

## STANDARDI ZA KOMPRESIJU SLIKE

Vremenom se pokazalo da je neophodno izvršiti standardizaciju metoda za kompresiju slike što bi omogućilo komunikaciju između korisnika koji koriste opremu i softver različitih proizvođača. Pokušaji velikih proizvođača opreme i softvera da nametnu neka svoja rešenja kao standarde nisu uspjeli, iako se i ta rješenja koriste u praksi (npr. GIF i TIFF grafički formati). Međutim, njihova primjena na kompresiju prirodnih slika, gray-scale ili u boji, ne daje dobre rezultate, jer su oni prvenstveno namijenjeni kompresiji crteža i tekstualnih dokumenata.

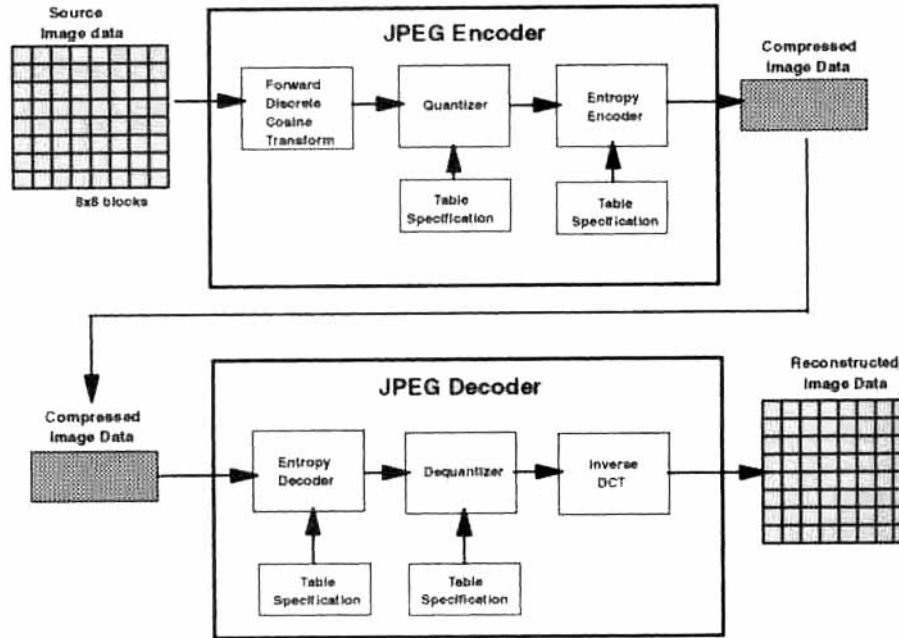
Internacionalna organizacija za standarde (ISO), Internacionalna elektrotehnička komisija (IEC) i Internacionalna telekomunikaciona unija (ITU) rade na standardizaciji metoda, hardvera i softvera za multimedijalne sisteme, videokonferencije, videotelefoniju i slične aplikacije. Do sada se pojavilo više standarda (JPEG za kompresiju mirnih slika, MPEG-1, MPEG-2, MPEG-4, H.261, H.263, za kompresiju video sekvenci, itd.).

### JPEG standard

JPEG (Joint Photographic Experts Group) je standard za kompresiju mirnih slika sa više nivoa svjetline, kao i slika u boji. To je zajednički standard tri međunarodne organizacije: ISO, IEC i ITU. On omogućava kompresiju bilo kakve digitalne slike, sive ili u boji, sa gubicima ili bez gubitaka, nezavisno od njene rezolucije. Postoje četiri načina kompresije po JPEG standardu, od kojih ćemo mi obraditi samo prvi navedeni:

- *Sekvencijalno DCT kodovanje*: slika se koduje jednim prolazom kroz sliku, s lijeva na desno, pa odozgo prema dole.
- *Progresivno DCT kodovanje*: slika se koduje u više prolaza, s ciljem da korisnik što prije dobije makar i grubu sliku, a zatim sve finiju i finiju strukturu slike.
- *Kodovanje bez gubitaka*: garantuje se perfektna rekonstrukcija originalne slike, ali je stepen kompresije mali.
- *Hijerarhijsko kodovanje*: slika se koduje različitim rezolucijama, pri čemu za rekonstrukciju slike niže rezolucije nije potrebno poznavanje podataka o kodovanju slike više rezolucije.

