

# MPEG-2 i MPEG-4 NORME ZA KODIRANJE VIDEOSIGNALA

Hoblaj J.<sup>1</sup>, Matković D.<sup>1</sup>

<sup>1</sup>Veleučilište u Varaždinu, Varaždin, Hrvatska

**Sažetak:** U ovome članku govori se o kodiranju nekomprimiranog SDTV (Standard Definition Television-televizija standardne razlučivosti) digitalnog videosignala prema MPEG-2 normama (ISO/IEC 13818). Unutar dokumenta ISO/IEC 13818 koji sadržava deset poglavlja, u poglavlju ISO/IEC 13818-2 definirano je kodiranje digitalnog videosignala. Njime su određeni postupci (profili i razine) za dobivanje toka podataka i načina kako dekoder mora interpretirati tok podataka. Normama nije određena konstrukcija i način rada koda. Ona je ostavljena proizvođačima uređaja i softvera da je zadrže kao svoju, zakonom zaštićenu, tehnologiju. Korisna posljedica ovoga je stalno natjecanje u razvijanju koda sve boljih karakteristika (viši stupanj kompresije uz zadanu kvalitetu televizijske slike). U nastavku je dan pregled videokodiranja prema MPEG-4 normama u kojemu su dane osnovne značajke i razlike u videokodiranju u odnosu na MPEG-2 norme te su spomenuti profili i razine MPEG-4 normi.

**Ključne riječi:** SDTV signal, videokodiranje, MPEG-2, MPEG-4

**Abstract :** This article discusses the encoding of uncompressed SDTV (Standard Definition Television) digital video according to MPEG-2 standard (ISO / IEC 13 818). Within ISO / IEC 13 818, which is documented in 10 parts, the part ISO / IEC 13818-2 defines the encoding of digital video. MPEG-2 devised a two-dimensional structure (profiles and levels) for classifying bit streams and decoders. Profiles define the tools that may be used. The MPEG-2 standards give very little information regarding the structure and operation of the encoder. It was left to device manufacturers and software developer to keep it as their own, proprietary, technology. A useful result is that there can be competition between different encoder designs which means that better designs can evolve (higher compression factor with a defined quality of television pictures). Further to is an overview of video coding according to MPEG-4 standard in which they are given the basic characteristics and differences in relation to video coding in MPEG-2 standard, and finally profiles and levels in MPEG-4 standard are mentioned.

**Key words:** SDTV signal, Video Coding Techniques, MPEG-2, MPEG-4

## 1. UVOD

Rasprostranjeno korištenje komprimiranog digitalnog videosignala omogućilo je napredak u digitalnoj tehnologiji. Normizacija je vrlo važna u razvoju zajedničkih metoda kompresije koje se koriste u novim uslugama i proizvodima. Najveći problem za uspješnu komunikaciju je ograničen kapacitet prijenosnog medija. Razvoj prijenosnih medija većeg kapaciteta teče sporije od razvoja elektroničkih sklopova, pa je posljedica ta da elektronika može procesirati mnogo više informacija nego što je medij može prenositi. Činjenica je da količina informacije koju je potrebno prenijeti neprestano raste, a prijenosni sustav ima ograničen kapacitet. Posljedica je zagušenje prijenosnog medija odnosno povećanje vremena potrebnog da bi se informacija uspješno prenijela. Jedno od rješenja je kompresija informacije. MPEG-2 normama se želi dobiti sustav kodiranja videosignala namijenjen za veliki raspon aplikacija. MPEG kompresija videosignala se koristi u mnogim proizvodima kao što su DVB-T prijammnici, DVD playeri, HDTV dekoderi i drugo. Prednost kod prijenosa komprimiranog videosignala je u tome što zahtijeva manje prostora za pohranu podataka te manje brzine prijenosa podataka. MPEG norma je jedna od najpopularnijih metoda kompresije videosignala zbog toga jer ne sadrži samo jednu normu već čitav niz normi pogodnih za kompresiju videosignala za različite aplikacije.

Digitalni nekomprimirani televizijski videosignali standardne kvalitete slike (SDTV-Standard Definition Television) imaju brzine prijenosa podataka od 270 Mbit/s, koja je prevelika za emitiranje te se mora smanjiti na 2 do 7 Mbit/s. Brzina prijenosa podataka kod nekomprimiranih digitalnih televizijskih videosignala visoke kvalitete slike (HDTV-High Definition Television) veća je od 1 Gbit/s i kodirani HDTV signal u MPEG-2 normi postiže brzinu prijenosa od 15 do 20 Mbit/s.

## 2. VIDEOKOMPRESIJA PREMA MPEG-2 NORMI

Podatke je moguće komprimirati tako da se uklone suvišne (redundantne) i nebitne (nevažne) informacije. Suvišne informacije su one koje se ponavljaju u toku podataka, koje ne nose novu informaciju, odnosno one

koje se na prijamnom mjestu mogu dobiti iz prethodno primljenih informacija. Smanjenje redundancije može se postići npr. kodiranjem dužine niza (RLC – Run-Length Coding). Umjesto emitiranja deset nula, informacija „deset nula“ može biti poslana pomoću specijalnog koda koji je mnogo kraći. Kod Morseove abecede također je primijenjen postupak smanjenja redundancije. Slova koje se češće ponavljaju u govoru (engleski jezik) kodiraju se manjom duljinom kodne riječi, a slova koja se rjeđe ponavljaju kodiraju se većom duljinom kodne riječi. Ova vrsta kodiranja zove se Huffmanovo kodiranje ili kodiranje s promjenjivom dužinom riječi (VLC-Variable-Length Coding). Nebitna (nevažna) informacija je tip informacije koja se ne može opaziti ljudskim osjetilima (sluh, vid). U slučaju videosignala to su komponente (detalji u slici po obliku, kontrastu i boji) koje se ljudskim vidom ne mogu zapaziti. Ljudskim vidom puno se bolje opažaju detalji s malim razlikama u svjetlini nego detalji s malim razlikama u boji. Zbog toga je moguće smanjiti „oštrinu boje“ u slici, odnosno bitno smanjiti frekvencijsku širinu signala boje bez utjecaja na vidljivu degradaciju slike. Isto tako, ljudskim vidom se slabije raspoznaju fine strukture u slici (visoke frekvencije signala) nego grube strukture (niske frekvencije signala), čime je ujedno određena i ovisnost osjetljivosti ljudskog vida od frekvencije šuma u slici. Ove osobine vida su iskorištene u postupcima kompresije kod JPEG i MPEG normi za kodiranje mirnih i pokretnih slika, na način da se grube strukture u slici kodiraju s puno većom točnošću (većim brojem bita) nego fine strukture u slici (manjim brojem bita). Ovakvim kodiranjem eliminira se nebitna (nevažna) količina informacije poznata pod nazivom subjektivna redundancija. Međutim, smanjivanje nevažne količine informacije dovodi do nepovratnog gubitka te količine informacije (ireverzibilni proces), za razliku od postupka kod kojega se smanjivanjem redundancije ne unose gubici (reverzibilni proces), odnosno dekodiranjem se dobiva ista količina informacije kao na ulazu u koder.

