



Diplomski studij

Informacijska i
komunikacijska tehnologija:

Obradba informacija
Telekomunikacije i informatika

Višemedijske komunikacije

5.

Informacijska svojstva i
kodiranje nepomične slike

Informacijska svojstva i kodiranje nepomične slike



Zavod za
telekomunikacije

- Izvori digitalne slike
- Ljudski vid i percepcija slike
- Računalni prikaz slike
- Principi kodiranja slike
- Kodiranje bez gubitaka
- Kodiranje sa gubicima
 - Diferencijalno kodiranje
 - Transformacijsko kodiranje: JPEG
 - Kodiranje valićima: JPEG 2000
 - Fraktalno kodiranje

Digitalna (nepomična, 2D) slika

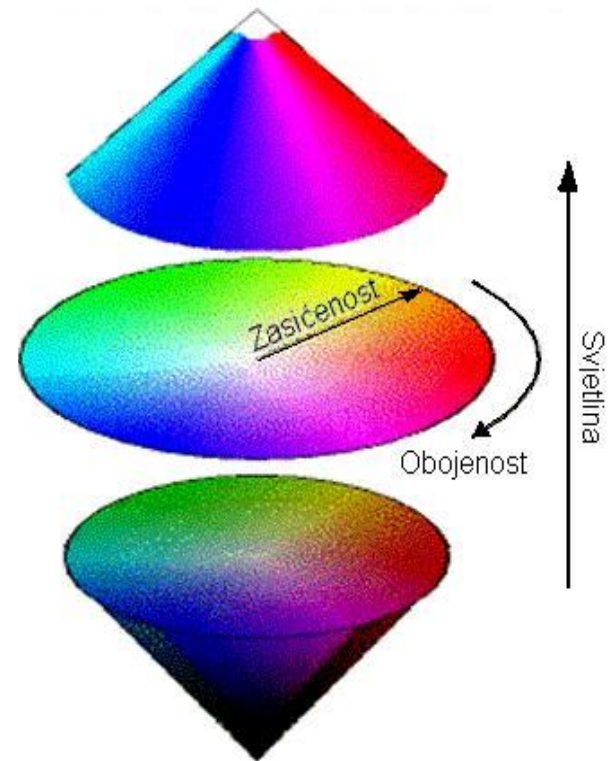
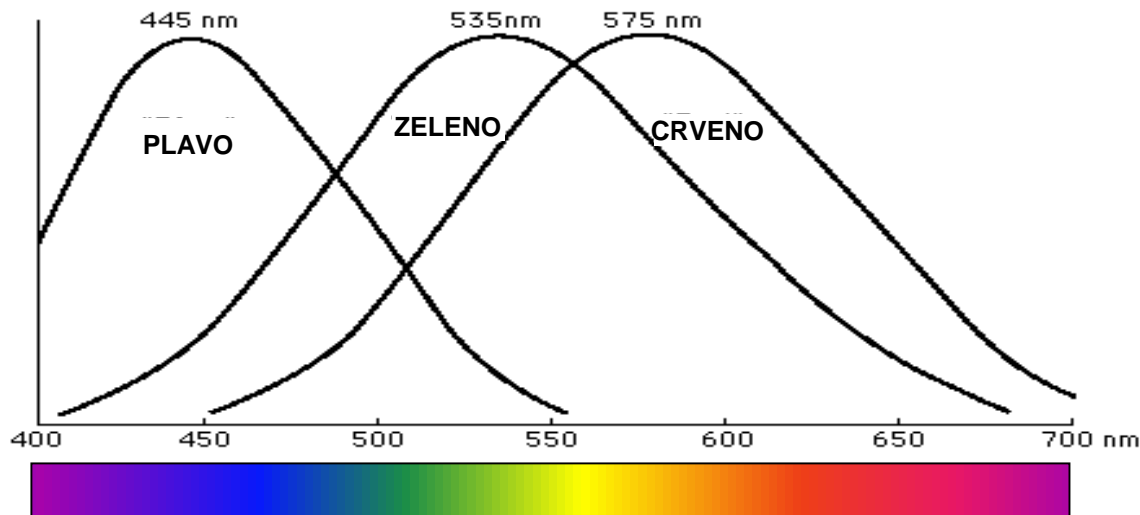


Zavod za
telekomunikacije

- Slika iz stvarnog svijeta (u digitalnom obliku):
 - slika snimljena digitalnom kamerom
 - slika prenesena na računalo preko optičkog čitača (*scanner*)
 - ...
- Slika stvorena pomoću računala:
 - crtež (vektorska grafika)
 - slika (obrađena digitalna slika, bitmap, računalna grafika)
 - fraktalna slika
 - graf
 - vizualizacija
 - ...
- Slika kao jedinica unutar animiranog filma ili videa

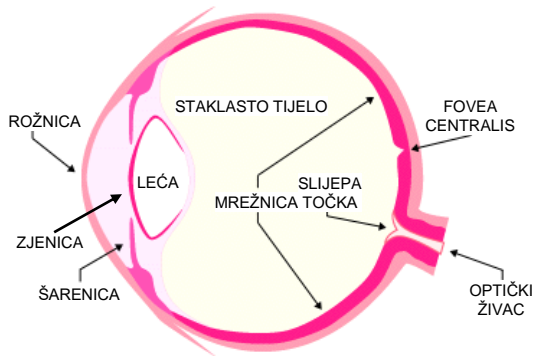
Percepcija slike

- Vidljiva svjetlost = elektromagnetski val
 - $\lambda = 330 - 770 \text{ nm}$, $c \approx 3 \times 10^8 \text{ m/s}$ (ovisno o gustoći medija)



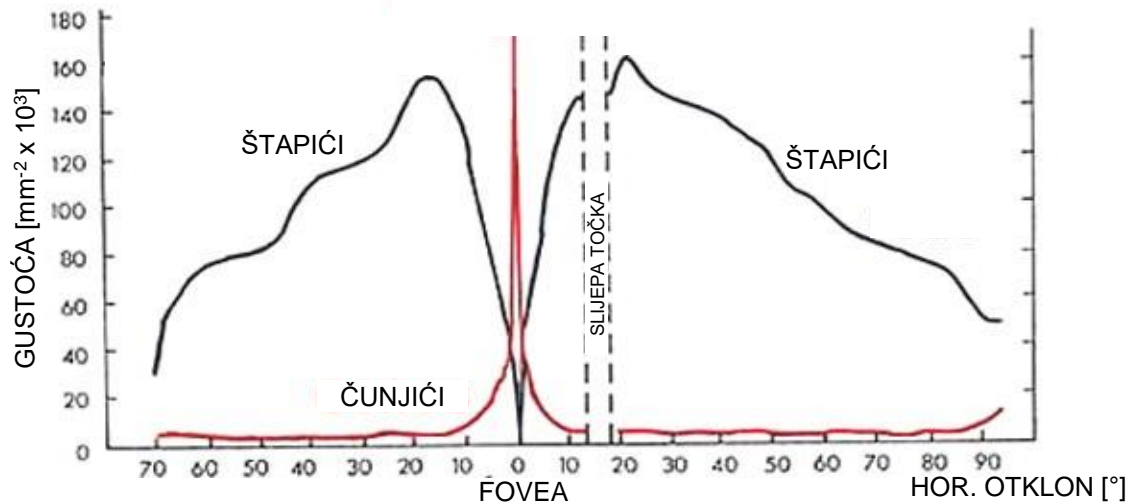
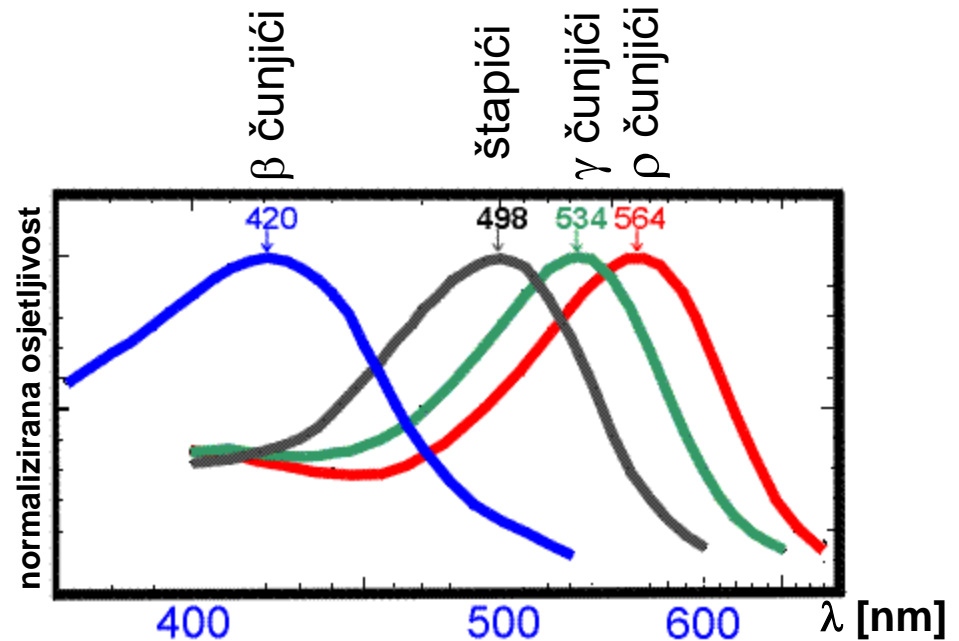
- tri svojstva boje koje ljudsko oko raspoznaje su **obojenost** (hue), **zasićenost** (saturation), i **svjetlina** (brightness)
 - to je osnova za HSB model boja u računalnoj grafici