Blok dijagram sekvencijalnog kodera i dekodera je prikazan na Slici 116.



Slika 116. Blok šema JPEG sekvencijalnog koda i dekodera

## JPEG koder

Koder se sastoji od tri osnovna bloka:

1. Blok za računanje direktne diskretne kosinusne transformacije (FDCT),
2. Kvantizer,
3. Entorpijski koder.

Na ulazu koda, originalne vrijednosti piksela, koje su pozitivni cijeli brojevi iz opsega  $[0, 2^p-1]$ , se pomijeraju u opseg  $[-2^{p-1}, 2^{p-1}-1]$ . Na primjer, za gray-scale slike kod kojih je  $p=8$ , originalni odmjerci iz opsega  $[0, 255]$  se oduzimanjem 128 pomijeraju u opseg  $[-128, +127]$ .

Zatim se slika dijeli na blokove dimenzija  $8 \times 8$ . Dimenzija blokova je odabrana na bazi opsežnih ispitivanja subjektivnog osjećaja kvaliteta kompresovane slike sa različitim deimenzijama blokova. Svaki od blokova se transformiše diskretnom kosinusnom transformacijom u skup od 64 DCT koeficijena koji predstavljaju "dvodimenzioni spektar" ulaznog signala:

$$F[k,l] = \frac{\alpha[k]\alpha[l]}{p} \sum_{m=0}^{p-1} \sum_{n=0}^{p-1} f[m,n] \cos \frac{\pi k(2m+1)}{2p} \cos \frac{\pi l(2n+1)}{2p}, \quad 0 \leq k, l \leq p-1$$

gdje je  $p = 8$  i

$$\alpha[j] = \begin{cases} 1/\sqrt{2}, & j = 0, \\ 1 & 1 \leq j \leq p-1. \end{cases}$$

Diskretna kosinusna transformacija je diskretna funkcija dvije prostorne dimenzije,  $k$  i  $l$ , koje nazivamo prostornim frekvencijama. Vrijednosti ove funkcije nazivamo DCT koeficijentima. Koeficijent  $F(0,0)$  se naziva DC koeficijent, dok se preostala 63 koeficijenta nazivaju AC koeficijenti. Za sive slike DCT koeficijenti su iz opsega vrijednosti  $[-1024, 1023]$ , što znači da u poređenju sa originalnim vrijednostima piksela slike trebamo 3 bita više za njihov zapis.

Za tipičan 8x8 blok većina DCT koeficijenata je jednaka nuli, ili ima vrijednost blisku nuli, odnosno znatno manju od koeficijenata na niskim prostornim frekvencijama, Slika 117, te ih ne treba kodovati. Ova osobina se koristi za postizanje kompresije slike.

140	144	147	140	140	155	179	171	185	-17	14	-8	23	-9	-13	-18
144	152	140	147	140	148	167	179	20	-34	26	-9	-10	10	13	6
152	155	136	167	163	162	152	172	-10	-23	-1	8	-18	3	-20	0
168	145	156	160	152	155	136	160	-8	-5	14	-14	-8	-2	-3	8
162	148	156	148	140	136	147	162	-3	8	7	1	-11	17	18	15
147	167	140	155	155	140	136	162	3	-2	-18	8	8	-3	0	-8
136	156	123	167	162	144	140	147	8	0	-2	3	-1	-7	-1	-1
148	155	136	165	152	147	147	136	0	-7	-2	1	1	4	-6	0