U kodiranju kod MPEG-2 norme provode se sljedeći koraci kako bi se postigao faktor kompresije podataka od 130 (kompresija 130:1):

1. 8 bitna rezolucija umjesto 10 bitne (uklanjanje nevažne informacije iz toka podataka),
2. izostavljanje horizontalnih i vertikalnih potisnih intervala (uklanjanje redundancije iz toka podataka),
3. smanjenje rezolucije boje u vertikalnom smjeru (4:2:0), (uklanjanje nevažne informacije iz toka podataka tj. uklanjanje subjektivne redundancije),
4. primjena postupka diferencijalne impulsokodne modulacije (uklanjanje redundancije (vremenska) između slika koje slijede u nizu jedna za drugom),
5. diskretna kosinusna transformacija s kvantizacijom (uklanjanje redundancije iz toka

podataka odnosno uklanjanje prostorne redundancije iz svake slike),

6. cik-cak skeniranje matrice koeficijenata uz primjenu kodiranja dužine niza (RLC-Run Length Coding), (uklanjanje redundancije iz toka podataka),
7. Huffmanovo kodiranje (VLC- Variable Length Coding- kod s promjenjivom dužinom kodne riječi), (uklanjanje redundancije iz toka podataka).

U analognoj televiziji, kada videosignal ima omjer signal/šum od 48 dB ili više (ponderiran i u odnosu na razinu vršnog bijelog), komponenta šuma je ispod praga percepcije ljudskog vida. Uz odgovarajuću pobudu A/D pretvarača, šum kvantizacije kod 8 bitne rezolucije je već ispod tog praga tako da je 10 bitna rezolucija YCbCr signala nepotrebna izvan studija. U studiju 10 bitna rezolucija daje bolje rezultate zbog toga jer je naknadna obrada signala lakša i daje bolje rezultate (digitalni videoefekti i tehnika kolorne podloge). Smanjenjem količine podataka korištenjem 8 bitne, a ne 10 bitne riječi po uzorku, a prema parametrima kodiranja u ITU-R BT.601 normi, znači smanjenje brzine prijenosa podataka od 20% ( $((10-8)/10=2/10=20\%)$ ). To je nepovratni gubitak informacije i originalni videosignal se ne može dobiti dekodiranjem na kraju prijama.

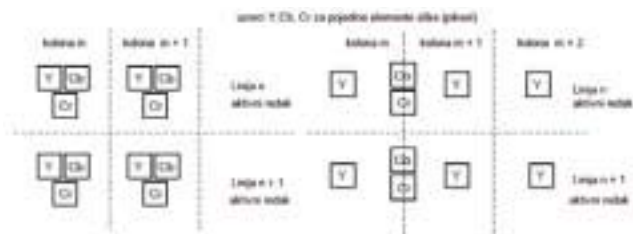
Horizontalni i vertikalni potisni intervali digitalnog televizijskog signala prema ITU-R BT.601 normi ne sadrže nikakve relevantne informacije, čak ni za teletekst. Horizontalni i vertikalni potisni intervali su kompletno izostavljeni kod MPEG-2 norme. Ta područja mogu sadržavati podatke kao što su tonski signali, ali se prema MPEG-2 normi oni prenose i kodiraju odvojeno. Horizontalni i vertikalni potisni intervali (slika 1.) i svi signali u njima mogu se regenerirati opet bez problema na kraju prijenosa.

Europski PAL signal ima 625 linija (redaka) u slici od kojih je vidljivo njih 575. Razlika od 50 linija u odnosu na 625 linija je 8% smanjenja protoka podataka kada je izostavljen vertikalni potisni interval. Duljina jedne linije je 64  $\mu$ s, ali aktivni (vidljivi) dio videolinije je samo 52  $\mu$ s gdje se uštedi 19% u protoku podataka kada je izostavljen horizontalni potisni interval. Budući da postoji određeno preklapanje kod ova dva reduciranja, ukupan postotak smanjenja redundancije je oko 25%.



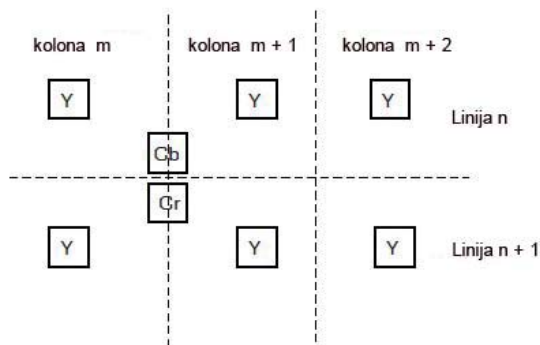
Slika 1. Vertikalni i horizontalni potisni intervali [1]

Dva različita krominantna signala Cb i Cr uzorkovana su s upola manjom frekvencijom uzorkovanja u usporedbi sa lumnantnim signalom Y. Osim toga, smanjena je i pojasna širina krominantnih Cb i Cr signala na 2,75 MHz u odnosu na pojasnu širinu lumnantnog signala od 5,75 MHz te je dobivena struktura uzorkovanja slike 4:2:2 (slika 2.).



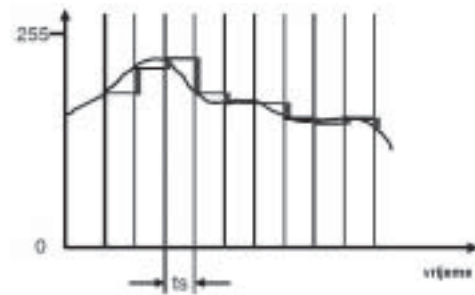
Slika 2. Struktura uzorkovanja slike 4:4:4 i 4:2:2 [1]

Kako se ljudskim vidom ne zamjećuje razlika u rezoluciji boje između horizontalnog i vertikalnog smjera, moguće je također smanjiti rezoluciju boje na pola u vertikalnom smjeru bez zamjetne degradacije slike. Kod MPEG-2 norme to je obično jedan od prvih koraka i signal tada ima 4:2:0 strukturu uzorkovanja (slika 3.). Svakom od četiri uzorka lumnantnog signala pridružen je jedan uzorak krominantnih Cb i Cr signala. Ovaj tip smanjenja rezultira 25% uštedom vrijednosti količine podataka.



