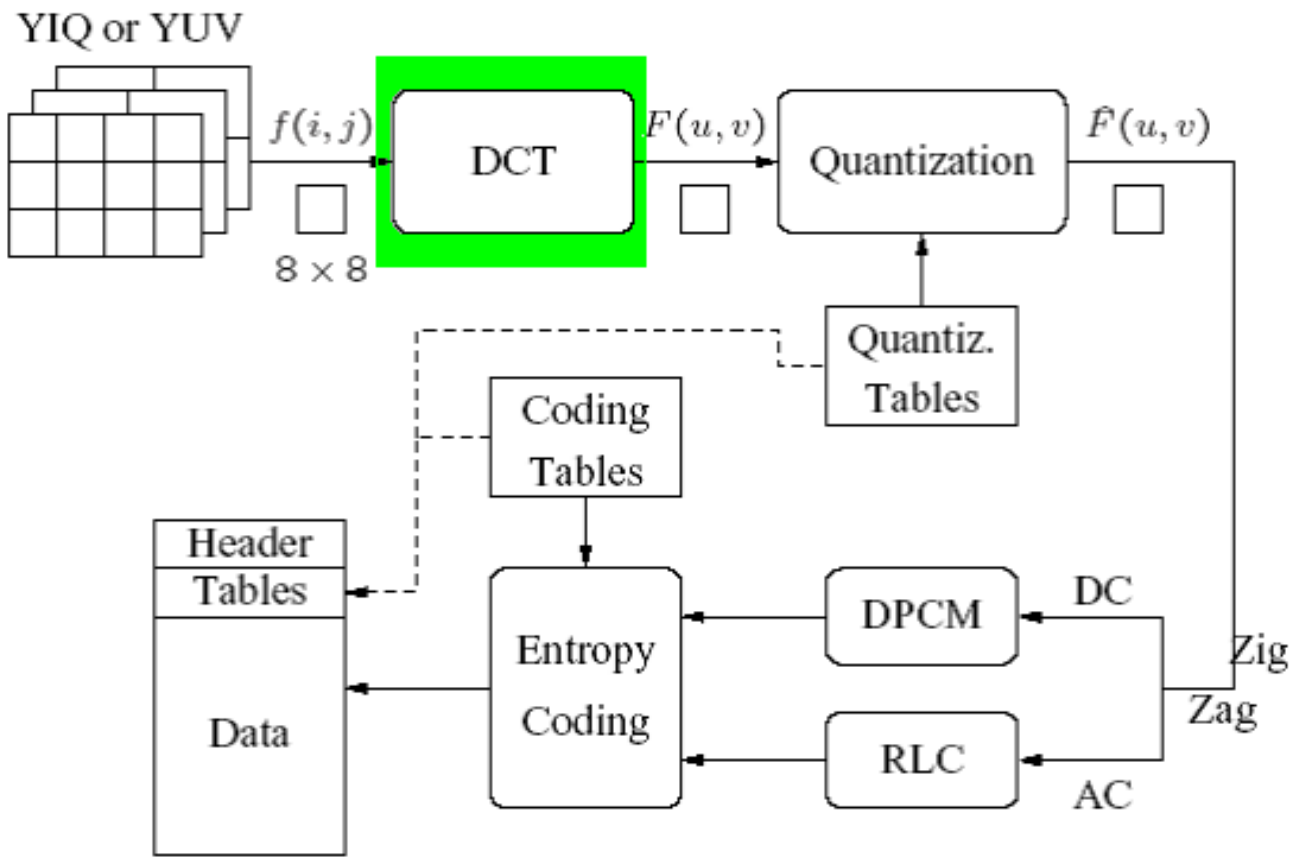
JPEG: Transform RGB Color Model to YUV or YIQ



JPEG: Transform RGB Color Model to YUV or YIQ

- Motivation
  - Partition the image data based on the visual acuity (敏锐度)
  - Human’s visual acuity is much greater for gray than for color
- YUV color model
  - Y: luminance （亮度）information
  - U and V: chrominance （色度） information
- YIQ color model
  - Y in YIQ is the same as in YUV
  - I and Q are a rotated version of U and V
- Chroma subsampling (4:2:0) is used in JPEG.

JPEG: DCT Coding for Image Block



Compression Scheme of 2D DCT for Image Data

- The DCT formalizes spatial frequency
  - With a measure of how much the image contents change in correspondence to the number of cycles of a cosine wave per block
- Motivation of DCT
  - Spatial redundancy of the image
    - Useful image contents change relatively slowly
    - DCT can concentrate the information within first several components
  - Humans are not sensitive to the loss of information of high spatial frequency

Original Grayscale Image

Recovered from F(0,0) ~ F(2,2)

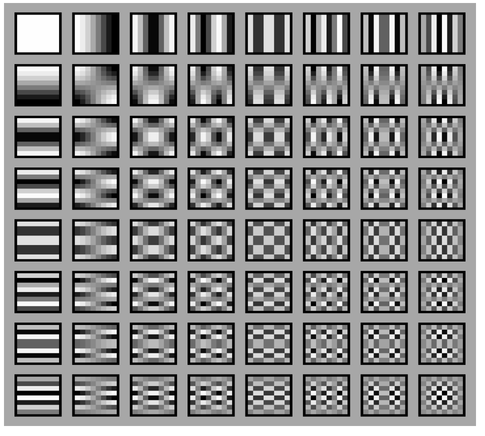
DCT Coding for Image Block

- Partition the image to the blocks of 8 pixels * 8 pixels
  - 2D DCT for each block in encoding part

