

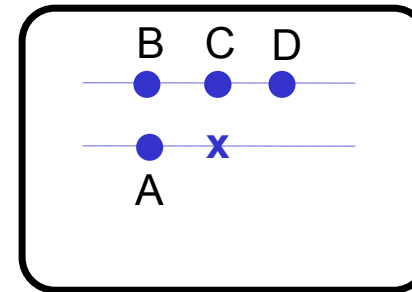
Postupci i norme za kompresiju slike i videosignala

Postupci kompresije

- kompresija (sažimanje)
 - procesiranje analognog ili digitalnog signala radi smanjenja ili uklanjanja redundancije (zalihosti, suvišnih informacija)
 - subjektivna redundancija
 - određena svojstvima ljudskog vizualnog sustava
 - ljudsko oko je osjetljivije na niske prostorne frekvencije nego na visoke (podsjetimo se oblika funkcije osjetljivosti na kontrast)
 - prostorna redundancija
 - javlja se kao posljedica postojanja korelacije (međuviznosti ili sličnosti) između elemenata slike u pojedinoj slici
 - vremenska redundancija
 - javlja se kao posljedica postojanja korelacije između uzastopnih slika u videosignalu
 - statistička redundancija
 - javlja se kao posljedica postojanja korelacije u nizu kodnih znakova
 - posljedica uklanjanja redundancije je smanjenje frekvencijske širine pojasa za prijenos signala, smanjenje brzine prijenosa (*bit-rate reduction*) te smanjenje prostora potrebnog za pohranjivanje signala

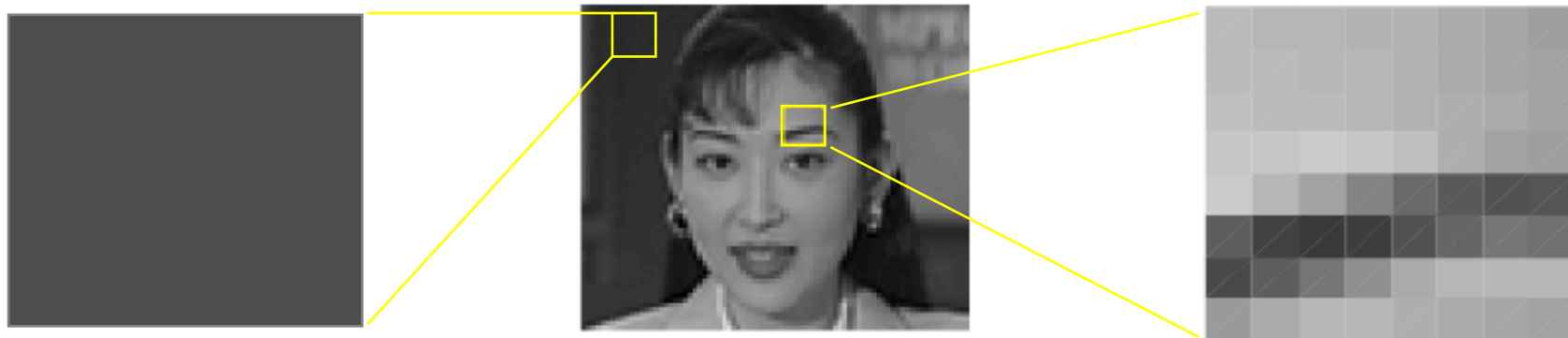
Postupci kompresije

- prostorna redundancija
 - postoji zbog korelacije susjednih elemenata slike unutar pojedine slike (x je koreliran s A, B, C i D)



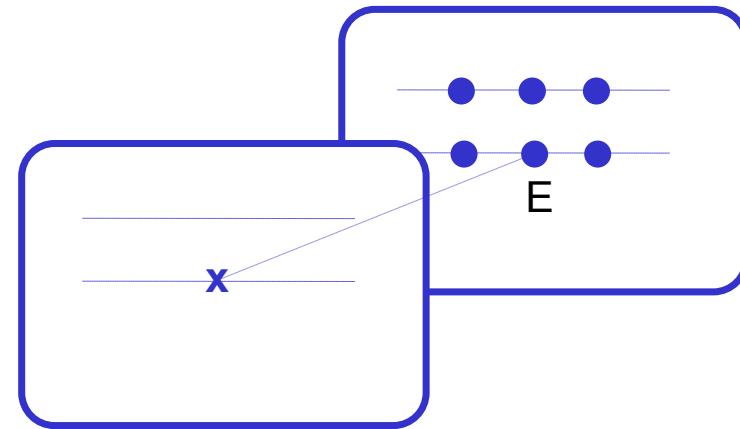
Prostorna korelacija

- susjedni elementi slike u slici imaju približno iste vrijednosti



Postupci kompresije

- vremenska redundancija
 - postoji zbog vremenske korelacije između elemenata slike uzastopnih slika (uzorak x iz trenutne slike je sličan uzorku E na odgovarajućem položaju u prethodnoj slici)



Vremenska korelacija

- susjedne slike u nizu slika su slične



Slika 1



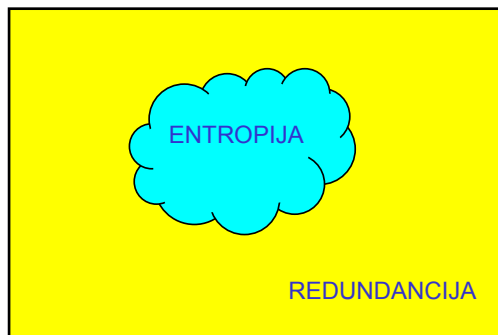
Slika 2



Slika 3

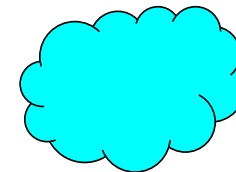
Postupci kompresije

- optimalna kompresija ili kompresija bez gubitaka
 - uređaj za komprimiranje uz dio redundancije izdvaja i cijelu entropiju
- ekscesna kompresija ili kompresija s gubicima
 - uređaj za komprimiranje uz redundanciju odbacuje i dio entropije
- savršena kompresija
 - uređaj za komprimiranje odbacuje redundanciju i izdvaja samo entropiju
 - u praksi bi takav uređaj bio vrlo složen i zahtijevao dugo vrijeme procesiranja



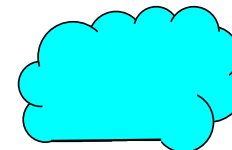
Savršena kompresija

- sva entropija odaslana
- sva redundancija uklonjena
- nema gubitka kvalitete



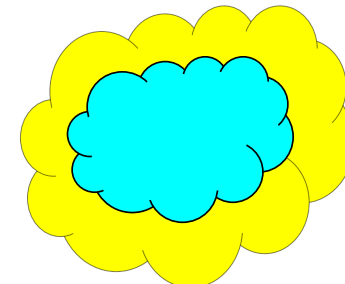
Ekscesna kompresija

- dio entropije izgubljen
- gubitak kvalitete



Optimalna kompresija

- sva entropija i dio redundancije odaslani
- nema gubitka kvalitete



Postupci kompresije

- pri kompresiji slike i videosignala potrebno je postići optimalan odnos stupnja kompresije, kvalitete slike i vremena procesiranja
 - stupanj kompresije (CR, *Compression Ratio*)
 - omjer ukupnog broja bita za kodiranje slike prije i nakon provedenog postupka kompresije (npr. 40 : 1)
 - prosječan broj bita za kodiranje uzorka slike nakon kompresije (npr. 0,2 bpp) koji se dobije kao omjer broja bita za kodiranje uzorka slike prije kompresije i stupnja kompresije (8 bpp / 40=0,2 bpp)

- kvaliteta slike

- srednja kvadratna pogreška (MSE, *Mean Square Error*)

$$MSE = \frac{1}{N} \cdot \sum_{i=0}^{N-1} (x_i - x_i')^2$$

- omjer signal/šum

- omjer vršne snage signala i efektivne snage šuma (PSNR, *Peak Signal-to-Noise Ratio*)

$$PSNR = 10 \log \frac{(2^n - 1)^2}{MSE}$$

N - broj elemenata slike u slici

x_i - vrijednost originalnog elementa slike

x_i' - vrijednost rekonstruiranog elementa slike

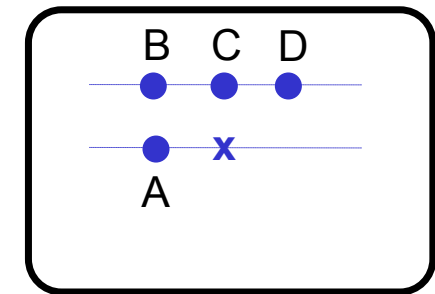
n - broj bita po uzorku

- subjektivna kvaliteta

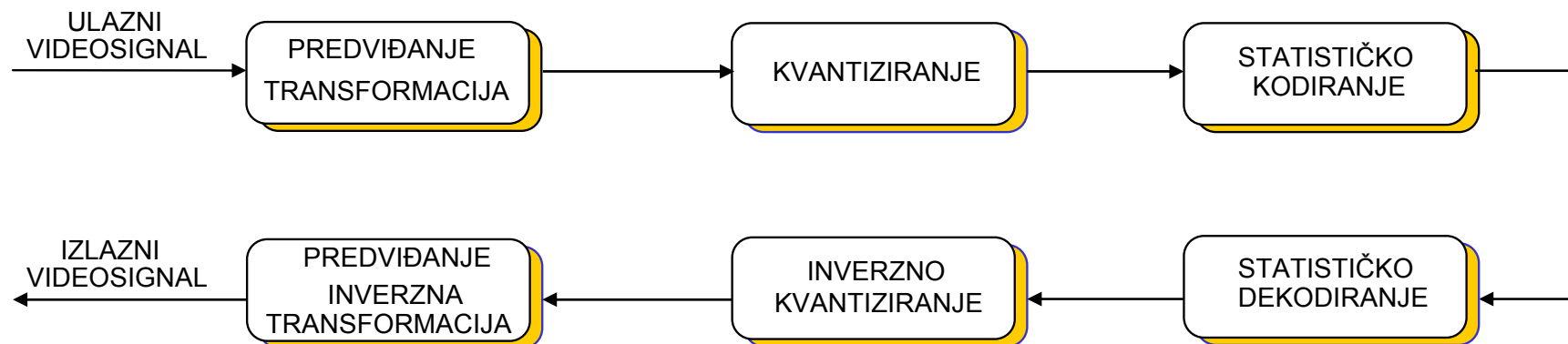
- srednja iskustvena vrijednost (MOS, *Mean Opinion Score*) dobivena ispitivanjima u kojima promatrači ocjenjuju kvalitetu slike u kontroliranim uvjetima

Postupci kompresije

- kodiranje unutar slike (*intraframe coding*)
 - slika se obrađuje neovisno o ostalim slikama u slijedu slika, a uklanja se prostorna i statistička redundancija
 - najčešće se rabi transformacijsko kodiranje
 - moguće je primijeniti kodiranje s predviđanjem (predikcijsko kodiranje) kod kojega se vrijednost pojedinog elementa slike koji treba biti kodiran (x) predviđa iz prethodno kodiranog elementa slike (A) u istoj liniji ili iz njemu najbližih elemenata slike (B, C, D) prethodne linije



Prostorna korelacija

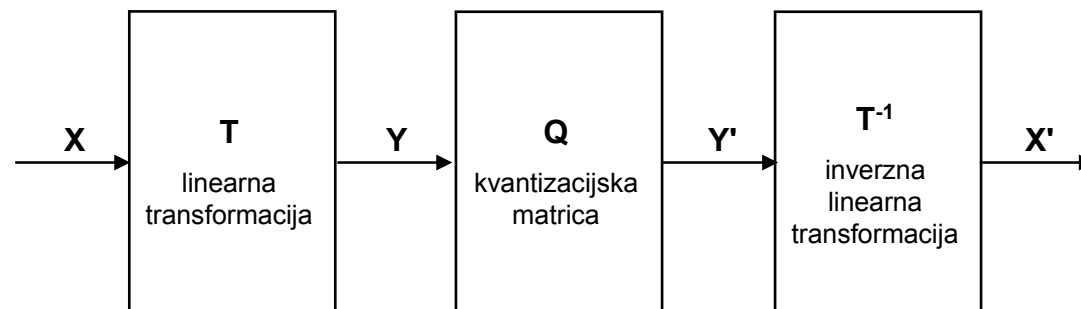


Transformacijsko kodiranje

- $\mathbf{X} = \{x_1, x_2, \dots, x_N\}^T$ je vektor uzorka ulaznog signala
 - između susjednih uzoraka u prirodnim slikama postoji visok stupanj korelacije
 - najveći dio snage prirodnih slika sadržan je u niskofrekvencijskom području
- primjenom linearne transformacije \mathbf{T} na vektor uzoraka ulaznog signala \mathbf{X} nastaje vektor transformiranih uzoraka $\mathbf{Y} = \mathbf{T}\mathbf{X}$
 - komponente vektora \mathbf{Y} nazivaju se transformacijski koeficijenti
- ukoliko linearna transformacija \mathbf{T} rezultira u komponentama vektora \mathbf{Y} koje su manje međusobno korelirane od komponenti vektora \mathbf{X} , vektor \mathbf{Y} može biti djelotvornije kodiran od \mathbf{X}
- inverzna transformacija \mathbf{T}^{-1} obnavlja ulaznu sekvencu podataka $\mathbf{X} = \mathbf{T}^{-1}\mathbf{Y}$
 - ako je matrica \mathbf{T} ortogonalna tada se inverzna transformacija dobije njezinim transponiranjem: $\mathbf{T}^{-1} = \mathbf{T}^T$

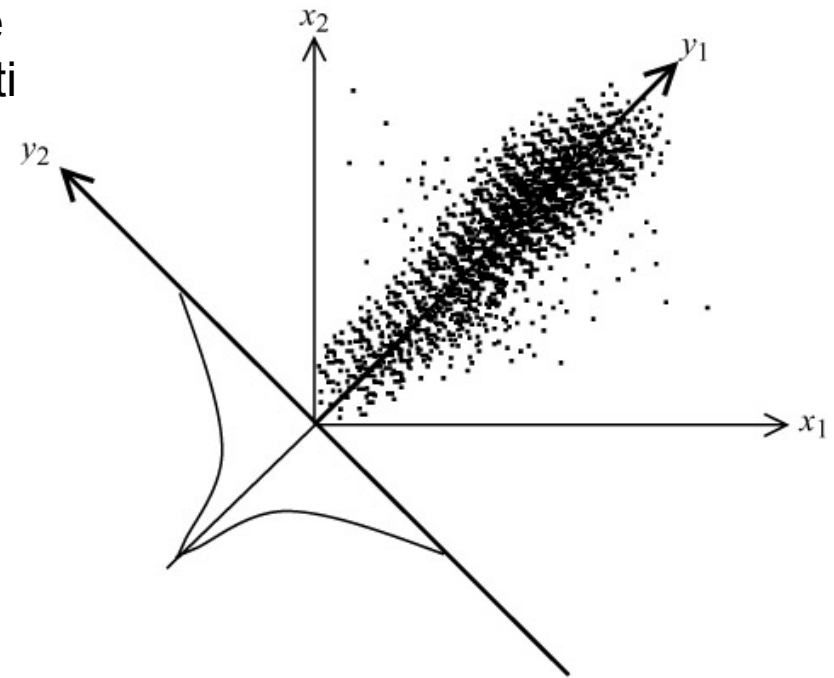
Transformacijsko kodiranje

- poželjna svojstva postupka transformacije
 - komponente vektora \mathbf{Y} su statistički neovisne
 - nakon transformacije veći dio snage zbijen je u mali broj komponenta
 - komponente vektora \mathbf{Y} s najvećim amplitudama smještene su u području niskih frekvencija
- ukoliko \mathbf{T} ispunjava navedena svojstva, informacija o signalu sadržana je u prvih nekoliko komponenti vektora \mathbf{Y}
 - manje značajne komponente mogu biti grubo kvantizirane ili izjednačene s 0 bez zamjetnog utjecaja na kvalitetu signala nakon inverzne transformacije
 - \mathbf{Y}' je vektor kvantiziranih transformacijskih koeficijenata, a \mathbf{X}' vektor rekonstruiranih uzorka ulaznog signala
 - kvantizacija je korak kojim se postiže kompresija i zbog kojeg je $\mathbf{X} \neq \mathbf{X}'$



