

# Modelovanje i procesiranje B frejmova kod H.264/AVC standarda

Zoran Miličević, Zoran Bojković

**Sadržaj** — H.264/AVC video kodni standard je značajno unapredio efikasnost kodovanja, ali i povećao kompleksnost samog koda. Da bi se redukovala ova kompleksnost, u radu je predložen algoritam koji kombinuje dva metoda za procesiranje B slike: rani inter skip i selektivni intra mod odluke. Rezultati eksperimenta pokazuju da je vreme potrebno za proces kodovanja redukovano 35-42% kroz proces rane identifikacije makroblokova koji će biti skipovani tokom procesa kodovanja, kao i kroz smanjenje broja kandidovanih modova.

**Ključne reči** — generalizovana B slika, H.264/AVC standard, Main profil, rate-distortion optimizacija, rana inter skip predikcija, selektivna intra predikcija.

## I. UVOD

H.264/AVC je video kodni standard koji je zajednički razvijen od strane eksperata iz JVT (Joint Video Team), odnosno ITU-T VCEG i ISO/IEC MPEG komiteta za standardizaciju [1], [2], [3], [4]. Ciljevi ovih napora u procesu standardizacije su bili povećanje efikasnosti kompresije i lakoća u mrežnom predstavljanju za interaktivne (video telefonija) i ne-interaktivne aplikacije (broadcast, streaming, memorisanje, video na zahtev). H.264/AVC video kodni standard može u značajnoj meri da poboljša efikasnost kompresije u poređenju sa prethodnim standardima. Zbog ovog poboljšanja efikasnosti kompresije, kao i povećanja fleksibilnosti kodovanja i prenosa, H.264/AVC ima i neophodan potencijal da omogući nove servise preko različitih tipova mreža. Međutim, dobici u performansama standarda imaju za cenu povećanje kompleksnosti izračunavanja. Poboljšanje efikasnosti kodovanja povećava kompleksnost koda. U skladu sa tim, H.264/AVC koristi različite metode za redukciju kompleksnosti implementacije.

Poboljšanje performansi kodovanja proističe uglavnom iz dela za predikciju. Intra predikcija značajno poboljšava performanse kodovanja koristeći prostornu redundansu između susednih makroblokova. Sa druge strane inter predikcija, koja koristi vremensku redundansu između frejmova, poboljšana je sa estimacijom pokreta sa preciznošću od  $\frac{1}{4}$  za vektore pokreta, blokovima

promenljive veličine, višestrukim referentnim slikama i poboljšanim prostorno/vremenskim direktnim modom.

Za određene aplikacije, H.264/AVC definiše sedam *Profila* [5]. Profili definišu skup kodnih alata i algoritama koji se mogu koristiti pri generisanju odgovarajućeg toka bita. Takođe, za svaki profil H.264/AVC definiše i skup nivoa i podnivoa.

Main (Glavni) profil pored podrške za I slike, P slike i entropijsko adaptivno kodovanje sa promenljivom kodnom reči (Context Arithmetic Variable Length Coding, CAVLC [6]) omogućava i podršku za interlaced video, inter kodovanje sa korišćenjem B slika i težinske predikcije, kao i binarno aritmetičko adaptivno entropijsko kodovanje (Context Adaptive Binary Arithmetic Coding, CABAC [7]). Aplikacije podržane sa Main profilom su: broadcast TV, video memorisanje i studijska distribucija.

Zahvaljujući svojim unapređenim performansama H.264/AVC je privukao značajnu pažnju industrije. Pored klasičnih aplikacijskih domena kao što su video konferencija i broadcasting TV sadržaja (satelitski, kablovski i zemaljski), sa poboljšanjem sposobnosti kompresije, H.264/AVC je obuhvatio i nove servise i tako otvorio nove tržišne mogućnosti u industriji. Drugi domen koji je privukao dosta interesovanja industrije je prenos i memorisanje sadržaja *Visoke Definicije* [8].

H.264/AVC se može primeniti i kod: korisničke opreme (npr. digitalne kamere, celularni telefoni...), videa preko IP mreža, digitalnog broadcasting-a visoke definicije (preko zemaljskih ili satelitskih kanala) i digitalnih sistema za memorisanje (DVD visoke rezolucije). Danas H.264/AVC omogućava i prenos IPTV preko digitalnih pretplatničkih linija i pasivnih optičkih mreža.

