The JPEG Standard

- JPEG **was developed by the *Joint Photographic Experts Group*, and was standard in 1992.**
- JPEG **is a** lossy **image compression method.**
- JPEG **employs the DCT (*Discrete Cosine Transform*) method to reduce the** *frequency* **redundancy in** *spatial* **domain**

Major Steps in JPEG coding

1. **Transform RGB to YIQ or YUV**
2. **DCT**
3. **Quantization**
4. **Zig-zag scan and run-length encoding**
5. **Entropy coding**

JPEG encoding

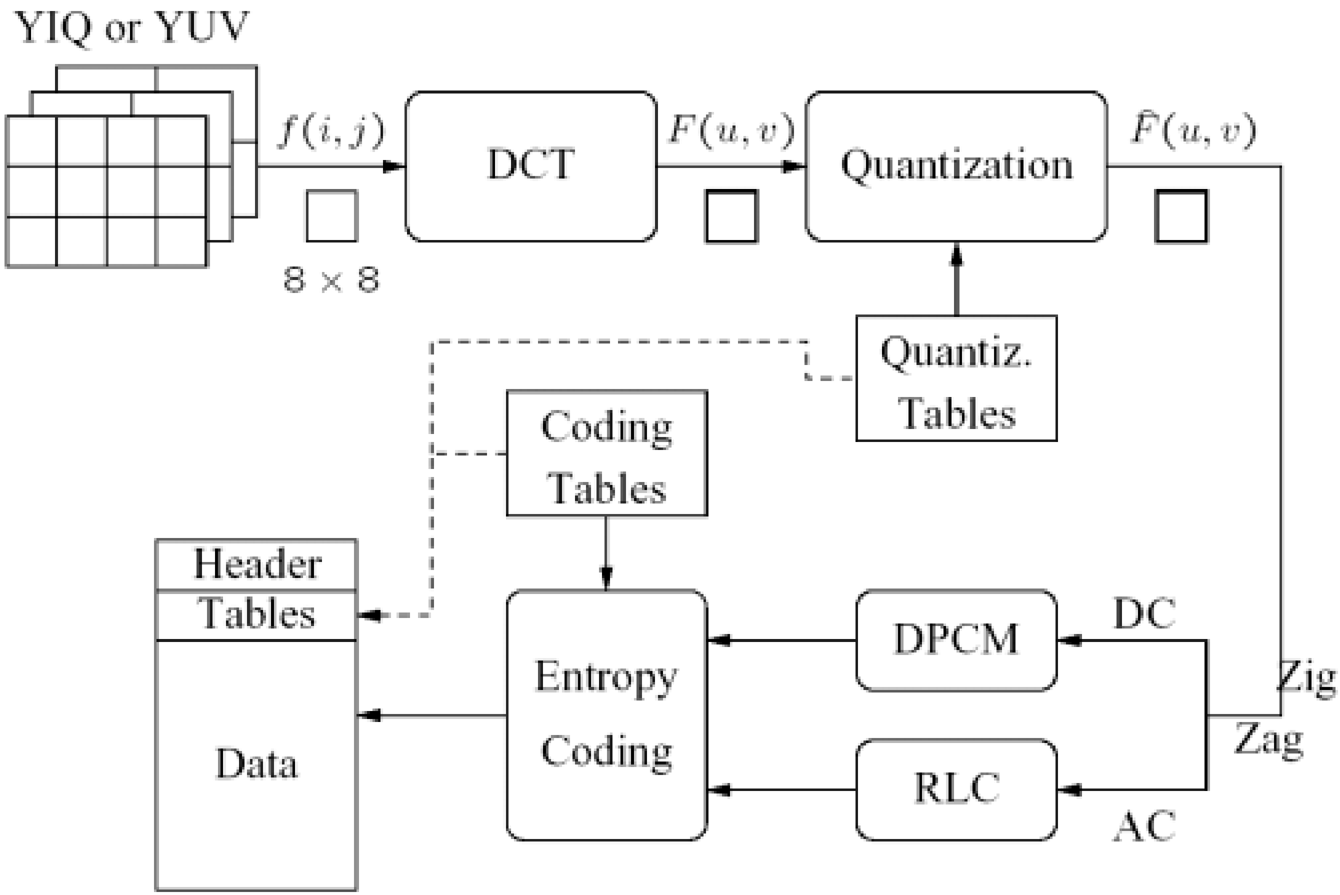


Fig. 9.1: Block diagram for JPEG encoder.