Transformacijsko kodiranje

- primjer transformacijskog kodiranja
 - x_1 i x_2 su susjedni uzorci slike koji mogu zauzeti bilo koju vrijednost između 0 i 255
 - zbog visokog stupnja korelacije točke koje označavaju pojavljivanje parova vrijednosti x_1 i x_2 grupirane su oko pravca 45°
 - ukoliko koordinate x_1x_2 rotiramo za 45° nastaje novi koordinatni sustav y_1y_2
 - duž osi y_1 točke koje označavaju pojavljivanje parova vrijednosti y_1 i y_2 imaju jednoliku distribuciju
 - duž osi y_2 točke koje označavaju pojavljivanje parova vrijednosti y_1 i y_2 su zbijene oko vrijednosti $y_2=0$
 - broj bita potreban za prikaz parametra y_1 jednako je velik kao broj bita potreban za prikaz x_1 ili x_2
 - y_2 zahtijeva manji broj bita tako da u cjelini uzorci prikazani u sustavu y_1y_2 mogu biti prikazani s manjim brojem bita po uzorku nego uzorci u sustavu x_1x_2



Transformacijsko kodiranje

- rotacija x_1x_2 koordinata za 45° predstavlja transformaciju vektora $\mathbf{X} = \{x_1, x_2\}^T$ transformacijskom matricom \mathbf{T}

$$\mathbf{T} = \begin{bmatrix} \cos 45 & \sin 45 \\ \sin 45 & -\cos 45 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

- komponente transformacijske matrice nazivaju su bazni vektori
- transformacijski koeficijenti y_1 i y_2 su tada

$$[y_1, y_2] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$y_1 = \frac{1}{\sqrt{2}}(x_1 + x_2) \quad y_2 = \frac{1}{\sqrt{2}}(x_1 - x_2)$$

- y_1 je istosmjerni (DC) koeficijent proporcionalan srednjoj vrijednosti od x_1 i x_2 , a y_2 predstavlja preostalu razliku
- normalizacijski faktor $1/\sqrt{2}$ osigurava da snaga signala uslijed transformacije neće biti promijenjena (Parsevalov teorem)
 - snaga komponenti u području elemenata slike ($x_1^2 + x_2^2$) jednaka je snazi komponenti u području transformacije ($y_1^2 + y_2^2$)

Transformacijsko kodiranje

- razmotrimo N koreliranih elemenata slike (umjesto dva kao u prethodnom primjeru)
- točke koje označavaju pojavljivanje N -torki x_1, x_2, \dots, x_N formiraju kuglu
- ako provedemo transformaciju na način da os y_1 leži na središnjoj dijagonali kugle, koeficijent y_1 postaje značajniji od ostalih jer prenosi najveći dio informacije, dok koeficijenti y_2, y_3, \dots, y_N prenose ostatak informacije
- transformacija N elemenata slike omogućava veću kompresiju ukoliko je N odabran tako da odgovara udaljenosti elemenata slike koji su još međusobno korelirani
- u određivanju djelotvornosti transformacijskog kodiranja veliki značaj ima oblik baznih vektora čijim odabirom je moguće postići koncentraciju snage signala u samo nekoliko transformacijskih koeficijenata

Optimalna transformacija

- optimalna transformacija
 - optimalna transformacijska matrica je ona koja u potpunosti dekorelira transformacijske koeficijente
 - \mathbf{R}_{xx} je autokorelacijska matrica ulaznih elemenata slike

$$\mathbf{R}_{xx} = E[\mathbf{X}\mathbf{X}^T] = \begin{bmatrix} R_{xx}(0) & R_{xx}(1) & R_{xx}(2) & \cdots & R_{xx}(N-1) \\ R_{xx}(1) & R_{xx}(0) & R_{xx}(1) & \cdots & R_{xx}(N-2) \\ R_{xx}(2) & R_{xx}(1) & R_{xx}(0) & \cdots & R_{xx}(N-3) \\ \vdots & \vdots & & \ddots & \vdots \\ R_{xx}(N-1) & \cdots & & \cdots & R_{xx}(0) \end{bmatrix}$$

- elementi matrice \mathbf{R}_{xx} su autokorelacijske funkcije

$$\mathbf{R}_{xx} = \{R_{xx}(|k-l|)\}; \quad k, l = 0, 1, \dots, N-1$$

$$R_{xx}(|k-l|) = E[X(n+l)X(n+k)]$$

Optimalna transformacija

- transformacijska matrica \mathbf{T} treba biti određena tako da komponente matrice \mathbf{Y} budu nekorelirane što znači da je autokorelacijska funkcija

$$R_{YY}(|k-l|) = E[X(n+l)X(n+k)] = 0 \text{ za } l \neq k$$

- autokorelacijska matrica \mathbf{R}_{YY} pod navedenim uvjetom postaje pozitivna dijagonalna matrica
- za \mathbf{R}_{YY} može se odrediti:
 - N ortogonalnih svojstvenih (vlastitih) vektora (*eigenvectors*) $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N$
 - N svojstvenih vrijednosti (*eigenvalues*) $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N \geq 0$
- Karhunen-Loèveova transformacija (KLT) definirana je kao

$$\mathbf{T} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N]^T$$

Optimalna transformacija

- za autokorelacijsku matricu transformacijskih koeficijenata \mathbf{R}_{YY} vrijedi:

$$\mathbf{R}_{YY} = E[\mathbf{Y}\mathbf{Y}^T] = E[\mathbf{T}\mathbf{X}\mathbf{X}^T\mathbf{T}] = \mathbf{T}\mathbf{R}_{XX}\mathbf{T}^T = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & & \\ 0 & 0 & \dots & \lambda_N \end{bmatrix}$$

- oblik autokorelacijske matrice \mathbf{R}_{YY} pokazuje da su komponente matrice \mathbf{Y} nekorelirane
- KLT potpuno dekorelira transformacijske koeficijente, dodjeljuje maksimalnu snagu minimalnom broju koeficijenta i postiže najveću djelotvornost kompresije
- nedostatak KLT je velika složenost proračunavanja jer se bazni vektori mijenjaju promjenom statističkih značajki signala

Diskretna kosinusna transformacija

- diskretna kosinusna transformacija (DCT, *Discrete Cosine Transform*)
 - svojim se svojstvima približava optimalnoj transformaciji
 - postoje brzi algoritmi za njezino proračunavanje
 - rabi se u gotovo svim normama za kodiranje slike i videosignala

- jednodimenzijska DCT (1D DCT)

$$F(u) = \sqrt{\frac{2}{M}} \cdot C(u) \cdot \sum_{i=0}^{M-1} f(i) \cdot \cos \frac{(2 \cdot i + 1) \cdot u \cdot \pi}{2 \cdot M}$$

$F(u)$ - transformacijski koeficijent

$f(i)$ - amplituda elementa slike

u - redni broj transformacijskog koeficijenta

i - redni broj ulaznog elementa slike

$$C(u) = \begin{cases} \frac{1}{\sqrt{2}} & \text{za } u = 0 \\ 1 & \text{za } u = 1, 2, \dots, M-1 \end{cases}$$

- DCT se može shvatiti kao razvoj funkcije $f(i)$ po ortogonalnim funkcijama

$$W(i, u) = \sqrt{\frac{2}{M}} \cdot C(u) \cdot \cos \frac{(2 \cdot i + 1) \cdot u \cdot \pi}{2 \cdot M}, \quad i, u = 0, 1, \dots, M-1$$

- uvjet ortogonalnosti:

$$\sum_{i=0}^{M-1} W(i, p) \cdot W(i, q) = \begin{cases} = 0 & \text{za } p \neq q \\ \neq 0 & \text{za } p = q \end{cases}$$

Diskretna kosinusna transformacija

- dvodimenzijaska DCT (2D DCT)

$$F(u,v) = \frac{2}{\sqrt{M \cdot N}} \cdot C(u) \cdot C(v) \cdot \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} f(i,j) \cdot \cos\left[\frac{(2 \cdot i + 1) \cdot u \cdot \pi}{2 \cdot M}\right] \cdot \cos\left[\frac{(2 \cdot j + 1) \cdot v \cdot \pi}{2 \cdot N}\right]$$

$F(u,v)$ - transformacijski koeficijent

$f(i,j)$ - amplitude elemenata slike u bloku

u, v - koordinate u području transformacije (prostorne frekvencije)

i, j - koordinate u području elemenata slike

$C(u)=C(v) = (1/2)^{1/2}$, za $u,v = 0$

$C(u)=C(v) = 1$, za $u = 1,2,...M-1, v = 1,2,...N-1$

$u=v=0$ - $F(0,0)$ - istosmjerni (DC, *Direct Current*) koeficijent

$u,v \neq 0$ - $F(u,v)$ - izmjenični (AC, *Alternating Current*) koeficijenti

- 2D inverzna DCT (2D IDCT)

$$f(i,j) = \frac{2}{\sqrt{M \cdot N}} \cdot \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} C(u) \cdot C(v) \cdot F(u,v) \cdot \cos\left[\frac{(2 \cdot i + 1) \cdot u \cdot \pi}{2 \cdot M}\right] \cdot \cos\left[\frac{(2 \cdot j + 1) \cdot v \cdot \pi}{2 \cdot N}\right]$$

Diskretna kosinusna transformacija

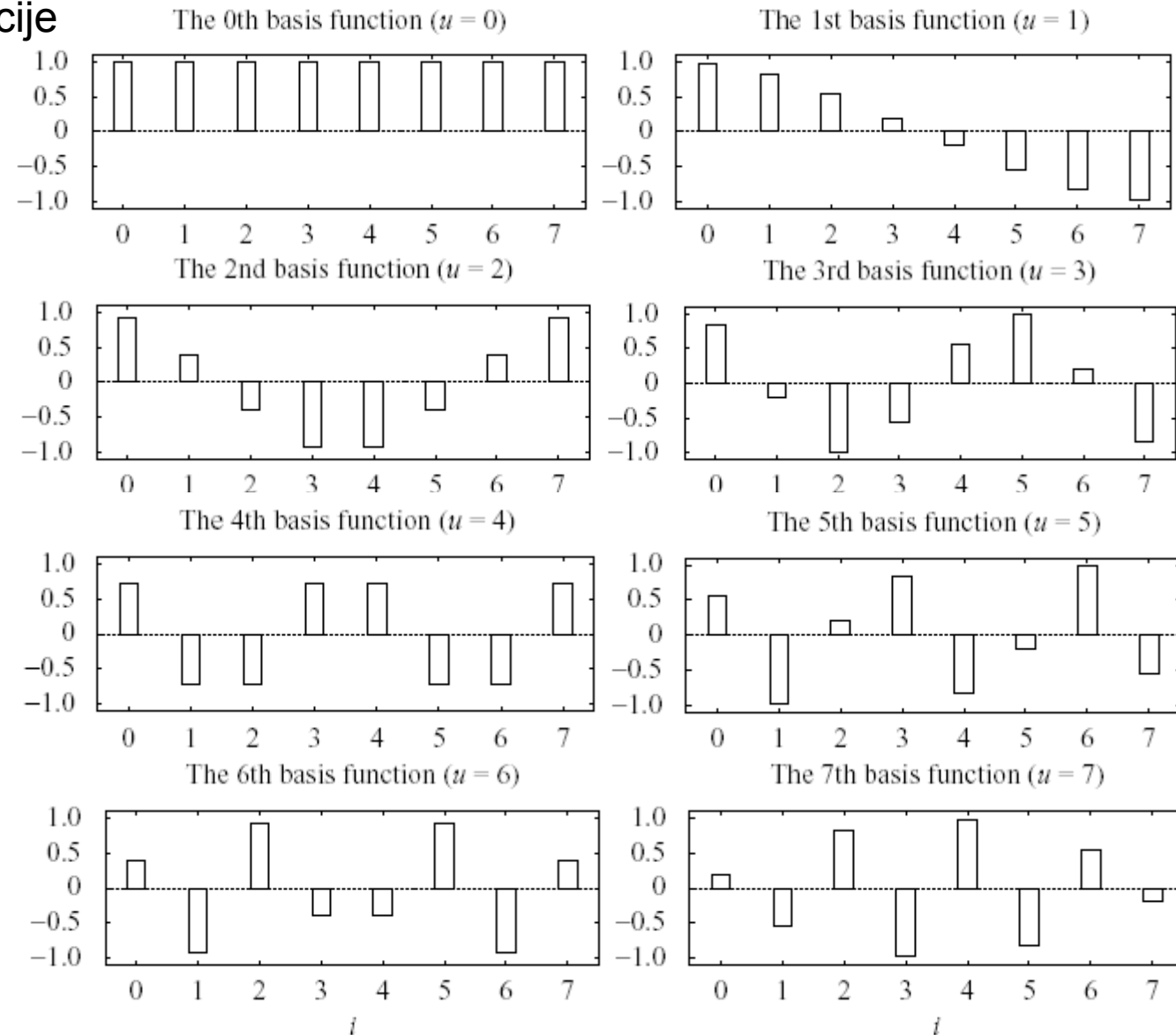
- DCT rastavlja ulazni signal na istosmjernu komponentu i izmjenične komponente komponente
 - DCT transformira uzorke ulaznog signala u frekvencijsku domenu u kojoj je provedba kompresije jednostavnija i djelotvornija
 - DCT koeficijenti imaju veći dinamički opseg od dinamičkog opsega ulaznog signala (DCT ne provodi kompresiju slike)
- IDCT rekonstruira ulazni signal
- u postupku transformacije i rekonstrukcije ulaznog signala DCT i IDCT rabe isti skup kosinusnih funkcija koje se nazivaju bazne funkcije
- bazne funkcije su ortogonalne i ortonormalne
 - bazne funkcije su ortonormalne ako su ortogonalne i ako ispunjavaju uvjet

$$\sum_{i=0}^{M-1} W(i,p) \cdot W(i,q) = 1 \quad \text{za } p = q$$

- ortonormalnost je poželjan uvjet jer osigurava da signal neće biti pojačan tijekom transformacije
- ortonormalnost se postiže odabirom $C(u)$ i $C(v)$

Diskretna kosinusna transformacija

- 1D DCT bazne funkcije
($i, u = 0, 1, \dots, 7$)