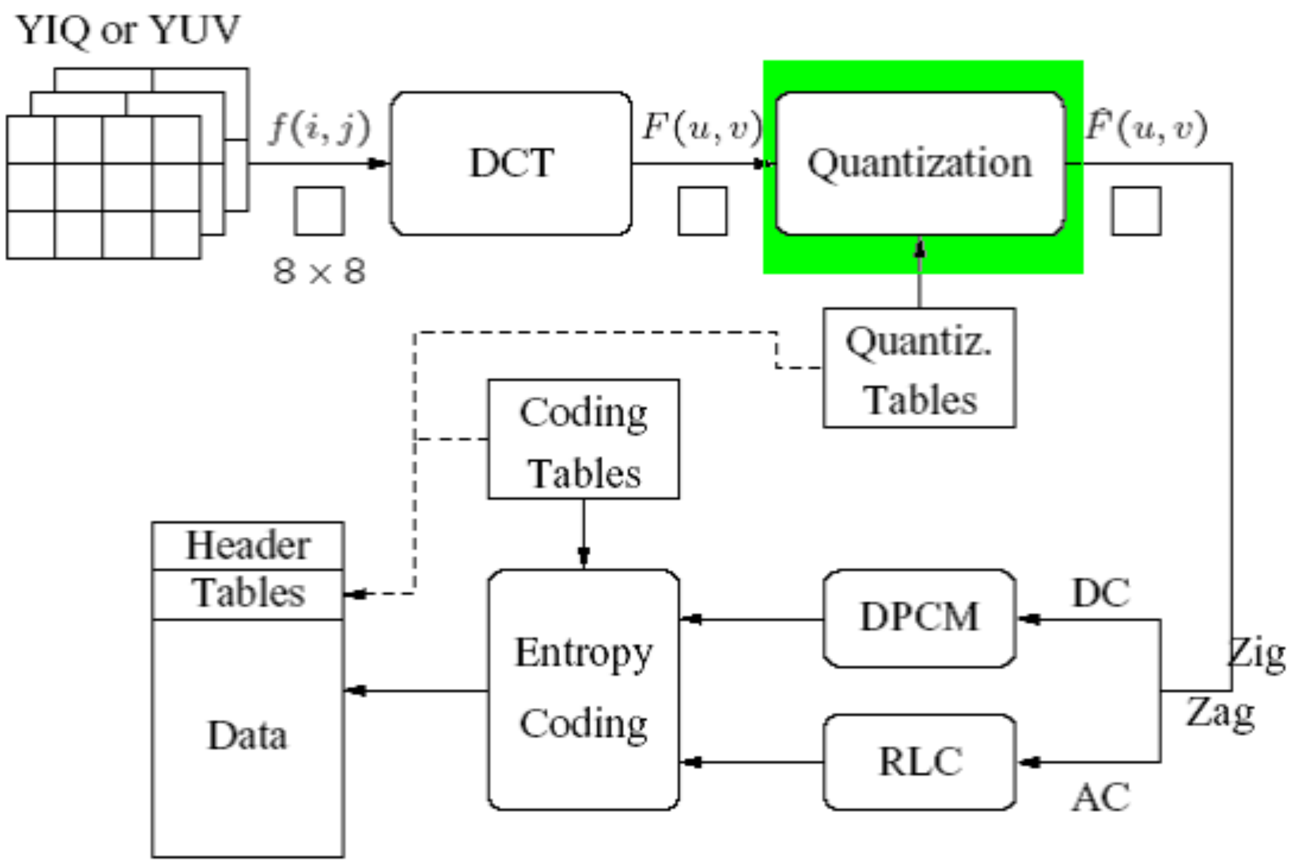
$$F(u,v) = \frac{C(u)\,C(v)}{4} \sum_{i=0}^{7} \sum_{j=0}^{7} \cos\frac{(2i+1)u\pi}{16} \cos\frac{(2j+1)v\pi}{16} f(i,j)$$

  - 2D IDCT for each block in decoding part

$$\tilde{f}(i,j) = \sum_{u=0}^{7} \sum_{v=0}^{7} \frac{C(u)\,C(v)}{4} \cos\frac{(2i+1)u\pi}{16} \cos\frac{(2j+1)v\pi}{16} F(u,v)$$



JPEG: Uniform Scalar Quantization



Uniform Scalar Quantization

- Motivation
  - Yield higher compression ratio
  - Main source for compression in JPEG
- Quantize DCT coefficients

$$\hat{F}(u,v) = round(\frac{F(u,v)}{Q(u,v)})$$

  - F(u,v) represents a DCT coefficients
  - Q(u,v) is a “quantization table" entry
  - $\hat{F}(u,v)$ represents the quantized DCT coefficients

Quantization Table

- Use different quantization tables for luminance information and chrominance information
- The entries of Q(u,v) tend to have larger values towards the lower right corner

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

The Luminance Quantization Table

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

The Chrominance Quantization Table

Quantization and Reconstruction

An 8 × 8 block from the Y image of 'Lena'

200	202	189	188	189	175	175	175
200	203	198	188	189	182	178	175
203	200	200	195	200	187	185	175
200	200	200	200	197	187	187	187
200	205	200	200	195	188	187	175
200	200	200	200	200	190	187	175
205	200	199	200	191	187	187	175
210	200	200	200	188	185	187	186