Ljudski vid



receptori u ljudskom oku:

- štapićaste stanice (osjetljive na svjetlost)
- čunjaste stanice (osjetljive na boju)
 - crveno ρ
 - plavo β
 - zeleno γ

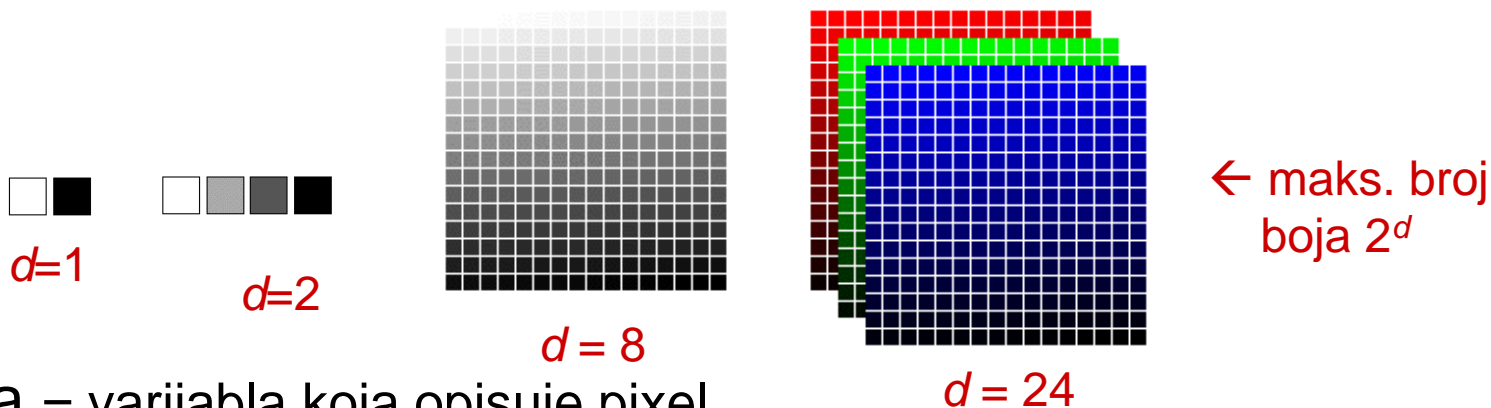


Računalni prikaz slike



Zavod za
telekomunikacije

- slika se promatra kao matrica obojanih točaka, odn. *pixela*
- pixel = *picture element*
 - razlikujemo *pixel slike* od "*pixela uređaja*" (*device pixel, dot*)!! Npr:
 - printer 600 dpi: kvadrat sa stranicom 1/600"
 - video monitor 72 dpi: kvadrat sa stranicom 1/72"
- rezolucija slike = dimenzije matrice pixela $N_1 \times N_2$
- dubina slike = broj bita (d) za opis pixela (odn. boju)

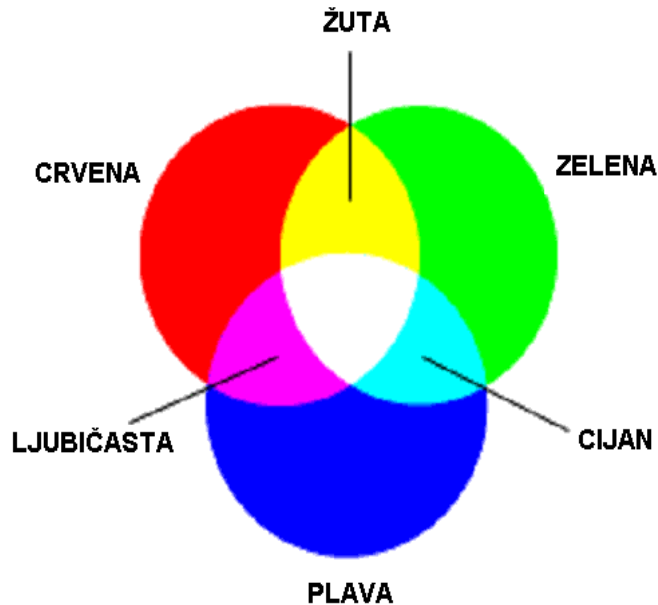


- boja = varijabla koja opisuje pixel
 - modeli boje za sliku: RGB, CMY, HSB, HSI, ...
 - modeli boje za video: YUV, YIQ, ...

Modeli boje za sliku

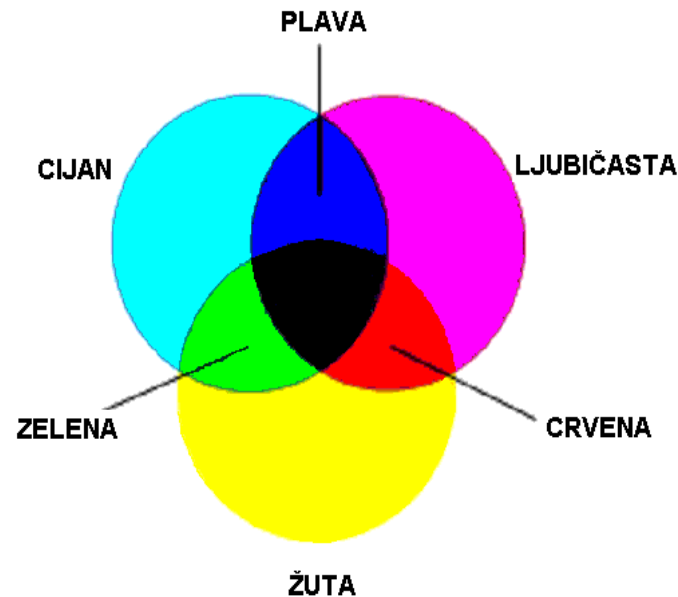
- RGB:

- Red – Green – Blue
- (Crvena – Zelena - Plava)
- svjetlo, zbrojive boje
- TV ekran, monitor u boji



- CMY:

- Cyan, Magenta, Yellow
- (Cijan – Ljubičasta - Žuta)
- pigment, oduzimljive boje
- pisač



- Brojne vlasničke i otvorene norme, npr:
 - BMP (Bitmap)
 - XBM (X11 Bitmap)
 - GIF (Graphics Interchange Format)
 - PNG (Portable Network Graphics)
 - TIFF (Tagged Image File Format)
 - **JPEG** (Joint Photographic Expert Group)
 - JPEG 2000

Kodiranje slike zasniva se na:



Zavod za
telekomunikacije

- Statističkim karakteristikama slike: kodiranje bez gubitaka
- Karakteristikama ljudskog sustava vida: kodiranje s neprimjetnim gubicima
- Sažimanju manje važnih elemenata slike prema nekom kriteriju: kodiranje s vidljivim gubicima
- Obično se radi o kombinaciji ovih ideja

- Koriste se metode entropijskog kodiranja (vidi prethodna predavanja)
- Slijedno kodiranje
 - Telefax (starija verzija)
- Huffman kodiranje
 - Telefax
- LZW metoda (metode rječnika)
 - GIF (Graphics Interchange Format) – bez gubitaka ako se koristi do 256 boja; pogodan za računalnu grafiku

Kodiranje sa gubicima



Zavod za
telekomunikacije

- Diferencijalno (prediktivno) kodiranje
- Transformacijsko kodiranje
 - JPEG
- Kodiranje valićima (wavelets)
 - JPEG 2000
- Fraktalno kodiranje
- Osnovna ideja:

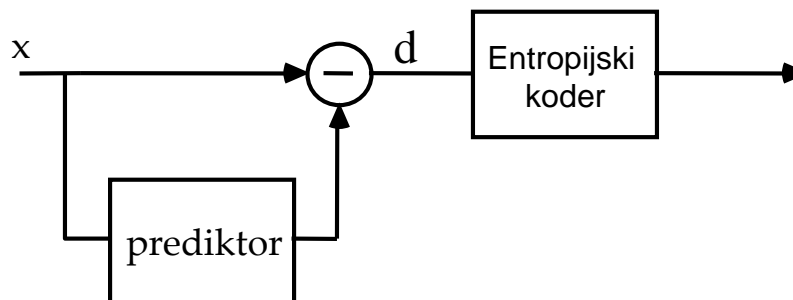


Diferencijalno (prediktivno) kodiranje



Zavod za
telekomunikacije

- Princip: vrijednost slijedećeg signala (pixela) predviđa se iz dosadašnjih vrijednosti, te se kodira razlika stvarnog i predviđenog

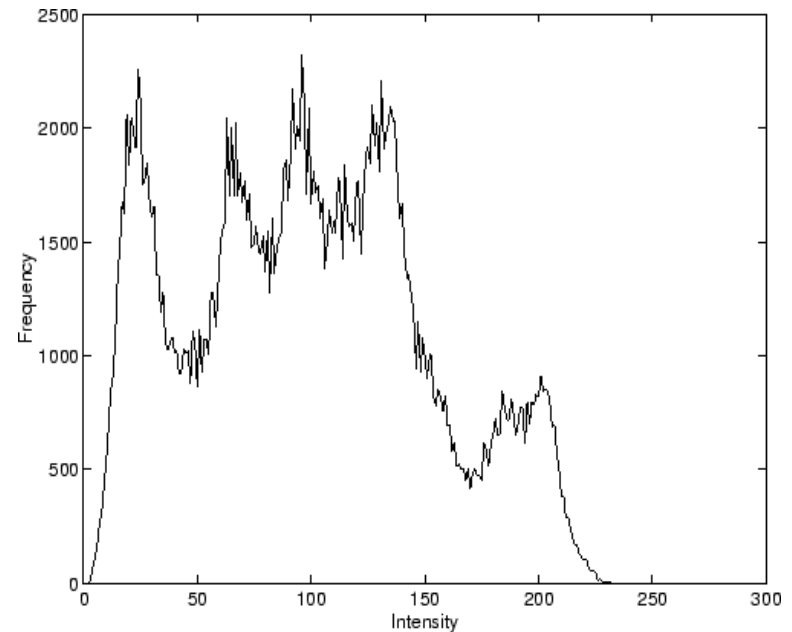


- Raspon amplituda diferencijalnog signala je povoljniji za kodiranje od originalne slike

Primjer: slika i njen histogram



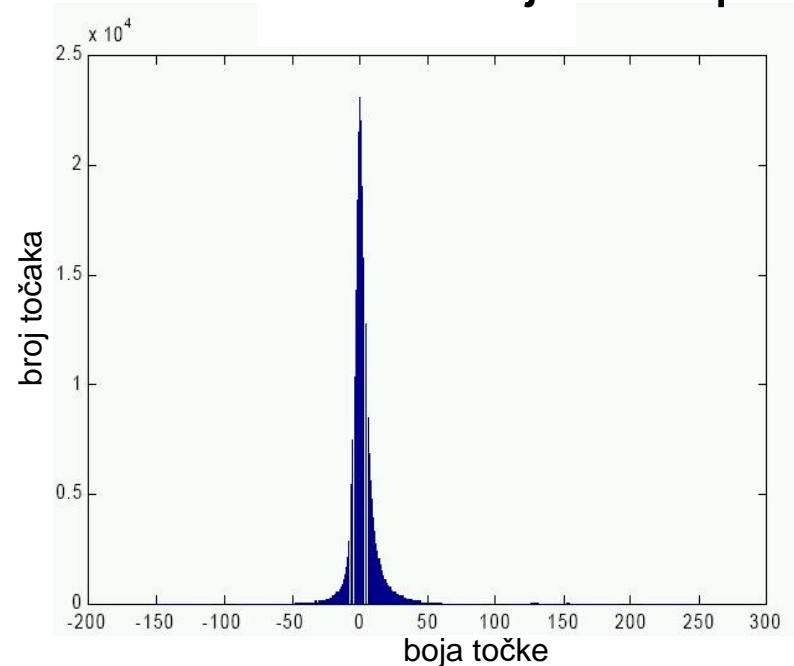
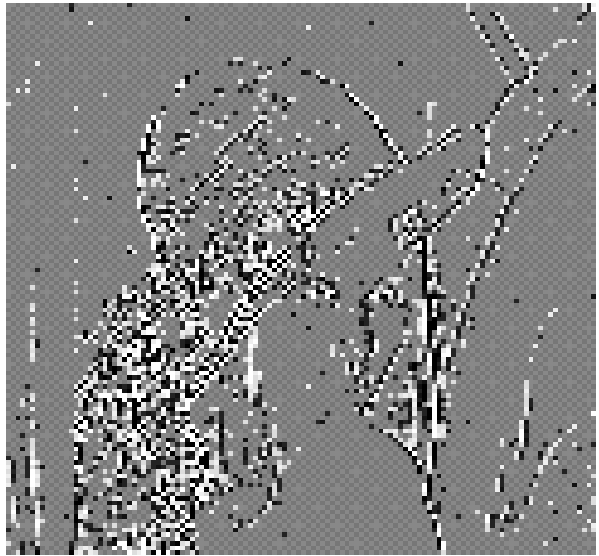
256*256*8 bita



Direktnim entropijskim kodiranjem može se postići 7 bit/pixel

Signal razlike

- Prikazan je signal razlike za sliku iz primjera uz jedan stupanj predviđanja – svodi se na razliku susjednih pixela



- Signal se ujednačuje, povoljnije za entropijsko kodiranje
- Entropijskim kodiranjem dobiva se 2.6 bit/pixel!