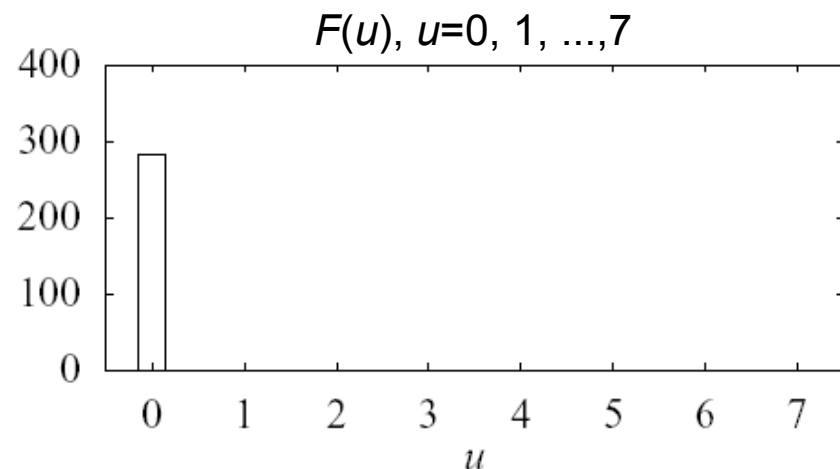
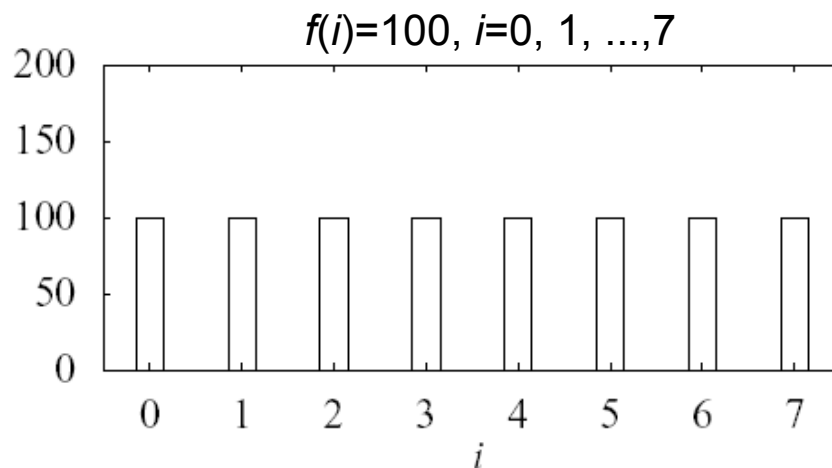
Diskretna kosinusna transformacija

- Primjer: Ulazni signal je istosmjerni signal s amplitudom 100, $f(i)=100, i=0, 1, \dots, 7$.
Potrebno je odrediti $F(u)$.

$$F(0) = \sqrt{\frac{2}{8}} \cdot \frac{1}{\sqrt{2}} \sum_{i=0}^7 100 \cdot \cos \frac{(2 \cdot i + 1) \cdot 0 \cdot \pi}{2 \cdot 8} = \frac{1}{2\sqrt{2}} 800 = 282,84$$

$$F(1) = \sqrt{\frac{2}{8}} \cdot 1 \cdot \sum_{i=0}^7 100 \cdot \cos \frac{(2 \cdot i + 1) \cdot 1 \cdot \pi}{2 \cdot 8} = \frac{1}{2} (100 \cos \frac{\pi}{16} + 100 \cos \frac{3\pi}{16} + 100 \cos \frac{5\pi}{16} + 100 \cos \frac{7\pi}{16} + 100 \cos \frac{9\pi}{16} + 100 \cos \frac{11\pi}{16} + 100 \cos \frac{13\pi}{16} + 100 \cos \frac{15\pi}{16}) = 0$$

$$F(2) = F(3) = F(4) = F(5) = F(6) = F(7) = 0$$



Diskretna kosinusna transformacija

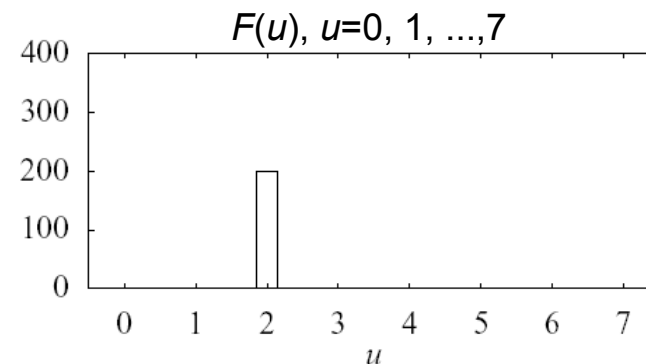
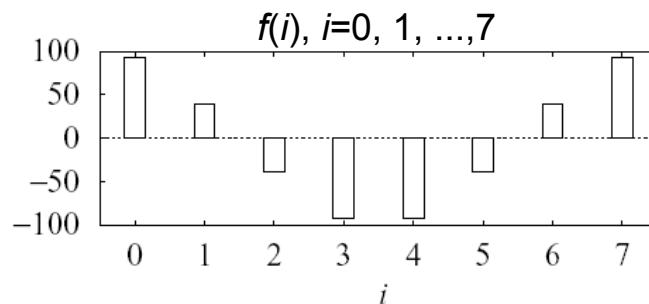
- Primjer: Ulazni signal svojim oblikom odgovara drugoj kosinusnog baznoj funkciji s amplitudom 100. Potrebno je odrediti $F(u)$.

$$f(i) = \sum_{i=0}^7 100 \cdot \cos \frac{(2 \cdot i + 1) \cdot 2 \cdot \pi}{16}$$

$$F(0) = \sqrt{\frac{2}{8}} \cdot \frac{1}{\sqrt{2}} \sum_{i=0}^7 100 \cdot \cos \frac{(2 \cdot i + 1) \cdot 2 \cdot \pi}{16} \cdot \cos \frac{(2 \cdot i + 1) \cdot 0 \cdot \pi}{2 \cdot 8} = \frac{1}{2\sqrt{2}} 100 (\cos \frac{\pi}{8} + \cos \frac{3\pi}{8} + \cos \frac{5\pi}{8} + \cos \frac{7\pi}{8} + 100 \cos \frac{9\pi}{8} + \cos \frac{11\pi}{8} + \cos \frac{13\pi}{8} + \cos \frac{15\pi}{8}) = 0$$

$$F(2) = \sqrt{\frac{2}{8}} \cdot 1 \cdot \sum_{i=0}^7 100 \cdot \cos \frac{(2 \cdot i + 1) \cdot 2 \cdot \pi}{16} \cdot \cos \frac{(2 \cdot i + 1) \cdot 2 \cdot \pi}{16} = \frac{1}{2} 100 (\cos \frac{\pi}{8} \cos \frac{\pi}{8} + \cos \frac{3\pi}{8} \cos \frac{3\pi}{8} + \cos \frac{5\pi}{8} \cos \frac{5\pi}{8} + \cos \frac{7\pi}{8} \cos \frac{7\pi}{8} + \cos \frac{9\pi}{8} \cos \frac{9\pi}{8} + \cos \frac{11\pi}{8} \cos \frac{11\pi}{8} + \cos \frac{13\pi}{8} \cos \frac{13\pi}{8} + \cos \frac{15\pi}{8} \cos \frac{15\pi}{8}) = \frac{1}{2} (1 + 1 + 1 + 1) = 200$$

$$F(1) = F(3) = F(4) = F(5) = F(6) = F(7) = 0$$



Diskretna kosinusna transformacija

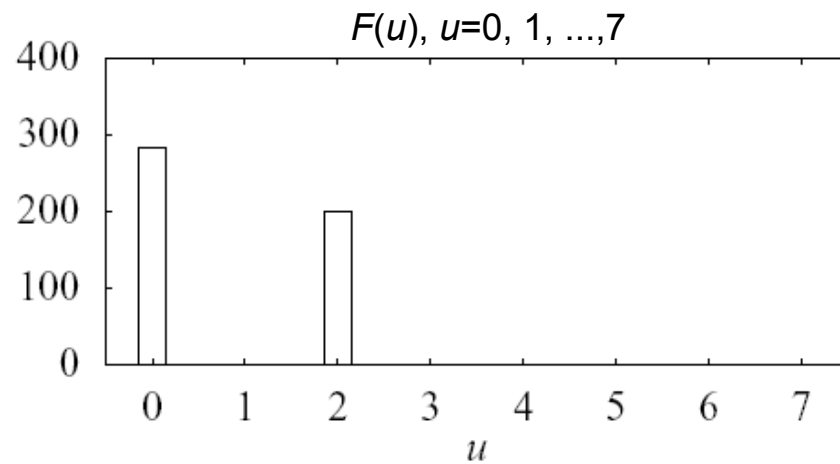
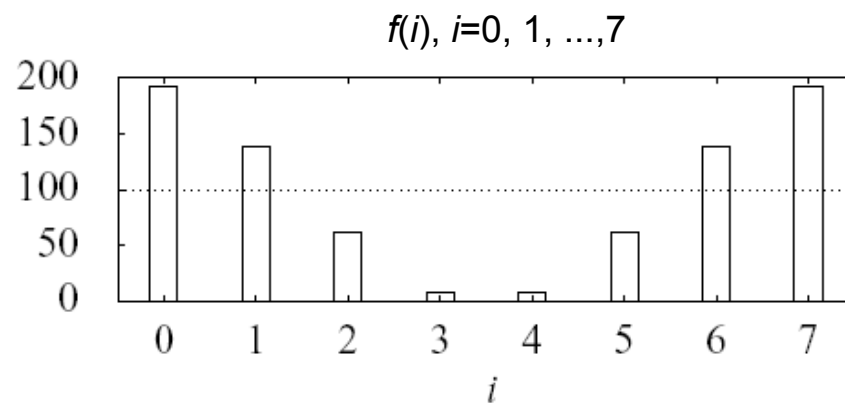
- Primjer: Ulazni signal je zbroj signala iz prethodna dva primjera. Potrebno je odrediti $F(u)$.

$$f(i) = \sum_{i=0}^7 100 + 100 \cdot \cos \frac{(2 \cdot i + 1) \cdot 2 \cdot \pi}{16}$$

$$F(0) = 282,84$$

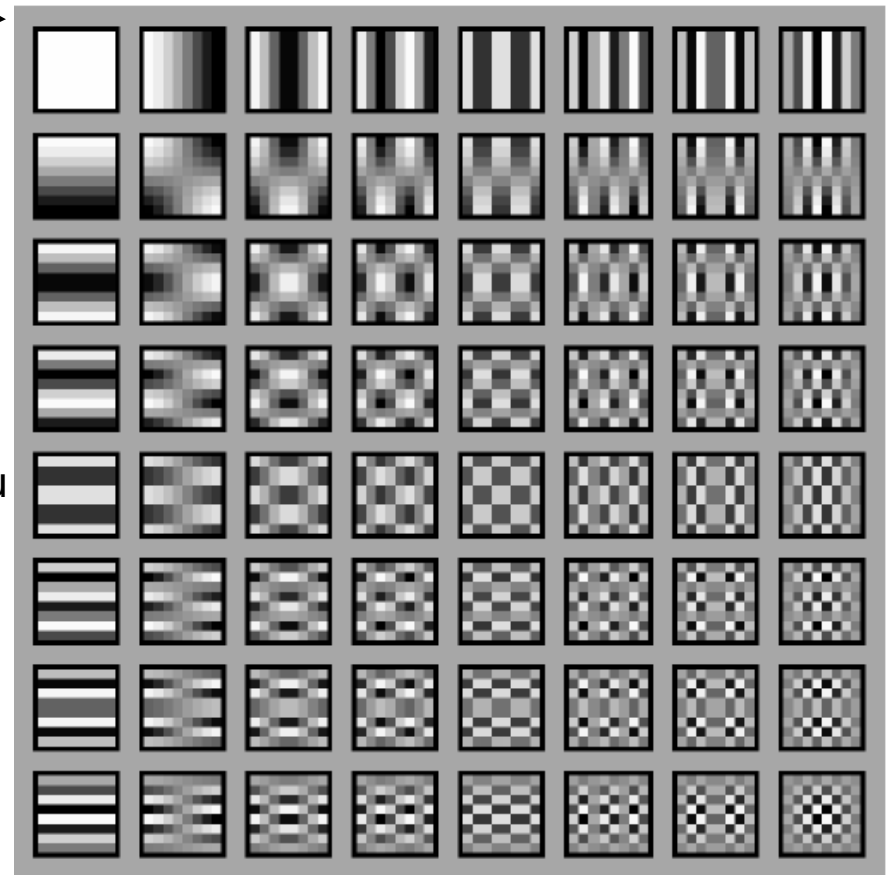
$$F(2) = 200$$

$$F(1) = F(3) = F(4) = F(5) = F(6) = F(7) = 0$$



Diskretna kosinusna transformacija

- u normama za kompresiju slike i videosignala rabi se 2D DCT s veličinom bloka 8x8 ($M=N=8$) elemenata slike
- 2D DCT bazne funkcije uz $M=N=8$ →
 - bijelo označava pozitivne vrijednosti, a crno negativne vrijednosti bazne funkcije
 - svakom DCT koeficijentu pridružena je jedna bazna funkcija
 - bazne funkcije određuju značenje pojedinih DCT koeficijenata $F(u,v)$ u odnosu na horizontalnu (u) i vertikalnu (v) frekvenciju
 - ovdje prikazane bazne funkcije predstavljaju uzorak elemenata slike (*pixel pattern*)



Diskretna kosinusna transformacija

- 8x8 2D DCT

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

$$F(0,0) = \frac{1}{8} \cdot \sum_{i=0}^7 \sum_{j=0}^7 f(i,j) \longrightarrow \text{osmerostruka srednja vrijednost elemenata slike u bloku}$$

- za proračunavanje svakog $F(u,v)$ koeficijenta potrebno je proći kroz sve vrijednosti elemenata slike $f(i,j)$
 - npr. za proračunavanje trećeg koeficijenta u drugom redu ($u=2, v=1$, $C(u)=C(v)=1$) potrebne su 64 operacije zbrajanja

$$F(2,1) = \frac{1}{4} \sum_{i=0}^7 \sum_{j=0}^7 f(i,j) \cdot \cos\left[\frac{(2 \cdot i + 1) \cdot 2 \cdot \pi}{16}\right] \cdot \cos\left[\frac{(2 \cdot j + 1) \cdot \pi}{16}\right]$$

Diskretna kosinusna transformacija

- 8x8 2D DCT može biti prikazana kao produkt dvije 1D DCT što smanjuje potreban broj operacija za proračunavanje DCT