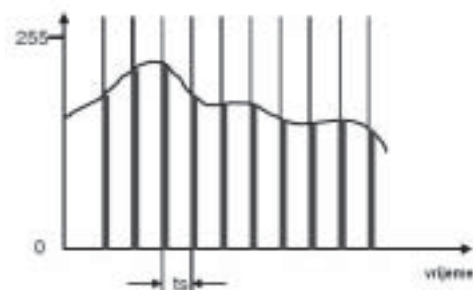
Slika 3. 4:2:0 struktura uzorkovanja slike [1]

Susjedne televizijske slike se vrlo malo razlikuju jedna od druge. One sadrže stacionarna područja koja se uopće ne mijenjaju iz slike u sliku. Postoje i područja koja samo mijenjaju svoju poziciju i postoje objekti koji se prvi put pojavljuju. Ako bi se svaka slika kompletno prenosila svaki put, neke emitirane informacije bi uvijek bile iste i rezultirale bi visokom brzinom prijenosa podataka. Zaključak koji se nameće je taj da je potrebno prenositi samo razlike između susjednih slika. Ovaj način smanjenja redundancije temelji se na odavno poznatom modulacijskom postupku diferencijalno impulsokodne modulacije (DPCM – Differential Pulse Code Modulation), (slika 4.).



Slika 4. Diferencijalno impulsokodna modulacija [1]

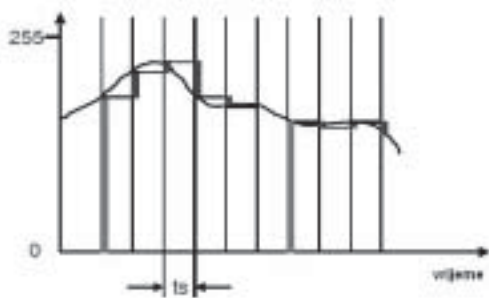
Ako je kontinuirani analogni signal uzorkovan i diskretiziran po amplitudi dobivaju se diskretne vrijednosti signala (brojevi) na jednako udaljenim vremenskim intervalima.



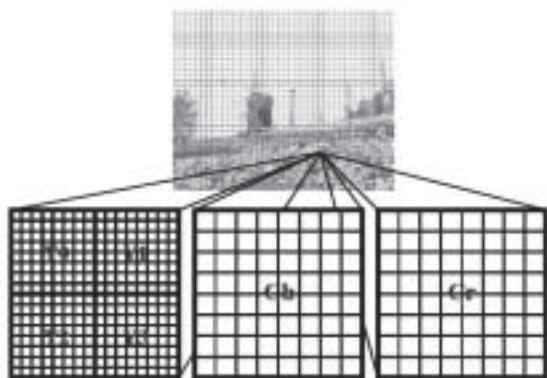
Slika 5. Impulsokodna modulacija, niz brojeva s vrijednostima između 0 i 255 [1]

Te vrijednosti se mogu prikazati kao impulsi na jednako udaljenim intervalima (slika 5.). Visina svakog impulsa nositelja informacije je diskretna. U stvarnosti razlike između susjednih uzoraka (PCM vrijednosti) nisu velike zbog prethodnog pojasnog ograničenja (niskopropusnog filtriranja) analognog signala. Ukoliko se umjesto PCM vrijednosti prenose samo njihove razlike smanjuje se i prenesena količina podataka (slika 4.). Sa smanjenjem količine podataka smanjuje se i brzina prijenosa podataka te se ekonomičnije koristi kapacitet prijenosnog kanala. Problem s uobičajenim DPCM-om je kod slučajnog uključivanja u tok podataka ili nakon pogrešaka u prijenosu kada treba duže vrijeme da demodulirani signal postigne oblik originalnog signala. Ovaj problem se rješava periodičnim prijenosom kompletnih PCM vrijednosti uzoraka, između kojih se prenose samo razlike između susjednih PCM uzoraka (slika 6.). Opisani princip u DPCM modulacijskom postupku iskorišten je u MPEG-1/-2 normama radi smanjenja redundancije između susjednih televizijskih slika (vremenska redundancija).





Slika 6. Diferencijalno impulsnokodna modulacija s periodičnim prijenosom referentnih (kompletnih PCM) vrijednosti [1]



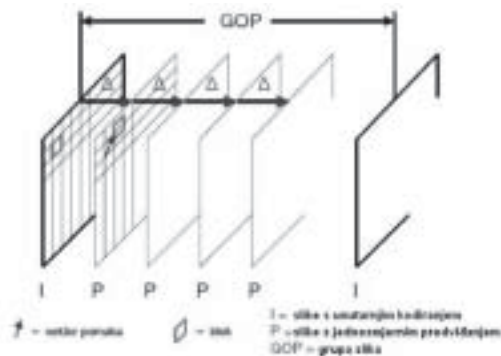
Slika 7. Struktura makrobloka (YCbCr) s 4:2:0 strukturom uzorkovanja [1]



Slika 8. Dijeljenje slike na makroblokove (16x16 uzoraka) [1]

Kodiranje televizijske slike (720x576) počinje podjelom slike na makroblokove od 16x16 luminantnih Y uzoraka (piksela) i 8x8 krominantnih Cb i Cr uzoraka (piksela) u slici (4:2:0 struktura uzorkovanja). Zbog 4:2:0 strukture uzorkovanja, 8x8 krominantni Cb i Cr uzorci prekrivaju površinu slike od 16x16 luminantnih uzoraka (makroblok), (slika 7.). Jedna slika se sastoji od velikog broja makroblokova (slika 8.), a horizontalni i vertikalni broj uzoraka slike je odabran tako da je djeljiv sa 16 i sa 8 (Y:720x576 piksela). Makroblok je najmanji dio slike

koji se koristi za procjenu i nadomještanje pokreta. Formiranje razlike vrijednosti uzoraka između susjednih slika zbog pomaka objekta u slici odvija se na razini makrobloka, tj. odgovarajući makroblok sljedeće slike je uvijek uspoređen s makroblokom prethodne slike. Preciznije rečeno, provjerava se u kom smjeru i za koliko se pomaknuo makroblok sljedeće slike u odnosu na makroblok iz prethodne slike. Ukoliko pomak postoji, u prijenos ide samo podatak o vektoru pomaka. Ako nema promjene položaja makrobloka prema istom u prijašnjoj slici u prijenos se ne šalje nikakav podatak o sadržaju makrobloka. Ukoliko se pored pomaka makrobloka dogodila i određena razlika prema makrobloku iz prethodne slike, tada se uz vektor pomaka šalje i razlika između makroblokova. Ovako kodirane slike s nadomještenim pokretom i pridruženom razlikom u odnosu na prethodnu sliku, ako je ima, nazivaju se P (Predicted)- slike. Naziv „predicted“ dolazi zbog procjenjivanja veličine pomaka objekta u trenutnoj slici (predicted-predvidjeti) prema njegovom položaju u referentnoj (prethodnoj) slici, npr. prethodnoj I slici, (slika 9.).



