

ANALIZA NEKIH POBOLJŠANJA PERFORMANSI KODERA KOD STANDARDA H.264/AVC

Zoran Miličević, *Uprava za vezu i informatiku u Beograd*
Zoran Bojković, *Saobraćajni fakultet u Beogradu*

Sadržaj - H.264/AVC je najnoviji video kodni standard koji značajno poboljšava performanse kompresije video signala u odnosu na postojeće standarde. Tokom 2004. godine ovom standardu je pridodat FRExt amandman koji poboljšava efikasnost kodovanja sadržaja. Eksperimentalni rezultati dobijeni simulacijom su bazirani na komparaciji dva različita metoda entropijskog kodovanja u okviru H.264/AVC standarda.

1. UVOD

H.264/AVC [1, 8] je najnoviji zajednički standard ITU Video Coding Expert Group (VCEG) i ISO/IEC Moving Picture Experts Group (MPEG). Osnovni cilj H.264/AVC je poboljšanje performansi kompresije video signala. H.264/AVC standardu prethodili su MPEG-1, MPEG-2, MPEG-4, H.261 i H.263. Tokom 2004. godine H.264/AVC dopunjen je sa novim amandmanom takozvanim Fidelity Range Extensions (FRExt, Amandman I) koji predstavlja ponovo veću efikasnost kodovanja, potencijalno čak sa odnosom 3:1 za neke ključne aplikacije.

Rad je koncipiran na sledeći način. Prvo poglavlje je uvodno. Drugo u najkraćem opisuje arhitekturu H.264/AVC standarda. Treće objašnjava novi FRExt amandman, profile i aplikacije. Četvrto detaljno predstavlja neke simulacione rezultate koji se odnose na analizu nekih poboljšanja performansi kodera. Peto poglavlje daje zaključna razmatranja.

2. ARHITEKTURA H.264/AVC

Mrežni apstraktni sloj NAL (*Network Abstraction Layer*) je specificiran za efikasno predstavljnje kodovanih video podataka, dozvoljavajući laku integraciju u različite protokole i multipleksne arhitekture. NAL obezbeđuje i informacije o zaglavlju i odgovarajući format za isporuku kodiranih video podataka pomoću transportnih nivoa ili medijuma za čuvanje. Osnovni koncepti NAL su: *NAL jedinica*, *tok bajtova*, *format paketa* koji se koristi od strane NAL jedinica, *set parametara* i *pristupne jedinice*.

Kodirani video podatak je organizovan u *NAL jedinice*. Prvi bajt svake NAL jedinice je bajt zaglavlja koji sadrži indikaciju tipa podataka u NAL jedinici, dok preostali bajtovi sadrže podatke o informaciji čiji tip je prikazan u zaglavlju. Za korišćenje u sistemima (H.320, MPEG-2 | H.222.0) H.264/AVC specifikacija definiše *format toka bajta*. U formatu toka bajta svaka NAL jedinica je prefiksovana sa specifičnim uzorkom od tri bita nazvanim *startni kodni prefiks*. Granice NAL jedinice mogu se tada odrediti pomoću pretraživanja kodiranog podatka za jedinstven slučaj startnog kodnog prefiksa. Korišćenje emulacije prevencije bajtova garantuje da su startni kodni prefiksi jedinstveno određeni početkom nove NAL jedinice. Kod drugih sistema (IP/RTP) kodirani podatak se prenosi u paketima koji su uokvireni

sistemske transportnim protokolom. Određivanje granica NAL jedinica unutar paketa može se uspostaviti bez korišćenja startnog kodnog prefiksa. NAL jedinice su klasifikovane u *VCL* i *ne-VCL*. VCL NAL jedinice sadrže podatke koji predstavljaju vrednosti uzoraka video slike, dok ne-VCL NAL jedinice sadrže neke pridružene dodatne informacije kao što su parametarski setovi i suplementarne povećane informacije.

Parametarski set je zamišljen da sadrži informaciju od koje se očekuje da se retko menja i nudi dekodiranje velikog broja VCL NAL jedinica. Postoje dva tipa parametarskih setova. *Parametarski set sekvence* koji je primenjen na seriju uzastopnih kodiranih video slika nazvanih kodirani video niz i *parametarski set slike* koji je primenjen za dekodiranje jedne ili više pojedinačnih slika u kodiranom video nizu.