Slika 117. Matrica vrijednosti originalne slike i DCT koeficijenti

U sljedećem bloku, kvantizeru, 64 DCT koeficijenta se kvantuju koristeći kvantizacione matrice (zavisno od aplikacije) koje redukuju amplitudu koeficijenata koji malo ili nimalo ne utiču na kvalitet slike, s ciljem da se poveća broj nultih vrijednosti u matrici DCT koeficijenata. Kvantizacijom se gube informacije koje nisu vizualno od značaja. Kvantovanje se vrši dijeljenjem svakog DCT koeficijenta sa odgovarajućim elementom neke od kvantizacionih matrica (Slika 118) i zaokruživanjem na najbliži celi broj:

$$F_q(k,l) = \text{Round} \left[ \frac{F(k,l)}{Q(k,l)} \right].$$

8	6	5	8	12	20	26	30
6	6	7	10	13	29	30	28
7	7	8	12	20	29	35	28
7	9	11	15	26	44	40	31
9	11	19	28	34	55	52	39
12	18	28	32	41	52	57	46
25	32	39	44	52	61	60	51
36	46	48	49	56	50	52	50

9	9	12	24	50	50	50	50
9	11	13	33	50	50	50	50
12	13	28	50	50	50	50	50
24	33	50	50	50	50	50	50
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50

16	17	18	19	20	21	22	23
17	18	19	20	21	22	23	24
18	19	20	21	22	23	24	25
19	20	21	22	23	24	25	26
20	21	22	23	24	25	26	27
21	22	23	24	25	26	27	28
22	23	24	25	26	27	28	29
23	24	25	26	27	28	29	30

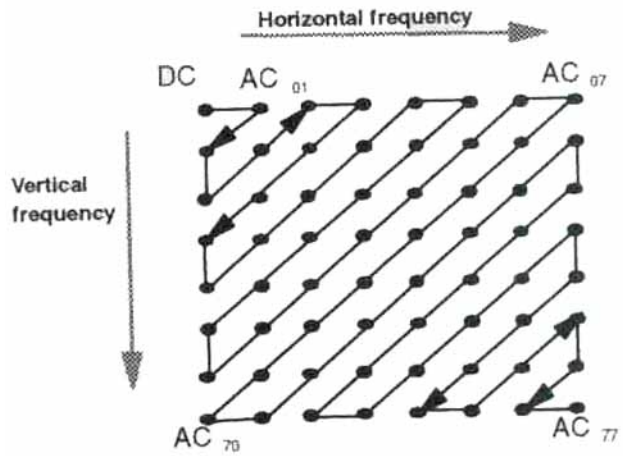
16	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

Slika 118. Standardne kvantizacione matrice

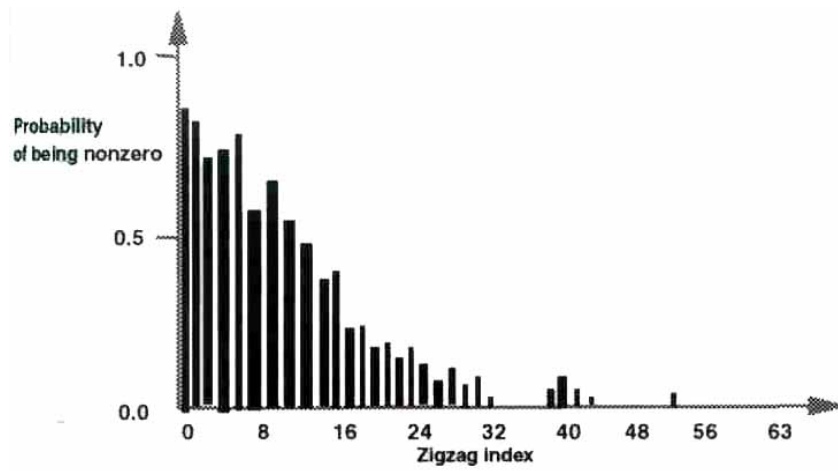
Kvantovanje se vrši sa različitim korakom kvantovanja za svaki DCT koeficijent. Za maksimalnu kompresiju slike bez vidljivih izobličenja, svaka vrijednost u kvantizacionoj matrici mora da bude izabrana tako da predstavlja prag percepcije odgovarajuće kosinusne bazisne funkcije. Sproveden je niz psihovizuelnih eksperimenata za određivanje najboljih vrednosti pragova. Sam standard ne propisuje sadržaj kvantizacionih matrica, ali daje određene tabele kao preporuke.

