فشرده‌سازی

محیط‌های چندرسانه‌ای

۱۱-۰۸۰-۱۳۳

(بخش هشتم)

دانشگاه شهید بهشتی

بهار ۱۳۹۴

احمد محمودی ازناوه

http://faculties.sbu.ac.ir/~a_mahmoudi/

فهرست مطالب

- اندازه‌گیری میزان اطلاعات
  - مفهوم آنتروپی
- انواع فشرده‌سازی
- چندی کردن(quantization)
- استاندارد فشرده‌سازی jpeg

اندازه‌گیری اطلاعات

- اگر تعدادی نماد در دست باشد، به وسیله‌ی رابطه‌ی شانون کمینه‌ی تعداد بیت‌ها برای ارسال نمادهای مورد نظر محاسبه می‌گردد.
- اگر M تعداد نمادها باشد:

$$Ent = -\sum_{i=1}^{M} \rho_i \log_2^{\rho_i} \qquad 0 \le \rho_i \le 1$$

$$Avrage\ \ Numbit = Ent$$

- اگر داشته باشیم:

$$0 \le Ent \le \log_2^M$$

A = 0.5 B = 0.2 C = 0.1 D = 0.1 E = 0.1

AAAAABBCDE

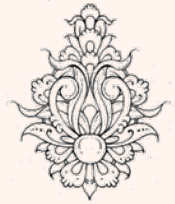
$$Ent = -[(0.5\log_2^{0.5} + 0.2\log_2^{0.2} + (0.1\log_2^{0.1})*3)]$$

$$Ent = -[-0.5 + (-0.46438) + (-0.9965)]$$

$$Ent = -[-1.9] = 1.9$$

ارسال اطلاعات

- با توجه به این‌که اطلاعات موجود یا نمادهای تولید شده در یک پیام دارای احتمال وقوع یکسان نیستند، می‌توان اطلاعات با احتمال وقوع کم‌تر را با تعداد بیت بیشتر و اطلاعات با احتمال وقوع بیشتر را با تعداد بیت کم‌تر ارسال کرد.

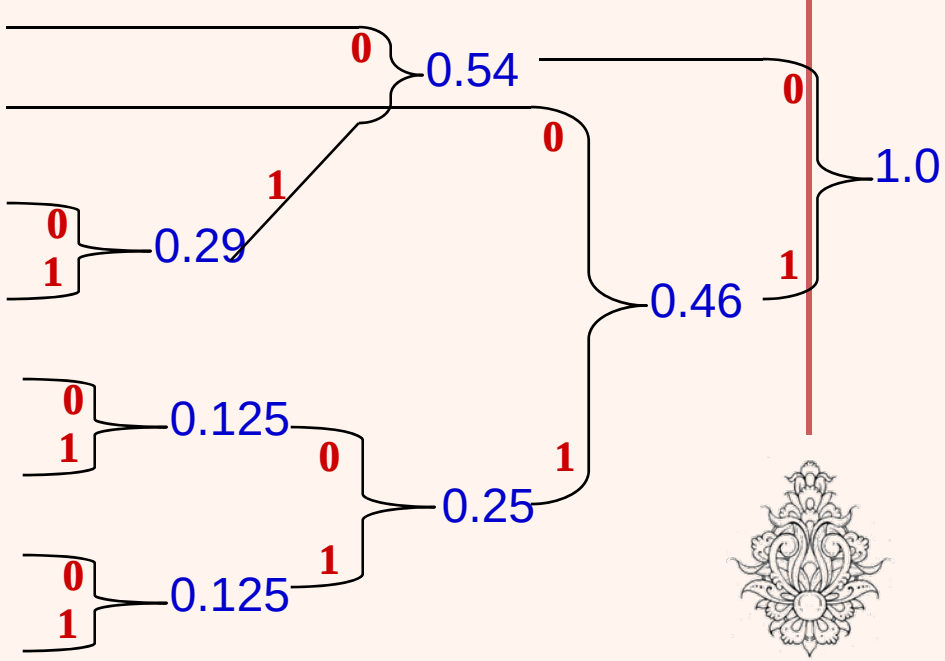


کدگذاری هافمن

000	00	S0	0.25
001	10	S1	0.21
010	010	S2	0.15
011	011	S3	0.14
100	1100	S4	0.0625
101	1101	S5	0.0625
110	1110	S6	0.0625
111	1111	S7	0.0625

Huffman

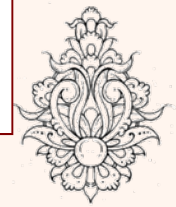
(trace from root)



نیازمندی‌های یک کدگذار

- به صورت یکتا قابل کدگشایی باشد.
- دارای پیشوند یکتا باشد، به عبارت دیگر برای کدگشایی نیازی به بررسی داده‌های بعدی نباشد.

- A good code should be:
  - Uniquely decodable
  - Instantaneously decodable – prefix code



مثال

Codebook 1
(a prefix code)

Symbol	Codeword
a1	0
a2	10
a3	110
a4	111

Codebook 2
(not a prefix code)

Symbol	Codeword
a1	0
a2	01
a3	100
a4	011

Bitstream: 0 0 1 1 0 1 0 1 1 0 1 0 0

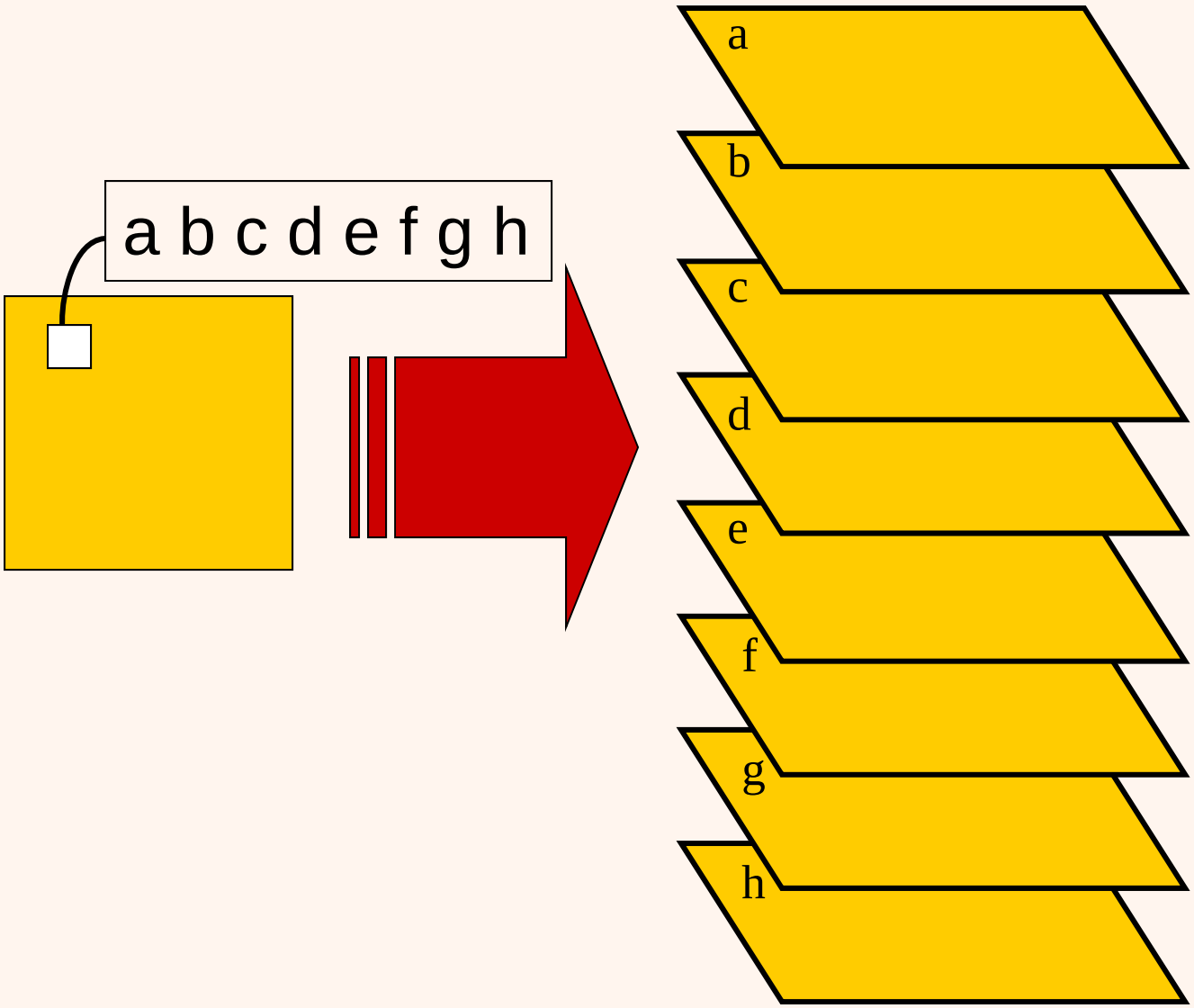
Decoded string based on codebook 1: 0|0|1 1 0|1 0|1 1 0|1 0|0--> a1 a1 a3 a2 a3 a2 a1
(can decode instantaneously)