Svojstva diferencijalnog kodiranja



Zavod za
telekomunikacije

- + Jednostavna implementacija
- + Može biti bez gubitaka (ovisno o tome da li se upotrebljava kvantizacija)
- Postiže se relativno slaba kompresija

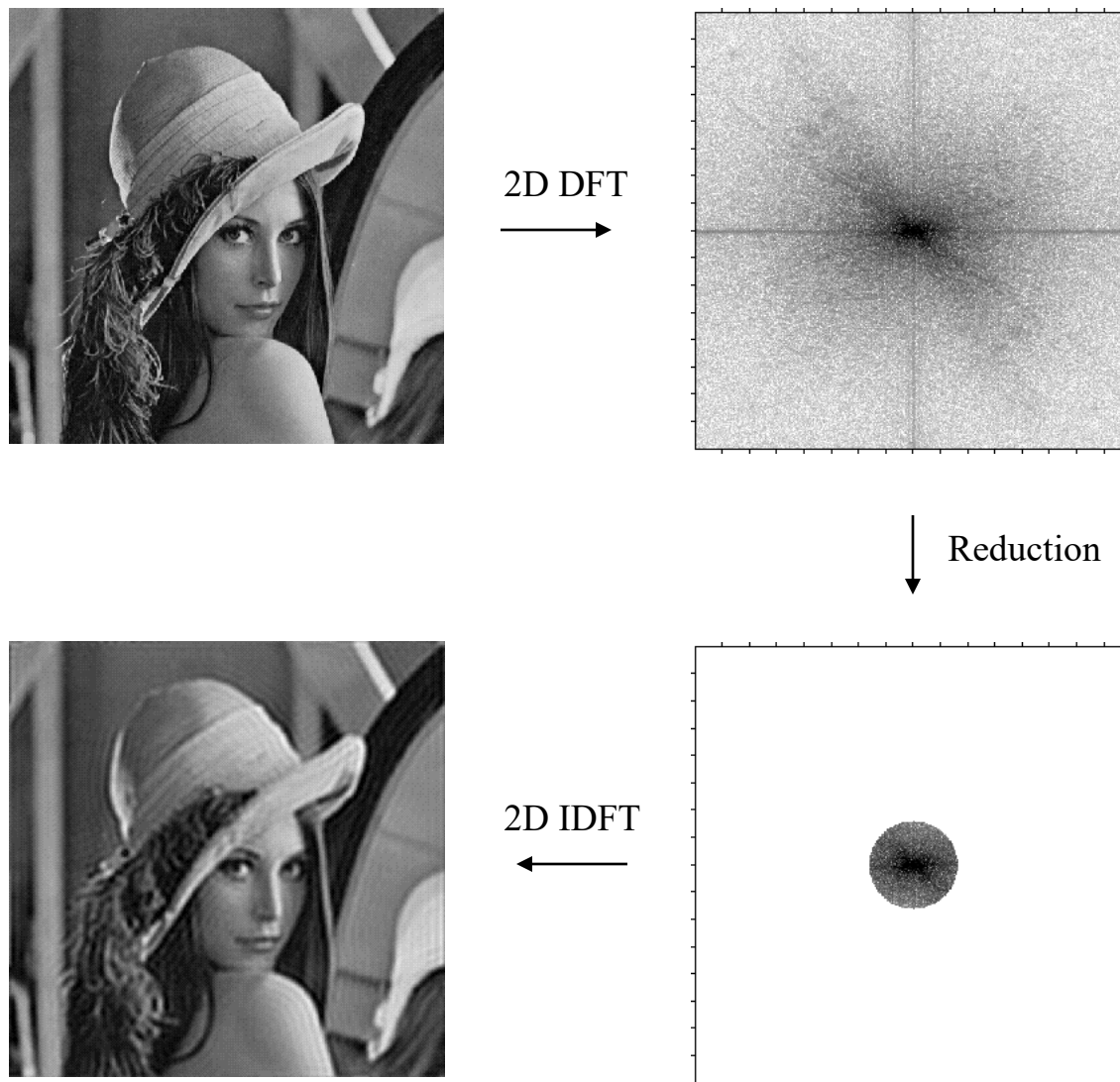
Transformacijsko kodiranje



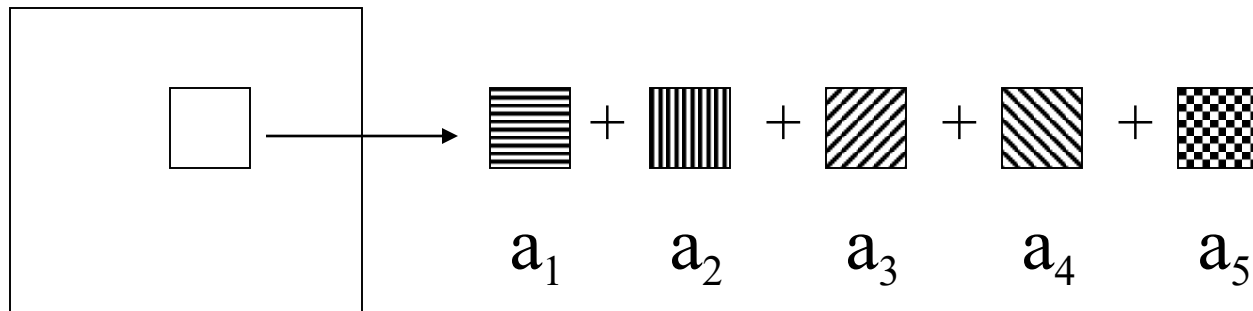
Zavod za
telekomunikacije

- Signal se transformira u prostorno frekvencijske komponente, te se one kodiraju
- Neke frekvencijske komponente slike pojavljuju se puno više od ostalih, što rezultira dobrim kodiranjem
- Metoda razvijena 70-tih godina
- Široka primjena kroz normu JPEG
- JPEG = Joint Photographic Experts Group

Primjer: Fourierova transformacija



1. Slika se dijeli na blokove veličine 8×8 točaka
2. Svaki blok se aproksimira sumom osnovnih DCT blokova, svaki sa svojim koeficijentom doprinosa
3. Koeficijenti doprinosa svakog osnovnog bloka se kodiraju.



Za potpunu rekonstrukciju potrebno je onoliko osnovnih blokova koliko ima točaka u bloku: $8 \times 8 = 64$