Main Steps in JPEG coding

- **Transform** RGB to YIQ or YUV
- **DCT**
- **Quantization**
- **Zig-zag** scan and **run-length** coding
- **Entropy** coding

DCT on image blocks

- **Each image is divided into 8 x 8 blocks.**
- **The** 2D DCT **is applied to each block image f(*i,j*), with output being the DCT coefficients F(*u,v*) for each block.**
- **Using blocks, however, has the effect of isolating each block from its neighboring context. This is why JPEG images look** *choppy* **(*blocky*) when a high *compression ratio* is specified by the user.**

Quantization

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

where F(u; v)* represents a DCT coefficient, *Q(u; v)* is a "quantization matrix" entry, and *^F(u; v)* represents the *quantized DCT coefficients

- **The quantization step is the main source for loss in JPEG compression.**
  - The entries of *Q(u; v)* tend to have **larger values** towards the lower-right corner.
  - This aims to introduce more loss at the higher spatial frequencies - a practice supported by Observations 1 and 2 (page 11)

Quantization

- **the default *Q* (*u; v*) values obtained from psychophysical studies with the goal of maximizing the compression ratio while [illegible] JPEG images**

9.1 The Luminance Quantization Table

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 Chrominance Quantization

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

JPEG compression on a smoothly block

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

$F(u, v)$

Fig 9.2: An image block at smoothly area. After DCT, only the DC and the first few AC components, showing low spatial frequencies, the other coefficients have small magnitudes

JPEG compression on a smoothly block

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

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

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

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

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

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

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

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

- ^F(u,v): quantized DCT, ~F(u,v): de-quantized DCT, ~f(i,j): reconstructed image, e(i,j): the error

JPEG compression on a textured 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

$F(u,v)$

An image block f(i,j) at rapidly changing area. In F(u,v), DCT of f(i,j): many more AC components show large magnitudes

-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

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

Figure 9.3 (p. 258): An image block at rapidly changing area. Many more AC components show large magnitudes, and the error e is also larger now than in Fig 9.2- **JPEG introduces more loss** if the image has quickly changing details

Pre-compression for Entropy coding

- **The third main step in JPEG is the** entropy coding**, there are some small additional data compression process steps:**
  - **Differential Pulse Code Modulation (DPCM)** on DC coefficients, for the DC values are large and are **unlikely** to change **drastically** in a short distance
  - **Run-length Coding (RLC)** on AC coefficients for many zeros in quantized AC coefficients

DPCM on DC coefficients

- Differential Pulse Code Modulation (DPCM)
- **Example:**
  - If the DC coefficients for the first 5 image blocks are 150,155, 149, 152, 144, then
  - the DPCM would produce 150, 5, -6, 3, -8,
  - According to $d_i = DC_{i+1} - DC_i$, and $d_0 = DC_0$.

Run-length Coding on AC coefficients

- Run-length Coding (RLC) **aims to turn the AC coefficients, *F*(*u*; *v*) values into sets of**