Decoded string based on codebook 2: 0|0 1 1|0 1|0 1 1|0|1 0 0-->a1 a4 a2 a4 a1 a3
(need to look ahead to decode)

شیط‌های پلدرسانه‌ای

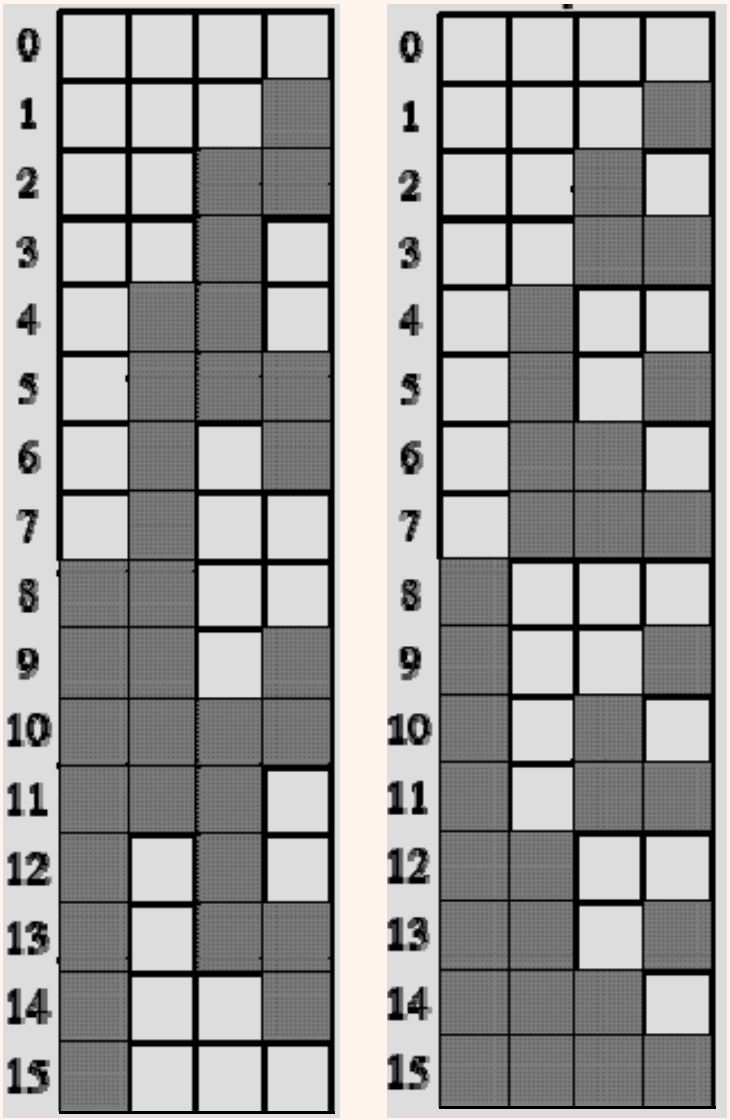
Bit Plane Coding

کدگذاری بیتی



حل مشکل با کد Gray

Gray code	Binary code



$$\{b_{m-1}, b_{m-2}, \ldots\ldots, b_1, b_0, \} \longrightarrow \{g_{m-1}, g_{m-2}, \ldots\ldots, g_1, g_0, \}$$

$$g_{m-1} = b_{m-1}$$

$$g_i = b_i \oplus b_{i+1} \qquad 0 \le i \le m\text{-}2$$

$$1110 \longrightarrow 1001$$

$$g_{m-1} = b_{m-1}$$

$$g_i = b_i \oplus b_{i+1} \qquad 0 \le i \le m\text{-}2$$

نمایش بیتی

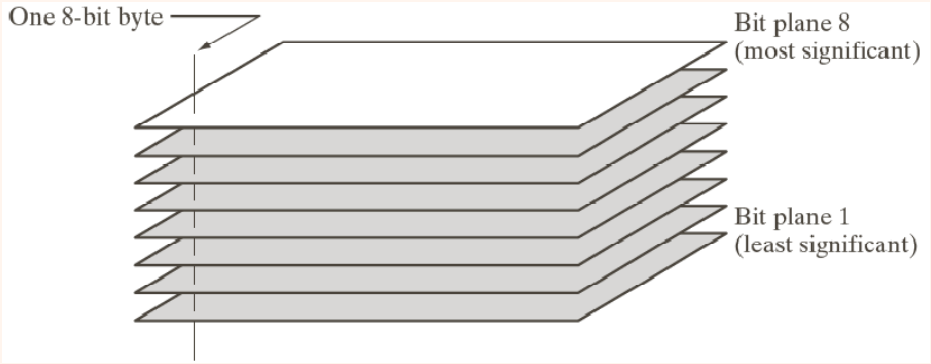
Bit Extraction

- به نمایش مقادیر روشنایی به صورت لایه‌هایی از تصاویر دودویی گفته می‌شود.
- با فرض داشتن $L=256$ داریم:

بیت پر ارزش

$$r = b_7 b_6 b_5 b_4 b_3 b_2 b_1 b_0 = \sum_{i=0}^{7} b_i \, 2^i$$

- متناظر با هر بیت یک تصویر ساخته می‌شود.