Jednodimenzionalna DCT

- DCT – diskretna kosinusna transformacija
- Promatramo jednodimenzionalnu sliku kao vektor dimenzije N pixela
- Za pixel na mjestu x , $0 \leq x < N$, $p(x)$ označava razinu sivog kodiranu s 8 bita (0 – bijelo, 255 – crno)



$$\mathbf{p} = [p(0) \ p(1) \ p(2) \ \dots \ p(x) \ \dots \ p(N-1)]$$

- Promatranu sliku možemo prikazati kao zbroj DCT funkcija s određenim težinama

$$p(x) = \sum_{f=0}^{N-1} S(f) \cdot DCT_f(x) \quad [1]$$

Osnovne DCT funkcije



Zavod za
telekomunikacije

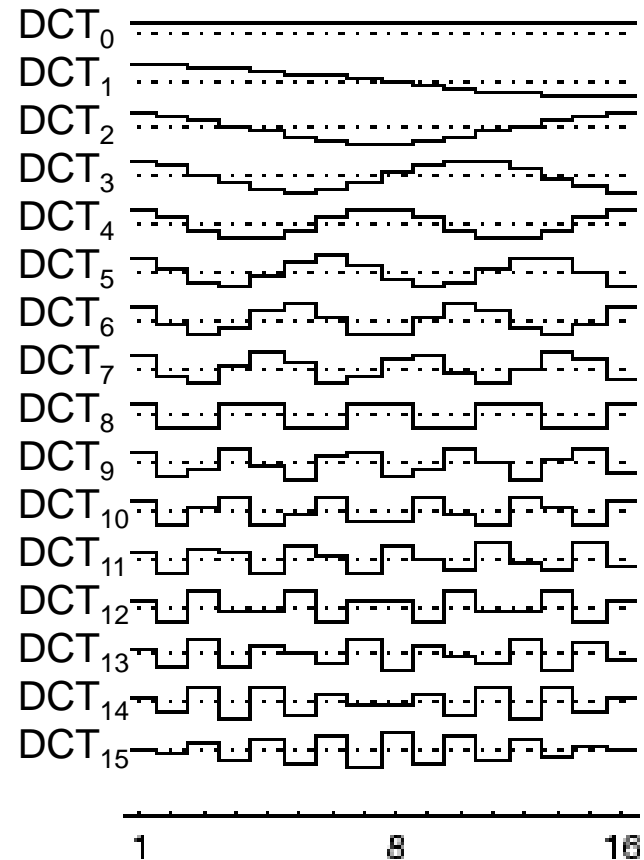
- U jednoj dimenziji:

$$DCT_f(x) = \sqrt{\frac{2}{N}} \cos\left[\frac{(2x+1) \cdot \pi \cdot f}{2N}\right], f > 0$$

$$DCT_0(x) = \sqrt{\frac{1}{N}}$$

- Primjer: N=16

- Uvrštavanjem u [1]:



$$p(x) = S(0) \sqrt{\frac{1}{N}} + \sum_{f=1}^{N-1} S(f) \cdot \sqrt{\frac{2}{N}} \cos\left[\frac{(2x+1) \cdot \pi \cdot f}{2N}\right] \quad [2]$$

- Cjelovitost
 - Težinski se zbroj ovih funkcija može naći za bilo koju kombinaciju od N piksela.
- Minimalnost
 - Niti jedna funkcija se ne može predstaviti zbrojem ostalih, tj. svih N je potrebno za cjelovitost.
- Jedinstvenost
 - Niti jedan drugi skup kosinus funkcija osim onih koje su u razmjeru s korištenim se ne može koristiti za opis bilo koje kombinacije od N piksela.

- Da bismo sliku prikazali DCT funkcijama, moramo izračunati koeficijente $S(f)$ za $0 \leq f \leq N-1$
- Prvi član u zbroju niza je srednja vrijednost, iz čega slijedi izraz za član $S(0)$:

$$\frac{S(0)}{\sqrt{N}} = \frac{1}{N} \sum_{x=0}^{N-1} p(x) \Rightarrow S(0) = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} p(x)$$

- Kada to uvrstimo u prethodni izraz, dobivamo za koeficijente $S(f)$:

$$S(f) = \sqrt{\frac{2}{N}} \cdot C(f) \cdot \sum_{x=0}^{N-1} p(x) \cdot \cos\left[\frac{(2x+1)\pi f}{2N}\right] \quad C(f) = \begin{cases} 1/\sqrt{2}; & f = 0 \\ 1; & f > 0 \end{cases} \quad [3]$$

- **Matrica DCT koeficijenata $S(f)$ opisuje sliku u frekvencijskoj domeni.**

Primjer (1/3)

- Promatramo konkretan primjer slike od $N=8$ pixela



$$p(x) = [240 \ 210 \ 180 \ 150 \ 120 \ 90 \ 60 \ 30]$$

- Prije primjene DCT radi se posmak razina (*level shift*), tj. od svih vrijednosti u vektoru razina sivog oduzimamo 128 kako bi vrijednosti bile simetrične u odnosu na 0 (interval od -128 do 127) i tako dobivamo vektor p'

$$p'(x) = [112 \ 82 \ 52 \ 22 \ -8 \ -38 \ -68 \ -98]$$

- Za $N=8$, koeficijente $S(f)$ računamo prema izrazu [3] i dobivamo:

$$S(f) = [19.89 \ 193.27 \ 0 \ 20.20 \ 0 \ 6.03 \ 0 \ 1.52]$$

Primjer (2/3)

- $\mathbf{S}(f)$ zaokružujemo na cijele brojeve i dobivamo $\mathbf{S}'(f)$:

$$\mathbf{S}'(f) = [\text{20} \quad \text{193} \quad \text{0} \quad \text{20} \quad \text{0} \quad \text{6} \quad \text{0} \quad \text{2}]$$

- Za provjeru, izračunajmo sada $p'(x)$ prema izrazu [2]:

$$p(x) = S(0) \sqrt{\frac{1}{N}} + \sum_{f=1}^{N-1} S(f) \cdot \sqrt{\frac{2}{N}} \cos \left[\frac{(2x+1) \cdot \pi \cdot f}{2N} \right]$$

$$p'(x) = [\text{111.89} \quad \text{81.86} \quad \text{52.29} \quad \text{21.86} \quad \text{-7.71} \quad \text{-38.15} \quad \text{-67.72} \quad \text{-97.75}]$$

- Nakon zaokruživanja i posmaka razina dobivamo originalnu funkciju $p(x)$.

$$p'(x) = [\text{112} \quad \text{82} \quad \text{52} \quad \text{22} \quad \text{-8} \quad \text{-38} \quad \text{-68} \quad \text{-98}]$$

$$p(x) = [\text{240} \quad \text{210} \quad \text{180} \quad \text{150} \quad \text{120} \quad \text{90} \quad \text{60} \quad \text{30}]$$

Ovisno o efektima zaokruživanja, DCT je samo približno reverzibilna!

Primjer (3/3)

- Utjecaj nižih frekvencijskih komponenata:

A $S'(f) = [20 \ 193 \ 0 \ 20 \ \underline{0 \ 6 \ 0 \ 2}]$

$p(x) = [240 \ 210 \ 180 \ 150 \ 120 \ 90 \ 60 \ 30]$

B $S'(f) = [20 \ 193 \ 0 \ 20 \ \underline{0 \ 0 \ 0 \ 0}]$

$p(x) = [238 \ 213 \ 179 \ 148 \ 122 \ 91 \ 57 \ 32]$

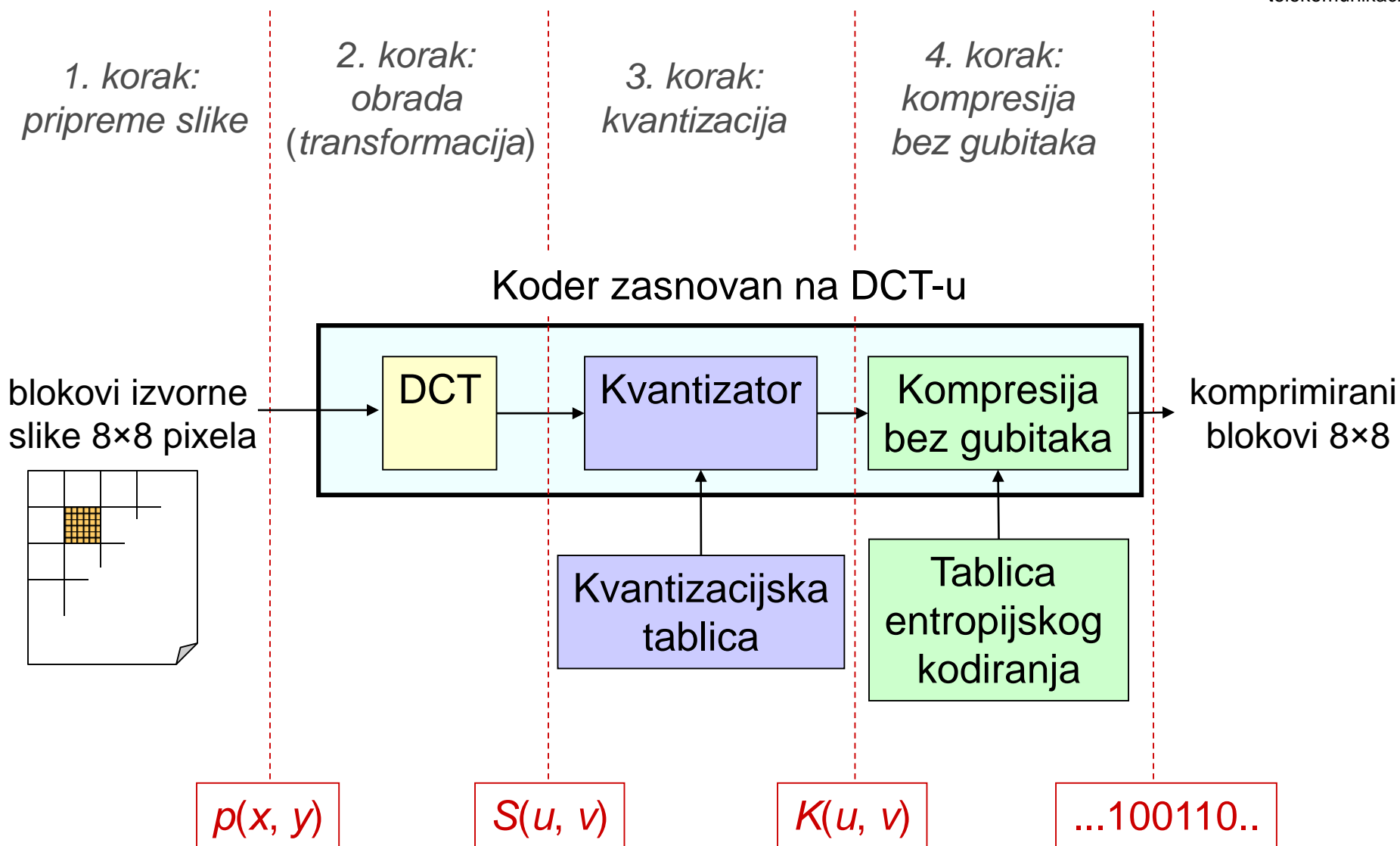
- Vizualno:

A  sa svim frekvencijama

B  bez nižih frekvencija

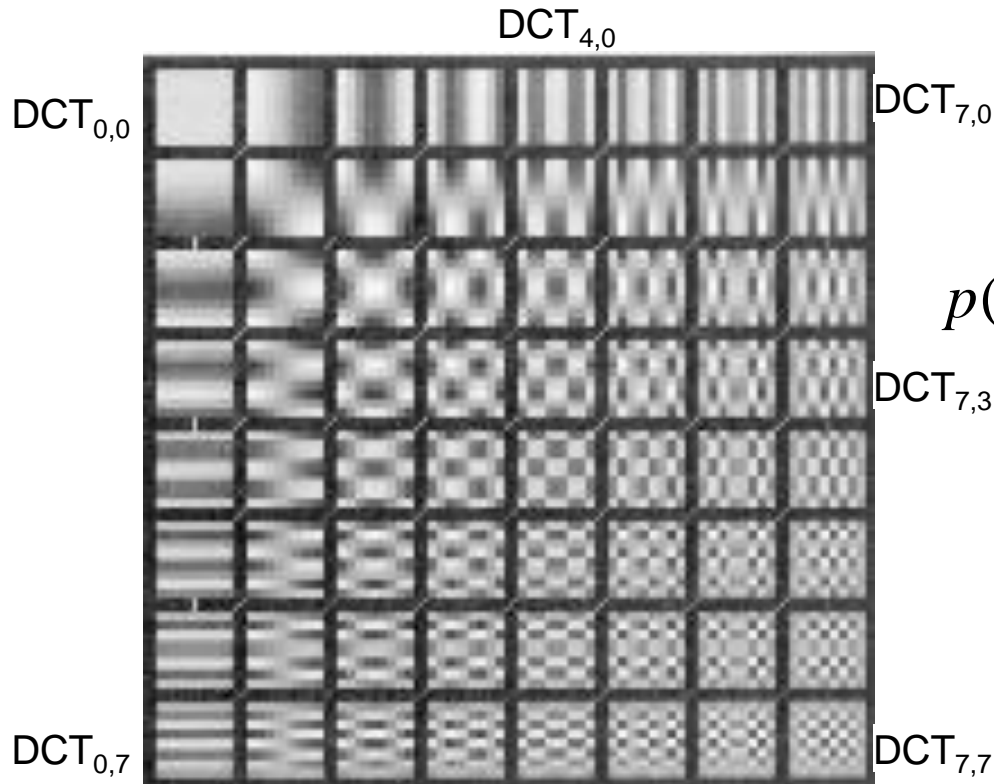
- Frekvencijske komponente na nižim frekvencijama su izraženije od onih na višim frekvencijama
 - Prva četiri koeficijenta u primjeru opisuju gotovo cijelu sliku
- Visoke frekvencije izražavaju zaista sitne detalje u slici
- Doprinos članova na višim frekvencijama je mali kada su razlike susjednih pixela relativno male, dakle kada slika ne sadrži puno detalja
- Za većinu slika, ovo je zaista slučaj, te se na tom svojstvu zasniva JPEG kompresija (i druge metode transformacijskog kodiranja)
- Premda smo ovo pokazali na primjeru 1D, ovo jednako važi za dvodimenzionalnu sliku

Blok shema JPEG koder



DCT transformacija slike (1/2)

- DCT funkcije u 2D dobivaju se množenjem 1D funkcija
- Osnovni blokovi DCT transformacije:



$$p(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} S(u, v) \cdot DCT_{u,v}(x, y)$$

DCT transformacija slike (2/2)

- DCT transformacija vrši preslikavanje niza vrijednosti piksela u niz koeficijenata težine osnovnih blokova
- DCT koeficijenti za svaki blok 8x8 pixela se računaju prema formuli:

$$S(u, v) = \frac{C(u) \cdot C(v)}{4} \cdot \sum_{x=0}^7 \sum_{y=0}^7 p(x, y) \cdot \cos\left[\frac{(2x+1)\pi \cdot u}{16}\right] \cdot \cos\left[\frac{(2y+1)\pi \cdot v}{16}\right] \quad C(f) = \begin{cases} 1/\sqrt{2}; & f = 0 \\ 1; & f > 0 \end{cases}$$

- primjer:

x/y	0	1	2	3	4	5	6	7
0	79	75	79	82	82	86	94	94
1	76	78	76	82	83	86	85	94
2	72	Slične vrijednosti!						82
3	74							79
4	73	70	75	67	78	78	79	85
5	69	63	68	69	75	78	82	80
6	76	76	71	71	67	79	80	83
7	72	77	78	69	75	75	78	78

Izvorišna matrica $p(x, y)$

	Najviša vrijednost $S(0, 0)$					Niske vrijednosti		
u/v	0	1	2	3	4	5	6	7
0	619	-29	8	2	1	-3	0	1
1	22	-6	-4	0	7	0	-2	-3
2	11	0	5	-4	-3	4	0	-3
3	2	-10	5	0	0	7	3	2
4	6	2	-1	-1	-3	0	0	8
5	1	2	1	2	0	2	-2	-2
6	-8	-2	-4	1	2	1	-1	1
7	-3	1	5	-2	1	-1	1	-3