$f(i, j)$

515	65	-12	4	1	2	-8	5
-16	3	2	0	0	-11	-2	3
-12	6	11	-1	3	0	1	-2
-8	3	-4	2	-2	-3	-5	-2
0	-2	7	-5	4	0	-1	-4
0	-3	-1	0	4	1	-1	0
3	-2	-3	3	3	-1	-1	3
-2	5	-2	4	-2	2	-3	0

$F(u, v)$

JPEG compression for a smooth image block.

$$\begin{array}{rrrrrrrr}
32 & 6 & -1 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}$$

$\widehat{F}(u,v)$

$$\begin{array}{rrrrrrrr}
512 & 66 & -10 & 0 & 0 & 0 & 0 & 0 \\
-12 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-14 & 0 & 16 & 0 & 0 & 0 & 0 & 0 \\
-14 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}$$

$\widetilde{F}(u,v)$

$$\begin{array}{rrrrrrrr}
199 & 196 & 191 & 186 & 182 & 178 & 177 & 176 \\
201 & 199 & 196 & 192 & 188 & 183 & 180 & 178 \\
203 & 203 & 202 & 200 & 195 & 189 & 183 & 180 \\
202 & 203 & 204 & 203 & 198 & 191 & 183 & 179 \\
200 & 201 & 202 & 201 & 196 & 189 & 182 & 177 \\
200 & 200 & 199 & 197 & 192 & 186 & 181 & 177 \\
204 & 202 & 199 & 195 & 190 & 186 & 183 & 181 \\
207 & 204 & 200 & 194 & 190 & 187 & 185 & 184
\end{array}$$

$\tilde{f}(i,j)$

$$\begin{array}{rrrrrrrr}
1 & 6 & -2 & 2 & 7 & -3 & -2 & -1 \\
-1 & 4 & 2 & -4 & 1 & -1 & -2 & -3 \\
0 & -3 & -2 & -5 & 5 & -2 & 2 & -5 \\
-2 & -3 & -4 & -3 & -1 & -4 & 4 & 8 \\
0 & 4 & -2 & -1 & -1 & -1 & 5 & -2 \\
0 & 0 & 1 & 3 & 8 & 4 & 6 & -2 \\
1 & -2 & 0 & 5 & 1 & 1 & 4 & -6 \\
3 & -4 & 0 & 6 & -2 & -2 & 2 & 2
\end{array}$$

$\epsilon(i,j) = f(i,j) - \tilde{f}(i,j)$

JPEG compression for a smooth image block.

Another 8×8 block from the Y image of 'Lena'

70	70	100	70	87	87	150	187
85	100	96	79	87	154	87	113
100	85	116	79	70	87	86	196
136	69	87	200	79	71	117	96
161	70	87	200	103	71	96	113
161	123	147	133	113	113	85	161
146	147	175	100	103	103	163	187
156	146	189	70	113	161	163	197

$f(i, j)$

-80	-40	89	-73	44	32	53	-3
-135	-59	-26	6	14	-3	-13	-28
47	-76	66	-3	-108	-78	33	59
-2	10	-18	0	33	11	-21	1
-1	-9	-22	8	32	65	-36	-1
5	-20	28	-46	3	24	-30	24
6	-20	37	-28	12	-35	33	17
-5	-23	33	-30	17	-5	-4	20

$F(u, v)$

JPEG compression for a textured image block.

-5	-4	9	-5	2	1	1	0
-11	-5	-2	0	1	0	0	-1
3	-6	4	0	-3	-1	0	1
0	1	-1	0	1	0	0	0
0	0	-1	0	0	1	0	0
0	-1	1	-1	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

$\hat{F}(u, v)$

-80	-44	90	-80	48	40	51	0
-132	-60	-28	0	26	0	0	-55
42	-78	64	0	-120	-57	0	56
0	17	-22	0	51	0	0	0
0	0	-37	0	0	109	0	0
0	-35	55	-64	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

$\tilde{F}(u, v)$

70	60	106	94	62	103	146	176
85	101	85	75	102	127	93	144
98	99	92	102	74	98	89	167
132	53	111	180	55	70	106	145
173	57	114	207	111	89	84	90
164	123	131	135	133	92	85	162
141	159	169	73	106	101	149	224
150	141	195	79	107	147	210	153