Za regulaciju stepena kompresije uveden je faktor kvaliteta  $Q$ , cio broj kojim se množe elementi kvantizacione matrice. Tako se omogućava jednostavno, ali grubo, zadavanje željenog stepena kompresije, pri čemu veće vrijednosti parametra  $Q$  odgovaraju grubljem kvantovanju, odnosno većem stepenu kompresije. Većem stepenu kompresije odgovara lošiji kvalitet rekonstruisane slike.

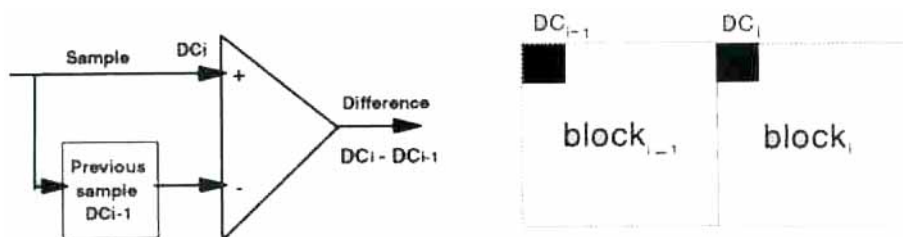
Nakon kvantizacije od 63 AC koeficijenta se formira jednodimenzionalna sekvenca, cik-cak redoslijedom, kao na Slici 119. Ovakav poredak koeficijenata olakšava entropijsko kodovanje jer su niskofrekventni koeficijenti, čija je vjerovatnoća da su različiti od nule veća nego kod visokofrekventnih, smješteni na početak. Ova činjenica je potvrđena nizom eksperimentalnih ispitivanja. Dobiveni rezultati pokazuju da je, u slučaju cik-cak poretka DCT koeficijenata, vjerovatnoća da su koeficijenti jednaki rayli;iti od nule monotono opadajuća funkcija indeksa, Slika 120. DC koeficijenti, koji predstavljaju srednju vrijednost piksela u bloku dimenzija 8x8, koduju se prediktivnim tehnikama, Slika 121, jer postoji velika prostorna korelacija između srednjih vrijednosti svjetline u susjednim blokovima. Zbog toga se kodovanjem razlike vrijednosti DC koeficijenata postiže veći stepen kompresije nego kodovanjem samih koeficijenata.



Slika 119. Cik-cak poredak AC koeficijenata



Slika 120. Vjerovatnoća da su DCT koeficijenti različiti od nule



Slika 121. Prediktivno kodovanje DC koeficijenata

Konačno, poslednji blok JPEG koda je entropijski koder, kojim se postiže dodatna kompresija kodujući kvantovane DCT koeficijente. Nakon kvantovanja, Hafmanov koder konvertuje DCT koeficijente u kompaktnu binarnu sekvencu kroz dva koraka: (1) formira se tabela simbola, (2) na osnovu Hafmanove tabele konvertuju se simboli u binarnu sekvencu.

U tabeli simbola, svaki AC koeficijent se predstavlja parom simbola:

- (DUŽINA\_NIZA, VELIČINA)
- (AMPLITUDA)

DUŽINA\_NIZA je broj koji pokazuje koliko AC koeficijenata nultih vrijednosti prethodi AC koeficijentu nenulte vrijednosti. Vrijednost za DUŽINU\_NIZA je u opsegu 0-15, te su potrebna 4 bita za njenu reprezentaciju.

VELIČINA je broj bita koji je potreban da se koduje AMPLITUDA. AMPLITUDA se koduje sa 0-10 bita, što znači da se neophodna 4 bita za kodovanje VELIČINE.

AMPLITUDA je vrijednost nenulatih AC koeficijenata koja se kreće u opsegu [-1024, 1023], te je potrebno maksimalno 10 bita za njeno kodovanje.

Na primjer, ako je sekvenca AC koeficijenata:

0,0,0,0,0,0,476

simbol kojim se predstavlja nenulti koeficijent 476 je

(6,9)(476)

gdje je: DUŽINA\_NIZA=6, VELIČINA=9 i AMPLITUDA=476.