Skup NAL jedinica specifičnog oblika odnosi se na *pristupnu jedinicu*. Dekodiranjem svake pristupne jedinice dobija se jedna dekodirana slika. Svaka pristupna jedinica ima sledeću strukturu: razdvajач pristupne jedinice, dodatna informacija (*Supplemental Enhancement Information*), primarna kodirana slika, redundantna kodirana slika, kraj sekvence i kraj toka.

Kodirana video sekvenca sadrži seriju pristupnih jedinica koje su sekvencijalne u toku NAL jedinica i koristi samo jedan parametarski set sekvence.

Video kodni sloj VCL (*Video Coding Layer*) je dizajniran na blok-baziranom hibridnom video koderu u kome svaka kodirana slika je predstavljena *makroblokovima* piksela.

Kodirana video sekvenca u H.264/AVC sastoji se od niza kodiranih slika [2]. Kodirana slika može predstavljati svaki deo okvira ili jedno polje. Video-kolor prostor korišćen od strane H.264/AVC razdvaja boju u tri komponente *Y*, *Cb* i *Cr*. Komponenta *Y* je *luma* i predstavlja osvetljenje. Dve *chroma* komponente *Cb*, *Cr* predstavljaju stepen devijacije boje od sive prema plavoj i crvenoj. Zato što je ljudsko čulo vida osetljivije na *luma* komponentu, H.264/AVC koristi strukturu sastavljenu od odbiraka u kojoj *chroma* komponenta ima jednu četvrtinu od broja odbiraka *luma* komponente (4:2:0) Svaki odbirak je kodovan sa 8 bita. Slika je izdeljena u *makroblobove* fiksne veličine od kojih svaki pokriva pravougaonu površinu slike od 16x16 odbiraka *luma* komponente i 8x8 odbiraka svake od dve *chroma* komponente. *Blok* je najmanja kodirana jedinica i sadrži podatke o kojeficijentima transformacije za predikciju grešaka.

Sekvence makroblokovia (*Slices*) se procesiraju da bi se analizirao raster kada se ne koristi fleksibilno raspoređivanje makroblokovia FMO (*Flexible Macroblok Ordering*). Slika može biti izdeljena u jednu ili više traka. FMO određuje

način na koji je slika izdijeljena u trake i makro-blokove korišćenjem koncepta grupe traka (*slice group*). Svaka grupa traka je skup makroblokova definisanih mapiranjem grupe makroblokova, koji je specificiran sadržajem parametarskog seta slike i informacijom iz zaglavlja traka. Makroblok za mapiranje grupe traka sadrži identifikacioni broj grupe traka za svaki makroblok u slici. Svaka grupa traka može biti izdijeljena u jednu ili više traka, tako da je traka niz makroblokova unutar iste grupe traka koja je procesirana da bi se analizao raster unutar skupa makroblokova određene grupe traka.

Svaka traka može biti kodirana korišćenjem različitih tipova koda: I, P, B i dva nova tipa SP i SI [3].

Da bi dostigao visoku efikasnost kodovanja, H.264/AVC standard dozvoljava koderu da napravi neku od sledećih odluka prilikom kodovanja okvira:

- da kombinuje dva polja zajedno i koduje ih kao jedno polje okvira (*frame mode*),
- da ne kombinuje dva polja zajedno i koduje ih kao odvojena kodna polja (*field mode*) i,
- da kombinuje dva polja zajedno i komprimuje ih kao jedno polje, ali kada koduje okvir da izdijeli dva vertikalna susedna bloka u svaki par od dva polja ili okvir makroblokova pre kodovanja istih.

Izbor između tri opcije može se napraviti adaptivno za svaki okvir u sekvenci. Izbor između prve dve opcije je nazvan PAFF (*Picture/Adaptive Frame/Filed*) kodovanje.

Odluka koda polja/okvira takođe može biti napravljena nezavisno za svaki vertikalni par makroblokova (16x32 luma regiona) u okviru. Ova kodna opcija naziva se MBAFF (*Macroblock-Adaptive Frame/Filed*) kodovanje[7].