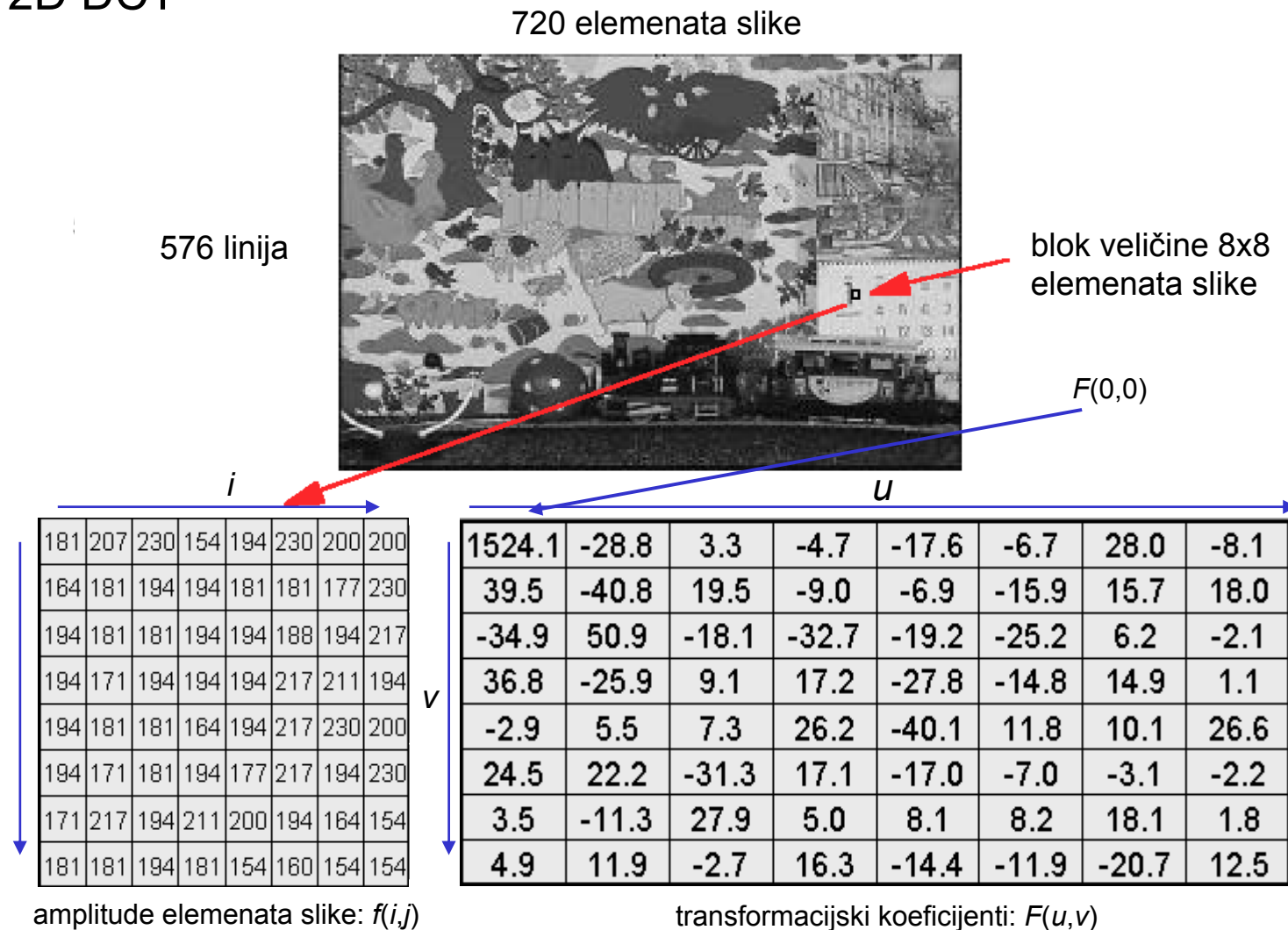
$$G(i,v) = \frac{1}{2} \cdot C(v) \cdot \sum_{j=0}^7 f(i,j) \cdot \cos\left[\frac{(2 \cdot j + 1) \cdot v \cdot \pi}{16}\right]$$

$$F(u,v) = \frac{1}{2} \cdot C(u) \cdot \sum_{i=0}^7 G(i,v) \cdot \cos\left[\frac{(2 \cdot i + 1) \cdot u \cdot \pi}{16}\right]$$

- za proračunavanje svakog $F(u,v)$ u tom slučaju potrebno je 16 operacija zbrajanja

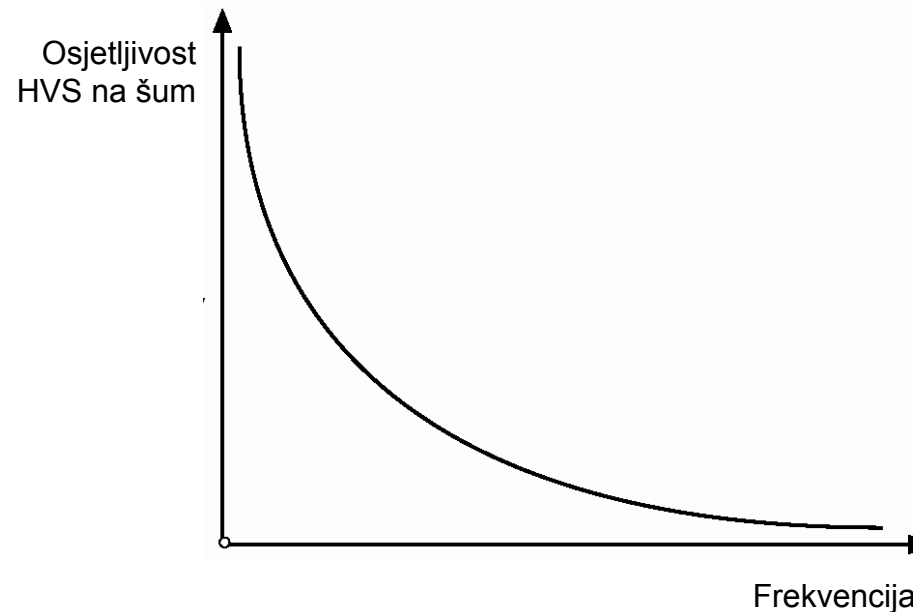
Diskretna kosinusna transformacija

- 8x8 2D DCT



Kvantizacija DCT koeficijenata

- cilj kvantizacije je postizanje kompresije
 - kvantizacijom se odbacuju informacije koje nisu važne za ljudski vizualni sustav (HVS, *Human Visual System*)
 - HVS je više osjetljiv na šum u području niskih frekvencija nego u području visokih frekvencija
 - HVS je vrlo osjetljiv na pogreške DC koeficijenta i niskofrekvencijskih (NF) AC koeficijenata
 - ljudski vizualni sustav nije osjetljiv na pogreške visokofrekvencijskih (VF) AC koeficijenata



Kvantizacija DCT koeficijenata

- kvantizacija se provodi uz pomoć kvantizacijske tablice sa 64 vrijednosti $q(u,v)$, s kojima se dijele DCT koeficijenti $F(u,v)$, a rezultat se zaokružuje na najbližu cjelobrojnu vrijednost
 - vrijednosti u kvantizacijskoj tablici $q(u,v)$ određuju korak kvantizacije za odgovarajući DCT koeficijent $F(u,v)$
 - vrijednosti $q(u,v)$ određene su u skladu sa značajkama HVS (manje su za DC koeficijent i NF AC koeficijente te rastu porastom prostorne frekvencije)
 - vrijednosti nastale kvantizacijom $S(u,v)$ su:

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

$S(u,v)$ - kvantizirana vrijednost DCT koeficijenta
 $Q(u,v)$ - koeficijent kvantiziranja (umnožak vrijednosti iz kvantizacijske tablice $q(u,v)$ i faktora skaliranja S)
 - promjenom faktora skaliranja S mijenja se stupanj kompresije i kvaliteta slike
- u dekoderu se obavlja inverzni postupak, tj. kvantizirani DCT koeficijent množi se s odgovarajućim koeficijentom kvantiziranja

$$F'(u,v) = S(u,v)Q(u,v)$$

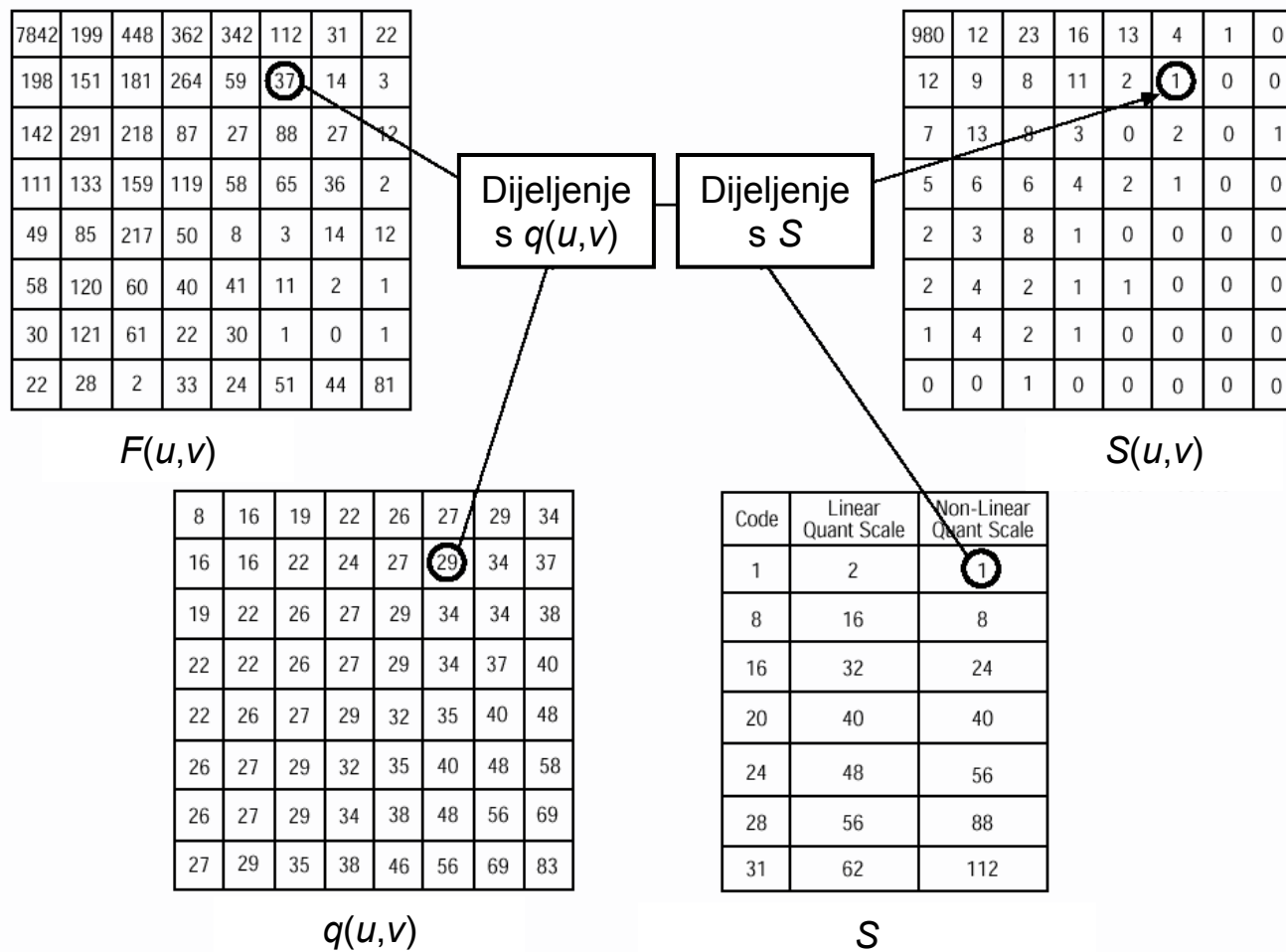
Kvantizacija DCT koeficijenata

Primjer kvantizacijske tablice $q(u,v)$
za luminantni blok (MPEG)

8	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

Kvantizacija DCT koeficijenata

- primjer kvantizacije



Kvaliteta slike

- preveliki stupanj kompresije u sustavima s DCT dovodi do izobličenja koje se očituje kao vidljivost rubova blokova u slici

Izvorna ispitna slika
Barbara



Dekodirana slika, stupanj
kompresije 20:1



Razlika izvorne i
dekodirane slike



Kodiranje između slika

- kodiranje između slika (*interframe coding*)
 - kodira se razlika slika i uklanja vremenska redundancija
 - do dekodera se prenosi slika A i slika $C=(A-B)$
 - slika C može biti djelotvornije kodirana primjenom kodiranja unutar slike od slike B
 - slika B se rekonstruira u dekoderu kao $(A+C)$
- ukoliko se radi o sekvenci s vrlo brzim pokretima, kodiranjem razlike uzastopnih slika ne može se postići velika kompresija
 - razlika između uzastopnih slika se smanjuje postupkom predviđanja i nadomještanja pokreta (*motion compensation*)



A



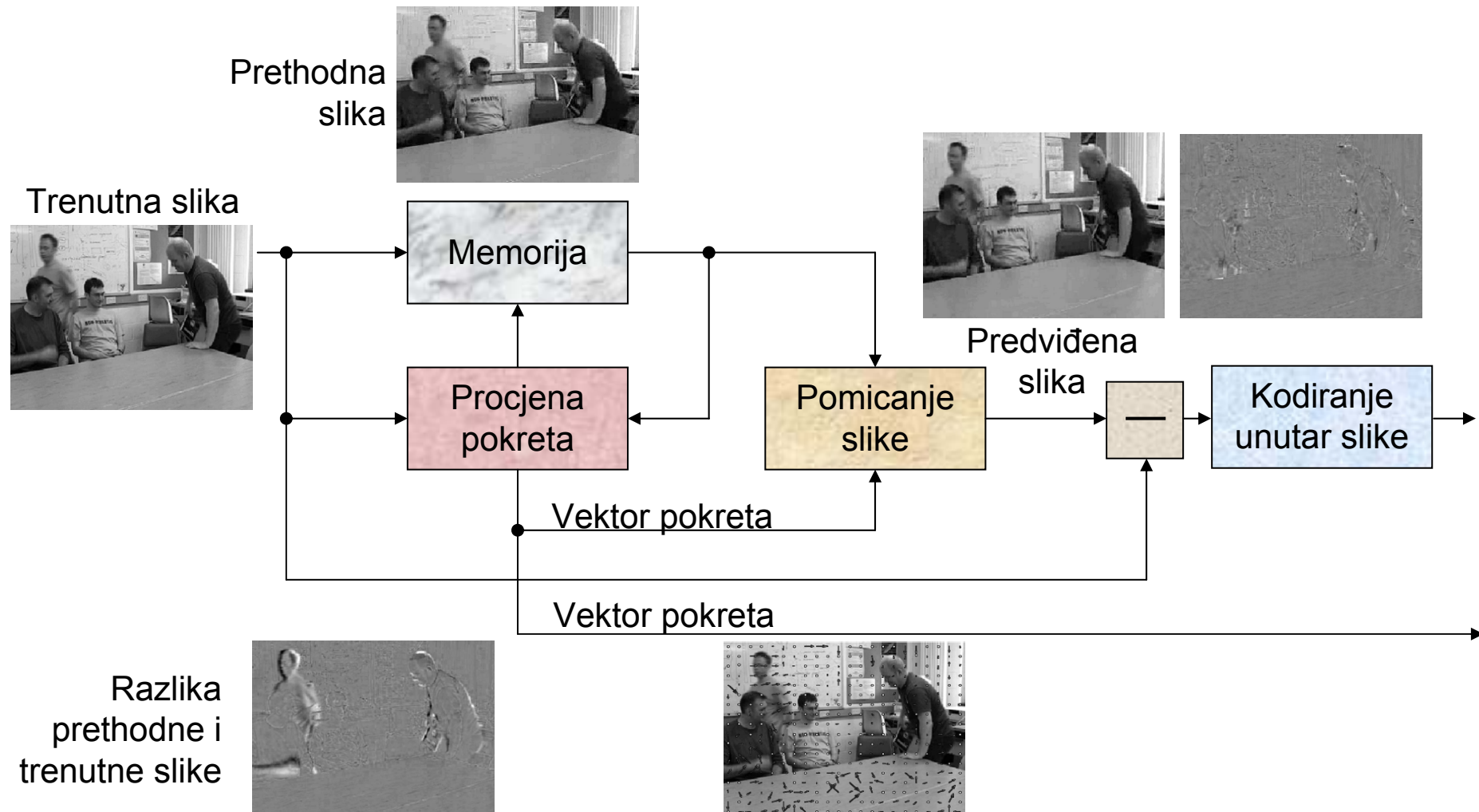
B



C

Kodiranje između slika

- kodiranje s predviđanjem uz nadomještanje pokreta



Kodiranje između slika

- kodiranje s predviđanjem uz nadomještanje pokreta
 - primjenjuje se za kodiranje razlike slika u sekvenci videosignala s visokom razinom pokreta koja uzrokuje pojavu velike razlike između uzastopnih slika
 - provodi se usporedba trenutne i prethodne slike, kako bi se odredio smjer i udaljenost pomaka pokretnih objekata između slika
 - smjer i udaljenost pomaka objekata iskazuje preko dvodimenzijskog vektora pokreta
 - koder koristi vektore pokreta za dobivanje predviđene slike, koja nastaje pomakom elemenata slike prethodne slike za vektore pokreta (na taj način smanjuje se razlika između trenutne i prethodne slike)
 - predviđena slika se oduzima od trenutno procesirane slike
 - do dekodera se prenosi razlika trenutne i predviđene slike, te vektori pokreta