نمایش دودویی و Gray

- جهت نمایش مدل bit plane از دو کد **دودویی** و gray استفاده می‌شود.
- بیت‌های کم ارزش همانند **نویز سفید** خواهند بود و می‌توان از آن‌ها صرفنظر کرد.
- استفاده از کد gray به این دلیل است که در نمایش دودویی کوچک‌ترین تغییرات می‌تواند همانند یک ضربه تلقی شود.
- به طور مثال دو پیکسل مجاور با اندازه‌ی ۱۲۷ و ۱۲۸ با این‌که یک مقدار تفاوت دارند در نمایش دودویی در تمام ۸ بیت اختلاف خواهند داشت.

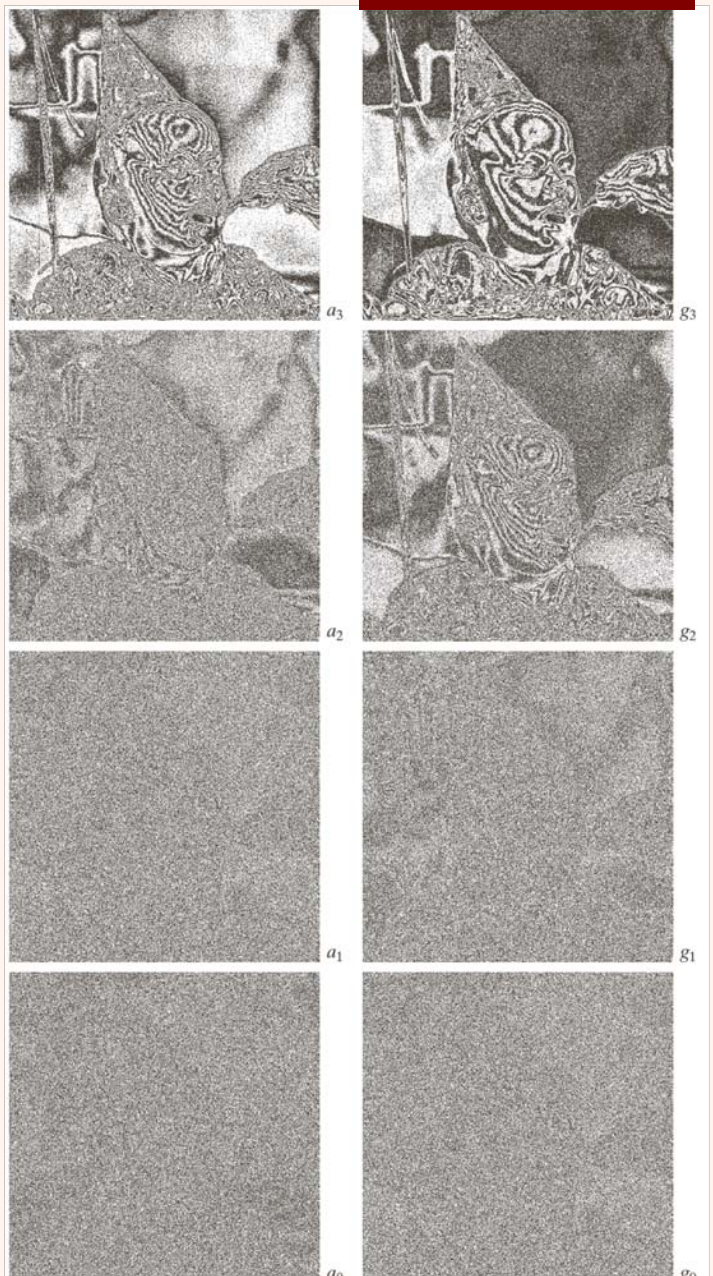
Gonzalez, R. C. and R. E. Woods (2002). Digital image processing

binary

Gray-coded

binary

Gray-coded



فشرده‌سازی

داده‌ی فشرده‌شده پس از بازیابی کاملا با داده‌ی اصلی یکسان است

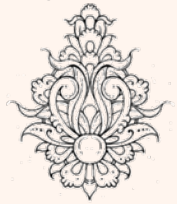
تصاویر پزشکی-تصاویر نظامی

داده‌ی فشرده‌شده پس از بازیابی با داده‌ی اصلی یکسان نیست

نرخ فشردگی بالا

کدگذاری دامنه‌ی تبدیل

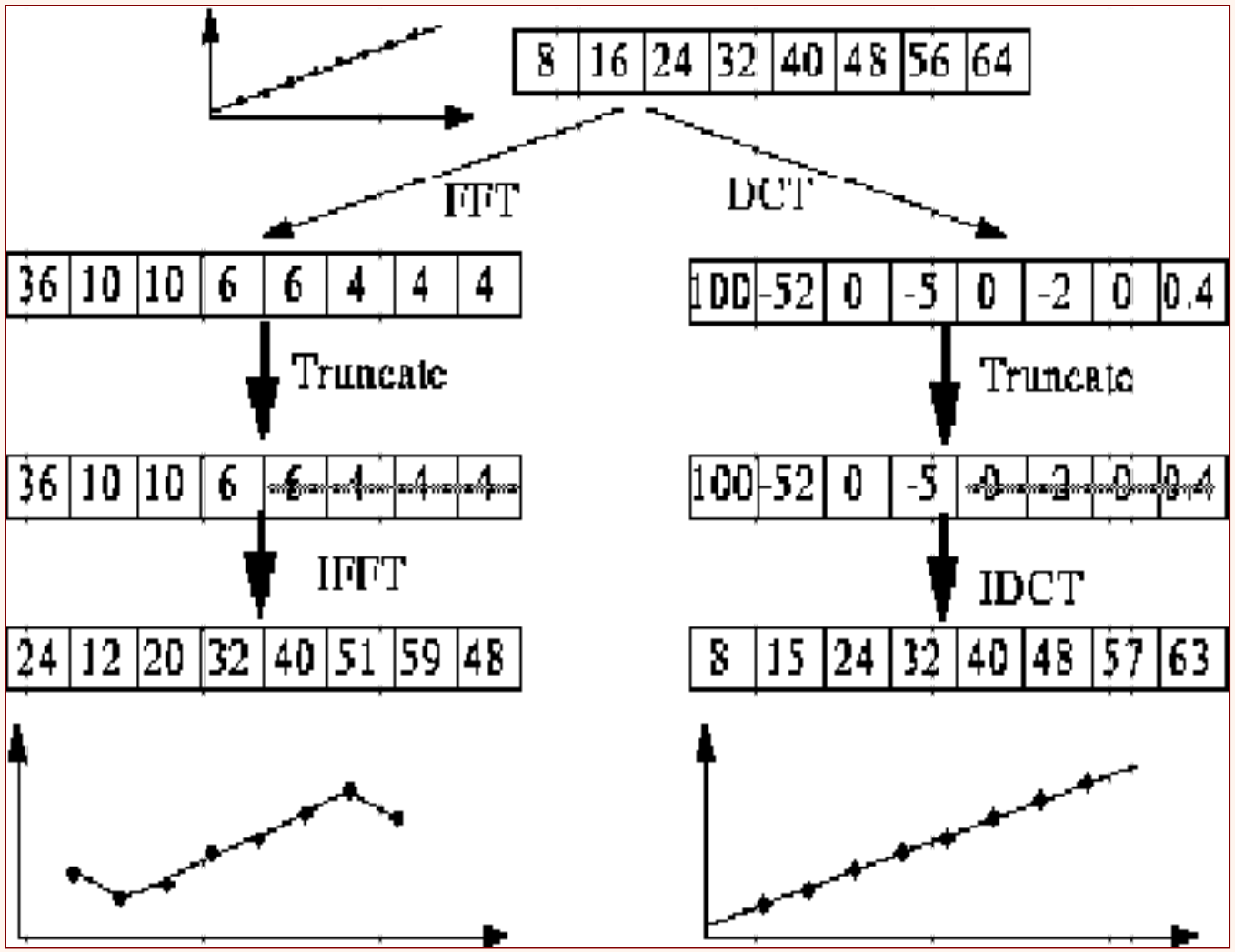
- مراحل
  - اعمال تبدیل به روی تصویر
  - کدگذاری ضرایب تبدیل
  - ارسال داده‌ها
  - بازسازی تصویر با استفاده از داده‌های ارسالی
- مزایا
  - انرژی در ضرایب محدودی فشرده‌شده است.
  - مقاومت در برابر خطاهای کانال ارسالی