Ako je DUŽINA\_NIZA veća od 15, simbol (15,0) se interpretira kao da je DUŽINA\_NIZA=16. Uzastopno se mogu pojaviti do tri ovakva simbola.

U sljedećem primjeru:

(15,0)(15,0)(7,4)(12)

DUŽINA\_NIZA je jednaka  $16+16+7=39$ , VELIČINA=4 i AMPLITUDA=12.

Simbolom (0,0) se završava svaki 8x8 blok.

Za DC koeficijente se tabela simbola sastoji od:

- VELIČINA
- AMPLITUDA

Kako se DC koeficijenti diferencijalno koduju, opseg vrijednosti je dvostruko veći nego kod AC koeficijenata, [-2048, 2047].

U sljedećem koraku se sekvenca simbola konvertuju u binarnu sekvencu. Simboli se zamjenjuju kodnim riječima promjenljive dužine, počevši od DC koeficijenta, a zatim AC koeficijenti u cik-cak poretku. Svaki simbol kojim je predstavljen AC ili DC koeficijent se koduje na osnovu Hafmanove tabele.

Na primjer, AC koeficijent predstavljen simbolom (1,4)(12) se koduje binarnom sekvencom (1111101101100), gdje je (111110110) kod za (1,4) očitao iz Hafmanove tabele (koja je preporučena JPEG standardom), slika 110, a (1100) je kod 12.

## JPEG dekodier

Kod JPEG sekvencijalnog dekodovanja, svi koraci iz procesa kodovanja se implementiraju obrnutim redoslijedom, slika 103. Prvo se binarna sekvenca konvertuje u sekvencu simbola koristeći Hafmanovu tabelu, a zatim se simboli konvertuju u DCT koeficijente. Zatim se vrši dekvantizacija po sljedećoj formuli:

$$F_q^i(u, v) = F_q(u, v) \times Q(u, v)$$

gdje su  $Q(u, v)$  kvantizacioni koeficijenti iz kvantizacione tabele.

Poslije toga se uradi inverzna diskretna kosinusna transformacija (IDT), kojom se slika iz 2D frekventnog domena prevodi u prostorni domen. Inverzna diskretna kosinusna transformacija je definisana sa:

$$f[m, n] = \frac{1}{4} \sum_{k=0}^7 \sum_{l=0}^7 \alpha[k] \alpha[l] F[k, l] \cos \frac{\pi k(2m+1)}{2p} \cos \frac{\pi l(2n+1)}{2p}$$

gdje je:

$$\alpha[j] = \begin{cases} 1/\sqrt{2}, & j = 0, \\ 1 & 1 \leq j \leq p-1. \end{cases}$$

Posljednji korak predstavlja pomijeranje dekompresovanih odmjera u opseg  $[0, 2^p-1]$ .

### Primjer

Za ilustraciju principa sekvencijalnog JPEG kodovanja, prikazaćemo proceduru korak po korak na jednom 8x8 bloku 8-bitnih uzoraka. Na slici 109 (a) su prikazane vrijednosti originalnog bloka (opseg vrijednosti je  $[0, 255]$ ), a na slici 109 (b) je isti blok nakon pomijeranja u opseg  $[-128, +127]$ . Primjenom diskretne kosinusne transformacije dobiju se DCT koeficijenti prikazani na slici 109 (c). Primijetimo da su skoro sve vrijednosti koeficijenata, izuzev onih na niskim frekvencijama, bliske nuli.