Kodiranje između slika

- kodiranje s predviđanjem uz nadomještanje pokreta

Prethodna slika



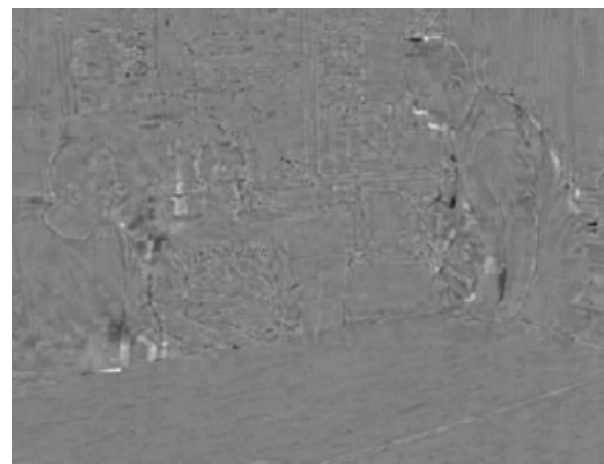
Trenutna slika



Razlika prethodne
i trenutne slike



Razlika trenutne
i predviđene
slike

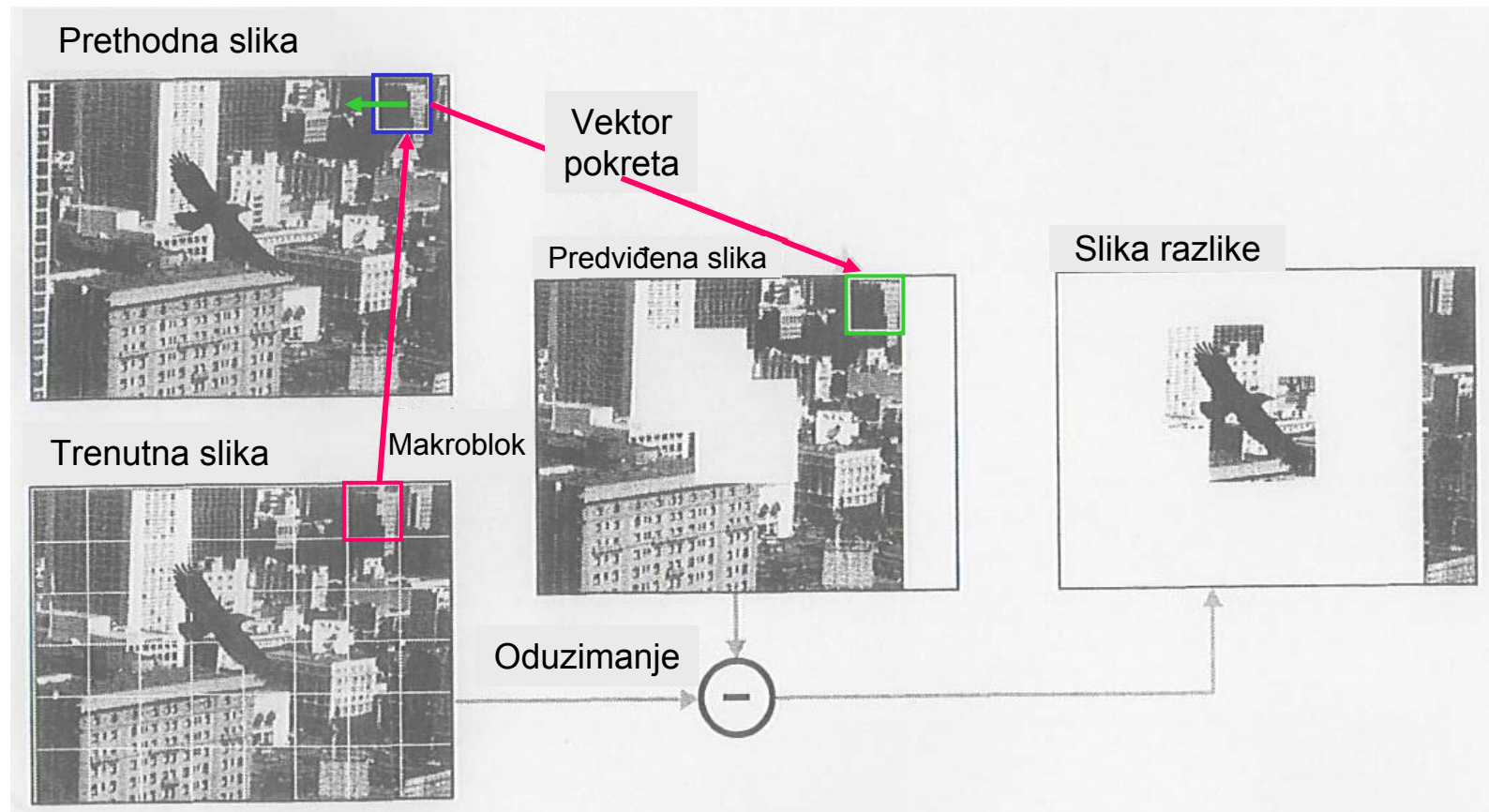


Kodiranje između slika

- nadomještanje pokreta se provodi za kvadratična područja uobičajene veličine 16x16 (makroblok)
 - makroblok iz trenutne slike traži se u prethodnoj slici
 - pretraživanje se obično ne provodi po cijeloj slici već se definira ograničeno područje pretraživanja (radi brže provedbe pretraživanja)
 - ukoliko određeni makroblok postoji u prethodnoj slici, izvodi se vektor pokreta koji pokazuje razliku položaja određenog makrobloka između prethodne i trenutne slike
 - ukoliko određenog makrobloka nema u prethodnoj slici ne provodi se predviđanje pokreta
- predviđena slika nastaje primjenom vektora pokreta na makroblokove u prethodnoj slici
- predviđena slika se oduzima od trenutne slike i do dekodera se prenosi slika razlike

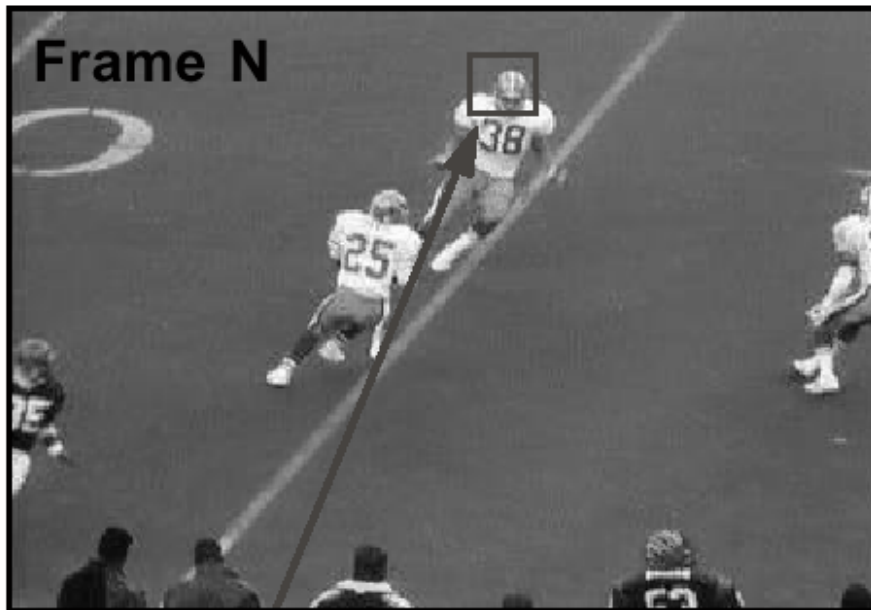
Kodiranje između slika

- predviđanje uz nadomještanje pokreta



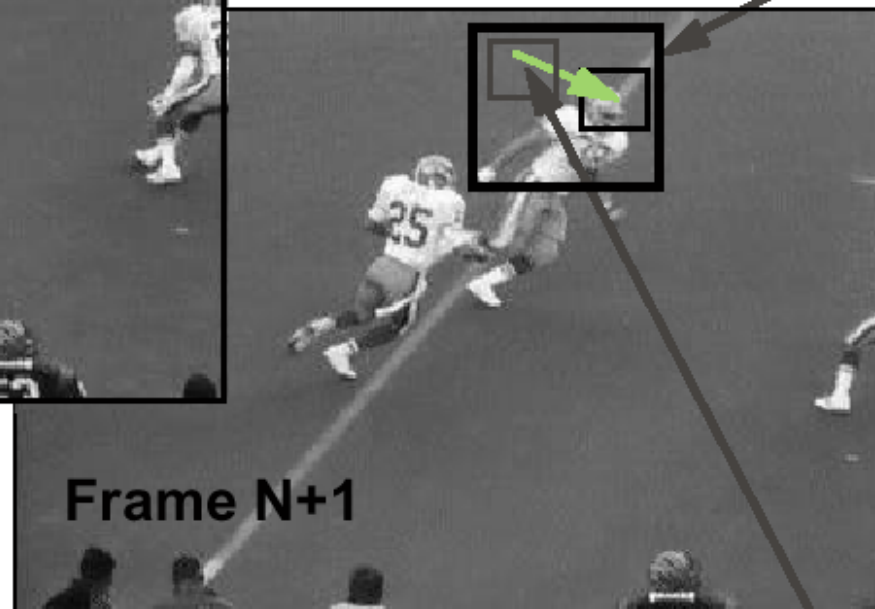
Kodiranje između slika

Prethodna slika



Makroblok 16x16
elemenata slike

Trenutna slika



Područje
pretraživanja

Vektor pokreta

Primjer rezultata kodiranja

- videosekvenca "Vectra", MPEG-2 kodiranje, 0,768 Mbit/s

1



2



3



4



5



6



Primjer rezultata kodiranja

- vektori pokreta

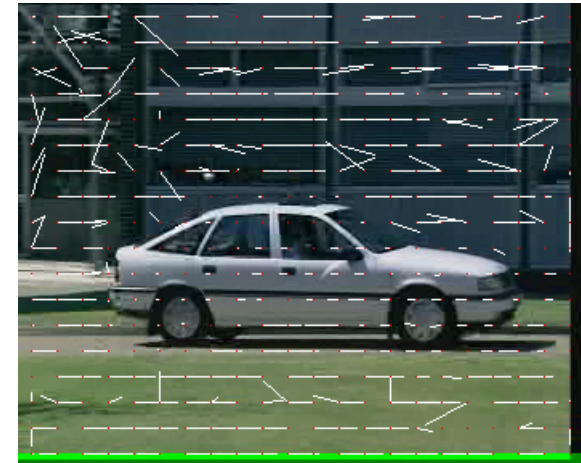
1



2



3



4



5



6



Primjer rezultata kodiranja

- slika razlike

1



2



3



4



5



6



Norme za kompresiju slike i videa

- normizacijska tijela

CCITT	International Telegraph and Telephone Consultative Committee
ITU	International Telecommunication Union
JPEG	Joint Photographic Experts Group
MPEG	Moving Picture Experts Group
ISO	International Standards Organization
IEC	International Electrotechnical Commission

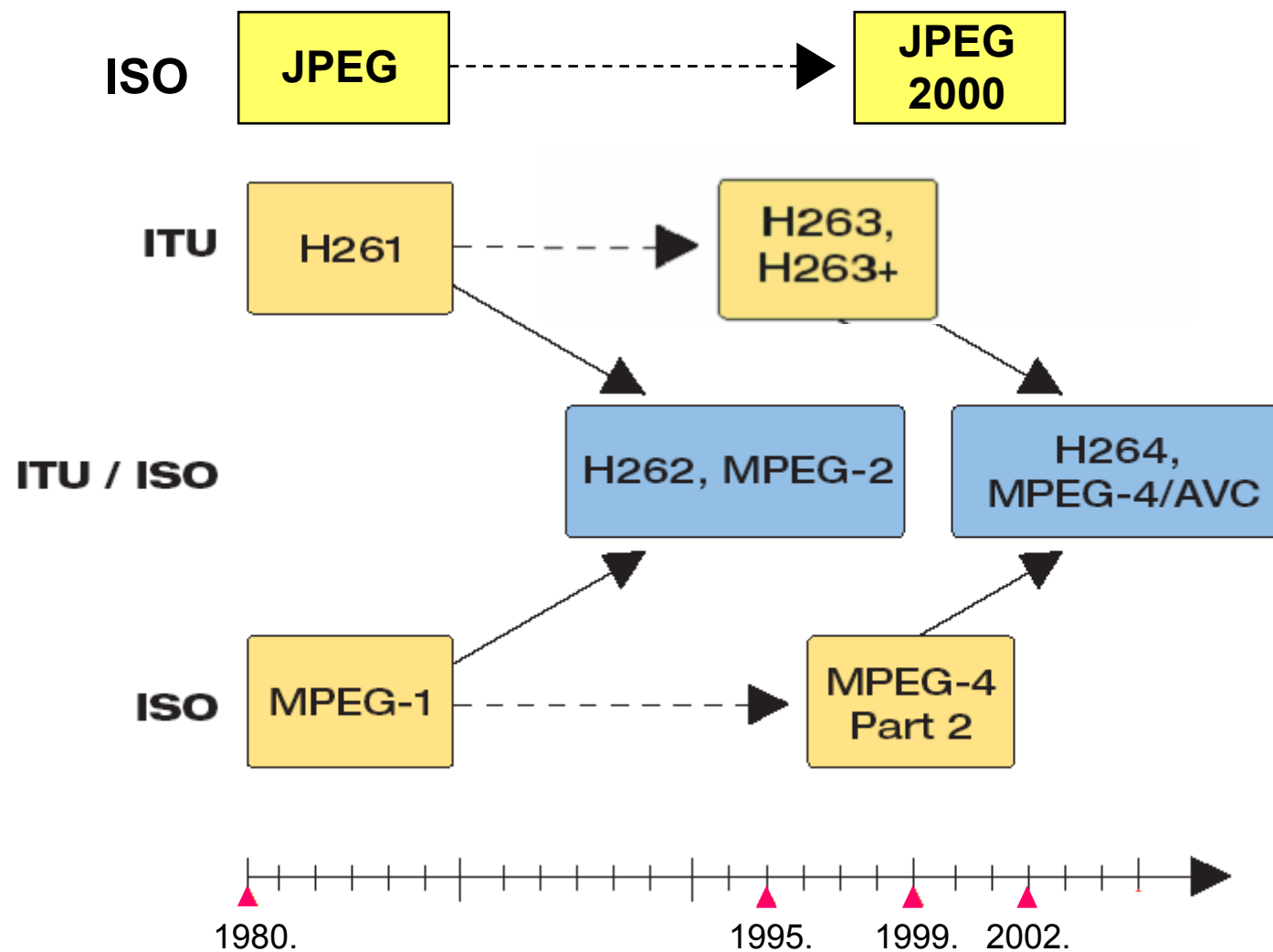
Norme za kompresiju slike i videa

- pregled normi za kompresiju

JPEG	1992	For still image coding, DCT based
JPEG-2000	2000	For still image coding, DWT based
H.261	1990	For videoconferencing, 64Kbps to 1.92 Mbps
MPEG-1	1991	For CD-ROM, 1.5 Mbps
MPEG-2 (H.262)	1994	For DTV, 2 to 15 Mbps, most extensively used
H.263	1995	For very low bit rate coding, below 64 Kbps
H.263+ (version 2)	1998	Add new optional features to H.263
MPEG-4	1999	For multimedia, content-based coding
MPEG-4 (version 2)	2000	Adds more tools to MPEG-4
H.263++	2000	Adds more optional features to H.263+
H.264	2004	Functionally different, much more efficient

