

SVEUČILIŠTE U ZAGREBU  
**FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA**

DIPLOMSKI RAD br. **2205**

**ODABIR PARAMETARA KOMPRESIJE VIDEOSIGNALA ZA  
TELEVIZIJU VISOKE KVALITETE**

**Dario Bojanjac**

Zagreb, svibanj 2009.



## **Sažetak**

U ovom diplomskom radu dan je pregled formata televizije visoke kvalitete (HDTV) propisanih trenutnim standardima. Objasnjen je teorijski postupak kompresije videosignalata, te na temelju toga i H.264 norma za kompresiju. Provedena su ispitivanja kvalitete slike temeljena na objektivnim i subjektivnim mjerama kvalitete. U ispitivanjima su korištene dvije videosekvence s potpuno različitim zahtjevima na postupak videokompresije.

# Sadržaj

1.	Uvod .....	1
2.	Formati videa .....	3
2.1.	Uvod .....	3
2.2.	Prostorno uzorkovanje .....	4
2.3.	Vremensko uzorkovanje .....	6
2.4.	Prikaz boje .....	7
2.4.1.	Uzorkovanje boje .....	8
3.	Videokompresija .....	10
3.1.	Uvod .....	10
3.2.	Princip videokompresije .....	11
3.3.	Postupak videokompresije .....	14
3.3.1.	Građa kodera .....	14
3.3.2.	Građa dekodera .....	16
3.3.3.	Dio za eliminaciju vremenske redundancije .....	16
3.4.	Nadomještanje pokreta .....	17
3.4.1.	Nadomještanje pokreta makroblokova .....	19
3.5.	Model slike .....	25
3.6.	Predikcijsko kodiranje .....	27
3.7.	Transformacijsko kodiranje .....	28
3.7.1.	DCT .....	29
3.8.	Entropijsko kodiranje .....	33
3.8.1.	Huffmanov kod .....	33
4.	H.264 standard za videokompresiju .....	34
4.1.	Uvod .....	34

4.2.	H.264 koder .....	34
4.3.	H.264 dekoder .....	36
4.4.	Kodiranje unutar slike .....	37
4.5.	Kodiranje predviđanjem iz prethodnih slika .....	37
4.6.	Dvosmjerno kodiranje .....	39
4.7.	Vektori pomaka .....	40
4.8.	Predviđanje vektora pomaka .....	43
4.9.	Transformacija i kvantizacija .....	44
4.10.	Entropijsko kodiranje .....	44
4.11.	Adaptivni rekonstrukcijski filter .....	45
4.12.	Razine u H.264 normi .....	45
5.	HDTV-televizija visoke kvalitete .....	48
5.1.	Uvod .....	48
5.2.	Formati slike .....	49
5.2.1.	720p format slike televizije visoke kvalitete .....	49
5.2.2.	1080p i 1080 i format slike televizije visoke kvalitete .....	50
5.3.	Omjer stranica slike u sustavu televizije visoke kvalitete .....	52
5.4.	Standardizacija kolorimetrije .....	53
6.	Mjere kvalitete slike .....	54
6.1.	Uvod .....	54
6.2.	PSNR .....	54
6.3.	VQM .....	57
6.4.	SSIM .....	58
7.	Mjerenja .....	61
7.1.	Uvod .....	61
7.2.	Videosekvence .....	61

7.3.	Postupak mjerenja kvalitete slike i rezultati .....	63
7.3.1.	Ispitna sekvenca Intotree 720p/50 .....	65
7.3.2.	Ispitna sekvenca Intotree 1080i/25.....	67
7.3.3.	Ispitna sekvenca Intotree 1080p/50 .....	69
7.3.4.	Ispitna sekvenca Crowdrun, 720p/50 .....	71
7.3.5.	Ispitna sekvenca Crowdrun 1080i/25 .....	73
7.3.6.	Ispitna sekvenca Crowdrun 1080p/50 .....	75
7.3.7.	Subjektivne ocjene kvalitete .....	77
8.	Zaključak .....	80
9.	Popis literature .....	82

## **Popis oznaka i kratica**

ATSC	Advanced Television Systems Committee
CABAC	Context-based Adaptive Binary Arithmetic Coding
CAVLC	Context-based Adaptive Variable Length Coding
CCD	Charge Coupled Device
DCT	Discrete Cosine Transform
DPCM	Differential Pulse-Code Modulation
DVB	Digital Video Broadcasting
DVD	Digital Versatile Disc
HDTV	High-Definition Television
ISDB	Integrated Services Digital Broadcasting
ISO	International Organization for Standardization
ITU	International Telecommunication Union
ITU-R	International Telecommunication Union-Radiocommunication
JVT	Joint Video Team
MPEG	Moving Picture Experts Group
PSNR	Peak Signal-to-Noise Ratio
SDTV	Standard-Definition Television
SMPTE	Society of Motion Picture and Television Engineers
SSIM	Structural Similarity
VQM	Video Quality Metrics

## **Popis tablica**

Tablica 3.1 Tablica Huffmanovog koda .....	33
Tablica 4.1 Razine u H.264 normi .....	47
Tablica 5.1 Pregled normi za analiziranje slike formata 720p .....	50
Tablica 5.2 Pregled normi za analiziranje slike formata 1080p .....	51
Tablica 5.3 Parametri kolorimetrije za HDTV sustav .....	53

## **Popis slika**

Slika 2.1 Prostorno i vremensko uzorkovanje videosignalna .....	3
Slika 2.2 Prostorno uzorkovanje videosignalna s nižom i višom frekvencijom .....	4
Slika 2.3 Slika 2.2 uzorkovana nižom frekvencijom.....	5
Slika 2.4 Slika 2.2 uzorkovana višom frekvencijom.....	5
Slika 2.5 Analiziranje slike s proredom.....	6
Slika 2.6 Podjela slike u dvije poluslike .....	6
Slika 2.7 Uzorkovanje krominantnih komponenti.....	9
Slika 3.1 Signal se sastoji od redundancije i entropije.....	11
Slika 3.2 Utjecaj stupnja kompresije.....	11
Slika 3.3 Jedna slika videosignalna s velikim homogenim područjima.....	12
Slika 3.4 Obrada pozadine NP filtrom .....	13
Slika 3.5 Dvije uzastopne slike u nizu .....	14
Slika 3.6 Koder.....	15
Slika 3.7 Dvije susjedne slike u videosekvenci.....	17
Slika 3.8 Rezidualna slika .....	17
Slika 3.9 Vektori pomaka svakog pojedinog elementa slike.....	18
Slika 3.10 Makroblok za format slike 4:2:0 .....	19
Slika 3.11 Dvije uzastopne sličice videosekvence.....	20
Slika 3.12 Rezidualna slika nastala oduzimanjem slike sa slike 3.11.....	20
Slika 3.13 Rezidualna slika, makroblokovi 16x16.....	21
Slika 3.14 Rezidualna slika, makroblokovi 8x8 .....	21
Slika 3.15 Rezidualna slika, makroblokovi 4x4 .....	22
Slika 3.16 Cjelobrojni, polovina i četvrtina razmaka između elemenata slike ...	23
Slika 3.17 Rezidualna slika, makroblokovi 4x4, polovina cjelobrojnog pomaka	23
Slika 3.18 Rezidualna slika, makroblokovi 4x4, četvrtina cjelobrojnog pomaka	24
Slika 3.19 Autokorelacijska funkcija prirodne slike.....	25
Slika 3.20 Autokorelacijska funkcija rezidualne slike.....	26
Slika 3.21 Prostorna predikcija (DPCM).....	27
Slika 3.22 Bazne funkcije za 8x8 DCT transformaciju.....	30
Slika 3.23 DCT koeficijenti, makroblokovi dimenzija 8x8 .....	31
Slika 3.24 Rezidualna slika .....	32
Slika 3.25 Cik-cak način analiziranja.....	32

Slika 4.1 H.264 koder .....	35
Slika 4.2 H.264 dekoder .....	37
Slika 4.3 Podjela makrobloka dimenzija 16x16 .....	38
Slika 4.4 Podjela makrobloka dimenzija 8x8 .....	38
Slika 4.5 Postupak predviđanja iz prethodnih i budućih slika .....	39
Slika 4.6 4x4 blok trenutne slike.....	40
Slika 4.7 Postupak interpolacije luminantnih komponenti.....	41
Slika 4.8 Interpolacija uzoraka za pomak od 1/4 .....	42
Slika 4.9 Određivanje krominantnih uzoraka na 1/8 udaljenosti .....	43
Slika 4.10 Profilu u H.264 normi.....	46
Slika 5.1 Usporedba veličine slike kod SDTV i HDTV standarda .....	49
Slika 5.2 Segmentacija slike .....	51
Slika 5.3 Usporedba omjera stranica u SDTV i HDTV normi .....	52
Slika 6.1 Usporedba elemenata rekonstruirane i izvorne slike .....	55
Slika 6.2 Blok shema VQM metode.....	57
Slika 6.3 Usporedba slika Brod s različitim distorzijama uz jednaki MSE.....	59
Slika 7.1 Prva slika videosekvence Crowdrun.....	62
Slika 7.2 Prva slika videosekvence Intotree .....	62
Slika 7.3 Nagib funkcije ovisnosti kvalitete slik o brzini prijenosa.....	64
Slika 7.4 Ovisnost PSNR koeficijenta o brzini prijenosa.....	65
Slika 7.5 Ovisnost VQM koeficijenta o brzini prijenosa .....	65
Slika 7.6 Ovisnost SSIM koeficijenta o brzini prijenosa .....	66
Slika 7.7 Ovisnost PSNR koeficijenta o brzini prijenosa.....	67
Slika 7.8 Ovisnost VQM koeficijenta o brzini prijenosa .....	67
Slika 7.9 Ovisnost SSIM koeficijenta o brzini prijenosa .....	68
Slika 7.10 Ovisnost PSNR koeficijenta o brzini prijenosa.....	69
Slika 7.11 Ovisnost VQM koeficijenta o brzini prijenosa .....	69
Slika 7.12 Ovisnost SSIM koeficijenta o brzini prijenosa .....	70
Slika 7.13 Ovisnost PSNR koeficijenta o brzini prijenosa.....	71
Slika 7.14 Ovisnost VQM koeficijenta o brzini prijenosa .....	71
Slika 7.15 Ovisnost SSIM koeficijenta o brzini prijenosa .....	72
Slika 7.16 Ovisnost PSNR koeficijenta o brzini prijenosa.....	73
Slika 7.17 Ovisnost VQM koeficijenta o brzini prijenosa .....	73

Slika 7.18 Ovisnost SSIM koeficijenta o brzini prijenosa .....	74
Slika 7.19 Ovisnost PSNR koeficijenta o brzini prijenosa.....	75
Slika 7.20 Ovisnost VQM koeficijenta o brzini prijenosa .....	75
Slika 7.21 Ovisnost SSIM koeficijenta o brzini prijenosa.....	76
Slika 7.22 Rezultati subjektivnog ispitivanja sekvence Crowdrun, 720p/50 .....	77
Slika 7.23 Rezultati subjektivnog ispitivanja sekvence Crowdrun, 1080i/25.....	78
Slika 7.24 Rezultati subjektivnog ispitivanja sekvence Crowdrun, 1080p/50 ...	78

## 1. Uvod

Zamislite sliku od prije stotinjak godina. Imate 40-ak godina i kćer spremnu za udaju. Organizirali ste veliku svadbu, pozvali najbližu rodbinu, veselite se, zabavljate i želite to podijeliti s ostalim ljudima koji nisu na vjenčanju. Na žalost, video kamera još nije izumljena i vi to ne možete snimiti, a sve i da nekako možete snimiti, ne postoji način na koji to možete nekome poslati.

U ovom banalnom primjeru vidimo koliko je zapravo proces snimanja i prikazivanja videa sadržan u svakom trenutku našeg života. Danas se snimaju gotovo svi trenuci života. Od rođenja djeteta, do krštenja, odlaska u školu, prvih utakmica, plesnih nastupa, glume, krizme, završetka škole, fakulteta, vjenčanja pa do kojekakvih izleta, roštilja, zabava... Sve to danas želimo snimiti, negdje pohraniti i nekome pokazati.

Svijet je postao globalno selo. Ljude ne veseli samo snimanje i pokazivanje tog materijala gostima i najbližoj rodbini. Svi mi imamo neke daleke rođake koje rijetko viđamo, ali ipak s njima želimo podijeliti naše sretne trenutke.

Izumom interneta razmjena podataka je postala kudikamo lakša i brža. Sve to što smo snimili ili fotografirali možemo u vrlo kratkom vremenu poslati na drugi kraj svijeta.

Nije ovdje samo riječ o osobnim događajima koje želimo snimiti. Danas postoji mnoštvo internet stranica na kojima ljudi mogu pokazati svoje filmove ili glazbu te tako steći poklonike na svim strana svijeta. Da svijet nije ovakvo globalno selo, jednu Susan Boyle, kućanicu iz zabačenog sela u Škotskoj i novootkrivenu glazbenu zvijezdu, ne bi vidjelo 36 000 000 ljudi u samo 2 dana. Od nepoznate osobe do svjetske zvijezde zahvaljujući videu i internetu.

Sve bi to bilo nemoguće bez digitalnih video kamera i postupaka kompresije. No, nije samo stvar u tehnologiji. Tehnologija postoji godinama. Poanta je što je ona prošla veliki put od sofisticirane tehnologije kojom su znali upravljati samo odabrani, do vrlo jednostavne tehnologije koju zna koristiti gotovo svatko.

Danas svako malo dijete zna upaliti kameru i nešto snimiti. Malo veće znanje potrebno je za slaganje tih sekvenci, ali i to postaje sve jednostavnije i jednostavnije. Sljedeći korak je neka jednostavna kompresija podataka i za

nekoliko minuta događaj u trajanju par sekundi, iz neke male ulice u predgrađu Zagreba, dolazi do najudaljenijih kutaka Zemlje.

Sve nam je to omogućilo pojednostavljenje tehnologije za snimanje i kompresiju videa. S pojavom prvih digitalnih kamera, pojavila se potreba za komprimiranjem snimljenog materijala. U prvim trenucima mediji za pohranu su bili mali a i brzine prijenosa mrežama su također bile niske. Zato su ljudi razvili postupke smanjivanja veličine videosekvenci kako bi ih mogli distribuirati internetom.

Svakim danom, tehnologija ide naprijed, rastu kapaciteti medija za pohranu, rastu brzine prijenosa, ali potreba za kompresijom i dalje postoji, jer želja za što bržim prijenosom, ili što kvalitetnijom slikom u sustavima s ograničenim resursima, i dalje postoji.

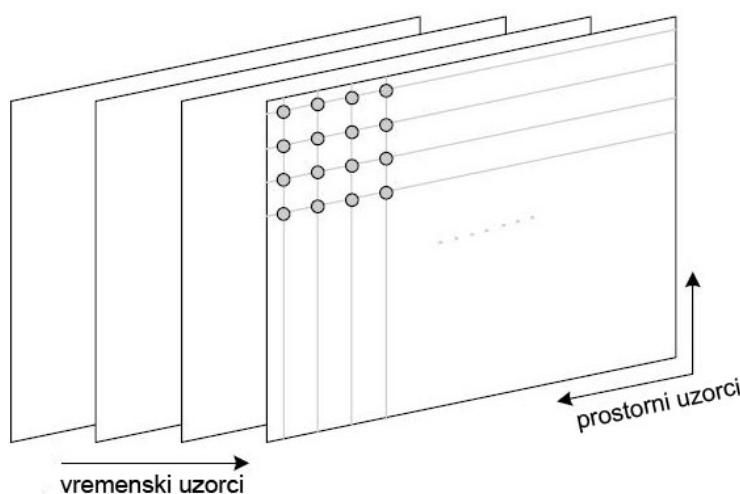
## 2. Formati videa

### 2.1. Uvod

Komprimiranje videa je proces kodiranja i dekodiranja video signala. Video signal nastaje u kameri vremenskim i prostornim uzorkovanjem određenog prizora. Na izlazu iz kamere dobijemo određeni niz sličica koje, prilikom prikazivanja određenom brzinom, stvaraju iluziju neprekidnog pokreta. Samim snimanjem scene učinili smo određenu prostornu kompresiju jer smo iz prirode izvadili jedan njezin segment, tj. snimali smo određeni, ograničeni dio prostora. Vremensku kompresiju određuje određeno vrijeme koje smo snimali. Općenito, videosekvenca na izlazu iz kamere uzorkovana je i prostorno i vremenski jer je to niz slika snimanih određenom brzinom (u intervalima od npr. 1/25 sekunde).

Svaka videosekvenca sastoji se od nekih vremenskih i nekih prostornih karakteristika. U prostorne karakteristike spadaju teksture, objekti, ljudi, itd. a u vremenske pokret, pomak kamere...

Prirodna scena je vremenski i prostorno kontinuirana. Da bismo prikazali prirodnu scenu u digitalnom obliku moramo ju vremenski i prostorno uzorkovati. Digitalni video je upravo prikaz uzorkovane video sekvence u digitalnom obliku. Možemo ga zamisliti kao niz sličica vremenski uzorkovanih, od kojih se svaka sličica prostorno uzorkuje i reprezentira matricom elemenata slike. Elementi matrice odgovaraju luminanciji ili boji na mjestu elementa slike.

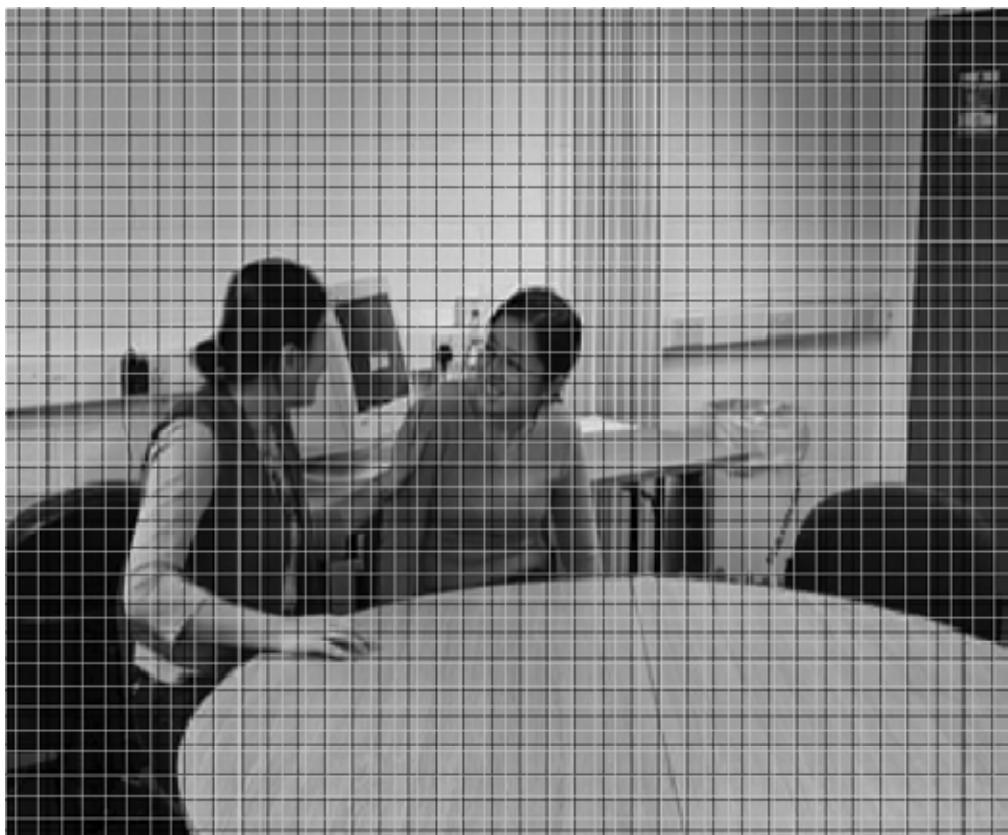


**Slika 2.1** Prostorno i vremensko uzorkovanje videosignal-a

Sliku u boji prikazujemo pomoću tri matrice od kojih je svaka matrica pojedine primarne boje.

## 2.2. Prostorno uzorkovanje

Na izlazu iz CCD (Charge-Couple Device) senzora dobijemo analogni video signal koji predstavlja sliku. Uzorkovanjem signala određenom frekvencijom, dobijemo uzorkovani signal koji ima vrijednosti samo u trenucima uzorkovanja. O frekvenciji uzorkovanja ovisi kvaliteta digitalne slike. Što je veća frekvencija uzorkovanja, to je moguće prikazati sitnije detalje u slici kao što vidimo na slici 2.2.



**Slika 2.2** Prostorno uzorkovanje videosignal-a s nižom (crna) i višom (siva) frekvencijom



**Slika 2.3** Slika 2.2 uzorkovana nižom frekvencijom (crna)

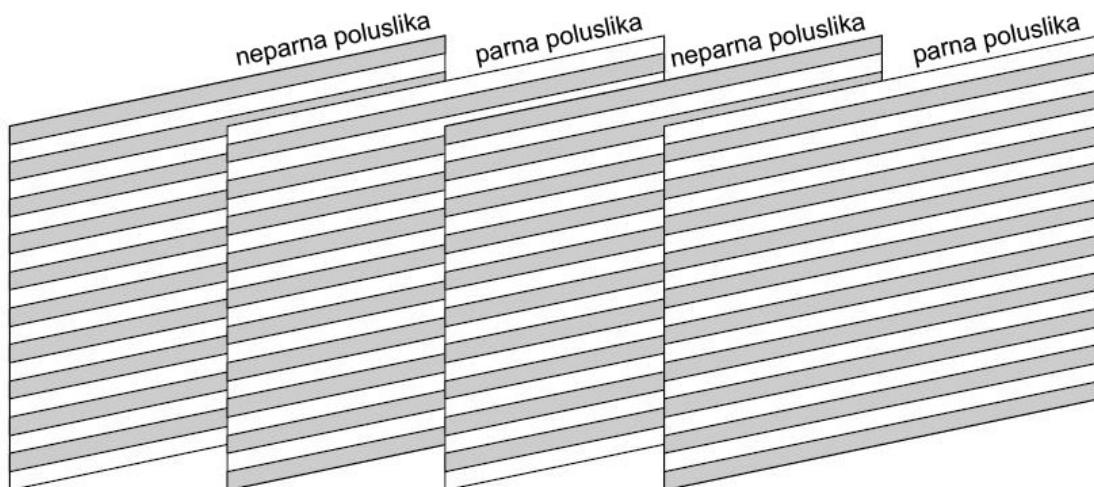


**Slika 2.4** Slika 2.2 uzorkovana višom frekvencijom (siva)

## 2.3. Vremensko uzorkovanje

Video se sastoji od neprekinutog pokreta. No, takav je zapis nemoguće napraviti, pa se radi vremensko uzorkovanje, tj. sličice se snimaju u određenim vremenskim intervalima. Veća frekvencija snimanja sličica rezultira fluidnijim pokretima. Uzorkovanje od 25 do 30 sličica u sekundi je standardno za televiziju i nakon analiziranja s proredom (koje udvostručuje frekvenciju), dobijemo čisti pokret.

Vremenski, scena može biti uzorkovana (analizirana) s proredom ili progresivno. Kod progresivnog analiziranja slika se analizira kao cjelina, dok se kod analiziranja s proredom slika dijeli u dvije poluslike i svaka poluslika se analizira zasebno. Poluslike se obično nazivaju parna i neparna, po rednom broju prve linije koju sadrže.



**Slika 2.5** Analiziranje slike s proredom



**Slika 2.6** Podjela slike u dvije poluslike

## 2.4. Prikaz boje

Boje u slici se mogu prikazivati u nekoliko različitih prostora boja od kojih je za televiziju najznačajniji YUV prostor boja.

Standardno, slika u boji se može prikazati pomoću tri matrice od kojih jedna sadrži informaciju za crvenu, druga za plavu, a treća za zelenu boju (RGB sustav). Te tri boje zovu se primari. Bilo koja druga boja može se prikazati kombinacijom ove tri boje.

Taj prikaz može se pojednostaviti u jedan manje zahtjevan prikaz za prijenos i pohranu a to je  $YC_bC_r$  (YUV) prikaz.

Ljudski vizualni sustav je osjetljiviji na promjene osvjetljenja nego na promjene boje. U RGB sustavu sve tri informacije o boji su jednak vijedne i ne možemo postići neku kompresiju podataka. Ako iz informacije o boji izvadimo luminantnu komponentu, dobit ćemo signal koji možemo kodirati s manje bitova od luminantnog signala.

Luminantni signal dobijemo kombinacijom primara na način:

$$E_Y = k_r E_R + k_g E_G + k_b E_B$$

Gdje su na  $k_r$ ,  $k_g$  i  $k_b$  težinski faktori i vrijedi:

$$k_r + k_g + k_b = 1$$

Informaciju o boji tada prenosimo kao razliku boje i luminantne komponente. Te komponentne nazivamo krominantne komponente

$$C_b = E_B - E_Y$$

$$C_g = E_G - E_Y$$

$$C_r = E_R - E_Y$$

Ovim postupkom došli smo do 4 signala, ali u njima imamo skrivenu redundanciju. Kako je luminantna komponenta sastavljena od težinskog zbroja krominantnih komponenti, dovoljno je prenosići luminantnu komponentu i dvije krominantne komponente. Tada na prijemu iz ta tri signala možemo u potpunosti rekonstruirati informacije o boji.