Rad je organizovan na sledeći način. Prvo poglavlje je uvodno. Drugo poglavlje sadrži detalje vezane za rate-distortion optimizaciju i značaj koncepta generalizovane B slike. Treće poglavlje, detaljno objašnjava predloženi kombinovani intra/inter algoritam. U četvrtom poglavlju prikazani su i obrazloženi eksperimentalni rezultati. Poslednje poglavlje sadrži zaključna razmatranja.

## II. RATE-DISTORTION OPTIMIZACIJA I KONCEPT GENERALIZOVANE B SLIKE

### A. Značaj Rate-Distortion Optimizacije (RDO)

Rate-Distortion Optimizacija (RDO) je veoma efikasan alat koji je prihvaćen od strane H.264/AVC kodeka za određivanje moda kodovanja za makroblokove i koji donosi veću efikasnost kodovanja. Međutim, zbog preciznijeg proračuna trenutne distorzije makroblokova i

Zoran Miličević, Uprava za telekomunikacije i informatiku GŠ VS u Beogradu, Raska 2, 11000 Beograd, Srbija (telefon: 381-64-1191125; e-mail: mmilicko@eunet.yu).

Zoran Bojković, Univerzitet u Beogradu, 11000 Beograd, Srbija (e-mail: zsbojkovic@yahoo.com).

kodovanih bita za svaki mod kodovanja, mod odluke koji se bazira na RDO sadrži veću kompleksnost izračunavanja.

Da bi se pronašli najbolji parametri za svaki makroblok, H.264/AVC referentni softver koduje sve kombinacije parametara i za svaku kombinaciju izračunava brzinu i distorziju za dati makroblok. Naime, kod sekvence kodovane H.264/AVC koderom, svaki makroblok može biti kodovan u jednom od velikog broja modova koji se vrednuju pre nego što se odgovarajući mod izabere. Ovo znači da Joint Model (JM) referentni koder izračunava tzv. R-D cost svih mogućih opcija kodovanja i bira mod kodovanja za dati makroblok. Izabrani mod ima minimalnu R-D cost. Drugim rečima, da bi postigao dobre rate-distortion performanse, RDO proces izbora moda procenjuje distorziju i bitksu brzinu svakog kandidovanog moda pre izbora moda koji minimizira rate-distortion cost funkciju za trenutni makroblok. Međutim, takav metod za mod odluke sa praktične tačke gledišta pokazuje izvesne nedostatke. Za zadati makroblok, koder mora imati informacije o zahtevanim bitima i rezultujućoj distorziji trenutnog moda kodovanja da bi izabrao najbolji mod kodovanja. Međutim, ova informacija je dostupna samo posle završetka procesa kodovanja. Zbog toga, sadašnji referentni softver sadrži kompleksan proces za pronalaženje samo najboljeg moda kodovanja [9].

### B. Koncept generalizovane B slike

U poređenju sa prethodnim video kodnim standardima koncept B slike kod H.264/AVC je generalizovan na nekoliko načina [8], [10]. Koncept generalizovanih B slika kod H.264/AVC omogućava da se svaki proizvoljan par referentnih slika iskoristi za predikciju svakog regiona. Za ovu svrhu za B isečke uspostavljaju se dva različita načina indeksiranja bafera sa više slika koji su nazvani prva ("list 0") i druga ("list 1") lista referentnih slika. U zavisnosti koja referentna slika je korišćena za formiranje signala predikcije, kod B isečaka istaknuta su tri različite tipa inter predikcije slike: "list 0", "list 1" i bi-prediktivni, gde poslednji koristi superpozicije signala predikcije liste 0 i liste 1 i koji je ključna karakteristika koju obezbeđuju B isečci. Sa sličnim deljenjem kao kod P isečaka, mogu se posebno izabrati tri različita tipa inter predikcije B isečaka za svaku makroblok ili pod-makroblok partciju. Takođe, B-isečci makroblokova ili pod-makroblokova mogu se kodovati u takozvanom *direktnom modu* bez potrebe za prenosom nekih dodatnih informacija o pokretu. Ako rezidualni podaci predikcije nisu prenešeni za direktno-kodovani makroblok nastaje skipovanje.