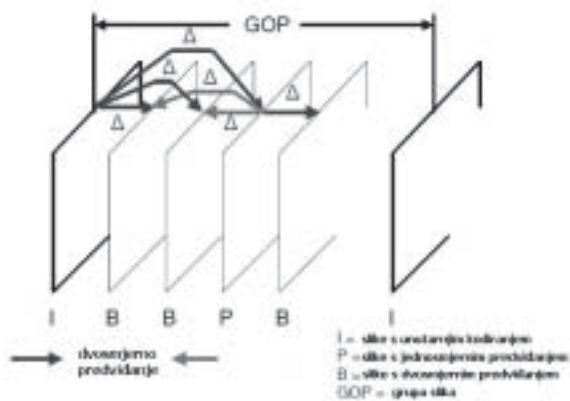
Slika 9. Slike s nadomještenim pokretom (pomakom) u odnosu na prethodnu sliku (P-slike) [1]

Ako je razlika između trenutnog i prethodnog makrobloka vrlo velika tada se cijeli trenutni makroblok ponovo kodira i prenosi (I-slike). Naziv dolazi od „Intra frame“ a odnosi se na kodiranje slike pri čemu se uklanja samo prostorna redundancija. Kod dekodiranja I-slike daju potpunu sliku neovisnu od sadržaja prethodne ili sljedeće slike, analogno kao kod demodulacije signala DPCM-a s referentnim PCM vrijednostima (slika 6.).

Osim jednosmjernog predviđanja pomaka u slici (P-slike), postoji i dvosmjerno predviđanje, tj. prema naprijed (razlika u pomaku u odnosu na sliku koja slijedi) i prema natrag (razlika u pomaku u odnosu na prethodnu sliku). Ovako kodirane slike s nadomještenim pokretom i pridruženom razlikom u odnosu na prethodnu i sljedeću sliku u nizu nazivaju se B (Bidirectional)-slike (slika 10.). Razlog uvođenja dvosmjernog predviđanja je puno manja vrijednost količine podataka kod B-slika u usporedbi s P-slikama ili I-slikama.

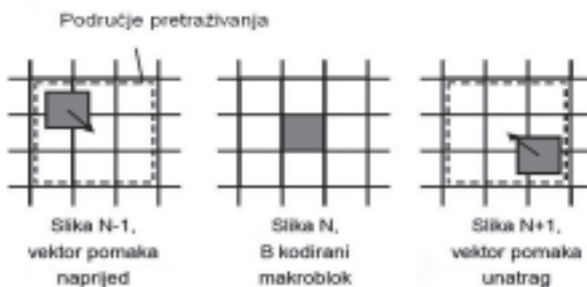
Procjena pokreta za dobivanje vektora pokreta dobiva se kodiranjem I- i B-slika ( $\Delta$ -delta slike) pretraživanjem

određenog područja unutar kojega se utvrđuje novi položaj traženog referentnog makrobloka.



Slika 10. Slike s dvosmjernim predviđanjem (B-slike) [1]

Novi položaj referentnog makrobloka određuje se prema prethodnoj slici (dobivanje P-slike) ili prema prethodnoj i sljedećoj slici (dobivanje B-slike). To se postiže traženjem podudaranja makroblokova (referentnog iz prethodne ili prethodne i sljedeće slike s novim položajem istog makrobloka), unutar područja pretraživanja oko referentnog makrobloka (slika 11.).

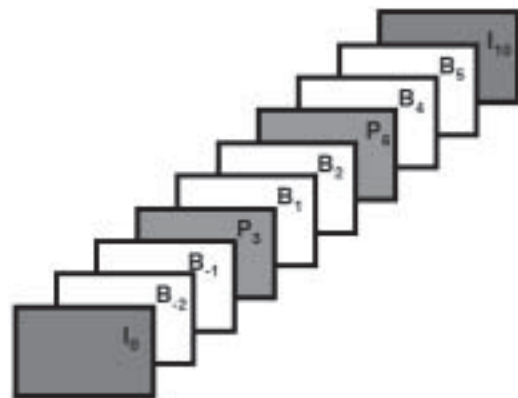


Slika 11. Vektori pokreta (pomaka) makrobloka prethodne slike i sljedeće slike u odnosu na trenutnu sliku [1]

Ako je traženi makroblok u sljedećoj slici na mjestu ispred, a kod dvosmjernog kodiranja i u prethodnoj slici na mjestu iza traženog makrobloka, izračunavaju se vektori pokreta za prvi ili oba slučaja (prema prethodnoj i prema sljedećoj slici) i prenose kao P- ili B-slike. Osim toga, svaki novi sadržaj u području pretraživanja koji nije rezultat pomaka makrobloka kodira se kao razlika između slika i prenosi zajedno s vektorom pokreta. Razlike između slika (bilo P- ili B- slike) komprimiraju se pomoću diskretne kosinusne transformacije (DCT), kvantizacije i entropijskog kodiranja.

Diskretna kosinusna transformacija je linearna transformacija vrijednosti uzoraka (piksela) u koeficijente (nove numeričke vrijednosti uzoraka), čijom je kvantizacijom omogućeno uklanjanje prostorne redundancije unutar slike.

Najmanja jedinica za prijenos kodiranih slika koja se može neovisno dekodirati naziva se grupom slika (GOP-Group of Pictures), (slika 12.), i sastoji se od određenog broja I-, P- i B-slika s točnim redoslijedom te uvijek počinje s I-slikom. GOP je obično duljine 12 slika i s redoslijedom I,B,B,P,B,B,P,B,B,P,B,B,I... B-slike se nalaze između I- i P-slike. Prije nego što se dekodira B-slika na kraju prijenaosa treba imati prethodne i sljedeće I- i P-slike. Prema MPEG normama, GOP struktura može biti varijabilna. Da se na strani dekodera izbjegne potreba za velikim memorijskim prostorom, mora se promijeniti redoslijed prijenaosa kodiranih slika unutar GOP jedinice. Iz tog razloga se slike prenose drukčijim redoslijedom od onoga kod reprodukcije, a koji je isti kao kod kodiranja.



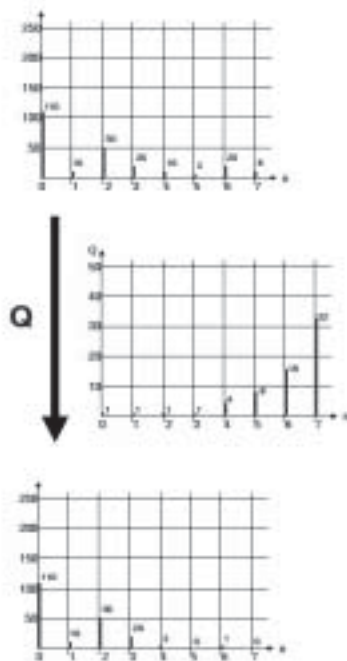
Slika 12. Redoslijed slika kod prijenaosa [1]

Umjesto originalnog redoslijeda: I0, B1, B2, P3, B4, B5, P6, B7, B8, P9 slike se prenose redoslijedom: I0, B-2, B-1, P3, B1, B2, P6, B4, B5, P9... (slika 12.). To znači da su P- i I-slike koje slijede iza B-slika dostupne na kraju prijenaosa prije B-slika. Sada je lako odrediti i izračunati memorijski prostor rezerviran na strani dekodera. Da bi se ponovo prijenosni redoslijed vratio u originalni, moraju se uz slike kodirati i prenositi pripadni redni brojevi slika. Za tu svrhu koriste se DTS (Decoding Time Stamp) vrijednosti sadržane u PES (Packetized Elementary Stream) zaglavlju toka podataka.