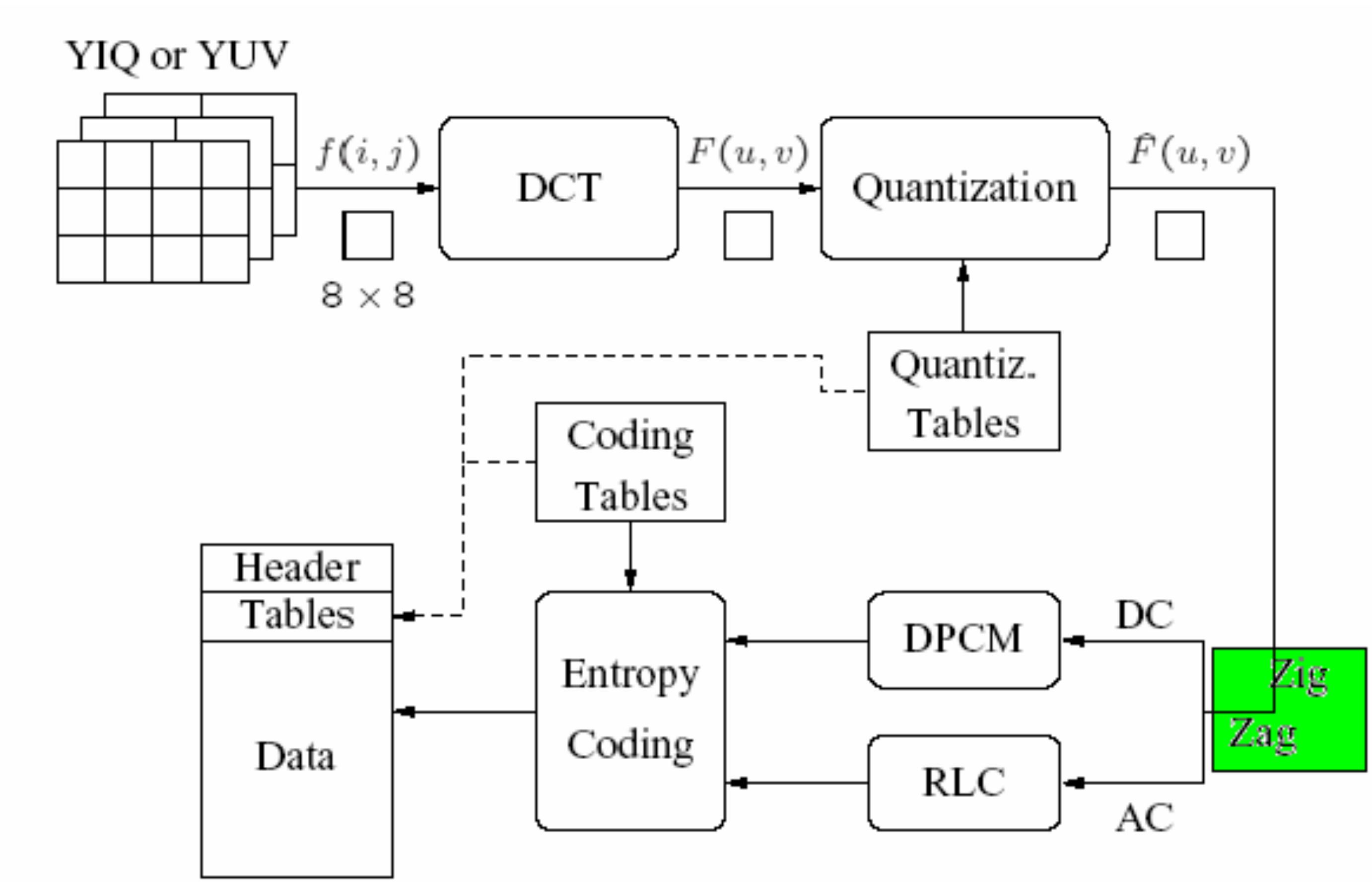
$\tilde{f}(i, j)$

0	10	-6	-24	25	-16	4	11
0	-1	11	4	-15	27	-6	-31
2	-14	24	-23	-4	-11	-3	29
4	16	-24	20	24	1	11	-49
-12	13	-27	-7	-8	-18	12	23
-3	0	16	-2	-20	21	0	-1
5	-12	6	27	-3	2	14	-37
6	5	-6	-9	6	14	-47	44

$\epsilon(i, j) = f(i, j) - \tilde{f}(i, j)$

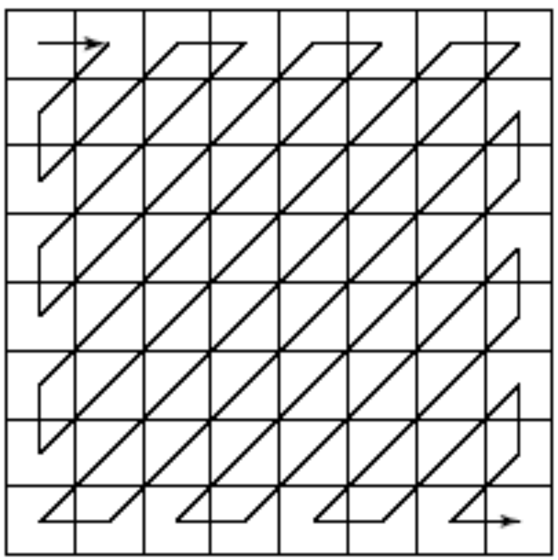
JPEG compression for a textured image block.