Norme za kompresiju slike i videa

- razvoj normi za kompresiju



Norme za kompresiju mirnih slika

- temelje se na primjeni transformacijskog kodiranja
- ISO/IEC IS 10918: *Digital Compression and Coding of Continuous Tone Still Images*, tzv. **JPEG** (*Joint Photographic Experts Group*) norma
 - temelji se na primjeni diskretne kosinusne transformacije (DCT, *Discrete Cosine Transform*)
 - namijenjena za područje 0,25-2 bpp
- ISO/IEC IS 15444-1: *JPEG2000 Image Coding System-Part 1: Core Coding System*, tzv. **JPEG2000** norma
 - temelji se na primjeni diskretne wavelet transformacije (DWT, *Discrete Wavelet Transform*)
 - pri višim stupnjevima kompresije postiže bolju kvalitetu slike u odnosu na JPEG

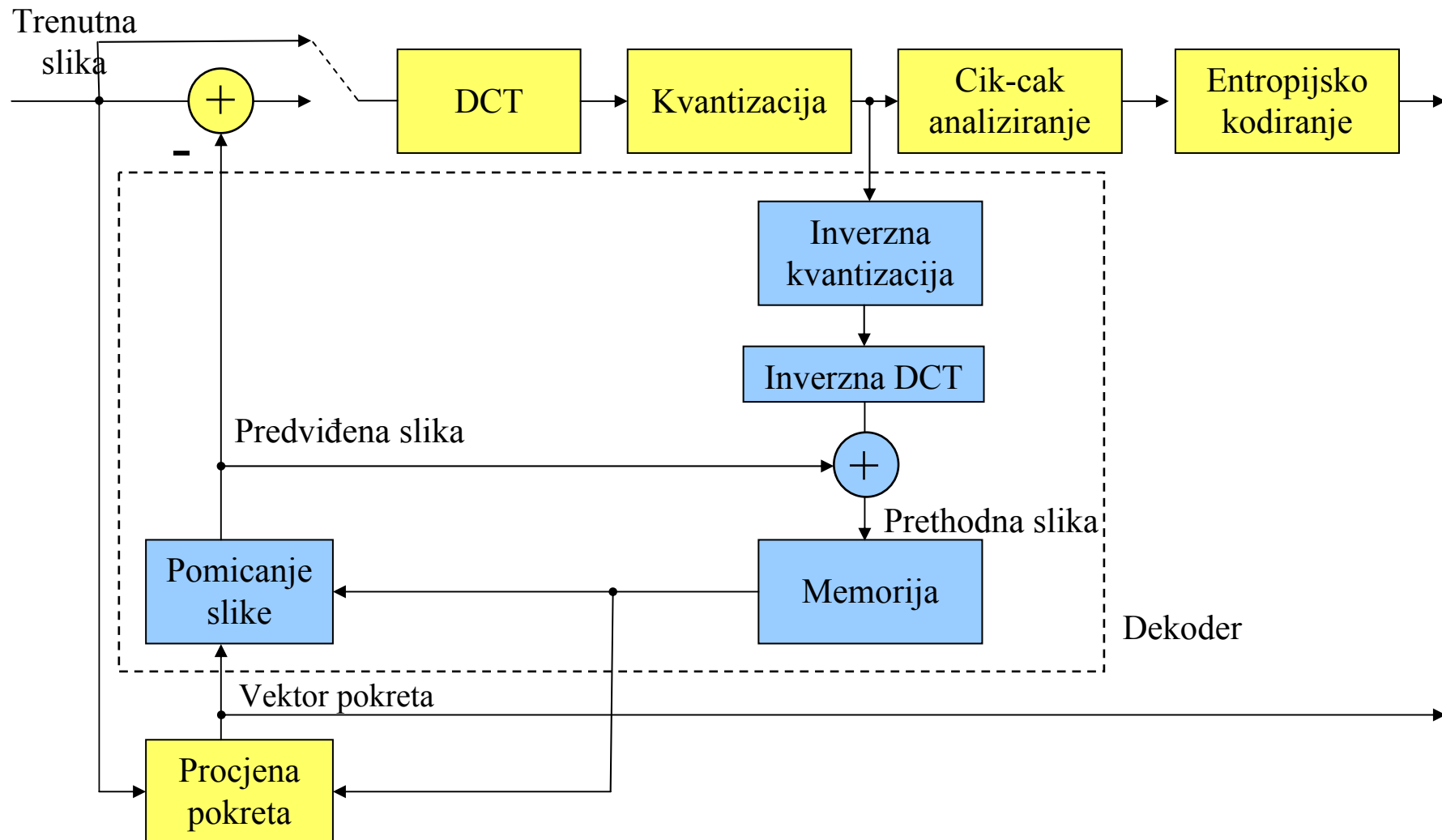
Norme za kompresiju videosignala

- ITU-T preporuka H.261

- Videokodek za audiovizualne usluge pri brzinama prijenosa od $p \times 64$ kb/s ("*Video Codec for Audiovisual Services at $p \times 64$ kb/s*") iz 1990.
- videokonferencijske i videotelefonske primjene putem digitalne mreže s integriranim uslugama (ISDN, *Broadband Integrated Services Digital Network*)
- brzine prijenosa: 64 kbit/s-2,048Mbit/s
 - brzine prijenosa pri kojima je $p < 6$ su namijenjene videotelefonskim primjenama
 - za brzine prijenosa za koje je $p > 6$ moguć prijenos složenih slika uz višu kvalitetu
- format ulazne slika
 - zajednički međufORMAT (CIF, *Common Intermediate Format*)
 - luminantni signal (352 elementa slike i 288 linija)
 - komponente krominantnog signala (176 elemenata slike i 144 linije)
 - dodatni format: QCIF (*Quarter CIF*) za videotelefonske signale
- maksimalna frekvencija izmjene slika: 30 Hz
- algoritam kodiranja (MC+DCT)
 - DCT za uklanjanje prostorne redundancije unutar pojedine slike
 - DPCM uz kompenzaciju pokreta (MC) za uklanjanje vremenske redundancije

Norme za kompresiju videosignala

- MC+DCT koder



Norme za kompresiju videosignala

- ITU-T preporuka H.263(v1)

- Kodiranje videosignala za komunikaciju niskom brzinom prijenosa ("*Video Coding for Low Bitrate Communication*") iz 1995.
- zamijenila je ITU-T preporuku H.261 u videokonferencijskim primjenama putem ISDN mreže, a rabi se i u videokonferencijskim sustavima putem paketskih mreža (LAN ili IP videokonferencije) te za prijenos videosignala putem Interneta, video na zahtjev i sl.
- ima poboljšane značajke algoritma za kompresiju u odnosu na H.261
 - mogućnost postizanja jednake kvalitete slike pri dvostruko nižim brzinama prijenosa
- brzine prijenosa: 20 kbit/s -12 Mbit/s
- formati ulazne slike: SQCIF (Sub-QCIF), QCIF, CIF, 4CIF, 16CIF
- algoritam kodiranja: MC+DCT

- 1998. usvojena je druga inačica preporuke ITU-T H.263 (H.263v2 ili H.263+), a 2000. treća inačica (H.263v3 ili H.263++)
 - zadržan je kompletan tehnički sadržaj prve inačice preporuke uz dodane nove opcije kojima se postiže veća djelotvornost postupka kompresije i otpornost na pogreške

Norme za kompresiju videosignala

- MPEG-1 norma

- ISO/IEC IS 11172 - Kodiranje pokretnih slika i pratećih audiosignala za digitalno pohranjivanje pri brzinama do 1,5 Mbit/s ("*Coding of Moving Pictures and Associated Audio for Digital Storage Media up to about 1.5 Mb/s*") iz 1993.
- algoritam kompresije videosignala: MC+DCT
 - uvodi se dvosmjerno predviđanje pokreta iz prethodnih i budućih slika (*bi-directional motion prediction*)
- za brzine prijenosa veće od 1 Mbit/s omogućava postizanje više kvalitete slike u odnosu na H.261
- podržava samo progresivno analiziranje
- ulazni format slike: SIF (*Source Input Format*)
 - 25 Hz
 - luminantni signal: 360 x 288 elemenata slike
 - komponente krominantnog signala: 180 x 144 elemenata slike
 - 30 Hz
 - luminantni signal: 360 x 240 elemenata slike
 - komponente krominantnog signala: 180 x 120 elemenata slike

Norme za kompresiju videosignala

- MPEG-2 norma

- ISO/IEC IS 13818 Informacijska tehnologija- Generičko kodiranje pokretnih slika i pratećih audiosignala ("*Generic Coding of Moving Pictures and Associated Audio*") iz 1995.
- drugi dio norme ISO/IEC IS 13818, koji se odnosi na kodiranje videosignala, usvojen je i kao ITU-T preporuka H.262
- nije ograničena s obzirom na formate ulaznog videosignala
 - podržava progresivno analiziranje i analiziranje s proredom
 - podržava 4:4:4, 4:2:2 i 4:2:0 formate
- uključuje u sebe parametre kodiranja propisane MPEG-1 normom
- podržava formate za televiziju visoke kvalitete (HDTV)
- pokriva tri područja
 - aspekti sustava (multipleksiranje i sinkronizacija), kodiranje videosignala, kodiranje audiosignala

Norme za kompresiju videosignala

- MPEG-4 norma

- ISO/IEC IS 14496 (Version 1 - 1999., Version 2 -2000., itd.): Kodiranje audio-vizualnih objekata ("*Coding of Audio-Visual Objects*")
 - norma se još razvija, a svaka nova inačica norme je kompatibilna s prethodnom inačicom (*backward-compatibility*)
 - pokriva više područja (kodiranje audio-informacija, kodiranje vizualnih informacija, sustav, povezivanje s komunikacijskim sustavima)
 - prikladna je za digitalnu televiziju, interaktivne grafičke primjene (*synthetic content*), interaktivne multimedijske primjene (www, distribucija i pristup sadržajima)
 - omogućava prijenos multimedijskih sadržaja uporabom različitih mreža: IP, ATM, mreže za mobilne komunikacije, PSTN, ISDN

- MPEG-4 Part 2

- ISO/IEC IS 14496: *Coding of Audio-Visual Objects - Part 2: Visual*
 - definira način kodiranja vizualnih informacija
 - brzine prijenosa: 5 kbit/s - 1 Gbit/s
 - kodiranje prirodnih i sintetičkih slika, te njihove kombinacije
 - nema ograničenja na formate ulaznog videosignala (od QCIF do vrlo visokih rezolucija (4kx4k))
 - podržava progresivno analiziranje i analiziranje s proredom

Norme za kompresiju videosignala

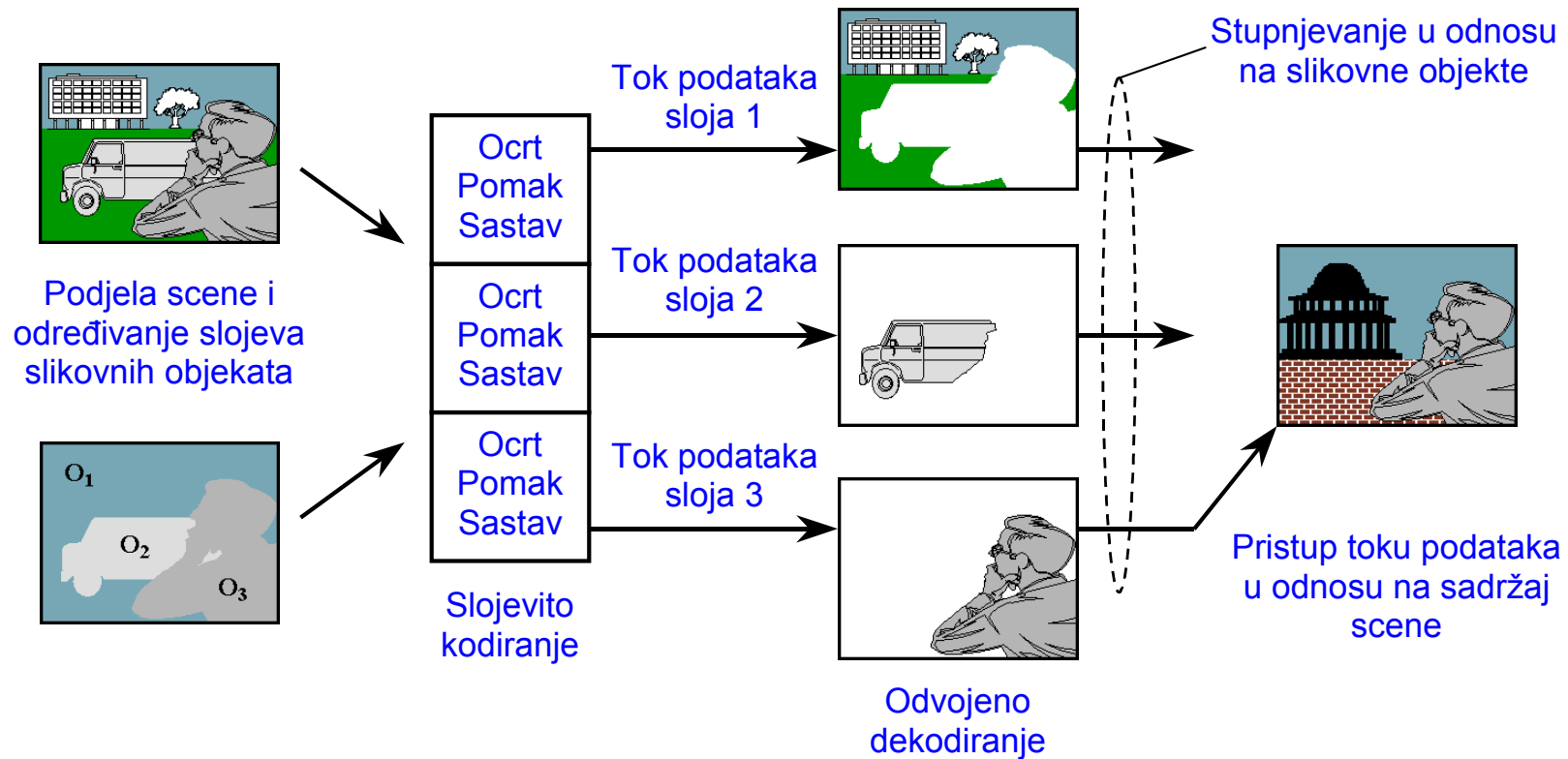
– MPEG-4 Part 2

- sadrži fleksibilne alate za kodiranje različitih vrsta vizualnih informacija
 - kodiranje standardnih videosignala kod kojih su slike pravokutnog oblika
 - kodiranje video-objekata pri čemu se slika dijeli na objekte proizvoljnog oblika koji se odvojeno kodiraju
 - kodiranje mirnih slika
 - hijerarhijsko kodiranje uz više rezolucijskih razina ili razina kvalitete
 - kodiranje videosignala vrhunske kvalitete za studijske primjene
- alati za kodiranje su organizirani u profile pri čemu svaki profil sadrži skupinu alata prilagođenu određenoj primjeni
 - profili se dijele u tri skupine u ovisnosti o tome da li sadrže alate za kodiranje prirodnih slika, sintetičkih slika ili njihove kombinacije (za sada 19 profila)
 - primjena većine alata je vrlo složena i onemogućava kodiranje u stvarnom vremenu
 - ukoliko se rabe samo alati koji omogućavaju aplikacije u stvarnom vremenu, postupak kodiranja postaje sličan MPEG-1 i MPEG-2 postupcima
- uporaba MPEG-4 norme (Part 2)
 - Internet Streaming Media Alliance (ISMA) koju čine Apple, Cisco, IBM, Kasenna, Philips i Sun su specificirale sustav za Internet streaming koji uključuje MPEG-4 Visual i MPEG-4 Audio
 - Microsoft's Windows Media sadrži MPEG-4 koder i dekodeer (MPEG-4 Simple Profile)
 - Apple's Quicktime version 6 podržava MPEG-4