Vrlo uspješna metoda za kompresiju mirnih slika koja je u upotrebi od kraja 80-tih godina, a koristi se kod digitalnih fotoaparata kojima se dobivaju slike vrhunske kvalitete, poznata je pod nazivom JPEG. JPEG je kratica za Joint Photographic Experts Group, a odnosi se na odbor za donošenje normi za kompresiju mirnih slika. Osnovni algoritam korišten u JPEG normi je diskretna kosinusna transformacija ili DCT (Discrete Cosine Transform). DCT isto tako ima ključnu ulogu u MPEG normama.

Slika se sastoji od grubih i finih struktura. Ljudski vid nije jednako osjetljiv za fine i grube strukture u slici. Moguće je dopustiti puno više šuma u visokofrekvencijskom području videosignala koje odgovara finim strukturama u slici nego u niskofrekvencijskom području videosignala koje odgovara grubim strukturama u slici. Zbog toga je već od

samih početaka mjerenja omjera signala/šuma u slici uključena osjetljivost ljudskog vida, a samo mjerenje vrednovano (ponderirano) je prema amplitudno-frekvencijskoj karakteristici osjetljivosti vida na šum u slici. Ova osobina vida iskorištena je u kodiranju slike (JPEG i MPEG norme). Tako se grube komponente slike kodiraju finijom kvantizacijom, a fine komponente slike grubom kvantizacijom, s ciljem smanjenja brzine prijenosa toka podataka.



Slika 13. Kvantizacija DCT koeficijenata [1]

Kako odvojiti grube komponente od finih komponenti slike. Rješenje je primjena onih linearnih integralnih transformacija putem kojih se može ostvariti prijelaz iz vremenske domene videosignala u frekvencijsku domenu (spektar videosignala). Diskretna kosinusna transformacija je poseban slučaj Fourierove transformacije, (i brze Fourierove transformacije), a pripada metodama numeričke matematike. Nakon DCT transformacije, vrijednosti uzoraka jednog retka iz bloka od 8x8 uzoraka (8x8 piksela) dobivaju nove vrijednosti (prvi dijagram na slici 13.). Ove vrijednosti sada predstavljaju koeficijente (frekvencijska domena), čije vrijednosti određuju udio kosinusnih valnih oblika na određenim frekvencijama. Koeficijent na prvom mjestu određuje udio istosmjerne komponente (DC-Direct Current), na drugome udio prvog poluharmonika, na trećemu i dalje do sedmog koeficijenta, cjelobrojne umnoške vrijednosti frekvencije poluharmonika i rastućih cijelih brojeva od 2 do 7 (AC-Alternating Current-AC koeficijenti). Kada bi se zbrojilo svih sedam valnih oblika čije su amplitude određene vrijednostima koeficijenata, s istosmjernom komponentom, rezultat bi bila vrijednost uzoraka jednog retka bloka, (8x8 piksela), od kojeg se i krenulo. Dobiveni koeficijenti se zatim kvantiziraju tj. dijele s određenim brojem, faktorom

kvantizacije. Što je veća vrijednost broja s kojim se dijeli, grublja je kvantizacija. Kako su koeficijentima određene grube i fine strukture u slici, mogu se primijeniti različiti faktori kvantizacije na pojedine koeficijente u smjeru od koeficijenta istosmjerne komponente do koeficijenta najviše frekvencije. Uzimajući u obzir osjetljivost vida od niskih prema višim frekvencijama u slici, kvantizacija će biti postavljena od fine prema gruboj, tj. od vrlo malih vrijednosti faktora za grube strukture u slici do velikih vrijednosti faktora za fine strukture u slici ( $Q(u,v)$  na slici 14.). Nakon kvantizacije nepovratno je izgubljena određena količina informacije iz slike.

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

$Q(v,u)$

Faktor skaliranja = 2

173	0	0	0	0	0	0	0
5	1	2	0	0	0	0	0
0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
-1	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0

$QF(v,u) = F(v,u) / Q(v,u) / \text{faktor skaliranja}$

Slika 14. Faktori kvantizacije  $Q(v,u)$ , skaliranja i vrijednosti koeficijenata (frekvencijska domena) [1]

100	190	220	235	244	236	222	211
196	209	222	231	229	215	198	186
144	154	170	184	190	190	185	180
162	164	166	167	165	161	157	154
195	191	195	188	179	178	173	161
174	166	161	156	160	170	183	192
174	160	138	113	112	115	125	133
152	138	119	105	104	115	133	146

$f(x,y)$

Slika 15. Blok od 8x8 uzoraka luminantnog signala jedne slike [1]

U praksi je kodiranje u JPEG i MPEG normi zasnovano na dvodimenzionalnoj DCT transformaciji. Slika je podijeljena na blokove od 8x8 piksela (uzorci luminantnog signala), (slika 15.). Svaki blok od 8x8 uzoraka se tada transformira u frekvencijsku domenu pomoću dvodimenzionalne DCT transformacije. Prije toga se vrijednost 128 najprije oduzima od svih

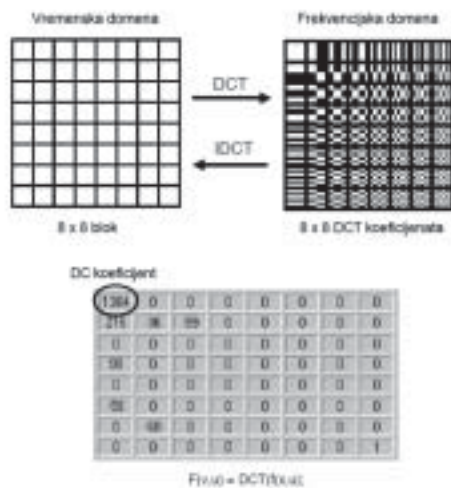


vrijednosti uzoraka kako bi se dobile vrijednosti s predznacima (slika 16.).

55	70	82	111	116	198	94	83
50	61	84	103	101	87	70	56
16	81	42	-36	62	62	57	52
-34	36	38	39	37	-33	29	26
67	63	57	22	50	50	51	53
46	40	33	28	32	42	55	64
46	32	30	4	16	13	-3	5
24	10	9	-25	-24	-15	5	18

Slika 16. Nove vrijednosti uzoraka dobivene oduzimanjem 128 od početnih vrijednosti [1]

Rezultat dvodimenzionalne DCT transformacije bloka od  $8 \times 8$  uzoraka je još jedan blok od  $8 \times 8$  uzoraka, ali sada u frekvencijskoj domeni (slika 17.).



