

ANALIZA VIDEO SEKVENCI PRI NISKIM PROTOCIMA BITA

Irina Reljin^{(1),(2)}, ireljin@ptt.yu, Marija Zajeganović-Ivančić⁽¹⁾, zmarija@eunet.yu

⁽¹⁾Viša tehnička PTT škola, Beograd, ⁽²⁾Elektrotehnički fakultet, Beograd

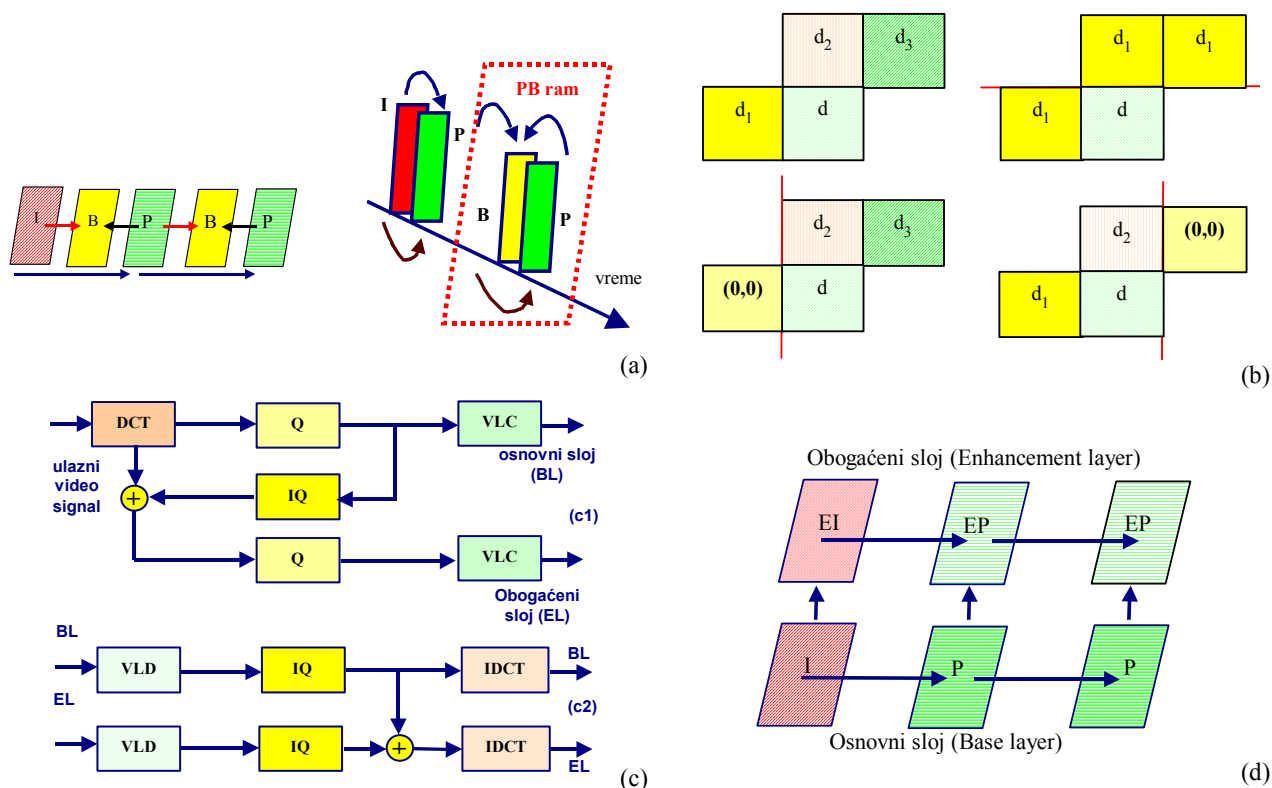
Sadržaj – U radu se analiziraju video sekvence kodovane prema preporukama H.261, H.263 i H.263+. Analiziran je broj bajtova po slici u video sekvenci, nelinearnim metodama: fraktalnom i multifraktalnom analizom.

1. UVOD

Za potrebe videotelefonijske i videokonferencijske razvijeni su standardi H.261 i H.263, koji dopuštaju vanredno male protoke bita (manje od 64 kb/s) tako da se video signali mogu prenositi i klasičnim telefonskim linijama i/ili mobilnom telefonijom. Za videotelefoniju i videokonferencije tipična scena je u osnovi statična, što dopušta uvođenje znatno većih

iznosa kompresije u odnosu na filmske/TV sekvence koje mogu da sadrže naglo promenljive scene.

U radu se analiziraju neke karakteristike video signala komprimovanih prema H.261, H.263 i H.263+ standardu. U poglavlju 2. se, ukratko, izlaže struktura ovih standarda [1-4]. U poglavlju 3. se razmatraju karakteristike video sekvenci filma «The Firm» koje su digitalizovane i komprimovane na Tehničkom univerzitetu u Berlinu, i dostupne na mreži [4]. Posmatrane su veličine pojedinačnih slika (izražene u bajtovima) u sekvenci, nakon kompresije. Izvršene su fraktalna i multifraktalna analiza sekvenci, prema [5-7], i izvedena poređenja, imajući u vidu strukturu sekvenci pri različitim kompresijama i za različite protoke bita.