DCT matrica $S(u, v)$

3: kvantizacija

- DCT koeficijenti se kvantiziraju zaokruživanjem na cjelobrojnu vrijednost omjera matrice koeficijenata **S** i kvantizacijske matrice **Q**; nastaje *kvantizirana DCT matrica K*:

$$K(u, v) = \text{round}\left(\frac{S(u, v)}{Q(u, v)}\right) = \left\lfloor \frac{S(u, v)}{Q(u, v)} + 0.5 \right\rfloor$$

- Q** mora biti tako izabrana da kvantizacija rezultira visokom kompresijom, ali bez primjetnog gubitka kvalitete
 - norme ne određuju, ali se preporučuje skup kvantizacijskih matrica **Q**

<i>u/v</i>	0	1	2	3	4	5	6	7
0	16	11	10	16	24	40	51	61
1	12	12	14	19	26	58	60	55
2	14	13	16	24	40	57	69	56
3	14	17	22	29	51	87	80	62
4	18	22	37	56	68	109	103	77
5	24	35	55	64	81	104	113	92
6	49	64	78	87	103	121	120	101
7	72	92	95	98	112	100	103	99

Kvantizacijska matrica $Q(u, v)$

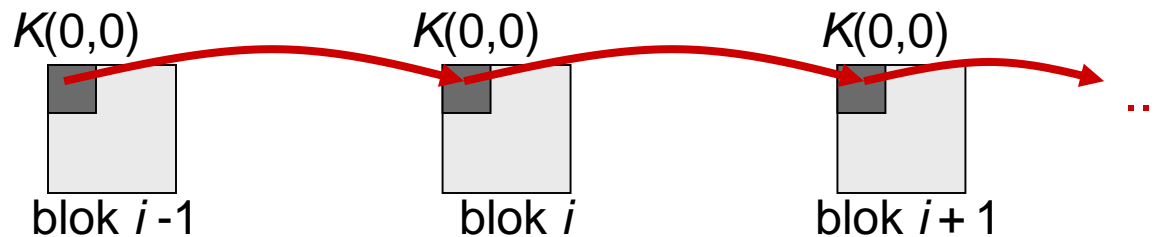
<i>u/v</i>	0	1	2	3	4	5	6	7
0	39	-3	1	0	0	0	0	0
1	2	-1	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0
3	0	-1	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0

Kvantizirana DCT matrica $K(u, v)$

Vrijednost 0

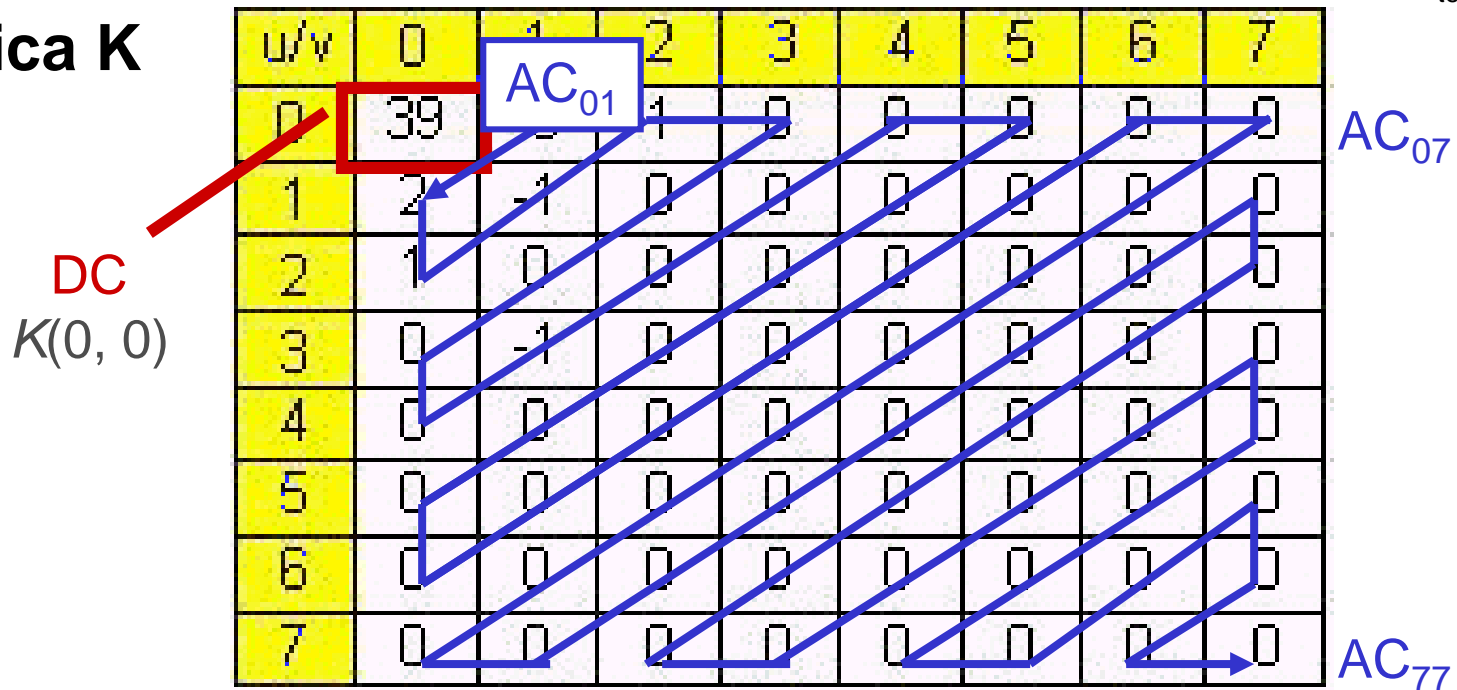
4: kompresija bez gubitaka

- Koeficijent $K(0,0)$ (DC komponenta) redovito sadrži najveći dio ukupne vrijednosti bloka, tj. nosi **najviše informacije** o bloku kojeg predstavlja
- Korelacija između vrijednosti koeficijenata $K(0,0)$ susjednih blokova u slici \rightarrow koristi se **diferencijalno kodiranje**



4: kompresija bez gubitaka

matrica K



- Koeficijenti redom: 39 -3 2 1 1 0 0 0 0 0 -1 0 0 0 0 0 0 0 0....0, što se kraće zapisuje kao:
39 -3 2 1 1 0 0 0 0 0 -1 **EOB (End of Block)**
- Nakon takvog zapisa svih blokova u slici slijeva nadesno i odozgo prema dolje, slijedi **entropijsko kodiranje**
 - može se upotrijebiti Huffmanovo ili aritmetičko kodiranje