  ***{#-zeros-to-skip , next non-zero value}*.**

- **Example of general RLC**

a b aaaa bbbbbbb → a1 b1 a4 b7

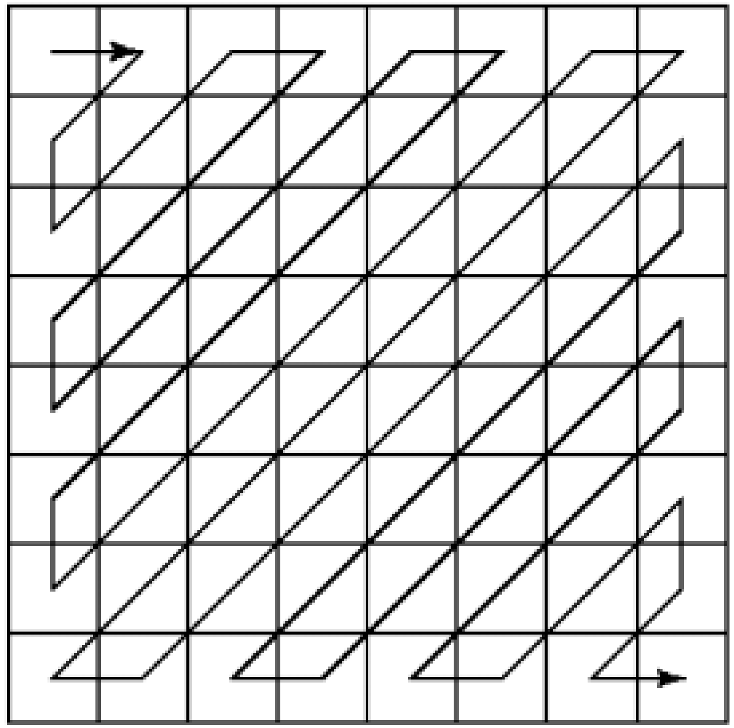
0000000 111111111 000000 11

→ 0,7 1,10 0,5 1,2

→ 7, 10, 5, 2 if 0 is always the start code

Run-length Coding on AC coefficients

- To make it most likely to hit a long run of zeros: a ***zig-zag scan*** is used to turn the 8x8 matrix $\hat{F}(u; v)$ into a *64-vector*



Entropy Coding

DC and AC coefficients finally undergo the 3rd *entropy coding* **step for further compression**

- **Use DC as example: A DPCM coded coefficient is represented by (**SIZE, AMPLITUDE**), where**
  - **SIZE** indicates how many bits are needed for representing the coefficient, and
  - **AMPLITUDE** contains the actual bits.
- **Ex: codes "150, 5, -6, 3, -8" will be turned into**

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

SIZE and AMPLITUDE

- SIZE **is then Huffman coded, since smaller SIZEs occur much more often**
- AMPLITUDE **is not Huffman coded, its value can change widely so Huffman coding has no appr**

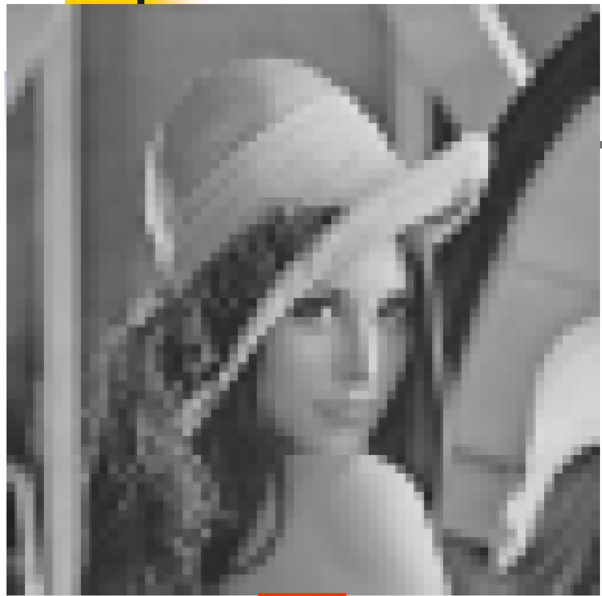
- Baseline entropy coding details - size and amplitude category

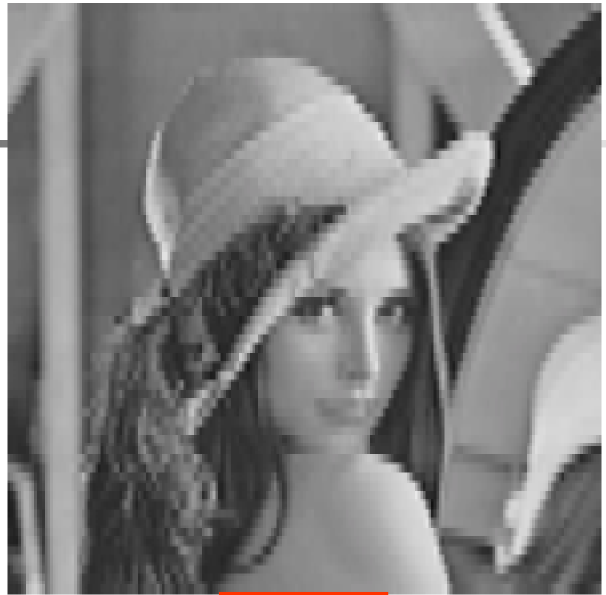
SIZE	AMPLITUDE
1	-1, 1
2	-3, -2, 2, 3
3	-7..-4, 4..7
4	-15..-8, 8..15
.	.
.	.
.	.
10	-1023..-512, 512..1023