JPEG: ZigZag Scan



Zig-Zag Scan for Quantized DCT Coefficients

◆ Zig-Zag scan



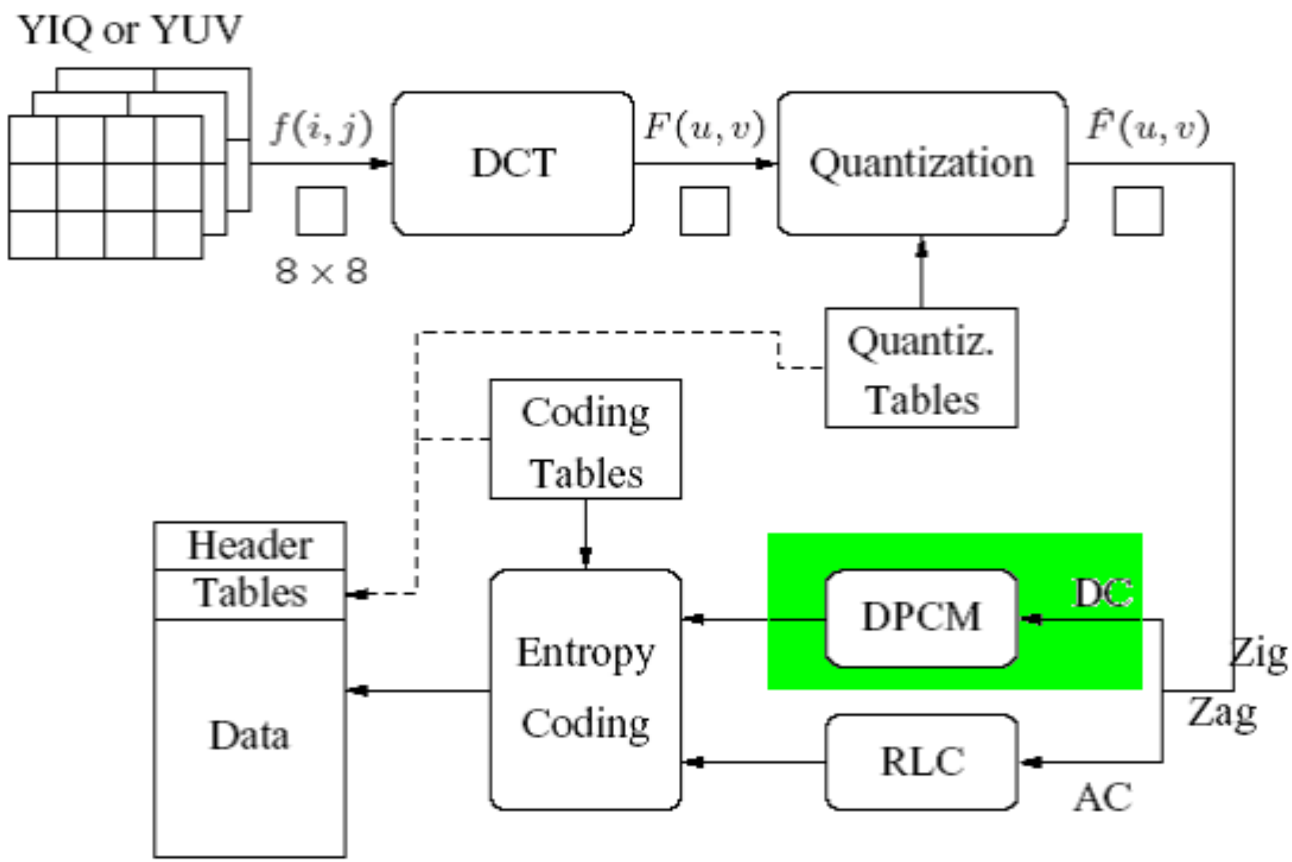
32	6	-1	0	0	0	0	0
-1	0	0	0	0	0	0	0
-1	0	1	0	0	0	0	0
-1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

$\hat{F}(u, v)$

Zigzag scan result:
(32, 6,-1,-1,0 –1,0,0,0,-1,0,0,1,0,0… … 0)

- To make it most likely to hit a long run of zeros: a *zig-zag scan* is used to turn the 8 × 8 matrix $\hat{F}$ (u, v) into a *64-vector*.