FFT /DCT

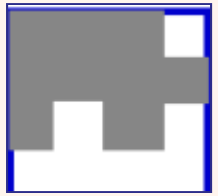
مقادیر ضرایب در DCT حقیقی ولی در DFT مختلطاند

تقریب بهتر در DCT با استفاده از ضرایب کم‌تر

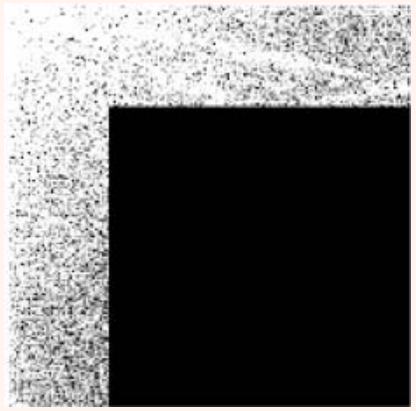


http://www.cs.cf.ac.uk/Dave/Multimedia/node231.html

- کدگذاری
  - ناحیه‌ای (Zonal Coding)
    - تنها ناحیه‌ای از ضرایب ارسال می‌شود.
  - آستانه‌ای (Threshold Coding)
    - تنها ضرایبی ارسال می‌شود که از میزان آستانه بالاتر باشد.



Zonal Coding



120	134	24	17
145	145	230	25
16	234	23	18
23	24	28	19

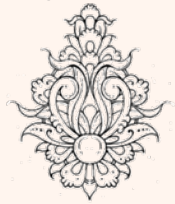
Compression!!!

306.2500	104.8835	-114.7500	-45.3384
100.0207	62.2961	-1.1577	30.1075
-111.7500	-6.2990	99.2500	5.0445
-55.7716	26.1075	-7.3678	-160.7961

چندی کردن

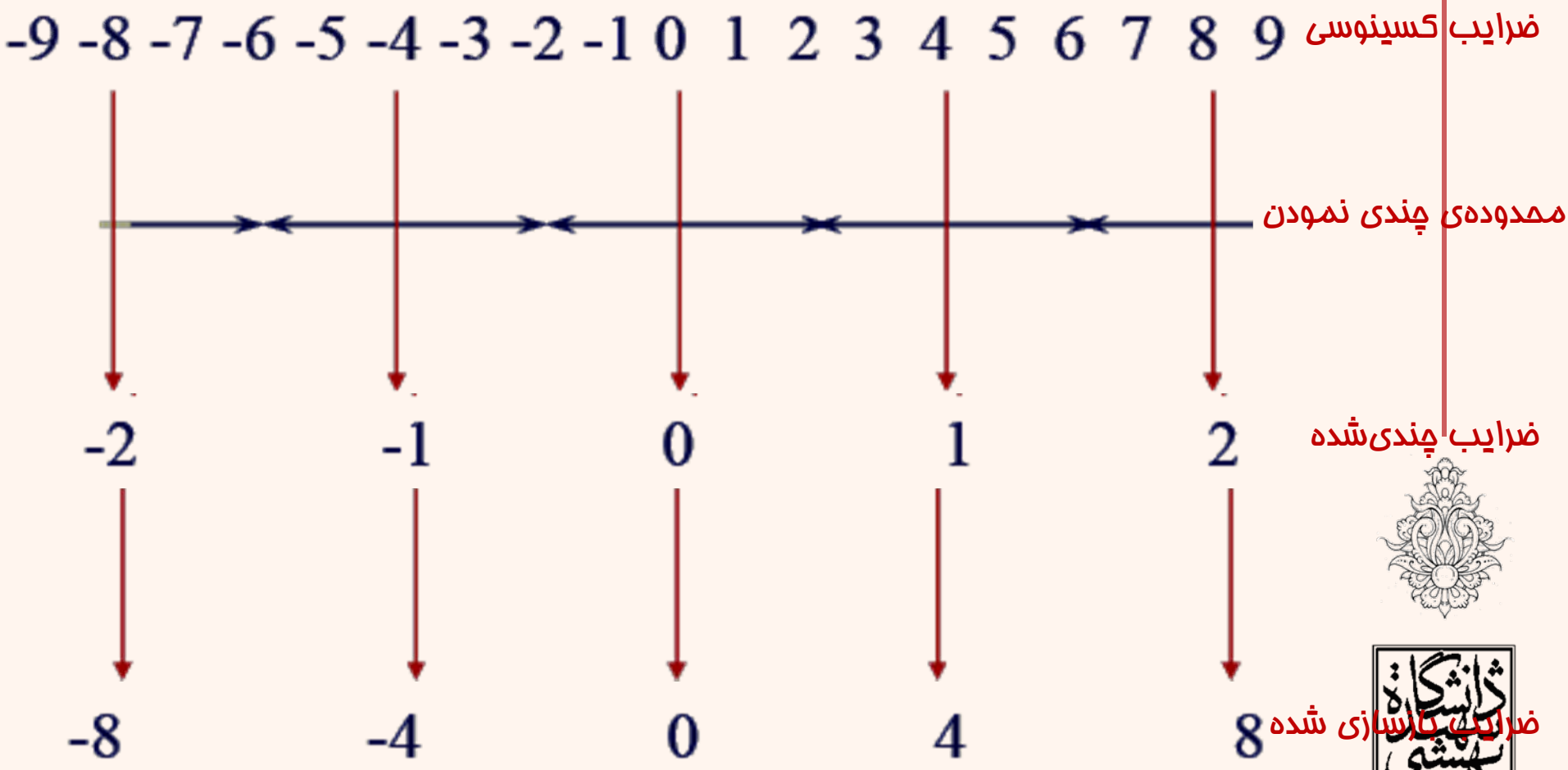
Quantization

- به وسیله‌ی چندی کردن از دقت ضرایب کسینوسی کاسته می‌شود به گونه‌ای که تبدیل به مقادیر صحیح گردند.
- این مساله باعث می‌شود که بسیاری از ضرایب فرکانس‌های بالا به صفر تبدیل گردند.
- آستانه‌ای که با استفاده از آن چندی شدن صورت می‌پذیرد به‌گونه‌ای انتخاب می‌شود که نتیجه‌ی بازسازی تا حد ممکن برای چشم انسان قابل رویت نباشد.