U luminantnoj komponenti najvećim je dijelom sadržana komponenta zelene boje. Krominantna komponenta koja odgovara razlici zelene boje i luminantne komponente sadrži malu energiju te je podložnija šumu od ostalih komponenti. Zato se u sustavu prenose komponente razlike crvene i plave boje, te luminantna komponenta. Na mjestu prijema tada iz te tri komponente možemo u potpunosti rekonstruirati informaciju o boji na način:

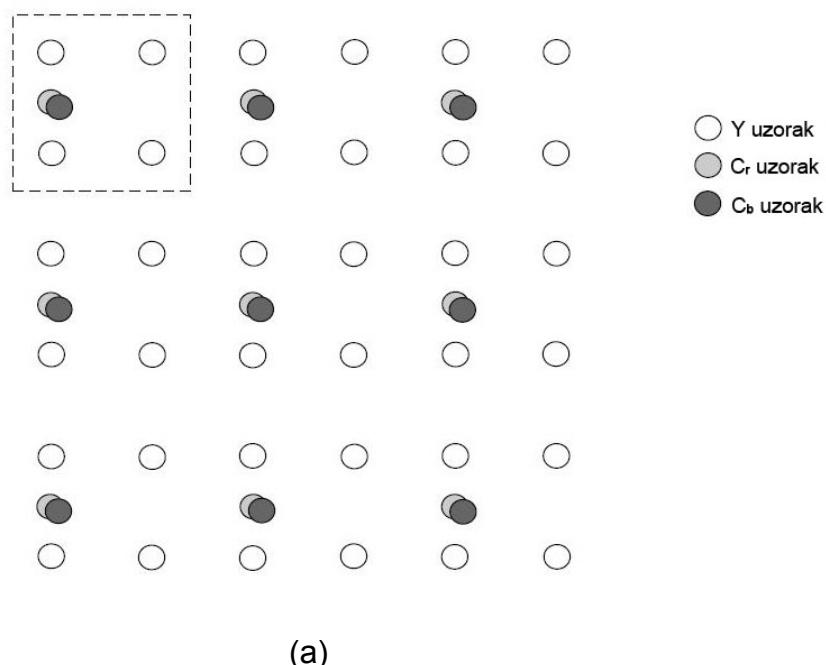
$$E_R = E_Y + 2(1 - k_r)C_r$$

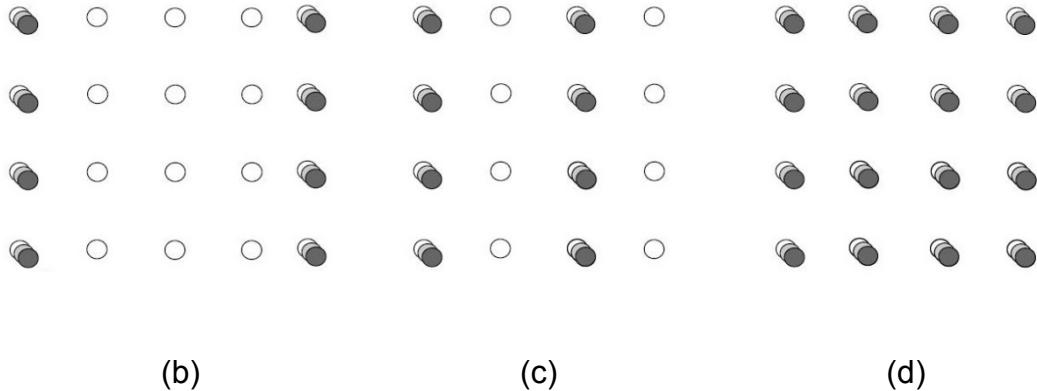
$$E_B = E_Y + 2(1 - k_b)C_b$$

$$E_G = E_Y - \frac{2k_b(1 - k_b)}{1 - k_b - k_r} C_b - \frac{2k_r(1 - k_r)}{1 - k_b - k_r} C_r$$

#### 2.4.1.Uzorkovanje boje

Ranije je rečeno da je ljudski vizualni sustav zbog svoje građe osjetljiviji na luminantnu komponentu nego na informaciju o boji. Zato je moguće uzorkovati krominantne komponente s manjom frekvencijom od luminantne bez vidljivog gubitka informacije. Tako razlikujemo nekoliko različitih načina uzorkovanja od kojih je daleko najpopularniji 4:2:0 način koji ima veliku primjenu u digitalnoj televiziji, videokonferencijama, pohrani filmova na DVD-u... Na slici 2.7 prikazani su različiti formati uzorkovanja krominantnih komponenti.





**Slika 2.7** Uzorkovanje krominantnih komponenti (a) 4:2:0 uzorkovanje (b) 4:1:1 uzorkovanje (c) 4:2:2 uzorkovanje (d) 4:4:4 uzorkovanje

Kod 4:1:1 formata, krominantne komponente uzorkovane su s četiri puta nižom frekvencijom, tj. na četiri luminantne komponente dolazi po jedna krominantna. Kod 4:2:2 formata krominantne komponente uzorkovane su dva puta nižom frekvencijom, a kod 4:4:4 formata uzorkovane su s jednakom frekvencijom kao i luminantna komponenta. U 4:2:0 formatu krominantne komponente su poduzorkovane dva puta u vertikalnom i u horizontalnom smjeru, pa na šesnaest luminantnih uzoraka dobijemo četiri krominantna uzorka.

### 3. Videokompresija

#### 3.1. Uvod

U zadnjih nekoliko godina brzina spajanja na internet je dramatično porasla. Do prije 5,6 godina na internet smo se spajali brzinama od 56kbit/s i imali problema pri učitavanju najjednostavnijih stranica, a o filmovima s npr. youtubea nismo niti pomicljali. Danas su te brzine enormno visoke i iznose nekoliko Mbit/s. Više nema problema s učitavanjem stranica, niti s gledanjem filmova na nekom od popularnih siteova za eksploraciju videa. Veća brzina spajanja rezultirala je i porastom medija za pohranu podataka, jer njezinim porastom raste i potreba za spremanjem veće količine materijala. Tako smo u roku par godina došli do memorija veličine stotina GB.

Uz sav taj porast brzina spajanja i medija za pohranu, i dalje postoji potreba za kompresijom videa, iako to na prvi pogled nije očigledno. Kompresija videa nam nudi dvije velike pogodnosti.

Kao prvo, kompresija nam omogućuje pohranu i upotrebu video materijala u uvjetima u kojima ne bismo mogli koristiti nekomprimirani materijal.

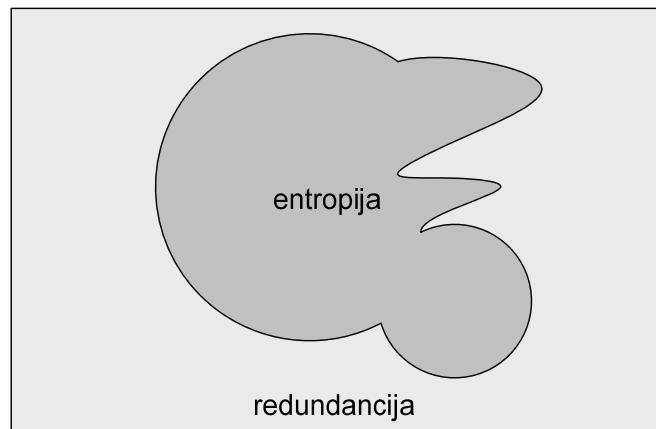
Kao primjer uzimamo DVD (Digital Versatile Disc) koji u svom standardnom dvoslojnem formatu zapisa (kapacitet mu je 8,54GB) ne bi mogao pohraniti niti nekoliko sekundi kvalitetnog nekomprimiranog videa. Niti pojava Blu-Ray diska nije puno pomaknula granicu. Kapacitet Blu-Ray diska iznosi oko 50GB u dvoslojnem formatu zapisa i to bi nam omogućilo možda minutu kvalitetnog nekomprimiranog videa. O dugometražnim filmovima u tom slučaju ne bi mogli niti razmišljati (Ben Hur bi nam zauzeo cca. 200 Blu-Ray diskova).

Druga prednost kompresije je vidljiva u efikasnosti prijenosa i pohrane video signala. Ako imamo određenu propusnost prijenosnog kanala, onda nam je atraktivnije slati komprimirani video visoke rezolucije, nego nekomprimirani nisko kvalitetni video.

Upravo zato će se video kompresija i dalje razvijati i koristiti iako brzine interneta i veličine medija svakodnevno eksponencijalno rastu.

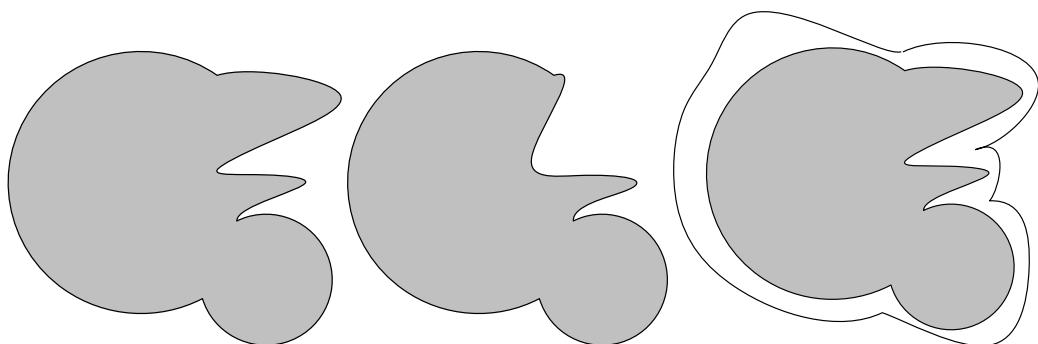
### 3.2. Princip videokompresije

Postupak kompresije je postupak u kojem smanjujemo količinu podataka u određenom signalu. Na taj način dolazimo do manje veličine signala i manje brzine prijenosa potrebne za prijenos kvalitetnog materijala.



**Slika 3.1** Signal se sastoji od redundancije i entropije

Svaki signal sastoji se od redundantnog dijela i entropije. Najjednostavnije kompresije postižu se eliminacijom redundantnog dijela signala. U postupcima kompresije bez gubitaka, eliminira se statistička redundancija, tako da u signalu ostaje samo njegov redundantni dio. U tom slučaju signal se može savršeno rekonstruirati na mjestu prijema. Na slici 3.2 vidimo utjecaj stupnja kompresije na kvalitetu slike. U prvom slučaju imamo savršeno odstranjenu redundanciju i rekonstruirana slika je identična izvornoj. U drugom slučaju smo odstranili i dio entropije, što će nam se odraziti na kvalitetu rekonstruirane slike, a u trećem slučaju šaljemo i dio redundancije.



**Slika 3.2** Utjecaj stupnja kompresije (a) šaljemo samo entropiju (b) dio entropije smo odbacili (c) šaljemo entropiju i dio redundancije

Iako ovaj postupak zvuči atraktivno, on to zapravo nije jer se na ovaj način mogu postići skromni rezultati postupka kompresije. Zato se svi efikasni postupci videokompresije baziraju na postupcima kompresije s gubicima. Na taj način postižu se veliki stupnjevi kompresije, ali rekonstruirani signal više nije jednak izvornom. Cilj takvih algoritama je postići velike stupnjeve kompresije uz minimalni distorziju signala. Algoritmi za kompresiju video signala baziraju se na uklanjanju vremenske i prostorne redundancije, uz filtriranje u frekvencijskoj domeni.

Slika 3.3 prikazuje jednu sliku videosignalata. Uokvirena područja slike sadrže veliku prostornu redundanciju jer sadrže minimalne promjene sadržaja. Određeni stupanj kompresije bi se mogao postići djelovanjem na ta područja.



**Slika 3.3** Jedna slika videosignalata s velikim homogenim područjima

Slika 3.4 prikazuje istu sliku nakon djelovanja niskopropusnog filtra. Ljudski vizualni sustav je osjetljiviji na niskofrekvencijske promjene. Koristeći tu informaciju, moguće je postići određeni stupanj kompresije filtriranjem visokofrekvencijskog sadržaja informacije. Iako je dobar dio informacije nepovratno izgubljen, slika je i dalje prepoznatljiva.



**Slika 3.4** Obrada pozadine NP filtrom

Na slici 3.5 prikazane su dvije uzastopne slike u videosignalu. Vidimo da je mala promjena između te dvije slike i da imamo popriličnu vremensku redundanciju jer su velike površine slike ostale nepromijenjene.



**Slika 3.5** Dvije uzastopne slike u nizu

U ovom primjeru vidljivo je da bi eliminacijom prostorne i vremenske redundancije, te visokofrekvenčkih komponenti postigli veći stupanj kompresije videosignal-a. U tom slučaju komprimirani signal ne bi bio jednak izvornom, tj. postupak kompresije bi bio s gubicima, ali ti gubici u odnosu na stupanj kompresije i primjenu ovog videosignal-a ne bi bili preveliki.

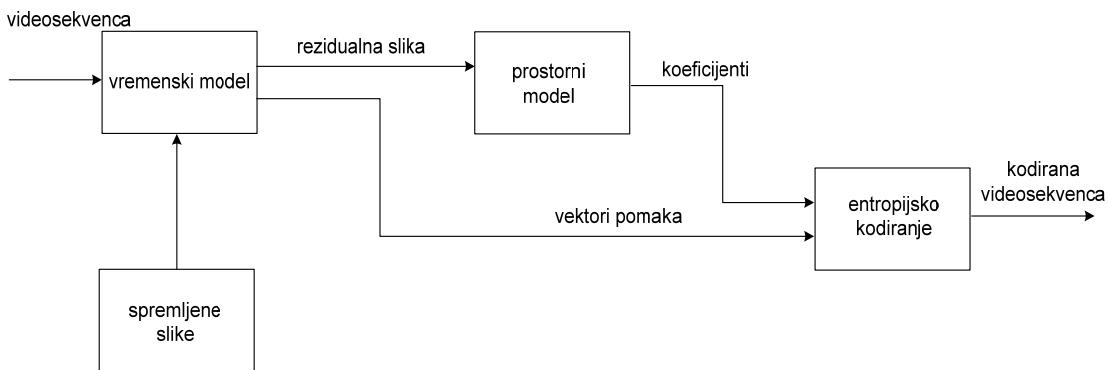
### 3.3. Postupak videokompresije

#### 3.3.1. Građa kodera

Proces kompresije odvija se u koderu. Kao što vidimo na slici 3.6 na ulaz kodera dovodimo videosekvencu, i nakon prolaska kroz nekoliko postupaka koji će biti objašnjeni u dalnjem tekstu, na izlazu dobivamo komprimirani video materijal.

Ako je komprimirana videosekvenca identična izvornoj, tada je naš koder provodi kompresiju bez gubitaka, a ako nije, onda on provodi postupak kompresije s gubicima.

Na koder stavljamo dva proturječna zahtjeva. Cilj nam je dobiti što kvalitetniji signal na izlazu uz što manju količinu podataka. Ta dva cilja obično razmatrama zasebno i onda odlučujemo što nam je bitnije, kvaliteta ili količina podataka i na taj način odlučujemo o stupnju kompresije koji ćemo koristiti. Također je moguće pronaći neki balans između ta dva proturječna zahtjeva i za svaku određenu primjenu videosignal-a odrediti potreban stupanj kompresije.

**Slika 3.6** Koder

Koder na slici 3.6 sastoji se od tri glavna dijela (dio za eliminaciju vremenske redundancije, dio za eliminaciju prostorne redundancije i dio za entropijsko kodiranje) od kojih svaki vrši jednu od tri glavne funkcije.

Prvi dio eliminira vremensku redundanciju iz ulaznog signala. On uspoređuje postojeću sliku s nizom referentnih/susjednih slika i na temelju njih predviđa sadržaj postojeće slike. Nakon toga od postojeće slike oduzima predviđenu sliku i na izlazu daje rezidualna sliku (sliku razlike). Osim nje, na izlazu daje i vektore pomaka koji nam opisuju nadomještanje pokreta u postojećoj slici.

Rezidualna slika je ulaz u dio za eliminaciju prostorne redundancije koji koristi sličnost između susjednih uzoraka u rezidualnoj slici. Obično se na rezidualnu sliku primjeni neka od transformacija, te se rezultat transformacije kvantizira. Transformacija prelazi u drugu domenu (najčešće frekvencijsku), u kojoj su uzorci rezidualne slike predstavljeni brojevima (koeficijentima transformacije). Koeficijenti se zatim kvantiziraju i eliminiraju, kako bismo dobili puno efikasniji prikaz rezidualne slike. Na taj smo način eliminirali visokofrekvencijske promjene u rezidualnoj slici, kojih zapravo i ne bi trebalo biti jer teško može doći do velike promjene između susjednih slika. Izlaz iz dijela za uklanjanje prostorne redundancije su kvantizirani koeficijenti rezidualne slike.

Parametri vremenskog modela (vektori pomaka) i prostornog modela (kvantizirani koeficijenti transformacije u frekvencijsku domenu) se zatim dovode na ulaz dijela za entropijsko kodiranje. U postupku entropijskog kodiranja eliminiramo statističku redundanciju iz niza podataka i na izlazu dobijemo

komprimirani niz podataka (koji se sastoji od kodiranih vektora pomaka, kodiranih rezidualnih koeficijenata i zaglavlja) spremam za pohranu ili prijenos.

### 3.3.2. Građa dekodera

Cilj dekodera je iz niza podataka rekonstruirati videosekvencu. Na ulaz dovodimo niz podataka i vršimo obrnuti proces od kodiranja. Kvantizirani koeficijenti i vektori pomaka su dekodirani postupkom entropijskog dekodiranja i u dijelu za prostornu rekonstrukciju rekonstruiramo rezidualnu sliku. Dekoder (njegov dio za vremensku rekonstrukciju) pomoću vektora pokreta i uz pomoć prethodno dekodirane slike rekonstruira postojeću sliku. Ovdje vidimo da je u procesu kodiranja potrebno poslati određenu količinu referentnih slika iz kojih ćemo zatim pomoći vektora pomaka i rezidualnih slika moći rekonstruirati videosekvencu.

### 3.3.3. Dio za eliminaciju vremenske redundancije

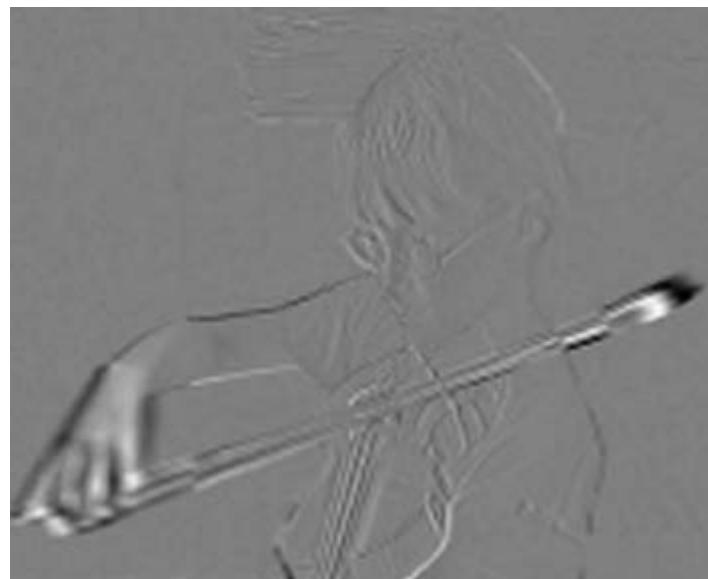
Uloga dijela za uklanjanje vremenske redundancije je smanjivanje redundancije između susjednih slika postupkom predviđanja slike iz određenog broja susjednih slika. Predviđanjem dolazimo do određene slike koja se zatim oduzima od postojeće slike i na taj način dobijemo rezidualnu sliku. Predviđanje se najčešće obavlja uz pomoć nekoliko prethodnih ili budućih slika.

Najjednostavniji način predviđanja je korištenje prethodne slike kao polazište za predviđanje trenutne slike. Na slici 3.7 vidimo dvije susjedne slike u jednoj videosekvenci. Prva slika je iskorištena kao polazište za predviđanje druge slike.

Na slici 3.8 vidimo rezidualnu sliku nastalu oduzimanjem ove dvije slike. U rezidualnoj slici siva boja predstavlja područja u kojima nema razlike između dvije slike, a bijela i crna predstavljaju mesta najveće razlike. Kao što vidimo, u rezidualnoj slici imamo određenu količinu energije signala predstavljenog bijelim i crnim dijelovima. Određenim postupkom uklanjanja te energije mogli bismo postići više stupnjeve kompresije.



**Slika 3.7** Dvije susjedne slike u videosekvenci



**Slika 3.8** Rezidualna slika nastala oduzimanjem dvije susjedne slike u videosekvenci

### 3.4. Nadomještanje pokreta

Promjenu između dvije susjedne slike u videosekvenci mogu uzrokovati promjena određenog objekta, paniranje kamere i promjena u svjetlosnim uvjetima. Prve dvije promjene uzrokuju pomak elemenata slike unutar slike i kao takve moguće ih je nadomjestiti pomakom tih istih elemenata. Na taj način dolazimo do vektora pomaka. Na slici 3.9 vidimo vektore pomaka za slike na slici 3.7



**Slika 3.9** Vektori pomaka svakog pojedinog elementa slike (poduzorkovani dva puta)

Na taj način možemo doći do efikasnijeg predviđanja pokreta, koristeći elemente slike iz prethodne slike i vektora pomaka između te dvije slike. Računanje vektora pomaka za svaki element slike zahtijevao bi previše računalnih resursa. Zato se predviđanje pokreta ne odvija nad svakim elementom slike, nego nad određenim dijelovima slike zvanim makroblokovi.

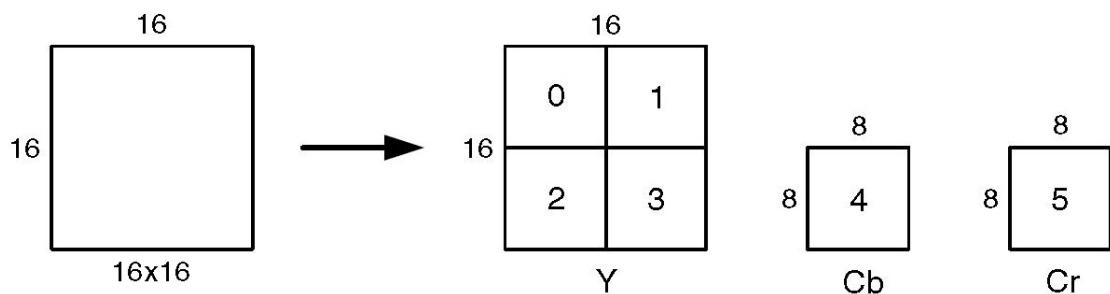
Slika se podijeli u određenu količinu makroblokova veličine  $M \times N$  elemenata slike i onda se nad njima vrši nadomještanje pokreta.

Blokovsko nadomještanje pokreta je popularno iz više razloga. Potrebno mu je manje računalnih resursa, manje računanja jer imamo manje blokova od elemenata slike, odgovara okviru slike koji je pravokutan, blokovska podjela odgovara i transformacijama koje također koriste pravokutne oblike. Iako je ponekad teško pravokutnim oblicima nadomjestiti pokret koji ne odgovara pravokutnom pomaku, postoje razne tehničke kojima se takve stvari olakšavaju.

### 3.4.1. Nadomještanje pokreta makroblokova

Slika se može podijeliti u nekoliko veličina makroblokova. Standardne veličine su 16x16, 8x8 i 4x4. Svaka od veličina ima svoje prednosti i mane. Manje veličine bolje nadomještaju pokrete ali ih ima više pa je ukupna količina podataka veća, dok većih makroblokova ima manje ali lošije nadomještaju određene promjene.

Za 4:2:0 format slike, makroblok 16x16 je prikazan na slici 3.10.



**Slika 3.10** Makroblok za format slike 4:2:0

Sastoji se od 256 luminantnih uzoraka i 2 puta po 64 krominantna uzorka.

Nakon podjele slike u makroblokove, za svaki određeni makroblok traži se određeni dio slike u koji on u sljedećoj slici dolazi. Nakon pronaleta područja, od tog dijela slike oduzima se makroblok iz prethodne slike i na taj način dobijemo rezidualni makroblok (luminantni i krominantni).

Postoji nekoliko načina nadomještanja pokreta, korištenjem prethodne slike, buduće, nekoliko prethodnih, nekoliko budućih, i jedne i druge itd. sve to pridonosi efikasnijoj kompresiji podataka.

Promjene između dvije slike rijetko prate podjelu na makroblokove veličine 16x16. Zato je pametno koristiti varijabilnu veličinu makrobloka koja bi odlučivala koja područja bi dijelila na veće a koja na manje makroblokove.



**Slika 3.11** Dvije uzastopne sličice videosekvene

Na slici 3.12 prikazana je rezidualna slika u kojoj je pohranjena određena količina energije.



**Slika 3.12** Rezidualna slika nastala oduzimanjem slika sa slike 3.11

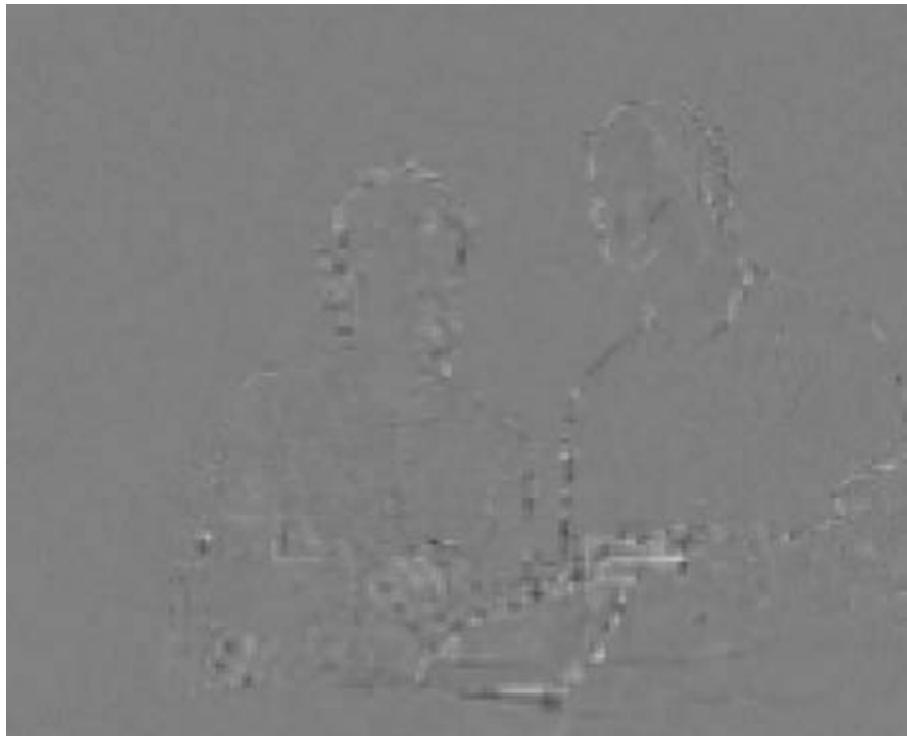
Na slikama 3.13, 3.14 i 3.15 prikazane su rezidualne slike s promjenjivom veličinom makroblokova.