U H.264, ako je makroblok kodovan u *intra modu*, predikcija je izvedena na bazi predhodnog kodovanja i rekonstruisana oko dostupnih odbiraka. Predikcija je tada oduzeta trenutnom bloku koji je ranije kodovan. Intra predikcija luma makrobloka može se formirati pomoću: (1) jedna predikcija za ceo 16x16 makroblok, sa 4 dostupna moda (vertikalni, horizontalni, DC, ravni); (2) šesnaest pojedinačnih predikcija sa 4x4 bloka, sa 9 dostupnih modova (DC i 8 usmerenih). *Chroma* intra predikcija podržava jedan tip predikcije za svaki 8x8 region. U ovom slučaju, samo 4 moda (vertikalni, horizontalni, DC, ravni) su dostupna.

Slično drugim standardima, *inter-prediktivno kodovanje* u H.264/AVC se oslanja na blok-baziranu kompenzaciju pokreta. Nove odlike uključuju korišćenje promenljive veličine bloka i vektora pokreta sa preciznošću 1/4 piksela bi-direkciono predikcije sa većim brojem inter-kodovanih modula.

Makroblok i pod-makroblok delovi. Interkodovani 16x16 makroblokovi mogu se segmentirati u manje makroblok delove veličine 16x8, 8x16 ili 8x8. Dodatno 8x8 delovi, takođe poznati kao pod-makroblokovi, se mogu dalje segmentirati u pod-makroblovske delove veličine 8x4, 4x8, i 4x4.

Podpiksel vektori pokreta. Svaki deo u interkodovanom makrobloku je prediktovan iz polja iste veličine referentne slike na osnovu vektora pokreta. H.264/AVC podržava preciznost od 1/4 piksela za vektore pokreta. U *luma*, interpolacije uzoraka na pola pozicije su generisane na srednjem elementu kašnjenja FIR filtra 6-tog reda sa koeficijentima 1/32 (1, -5, 20, 20, -5, 1). Nakon toga, pikseli na četvrtini pozicije se dobijaju na srednjim odabiranjem cele i polovine pozicije. Odgovarajući 1/8 chroma odbirci su dobijeni pomoću linearne interpolacije.

Modovi inter predikcije. H.264 standard podržava četiri moda inter predikcije: prediktivni, biprediktivni, skip i direktni mod [4].

Višestruke referentne slike. Makroblokovi u trenutnoj slici se mogu prediktovani od višestrukih referentnih slika. Višestruke referentne slike su najkorisnije za sekvence sa haotičnim pokretom. Ova nova osobina uglavnom povećava računarsku složenost koda i zahteva dodatno korišćenje memorije. Većina dobiti primenom višestrukih referentnih slika se postiže korišćenjem 2-5 referentnih okvira. Tako je pokazano da korišćenjem 5 referentnih okvira za predikciju, smanjuje se 5-10 % bitski protok u odnosu na korišćenje samo 1 referentnog okvira.

H. 264/AVC standard je jedinstven u korišćenju prostorne 4x4 celobrojne transformacije kao aproksimacije diskretne kosinusne transformacije DCT. Korišćenjem 16-bitne aritmetike u izračunavanju transformacija, smanjuje se značajno složenost izračunavanja.

Analiza (skeniranje) omogućava da se korišćenjem dva načina skeniranja zig-zag i naizmeničnog, 2D koeficijenti transformacije konvertuju u 1D sekvencu za kvantizaciju i kodovanje.

Kvantizacija koeficijentata transformacije kod H.264/AVC se izvodi sa logaritamskom kontrolom veličine koraka, proširenim opsegom veličine koraka i manjom veličinom koraka za chroma komponente.

Suprotno drugim kodnim standardima, H.264/AVC podržava dva metode entropijskog kodovanja: binarno aritmetičko adaptivno CABAC (*Context-Adaptive Binary Arithmetic Coding*) i adaptivno kodovanje sa promeljivom kodnom reči CAVLC (*Context-Adaptive Variable Length Coding*). Kod CABAC koda, modelovanje konteksta i intersimbolske redukcije obezbeđuje uslovne verovatnoće simbola[5]. Aritmetički koderi su posebno korisni za verovatnoću simbola veću od 0.5, koja se često javlja kada je korišćeno modeliranje konteksta. Deljenje bita u odnosu na minimum 1 bit za VCL se može koristiti u slučaju aritmetičkih koda. Adaptivnost je postignuta uzimanjem u obzir kumulativnih verovatnoća već kodiranih simbola, koji dovodi do boljeg podešavanja aritmetičkog koda za trenutnu statistiku simbola. Kao primer, moguće je da statistika varira za različite sekvence i za tokove bita. CAVLC je bio svrsishodan za poboljšanje performansi VCL bez promene kompleksnosti CABAC-a. CAVLC dodaje modele konteksta samo na koeficijente transformacije kodovanja. Drugi elementi, kao što su vektori pokreta, tipovi makrobloka, su kodirani sa exp-Golomb kodovima.