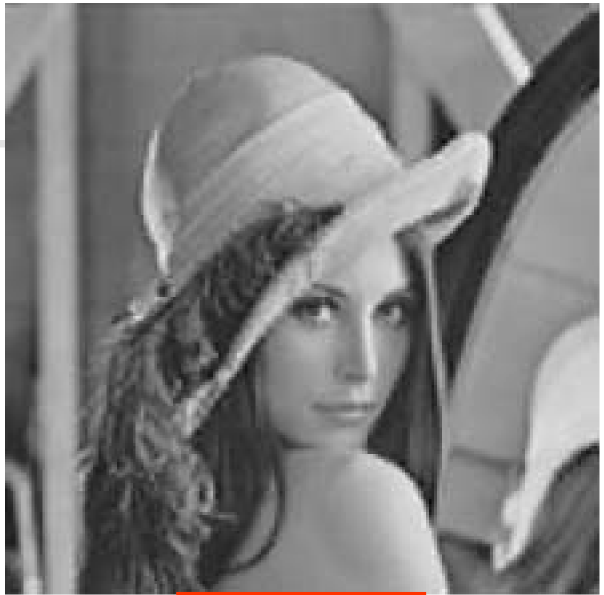
JPEG display Modes

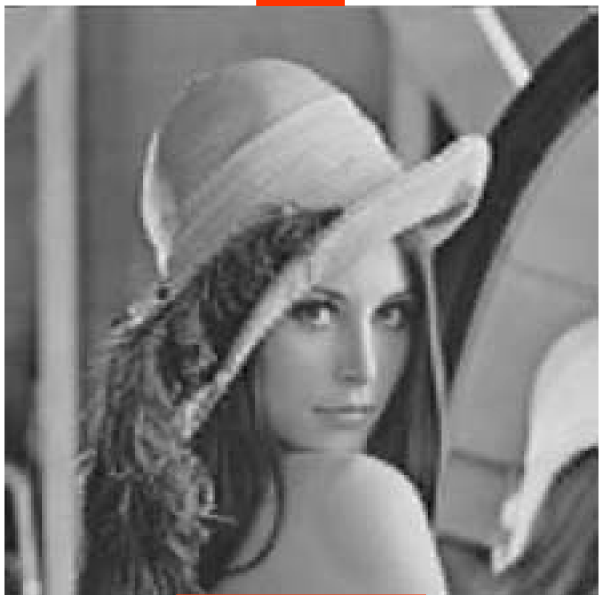
- *Sequential Mode*: **the default mode, each gray level or color image component is encoded in a single left-to-right, top-to-bottom scan.**
- *Progressive Mode*: **delivers low quality versions of the image quickly, followed by higher quality passes:**
  - **Spectral selection**: using the advantage of spectral characteristics of DCT coefficients: higher AC components provide only detail information.
  - **Successive approximation**: (1) encode the first few MSBs of DCT coeffs,e.g., bits 7,6,5, 4 then, (2) a few more less-significant bits, e.g., bits 3, 2, ., and finally the LSB, bit 0.