## III. KOMBINOVANI ALGORITAM ZA INTRA/INTER PREDIKCIJU

### A. Skip algoritam za inter predikciju

Inter skip algoritam služi za smanjenje procesa izračunavanja kroz identifikaciju da li makroblok skipovati ili ne. Pre procesa kodovanja svakog makrobloka, koder procenjuje rate-distortion gubitke za kodovane ili skipovane makroblokove. Bazirajući se na

ovim procenama, makroblok se koduje kao normalan (koder procesira makroblok i selektuje odgovarajući mod kodovanja) ili skipovan (proces kodovanja se ne izvršava). Skip model predikcije ima za cilj da redukuje vreme izračunavanja dok održava ili unapređuje rate-distortion performanse [11].

Algoritam za inter skip predikciju baziran je na sledećim ograničenjima:

- Main Profilu sa odgovarajućim alatima,
- Algoritmu za skip predikciju samo B slike,
- Rate distortion modu optimizacije koji je isključen.

Ako je  $M_i$  mod kodovanja (jedan od  $K$  mogućih modova) koji je izabran od strane koder za makroblok  $X_i$  i ako  $M_i=skip$  označava "skip" mod kodovanja. Rate-distortion cost kodovanja ili skipovanja makrobloka se definiše kao [12]:

$$J(X_i) = D(X_i, M_i) + \lambda R(X_i, M_i) \quad (1)$$

Dok je rate-distortion cost skipovanog makrobloka:

$$J(X_i, K) = D(X_i, K) \quad (2)$$

gde je  $\lambda$  težinski parameter (Lagranžov multiplikator). Brzina skipovanih makroblokova približno je jednaka nuli.  $I$ -ti macroblock se može se skipovati (ne kodovati) ako:

$$D(X_i, M_i) + \lambda R(X_i, M_i) \geq D(X_i, K) \quad (3)$$

$D(X_i, K)$  je distorzija (izmerena kao srednja kvadratna greška (Mean Square Error, MSE)) između tekućeg makrobloka i makrobloka kompenzacije pokreta (sa nula pomeraja vektora predikcije MVP) referentne slike.

$D(X_i, M_i)$  je srednja kvadratna greška između tekućeg makrobloka i dekodovanog rekonstruisanog makrobloka, dok je  $R(X_i, M_i)$  broj bita neophodan za kodovanje tekućeg makrobloka sa korišćenjem  $M_i \in \{1, \dots, K-1\}$  moda kodovanja. Stvarne vrednosti  $D(X_i, M_i)$  i  $R(X_i, M_i)$  nisu dostupne pre procesa kodovanja i za procenu svakog od ovih parametara za svaki makroblok u tekućem frejmu ( $n$ ) koriste se sledeći modeli:

$$\widehat{D}^{(n)}(X_i^{(n)}, M_i^{(n)}) = \alpha_d D^{(n-1)}(X_i^{(n-1)}, M_i^{(n-1)}) \quad (4)$$

i

$$\widehat{R}^{(n)}(X_i^{(n)}, M_i^{(n)}) = \alpha_r R^{(n-1)}(X_i^{(n-1)}, M_i^{(n-1)}) \quad (5)$$

$D^{(n-1)}(X_i^{(n-1)}, M_i^{(n-1)})$  je srednja kvadranta greška između originalnog i rekonstruisanog makrobloka na istoj poziciji u prethodnom frejmu ( $n-1$ ), dok je  $R^{(n-1)}(X_i^{(n-1)}, M_i^{(n-1)})$  broj bita koji je potreban za kodovanje makrobloka na istoj poziciji u prethodnom frejmu, dok su  $\alpha_d$  i  $\alpha_r$  fiksni parametri.

Lagranžov multiplikator  $\lambda$  je modelovan kao funkcija od  $QP$  i aktivnosti izvora  $F_i$  (gde je  $F_i$  = srednja MSE x srednji broj bita (po makrobloku)):

$$\widehat{\lambda} = \left( 7.374 \cdot 10^{-8} F_i + 5.239 \cdot 10^{-5} \right) \exp \left[ -3.688 \cdot 10^{-5} F_i + 0.3203 \right] \cdot QP \quad (6)$$

Algoritam za skip predikciju ima sledeću proceduru [12]:

1. Za svaki makroblok, izračunati  $D^n(X_i, K)$  i pročitati prethodno memorisane vrednosti  $D^{(n-1)}(X_i^{(n-1)}, M_i^{(n-1)})$  i  $R^{(n-1)}(X_i^{(n-1)}, M_i^{(n-1)})$ .