H.264/AVC standard uključuje deblocking filter u petlji dekodovanja. Deblocking filter je potpuno definisan i mora uzeti u obzir i koder i dekodek. Kao filter u petlji, on poboljšava kvalitet dekodovane slike, pokazujući superiorne performanse nasuprot post-procesnim filterima. Deblocking filter operiše sa vertikalnim i horizontalnim ivicama između blokova sa 4x4 piksela unutar prediktivne petlje za pomeranje predmeta uzrokovanog sa kvantizacijom. Visokom adaptacijom na sadržaj, glavna procedura filtriranja pomera blokirane predmete. Na nivou traka, globalno filtriranje je podešeno za pojedinačne karakteristike video sekvence. Na nivou ivice, filtriranje je podešeno zavisno od inter/intra moda odluke, vektora pokreta i kodiranih ostataka. Konačno, na nivou piksela, kvantizer zavisi od praga definisanog za isključenje procesa filtriranja za svaki pojedinačni odabirak.

3. FRExt AMANDMAN, PROFILI I APLIKACIJE

Apostrofirajući porebu za najzahtevnijim aplikacijama kao što su aplikacije doprinosa u sadržaju, raspodele sadržaja, studiskog editovanja i post-procesne aplikacije, a u kontinuitetu sa projektom, JVT (Joint Video Team) je dodao novu ekstenziju u karakteristike originalnog standarda. Ovaj napor iziskivao je godinu dana za kompletiranje - počevši od Maja 2003 da bi konačana specifikacija bila kompletirana u Julu 2004. godine, dok je period ispitivanja završen u Avgustu 2004. godine. Ove ekstenzije, u originalu nazvane "profesionalne" ekstenzije, su preimenovane u "Fidelity Range Extensions" (FRExt) [9].

U periodu projektovanja FRExt amandmana, JVT je mogao da se ponovno vrati na ispitivanje nekoliko prethodnih tehničkih predloga koji nisu bili obuhvaćeni početnim standardom zbog vremenskih ograničenja i neizvesnosti oko dobiti. Sa dodatkom vremena datog za projekat oko ekstenzija, JVT-u je omogućeno da obuhvati neke od tih karakteristika u novim ekstenzijama i to:

- podršku za adaptivnu veličinu bloka za rezidualnu prostornu frekventnu transformaciju,
- podršku koderu - specifične percepcije - baziranom na kvantizacionim skalirajućim matricama i,
- podršku efikasnijem smanjenju gubitaka predstavljenim na specifičnim regionima u video sadržaju.

Profili definišu skup kodnih alata i algoritama koji se mogu koristiti pri generisanju odgovarajućeg toka bita. H.264/AVC standard je definisao tri profila [6]: BL (*Base Line*), MP (*Main Profile*) i EP (*Extended Profile*).

Za sve profile H.264/AVC definisani su i skup nivoa i podnivoa.

FRExt projekat je uveo četiri nova profila koji su zajedno nazvani *High* profili:

- High profil (HP), podržava 8-bitni video sa 4:2:0 odabiranjem,
- High 10 profil (Hi10P) podržava 4:2:0 video sa do 10 bita za predstavljanje tačnosti po odabirku,
- High 4:2:2 profil (H422P), podržava do 4:2:2 hroma odabiraka i do 10 bita po odabirku i,
- High 4:4:4 profil (H444P), podržava do 4:4:4 hroma odabiraka i do 12 bita po odabirku.

Svi navedeni tipovi High profila podržavaju sve karakteristike prethodnog Glavnog (Main) profila i dodatnu podršku za adaptivnu transformaciju veličine-bloka i perceptualne kvantizacione skalirajuće matrice.

Aplikacije sa Baseline (Osnovnim) profilom su video konferencija i video telefonija, Main (Glavnim) profilom broadcast video i Extended (Proširenim) profilom streaming video. High profil je inkorporiran u nekoliko važnih aplikativnih specifikacija, posebno obuhvatajući HD-DVD specifikaciju DVD Foruma, BD-ROM Video specifikaciju Blu-ray Disc Association, DVB (digital video broadcast) standard za Evropsku broadcast televiziju.