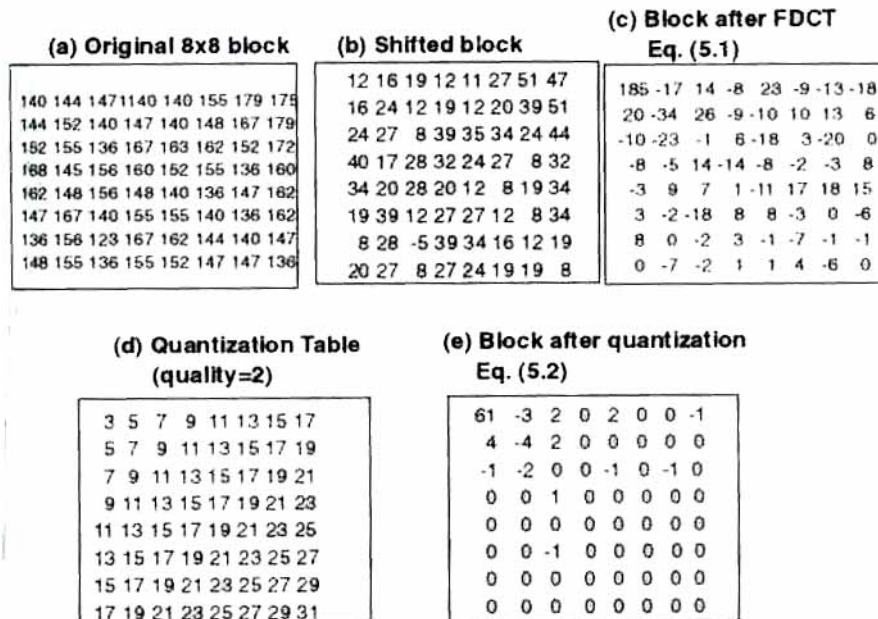
Za generisanje kvantizacione tabele se preporučuje sljedeći program:

```

for(i = 0; i < N; i++)
  for(j = 0; j < N; j++)
    Q(i,j) = 1+[(1+i+j) x Quality];

```

Pri tome faktor *Quality* specificira kvalitet rekonstruisane slike. *Quality* = 1 daje najbolji kalitet, ali i najniži stepen kompresije. U ovom primjeru je korišten faktor kvaliteta 2, i kvantizaciona tabela je prikazana na Slici 122 (d). Kvantovani DCT koeficijenti su prikazani na Slici 122 (e). Primijetimo da je većina AC koeficijenata na višim frekvencijama jednaka nuli. Sada se od DCT koeficijenata formira cik-cak sekvenca, prikazana na Slici 122 (f). Koeficijent različiti od nule su grupisani na početku sekvence, a zatim slijede duge sekvence nula. Zbog načina generisanja simbola, postiže se veća kompresija nego kad bi imali kraće sekvence nula isprekidane vrijednostima različitim od nule. Tabela simbola je prikazana na Slici 122 (g). Konačno, nakon implementacije Hafmanovog kodera, binarna sekvenca je prikazana na Slici 122 (e).



**(f) Zig-zag sequence**

61,-3,4,-1,-4,2,0,2,-2,0,0,0,0,0,2,0,0,0,1,0,0,0,0,0,0,-1,0,0,-1,0,0,  
0,0,-1,0,0,0,0,0,0,-1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0

**(g) Intermediate symbol sequence**

(6)(61),(0,2)(-3),(0,3)(4),(0,1)(-1),(0,3)(-4),(0,2)(2),(1,2)(2),(0,2)(-2),  
(0,2)(-2),(5,2)(2),(3,1)(1),(6,1)(-1),(2,1)(-1),(4,1)(-1),(7,1)(-1),(0,0)

**(e) Encoded bit sequence (total 98 bits)**

(110)(111101) (01)(00) (100)(100) (00)(0) (100)(001) (01)(10)  
(11011)(10) (01)(01) (01)(01) (1111110111)(10) (111010)(1)  
(1111011)(0) (11100)(0) (111011)(0) (11111010)(0) (1010)

Slika 122. [16] Ilustracija JPEG kodovanja



Dio Hafmanove tabele je prikazan na slici 123.

Zero Run	Category	Code length	Codeword
0	1	2	00
0	2	2	01
0	3	3	100
0	4	4	1011
0	5	5	11010
0	6	6	111000
0	7	7	1111000
.	.	.	.
1	1	4	1100
1	2	6	111001
1	3	7	1111001
1	4	9	111110110
.	.	.	.
.	.	.	.
2	1	5	11011
2	2	8	11111000
.	.	.	.
.	.	.	.
3	1	6	111010
3	2	9	111110111
.	.	.	.
.	.	.	.
4	1	6	111011
5	1	7	1111010
6	1	7	1111011
7	1	8	11111001
8	1	8	11111010
9	1	9	111111000
10	1	9	111111001
11	1	9	111111010
.	.	.	.
.	.	.	.
EOT	.	4	1010

Slika 123. Hafmanova tabela za JPEG kodovanje

Napomenimo da je DC koeficijent kodovan kao koeficijent prvog bloka, što znači da nije korišteno prediktivno kodovanje kao za DC koeficijente iz ostalih blokova, već je kodovan direktno.