Primjer: JPEG

- Slika 516 x 516 točaka
- Kompresija 1:16



256 KB = 8 bit/pixel



16 KB = 0.5 bit/pixel

- + Visoka kompresija

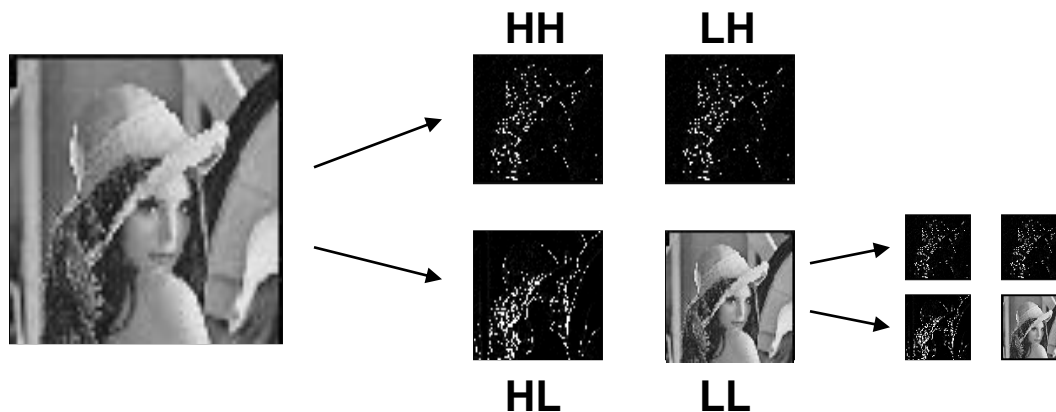
(odlična kvaliteta slike do 0.25 bit/pixel)

- + Kvaliteta se može regulirati

- Pri većim kompresijama postaju vidljive granice blokova

Kodiranje valićima (wavelet coding)

- Modernija metoda
- Slika se iterativno dijeli u prostorno-frekvencijske pojaseve u horizontalnom i vertikalnom smjeru
- Podjela se obavlja zbirkama filtara
- Prilikom kodiranja, manje bitova za komponente s manje energije



Kodiranje valićima: svojstva



Zavod za
telekomunikacije

- Visoka kompresija
- „Prirodna“ distorzija (zamućenost)
- Prednosti u odnosu na transformacijsko kodiranje:
 - Manja složenost
 - Nema blok efekta
 - Mogućnost progresivnog primanja i stvaranja slike
- Osnova nove norme JPEG 2000

JPEG 2000 vs JPEG (1/2)



Zavod za
telekomunikacije

- Veći dinamički raspon (16-32 bit/pixel)
- Bolja kompresija (25-30% manje podataka)
- Progresivno slanje podataka
- Kodiranje bez gubitaka prema područjima interesa (Region-of-interest, ROI)
- Međunarodna norma 2001

JPEG 2000 vs JPEG (2/2)

- Usporedba rezultata uz jednaku kompresiju – 1:64



JPEG

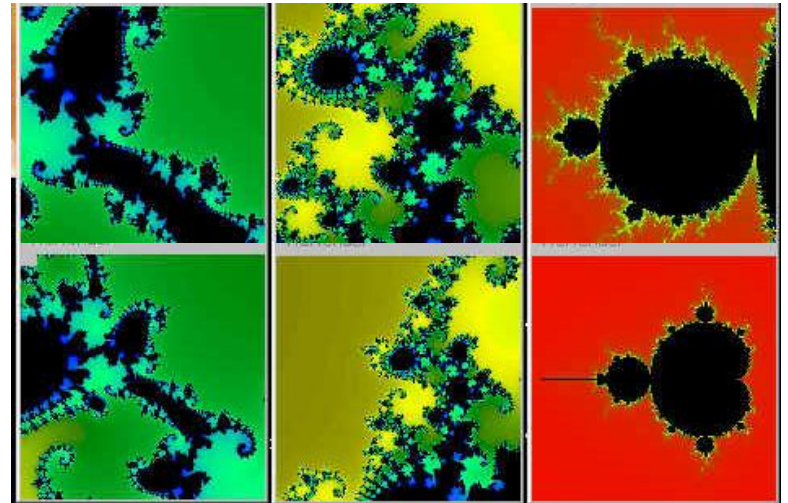


JPEG2000

- Princip analize i sinteze
 - Umjesto kodiranja samih podataka, kodiraju se parametri neke **funkcije** koja će generirati te podatke
 - Isti princip kao kod koda govora zasnovanih na modelu
 - Nažalost, za općenitu sliku je teško pronaći model, odnosno funkciju koja je opisuje
- Ideja: samo-sličnost
 - Dijelovi slike međusobno slični, uz transformaciju
 - Pretpostavka je da se ovo svojstvo može iskoristiti za prikaz slike pomoću **fraktala**, koji pokazuju slična svojstva

Fraktali

- Benoit Mandelbrot, 1975
- Fragmentirani, nepravilni geometrijski objekti koji pokazuju svojstvo samo-sličnosti
- Obično stvoreni rekurzivnim ponavljanjem određene funkcije
- U svakoj iteraciji objekt je transformirana verzija objekta iz prošle iteracije
- Proizvoljan nivo detalja



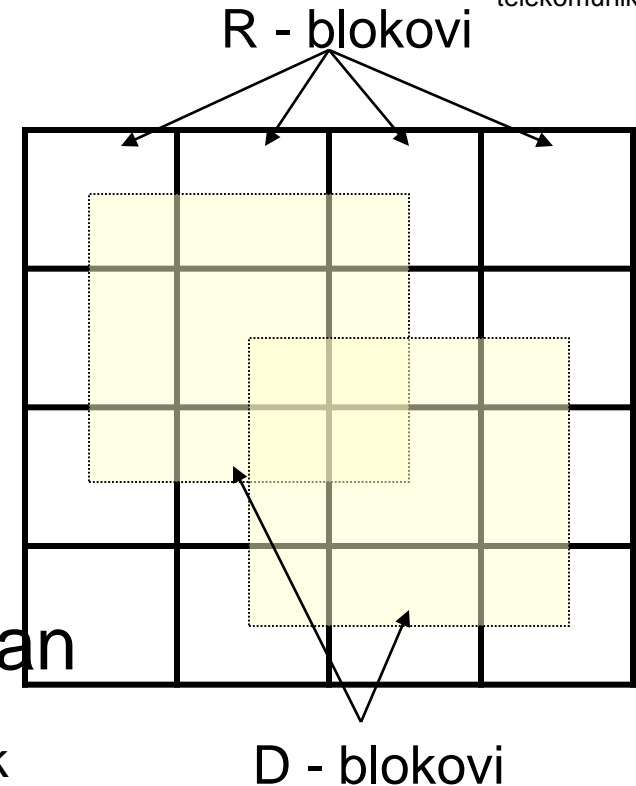
Funkcija s fiksnom točkom



Zavod za
telekomunikacije

- $f(x) = x$; za točno jednu vrijednost $x=x_0$ (vrijednost funkcije će biti ista kao onu koju uvrštavamo)
- Npr. $F(x) = ax + b$; za $a = 0.5$ i $b = 1 \Rightarrow x_0 = 2$, $F(2) = 2$
- Počevši od bilo koje početne vrijednosti, iteracijom ćemo uvijek stići do $x_0 = 2$
- Vrijednost x_0 možemo prikazati parom parametara a , b ako smo prethodno fiksirali oblik funkcije $F = ax + b$
- Za sliku (skup točaka) I , $F(I) = I$
 - Ako se parametri funkcije F mogu prikazati manjim brojem bita nego sama slika, postiže se kompresija

- Arnaud Jacquin (1989)
- Podjela slike
 - R – blokovi uniformno pokrivaju čitavu sliku $\bigcup_k R_k = I$
 - D – blokovi su veći, mogu biti bilo gdje unutar slike i ne pokrivaju je
- Za svaki R – blok, nalazi se jedan D – blok i funkcija f_k , $f_k(D_k) = R_k$
- Skup svih funkcija f_k definira čitavu sliku



$$\bigcup_k f_k = F \rightarrow F(\hat{I}) = \hat{I}$$

- Komplicirano
- Neujednačen stupanj kompresije
 - 4:1 do 100:1, jako ovisno o slici
- U početku se od fraktala puno očekivalo, no druge metode (najviše valići!) su uspješnije
- Princip je privlačan, moguće je da dođe do daljnjih napredaka i da metoda ipak zaživi u praksi