2. Izračunati faktor aktivnosti  $F_i$  za tekući makroblok:

$$F_i = D^{(n-1)}(X_i^{(n-1)}, M_i^{(n-1)}) \cdot R^{(n-1)}(X_i^{(n-1)}, M_i^{(n-1)}) \quad (7)$$

3. Izračunati  $\hat{\lambda}$  koristeći jednačinu “(6)”.

4. Izabрати “skip” mod kodovanja ako je sledeći izraz tačan:

$$D^{(n-1)}(X_i^{(n-1)}, M_i^{(n-1)}) + 0.5 \hat{\lambda} R^{(n-1)}(X_i^{(n-1)}, M_i^{(n-1)}) \geq D^{(n)}(X_i, K) \quad (8)$$

5. Ako je izabran “skip” mod, ne vršiti dalje kodovanje i makroblok označiti kao “skipovan”. Ako je izabran mod “kodovati”, kodovati makroblok po standardnoj proceduri.

#### B. Algoritam za selektivnu intra predikciju

Intra predikcija koristi prostornu korelaciju između blokova u okviru slike za redukciju prostorne redundanse. Svaka slika je podeljena u makrobloke veličine 16x16 piksela i svaki makroblok se sastoji od lumentne i chroma komponente. Makroblok veličine 16x16 piksela za lumentne komponente može se izdeliti u šestnaest blokova veličine 4x4. Chroma komponente su prediktovane pomoću 8x8 bloka sa sličnom tehnikom za predikciju kao kod 16x16 lumentnih blokova. Postoji 9 modova predikcije za lumentni blok veličine 4x4 i 4 moda predikcije za lumentni 16x16 blok. Za chroma komponente, postoji 4 moda predikcije koji su primenjeni na dva 8x8 chroma bloka (U i V).

H.264/AVC koduje makrobloke sa ponavljanjem svih lumentnih intra odluka za svaki mogući chroma intra mod predikcije da bi se postigla najbolja efikasnost. Broj kombinacija modova za lumentne i chroma komponente u okviru makrobloka je  $C8 \times (L4 \times L16 + L16)$ , gde C8, L4 i L16 predstavljaju broj modova za predikciju chroma komponenti i 4x4 i 16x16 lumentne komponente. To znači da je za makroblok potrebno izvršiti  $4 \times (9 \times 16 + 4) = 592$  različita RDO proračuna pre nego što se odredi najbolji RDO mod.

Selektivna metoda intra predikcije bazira se na činjenici da dominantan smer predikcije većeg bloka je sličan kao kod manjeg bloka [13]. Najbolji mod predikcije kod 4x4 lumentnog bloka u okviru 16x16 bloka ima iste smerove kao i taj 16x16 lumentni blok. Proračun za lumentnu i chroma intra predikciju može se umanjiti na osnovu opštih informacija o ivicama dobijenih iz rezultata kod 16x16 intra predikcije. U najboljem slučaju broj mogućih kombinacija modova za svaki makroblok je samo  $1 \times (4 \times 16 + 4) = 68$ , dok trenutni proračun RDO kod H.264/AVC zahteva  $4 \times (9 \times 16 + 4) = 592$ . Broj modova kandidata za mod odluke kod selektivne intra predikcije je prikazan u Tabeli 1.

TABELA 1. BROJ MODOVA KANDIDATA

Komponente slike	Veličina bloka	Ukupan broj modova	Broj modova kandidata
Lumentna komponenta (Y)	4x4	9	4 to 7

Lumentna komponenta (Y)	16x16	4	4
Chroma komponente (U, V)	8x8	4	1 do 2

#### IV. EKSPERIMENTALNI REZULTATI I DISKUSIJA

Predloženi kombinovani algoritam je inkorporiran u referentni H.264/AVC kodek (JVT - JM 10.2). Eksperiment je urađeni uz korišćenje originalne i modifikovane verzije (verzija JM 10.2M) referentnog test softvera JM 10.2 [14]. Rezultati eksperimenta su dobijeni pod sledećim opštim uslovima:

- Main profil sa CABAC metodom entropijskog kodovanja,
- vrednost opsega pretraživanja MV je 32,
- vrednost parametra kvantizacije je 40,
- broj referentnih slika je 5,
- koristi se Hadamarova transformacija i
- IBBP format.