**Slika 3.13** Rezidualna slika, makroblokovi 16x16



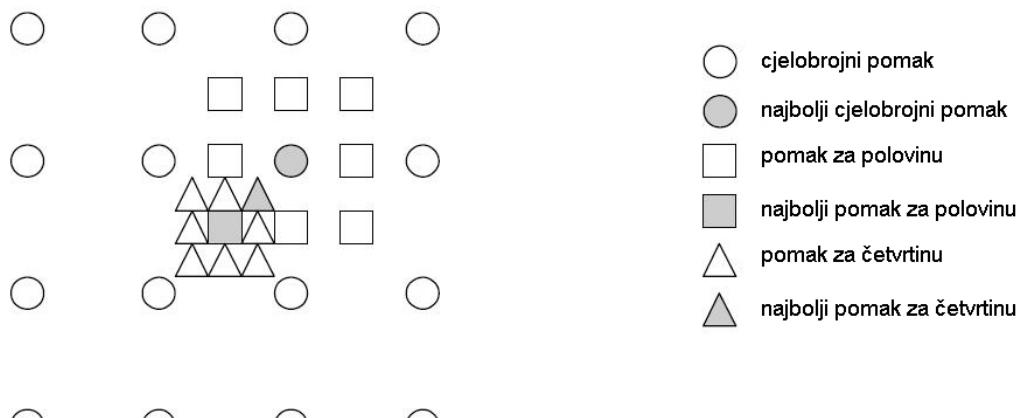
**Slika 3.14** Rezidualna slika, makroblokovi 8x8



**Slika 3.15** Rezidualna slika, makroblokovi 4x4

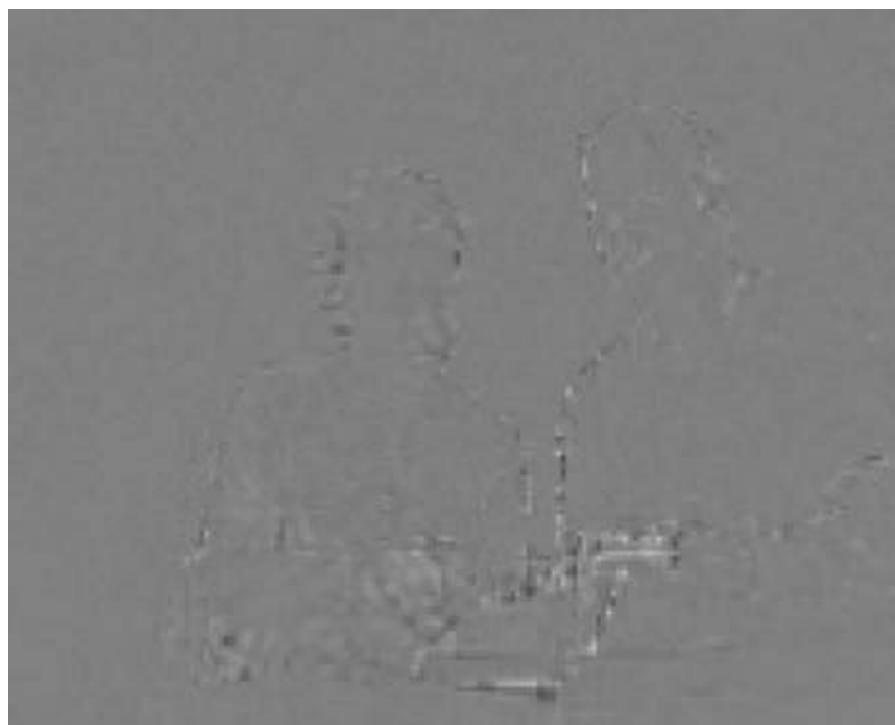
Vidimo da manji makroblokovi kao rezultat daju rezidualne slike s manje energije, ali imamo veći tok podataka koji odgovara vektorima pomaka, dok nam veći makroblokovi daju veću energiju u rezidualnoj slici uz manji broj vektora pomaka. Kompromis je korištenje promjenjive i adaptivne veličine makrobloka, što je i implementirano u H.264 postupku kompresije.

Određeno poboljšanje postupka kompresije postiže se necjelobrojnim vektorom pomaka. Tako imamo pomak od cijelog broja, pola elementa slike ili četvrtinu elementa slike. Kao što je pokazano na slici 3.16 koder prvo pronalazi najbolji pomak za cjelobrojni vektor pomaka. Zatim prelazi na pola vektora pomaka, a onda na četvrtinu vektora pomaka i uzima onaj koji najviše odgovara.

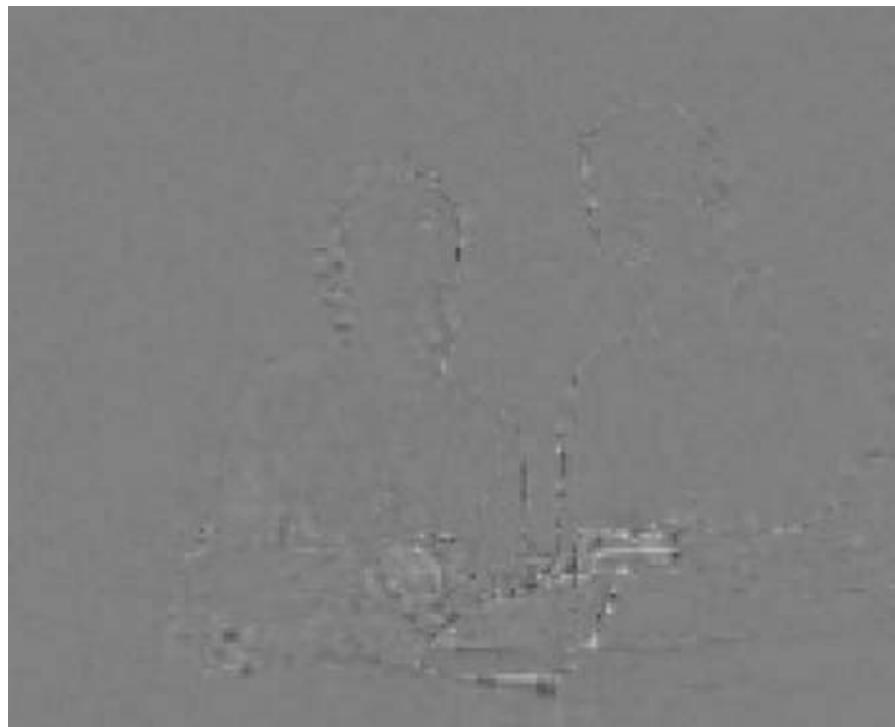


**Slika 3.16** Cjelobrojni, polovina i četvrtina razmaka između elemenata slike

Na slikama 3.17 i 3.18 možemo vidjeti kako se energija rezidualnog signala smanjuje za necjelobrojni vektor pomaka kompenzacije pokreta.



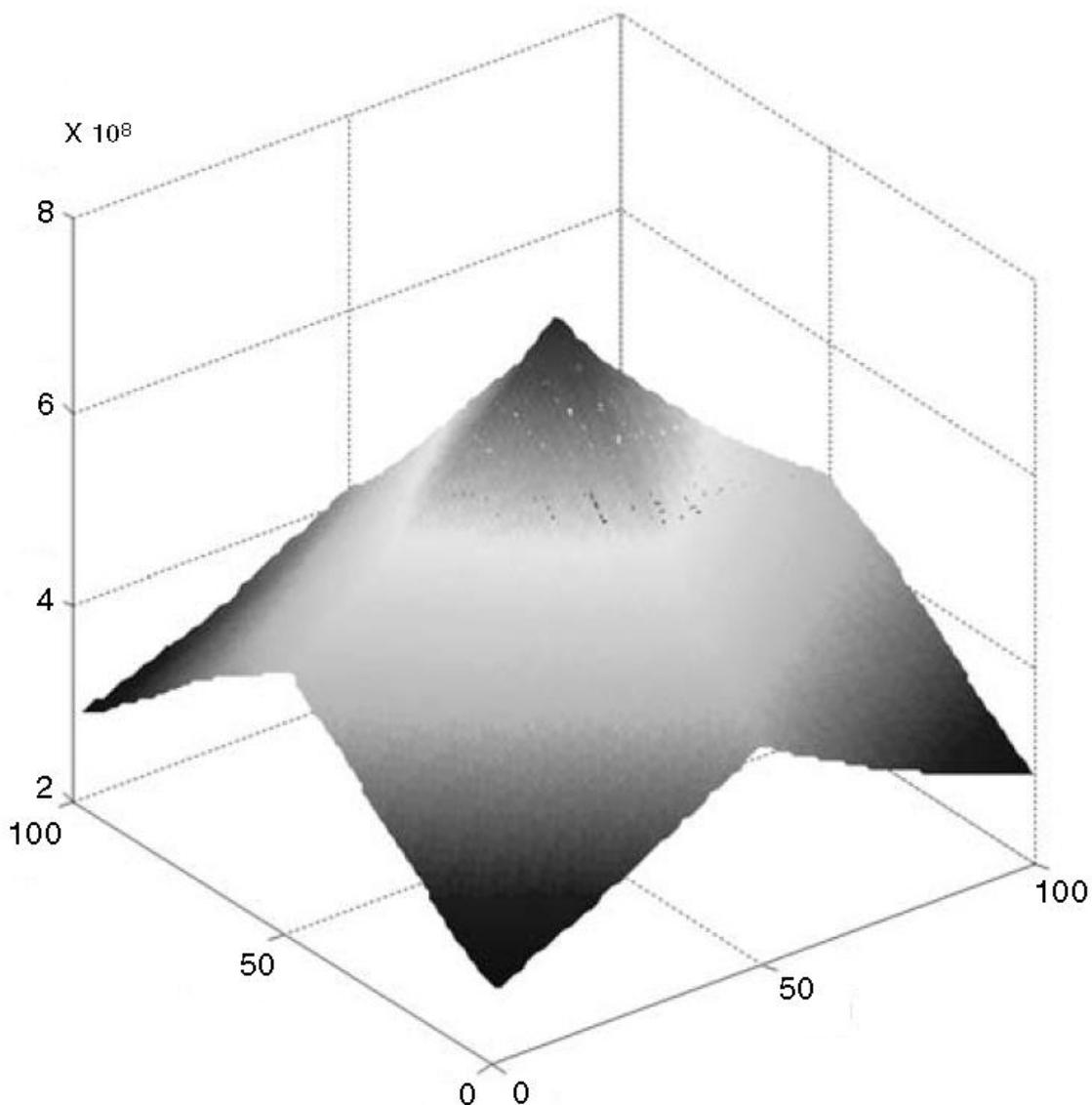
**Slika 3.17** Rezidualna slika, makroblokovi 4x4 i polovina cjelobrojnog pomaka



**Slika 3.18** Rezidualna slika, makroblokovi  $4 \times 4$  i četvrtina cijelobrojnog pomaka

### 3.5. Model slike

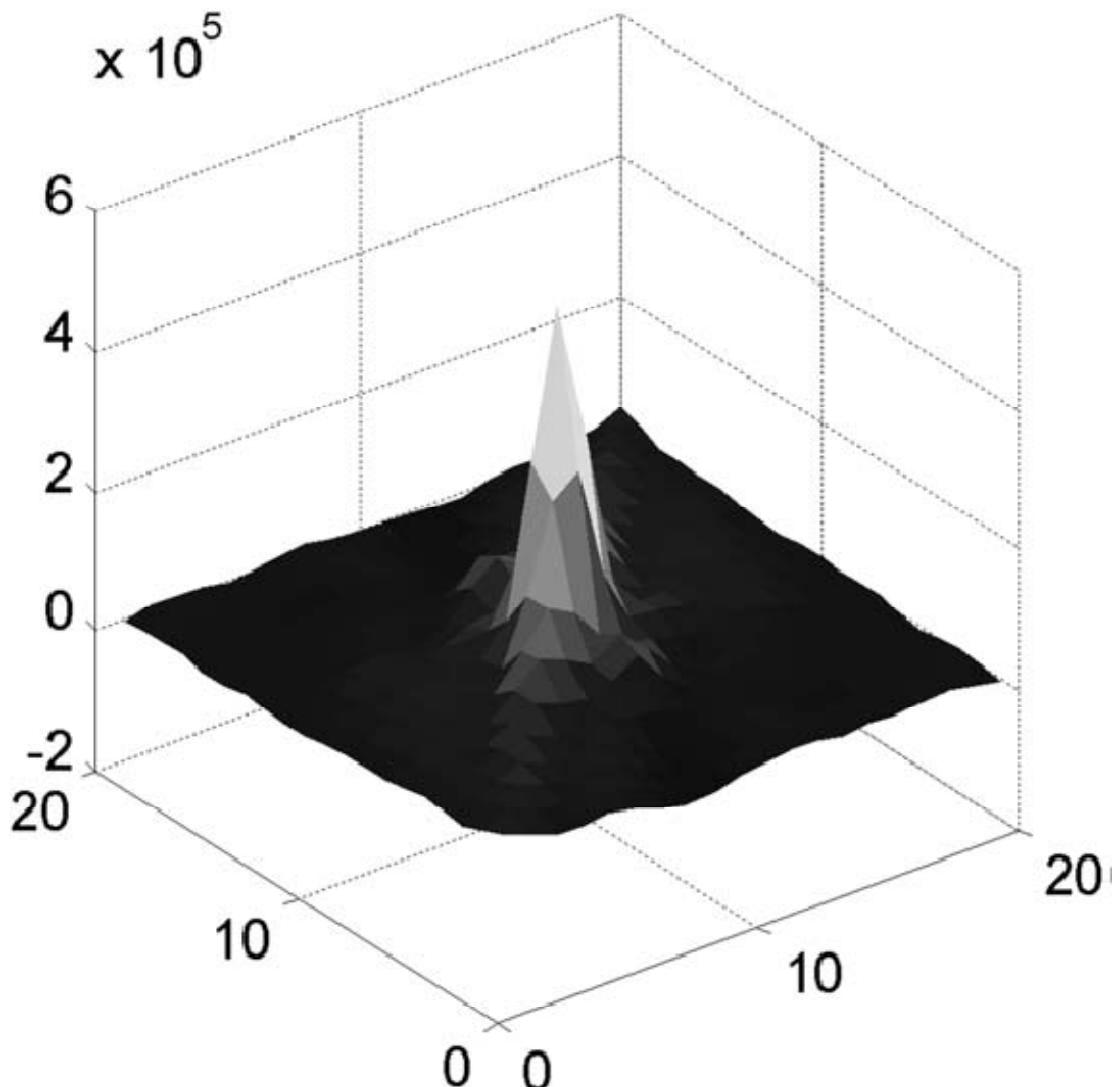
Prirodne slike (a time i videosekvene) su autokorelirane i zbog toga ih je teško komprimirati u originalnoj formi. Na slici 3.19 prikazana je dvodimenzionalna autokorelacijska funkcija prirodne slike. Možemo vidjeti kako su susjedni elementi slike poprilično slični.



**Slika 3.19** Autokorelacijska funkcija prirodne slike

Na slici 3.20 prikazana je autokorelacijska funkcija rezidualne slike u kojoj je korištena kompenzacija pokreta. Kao što vidimo na slici, sličnost susjednih elemenata slike eksponencijalno opada kako se udaljujemo od početnog elementa slike, tj. susjedni elementi slike nisu previše korelirani. Efikasno nadomještanje

pokreta smanjuje korelaciju između elemenata rezidualne slike i čini ju pogodnijom za kompresiju.



**Slika 3.20** Autokorelacijska funkcija rezidualne slike

Uloga modela slike je dekoreliranje elemenata slike i pretvorba slike u oblik pogodan za efikasno komprimiranje korištenjem entropijskog kodiranja. Najčešće korišteni modeli slike obavljaju tri funkcije (transformaciju, kvantizaciju i preslagivanje podataka)

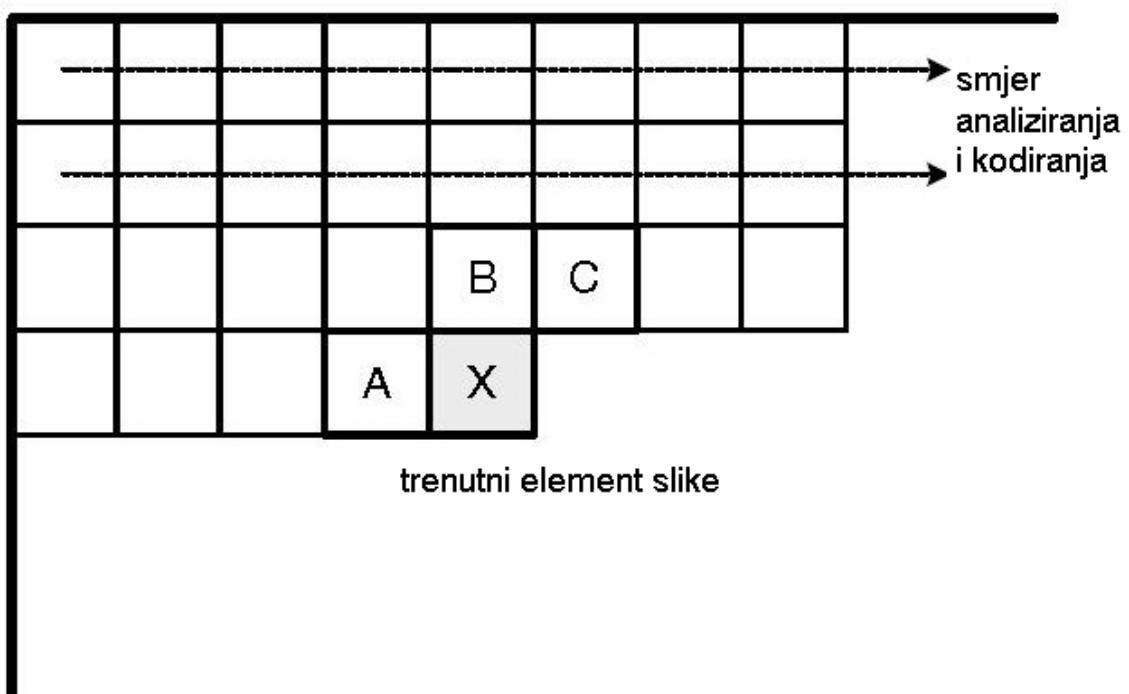
### 3.6. Predikcijsko kodiranje

Nadomještanje pokreta je jedan od primjera predikcijskog kodiranja. U njemu koder predviđa izgled sličice na temelju prethodno dekodiranih i rezidualnih slika. Ako je funkcija dobro obavljena, onda je energija u rezidualnoj slici mala, tj. može se kodirati s manjim brojem bitova.

Na sličan način, predikcijsko kodiranje se može obaviti i unutar jedne slike na temelju prethodno komprimiranih dijelova te iste slike. Na taj način dolazimo do puno manjeg niza podataka kojeg trebamo pohraniti ili poslati, tj. imamo efikasniju kompresiju.

Prostorna kompresija se ponekad naziva i DPCM (Differential Pulse Code Modulation) po modulacijskom postupku u radiokomunikacijama.

Slika 3.21 prikazuje nam element slike  $X$  koji trebamo kodirati. Slika je analizirana i kodirana u linijama i na taj način je njegovo susjedstvo (elementi  $A$ ,  $B$  i  $C$ ) ranije kodirano.



**Slika 3.21** Prostorna predikcija (DPCM)

Određenim postupkom koder je došao do vrijednosti za element  $X$  koji iznosi:

$$P(X) = \frac{2A + B + C}{4}$$

Zatim od elementa  $X$  oduzima dobivenu vrijednost i dolazimo do rezidualnog elementa

$$R(X) = X - P(X)$$

U dekoderu obavljamo suprotan proces. Dekoder dolazi do iste predikcije za  $P(X)$

$$P(X) = \frac{2A + B + C}{4}$$

I na temelju toga rekonstruira element slike  $X$

$$X = R(X) + P(X)$$

Efikasnost predikcijskog kodiranja ovisi o kvaliteti predviđanja za  $P(X)$ . Ako je vrijednost  $P(X)$  bliska vrijednosti  $X$ , tada je u rezidualnom elementu  $R(X)$  pohranjena mala količina energije i on se može efikasno kodirati.

Problem metode je težina postupka predviđanja. Gotovo je nemoguće napraviti koder koji bi jednako dobro predviđao sve dijelove slike, pa se ide na adaptivni način predviđanja koji bi na različite načine predviđao različite dijelove slike.

### 3.7. Transformacijsko kodiranje

kodiranje se uobičajeno obavlja nad rezidualnim slikama i svrha mu je prelazak u drugu domenu. Transformaciju odabiremo tako da:

1. u drugoj domeni podaci budu što je moguće više dekorelirani, i većina energije signala mora biti skoncentrirana u što manjem broju elemenata
2. bude inverzna
3. koristi što manje računalnih resursa

Transformacije slike dijele se na dvije osnovne grupe. Jedna grupa je blokovski orijentirana, tj. transformacija se obavlja nad blokovima, dok se u drugoj grupi transformacija obavlja nad cijelom slikom.

U prvu grupu ulaze Karhunen-Loeveova transformacija, SVD transformacija (Singulare Value Decomposition) i diskretna kosinusna transformacija (DCT). Svaka od naboranih transformacija obavlja se nad makroblokom slike dimenzija NxN. Transformacije temeljene na podjeli slike u blokove su računalno manje zahtjevne i odgovaraju blokovskoj podjeli slike kod nadomještanja pokreta. Mana im je ponekad vidljiva blokovska podjela slike kod većih stupnjeva kompresije.

U drugoj grupi je diskretna wavelet transformacija DWT (Discrete Wavelet Transform) koja mogućnostima nadilazi transformacije koje djeluju nad blokovima, ali je računalno zahtjevnija i ne odgovara joj blokovska podjela kod nadomještanja pokreta.

Za primjene u kompresiji video signala koristi se diskretna kosinusna transformacija.

### 3.7.1.DCT

Diskretna kosinusna transformacija je najčešće korištena transformacija u postupcima kompresije videosignalata. Ona djeluje nad blokom  $\mathbf{X}$ , dimenzija NxN (rezidualne ili čiste slike). Na taj način dolazimo do bloka koeficijenata  $\mathbf{Y}$ , također dimenzija NxN.

Operacije transformacije DCT i njezine inverzne transformacije IDCT mogu se opisati matričnom algebrrom, tj. transformacijama operatora transformacije kojim prelazimo iz jednog vektorskog prostora u drugi (iz prostora funkcije u prostor transformacije, tj. iz prostora signala, vremenske domene u prostor transformacije, frekvencijsku domenu) gdje nam je matrica  $\mathbf{A}$  matrica transformacije.

Transformacija DCT kraće se zapisuje:

$$\mathbf{Y} = \mathbf{AXA}^T$$

a inverzna transformacija dana je s:

$$\mathbf{X} = \mathbf{A}^T \mathbf{YA}$$

Gdje je  $\mathbf{X}$  matrica uzorka,  $\mathbf{Y}$  matrica koeficijenata a  $\mathbf{A}$  transformacijska matrica dimenzija NxN.

Elementi matrice  $\mathbf{A}$  su:

$$A_{ij} = C_i \cos \frac{(2j + i)i\pi}{2N}$$

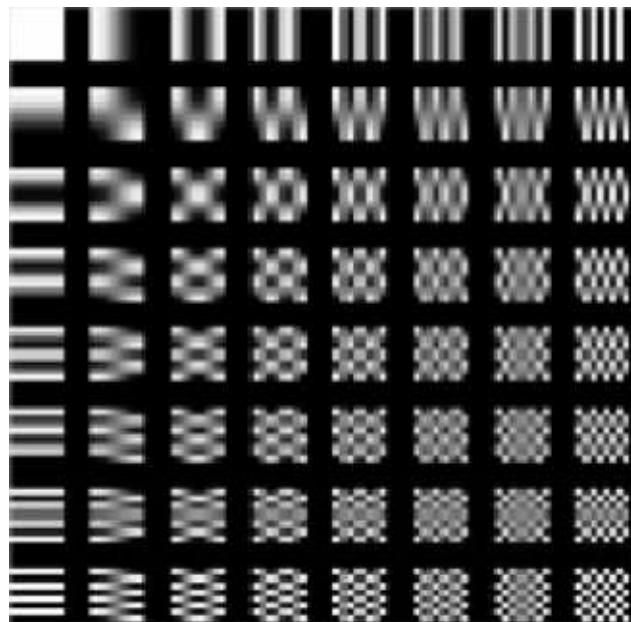
$$C_i = \sqrt{\frac{1}{N}} \quad (i = 0) \quad C_i = \sqrt{\frac{2}{N}} \quad (i \neq 0)$$

Transformacija se za lakšu implementaciju može zapisati u formi sume kao:

$$Y_{xy} = C_x C_y \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} X_{ij} \cos \frac{(2j + 1)y\pi}{2N} \cos \frac{(2i + 1)x\pi}{2N}$$

$$X_{ij} = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} C_x C_y Y_{xy} \cos \frac{(2j + 1)y\pi}{2N} \cos \frac{(2i + 1)x\pi}{2N}$$

Nakon djelovanja dvodimenzionalne diskretne kosinusne transformacije dobijemo matricu koeficijenata dimenzije NxN koja nam predstavlja određeni blok slike u DCT domeni (analogno frekvencijskoj domeni).



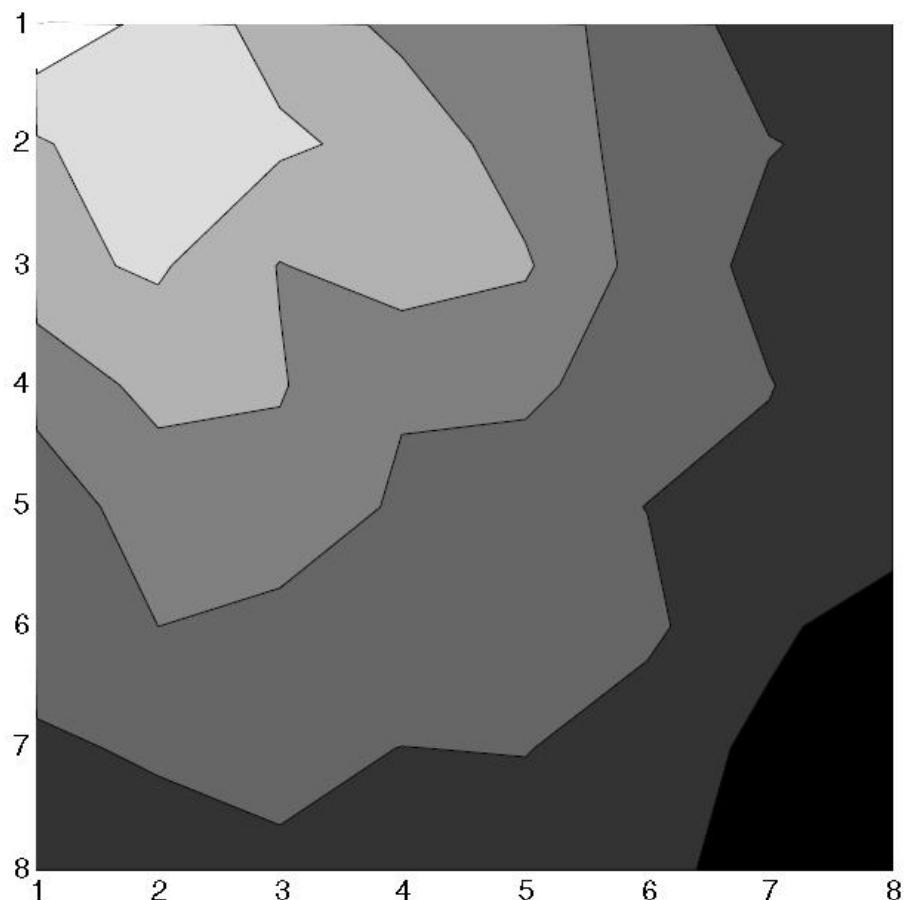
**Slika 3.22** Bazne funkcije za 8x8 DCT transformaciju

Na slici 3.22 prikazane su bazne funkcije za 8x8 DCT transformaciju. Na slici vidimo da su u gornjem lijevom kutu stacionirane niskofrekvenčne bazne

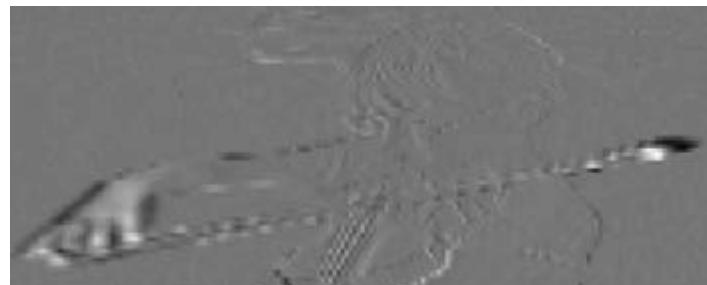
funkcije, dok frekvencija izmjene u horizontalnom i vertikalnom smjeru raste kako idemo prema dolnjem desnom kutu. Svaka slika može se predstaviti kao kombinacija baznih funkcija s određenim težinskim faktorom.