مثال

Quantization=4



$$\begin{bmatrix} -415 & -33 & -58 & 35 & 58 & -51 & -15 & -12 \\ 5 & -34 & 49 & 18 & 27 & 1 & -5 & 3 \\ -46 & 14 & 80 & -35 & -50 & 19 & 7 & -18 \\ -53 & 21 & 34 & -20 & 2 & 34 & 36 & 12 \\ 9 & -2 & 9 & -5 & -32 & -15 & 45 & 37 \\ -8 & 15 & -16 & 7 & -8 & 11 & 4 & 7 \\ 19 & -28 & -2 & -26 & -2 & 7 & -44 & -21 \\ 18 & 25 & -12 & -44 & 35 & 48 & -37 & -3 \end{bmatrix}$$

ضرایب DCT

$$\text{round}\left(\frac{-415}{16}\right) = \text{round}(-25.9375) = -26$$

ماتریس چندی کننده

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

ماتریس چندی کننده استاندارد برای بازسازی استفاده خواهد شد.

$$\begin{bmatrix} -26 & -3 & -6 & 2 & 2 & -1 & 0 & 0 \\ 0 & -3 & 4 & 1 & 1 & 0 & 0 & 0 \\ -3 & 1 & 5 & -1 & -1 & 0 & 0 & 0 \\ -4 & 1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

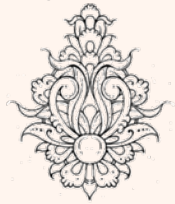
ماتریس نتیجه‌ی چندی شده

محیط‌های چندرسانه‌ای

مراحل الگوریتم jpeg

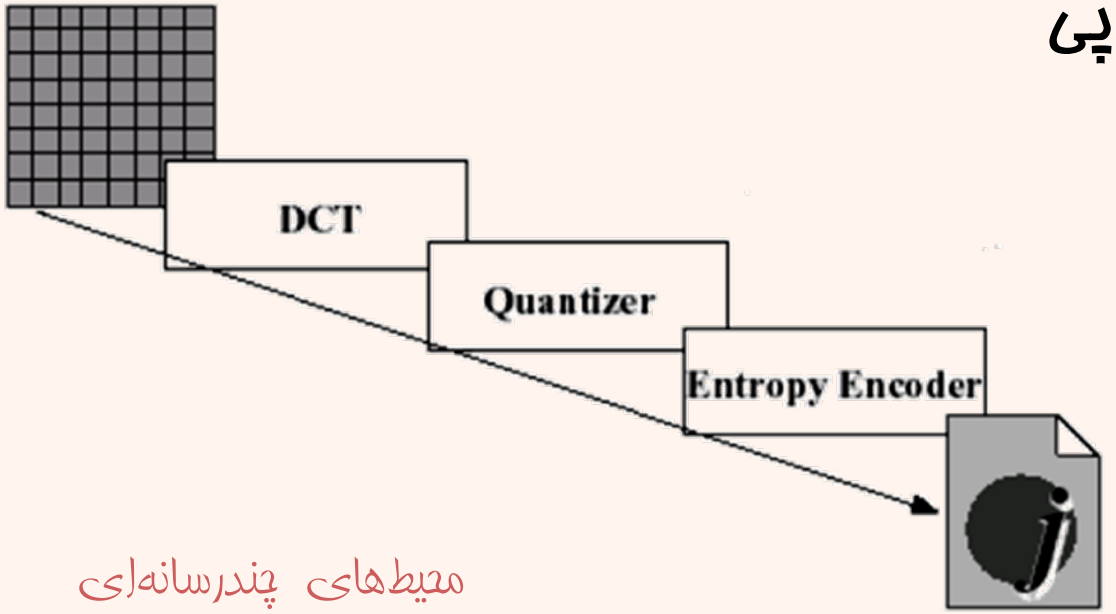
- **بلوک‌بندی تصاویر**
  - تصویر $f(m, n)$ ← زیر تصویرهای ۸×۸ $f_i(m, n)$
  - فرض می‌کنیم مقادیر روشنایی $[0,2^L\text{-}1]$
- **هر f_i با رابطه‌ی زیر به g_i تبدیل می‌شود:**

$$g_i(m, n)=f_i(m, n)-2^{L-1}$$



مراحل الگوریتم jpeg (ادامه...)

- چگونگی الگوریتم فشرده‌سازی
  - بلوک‌بندی تصویر
  - اعمال تبدیل گسسته‌ی کسینوسی
  - چندی نمودن ضرایب
  - پویش زیگزاگ ضرایب
  - کدگذاری آنتروپی
  - ارسال



مراحل الگوریتم jpeg (ادامه...)

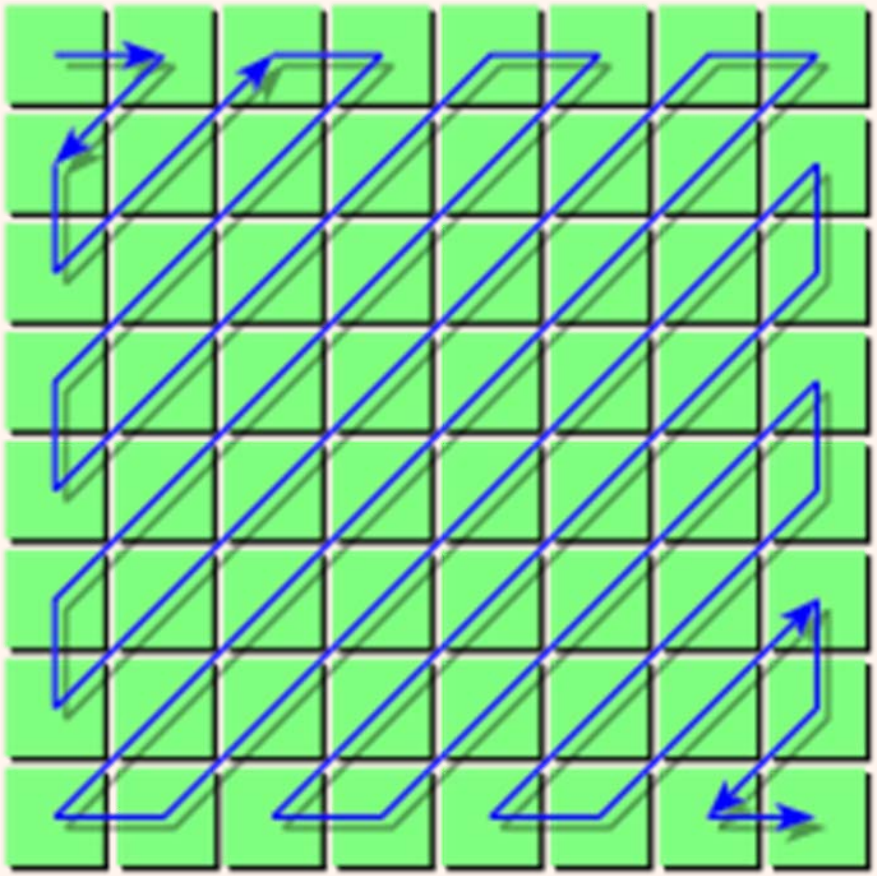
- $g_i(m, n)$ ورودی تبدیل DCT هستند که نتیجه‌ای همانند $D_i(k_1, k_2)$ دارند.
- ماتریس چندی‌کننده‌ای در نظر گرفته شده رابطه‌ی زیر به دست می‌آید:

$$D_{iq}(k1, k2) = round[\frac{D_i(k1, k2)}{T(k1, k2) \times \alpha}]$$

میزان ریزش

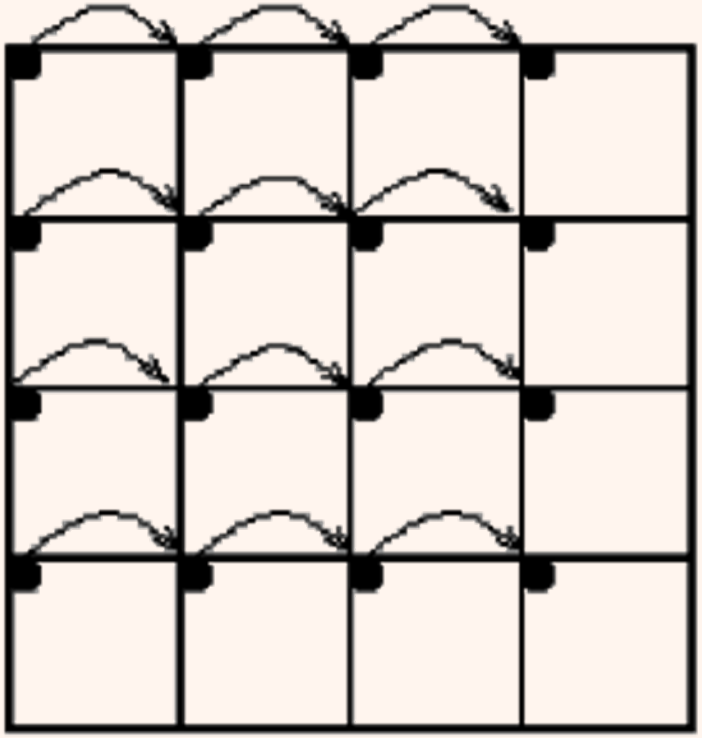
- مرتب نمودن زیگزاگ
- رمزگذاری مقادیر
  - DC –
  - AC –

پویش ضرایب



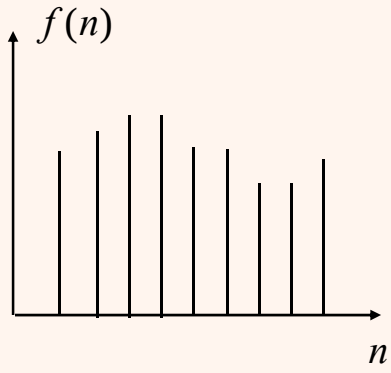
پویش ضرایب به صورت زیگزاگ صورت می‌پذیرد

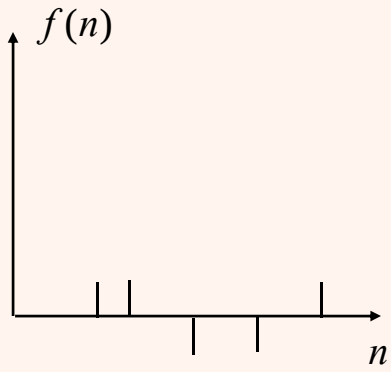
کدگذاری مؤلفه‌های DC



اختلاف مولفه‌های DC بلوک‌های متوالی کد می‌شوند

Differential Pulse Code Modulation (DPCM)





$$f(n) = 156,157,158,158,156,156,154,154,155$$

$$\delta f(n) = 156,1,1,0,-2,0,-2,0,1$$

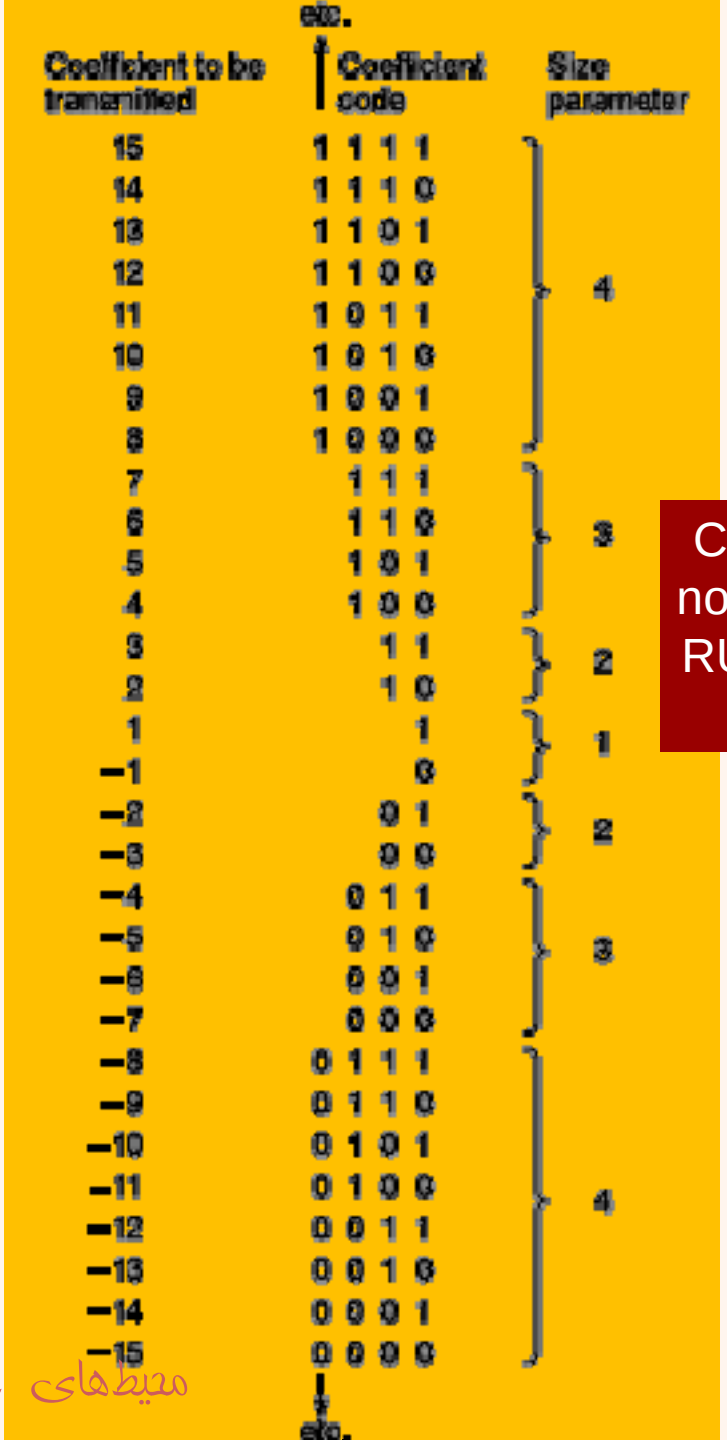
Run-Level coding

- در این شیوه‌ی کد نمودن از تعداد صفرها و عناصر غیر صفر استفاده می‌شود:

102	-33	-3	-4	-2	-1	0	0
21	-2	-3	0	-1	0	0	0
-3	0	1	0	0	0	0	0
2	0	0	0	1	0	0	0
1	0	0	0	0	0	0	0
-2	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

[-33, 21, -3, -2, -3, -4, -3, 0,
2, 1, 0, 1, 0, -2, -1, -1, 0, 0, 0,
-2, 0, 0, 0, 0, 0, 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, 0, 0, 0,
0, 0, 0, 0, 0,0]