Za testiranje je korišćena sledeća sistemna platforma: AMD Pentijum 4 procesor brzine 2.53 GHz, 512 MB RAM-a i operativni sistem Microsoft Windows XP.

Ideja je bila da se izvrše testiranja i poređenja različitih softverskih verzija u nameri da se prikažu poboljšanja koja su dobijena korišćenjem verzije JM 10.2M.

U eksperimentu se koristili prvih 50 frejmova 6 različitih test sekvenci u QCIF formatu (Claire, Coastguard, Container, Grandmother, Newa i Salesman). Test sekvence su izabrane prema razlici u sadržaju i pokretu.

Za poređenje i analizu izlaznih rezultata kod svih eksperimenata koriste se tri dobro poznata faktora: odnos signal/šum (SNR) za lumentnu (Y) komponentu slike, brzina prenosa (bitska brzina) i računarsko vreme. Računarsko vreme je dobijeno kao izlazna vrednost nakon svakog testiranja.

Izmeren je odnos signal/šum samo za lumentnu komponentu slike zato što je ljudski vizuelni sistem osetljiviji na lumentnu nego na chroma (Cb i Cr) komponente slike.

Uzimajući u obzir činjenicu da video kodni standard H.264/AVC sadrži dve metode entropijskog kodovanja t.j. CABAC i CAVLC, koristi se CABAC zato što ovaj metod daje mnogo bolje rezultate nego CAVLC u odnosu na video sadržaj [2].

Takođe, koristi se Hadamard transformacija zato što ona ima bolje performanse kod video kodovanja u poređenju sa drugim vrstama transformacija [2].

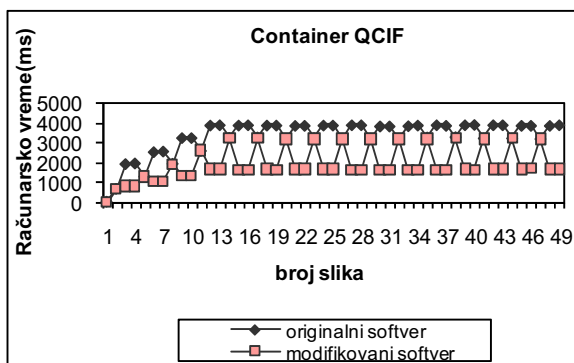
Tabela 2 pokazuje performanse predloženog kombinovanog algoritma za procesiranje B slike u okviru IBBP formata. Kada je broj referentnih frejmova 5, sa predloženim algoritmom, ušteda u računarskom vremenu je u rasponu od 35% do 42%. Računarsko vreme je smanjeno u proseku oko -39%. To ukazuje da je modifikovani H.264/AVC koder brži od referentnog. Međutim, treba napomenuti da postoje neznatni gubici u vrednostima SNR-a (u proseku samo 0,03 (dB)), kao i da

su se vrednosti brzine prenosa neznatno povećale (od -0,53 do 2,8%).

TABELA 2. REZULTATI EKSPERIMENTA

Test sekvence	Računarsko vreme (%)	SNR (dB)	Brzina prenosa (%)
Claire	-35,41	0,03	0,24
Coastguard	-38,01	0,04	2,87
Container	-42,26	0,00	-0,53
Grandmother	-39,88	0,00	0,89
News	-39,00	0,07	2,52
Salesman	-34,53	0,04	1,71
Prosečna vrednost	-38,18	0,03	1,28

Od svih testiranih sekvenci Container sekvenca je pokazala najbolje performanse u smislu računarskog vremena (-42,26%), brzine prenosa (-0,53 %) i SNR, kada se primeni predložena kodna šema. Sl. 1 prikazuje smanjenje u računarskom vremenu za Container test sekvencu.