Sam postupak DCT ne unosi nikakvu kompresiju u sustav jer on predstavlja samo prelazak u drugo područje s jednakim brojem koeficijenata.

Prirodne slike podijeljene u blokove, na razini blokova ne sadrže puno visokih frekvencija. Koeficijenti DCT-a u tom slučaju su stacionirani oko gornjeg lijevog kuta (niskih frekvencija) dok su oni koji pripadaju visokim frekvencijama izrijekom mali ili jednaki nuli, kao što vidimo na slici 3.23 koja predstavlja DCT transformaciju slike 3.24.



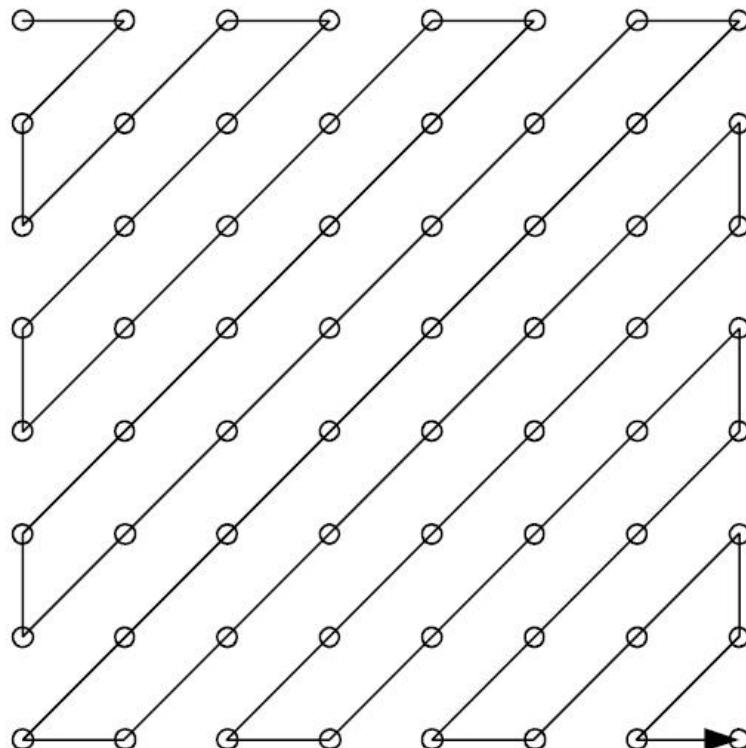
**Slika 3.23** DCT koeficijenti, makroblokovi dimenzija 8x8



**Slika 3.24** Rezidualna slika

Ljudski vizualni sustav bolje percipira niske prostorne frekvencije a kako te frekvencije rastu, percepcija opada. Upravo iz tog razloga moguće je odbaciti određenu količinu visokofrekvencijskih komponenti i tako učiniti određenu kompresiju videosignalata bez vidljivije distorzije slike.

Uz eliminiranje visokofrekvencijskih komponenti moguće je pridonijeti stupnju kompresije i na malo drugačiji način analiziranja bloka. Blok nećemo analizirati u vodoravnom smjeru nego cik-cak načinom analiziranja (slika 3.25). Prvo ćemo pokupiti sve koeficijente koji odgovaraju niskim frekvencijama a zatim ćemo grupirati koeficijente koji odgovaraju visokim frekvencijama. Na taj način dobivamo niz s velikom količinom nula, pa je eliminacijom tih nula moguće smanjiti niz podataka i doći do efikasnijeg stupnja kompresije.



**Slika 3.25** Cik-cak način analiziranja

### 3.8. Entropijsko kodiranje

Entropijsko kodiranje je jedan od načina kompresije podataka bez gubitaka. Temelji se na statističkim metodama procjene vjerojatnosti pojavljivanja određenih simbola. Simbolima koji se pojavljuju češće dodjeljuje se kraća kodna riječ, dok se onima koji se pojavljuju rjeđe dodjeljuje dulja kodna riječ. Duljina kodne riječi  $L_c$  proporcionalna je vjerojatnosti pojavljivanja simbola a dana je formulom:

$$L_c = \log_b \frac{1}{P}$$

gdje je  $P$  vjerojatnost pojavljivanja ulaznog simbola, a  $b$  broj simbola korištenih za stvaranje kodne riječi.

Najpoznatije metode entropijskog kodiranja su Huffmanovo i aritmetičko kodiranje.

#### 3.8.1. Huffmanov kod

Na osnovu statističke procjene pojavljivanja određog simbola, svakom simbolu pridružena je kodna riječ promjenljive dužine, i to na način da se simbolu koji se učestalije ponavlja dodjeljuje kraća kodna riječ a onom čija je frekvencija pojavljivanja manja, duža kodna riječ. Iz ovoga vidimo da je za svaku određenu implementaciju Huffmannovog koda potrebno proračunati kodnu tablicu. U tablici 2.1 prikazana je tablica Huffmannovog koda za neke vrijednosti vektora pomaka i njima pripadne vjerojatnosti pojavljivanja.

Iznos Vektora	Vjerojatnost	Kodna riječ	Broj bita	Optimalan broj bita [ $\log_2(1/P)$ ]
-2	0.1	000	3	3.32
-1	0.2	010	3	2.32
0	0.4	1	1	1.32
1	0.2	011	3	2.32
2	0.1	001	3	3.32

**Tablica 3.1** Tablica Huffmannovog koda

## 4. H.264 standard za videokompresiju

### 4.1. Uvod

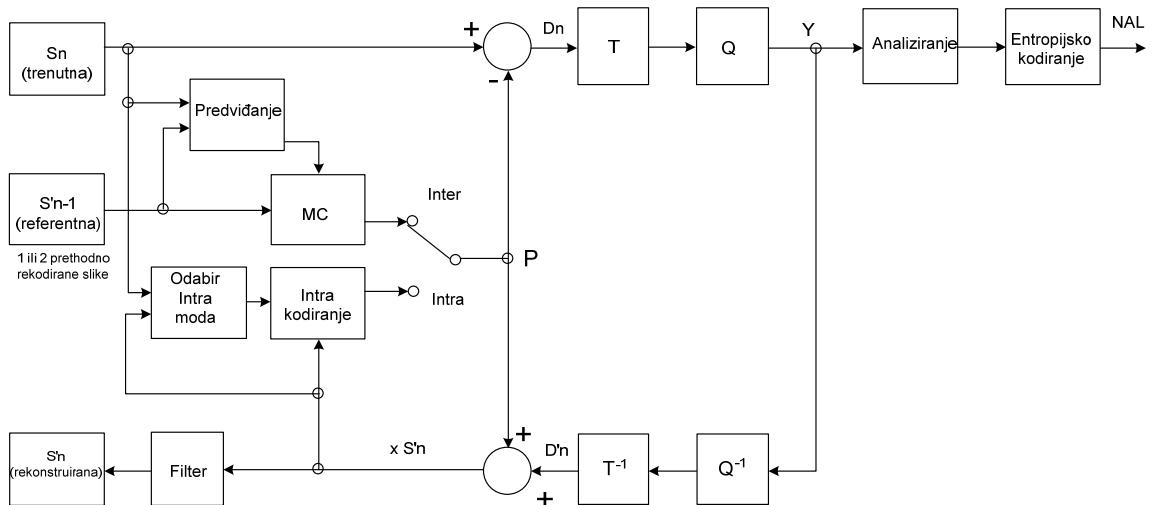
H.264 je standard za videokompresiju, ekvivalentan standardima MPEG4 Part 10 i MPEG4 AVC (Advanced Video Coding, napredno kodiranje videa), razvijen od strane ITU-T VCEG grupe (Video Coding Experts Group) i ISO/IEC MPEG (Moving Picture Experts Groupe) grupe. Namjera im je bila stvoriti standard za kompresiju videa koji bi bio u stanju komprimirati video u jednaku kvalitetu kao i prethodni standardi ali s manjim tokom podataka i bez velikog porasta kompleksnosti postupka kompresije. H.264 standard daje jednaku kvalitetu kodiranog videa kao MPEG2, H.263 ili MPEG4 Part 2 uz dvostruko manje brzine prijenosa. Osim poboljšanja efikasnosti postupka kompresije, cilj je bio i poboljšanje primjene standarda u širokom spektru aplikacija (radiodifuzija, vjeokonferencijski prijenosi, video na zahtjev, live streaming, multimedijiske poruke...).

### 4.2. H.264 koder

H.264 je standard za hibridno kodiranje. Koriste se, u prethodnom poglavlju spomenuti, postupci za uklanjanje prostorne i vremenske redundancije, uklanjanje frekvencijske redundancije transformacijom u frekvencijsku domenu, te na kraju entropijsko kodiranje i prilagodba toka podataka na različite sustave.

U H.264 normi podržano je kodiranje progresivno analiziranih slika i slika analiziranih s proredom.

Na slici 4.1 prikazan je proces kodiranja. Ne smanjujući općenitost, opisat ćemo kodiranje slike podijeljene u makroblokove dimenzije 16x16. U tom slučaju makroblok luminantnih komponenti je dimenzije 16x16, a makroblok krominantnih komponenti je dimenzije 8x8.


**Slika 4.1** H.264 koder

U koderu imamo dva toka podataka. Tok podataka s lijeva na desno je tok kodiranja, a tok s desna na lijevo je tok rekonstrukcije.

Na ulaz kodera dovodimo trenutnu sliku (ili polusliku u slučaju analiziranja s proredom). Ona se uspoređuje s prethodnom slikom (referentnom) i vrši se predikcija. Na izlazu dobijemo makroblok predikcije  $P'$  koji se uspoređuje s trenutnim makroblokom. To se obavlja za cijelu sliku. Rezultati usporedbe proslijedjuju se u dio za transformaciju u ferekvencijsku transformaciju. Koeficijenti nastali transformacijom se kvantiziraju i obrađuju te zatim proslijedjuju sustavu za entropijsko kodiranje koje je završni dio obrade. Na izlazu iz kodera dobijemo tok podataka koji odgovara kodiranoj slici.

Osim kodiranja slike na temelju predviđanaj pokreta iz referentnih slika, slike se mogu kodirati unutar sebe, tj. bez predviđanja iz referentnih slika. Na svakih nekoliko odaslanih slika kodiranih na temelju predviđanja pokreta iz referentnih slika, moramo odaslati i jednu sliku kodiranu unutar sebe koja će nam biti referentna za buduća predviđanja. Na taj se način smanjuje akumulacija greške.

Slike kodirane unutar sebe kodiraju se također na principu predviđanja, ali makroblokovi se predviđaju na temelju susjednih prethodno kodiranih i lokalno rekonstruiranih makroblokova. Nakon prvog dijela koji je drugačiji u odnosu na kodiranje s predviđanjem pokreta iz referentnih slika, nastavak procesa je identičan.

Tok podataka s desna na lijevo odgovara toku rekonstrukcije. Kako se koji blok transformira i kvantizira, tako se vraća natrag u suprotnom smjeru te se vrše inverzni postupci kvantiziranja i transformiranja. Na taj se način dobiju makroblokovi trenutne slike nakon rekonstrukcije koji služe kao referentni susjedni makroblokovi pri kodiranju slike unutar sebe.

U H.264 standardu za kompresiju videa koristi se pet različitih vrsta slika:

I slike – makroblokovi su kodirani predviđanjem pokreta unutar slike,

P slike – makroblokovi su kodirani predviđanjem uz nadomještanje pokreta pomoću prethodno kodiranih referentnih slika,

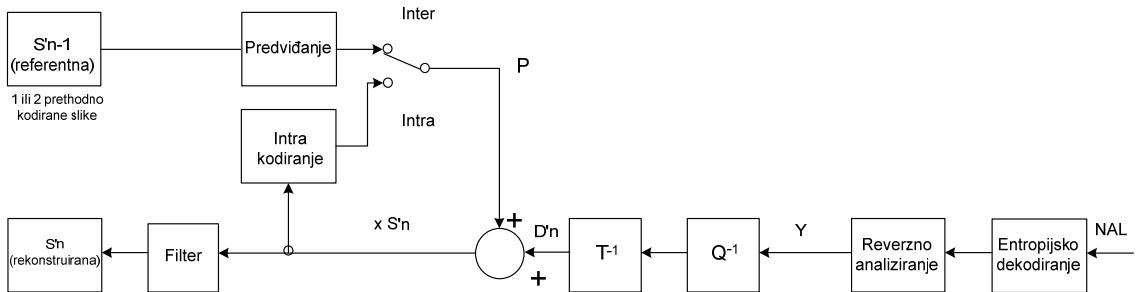
B slike – makroblokovi su kodirani predviđanjem uz nadomještanje pokreta iz prethodnih, ali i budućih slika,

SP (Switching Predictive) slike – sadrže makroblokove kodirane na I i P način,

SI (Switching Intra) slike – sadrže makroblokove kodirane na I i P način ali i poseban oblik makroblokova, SI makroblokove (poseban slučaj makroblokova kodiranih unutar same slike).

### 4.3. H.264 dekoder

Proces obrnut kodiranju zove se dekodiranje i obavlja se u H.264 dekoderu kojemu na ulaz dovodimo kodirani tok podataka. Na ulazu se obavlja entropijsko dekodiranje (dobijemo kvantizirane transformacijske koeficijente) te procesi koji su inverzni procesima kvantiziranja i transformacije. Na izlazu iz procesa inverzne transformacije, dobijemo prediskocijski blok kojem dodajemo prenesenu razliku između predviđene i trenutne slike kako bi se rekonstruirala referentna slika. Tako dobivena slika se filtrira kako bismo smanjili utjecaj blokovske podjele slike. Na slici 4.2 prikazan je postupak dekodiranja videosekvence kodirane H.264 postupkom.

**Slika 4.2** H.264 dekoder

#### 4.4. Kodiranje unutar slike

Za kompresiju I slika koristi se kodiranje unutar slike (Intra kodiranje), kako bi se smanjio tok podataka i iskoristila korelacija između pojedinih elemenata slike. Ona se obično koristi za kodiranje prve slike u nizu koja kasnije služi kao referentna slika. Nakon svakih nekoliko slika dobivenih predviđanjem, ubaci se jedna I slika kako bi se smanjila akumulacija pogreške.

Na temelju prethodno kodiranih i rekonstruiranih dijelova slike predviđa se predikcijski blok za luminantne i krominantne komponente. Ostatak postupka je jednak općenitom postupku za videokompresiju.

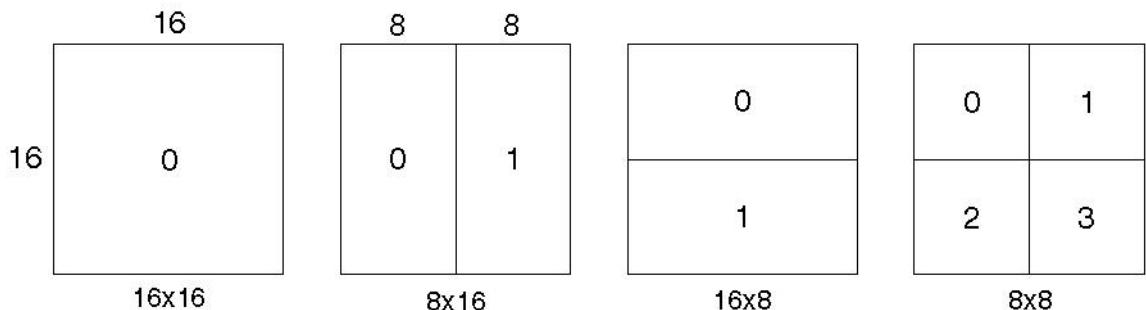
Ovisno o sadržaju slike, predikcijski blok može se formirati u tri veličine ( $4 \times 4$ ,  $8 \times 8$  i  $16 \times 16$ ).

Osim predviđanja, I slike se mogu kodirati i PCM načinom. On omogućuje slanje izvornih uzoraka slike, za neke nepravilne dijelove slike koje je nemoguće dobro predvidjeti iz prethodno kodiranih dijelova. Ovakav način kodiranja određuje gornju granicu količine bitova potrebnih za kodiranje makrobloka.

#### 4.5. Kodiranje predviđanjem iz prethodnih slika

Kodiranje predviđanjem iz prethodnih slika je postupak kodiranja koji koristi vremensku korelaciju između uzastopnih sličica u videosekvenci. Postupak predviđanja se provodi u odnosu na nekoliko prethodno kodiranih slika (referentnih slika). Tako nastali makroblok oduzima se od trenutnog makrobloka i rezultantni makroblok se dalje obrađuje. H.264 norma podržava različite veličine

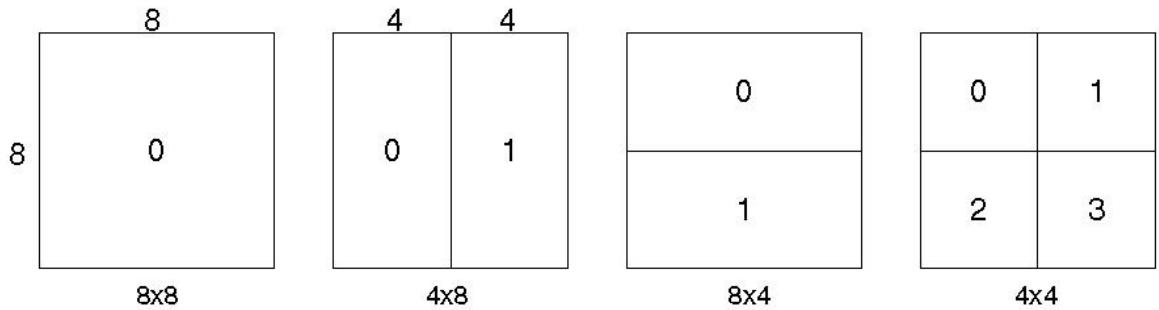
makroblokova unutar predviđenog makrobloka, od kojih svaki generira vlastiti vektor pomaka. Pored kodiranih vektora pomaka i uzoraka slike, do dekodera se prenosi i informacija o odabranoj veličini makrobloka. Kao što vidimo na slici 4.3 makroblokovi veličine 16x16 dijele se na četiri klase makroblokova.



**Slika 4.3** Podjela makrobloka dimenzija 16x16

Upotreba manjih blokova rezultira većom točnošću predviđanja jer je manjim blokovima moguće uspješnije nadomjestiti pokret. Tim načinom smanjuje se broj bitova kojima je potrebno kodirati rezidualne makroblokove (makroblokove nastale oduzimanjem predviđenih od trenutnih), ali se povećava količina vektora pomaka, pa je potrebno pronaći ravnotežu.

Odabirom makroblokova veličine 8x8 moguće je dijeliti postojeći makroblok na još manje makroblokove kao što je i prikazano na slici 4.4.

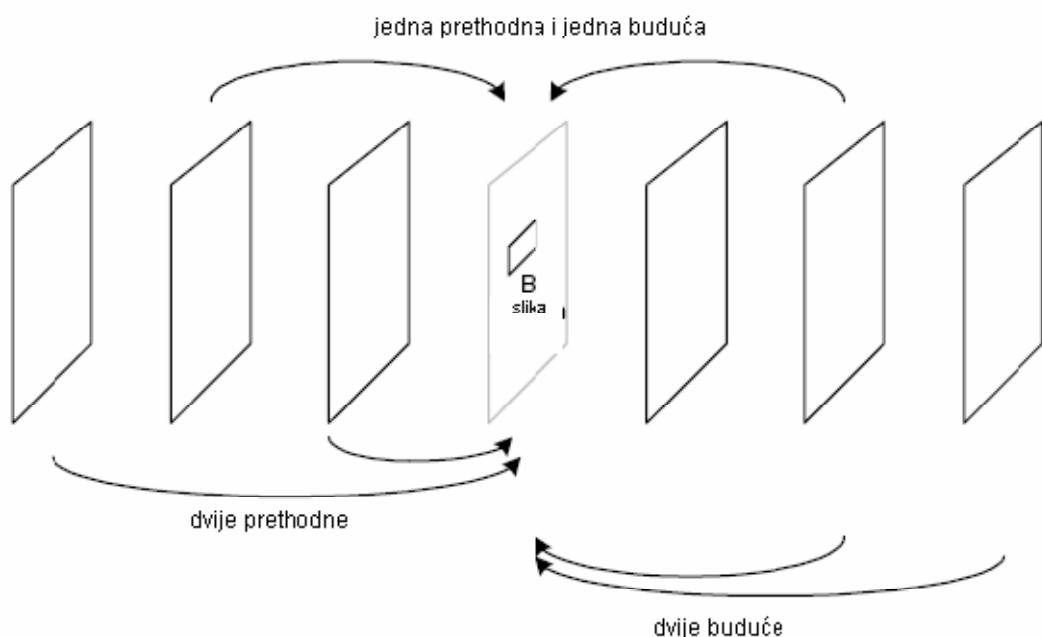


**Slika 4.4** Podjela makrobloka dimenzija 8x8

Veličina krominantnih makroblokova ovisi o načinu uzorkovanja informacije o boji.

## 4.6. Dvosmjerno kodiranje

Novost u H.264 normi je postojanje B slika. U njima se pokret ne predviđa samo iz prethodnih slika nego se kao referentne slike koriste i slike koje dolaze nakon trenutne slike. Zato se taj način kodiranja naziva dvosmjerno kodiranje. Na taj način se povećava efikasnost postupka predviđanja, a kao rezultat dobijemo B slike u kojima je snaga signala još manja nego kod P slika. Ostatak procesa kodiranja nakon procesa predviđanja odgovara ranije opisanom postupku koji je kod svih slika jednak. Na slici 4.5 vidimo postupak predviđanja iz prethodnih i budućih slika.



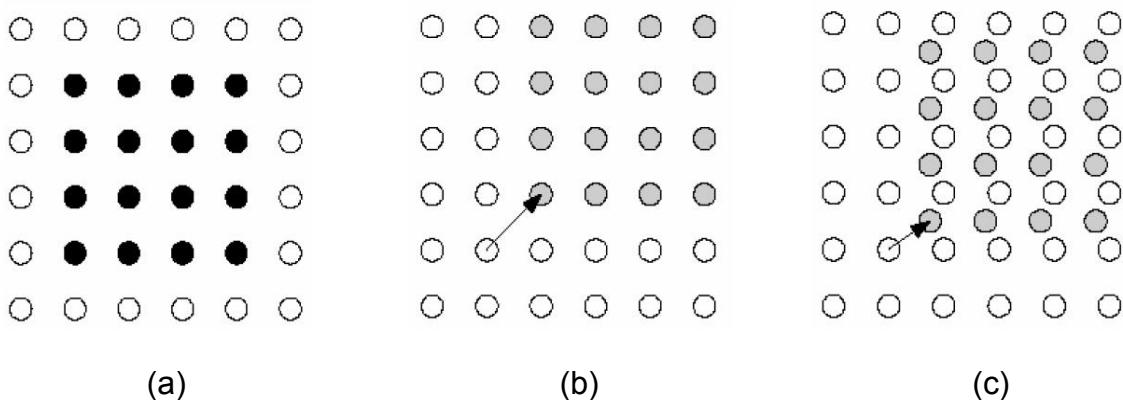
**Slika 4.5** Postupak predviđanja iz prethodnih i budućih slika

Za dvosmjerno kodiranje H.264 norma koristi 2 referentna spremnika za pohranu prethodno kodiranih slika. Spremniči sadrže slike koje dolaze prije i poslije slike koja se trenutno kodira.

## 4.7. Vektori pomaka

Predviđanje trenutnog makrobloka vrši se pomoći makrobloka iste veličine koji se nalazi u slici koja ima ulogu referentne slike. Prostorni pomak između te dvije pozicije odgovara vektoru pomaka.

U H.264 normi pomak između dva makrobloka ne mora odgovarati cjelobrojnoj veličini. Na slici 4.6 možemo vidjeti dva pomaka. Jedan odgovara cjelobrojnoj veličini a drugi necjelobrojnoj. Ako nam pomak ne odgovara cjelobrojnoj veličini, tada se interpolacijom između susjednih elemenata slike dolazi do vrijednosti određenog elementa slike.