(0,-33) (0,21) (0,-3) (0,-2) (0,3)(0,-4) (0,-3) (1,2) (0,1) (1,1) (1,-2) (0,-1) (0,-1) (3,-2) (11,1)



(RUN, CAT)

CAT is the category for the amplitude of a nonzero coefficient in the zigzag order, and RUN is the number of zeros preceding this nonzero coefficient

گروه‌بندی مقادیر

Range	DC Difference Category	AC Category
0	0	N/A
−1, 1	1	1
−3, −2, 2, 3	2	2
−7, …, −4, 4, …, 7	3	3
−15, …, −8, 8, …, 15	4	4
−31, …, −16, 16, …, 31	5	5
−63, …, −32, 32, …, 63	6	6
−127, …, −64, 64, …, 127	7	7
−255, …, −128, 128, …, 255	8	8
−511, …, −256, 256, …, 511	9	9
−1023, …, −512, 512, …, 1023	A	A
−2047, …, −1024, 1024, …, 2047	B	B
−4095, …, −2048, 2048, …, 4095	C	C
−8191, …, −4096, 4096, …, 8191	D	D
−16383, …, −8192, 8192, …, 16383	E	E
−32767, …, −16384, 16384, …, 32767	F	N/A

Table is from slides at Gonzalez/ Woods DIP book website (Chapter 8)

Table is from slides at Gonzalez/ Woods DIP book website (Chapter 8)

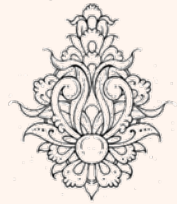
ضرایب DC

Category	Base Code	Length	Category	Base Code	Length
0	010	3	6	1110	10
1	011	4	7	11110	12
2	100	5	8	111110	14
3	00	5	9	1111110	16
4	101	7	A	11111110	18
5	110	8	B	111111110	20

ضرایب AC

Run/ Category	Base Code	Length	Run/ Category	Base Code	Length
0/0	1010 (=EOB)	4			
0/1	00	3	8/1	11111010	9
0/2	01	4	8/2	1111111111000000	17
0/3	100	6	8/3	1111111110110111	19
0/4	1011	8	8/4	1111111110111000	20
0/5	11010	10	8/5	1111111110111001	21
0/6	111000	12	8/6	1111111110111010	22
0/7	1111000	14	8/7	1111111110111011	23
0/8	111110110	18	8/8	1111111110111100	24
0/9	1111111110000010	25	8/9	1111111110111101	25
0/A	1111111110000011	26	8/A	1111111110111110	26
1/1	1100	5	9/1	111111000	10
1/2	111001	8	9/2	1111111110111111	18
1/3	1111001	10	9/3	1111111111000000	19
1/4	111110110	13	9/4	1111111111000001	20

سطوح چندرسانه‌ای



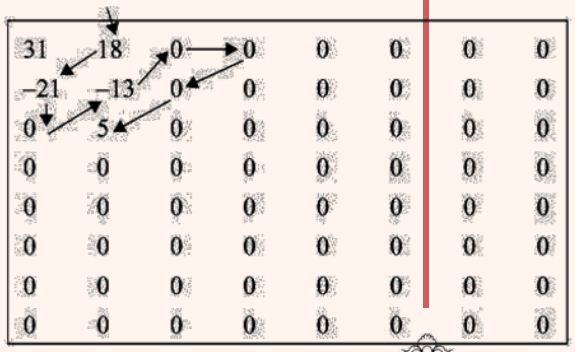
Codeword for the DC coefficient

Category	Base Code	Length
0	010	3
1	011	4
2	100	5
3	00	5
4	101	7
5	110	8

DC

DIFF = 31 - 28 = 3 → CAT = 2

CAT is 100 → DIFF = 3 > 0 → the DC coefficient is 10011



Codeword for the AC coefficients

AC=18

the codeword for (0, 5) is 11010 → 1101010010

AC

Run/ Category	Base
0/0	1010 (=
0/1	00
0/2	01
0/3	100
0/4	1011
0/5	1010

AC=-21

the codeword for (0, 5) is 11010 → 1101001010

Range	DC Difference Category	AC Category
0	0	N/A
−1, 1	1	1
−3, −2, 2, 3	2	2
−7, ..., −4, 4, ..., 7	3	3
−15, ..., −8, 8, ..., 15	4	4
−31, ..., −16, 16, ..., 31	5	5

محیط‌های چندرسانه‌ای

کدگذاری (ادامه...)

-26 -3 1 -3 -2 -6 2 -4 1 -4 1 1 5 0 2 0 0 -1 2 0 0 0 0 0 -1 -1 EOB

-3	(0/2)=0100
1	(0/1)=001
-3	(0/2)=0100
-2	(0/2)=0101
-6	(0/3)=100001
2	(0/2)=0110

-4	(0/3)=100011
1	(0/1)=001
-4	(0/3)=100011
1	(0/1)=001
1	(0/1)=001
5	(0/3)=100101

2	(1/2)=11100110
-1	(2/1)=110110
2	(0/2)=0110
-1	(5/1)=11110100
-1	(0/1)=000
EOB	1010

قطار بیت ارسالی

0100 001 0100 0101 100001 0110 100011 001 100011 001 001 100101 11100110
110110 0110 11110100 000 1010

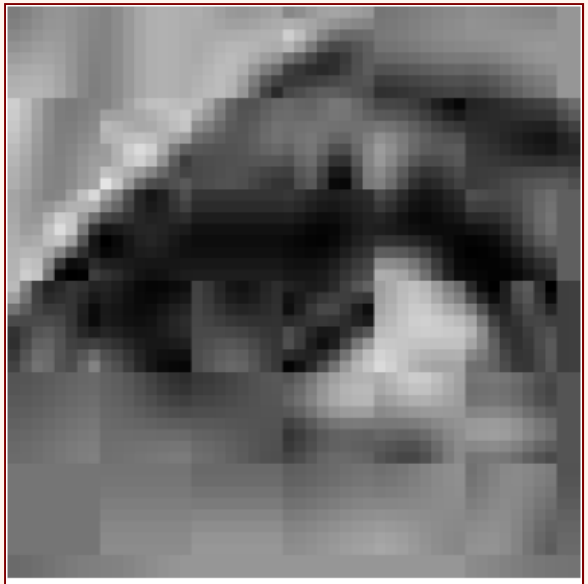
تعداد بیت‌های ارسالی 85 بیت

Ghanbari, M. a (2003). Standard codecs: image compression to advanced video coding, Institution of Electrical Engineers.