4. REZULTATI SIMULACIJE

U ovom delu rada predstavljenu su eksperimentalni rezultati do kojih se došlo korišćenjem simulacionog softvera JVT verzije JM 9.0 u koju je implementiran i Fidelity Range Extensions (FRExt) amandman.

Za sva eksperimente korišćene su tri test sekvence (foreman, bus, stefan) u CIF formatu (352x288) i 4:2:0 High profilu. Sva testiranja su izvedena sa CAVLC i CABAC metodom entropijskog kodovanja. Treba napomenuti da se kod svih testiranja koriste sledeći ulazni parametri: jedna B slika, pet referentnih slika, RD optimizacija, Hadamard-ova transformacijom, prostorno usmeravanje uz primenu kodovanja na bazi okvira (frame moda), odnosno kodovanja na bazi polja (field moda) i na kraju adaptivnog kodovanja okvira i polja (adaptive frame/field mod). Zavisno od izabranog moda, a radi upoređenja i analize izlaznih rezultata kod obe metode entropijskog kodovanja uzeti su u obzir kao ključni faktori odnos signal/šum za luma komponentu slike i bitski protoci.

Tabela 1 (a) prikazuje simulacione rezultati kada se koristi kodovanje na bazi okvira (frame mod) i oba metoda entropijskog kodovanja. Sa stanovišta bitske brzine uočljivo je da CABAC metoda daje nešto bolje rezultate od CAVLC metode (vidljivo je smanjenje bitske brzine u opsegu od 3,4 do 6 %). Prisutna je i činjenica da je došlo i do neznatnog povećanja (odnosno smanjenja kod *foreman* test sekvence) odnosa signal/šum, kao prateće pojave smanjenja bitske brzine kod izabranih CIF test sekvenci.

Kada se umesto kodovanja na bazi okvira koristi kodovanje na bazi polja (field mod) kao u tabeli 1 (b) uočljivo je da redukcija bitske brzine kod *stefan* test sekvence preko 8 %, dok je kod *bus* test sekvence to 5,7 %, a kod *foreman* test sekvence 4,7 %. Sa stanovišta odnosa signal/šum vidljive su znatno bolje vrednosti kod CABAC nego kod CAVLC metode.

U tabeli 1 (c) prikazani su rezultati koji su dobijeni kada se koristi kodovanje na bazi adaptivnog kodovanja okvira i polja (adaptive frame/field). Vidljivo je da je kao i kod prethodna dva eksperimenta i dalje prisutna redukcija bitske brzine kao i nešto bolji odnos signal/šuma kod CABAC metode u odnosu na drugu metodu entropijskog kodovanja.

Tabela 1. a, b, c. Eksperimentalni rezultati kod CABAC i CAVLC metode entropijskog kodovanja.

a)

CIF, 1 B-pic., 5 Ref. pic., Spatial direct, RDOpt., Hadamard, Frame				
Sequence	CAVLC		CABAC	
	Y(SNR)	Bit-rate (kbps)	Y(SNR)	Bit-rate (kbps)
Foreman	37,45	826,40	37,42	776,88 (-5,99%)
Bus	35,01	1938,48	35,04	1872,00 (-3,42%)
Stefan	35,81	1973,44	35,81	1877,84 (-4,84%)

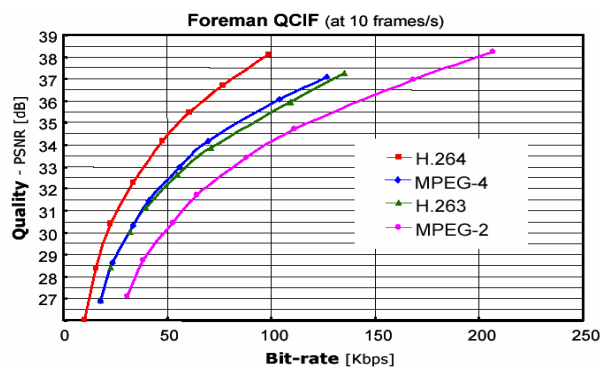
b)

CIF, 1 B-pic., 5 Ref. pic., Spatial direct, RDOpt., Hadamard, Field				
Sequence	CAVLC		CABAC	
	Y(SNR)	Bit-rate (kbps)	Y(SNR)	Bit-rate (kbps)
Foreman	36,40	798,64	36,41	760,96 (-4,71%)
Bus	33,91	2058,56	33,97	1940,56 (-5,73%)
Stefan	34,82	2170,56	34,88	1982,16 (-8,67%)

c)