Norme za kompresiju videosignala

– MPEG-4 Part 2

- dijeljenje slike na slojeve slikovnih objekata (VOP, *Video Object Plane*) te odvojeno kodiranje i dekodiranje slojeva



Norme za kompresiju videosignala

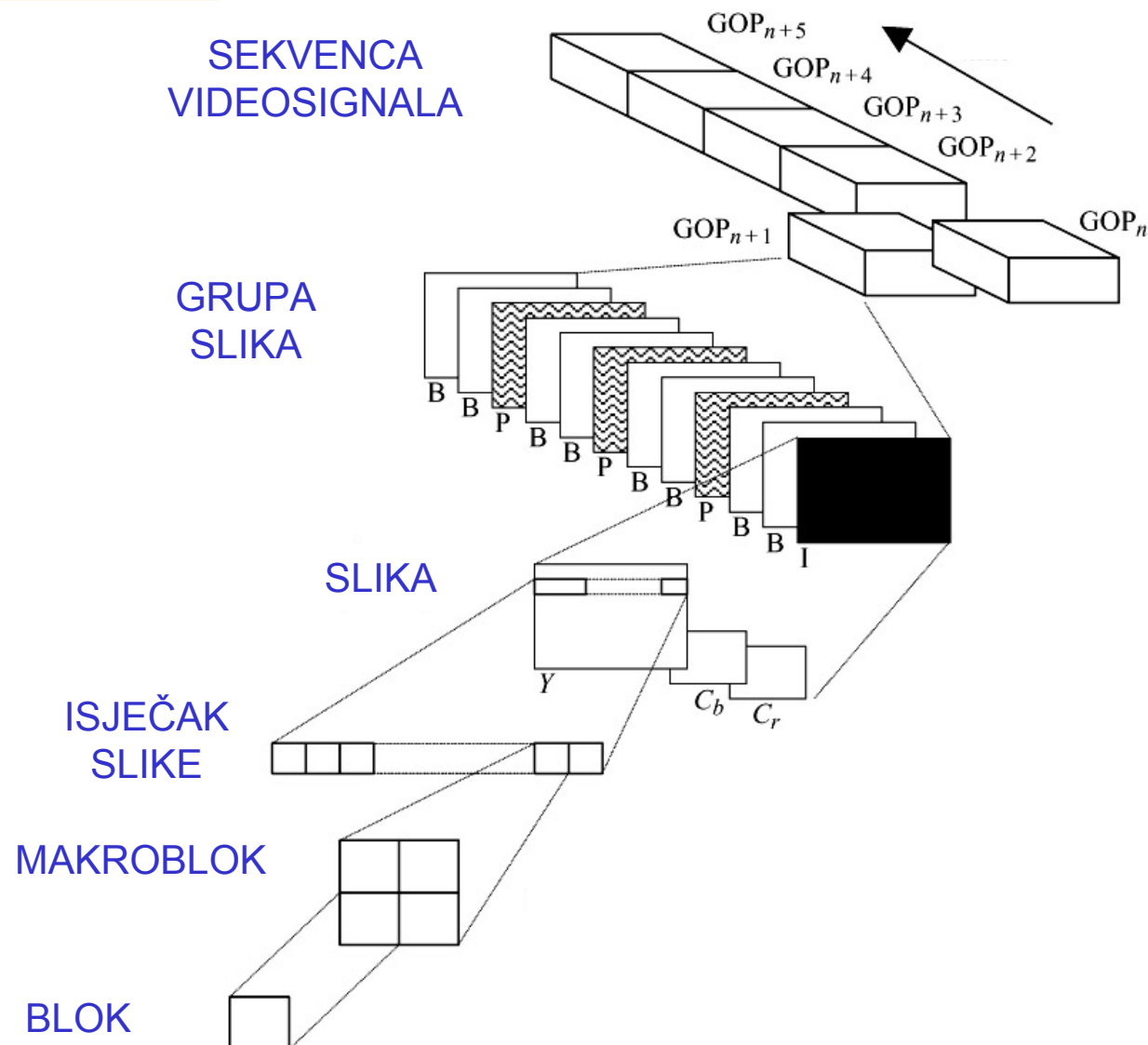
– MPEG-4 Part 10

- ISO/IEC IS 14496 *Part 10* i ITU-T Rec. H.264: Poboljšano kodiranje videosignala ("*Advanced Video Coding*") iz 2002.
 - norma se popularno naziva H.264/AVC
 - prikladna je za primjenu u sustavima kabelaške televizije, DVB sustavima, DSL mrežama, sustavima za nadzor, paketskim mrežama, UMTS sustavu, itd.
 - algoritam kodiranja je kombinacija postupka predviđanja između slika i transformacijskog kodiranja
 - povećanje učinkovitosti kompresije u odnosu na MPEG-2
 - » fleksibilniji i složeniji postupci predviđanja pokreta, povećana točnost određivanja vektora pokreta, poboljšano entropijsko kodiranje
 - » u usporedbi s MPEG-2 normom za postizanje jednake kvalitete slike potrebna je u prosjeku dvostruko niža brzina prijenosa
 - povećana složenost koda (8x složeniji od MPEG-2 koda) i dekodera

Model MPEG-2 kodiranja

- model MPEG-2 kodiranja je organiziran u šest slojeva
 - BLOK
 - najmanja jedinica kodiranja od 8x8 elementa slike
 - MAKROBLOK
 - osnovna jedinica za kodiranje s predviđanjem pokreta od 16x16 elemenata slike
 - ISJEČAK SLIKE (*Slice*)
 - horizontalni niz makroblokova
 - SLIKA
 - osnovna jedinica u MPEG postupku kodiranja
 - GRUPA SLIKA (GOP, *Group of Pictures*)
 - niz od jedne ili više slika
 - SEKVENCA
 - niz od određenog broja grupa slika

Model MPEG-2 kodiranja

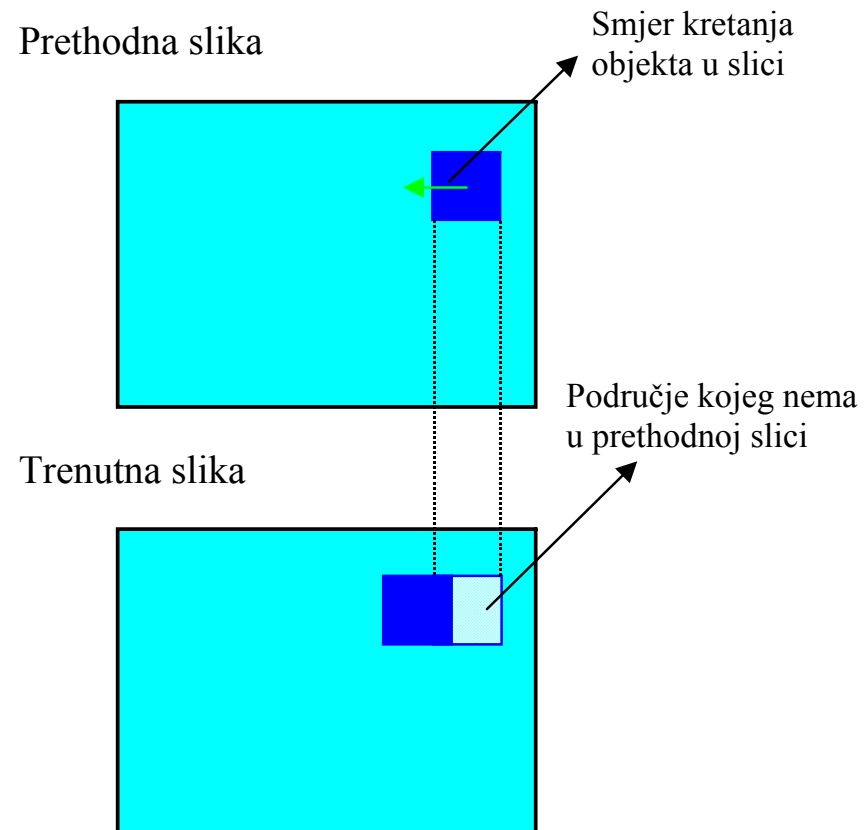


Model MPEG-2 kodiranja

- SLIKA
 - osnovna jedinica u MPEG postupku kodiranja
 - slike se dijele na:
 - I-slike (*Intra Coded Pictures*)
 - samostalno kodirane kodiranjem unutar slike primjenom DCT
 - postupak kodiranja vrlo sličan JPEG postupku
 - P-slike (*Predictive Coded Pictures*)
 - kodirane su s predviđanjem uz nadomještanje pokreta (MC, *Motion Compensated prediction*)
 - kao polazne slike u postupku predviđanja koriste se prethodne I-slike i P-slike
 - B-slike (*Bidirectionally Coded Pictures*)
 - kodirane su s predviđanjem uz nadomještanje pokreta
 - kao polazne slike u postupku predviđanja koriste se i prethodne i buduće I-slike i P-slike

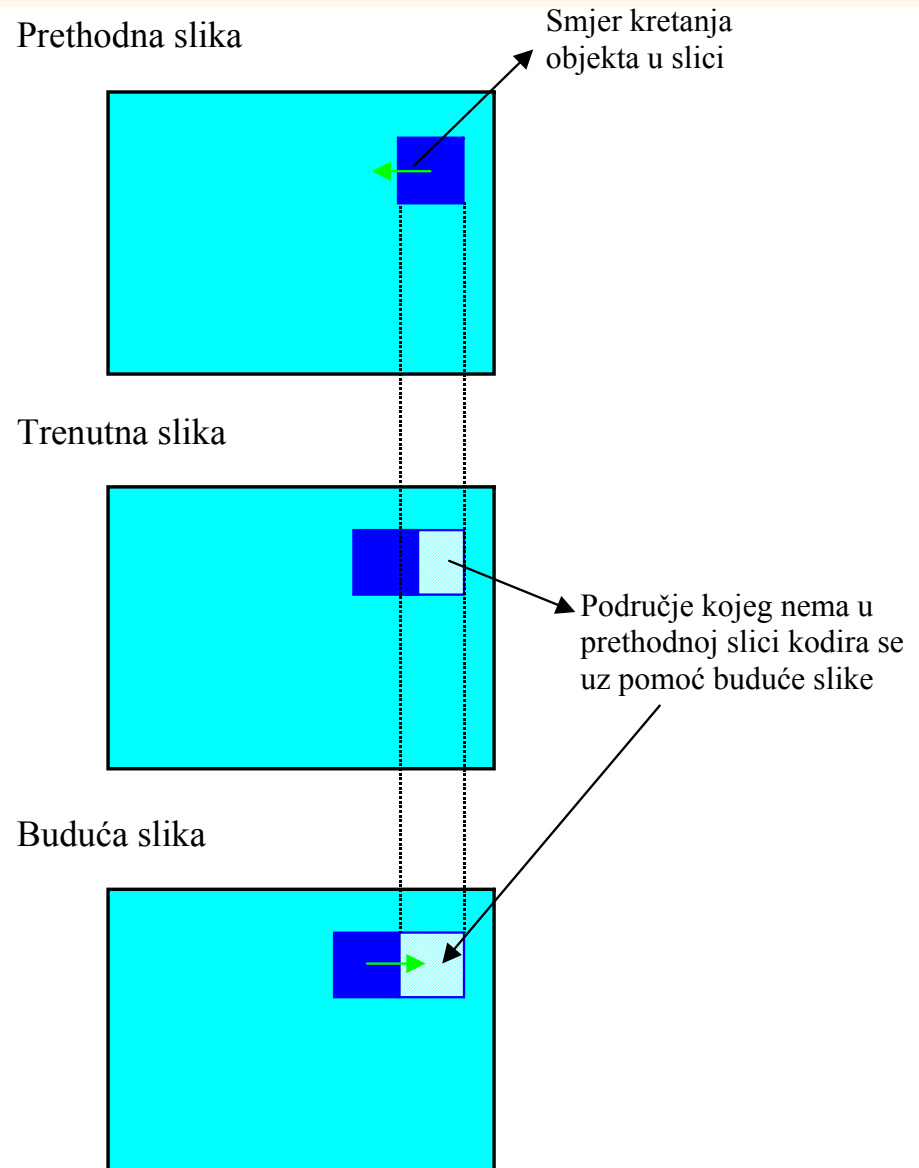
Model MPEG-2 kodiranja

- za dobivanje P-slika rabi se jednosmjerno predviđanje pokreta
 - predviđanje pokreta se provodi samo iz prethodnih slika
 - područja kojih nema u prethodnoj slici pojavljuju se u slici razlike i moraju se kodirati



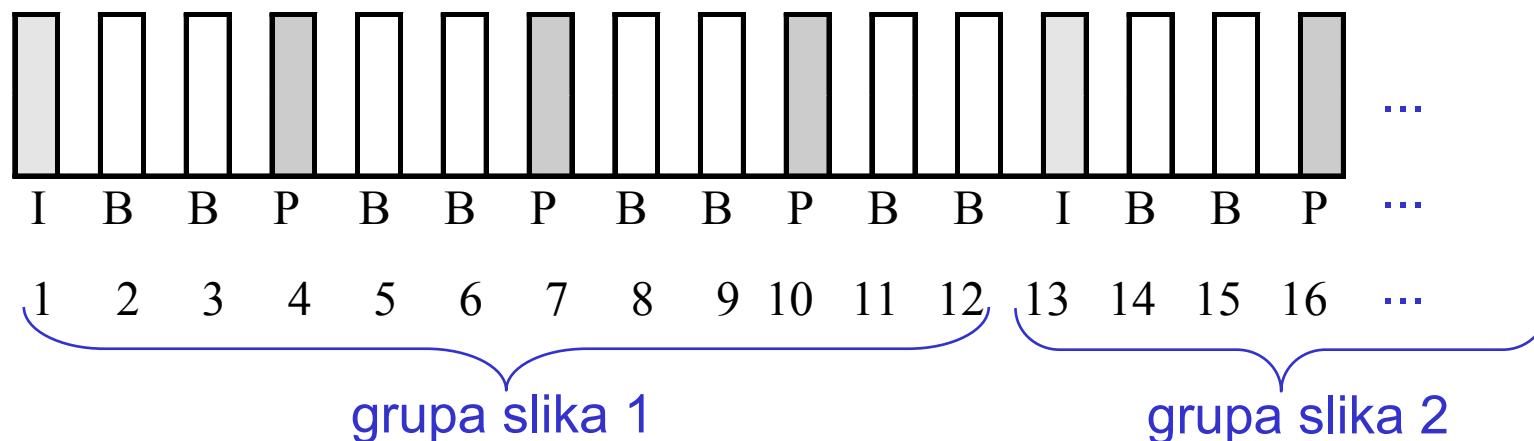
Model MPEG-2 kodiranja

- za dobivanje B-slika rabi se dvosmjerno predviđanje pokreta
 - za B-sliku predviđanje pokreta se provodi i iz prethodnih i iz budućih slika I- ili P-slika
 - buduća slika se rabi kodiranja područja trenutne slike kojih nema u prethodnoj slici
 - postižu se veći stupnjevi kompresije u odnosu na jednosmjerno predviđanje pokreta
- dvosmjerno predviđanje pokreta zahtijeva promjenu rasporeda slika u grupi slika
 - do dekodera mora prvo stići buduća slika, a iza nje trenutna slika jer je buduća slika potrebna u procesu dekodiranja trenutne slike, koja je kodirana kao B-slika