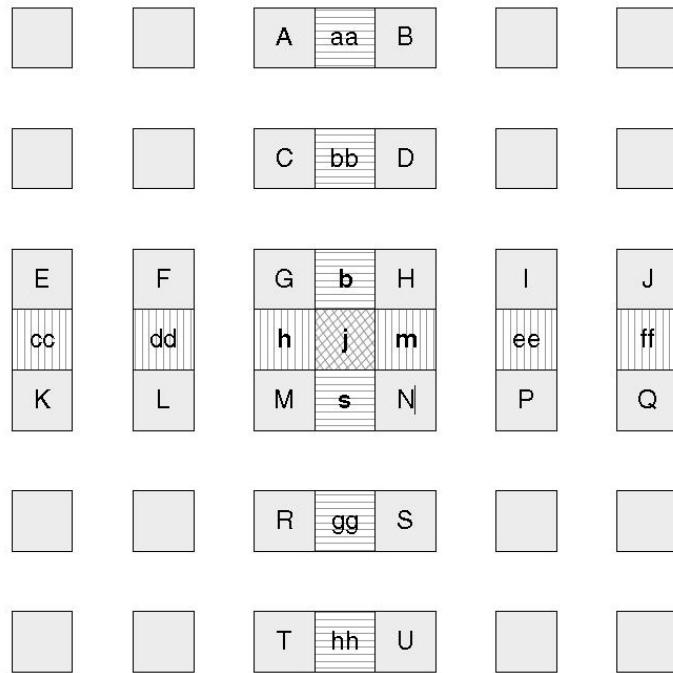


**Slika 4.6** (a) 4x4 blok trenutne slike (b) cjelobrojni pomak (c) necjelobrojni pomak

H.264 standard u procesu predviđanja koristi interpolacijske tehnike s točnošću od 1/4 za luminantne uzorke i 1/8 za krominantne uzorke.

Interpolacija se provodi u dva koraka. U prvom koraku određuje se vrijednost uzorka na polovici udaljenosti dva cjelobrojna uzorka u horizontalnom i vertikalno smjeru. Interpolacija se provodi upotrebom digitalnog filtra koji daje težinsku sumu šest susjednih uzoraka koji se nalaze na cjelobrojnim udaljenostima slike.

Na slici 4.7 prikazan je postupak interpolacije luminantnih komponenti koje se nalaze na polovini cjelobrojnog pomaka.



**Slika 4.7** Postupak interpolacije luminantnih komponenti

Cjelobrojni uzorci označeni su cijelim brojevima, a uzorci označeni malim slovima predstavljaju uzorke na polovici udaljenosti. Računaju se primjenom filtra na sljedeći način:

$$b = \text{round} \left( \frac{E - 5F + 20G + 20H - 5I + J}{32} \right)$$

$$h = \text{round} \left( \frac{A - 5C + 20G + 20M - 5R + T}{32} \right)$$

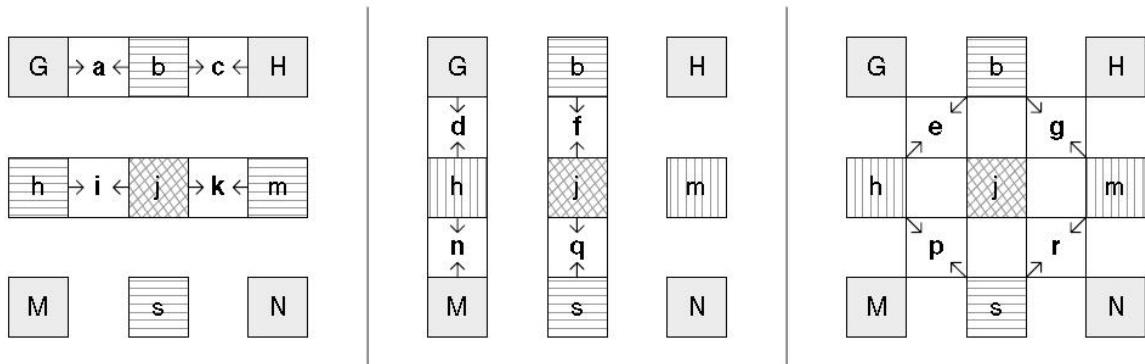
$$m = \text{round} \left( \frac{B - 5D + 20H + 20N - 5S + U}{32} \right)$$

$$s = \text{round} \left( \frac{K - 5L + 20M + 20N - 5P + Q}{32} \right)$$

Uzorak  $j$  određuje se interpolacijom iz uzoraka  $cc$ ,  $dd$ ,  $h$ ,  $m$ ,  $ee$  i  $ff$  ili  $aa$ ,  $bb$ ,  $b$ ,  $s$ ,  $gg$  i  $hh$ .

U drugom interpolacijskom koraku određuju se vrijednosti uzorka na četvrtini udaljenosti između cjelobrojnih udaljenosti. Vrijednosti se određuju linearnom interpolacijom pomoću uzorka dobivenih nakon prvog koraka interpolacije.

Na slici 4.8 vidimo uzorce pomaknute za 1/4 udaljenosti između cjelobrojnih uzorka izvorne slike.



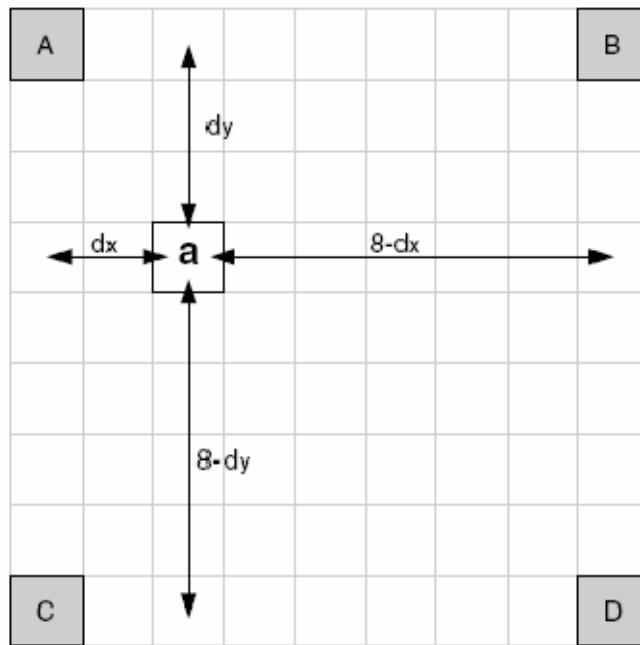
**Slika 4.8** Interpolacija uzoraka za pomak od 1/4

Uzorci  $a, c, i$  i  $k$  predstavljaju uzorce s 1/4 pomaka u horizontalnom smjeru, dok uzorci  $d, f, n$  i  $g$  predstavljaju uzorce s 1/4 pomaka u vertikalnom smjeru. Vrijednosti uzorka računaju se linearnom interpolacijom:

$$a = \text{round}\left(\frac{G + b}{2}\right)$$

Preostali uzorci ( $e, g, p$  i  $r$ ) računaju se linearnom interpolacijom iz dijagonalno suprotnih uzorka.

Kod krominantnih uzorka pomak ovisi o primijenjenom uzorkovanju. Za 4:2:0 uzorkovanje rezoluciji vektora pomaka od 1/4 za luminantne komponente odgovara pomak od 1/8 za krominantne komponente.



**Slika 4.9** Određivanje krominantnih uzoraka na 1/8 udaljenosti

Na slici 4.9 vidimo položaj cjelobrojnih krominantnih uzoraka i položaj pomaknutog krominantnog uzorka za 1/8 udaljenosti. Krominantni uzorak se tada interpolira kao :

$$a = \text{round}\left(\frac{(8 - d_x)(8 - d_y)A + d_x(8 - d_y)B + (8 - d_x)d_yC + d_xd_yD}{64}\right)$$

Na slici imamo  $d_x=2$  i  $d_y=3$ , pa prema tome

$$a = \left(\frac{30A + 10B + 18C + 6D}{64}\right)$$

## 4.8. Predviđanje vektora pomaka

Kodiranjem i prenošenjem svakog pojedinog vektora pomaka, za svaki makroblok, raste tok podataka (pogotovo u slučaju malih makroblokova). Koristeći činjenicu da su prirodne slike strogo korelirane, možemo predvidjeti vektore pokreta na temelju prethodno kodiranih vektora pokreta jer su pokreti susjednih

makroblokova visoko korelirani. Na taj način dolazimo do smanjenja toka podataka. Predviđeni vektor pomaka formira se iz prethodno kodiranih vektora pomaka i to na način da on predstavlja srednju vrijednost njihovih pomaka. Tako dobiveni vektor pomaka oduzima se od trenutnog vektora pomaka i njihova razlika se kodira i prenosi.

## 4.9. Transformacija i kvantizacija

Nakon postupka inter i intra kodiranja, dobivene rezidualne slike transformiraju se u frekvencijsku domenu, gdje se odbacivanjem visokih frekvencija uklanja frekvencijska redundancija. Za razliku od prethodnih normi, u H.264 normi koristi se cjelobrojna transformacija. Na taj način nema gubitka točnosti, a postupak dekodiranja je u potpunosti točan.

Procesom kvantizacije dodatno se smanjuje količina bitova potrebnih za prijenos video sekvence. U H.264 normi koristi se skalarna kvantizacija. Rezultat dijeljenja transformacijskih koeficijenata i kvantacijskih parametara aproksimira se cjelobrojnom vrijednošću. Zbog cjelobrojne aproksimacije rezultata dijeljenja, kvantizacija predstavlja proces s gubicima.

## 4.10. Entropijsko kodiranje

Entropijsko kodiranje je jedan od načina kompresije podataka bez gubitaka. Temelji se na statističkim metodama procjene vjerojatnosti pojavljivanja određenih simbola. Simbolima koji se pojavljuju češće dodjeljuje se kraća kodna riječ, dok se onima koji se pojavljuju rjeđe dodjeljuje dulja kodna riječ. Smanjenje broja bitova očituje se u tome da se elementima koji se češće pojavljuju dodjeljuju kraći kodni nizovi, dok se dulji kodni nizovi dodjeljuju elementima koji imaju manju vjerojatnost pojavljivanja. H.264 norma koristi efikasnije algoritme entropijskog kodiranja, adaptivno entropijsko kodiranje. Koriste se različite kodne tablice, u ovisnosti o karakteristikama konteksta kodiranja.

U H.264 normi prisutne su dvije norme entropijskog kodiranja. Prva metoda predstavlja tehniku kodiranja promjenjivom dužinom niza, VLC (Variable Length Coding) koja koristi jednu tablicu definiranih kodova (exp-Golomb kodova) i služi za kodiranje svih elemenata sintakse osim kvantiziranih transformacijskih koeficijenata. Za kodiranje kvantiziranih transformacijskih koeficijenata koristi se

CAVLC (Context Adaptive Variable Length Coding) metodom, u kojoj dužina kodnih riječi ovisi o prethodno kodiranim elementima sintakse (koeficijentima).

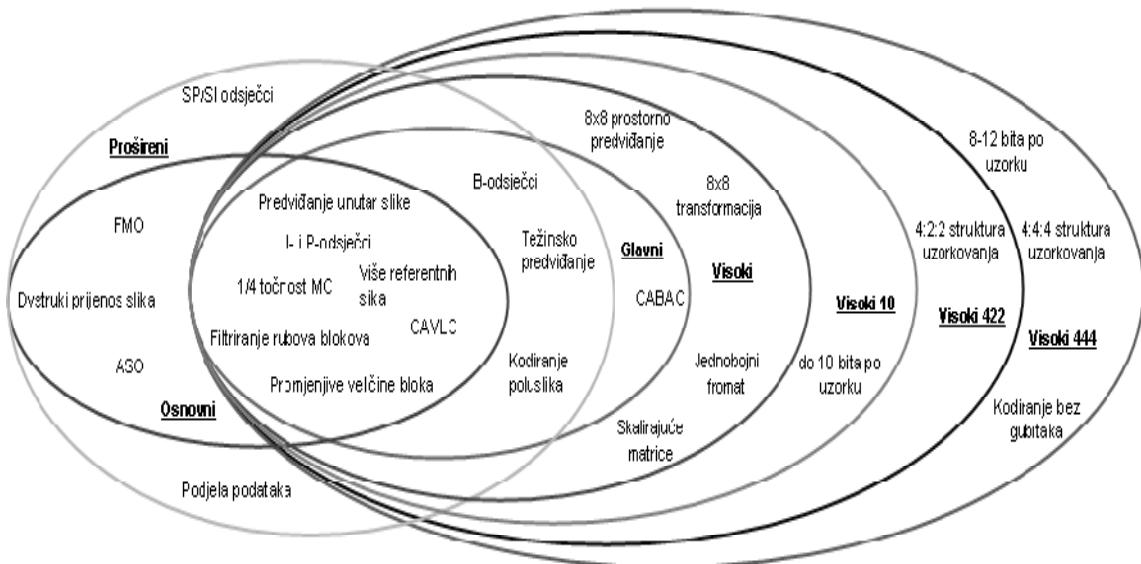
Drugi norma je CABAC (Context Adaptive Binary Arithmetic Coding) koja predstavlja kompleksnu tehniku binarnog aritmetičkog kodiranja niza ovisnog o kontekstu. CABAC je efikasnija metoda kodiranja jer ima veću redukciju podataka, ali je računalno kompleksnija i zahtjeva više vremena za kodiranje. Prednosti CABAC metode kodiranja mogu se vidjeti tek pri višim brzinama prijenosa.

## 4.11. Adaptivni rekonstrukcijski filter

Podjela slike na makroblokove, te proces kompresije koji se obavlja na razini blokova, nakon rekonstrukcije dovodi do vidljive distorzije u slici. Pri višim stupnjevima kompresije vidljiva je blokovska struktura slike jer su rubni elementi obično rekonstruirani s manjom točnošću nego unutarnji elementi blokova. Kako bi se smanjila vidljivost distorzije rekonstruirane slike, koristi se adaptivni rekonstrukcijski filter. Primjenjuje se u koderu nakon inverzne transformacije koju slijedi proces rekonstrukcije i pohrane za buduća predviđanja, te u dekoderu prije rekonstrukcije i prikaza slike. U koderu se na taj način dobiva bolja kompresija uz manje vidljivu distorziju. Ideja postupka filtriranja je pronaći, odrediti i ukloniti velike razlike između susjednih, rubnih dijelova blokova.

## 4.12. Razine u H.264 normi

H.264 norma dizajnirana je za širok raspon aplikacija, a različite aplikacije imaju različite zahtjeve na postupak kompresije. Kako bi se postigla veća adaptivnost H.264 norme, definirane su razne razine (profili) u H.264 normi. Profilom se određuje stupanj kompleksnosti sustava prikladan za određene aplikacije. U H.264 normi definirano je sedam profila



**Slika 4.10 Profil u H.264 normi**

Osnovni profil (Baseline) podržava sve osobine H.264 norme osim B slika, težinskog predviđanja, analiziranja s proredom, CABAC entropijskog kodiranja, SP/SI odsječke. Koristi se u videokonferencijskim uslugama i aplikacijama u mobilnim telefonima, koje zahtijevaju malo kašnjenje i brz postupak kodiranja.

Glavni profil (Main) koristi se u radiodifuziji i pohrani SDTV i HDTV signala.

Prošireni profil (Extended) ne podržava CABAC entropijsko kodiranje i sliku analiziranu s proredom. Koristi se za prijenos putem interneta i mobilne telefonije

Visoki profil (High) koristi se za radiodifuziju visokokvalitetnog televizijskog signala i za pohranu istog na rezličite medije (Blu Ray diskove). Pojavom visokog profila, glavni profil je izgubio na važnosti jer visoki profil omoguće sve što i glavni uz neke dodatke i efikasniji proces kompresije.

Visoki 10P profil (Hi10P) predstavlja nadogradnju na visoki profil. Podržava 4:2:0 format slike s 10 bita po uzorku

Visoki 422P profil (Hi422P) namijenjen je profesionalnim aplikacijama koje koriste analiziranje s proredom. Podržava 4:2:2 video format s 10 bita po uzorku.

Visoki 444P profil (Hi444P) podržava kompresiju slike kao tri odvojene ravnine boja, efikasnije kodiranje bez gubitaka i 4:4:4 format s ukupno 14 bita po uzorku.

Osim ovih profila, definirana su još četiri profila koji koriste samo kodiranje unutar slike, a primjenjuju se u profesionalnim sustavima za snimanje i obradu videosignalova (High 10 Intra Profile, High 4:2:2 Intra Profile, High 4:4:4 Intra Profile i CAVLC 4:4:4 Intra Profile).

Osim profila, u H.264 normi, definirane su i razine po formatima i primjenama.

Razina	Format	Frekvencija izmjene slike	Maksimalna brzina prijenosa	Maksimalni broj referentnih slika
1	QCIF	15	64 kbit/s	4
1b	QCIF	15	128 kbit/s	4
1.1	CIF; QCIF	7.5(CIF);30(QCIF)	192 kbit/s	2(CIF); 9(QCIF)
1.2	CIF	15	384 kbit/s	6
1.3	CIF	30	768 kbit/s	6
2	CIF	30	2 Mbit/s	6
2.1	SD (576i, 480i)	25; 30	4 Mbit/s	6
2.2	SD	15	4 Mbit/s	5
3	SD	25; 30	10 Mbit/s	5
3.1	HD (720p)	30	14 Mbit/s	5
3.2	HD (720p)	60	20 Mbit/s	4
4	HD (720p i 1080i)	60p; 30i	20 Mbit/s	4
4.1	HD (720p i 1080i)	60p; 30i	50 Mbit/s	4
4.2	HD (1080p)	60p	50 Mbit/s	4
5	2k x 1k	72	135 Mbit/s	5
5.1	2k x 1k; 4k x 2k	120; 30	240 Mbit/s	5

**Tablica 4.1** Razine u H.264 normi

## 5. HDTV-televizija visoke kvalitete

### 5.1. Uvod

Termin HDTV (High Definition Television, televizije visoke kvalitete) označava emitiranje televizijske slike veće kvalitete od standardne televizije (SDTV, Standard Definition Television). Pod pojmom veće kvalitete podrazumijeva se viša rezolucija, veći broj analizirajućih linija, progresivno analiziranje, veći omjer stranica slike, poboljšane kolorimetrijske značajke, redefiniran luminantni signal...

Danas su u svijetu prisutna tri sustava zemaljske digitalne televizije ATSC (Sjeverna i Južna Amerika), ISDB-T (Japan) i DVB-T (Europa) koji služe za prijenos SDTV i HDTV signala. Kako bi se olakšala međunarodna razmjena, prijenos i prilagodba televizijskog sadržaja između različitih formata, od strane SMPTE (Society for Motion Pictures and Television Engineers) i ITU-R (International Telecommunication Union – Radiocommunication) su propisane različite norme.

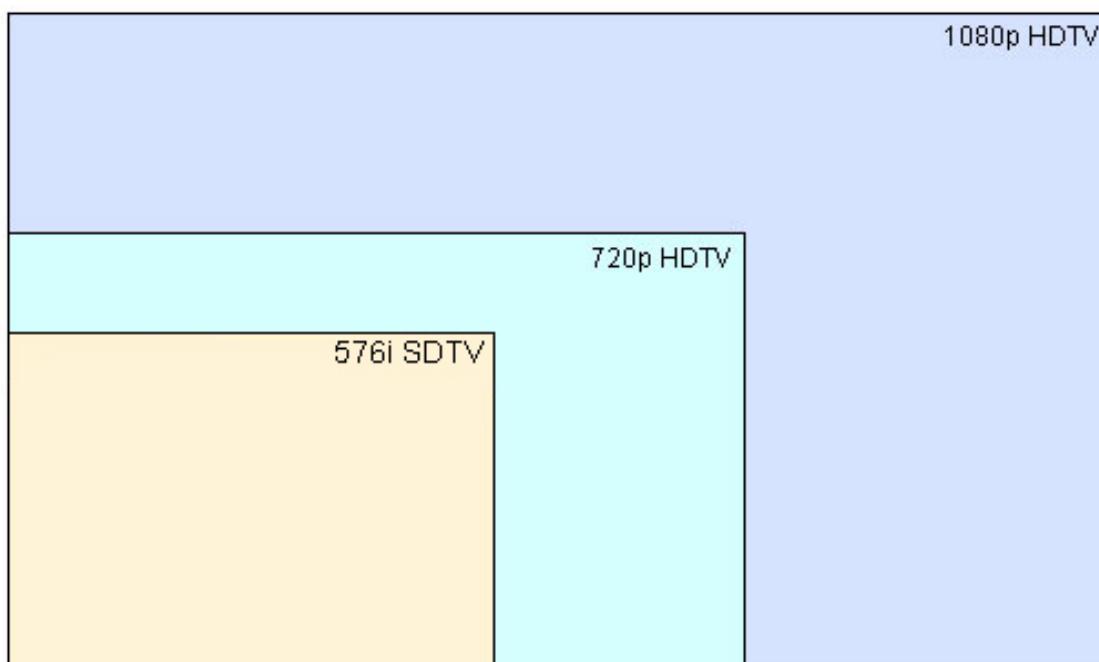
SMPTE organizacija je propisala dvije norme. Prva norma je SMPTE 296M (1280x720 Progressive Image Sample Structure – Analogue and Digital Representation and Analogue Interface), propisana 2001. godine za formate videosignalova od 720 linija od kojih svaka sadrži 1280 uzoraka. Druga norma je SMPTE 274M (1920x1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequences for Multiple Picture Rates), propisana 2003. godine za formate videosignalova od 1080 linija od kojih svaka linija sadrži 1920 uzoraka.

Od strane ITU-R propisana je jedna norma ITU-R BT.709 u kojoj su definirane vrijednosti parametara za produkciju i međunarodnu programsku razmjenu video signalova visoke kvalitete. U normi je opisan format od 1080 linija, od kojih svaka sadrži 1920 uzoraka.

Parametri sustava televizije visoke kvalitete opisani u preporuci ITU-R BT.709 jednaki su parametrima u normi SMPTE 274M.

## 5.2. Formati slike

U prethodnom poglavlju vidjeli smo da postoji nekoliko normi za analiziranje slike u sustavu televizije visoke kvalitete. One se mogu podijeliti u dvije glavne grupe. Prva grupa se odnosi na formate slike od 720 linija i 1280 uzoraka po svakoj liniji, dok se druga grupa odnosi na formate slike od 1080 linija sa po 1920 uzoraka po svakoj liniji. Na slici 5.1 prikazana je usporedba veličine slike kod televizije standardne kvalitete i televizije visoke kvalitete.



**Slika 5.1** Usporedba veličine slike kod SDTV i HDTV standarda

### 5.2.1.720p format slike televizije visoke kvalitete

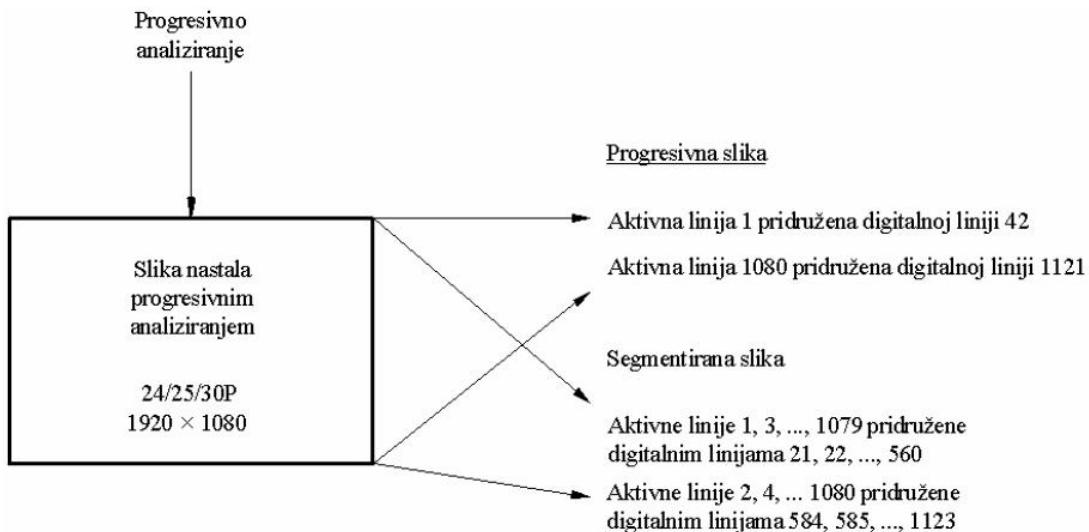
Naziv 720p koristi se za format slike od 720 linija i 1280 uzoraka po svakoj liniji, progresivno analiziranih. Ukupan broj elemenata slike iznosi 921 600. Okvir zapravo sadrži 750 horizontalnih linija. Slika je pohranjena u 720 od tih 750 linija (njih zovemo aktivne linije), dok su u preostalim linijama pohranjene nevidljive informacije (sincronizacija, teletekst...). Frekvencija uzorkovanja signala iznosi 74.25 MHZ, a tolerancija frekvencije izmjene slike iznosi  $\pm 10$  ppm. U tablici 5.1 prikazane su norme za analiziranje slike formata 720p propisane SMPTE 296M normom.

Oznaka	Aktivnih luminantnih uzoraka po liniji	Aktivnih linija po slici	Vertikalna frekvencija [Hz]	Frekvencija uzorkovanja luminantne komponente [MHz]	Ukupni broj luminantnih uzoraka po liniji	Ukupan broj linija po slici
1280x720/60	1280	720	60	74.25	1650	750
1280x720/59.94	1280	720	59.94	74.18	1650	750
1280x720/50	1280	720	50	74.25	1980	750
1280x720/30	1280	720	30	74.25	3300	750
1280x720/29.97	1280	720	29.97	74.18	3300	750
1280x720/25	1280	720	25	74.25	3960	750
1280x720/24	1280	720	24	74.25	4125	750
1280x720/23.98	1280	720	23.98	74.18	4125	750

**Tablica 5.1** Pregled normi za analiziranje slike formata 720p

### 5.2.2.1080p i 1080 i format slike televizije visoke kvalitete

Norme SMPTE 274M i ITU-R BT.709 propisuju parametre za analiziranje slike s 1080 linija koje sadrže svaka po 1920 uzoraka. Osim progresivnog analiziranja omogućeno je i analiziranje slike s proredom. Osim toga, u slučaju progresivnog analiziranja slike, omogućuju segmentaciju slike na poluslike, pri čemu poluslike nisu analizirane u različitim vremenskim trenucima što je slučaj kod analiziranja s proredom. U tom slučaju, kao i kod analiziranja s proredom, parne linije pripadaju jednoj poluslici a neparne drugoj. Segmentacija je moguća u samo četiri slučaja i to za vertikalne frekvencije od 24, 25, 29.97 i 30 Hz. U slučaju segmentacije, slike nose oznaku PsF (Progressive Segmented Frame, progresivno analizirane i segmentirane slike). Slika 5.2 prikazuje proces segmentacije progresivno analizirane slike.