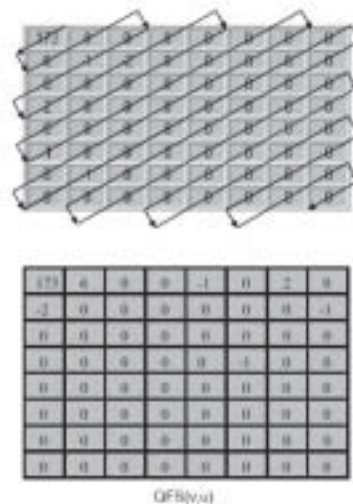
Slika 17. Koeficijenti  $F(v,u)$  dobiveni nakon dvodimenzionalne DCT transformacije [1]

Prvi koeficijent prvog reda je DC koeficijent koji odgovara DC komponenti cijelog bloka. Vrijednošću DC komponente određena je srednja vrijednost svjetline bloka. Drugi koeficijent odgovara intenzitetu grubih struktura slike u horizontalnom smjeru, a zadnji koeficijent prvog reda odgovara intenzitetu finih struktura slike u horizontalnom smjeru. Prvi stupac bloka od  $8 \times 8$  uzoraka od vrha do dna, na isti način kao kod horizontalnog smjera, sadrži intenzitete od grubih struktura slike do finih struktura slike u vertikalnom smjeru. Koeficijenti od grubih do finih struktura slike koji su kombinacije horizontalnih i vertikalnih frekvencija leže u dijagonalnom smjeru bloka.

Sljedeći korak je kvantizacija dobivenih koeficijenata  $F(v,u)$ . Svi koeficijenti se dijele s odgovarajućim faktorom kvantizacije, (tablica  $Q(u,v)$  na slici 14.). MPEG normom definirana je tablica kvantizacije koja može biti zamijenjena i nekom drugom na strani koda. Nova tablica se tada mora poslati dekoderu unutar toka

podataka. Kvantizacija rezultira velikim brojem koeficijenata čije su vrijednosti nula. Nakon kvantizacije, matrica ( $QF(v,u)$  slika 14.) je relativno simetrična s dijagonalom od gore lijevo do dolje desno. Matrica se očitava cik-cak skeniranjem kojim se tada stvara mnogo susjednih nula, (slika 18.). Primjenom kodiranja dužine niza (RLC) dolazi se do velikog smanjenja količine podataka, (gornji blok na slici 19.).

Uz 4:2:0 strukturu uzorkovanja, četiri Y bloka od  $8 \times 8$  uzoraka i po jedan blok od  $8 \times 8$  Cb i  $8 \times 8$  Cr uzoraka tvore jedan makroblok od  $16 \times 16$  elemenata slike u boji (slika 7.) Kvantizacija Y, Cb i Cr uzoraka se može mijenjati pomoću faktora skaliranja od makrobloka do makrobloka, (slika 14). Faktor skaliranja je određeni broj koji množenjem mijenja sve faktore kvantizacije, bilo u tablicama MPEG normi ili tablicama kvantizacije stvorenim od strane koda. Kompletna se tablica kvantizacije može zamijeniti na razini sekvence koja se sastoji od određenog broja GOP jedinica, u određenom vremenu u toku podataka.



Slika 18. Cik-cak skeniranje matrice  $QF(v,u)$  [1]

Dvodimenzionalna DCT transformacija s kvantizacijom koeficijenata primjenjuje se za makroblok Y uzoraka i za makroblok Cb i Cr uzoraka. U slučaju I-slika, svi makroblokovski su kodirani na već opisan način. U slučaju P- i B-slika se na razlike između uzoraka makroblokova dvije susjedne slike primjenjuje DCT transformacija (na svaki od četiri bloka), kvantizacija i entropijsko kodiranje (RLC i VLC), te se zajedno s vektorom pomaka prenosi u toku podataka. Razlika između dva makrobloka dobiva se tako da se makroblok prethodne slike pomakne za iznos vektora pomaka u novi položaj, pa se zatim izračuna razlika s makroblokom trenutne slike. Isti postupak se primjenjuje kod dvosmjernog kodiranja (B-slike).

Nakon cik-cak skeniranja (slika 18.), kvantiziranih DCT koeficijenata, dobiva se veliki broj susjednih nula. Umjesto tih nula, prenosi se samo podatak o njihovom broju koristeći RLC (Run-Length Coding) kodiranje. Ovaj tip smanjenja redundancije, uz pomoć DCT

transformacije i kvantizacije, osigurava veliki faktor kompresije podataka.



Slika 19. Kodiranje dužine niza (RLC) i Huffmanovo kodiranje [1]

Huffmanovo kodiranje se često koristi za kompresiju signala bez gubitaka, uz faktor kompresije 2:1. Huffmanovo kodiranje spada u postupke kompresije s promjenjivom duljinom kodne riječi (VLC, Variable Length Coding). Sam postupak kompresije temelji se na tome da se uštede bitovi i smanji brzina toka podataka. Kombinacije bitova koje se najčešće javljaju kodiraju se kratkim kodnim riječima, a one koje se rijetko pojavljuju, kodiraju se dužim kodnim riječima. Kodovi se pridjeljuju po Huffmanovoj tablici (slika 19.).

### 3. VIDEOKODIRANJE PREMA NORMI MPEG-4 PART 10 AVC (ADVANCED VIDEO CODING)

U usporedbi s MPEG-2 normama, znatno poboljšani MPEG-4 Part 10 AVC (H.264) videokodek omogućava smanjenje brzine prijenosa podataka od 30 do 50% uz bolju kvalitetu slike. To znači da se SDTV digitalni signal sada može prenositi brzinama od 1,5 do 3 Mbit/s. Radi usporedbe, brzina prijenosa kod MPEG-2 norme je od 2 do 7 Mbit/s a brzina kod nekomprimiranog SDTV digitalnog signala 270 Mbit/s. Koristeći MPEG-4 normu, HDTV digitalnom signalu se može smanjiti brzina prijenosa na oko 10 Mbit/s, u usporedbi s brzinama nekomprimiranog digitalnog signala vrijednosti oko 1,5 Gbit/s. Kod MPEG-2 norme brzina prijenosa istog iznosi oko 20 Mbit/s.

Glavna obilježja MPEG-4 Part 10 AVC (H.264):

1. podržani su 8, 10 i 12 bitni uzorci YCbCr komponentnog digitalnog videosignala strukture uzorkovanja 4:2:0, 4:2:2 i 4:4:4,
2. do maksimalno 16 referentnih slika,
3. povećana točnost kod nadomještanja pokreta (na 1/4 piksela za razliku 1/2 kod MPEG-2 normi),
4. korištenje SP (Switching P) i SI (Switching I) isječaka slike s kojima je omogućeno brzo prekapčanje dekodera s jednog na drugi stream,

5. cjelobrojna i Hadamardova transformacija umjesto DCT transformacije (veličina blokova 4x4 odnosno 2x2 piksela),
6. fleksibilnija struktura makroblokova (16x16, 16x8, 8x16, 8x4, 4x8, 4x4),
7. 52 moguća seta kvantizacijskih tablica,
8. In-loop de-blocking filter (uklanjanje izobličenja u vidu vidljivih blokova u slici),
9. redoslijed isječaka slike kod prijenosa može biti proizvoljan (osigurava bolju zaštitu toka podataka od grešaka u prijenosu),
10. entropijsko kodiranje; VLC i kontekstno prilagodljivo binarno aritmetičko kodiranje (CABAC).