Model MPEG-2 kodiranja

- tipična grupa slika u MPEG-2 normi sastoji se od 12 ili 15 slika
- stvarni raspored slika u grupi slika koja se sastoji od 12 slika:



- dekodiranje B-slika traži prisustvo i prošlih i budućih I- i P-slika
- prijenosni raspored slika u grupi slika prilagođen je postupku dekodiranja i razlikuje se od stvarnog rasporeda:

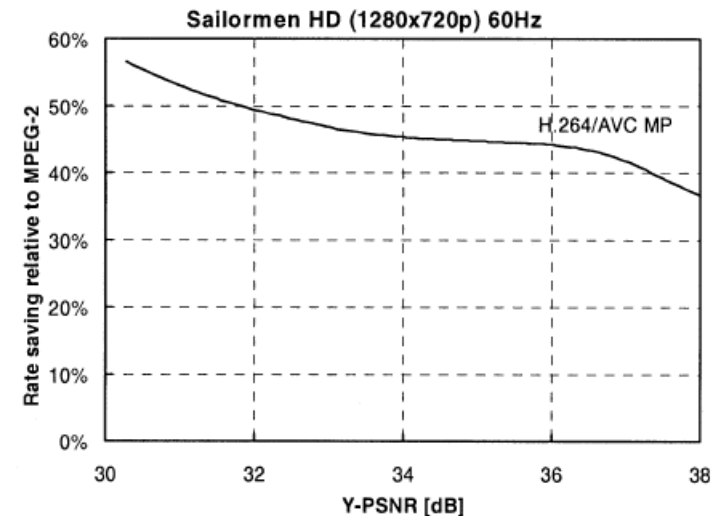
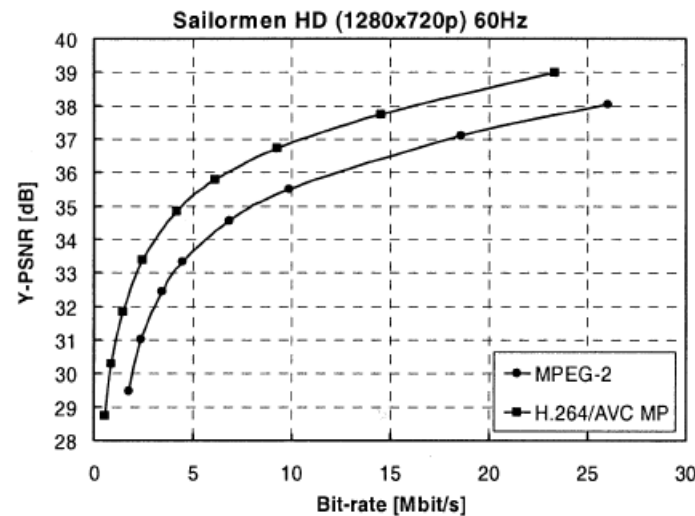
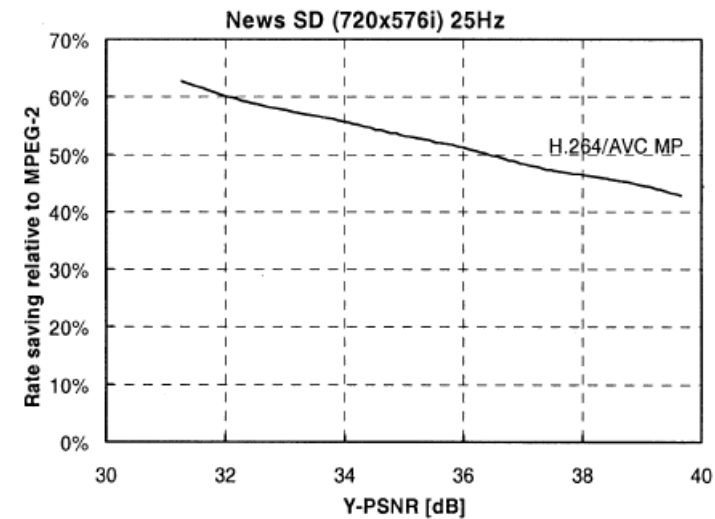
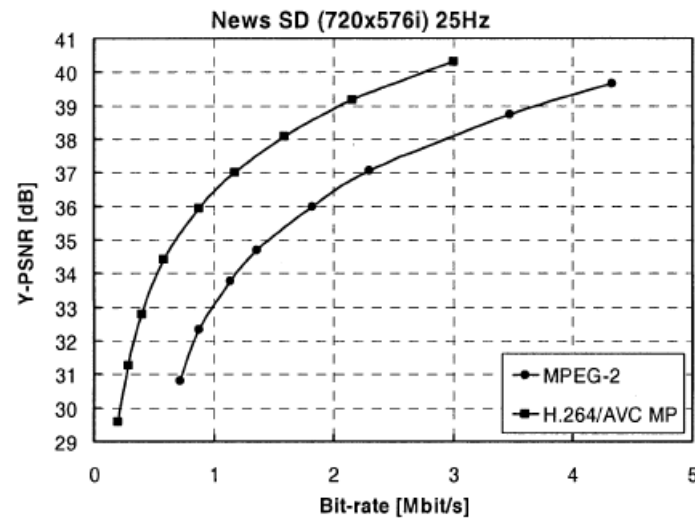
1	4	2	3	7	5	6	10	8	9	13	11	12	16	14	15
I	P	B	B	P	B	B	P	B	B	I	B	B	P	B	B

Model MPEG-2 kodiranja

- uporabom kodiranja s predviđanjem u kome se rabe P-slike i B-slike, te relativno dugih grupa slika (12 ili 15 slika) povećava se djelotvornost koda u uklanjanju redundancije te stupanj kompresije
 - P-slike sadrže približno $\frac{1}{2}$ količine podataka I-slika
 - B-slike sadrže približno $\frac{1}{4}$ količine podataka I-slika
- GOP sastavljena samo od I-slika zahtjeva veću brzinu prijenosa od GOP sastavljene od I, P i B-slika za istu kvalitetu slike
- smanjenje brzine prijenosa uz održavanje iste kvalitete moguće je postići promjenom strukture GOP, tj. uključivanjem većeg broja P i B-slika
- B-slike imaju veću pogrešku rekonstrukcije nego I-slike i P-slike
 - B-slike se ne pohranjuju u memoriji slike koda jer se ne koriste kao referentne slike u postupku predviđanja koji slijedi, tako da ne postoji opasnost od propagacije pogrešaka koje se pojavljuju u B-slikama

Usporedba normi

- usporedba MPEG-2 i H.264/AVC



Usporedba normi

- temeljne razlike H.264/AVC postupka kodiranja u odnosu na MPEG-2 normu
 - umjesto DCT koristi se cjelobrojna transformacija
 - transformacija je primjenjuje na blokove veličine 4x4 ili 2x2
 - za I-slike se rabi predviđanje unutar slike (*Intra Mode*) pri čemu se već kodirani blokovi slike rabe za predviđanje sljedećih blokova iste slike
 - pri kodiranju P-slika i B-slika može se rabiti više od jedne prethodne i više od jedne buduće slike koje se prethodno pohranjuju u memoriju
 - u postupku nadomještanja pokreta rabe se različite klase makroblokova koji se dijele na blokove različitih veličina (klase makroblokova: Inter16x16, Inter16x8, Inter8x16 i Inter8x8) radi postizanja što bolje djelotvornosti kodiranja
 - povećana točnost predviđanja vektora pokreta
 - kako bi se smanjila vidljivost rubova blokova u slici u koderu i dekoderu se u petlju za predviđanje dodaje adaptivni filter za poništavanje vidljivosti rubova blokova u slici (*deblocking filter*)
 - povećana je djelotvornost entropijskog kodiranja uporabom modeliranja konteksta

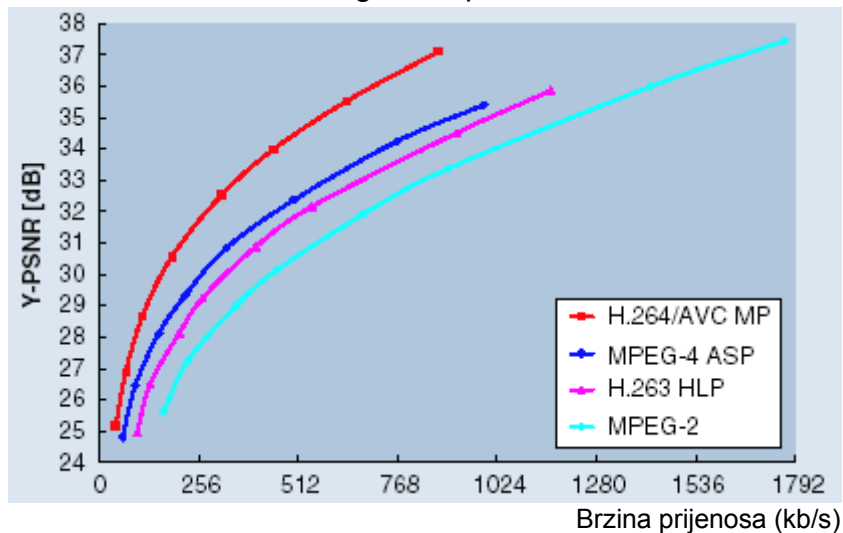
Usporedba normi

- prednost H.264/AVC u odnosu na MPEG-2
 - H.264/AVC omogućava povećanu djelotvornost kodiranja (*coding efficiency*) u odnosu na MPEG-2 i ostale norme za kodiranje videosignala
 - djelotvornost kodiranja se odnosi na redukciju brzine prijenosa uz održavanje visoke subjektivne kvalitete slike
 - za postizanje određene kvalitete slike (definirane omjerom PSNR) moguće je rabiti 2-2,5 puta nižu brzinu prijenosa u odnosu na MPEG-2
- nedostatak H.264/AVC u odnosu na MPEG-2
 - povećana složenost koda (8x složeniji od MPEG-2 koda) i dekodera
 - povećano je kašnjenje uneseno postupkom kodiranja i dekodiranja

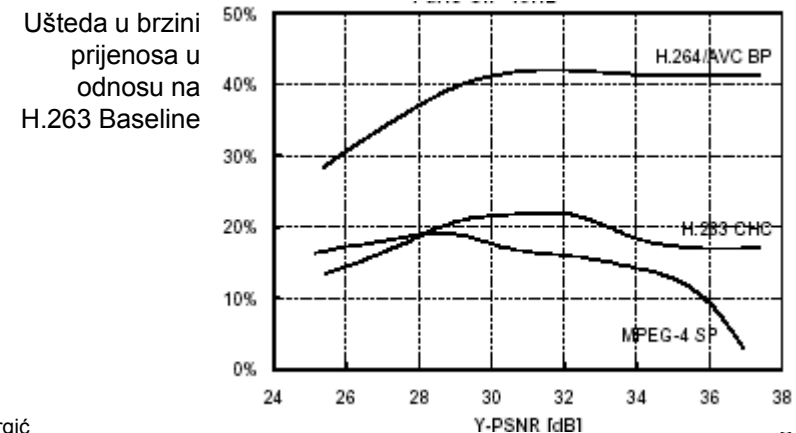
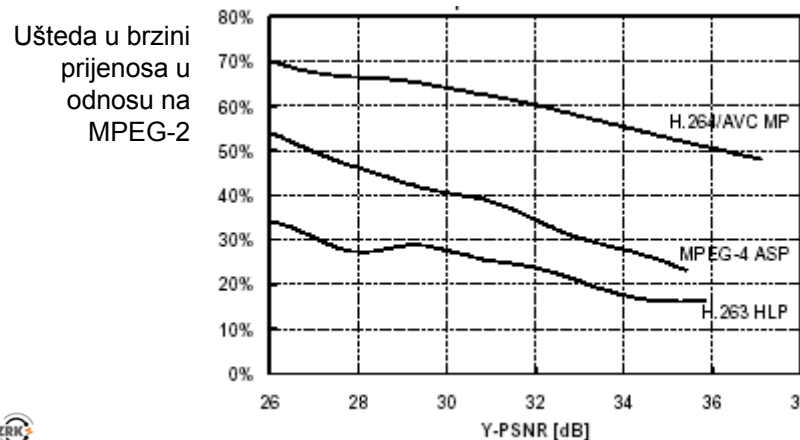
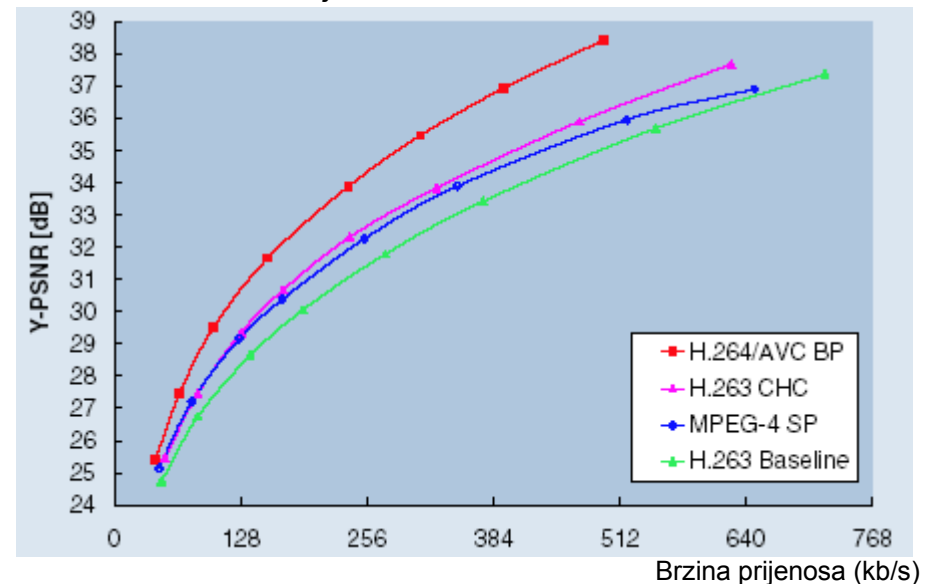
Usporedba normi

- usporedba H.264/AVC s ostalim relevantnim normama

Video streaming - Tempete, CIF, 15 Hz



Videokonferencijska sekvenca - Paris, CIF, 15 Hz



Usporedba normi

- usporedba H.264/AVC i Motion JPEG2000