CIF, 1 B-pic., 5 Ref. pic., RDOpt., Spatial direct, Hadamard, Adaptive Frame/Field				
Sequence	CAVLC		CABAC	
	Y(SNR)	Bit-rate (kbps)	Y(SNR)	Bit-rate (kbps)
Foreman	37,46	834,08	37,48	780,64 (-6,40%)
Bus	34,99	1938,88	35,06	1882,96 (-2,88%)
Stefan	35,82	1978,88	35,84	1875,92 (-5,20%)

Slika 1. je prikazuje poređenje H.264/AVC sa drugim standardima. Može se uočiti relativno brzo dostizanje visokog kvaliteta kodovanog video sadržaja (PSNR od 38 dB) sa manjom bitskom brzinom (100 kb/s) [7].



Sl 1. PSNR performanse H.264/AVC u poređenju sa drugim video kodnim standardima[7].

5. ZAKLJUČAK

Na osnovu analize rezultata do kojih se došlo simulacijom sa različitim video test sekvencama nepobitno se može zaključiti da je CABAC metod entropijskog kodovanja u pogledu performansi bolja od CAVLC na šta ukazuje i međusobno poredjenje nekih ključnih faktora kao što su odnos siglan/šum i bitski protoci nakon kodovanja odnosno dekodovanja. Sve nas ovo navodi na zaključak, a u isto vreme i potvrđuje prehodne konstatacije, da H.264/AVC predstavlja napredak u standardizaciji video kodnih tehnologija nudeći efikasniju predikciju pokreta, korišćenje blokova male veličine, adaptivno debloking filtriranje u petlji i poboljšanje metoda entropijskog kodovanja, što sve zajedno

obežbeđuju otprilike 50% uštede na bitskom protoku za jednaki perceptualni kvalitet.

LITERATURA

- [1] K.R.Rao, Z.S.Bojkovic, D.A.Milovanovic, *Multimedia communication systems: techniques, standards and networks*, Prentice Hall, New York, 2002.
- [2] Joint Video Team of ITU-T and ISO/IEC JTC 1, *Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification* (ITU-T Rec. H.264 [ISO/IEC 14496-10 AVC), Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, JVT-G050, March 2003.
- [3] M. Karczewicz and R. Kurceren, "The SP and SI frames design for H.264/AVC," *IEEE Transactions on Circuits and Systems for Video Technology*, vol.13, no.7, pp.637-644, 2003.
- [4] M Flierl and B. Girod, *Video coding with supermodes motion-compensated signals. Applications to H.264 and Beyond*, vol. 70 of Series in Engineering and Computer Science, Kluwer Academic Publishers, Boston, Mass, USA, 2004.
- [5] D. Marple, H. Schwrts, and T. Wiegand, "Context-based adaptive binary arithmetic coding in the H.264/AVC video compression standard", *IEEE Transactions on Circuits and System for Video Technology*, vol.13, no.7, pp.620-636, 2003.
- [6] T. Weigand, G. Sullivan, G. Bjontegaard and A. Luthra, "Overview of the H.264/AVC Video Coding Standard", *IEEE Transactions on Circuits and System for Video Technology*, vol.13, no.7, pp.560-576, 2003.
- [7] Z. Miličević, D. Milovanović, Z. Bojković "Aplikacija i arhitektura standarda H.264/AVC", *XII Telekomunikacioni forum TELFOR*, Beograd, 2004.
- [8] A. Puri ,H. Chen, A. Luthra, "Video Coding using the H.264/MPEG-4 AVC compression standard", *Signal Processing: Image Communication 19*, pp.793-846, 2004.
- [9] G. Sullivan, P. Topiwala and A. Luthra, "Overview of the H.264/AVC Advanced Video Coding Standard: Overview and introduction to the Fidelity Range Extensions", *SIPE Conference on Applications of Digital Image Processing*, August, 2004.

Abstract - H.264/AVC is newest video coding standard which has achieved a significant improvement coding performance relative to existing standards. During 2004 FRExt amendment was added to this standard for improvement coding content efficiency. Simulation results compare two different methods of entropy coding included with H.264/AVC standard.

ANALYSIS SOME CODER PERFORMANCE IMPROVEMENT IN H.264/AVC STANDARD

Z.Milicevic, Z.Bojkovic