Original
bpp = 8.00 mse = 0.00

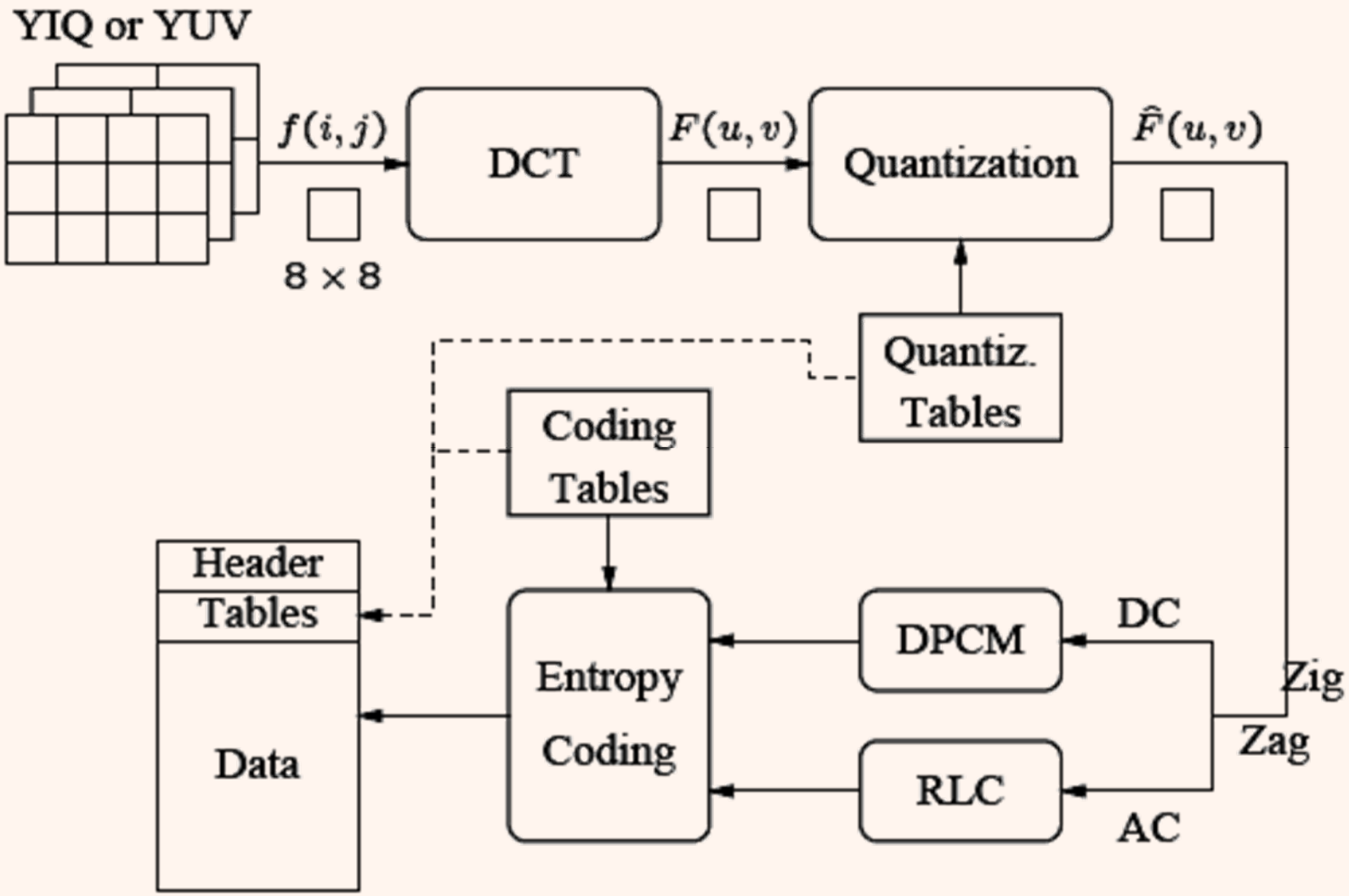
JPEG
bpp = 1.00 mse = 17.26

JPEG
bpp = 0.50, mse = 33.08

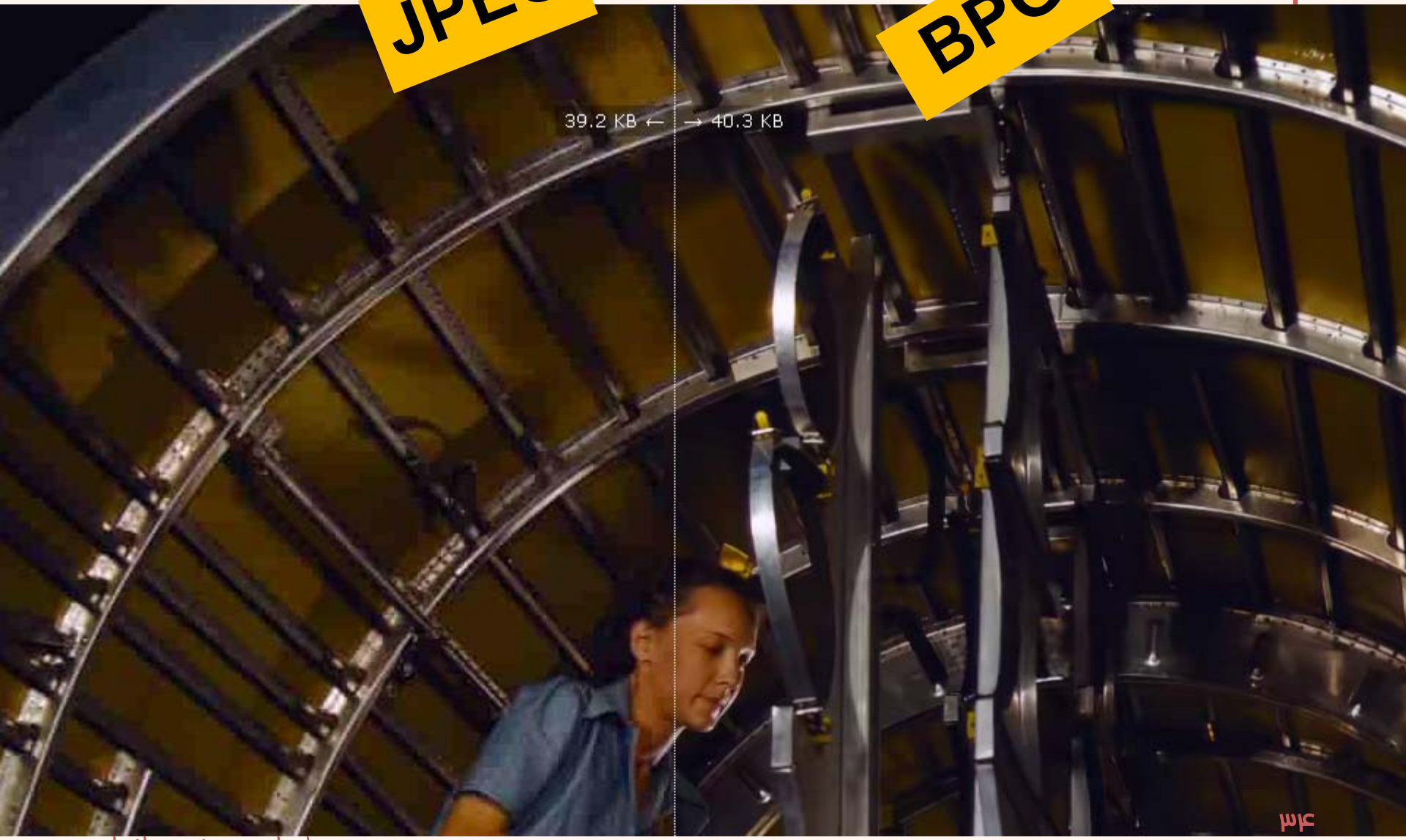


JPEG
bpp = 0.25, mse = 79.11

Block diagram for JPEG encoder.



استانداردهای جدید



استانداردهای جدید

The blocking, color banding and aliasing artifacts of heavily compressed JPEG (left), while the heavily compressed BPG (right) looks much smoother

محیط‌های چندرسانه‌ای