H.264 norma koristi YCbCr model boja, podržavajući 4:2:0, 4:2:2 i 4:4:4 strukture uzorkovanja. Uz 4:2:2 i 4:4:4 strukture uzorkovanja povećana je rezolucija boje u slici u odnosu na 4:2:0 strukturu uzorkovanja, što rezultira boljom kvalitetom slike. U odnosu na 8-bitne uzorke YCbCr digitalnog komponentnog videosignala kod MPEG-2 norme, H.264 norma podržava 10-bitne i 12-bitne uzorke kojima se osigurava daljnje poboljšanje kvalitete slike.

Kod H.264 norme, uvedena je dodatna segmentacija makroblokova, pa je ukupno osam veličina (16x16, 16x8, 8x16, 8x8, 8x4, 4x8, 4x4). Takva fina segmentacija dovodi do potencijalno velikog broja vektora pokreta po makrobloku, (do 32), i broja blokova čiji se uzorci izračunavaju interpolacijom, (do 96). Da bi se ograničila kompleksnost koda/dekoda, postoje ograničenja u broju vektora pomaka koji se koriste za dva uzastopna makrobloka. Poboľšano je prikrivanje pogrešaka kod prijenosa fleksibilnijim raspoređivanjem makroblokova (FMO-Flexible Macroblock Ordering) unutar isječaka slike tako što se makroblokovima mogu dodijeliti drugim isječcima koji slijede isječak koji se trenutno popunjava makroblokovima očitavanjem slike. Time se smanjuje vjerojatnost da pogreška utječe na veće prostorno područje slike, te se poboljšava prikrivanje pogreške upotrebom susjednih ispravnih makroblokova za predviđanje neispravnog makrobloka.

In-Loop De-blocking Filter uklanja izobličenja u vidu vidljivosti blokova koja proizlaze zbog razlika u vrijednostima susjednih piksela, na granicama blokova i makroblokova, zbog razlike u nadomještanju pomaka kod susjednih blokova ili različitim tablicama kvantizacija susjednih blokova. Taj filter isto tako modificira razlike susjednih piksela dvaju blokova ili makroblokova u zavisnosti od sadržaja obuhvaćenih blokovima.

Isječak slike je niz makroblokova koji se očitavaju s lijeve prema desnoj strani slike. Isječak je najmanja jedinica koja nosi podatak za sinkronizaciju rada koda i dekodera. U slučaju pogreške kod jednog isječaka dekodiranje se nastavlja sljedećim ispravnim isječcima (nema utjecaja pogreške na sljedeće isječke). Ova fleksibilnost omogućuje proširenje I-, P- i B-slika na nižu razinu, razinu isječaka, čime se dobivaju I-, P- i B-tipovi isječaka. Proizvoljni redoslijed isječaka (ASO- Arbitrary Slice Ordering) omogućuje isječcima da se prenose i



primaju izvan poretka očitavanja slike. Time se poboljšavaju nedostaci većeg kašnjenja u prijenosu podataka u videokonferencijama i internet aplikacijama. Redundantni isječci su također poželjni kod zaštite podataka od pogrešaka, te su ovom normom i oni predviđeni. Ovi alternativni podaci se mogu upotrijebiti za ispravljanje pogrešaka u prijenosu kod bilo kojeg makrobloka.

Osim I, P i B isječaka, H.264 normom definirani su SP i SI isječci. Kod SP isječaka koristi se jednosmjerno nadomještanje pokreta. Oni koriste vremensku redundanciju za rekonstrukciju oštećenog isječaka čak i uz pomoć drugih referentnih isječaka. Kod SI isječaka uklonjena je samo prostorna redundancija te oni služe u rekonstrukciji odgovarajućeg u nizu SP isječaka.

Kada procjena pokreta nije efikasna, koristi se kodiranje unutar makrobloka (Intra Prediction) na način da se na temelju već kodiranih blokova (i puštenih u prijenos) izračunavaju vrijednosti uzoraka susjednog nekodiranog bloka određenom metodom te oduzimanjem tih vrijednosti od stvarnih vrijednosti trenutnog bloka samo ta razlika kodira i propusti u prijenos. S razlikom se obvezno šalje i oznaka metode izračunatih uzoraka bloka. Ovaj postupak kodiranja pomoću I- makroblokova koristan je kada slika sadrži velike jednolike pozadine koje su izvori velikih prostornih redundancija.

Točnost nadomještanja pokreta (Inter Prediction) je poboljšana s 1/2 piksela točnosti korištene u mnogim prijašnjim videokodecima na 1/4 piksela. H.264 norma podržava istu točnost od 1/4 piksela koja je korištena u posljednjim MPEG-4 videokodecima.

H.264 normom omogućeno je stvaranje i korištenje više referentnih slika. Ovime se povećava kompresija poboljšanjem procesa nadomještanja pokreta i povećava zaštita od pogrešaka zbog mogućnosti korištenja drugih referentnih slika u slučaju kada je jedna izgubljena. Jedan makroblok može koristiti do 8 referentnih slika (do 3 kod HDTV).

Izračunavanje razlike između uzoraka makroblokova na različitim vremenskim udaljenostima, između trenutne i referentne slike, postiže se usrednjavanjem s podesivim težinskim parametrima. Ti parametri mogu biti umetnuti unutar toka podataka ili ih dekodek može implicitno izvoditi iz vremenskih oznaka.

H.264 norma koristi jednostavnu 4x4 cjelobrojnu transformaciju. Dodatna 2x2 transformacija se primjenjuje na četiri CbCr DC koeficijenta. Intra 16x16 makroblokovi imaju dodatnu 4x4 transformaciju primjenjenu na šesnaest Y DC koeficijenata. Pojava blokova unutar slike i pojava kontura oko objekata u slici je smanjena kao rezultat korištenja manjih veličina blokova u H.264 normi.

Za kvantizaciju, norma H.264 koristi 52 faktora skaliranja s jednolikim korakom kvantizacije od oko 12,5%. Kvantizirani koeficijenti se očitavaju (skeniraju) od niskih frekvencija prema visokim na jedan od dva moguća načina.