JPEG: DPCM for DC Coefficients



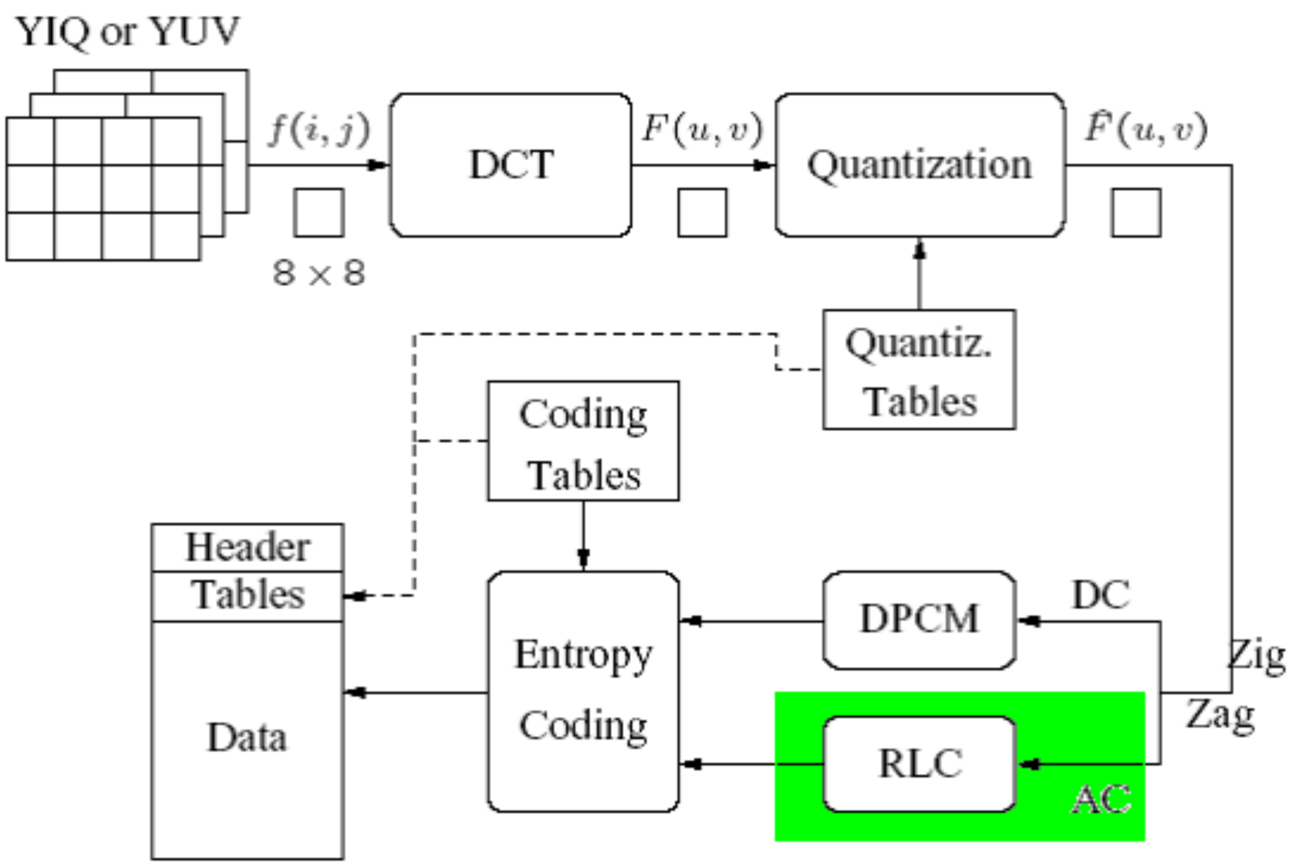
DPCM for DC Coefficients

- Quantized DCT coefficients
  - Direct current (DC) coefficient ~ F(0,0)
  - Alternate current (AC) coefficient ~ other F(u,v)
  - DPCM is the coding method for DC coefficients
  - Differential Pulse Code Modulation （差分脉冲编码调制）

- Example of encoding DC coefficients
  - assuming $d_i = DC_{i+1} - DC_i$, and $d_0 = DC_0$.
  - DC coefficients for the first 5 image blocks: 150,155, 149, 152, 144
  - Coding result: 150, 5, -6, 3, -8

- Reference reading
  - 中文版 P106 (P172 英文版)

JPEG: RLC for AC Coefficients



Run-length Coding (RLC, 游长编码) on AC coefficients

- RLC aims to turn the $\hat{F}(u, v)$ values into sets {#-zeros-to- skip , next non-zero value}.

32	6	-1	0	0	0	0	0
-1	0	0	0	0	0	0	0
-1	0	1	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
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