Sl. 1. Redukcija u računarskom vremenu kod Container test sekvence.

#### IV. ZAKLJUČAK

H.264/AVC koder ima značajnu kompleksnost izračunavanja zato što selektuje najbolji mod kodovanja korišćenjem rate-distortion optimizacije (RDO), da bi iskoristio potpunu prednost selektovanog moda uz uslov maksimiziranja kvaliteta kodovanja i minimiziranja bita sa podacima. RDO mod odluke detaljno pretražuje svaki mogući mod za svaki makroblok da bi pronašao mod koji ima minimalnu tzv. rate-distortion cost. Ova tehnika optimizacije ima veoma veliku kompleksnost.

Da se minimizirala RD cost, ovaj rad poredi dva test modela softvera (originalni i modifikovani) za kombinovani rani inter skip i selektivni intra metod predikcije kod H.264/AVC koder, kada se procesira i analizira B slika. Rezultati eksperimenta pokazuju da je vreme potrebno za proces kodovanja redukovano 35-42% kroz postupak rane identifikacije makrobloka koji će verovatno biti skipovani tokom procesa kodovanja, kao i kroz smanjenje broja kandidovanih modova kod intra

predikcije. Pri tome ne postoje značajni gubici u rate-distortion performansama.

#### LITERATURA

- [1] 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.
- [2] 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.
- [3] K.R.Rao, Z.S.Bojkovic, D.A.Milovanovic, *Multimedia communication systems: techniques, standards and networks*, Prentice Hall PTR, 2002.
- [4] Iain E. G. Richardson, *H.264 and MPEG-4 video compression video coding for next-generation multimedia*, John Wiley & Sons Ltd., The Atrium, Southern Gate, Chichester, West Sussex PO 19 8SQ, England, 2003.
- [5] S. Kwon, A. Tamhankar, K. R. Rao, Overview of the H.264/MPEG-4 part 10, *Journal of Visual Communication and Image Representation* 17, pp.186-216, 2006.
- [6] G. Bjontegaard and K. Lillevold, Context-adaptive VLC coding of coefficients, JVT document C028, Fairfax, May 2002.
- [7] 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, July 2003.
- [8] D. Marple, T. Weingard, G. J. Sullivan, The H.264/MPEG4 Advanced Video Coding Standard and its Applications, *IEEE Communication Magazin*, pp. 134-143, 2006.
- [9] I. Choi, J. Lee and B. Jeon, Fast Coding Mode Selection With Rate-Distortion Optimization for MPEG-4 part-10 AVC/H.264, *IEEE Transactions on Circuits and System for Video Technology*, vol. 16, no. 12, pp. 1557-1561, December 2006.
- [10] M. Flierl and B. Girod, Generalized B Pictures and the draft JVT/H.264 Video Compression Standard, *IEEE Transactions on Circuits and System for Video Technology*, vol.13, no.7, pp. 587-597, 2003.
- [11] Z. Milicevic, Z. Bojkovic, Computational time reduction using low complexity skip prediction for H.264/AVC standard, *WSEAS Transactions on Information Science and Applications*, is. 1, vol. 14, pp. 59-63, 2007.
- [12] C. S. Kannangara, I. E. G. Richardson, M. Bystrom, J. R. Solera, Y. Zhao, A. MacLennan, and R. Cooney, Low-complexity skip prediction for H.264 through lagrangian cost estimation, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 2, pp. 202-208, February 2006.
- [13] J. S. Park, and H. J. Song, Selective Intra Prediction Mode Decision for H.264/AVC Encoders, *Transactions on Engineering, Computing and Technology*, vol. 13, pp. 51-55, 2006.
- [14] <http://bs.hhi.de/~suehring/tml/download/JM.10.2>

#### ABSTRACT

The H.264/AVC video coding standard is significantly improved compression performance, but coder complexity is increased, too. In order to reduce the complexity, the combination of two methods for B pictures processing i.e. early skip mode decision and selective intra mode decision is proposed. Experimental results show that the coding time is reduced by 35-41 % through early identification of macroblocks that are to be skipped during the coding process as well as through reducing the number of candidate modes.

#### MODELING AND B FRAME PROCESSING IN H.264/AVC STANDARD

Zoran Milicevic, Zoran Bojkovic