Za ovaj primjer, stepen kompresije je:

$$C_R = \frac{\text{originalan broj bita}}{\text{kodovani broj bita}} = \frac{64 \times 8}{98} = \frac{512}{98} = 5.22$$

a broj bita po pikselu iznosi:

$$N_b = \frac{\text{kodovani broj bita}}{\text{broj piksela}} = \frac{98}{64} = 1.53$$

## Mjera kompresije

Kad govorimo o preciznosti slike, mislimo na broj nivoa svjetline koje je moguće prikazati. Pri tome se preciznost izražava kao broj bita/pikselu. Na osnovu ovog parametra, može se izvršiti klasifikacija originalne (nekomprimovane) slika na:

- *binarne slike* predstavljene sa 2 bita/pikselu,
- *računarsku grafiku* sa 4 bita/pikselu,
- *grayscale* slike sa 8 bita/pikselu,
- *slike u boji* sa 16, 24 ili više bita/pikselu.

Osnovna mjera pri poređenju različitih algoritama za kompresiju je:

$$C_R = \frac{\text{originalna količina podataka}}{\text{količina podataka nakon kompresije}}$$

Prilikom kompresije slike mora se tražiti kompromisno rješenje između stepena kompresije i kvaliteta slike. Pri tome, odnos kompresije i kvaliteta varira zavisno od karakteristika slike i od toga šta se nalazi na sceni. Kao mjera kvaliteta slike usvaja se broj bita po pikselu u komprimovanoj slici, definisan kao odnos ukupnog broja bita u komprimovanoj slici i broja piksela:

$$N_b = \frac{\text{broj bita nakon kodovanja}}{\text{broj piksela}}$$

Klasifikacija je prikazana u sljedećoj tabeli:

Broj bita/pikselu	Kvalitet slike
0,25 – 0,5	Srednji do dobar
0,5 - 0,75	Dobar do veoma dobar
0,75 – 1,5	Izuzetan
1,5 – 2	Najčešće nema razlike u odnosu na original

Druga statistička mjera koja se može koristiti pri razvoju algoritama za kompresiju slike je srednjekvadratna greška:

$$RMS = \frac{1}{n} \sqrt{\sum_{i=1}^n (X_i - \hat{X}_i)^2}$$

gdje je:

$X_i$  – originalna vrijednost piksela,

$\hat{X}_i$  – vrijednost piksela nakon kompresije,

$n$  – ukupan broj piksela.

U većini slučajeva, ali ne i uvijek, kvalitet slike je bolji ako je srednjekvadratna greška manja.

*Primjer:*

Na Slici 124 (a) je prikazana gray-scale slika *Cvijeće*, dimenzija 500×362 piksela i kodovana sa 8 bita/pikselu. Nakon JPEG kompresije, na Slici 124 (b) je predstavljena slika komprimovana 4 puta, na Slici 124 (c) je stepen kompresije jednak 10, dok je na Slici 124 (d) prikazana slika koja je komprimovana 22 puta. Tek kad stepen kompresije pređe 20 mogu se uočiti vidljiva izobličenja. Kod slika sa manje detalja stepen kompresije može biti i veći od 20 a da se izobličenja ne primjećuju. Međutim, kod slika sa mnogo detalja, (npr. otiska prsta), izobličenja se primjećuju već kad stepen kompresije pređe vrijednost 10.



(a)



(b)



(c)



(d)

Slika 124. [15] JPEG kompresija sive slike *Cvijeće*: (a) Originalna slika.  
(b) Slika komprimovana 4 puta. (c) Slika komprimovana 10 puta.  
(d) Slika komprimovana 22 puta.