$\hat{F}(u, v)$

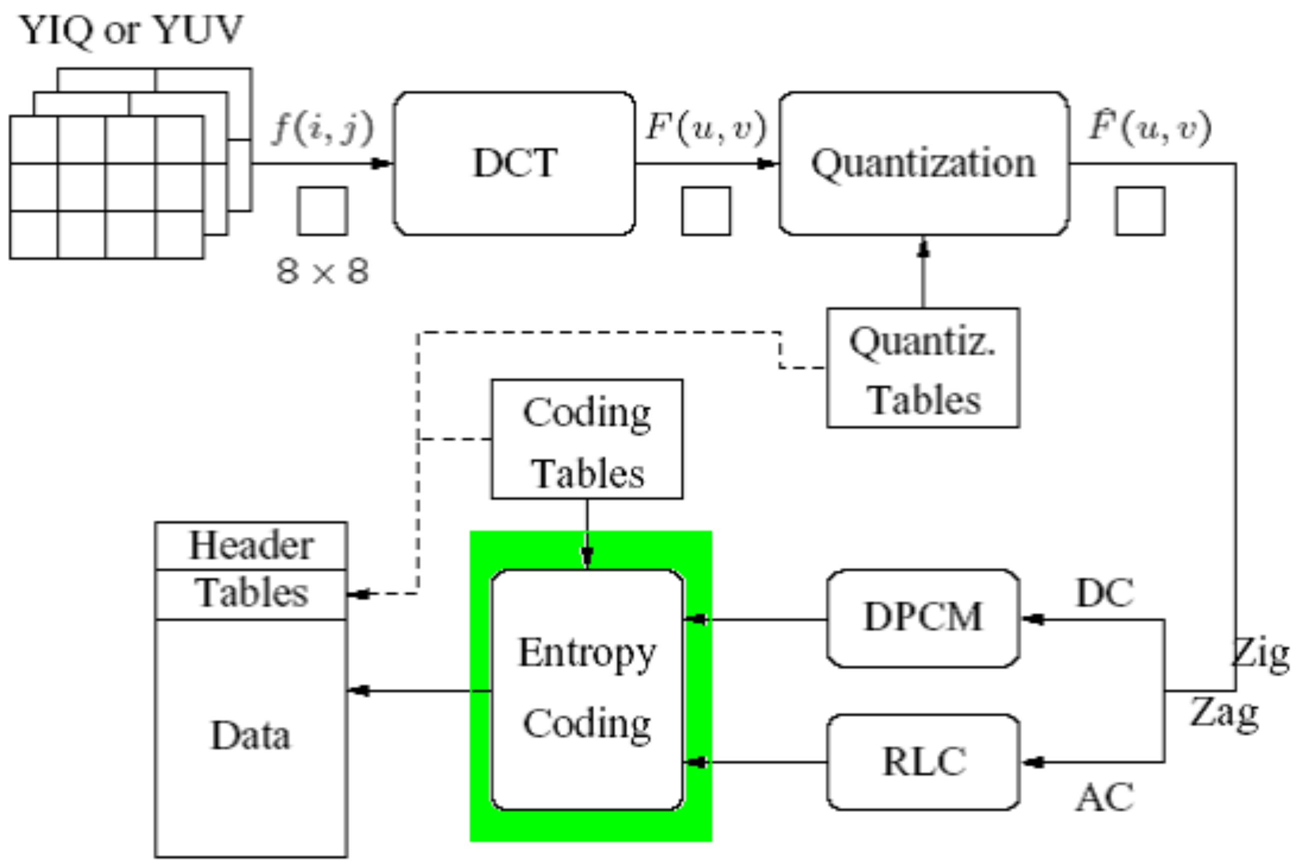
Zigzag scan result

(32, 6,-1,-1,0 –1,0,0,0,-1,0,0,1,0,0… … 0)

Modified Run-Length coding result

(0,6)(0,-1)(0,-1)(1,-1)(3,-1)(2,1)(0,0)

JPEG: Entropy Coding



Entropy Coding

- The DC and AC coefficients finally undergo an entropy coding step to gain a possible further compression
- Each DPCM coded DC coefficient is represented by (SIZE, AMPLITUDE)
  - DC coefficients for the first 5 image blocks: 150,155, 149, 152, 144
  - Coding result: 150, 5, -6, 3, -8
- In (SIZE, AMPLITUDE) format：

（8，10010110）（3，101）（3，001）（2，11）（4，0111）

- 负数表示为正数的反码，以0开头的都为负数

- SIZE is Huffman coded

Entropy Coding

- Each RLC coded AC coefficient (RUNLENGTH, VALUE) is represented by (RUNLENGTH, SIZE, AMPLITUDE)

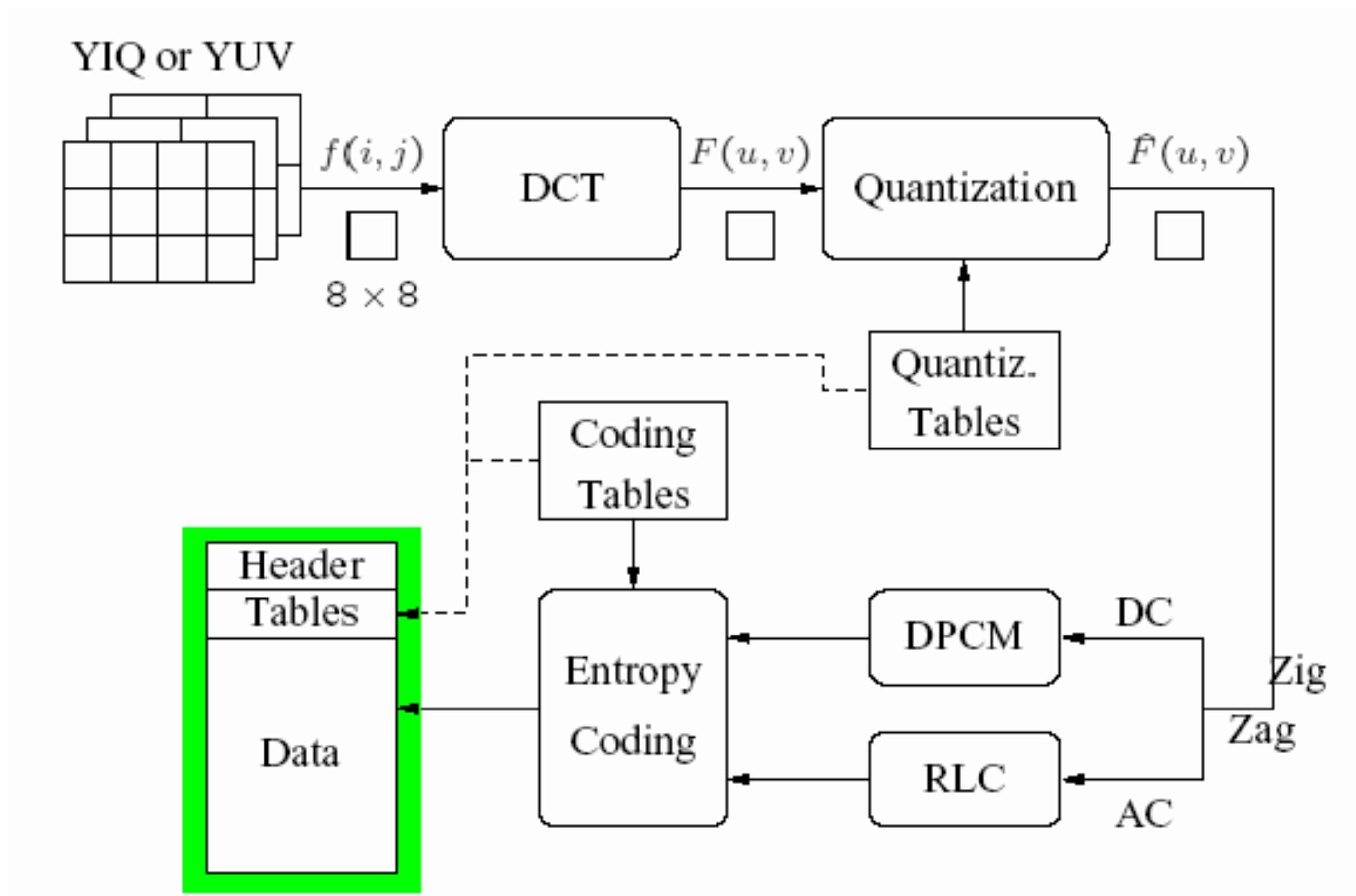
  Zigzag scan result
  (32, 6,-1,-1,0 –1,0,0,0,-1,0,0,1,0,0… … 0)