**Slika 5.2 Segmentacija slike**

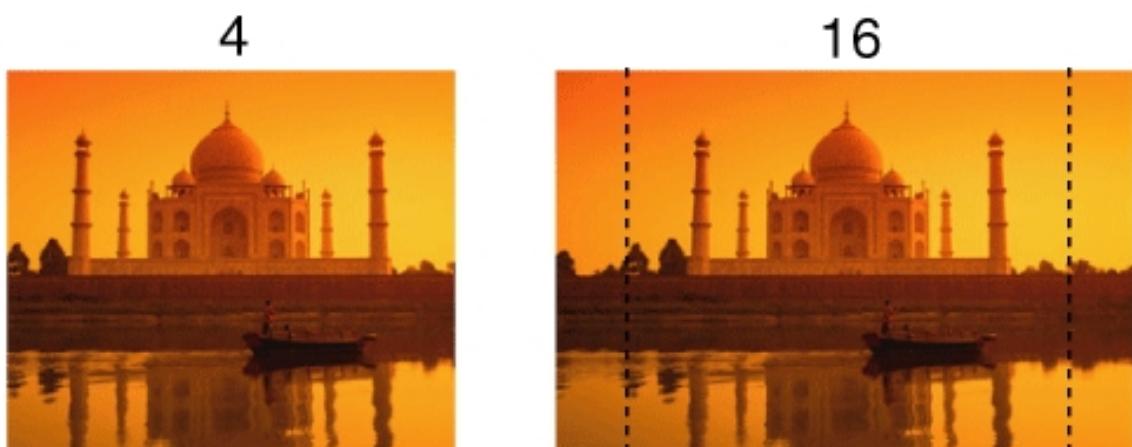
Oznaka	Aktivnih luminantnih uzoraka po liniji	Aktivnih linija po slici	Vertikalna frekvencija [Hz]	Frekvencija uzorkovanja luminantne komponente [MHz]	Ukupni broj luminantnih uzoraka po liniji	Ukupan broj linija po slici
1920/1080/60p	1920	1080	60	148.5	2200	1125
1920/1080/59.94p	1920	1080	59.94	148.35	2200	1125
1920/1080/50p	1920	1080	50	148.5	2640	1125
1920/1080/60i	1920	1080	30	74.25	2200	1125
1920/1080/30PsF						
1920/1080/59.94i	1920	1080	29.97	74.18	2200	1125
1920/1080/29.97PsF						
1920/1080/50i	1920	1080	25	74.25	2640	1125
1920/1080/25PsF						
1920/1080/30p	1920	1080	30	74.25	2200	1125
1920/1080/29.97p	1920	1080	29.97	74.18	2200	1125
1920/1080/25p	1920	1080	25	74.25	2640	1125
1920/1080/24p	1920	1080	24	74.25	2750	1125
1920/1080/24PsF						
1920/1080/23.98	1920	1080	23.98	74.18	2750	1125

**Tablica 5.2 Pregled normi za analiziranje slike formata 1080p**

### 5.3. Omjer stranica slike u sustavu televizije visoke kvalitete

Osim povećanja rezolucije slike, glavna vidljiva promjena u sustavu televizije visoke kvalitete odnosi se na omjer stranica slike. Omjer stranica (Aspect Ratio, AR) definiran je kao omjer širine i visine slike. U SDTV sustavu omjer stranica je bio 4:3 (1.33:1). Legenda kaže da je taj omjer stranica uzet jer je Edison nezadovoljan prevelikim omjerom stranica 70mm filma Georga Eastmana, rekao šegrtu da ga izreže na manje komade. Pri tome je prstima formirao kvadrat koji je imao omjer stranica 4:3. Tako je službeni omjer stranica postao 4:3, dok su filmski snimatelji koristili puno veći omjer stranica (1.66:1, 1.77:1, 2.35:1...) koji je odgovarao građi ljudskog vizualnog sustava (vidno polje je šire horizontalno nego vertikalno).

Kako bi se slika približila omjerima slike filmske vrpce u sustavu televizije visoke kvalitete koristi se omjer stranica 16:9 (1.77:1). Povećanje omjera stranica pridonosi 33% povećanju površine slike. Na slici 5.3 vidimo odnos omjera stranica u HDTV i SDTV sustavu.



Slika 5.3 Usporedba omjera stranica u SDTV i HDTV normi

## 5.4. Standardizacija kolorimetrije

Različiti standardi televizije standardne kvalitete nisu se razlikovali samo u broju linija i načinu analiziranja nego su im i kolorimetrijske značajke bile u potpunosti različite. Različitost parametara kolorimetrije pridonjela je težoj razmjeni videosadržaja između različitih sustava. Kako bi se u budućnosti izbjegla takva greška, donesena je ITU-R BT.1361 norma kojom se propisuju jedinstveni parametri kolorimetrije budućih televizijskih sustava. Tako su u HDTV-u u odnosu na SDTV unesene neke korekcije sustava kolorimetrije. Promjene se odnose na novu definiciju primara, a time i na novu definiciju luminantnih i krominantnih komponenti. U tablici 5.3 prikazani su propisani parametri kolorimetrije sustava televizije visoke kvalitete.

Parametar	Vrijednost	
	1125/60/2:1	1250/50/2:1
Optoelektroničke transformacijske karakteristike prije nelinearne predkorekcije	x	y
Kolorimetrijske kordinate (CIE, 1931) Primari: - Crvena (R) - Zelena (G) - Plava (B)	0.640 0.300 0.150	0.330 0.600 0.060
Kolorimetrijski parametri za jednake primarne signale $E_R = E_G = E_B$		D <sub>65</sub>
	x	y
	0.3127	0.3290
Vrijednost luminantnog signala $E'_Y$	$E'_Y = 0.2126 E'_R + 0.7152 E'_G + 0.0722 E'_B$	
Vrijednosti krominantnih signala $E'_{CB}$ i $E'_{CR}$	$E'_{CB} = (E'_B - E'_Y) / 1.8556$ $E'_{CR} = (E'_R - E'_Y) / 1.5748$	

**Tablica 5.3** Parametri kolorimetrije za HDTV sustav

## 6. Mjere kvalitete slike

### 6.1. Uvod

Uloga mjerjenja kvalitete slike je procjena kvalitete slike nakon određene računalne obrade. Pod računalnom obradom misli se na svaku obradu slike i dodatno, kompresiju slike jer ti procesi unose određenu distorziju u sliku i time joj smanjuju kvalitetu. Procjena kvalitete slike je vrlo važan proces jer se na temelju njega odabire prikladan medij za pohranu, prijenos ili reprodukciju iste.

Metode mjerjenja kvalitete slike dijele se u dvije grupe. Jednu grupu čine subjektivne procjene kvalitete. One su vremenski zahtjevne i složene, a sastoje se od ispitivanja velikog broja pojedinaca koji ocjenjuju kvalitetu slike na temelju vlastite procjene. Na procjenu svakog pojedinca utječe cijeli niz faktora, od odabira sekvenci, raspoloženja promatrača do uvjeta promatranja. Zbog toga su subjektivne metode skupe i rijetke.

Objektivne metode bazirane su na matematičkim metodama. Neusporedivo su brže, lakše i jeftinije za implementaciju. Svrha im je razvoj kvantitativnih metoda za procjenu doživljaja kvalitete slike. Mana im je nemogućnost ocjene doživljaja kvalitete slike jer su temeljene isključivo na matematičkim modelima, koji nisu u mogućnosti procijeniti subjektivan doživljaj slike i zato su često u raskoraku sa subjektivnim metodama mjerjenja.

### 6.2. PSNR

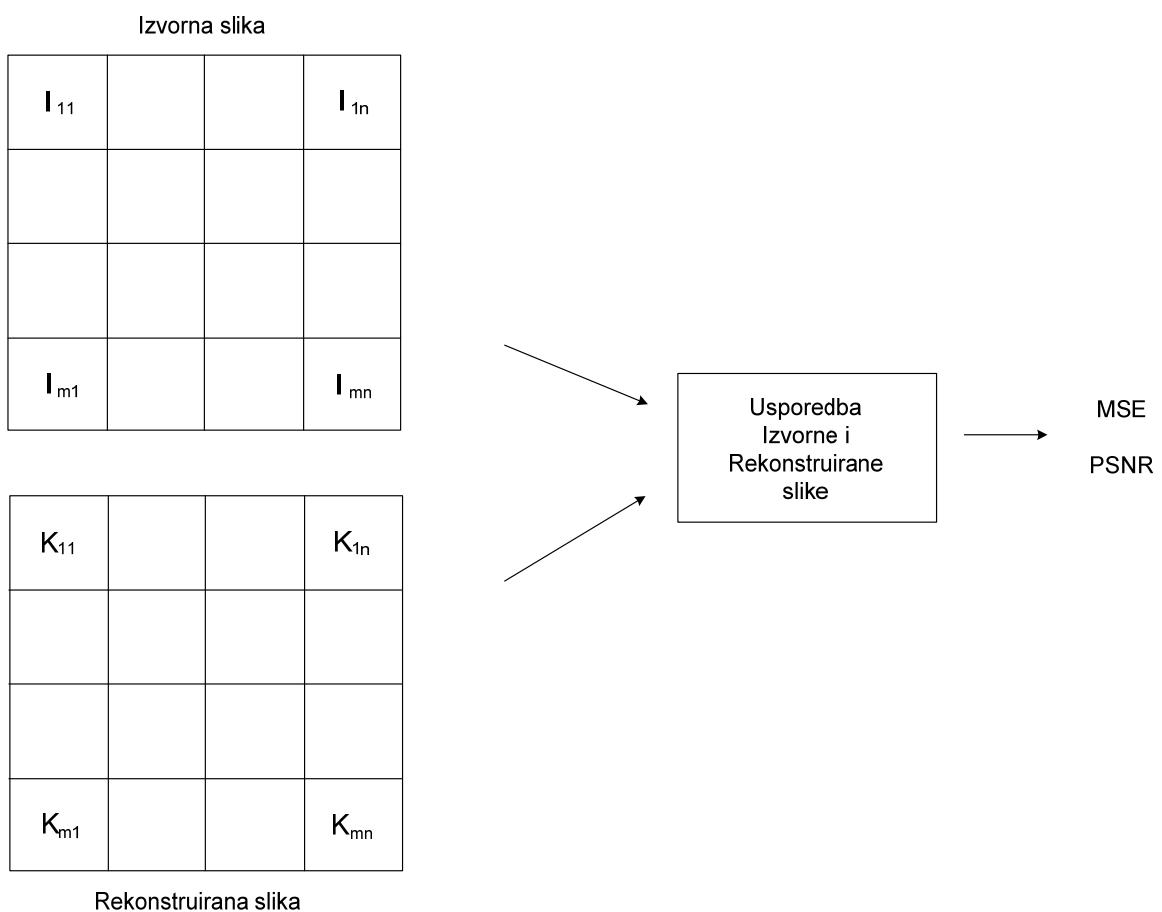
Vršna vrijednost odnosa signala i šuma (PSNR, Peak Signal-to-Noise Ratio) je objektivna mjera kvalitete rekonstruiranog signala. Daje nam omjer maksimalne snage signala i snage šuma. Kako većina signala ima vrlo veliki dinamički raspon (raspon između najveće i najmanje vrijednosti), PSNR izražava u logaritamskom omjeru.

PSNR se najčešće koristi kao objektivna mjera kvalitete rekonstruiranog signala nakon kompresije s gubicima (u našem slučaju kompresije videa). Kod kompresije videa ili slike, snaga signala je vrijednost pojedinog elementa slike u nekomprimiranoj slici, a snaga šuma predstavljena je srednjom kvadratnom

pogreškom između uzorka izvorne (nekomprimirane) i rekonstruirane (komprimirane) slike.

Videosignal se sastoji od niza slika od kojih se svaka može prikazati matricom čiji elementi predstavljaju vrijednost uzorka slike. U našem slučaju riječ je o monokromatskim slikama jer su predstavljene samo jednom matricom s vrijednostima luminancije na svakom elementu slike. U slučaju slika u boji, slike se rastavljaju na 3 matrice.

Slika 6.1 prikazuje dvije slike  $I$  i  $K$  veličine  $m \times n$ , od kojih je slika  $I$  izvorna slika a slika  $K$  je rekonstruirana slika nakon kompresije. Određivanjem razlike između elemenata izvorne i rekonstruirane slike moguće je procijeniti kvalitetu rekonstruirane slike a ujedno i kvalitetu načina kompresije.



**Slika 6.1** Usporedba elemenata rekonstruirane i izvorne slike

Do razlike dolazi pri kompresiji jer kompresija narušava kvalitetu slike na uštrb količine podataka. Veći stupanj kompresije odlikuje se većim narušavanjem kvalitete slike a s time i većom razlikom između slika.

PSNR se za dvije monokromatske slike ( $I$  i  $K$  dimenzija  $m \times n$ ) definira prema :

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

gdje je  $n$  broj bita korištenih pri uzorkovanju izvorne slike, dok MSE predstavlja srednju kvadratnu pogrešku između uzoraka izvorne (nekompimirane) slike  $I$  i rekonstruirane (kompimirane) slike  $K$  a definiran je relacijom:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|I(i,j) - K(i,j)\|^2$$

Broj  $(2^n - 1)$  predstavlja maksimalnu vrijednost koju element slike može postići a ona ovisi o broju bitova s kojima je postojeća slika uzorkovana. Za standardnih 8 bitova ta vrijednost iznosi 255.

Za slike u boji (koje sadrže tri vrijednosti za svaki element slike) PSNR je ekvivalentno definiran, ali se MSE računa po sve tri komponente ali se na kraju još dijeli s tri.

Tipične vrijednost PSNR-a nakon kompresije s gubicima kreću se između 30 i 50dB, gdje nam veća vrijednost označava bolju kvalitetu slike.

Pri usporedbi dvije identične slike, MSE će biti jednak nuli (jer nema razlika koje nam unosi kompresija) pa nam tad PSNR teži u beskonačno. Ako uspoređujemo dvije slike od kojih je jedna u potpunosti crna, PSNR nam ne iznosi nula. Da bi dobili vrijednost PSNR-a od 0dB moramo usporediti dvije slike od kojih će jedna biti u potpunosti bijela a druga crna.

Zbog jednostavne implementacije i relativno jednostavnog izračuna (velika brzina izračunjavanja), PSNR je vrlo popularna metoda objektivnog mjerjenja kvalitete slike. No, ona ne predstavlja idealnu ocjenu kvalitete budući dobiveni rezultati nisu u potpunosti korelirani s rezultatima subjektivnih ocjena kvalitete. Pri usporedbi događa se da rekonstruirana slika subjektivno bude ocijenjena kao bolja (ili bliža originalu, oku ugodnija) a da ima manju vrijednost PSNR-a od neke druge slike koja je subjektivnim mjerama ocijenjena kao lošija. Zato metoda

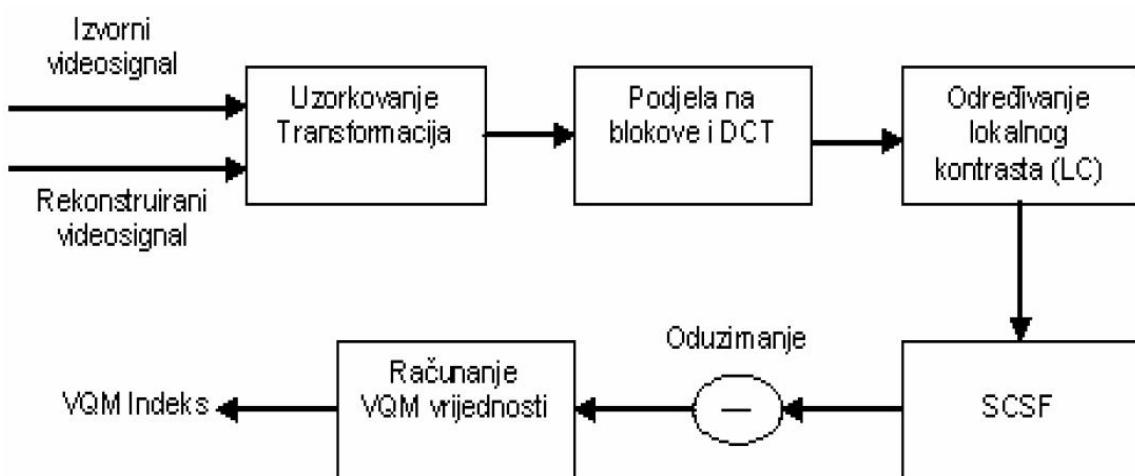
usporedbe slika PSNR-om ima smisla samo pri usporedbi identičnih slika pri različitim stupnjevima kompresije, dok pri usporedbi različitih slika metoda pada u vodu.

### 6.3. VQM

VQM (Video Quality Metric) je objektivna metoda ocjene kvalitete slike koja ima veću korelaciju sa subjektivnim metodama ocjene kvalitete slike. Pri ocjeni kvalitete slike, kao i kod PSNR-a, potrebna je izvorna slika.

VQM metoda je modificirana verzija DVQ (Digital Video Quality) metode temeljene na jednostavnom i brzom algoritmu koji oponaša karakteristike osjetljivosti ljudskog vizualnog sustava. Ljudski vizualni sustav osjetljiviji je pri manjim vremenskim i prostornim promjenama (frekvencijama), dok s brzinom porasta promjena gubi na osjetljivosti.

VQM uzima originalnu i komprimiranu sekvencu i nad njima provodi mjeru koja se temelji na DCT transformaciji. Metoda simulira proces obrade vizualne informacije ljudskog vizualnog sustava koji se sastoji od adaptacije na osvijetljenost, separaciju krominantnih i luminantnih kanala, vremenskog i prostornog filtriranja, određivanje prostornih frekvencija, maskiranja kontrastom...



**Slika 6.2** Blok shema VQM metode

transformacija u igru ulazi pri filtriranju prostornih frekvencija.

Ocjena kvalitete slike provodi se usporedbom izvornog (nekompresionog) i rekonstruiranog (kompresionog) videosignala u 5 koraka.

1. uzorkovanje, transformacija boje, deinterlacing, reverzna gama korekcija
2. podjela sekvenci na makroblokove nad kojima se provodi DCT transformacija
3. koeficijenti DCT transformacije iz prethodnog koraka transformiraju se u koeficijente lokalnog kontrasta koji predstavljaju omjer DCT koeficijenata i DC koeficijenta određenog makrobloka
4. simulacija funkcije osjetljivosti ljudskog vizualnog sustava na kontrast primjenom dviju SCSF (Spatial Contrast Sensitivity Function) matrica, jedne za statične a druge za dinamične sekvence
5. oduzimaju se vrijednosti koeficijenata izvorne i rekonstruirane (komprimirane) slike te se računaju srednje i maksimalne vrijednosti njihovih razlika. Zbroj tih dviju vrijednosti predstavlja VQM indeks kvalitete slike

Rezultat usporedbe je VQM indeks pri čemu vrijednost nula predstavlja identičnost izvorne i rekonstruirane sekvence, dok porast VQM indeksa označava sve veću i veću degradaciju rekonstruirane slike.

## 6.4. SSIM

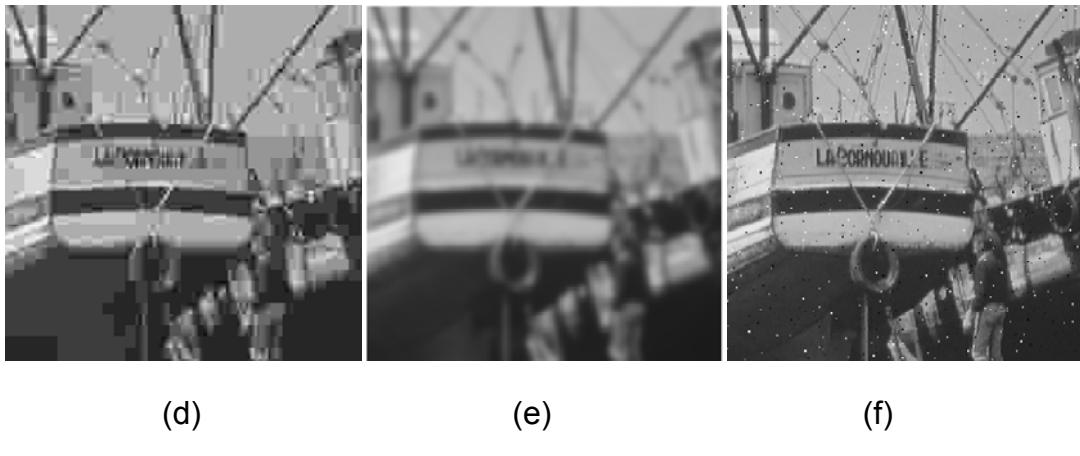
Metoda strukturalne sličnosti (SSIM, Structural Similarity Method) je objektivna metoda ocjene kvalitete slike koja određuje stupanj izobličenja strukture rekonstruirane (komprimirane) slike. Prirodne slike su strogo strukturirane, tj. elementi slike su strogo povezani iz njihove konfiguracije moguće je rekonstruirati objekt u slici (oni nose važnu informaciju o strukturi objekta). Ljudski vizualni sustav je prilagođen ekstrakciji strukturalnih informacija iz svog vidnog polja. Pretpostavka je da bi ocjena kvalitete bazirana na strukturalnim promjenama bila bliža subjektivnim ocjenama. Prethodne objektivne metode ocjene kvalitete slike daju ocjenu kvalitete neovisno o strukturi slike i zato su slabije korelirane sa subjektivnim metodama.



(a)

(b)

(c)



(d)

(e)

(f)

**Slika 6.3** Usporedba slika Brod s različitim distorzijama uz jednaki MSE=210  
 (a) originalna slika (b) podizanje kontrasta SSIM = 0.9168 (c) mean-shift SSIM = 0.9900 (d) JPEG kompresija SSIM = 0.6949 (e) zamućenje SSIM = 0.7052 (f) sol-papar šum SSIM = 0.7748

Slika 5.3 prikazuje sliku Brod komprimiranu na pet različitih načina uz uvjet da joj MSE iznosi oko 210. Vizualno su te slike poprilično različite, ali njihov PSNR i VQM odudaraju od subjektivnih ocjena kvalitete. Dok je PSNR za svaku sliku jednak i time u potpunosti odudara od subjektivnih procjena, VQM još donekle prati subjektivne ocjene. Na ovom primjeru vidimo da nestrukturalni pristup ocjeni kvalitete slike odudara od vizualne percepције istih.

SSIM metoda usporedbe kvalitete slike dobije se iz mjernog sustava na slici.

Na ulaz sustava dovedene su dvije slike. Ocjena kvalitete ovisi o degradacijama strukture slike, luminancije i kontrasta. Zato je ocjena kvalitete podijeljenja u tri dijela.

Prvo se uspoređuju luminancije signala i dobijemo vrijednosti  $\mu_x$  i  $\mu_y$  koje odgovaraju srednjoj vrijednosti luminantnih signala

$$\mu_x = \frac{1}{n} \sum_{i=0}^{n-1} x_i$$

Usporedba luminancije je funkcija od  $\mu_x$  i  $\mu_y$

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{{\mu_x}^2 + {\mu_y}^2 + C_1}$$

$$C_1 = (K_1 L)^2$$

$$L = 2^n - 1$$

$$K_1 \ll 1$$

gdje je  $L$  dinamički raspon vrijednost (za 8 bitnu monokromatsku sliku iznosi 255) a  $K_1$  je konstanta.

Tada se iz signala oduzme srednja vrijednost luminancije i dobijemo signal  $x - \mu_x$ . Na postojećem signalu izračuna se standardna devijacija i dobiju se signali  $\sigma_x$  i  $\sigma_y$ .

$$\sigma_x = \left( \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu_x)^2 \right)^{\frac{1}{2}}$$

Usporedba kontrasta je upravo usporedba signala  $\sigma_x$  i  $\sigma_y$ .

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{{\sigma_x}^2 + {\sigma_y}^2 + C_2}$$

$$C_2 = (K_2 L)^2$$

$$L = 2^n - 1$$

$$K_2 \ll 1$$

Na kraju se signal normalizira na standardnu devijaciju, tako da signali pri strukturalnom uspoređivanju imaju jednaku standardnu devijaciju. Tako uspoređeni signali nam daju strukturalnu usporedbu.

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$$

$$\sigma_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y)$$

$$C_3 = \frac{C_2}{2}$$

Na taj način dobiveni signali kombiniraju se u jednu mjeru.

$$SSIM = l(x, y)c(x, y)s(x, y)$$

Važno je primijetiti da su ove tri komponente u potpunosti nezavisne i da promjena jedne od njih ne utječe na promjenu mjere preostalih komponenti.

## 7. Mjerenja

### 7.1. Uvod

Promjenom parametara slike u televizijskim sustavima i prelaskom na sustave televizije visoke kvalitete, osim poboljšanja kvalitete slike, porasli su i zahtjevi na prijenos i pohranu videosignalna. Postupak MPEG-2 kompresije videosignalna se u tim novim uvjetima pokazao kao nedovoljno efikasan, a H.264 postupak se pokazao kao puno efikasniji te se počeo primjenjivati u kompresiji videosignalna visoke kvalitete.

Utjecaj H.264 postupka kompresije videosignalna ispitali smo na dvije videosekvence dostupne u sva tri formata slike visoke kvalitete (720p, 1080i i 1080p).

### 7.2. Videosekvence

Pri usporedbi analizirajućih normi slika namijenjenih sustavima televizije visoke kvalitete kodiranih H.264 postupkom korištene su video sekvene iz višeformatnog niza visoko kvalitetnih sekvenci švedske televizije. Testne sekvene dobivene su digitaliziranjem visokokvalitetnog izvornog materijala snimanog profesionalnom filmskom kamerom na 65mm film. Izvorni materijal konvertiran je digitalni oblik skeniranjem negativa veličine 50.94x22.34mm, upotrijebljenog za dobivanje mastera veličine 3840x2160 elemenata slike, brzinom analiziranja od 50 slika u sekundi. Elementi slike kodirani su s 16 bita po uzorku. Korekcijska vrijednost boje izvršena je prema ITU-R BT.709 normi. Digitalni master materijal rezolucije 3840x2160 elemenata slike, konvertiran je u tri različita formata (720p/50, 1080i/25 i 1080p/50), odnosno u tri različite norme za analiziranje slike.

Pri testiranju H.264 postupka korištene su dvije videosekvence s potpuno različitim zahtjevima na proces kompresije. Video sekvenca Crowdrun je vrlo zahtjevna videosekvencia za proces kompresije. U sekvenci vidimo mnoštvo ljudi kako trči preko jedne livade, a u pozadini vidimo šumu. Kamera se cijelo vrijeme lagano kreće u smjeru ljudi koji trče. U sceni imamo jako puno objekata koji se kreću (ljudi koji trče) i jako puno različitih boja (svi ljudi su drugačije odjeveni), te je to razlog za težinu postupka kompresije. Na slici 7.1 vidimo prvu sliku videosekvence Crowdrun.