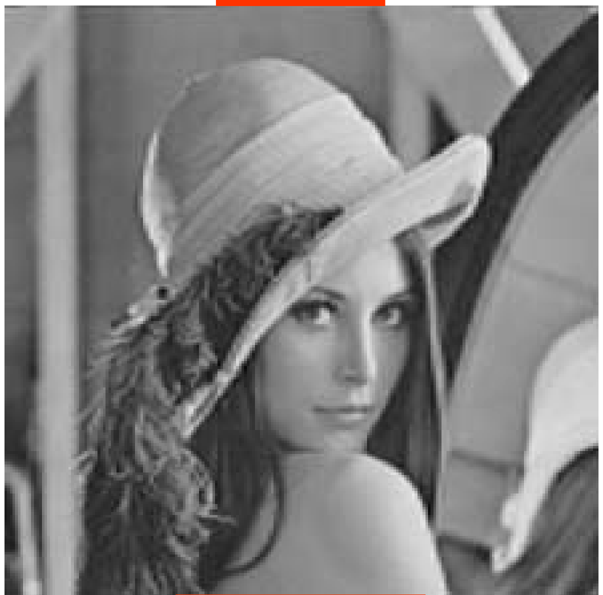
Progressive display Spectral approach

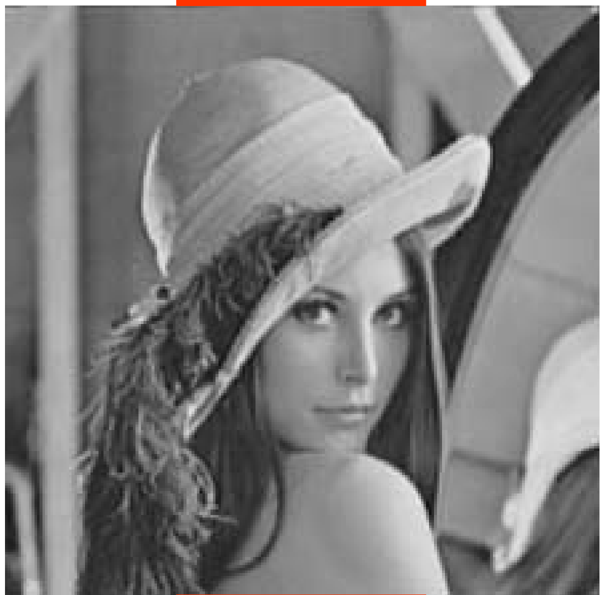

DC


DC+AC1


DC+AC1+AC2


DC+AC1~AC3

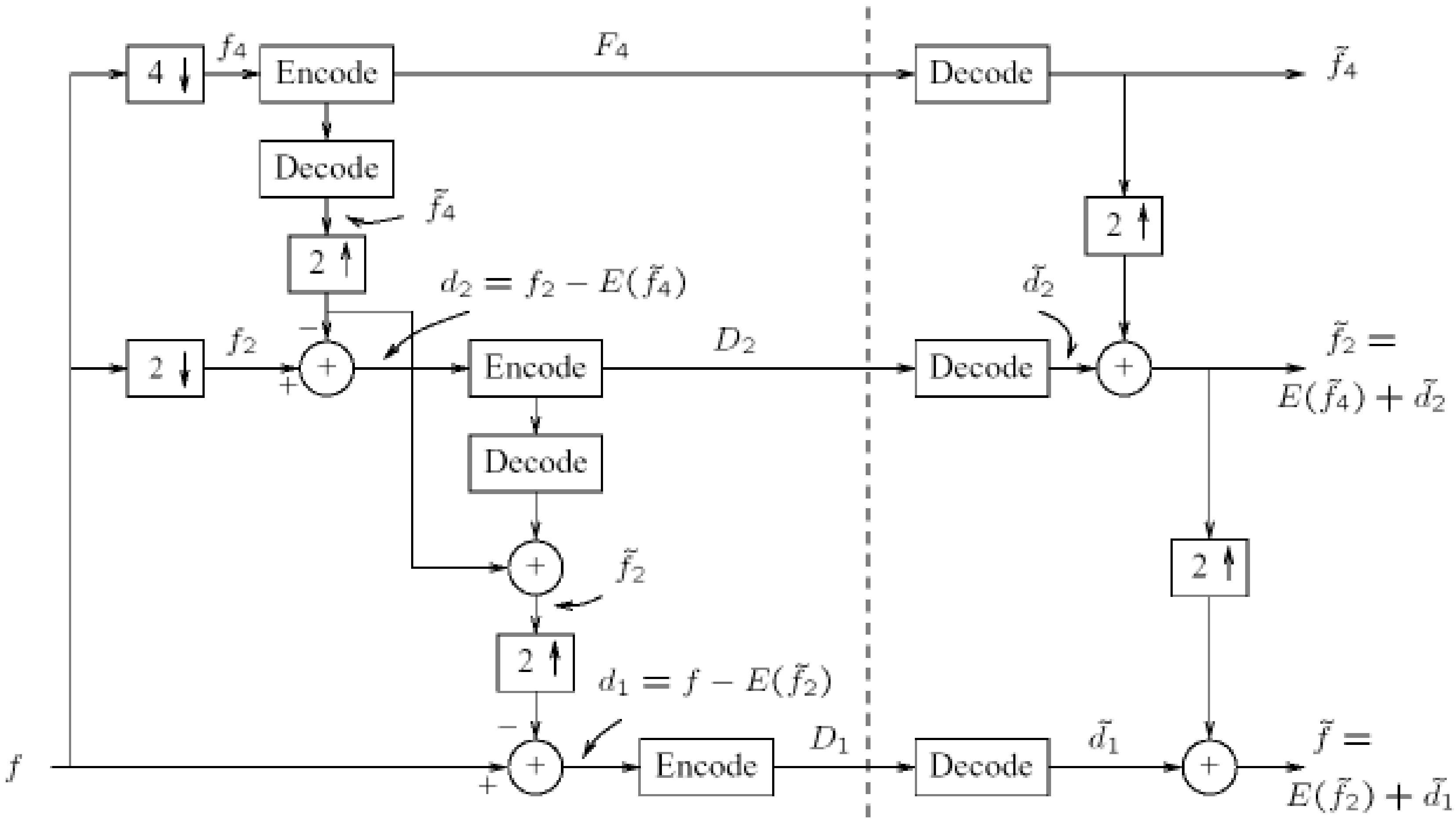

DC+AC1~AC4


DC+AC1~AC5

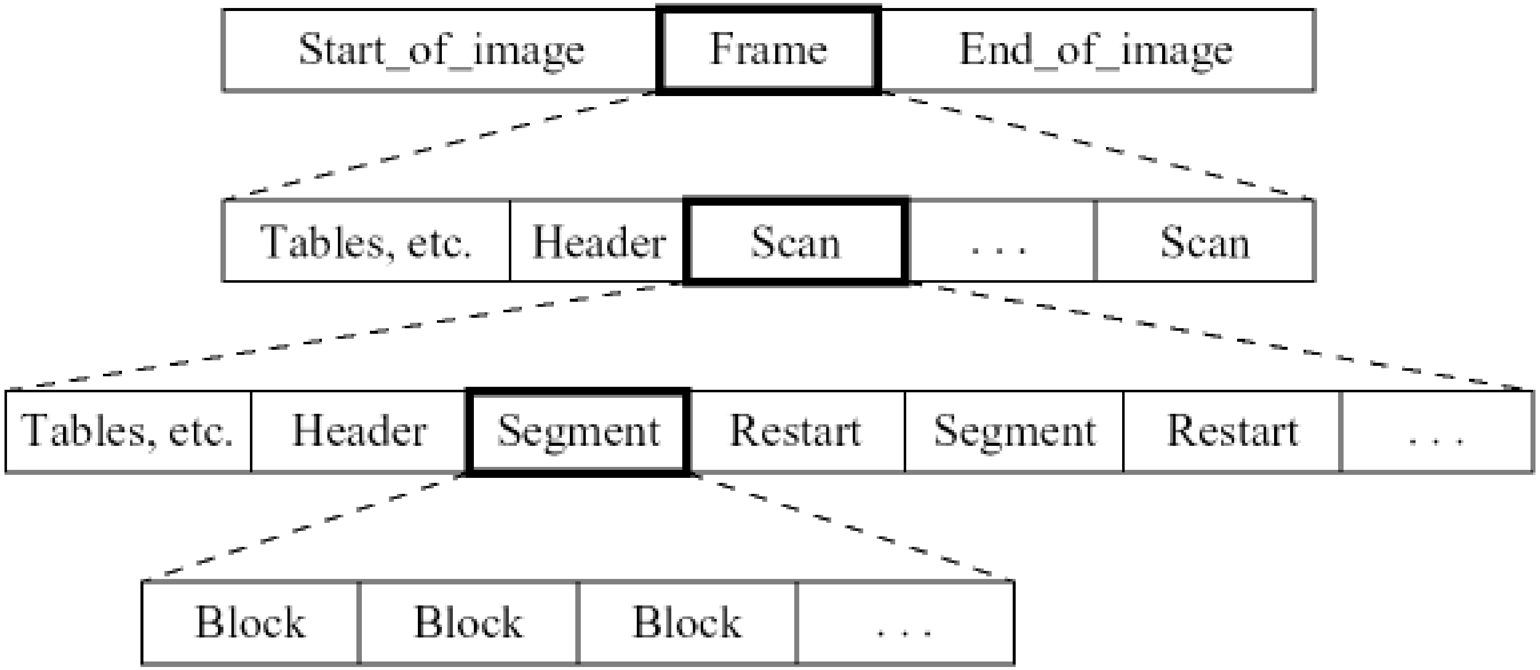
JPEG display Modes (c)

- *Hierarchical Mode*: **delivers compressed low-pass filtered image first, then successively higher resolutions of additional details (differences from the lower resolution images).**

Block diagram for Hierarchical JPEG



Glance at the JPEG Bitstream



- a ***frame*** is a picture;
- a ***scan*** is a pass thro pixels, i.e., the red component
- a ***segment*** is a group of blocks and
- a ***block*** consists of 8x8 pixels