Nakon kvantizacije i cik-cak skeniranja, norma H.264 koristi dvije vrste entropijskog kodiranja: kodiranje s promjenjivom dužinom riječi (VLC) i adaptivno

kontekstno binarno kodiranje (Context Adaptive Binary Arithmetic Coding-CABAC). Za sve, osim za koeficijente transformacije, norma H.264 koristi jednu univerzalnu VLC (UVLC) tablicu koja koristi set beskonačno proširivih kodnih riječi (Exponential Golomb). Umjesto kodiranja pomoću višestrukih VLC tablica kao kod ostalih videokodeka, kodiranje ide pridruživanjem vrijednosti iz jedne UVLC tablice čije se vrijednosti mijenjaju prema statističkim obilježjima podataka koji se kodiraju. Za koeficijente transformacije, koji koriste najveću širinu frekvencijskog pojasa, norma H.264 koristi adaptivno kontekstno kodiranje s promjenjivom dužinom riječi (Context Adaptive Variable Length Coding-CAVLC). Na temelju prethodno obrađenih podataka odabire se najbolja VLC tablica. Dodatna učinkovitost u kodiranju (5-10%) postiže se korištenjem CABAC-a. CABAC-om se stalno ažuriraju statistička obilježja dolaznih podataka, te se u realnom vremenu stalno podešava algoritam (proces nazvan kontekstno modeliranje).

### *Profili i razine*

Slično kao i kod drugih videokodeka, profilima je određena sintaksa (odnosno alati, kao npr. tipovi slika), a razinama su definirani različiti parametri (rezolucija (raster) slike, broj slika u sekundi, brzina prijenosa itd.) toka podataka, te način na koji ih dekodek mora obraditi.

### **Osnovni profil (Baseline Profile - BP)**

Osnovni profil je namijenjen za videosadržaj s progresivnom analizom slike kao što su videokonferencija, videosadržaj preko IP-a i mobilne aplikacije. Kod osnovnog profila uključeni su sljedeći alati:

- I i P tipovi isječaka, točnost nadomještanja pokreta od 1/4 piksela,
- UVLC (Universal variable-length coding) i CAVLC (Context-adaptive variable-length coding) entropijsko kodiranje,
- proizvoljni redosljed isječaka (ASO),
- fleksibilno raspoređivanje makroblokova (FMO),
- redundantni isječci (RS),
- struktura uzorkovanja 4:2:0, YCbCr komponentnog videosignala.

Važno je naglasiti da osnovni profil nije dio glavnog profila. Kod mnogih rješenja koriste se svi alati osim ASO ili FMO alata. Ova rješenja su dio glavnog profila i mnogo su jednostavnija u primjenama.

### **Prošireni profil (Extended Profile - XP)**

Namijenjen je za streaming mobilnih i internet aplikacija i kod njega su uključeni sljedeći alati:

- B, SP i SI vrste isječaka,

- razvrstavanje podataka prema njihovom značaju radi veće zaštite unutar isječaka (Slice data partitioning),
- nadomještanje pomaka s podesivim težinskim parametrima (Weighted prediction).

### Glavni profil (Main Profile - MP)

Glavni profil je namijenjen za širok raspon televizijskih aplikacija. Dodatni alati uz alate osnovnog profila su:

- kodiranje slike s proredom,
- B vrste isječaka slike,
- CABAC entropijsko kodiranje,
- nadomještanje pomaka s podesivim težinskim parametrima,
- 10 i 12 bitni uzorci YCbCr komponentnog digitalnog videosignala, strukture uzorkovanja 4:2:2 i 4:4:4,
- ASO, FMO i RS nisu podržani.

### Vršni profil (High Profile - HP)

Nakon što je početna specifikacija dovršena, dodano je proširenje vezano uz poboljšanje kvalitete slike ( Fidelity Range Extension-FRExt). Proširenje je rezultiralo sa četiri dodatna profila:

- High Profile (HP) - podržava odabir veličine bloka 4x4 ili 8x8 uzoraka luminantnog signala te odabir tablice faktora skaliranja prema vrijednosti frekvencija za pojedine koeficijente,
- High 10 Profile (Hi10P) - podržava 9 ili 10 bitne uzorke komponentnog digitalnog videosignala YCbCr sa strukturom uzorkovanja 4:2:0,
- High 4:2:2 Profile (Hi422P),
- High 4:4:4 Profile (Hi444P) - podržava 11 ili 12 bitne uzorke komponentnog digitalnog videosignala YCbCr ili RGB, strukture uzorkovanja 4:4:4.

Unutar petnaest razina definirane su rezolucije slike (raster) od 176x144 do 4096x2048 piksela, frekvencije izmjene slika od 15 slika/s do 120 slika/s, maksimalan broj makroblokova u sekundi od 1485 do 983040, maksimalni broj vektora pomaka između dva susedna makrobloka od 16 do 32, maksimalni broj referentnih slika od 4 do 9 te brzina prijenosa toka podataka od 64 kBit/s do 240 Mbit/s.

## 4. ZAKLJUČAK

Gotovo dva desetljeća MPEG norme definiraju postupke kodiranja i prijenosa digitalnog videosignala u području videotelefonijske i videokonferencijske, te pohranjivanja na CD medije. Unazad nekoliko godina prijenos se počinje znatnije ostvarivati putem zemaljske radiodifuzije, satelitskim prijenosom i kabelskom distribucijom, zatim putem interneta, najprije preuzimanjem videosadržaja kao datoteka, a zatim streamingom putem javnih i privatnih, fiksnih i mobilnih telekomunikacijskih mreža.

Potpuno nove mogućnosti javljaju se upotrebom mobilnih telefona kao fotoaparata, videokamera, reproduktora videosadržaja, navigacijskih uređaja, pa sve do pretraživanja interneta i korištenja usluga prijenosa video i audiosignala IP tehnologijom. Razvoj MPEG normi omogućio je pohranjivanje filmova i drugih videosadržaja na DVD i Blue-Ray Disc medije. Nagli porast potrošačkog tržišta koje traži vrhunske odlike proizvoda baziranih na videotehnologiji predstavlja značajne izvore prihoda za televizijske i telekomunikacijske kompanije, te za proizvođače televizijske, profesionalne i potrošačke (zabavne) elektronike i softvera. Ključnu ulogu na ovakvom tržištu imaju norme. Normizacija postupaka kodiranja, prijenosa i pohranjivanja videosignala (formata) ima brojne koristi u odnosu na nenormirane, vlasničke formate. Neke od njih su pojednostavljeno korištenje uređaja raznih proizvođača (npr. korištenje istih normi izdavačkih kompanija i proizvođača uređaja), uvođenje novih tehnika kodiranja koje mogu povrijediti postojeće rješenje zaštićeno patentom, dok je primjena tehnika i algoritama ugrađenih u određenu normu jasno regulirana cijenom prava korištenja. Unatoč stalnim raspravama o koristima videokodeka bez naknade za autorska prava nasuprot normiranim industrijskim videokodecima, normizacija videokodeka je vrlo važna za mnoge industrije.

## 5. LITERATURA

- [1] Fischer, Walter, *Digital Video and Audio Broadcasting Technology* (Second edition), Springer-Verlag, Berlin, Heidelberg, 2008.
- [2] Jack, Keith, *Video Demystified* (Fifth edition), Elsevier Inc., 2007.