**Slika 7.1** Prva slika videosekvence Crowdrun

Videosekvenca Intotree je u potpunosti različita. U njoj vidimo kuću u širokom planu i kamera lagano kreće prema njoj, dok ne dođe vrlo blizu, a onda skrene u desno prema šumi. Za razliku od prethodne sekvence, sekvanca Intotree je u puno manje zahtjevna i kao takva ne predstavlja neki problem H.264 postupku kompresije videosignalata. Na slici 7.2 vidimo prvu sliku videsekvence Intotree.



**Slika 7.2** Prva slika videosekvence Intotree

Sekvence su izabrane s namjerom da pretzentiraju dvije potpuno različite realne programske sadržaje. Kvaliteta dekodiranog i rekonstruiranog video signala, ovisi o sadržaju. Sekvence s puno kretanja i puno detalja se teže komprimiraju. Uz istu količinu podataka kao i sekvene s puno manje detalja i kretanja, sekvene s više kretanja i detalja će biti manje kvalitetne nakon rekonstrukcije.

Za sve realne programske sadržaje može se odrediti stupanj kritičnosti kodiranja. Procjena kritičnosti kodiranja korištenih ispitnih sekvenci određena je usporedbom s preko dvije stotine referentnih sekvenci različitog sadržaja (sportske, filmske, informativne...). Pri uspoređivanju kritičnosti uspoređuje se brzina podataka ispitnih i referentnih sekvenci, uz iste parametre kompresije. Kritičnost od 50% značida je 50% referentnih sekvenci kodirano s manjom brzinom podataka od ispitne sekvence uz istu kvalitetu nakon rekonstrukcije. Na osnovu ove klasifikacije, ispitna sekvenca Crowdrun ima kritičnost kodiranja od čak 95.3%, dok ispitna sekvenca Intotree od 28%.

### 7.3. Postupak mjerenja kvalitete slike i rezultati

Mjerenje kvalitete kodiranih videosekvenci izvedeno je pomoću programa Elecard Video Quality Estimator koji vrši usporedbu izvorne nekomprimirane videosekvence i komprimirane videosekvence. Kako bi se omogućila usporedba videosekvenci, obje moraju imati 4:2:0 strukturu uzorkovanja krominantnih komponenti (YV12 struktura). Program omogućuje usporedbu i davanje objektivnih ocjena kvalitete slike, od kojih su nam najbitnije ocjene izražene u obliku PSNR, VQM i SSIM veličina.

Kao referentne sekvene poslužile su izvorne nekomprimirane videosekvence.

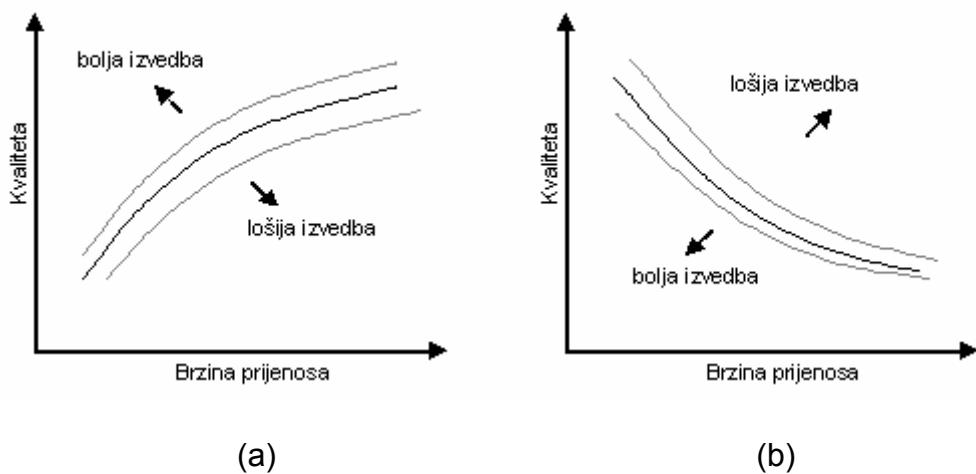
Ispitivanje je provedeno za svaku brzinu prijenosa, u oba profila H.264 postupka kompresije podataka, što jer ukupno dalo 72 mjerenja.

Uz objektivne ocjene kvalitete, u obzir su uzete i subjektivne ocjene kvalitete koje su proveli Hans Hoffmann, *Member, IEEE*, Takebumi Itagaki, *Member, IEEE*, David Wood, and Alois Bock 2006.godine.

Za ocjenu kvalitete videosekvenci korišteni su PSNR, VQM i SSIM postupak. Za svaku pojedinu kombinaciju parametara kompresije, rezultati su predočeni u obliku

dijagrama koji predstavljaju funkciju ovisnosti objektivne mjere kvalitete slike o određenoj brzini prijenosa (8 Mbit/s, 10 Mbit/s, 12 Mbit/s, 14 Mbit/s, 16 Mbit/s i 18 Mbit/s).

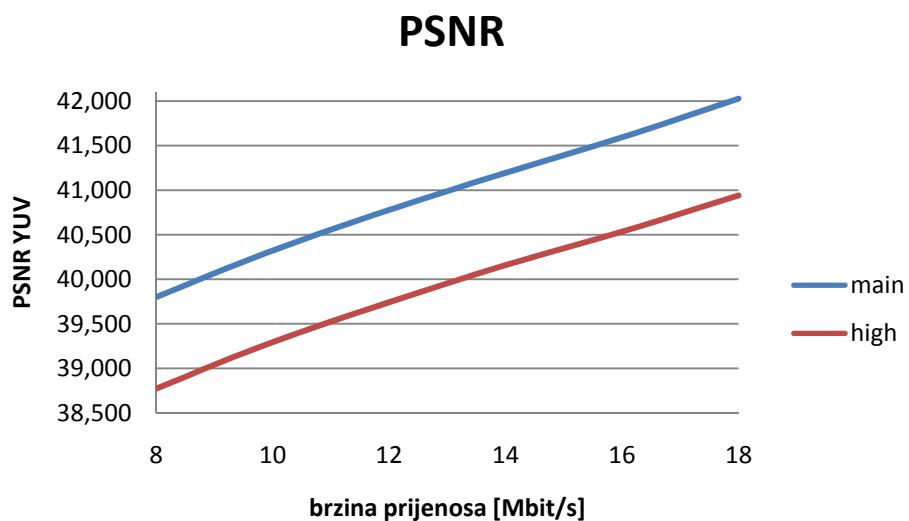
S porastom brzine prijenosa krivulja PSNR i SSIM vrijednosti ima pozitivnu promjenu, tj. povećanju brzine prijenosa odgovara porast PSNR i SSIM vrijednosti, što ujedno odgovara i porastu kvalitete rekonstruiranih (a prethodno kodiranih) videosekvenci. VQM vrijednost s porastom brzine prijenosa opada, tj. veće VQM vrijednosti odgovaraju slabijoj kvaliteti slike, tj. manjoj sličnosti s izvornom sekvencom.



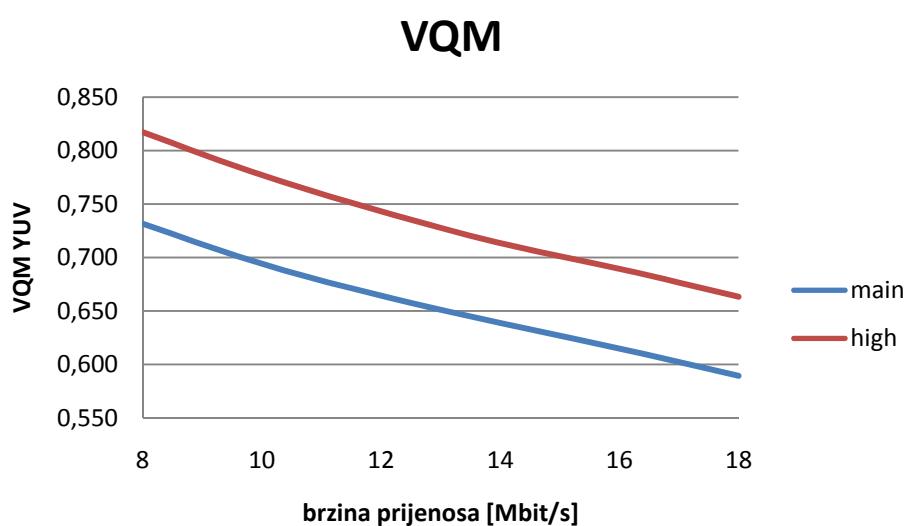
**Slika 7.3** Nagib funkcije ovisnosti kvalitete slike o brzini prijenosa (a) PSNR i SSIM veličine (b) VQM veličine

### 7.3.1. Ispitna sekvenca Intotree 720p/50

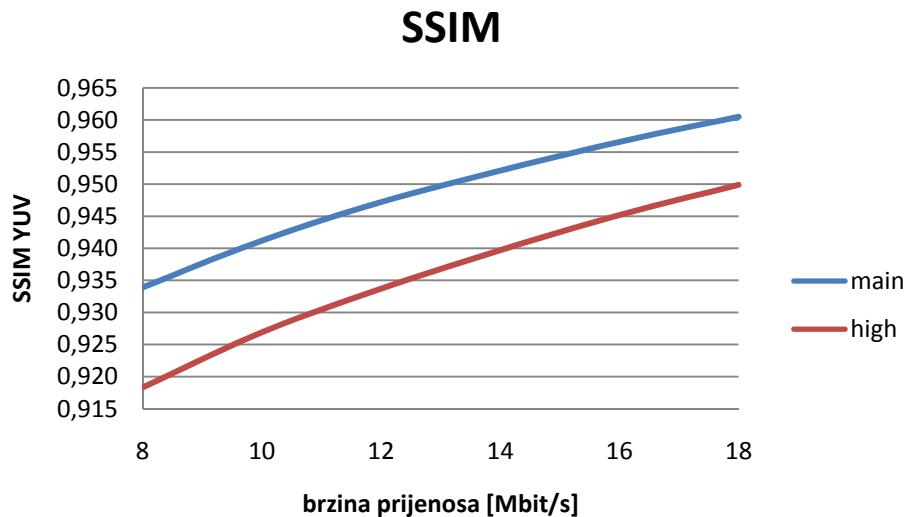
Rezultati objektivne ocjene kvalitete za videosekvencu Inotree formata 720p/50 prikazani su na slikama. Na slici 7.4 prikazana je ovisnost odnosa signal-šum (PSNR) o brzini prijenosa. Slika 7.5 prikazuje ovisnost VQM koeficijenta o brzini prijenosa, a slika 7.6 ovisnost SSIM koeficijenta o brzini prijenosa. Na oba dijagrama prikazani su rezultati postupka kompresije u glavnom i visokom profilu H.264 kodera.



**Slika 7.4** Ovisnost PSNR koeficijenta o brzini prijenosa



**Slika 7.5** Ovisnost VQM koeficijenta o brzini prijenosa

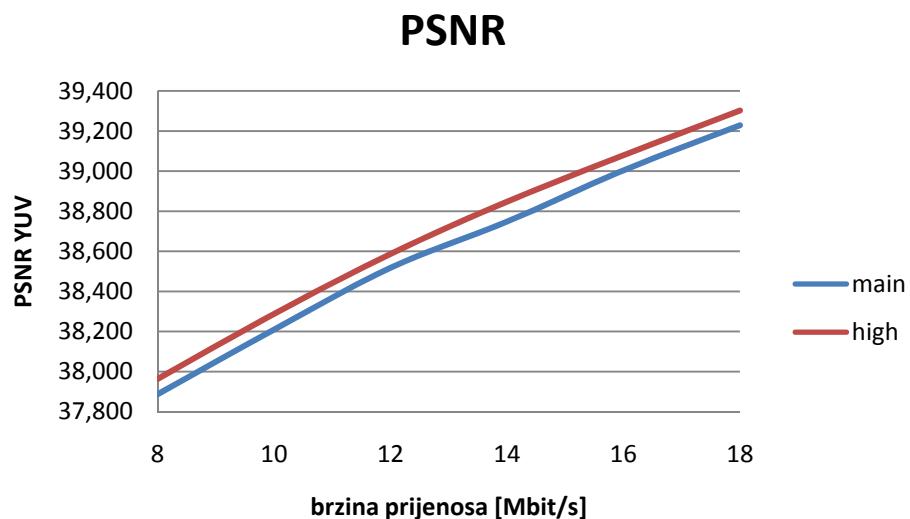


**Slika 7.6** Ovisnost SSIM koeficijenta o brzini prijenosa

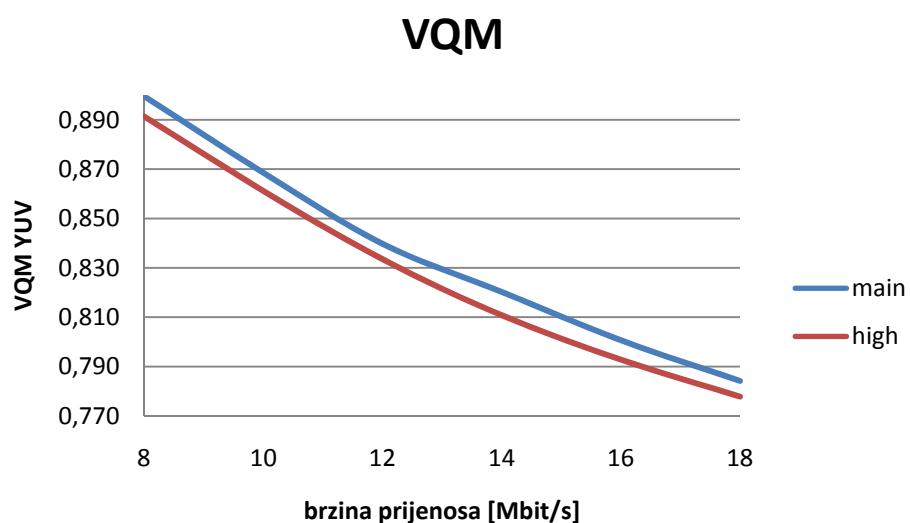
720p format slike daje nešto bolje rezultate pri kompresiji u glavnom profilu. Iako su ovo objektivna mjerena kvalitete, razlike se povećanjem brzine prijenosa ne mijenjaju previše.

### 7.3.2. Ispitna sekvenca Intotree 1080i/25

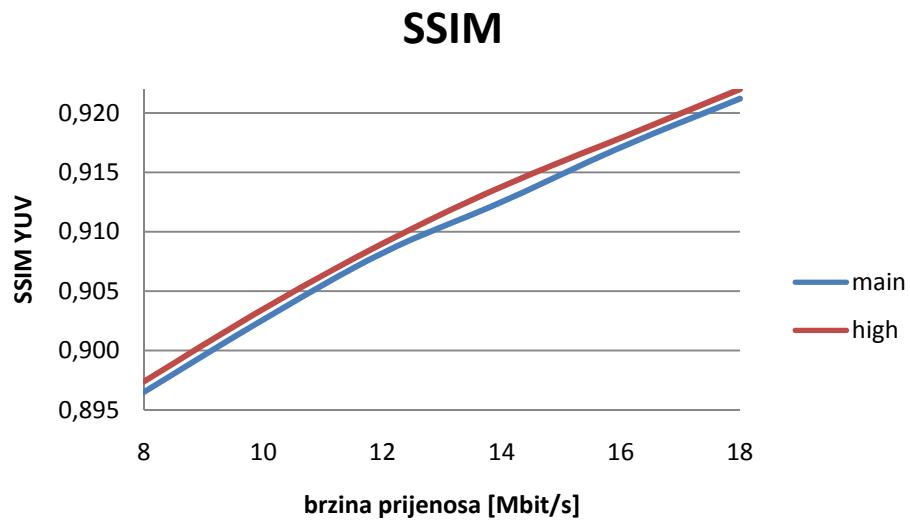
Rezultati objektivne ocjene kvalitete za videosekvencu Intotree formata 1080i/25 prikazani su na slikama. Na slici 7.7 prikazana je ovisnost odnosa signal-šum o brzini prijenosa. Slika 7.8 prikazuje ovisnost VQM koeficijenta o brzini prijenosa, a slika 7.9 ovisnost SSIM koeficijenta o brzini prijenosa. Na oba dijagrama prikazani su rezultati postupka kompresije u glavnom i visokom profilu H.264 kodera.



Slika 7.7 Ovisnost PSNR koeficijenta o brzini prijenosa



Slika 7.8 Ovisnost VQM koeficijenta o brzini prijenosa

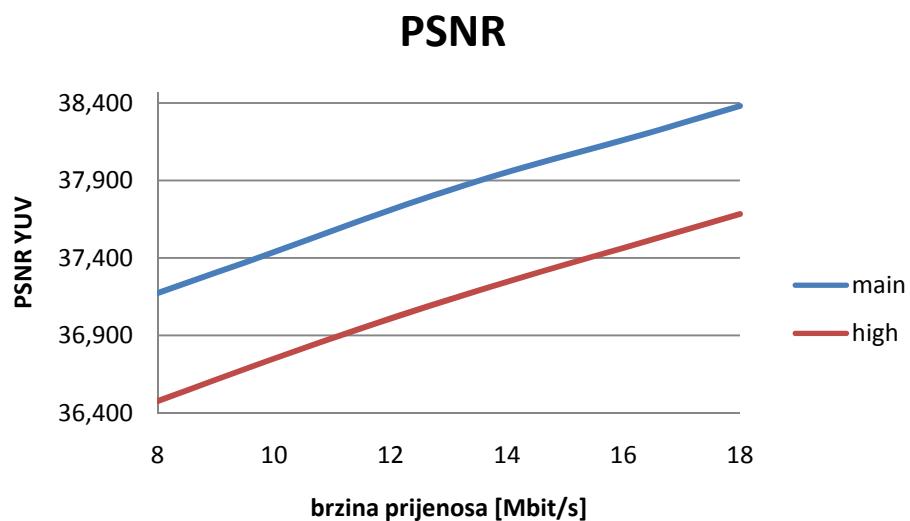


**Slika 7.9** Ovisnost SSIM koeficijenta o brzini prijenosa

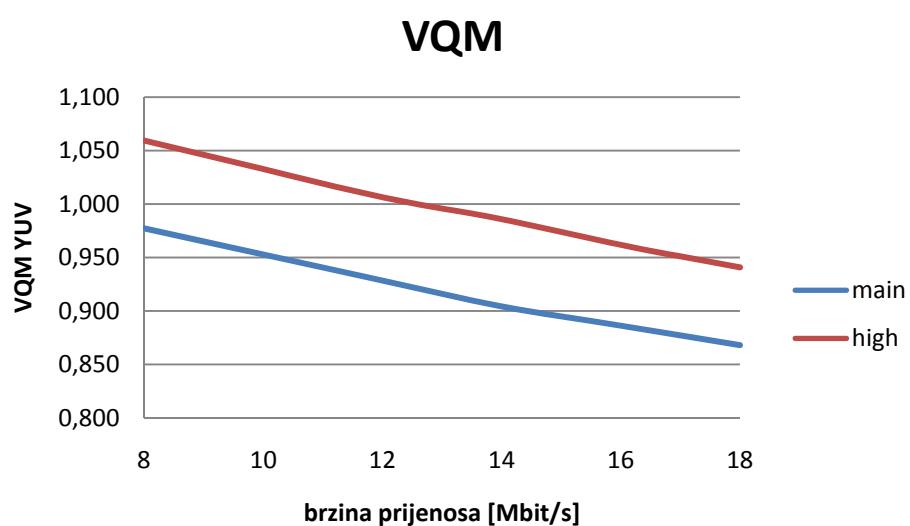
Za razliku od 720p formata slike koji je pokazao višu kvalitetu pri kompresiji u glavnom profilu, 1080i format koristi sve prednosti visokog profila i pokazuje nešto višu kvalitetu pri kompresiji u visokom profilu.

### 7.3.3. Ispitna sekvenca Intotree 1080p/50

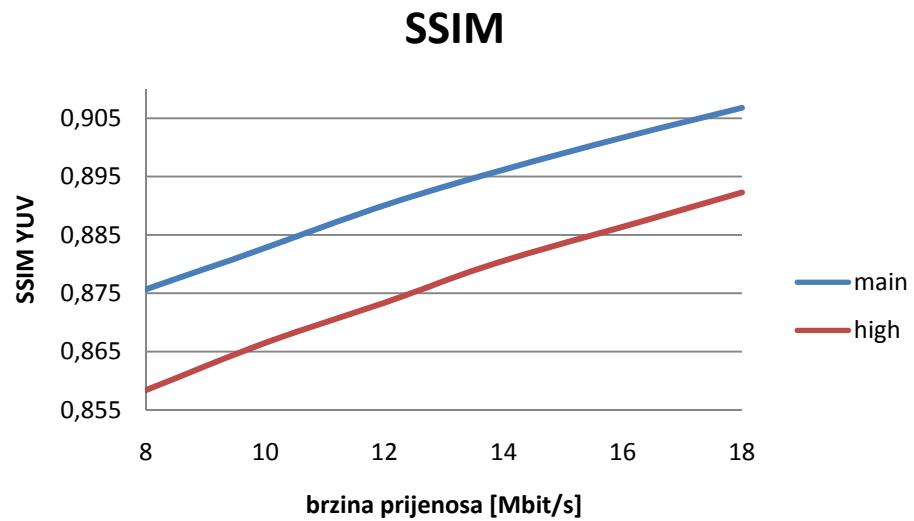
Rezultati objektivne ocjene kvalitete za videosekvencu Intotree formata 1080p/50 prikazani su na slikama. Na slici 7.10 prikazana je ovisnost odnosa signal-šum (PSNR) o brzini prijenosa. Slika 7.11 prikazuje ovisnost VQM koeficijenta o brzini prijenosa, a slika 7.12 ovisnost SSIM koeficijenta o brzini prijenosa. Na oba dijagrama prikazani su rezultati postupka kompresije u glavnom i visokom profilu H.264 kodera.



**Slika 7.10** Ovisnost PSNR koeficijenta o brzini prijenosa



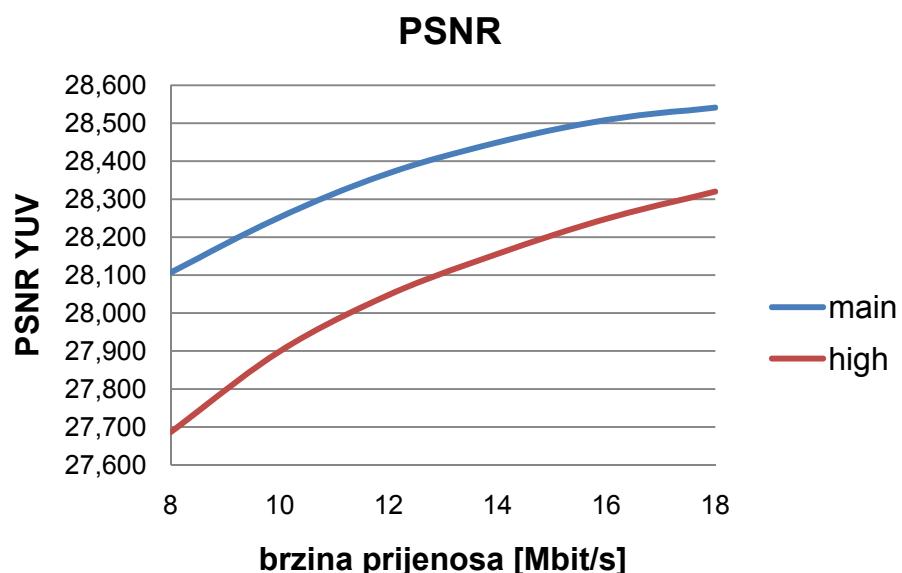
**Slika 7.11** Ovisnost VQM koeficijenta o brzini prijenosa



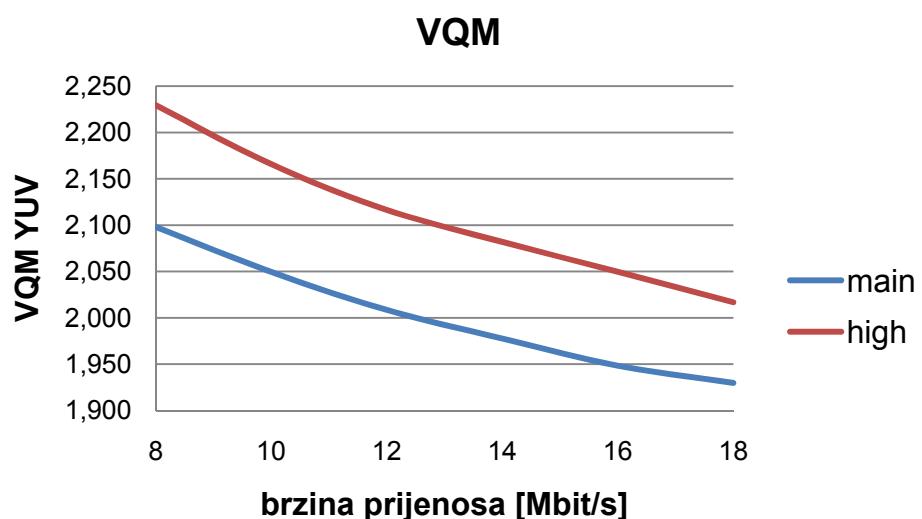
**Slika 7.12** Ovisnost SSIM koeficijenta o brzini prijenosa

### 7.3.4. Ispitna sekvenca Crowdrun, 720p/50

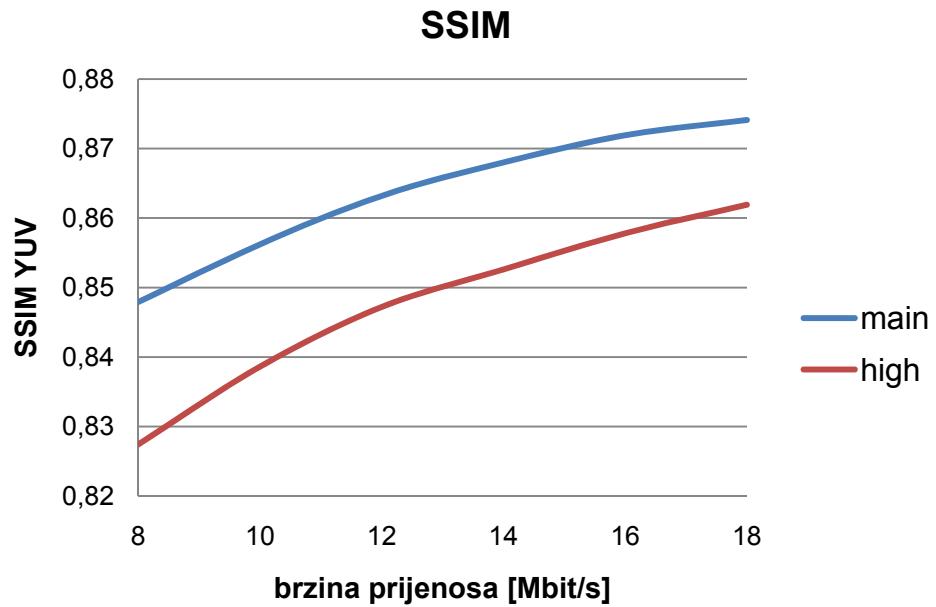
Rezultati objektivne ocjene kvalitete za videosekvencu Crowdrun formata 720p/50 prikazani su na slikama. Na slici 7.13 prikazana je ovisnost odnosa signal-šum (PSNR) o brzini prijenosa. Slika 7.14 prikazuje ovisnost VQM koeficijenta o brzini prijenosa, a slika 7.15 ovisnost SSIM koeficijenta o brzini prijenosa. Na oba dijagrama prikazani su rezultati postupka kompresije u glavnom i visokom profilu H.264 kodera.