  Modified Run-Length coding result
  (0,6)(0,-1)(0,-1)(1,-1)(3,-1)(2,1)(0,0)

- In (RUNLENGTH, SIZE, AMPLITUDE) format
  (0,3,110)(0,1,0)(0,1,0)(1,1,0)(3,1,0)(2,1,1)(0,0,0)

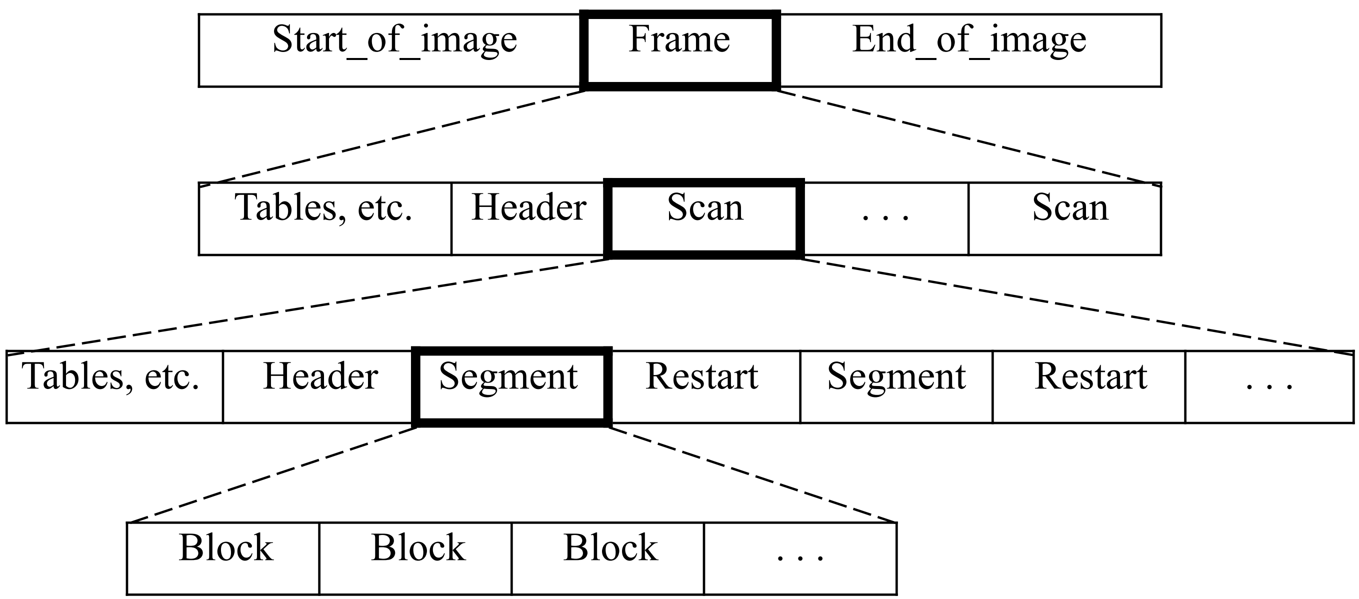
- RUNLENGTH, SIZE各用4位表示，整合成一个字节，用哈夫曼编码
- AMPLITUDE use actual values

JPEG File Generation



JPEG File Generation

A Glance at the JPEG Bitstream



JPEG bitstream.

练习

1030	130	-24	8	2	4	-16	10
-32	6	4	0	0	-22	-4	6
-24	12	22	-2	6	0	2	-2
-16	6	-8	4	-4	-6	-10	-4
0	-4	14	-10	8	0	-2	-8
0	-3	-2	0	8	2	-2	0
6	-4	-6	6	6	-2	-2	-6
-4	10	-4	8	-4	4	-6	0

F(u,v)

$Q(u,v)$

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

$$\hat{F}(u,v) = round\left(\frac{F(u,v)}{Q(u,v)}\right)$$

量化结果? $\hat{F}(u,v)$

Z字扫描 $\hat{F}(u,v)$ 的结果？

AC系数的游长编码？

练习

64	12	-2	1	0	0	0	0
-3	1	0	0	0	0	0	0
-2	1	1	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
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

量化结果 $\hat{F}(u,v)$

Z字扫描的结果？

64， 12， -3， -2， 1， -2，1，0，1，-1，0，0，1，0，...0

AC系数的游长编码？

（0,12）（0，-3）（0，-2）（0,1）（0，-2）（0,1）（1,1）（0，-1）（2,1）（0,0）

练习

DC系数：140, 75, 90, 120, 80, 45

DC系数的DPCM 编码？

140， -65，15，30，-40，-35

(SIZE, AMPLITUDE), 表示？

（8,10001100）（7，0111110）（4，1111）（5,11110）（6，010111）（6, 011100）

扩展

- 阅读第9章 9.2 The JPEG2000 Standard，介绍JPEG2000相对于标准JPEG的改进。