Slika 7.13 Ovisnost PSNR koeficijenta o brzini prijenosa



Slika 7.14 Ovisnost VQM koeficijenta o brzini prijenosa

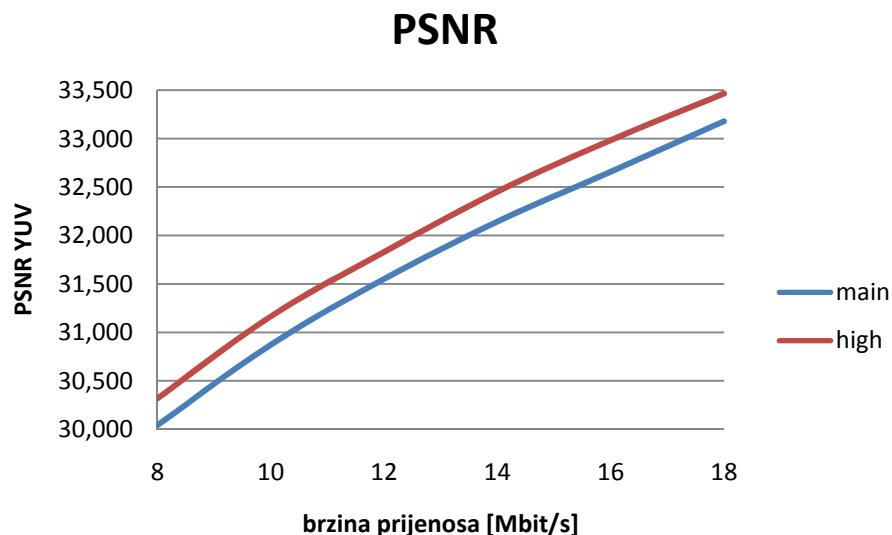


**Slika 7.15** Ovisnost SSIM koeficijenta o brzini prijenosa

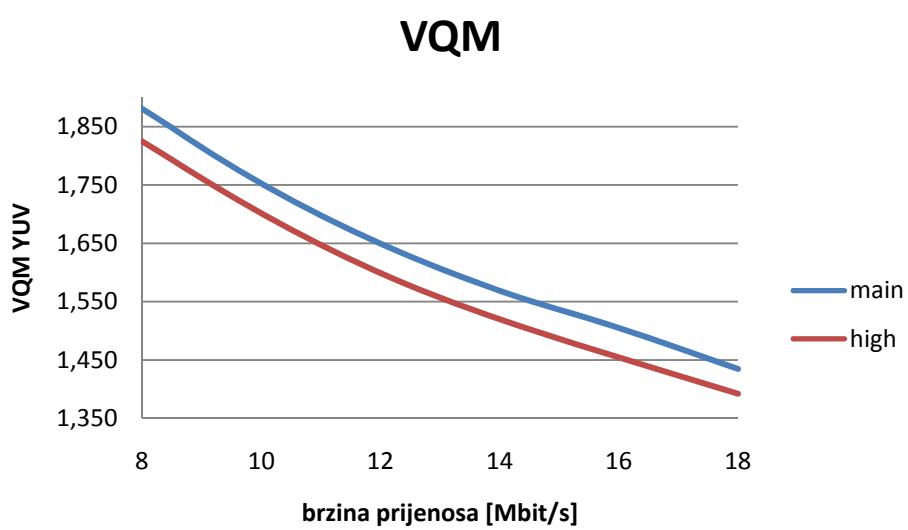
Iako visoki profil donosi puno prednosti kod kompresije videosekvenci, sva tri formata slike su se kod objektivnih ocjena pokazala boljima uz kompresiju u glavnom profilu.

### 7.3.5. Ispitna sekvenca Crowdrun 1080i/25

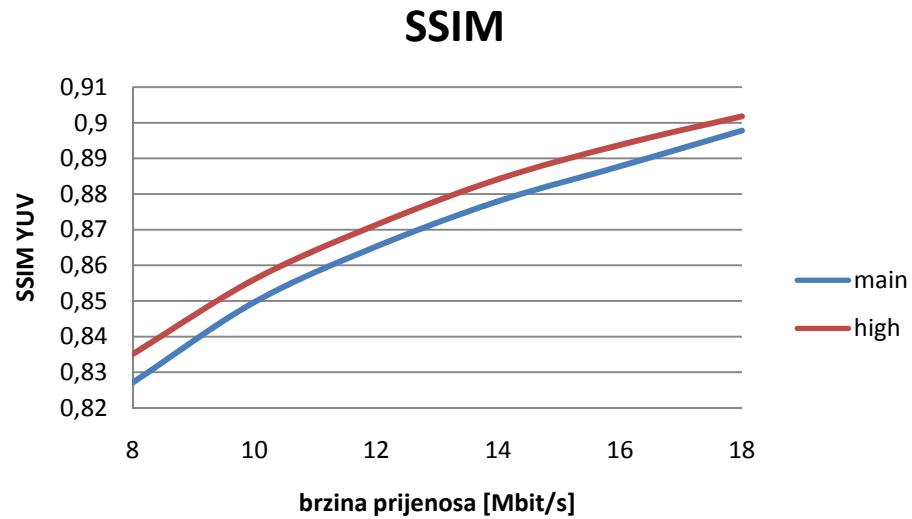
Rezultati objektivne ocjene kvalitete za videosekvencu Crowdrun formata 1080i/25 prikazani su na slikama. Na slici 7.16 prikazana je ovisnost odnosa signal-šum (PSNR) o brzini prijenosa. Slika 7.17 prikazuje ovisnost VQM koeficijenta o brzini prijenosa, a slika 7.18 ovisnost SSIM koeficijenta o brzini prijenosa. Na oba dijagrama prikazani su rezultati postupka kompresije u glavnom i visokom profilu H.264 kodera.



**Slika 7.16** Ovisnost PSNR koeficijenta o brzini prijenosa



**Slika 7.17** Ovisnost VQM koeficijenta o brzini prijenosa

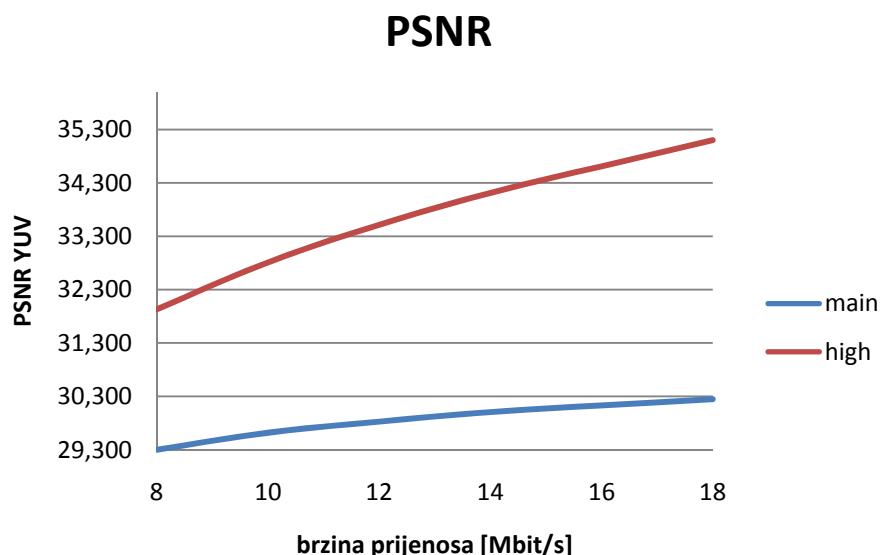


**Slika 7.18** Ovisnost SSIM koeficijenta o brzini prijenosa

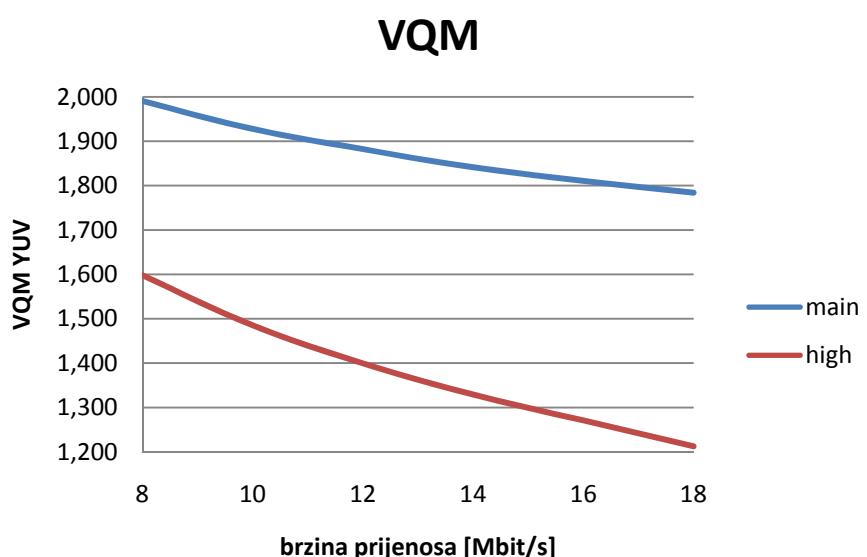
Za razliku od 720p formata slike koji je pokazao višu kvalitetu pri kompresiji u glavnom profilu, 1080i format koristi sve prednosti visokog profila i pokazuje nešto višu kvalitetu pri kompresiji u visokom profilu.

### 7.3.6. Ispitna sekvenca Crowdrun 1080p/50

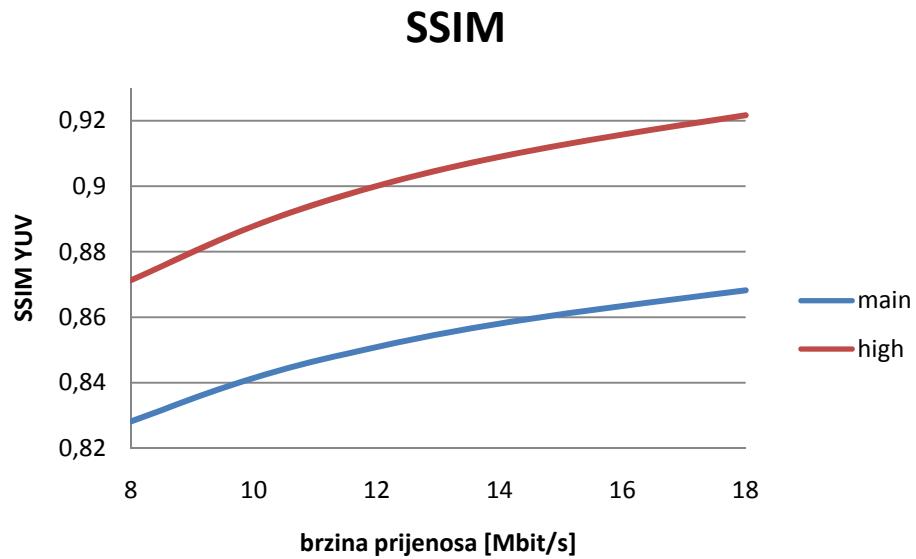
Rezultati objektivne ocjene kvalitete za videosekvencu Crowdrun formata 1080p/50 prikazani su na slikama. Na slici 7.19 prikazana je ovisnost odnosa signal-šum (PSNR) o brzini prijenosa. Slika 7.20 prikazuje ovisnost VQM koeficijenta o brzini prijenosa, a slika 7.21 ovisnost SSIM koeficijenta o brzini prijenosa. Na oba dijagrama prikazani su rezultati postupka kompresije u glavnom i visokom profilu H.264 kodera.



**Slika 7.19** Ovisnost PSNR koeficijenta o brzini prijenosa



**Slika 7.20** Ovisnost VQM koeficijenta o brzini prijenosa



**Slika 7.21** Ovisnost SSIM koeficijenta o brzini prijenosa

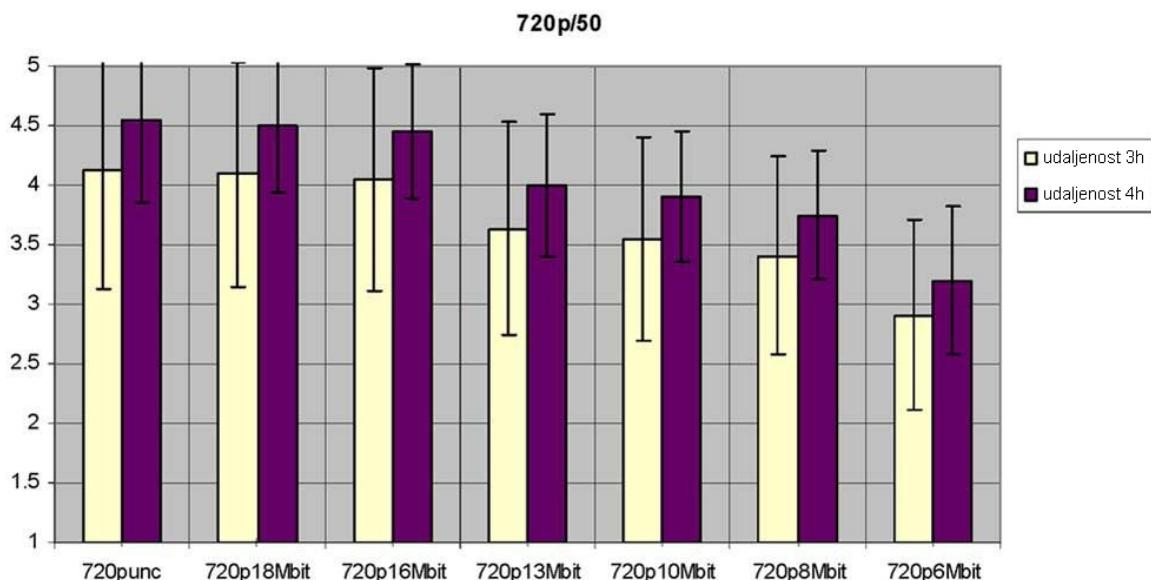
Kao i kod 1080i/25 formata, kompresija videosekvence u visokom profilu rezultirala je boljim ocjenama kod objektivnih mjera kvalitete.

### 7.3.7. Subjektivne ocjene kvalitete

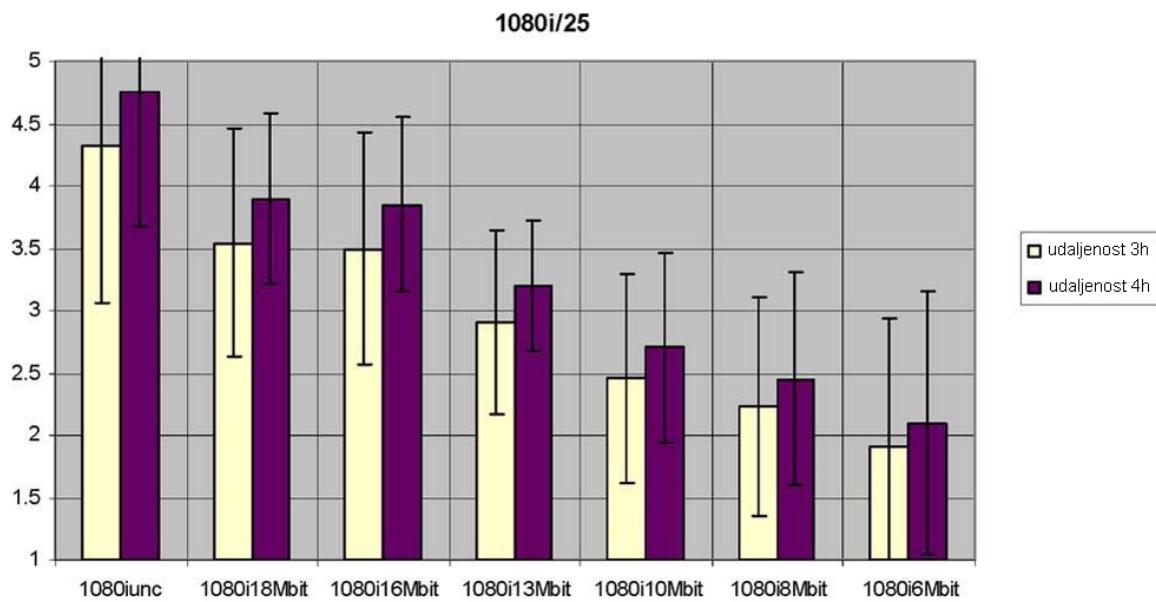
Subjektivne ocjene kvalitete slike za videosekvencu Crowdrun, donesene su na temelju rada kojeg su objavili Hans Hoffmann, Member, IEEE, Takebumi Itagaki, Member, IEEE, David Wood, and Alois Bock 2006. godine u IEEE TRANSACTIONS ON BROADCASTING, VOL. 52, NO. 4.

Proveli su subjektivna istraživanja ocjene kvalitete slike za sva tri formata slike pri brzinama prijenosa od 6 Mbit/s do 18 Mbit/s. Videosekvence su promatrane na udaljenostima od tri i četiri visine slike (3h i 4h), a ocijenjivane su ocjenama od 1 do 5, gdje 1 odgovara najslabijoj kvaliteti slike a 5 najvišoj kvaliteti slike.

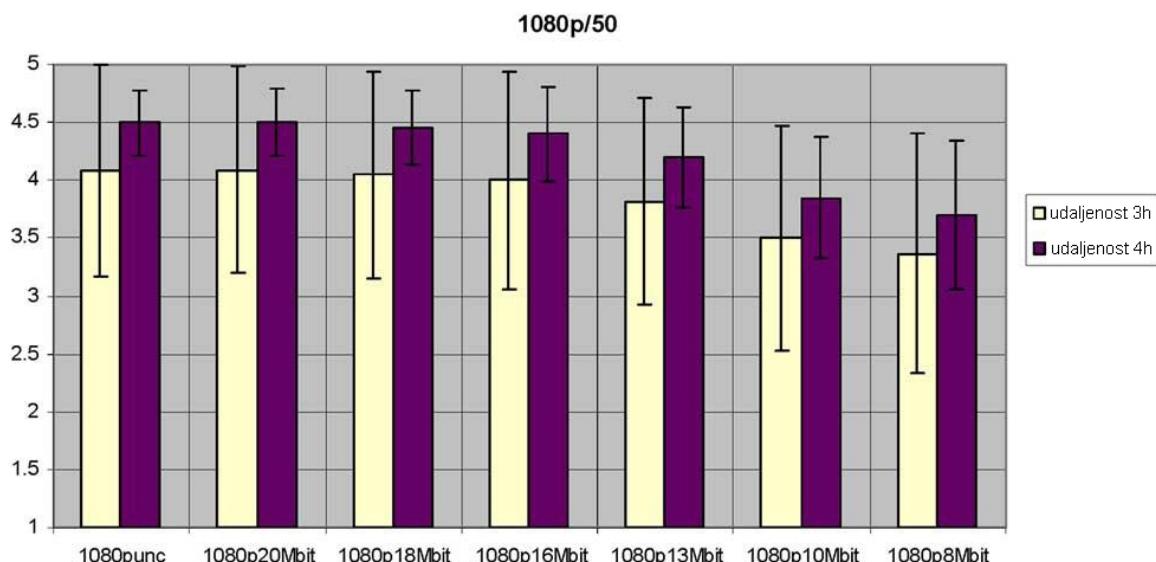
Na slikama 7.22, 7.23 i 7.24 prikazani su dobiveni rezultati.



**Slika 7.22** Rezultati subjektivnog ispitivanja sekvence Crowdrun, format 720p/50



**Slika 7.23** Rezultati subjektivnog ispitivanja sekvence Crowdrun, format  
1080i/25



**Slika 7.24** Rezultati subjektivnog ispitivanja sekvence Crowdrun, format  
1080p/50

Na temelju dobivenih rezultata vidimo da su subjektivne ocjene kvalitete slike u raskoraku s objektivnim ocjenama. Objektivne ocjene oslanjaju se na matematičke modele i kao takve, ne mogu u potpunosti ocijeniti vizualni doživljaj kvalitete slike. Najveća razlika vidi se u 1080i/25 formatu slike. Matematički modeli daju dobre ocjene za ovaj format slike i stavljuju ga negdje u sredinu, između 720p/50 i 1080p/50 formata slike, dok je po subjektivnim ocjenama kvalitete on uvjerljivo najlošiji format i dobiva uvjerljivo najlošije ocjene kako opada brzina prijenosa. Po subjektivnim ocjenama kvalitete slike, 720p/50 i 1080p/50 su vrlo slični formati i za jednake brzine prijenosa dobivaju približno jednake ocjene.

## 8. Zaključak

H.264 postupak kompresije napredniji je postupak kompresije od prethodnih postupaka. Ima veliku primjenu u distribuciji i pohrani kako visokokvalitetnog videosignalna, tako i onog niske kvalitete. Prednost H.264 postupka kompresije u odnosu na prijašnje postupke je u velikoj adaptivnosti. H.264 norma sadrži sedam različitih profila koji imaju specifičnu namjenu. Na taj način postiže se optimalniji stupanj kompresije za određene primjene.

U postupku mjerjenja komprimirali smo dvije videosekvence s dijametralno suprotnim zahtjevima na postupak kompresije. Videosekvenca Intotree je sekvenca s malo dinamike i kao takva poželjna je za kompresiju, tj. mogu se dobiti visoki stupnjevi kompresije s vrlo malo vidljive distorzije u slici. Objektivne mjere kvalitete prikazuju vrlo visoki stupanj kvalitete dekodirane slike (SSIM vrlo blizu 1, VQM ispod 1, a PSNR oko 35dB).

Za razliku od videosekvence Intotree, videosekvenca Crowdrun je vrlo zahtjevna za kompresiju. Sadrži vrlo visoki stupanj dinamike, promjene boje, kretanje kamerom. Stupanj kompleksnosti postupka kodiranja iznosi preko 95%. Na temelju subjektivnih i objektivnih postupaka mjerjenja kvalitete slike vide se sve prednosti H.264 postupka kompresije. Glavni profil omogućuje nam kvalitetnu kompresiju 720p/50 formata, dok se u 1080p/50 i 1080i/50 ne pokazuje najbolje. U visokom profilu transformacijsko kodiranje se provodi na manjem bloku. Upravo iz tog razloga, visoki profil daje puno bolje objektivne mjere kvalitete nego glavni profil.

Osim objektivnih mjera kvalitete, tu su i subjektivne mjere kvalitete. Nemoguće je matematičkim modelima nadomjestiti i simulirati doživljaj viđenog. To je razlog zbog kojeg subjektivne mjere kvalitete odudaraju od objektivnih. U subjektivnim mjerama kvalitete daleko najlošije prolazi 1080i/25 format. Kod pojave prvih televizijskih sustava problem je bila visoka vertikalna frekvencija, te se ona nadomještala analiziranjem s proredom. Takav način analiziranja pokazuje velike slabosti kod scena s vrlo visokom dinamikom i brzim pokretima. Upravo to je razlog subjektivno slabijim ocjenama od preostala dva formata, iako objektivne mjere kvalitete to ne pokazuju. 1080p/50 format u obje mjere pokazuje vrlo visok stupanj kvalitete. Pri visokim brzinama prijenosa, u potpunosti iskorištava sve

prednosti H.264 postupka i ostvaruje vrlo visoke objektivne mjere (PSNR, SSIM i VQM), dok se prednost u višoj rezoluciji najbolje vidi u objektivnim mjerama. Iako 1080p/50 format zahtijeva visoke brzine prijenosa, i pri niskim brzinama prijenosa, ostvaruje relativno visoke subjektivne ocjene. Iako se glavni profil nije pokazao pun lošijim od visokog profila, a u nekim postupcima i neznatno boljim, upotreba glavnog profila izgubila je svrhu pojavom visokog profila. Visoki profil danas zauzima mjesto glavnog profilu u pohrani i distribuciji visokokvalitetnog videosignalna.

## 9. Popis literature

1. I.E.G. Richardson: *H.264 and MPEG-4 Video Compression*, John Wiley & Sons, 2003.
2. G.Gvozden: *Usporedba normi za analiziranje slike u sustavima televizije visoke kvalitete*, magistarski rad
3. Z.Wang, A.C.Bovik, H.R.Shek et al: *Image Quality Assessment : From Visibility to Structural Similarity*, IEEE TRANSACTIONS ON IMAGE PROCESSING, Vol 13, No4, April 2004
4. H.Hoffman, T.Itagaki, D.Wood et al: *Studies on the Bit Rate Requirements for a HDTV Format With 1920x1080 pixel Resolution, Progressive Scanning at 50 Hz Frame Rate Targeting Large Flat Panel Displays*, IEEE TRANSACTIONS ON BROADCASTING, Vol52, No4, December 2006
5. T.Oelbaum, H.Schwarz, M.Wien et al: *Subjective Evaluation on the SVC Extension of H.264/AVC*
6. G.Gvozden, E.Dumic, S.Grgic: *Exploring the Characteristics of High Definition Television*, Springer
7. D.Bardone, E.S.G.Carotti, J.C De Martin: *Adaptive Golomb Codes for Level Binarization in the H.264/AVC FReExt Lossless Mode*
8. C.Poynton: *Digital video and HDTV*, Elsevier Scienc, 2003.
9. L.Haglud: *The SVT High Definition Multi Format Test Set*, 2006
10. A.Sadka: *Compressed Video Communications*, John Wiley & Sons, 2002.
11. S.Juric: *Obrada digitalnog videosignalata u televizijskim kamerama*, magistarski rad
12. A.B.Watson, J.Hu, J.F.McGowan III: *DVQ – A digital video quality metric based on human vision*
13. M.Ouaret, F.Dufaux, T.Ebrahimi: *On Comparing Image and Video Compression Algorithms*
14. M.Rboz: *Primjena kompresije videosignalata u digitalnim kamerama*, diplomski rad

1)