Slika 1. (a) Formiranje P, B i PB slika; (b) Formiranje vektora pokreta; (c) Koder/dekoder osnovnog i obogaćenog sloja u H.263+ standardu; (d) Struktura sekvence slika u osnovnom i obogaćenom sloju – SNR skalabilnost.

2. RAZVOJ H.263 STANDARDA

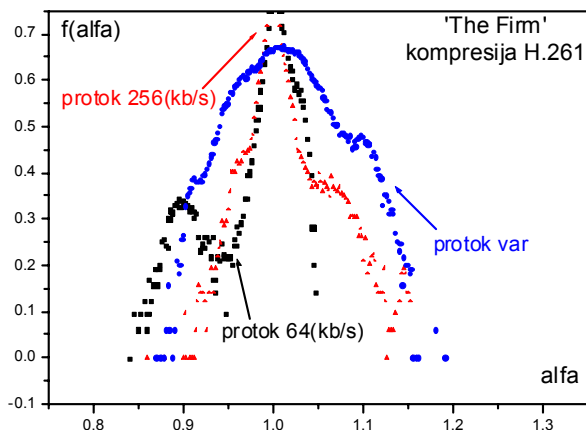
Standard H.261, je namenjen za protoke bita iznosa $m \cdot 384$ kb/s, gde je m između 1 i 5 i predviđen je za ISDN. Od 1988. cilj je pomeren na protoke bita vrednosti $p \cdot 64$ kb/s, gde je p od 1 do 30, tako da je ova preporuka poznata i kao $p \cdot 64$ standard [3]. Dalje usavršavanje video kompresije dovelo je do preporuke H.263 usvojene 1995. godine, namenjene prenosu pri protocima bita nižim od 64 kb/s, tj. prenosu standardnom telefonskom ili mobilnom mrežom. Tokom razvoja standarda postavljena su dva cilja: da se unapredi postojeći standard H.261, bez bitnijih izmena njegovih osnovnih postavki, i da se projektuje standard za

video kompresiju koji ne mora da sadrži postavke prethodnog, već može imati i druga rešenja. Kao rezultat toga razvijen je standard H.263, kao i njegove kasnije verzije H.263+ i H.263++.

Kodovanje prema preporuci H.261 koristi mešovitu kompresiju: u prostornom domenu – unutar jedne slike, poznatu i kao *intra-frame* kompresija, koja se bazira na prostornoj redundansi piksela slike, i kompresiju u vremenskom domenu – između susednih slika sekvence, poznatu kao *inter-frame* kompresija, koja se bazira na vremenskoj redundansi piksela susednih slika, uz korišćenje predikcije pokreta. Za smanjenje prostorne redundanse se koristi diskretna kosinusna transformacija (DCT) nad

blokovima dimenzije 8x8 piksela, kvantovanje (Q) računanjem osobina vizuelnog sistema, *run-length* kodovanje i kodovanje sa promenljivom dužinom kodne reči (VLC, *Variable Length Coding*). Intraframe kodovane slike se označavaju kao I-slike. Dodatno je uvedena i predikcija narednih slika, na osnovu postupka poznatog kao kompenzacija pokreta (*motion compensation*). Podaci prethodne slike se koriste za predikciju blokova aktuelne slike, P-slike. Koduje se samo razlika aktuelnih i predviđenih blokova slike, koja je, uobičajeno, mala, čime se postižu znatne uštede u količini informacija.

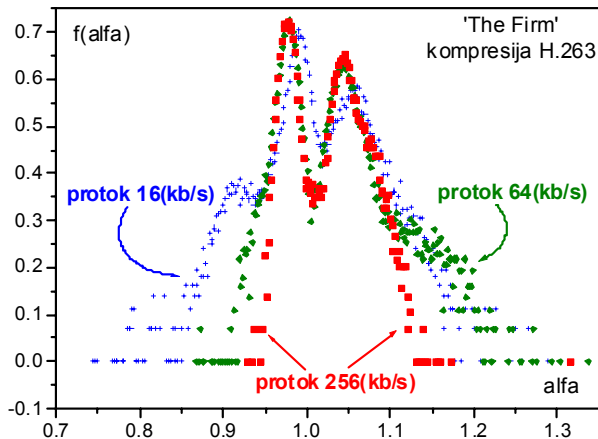
Dalje povećanje stepena kompresije se može ostvariti bi-direkcionom predikcijom slika – uvođenjem tzv. B-slika. Ove slike se koduju na osnovu predikcije pokreta prema najbližoj prethodnoj (*forward prediction*) i narednoj (*backward prediction*) slici, bilo da je ona I- ili P-slika. B-slike koriste oko 50% manje bita u odnosu na P-slike. Ove slike se ne koriste kao referentne, tako da ne unose grešku. Međutim, kako se pri *backward* predikciji koristi naredna slika u sekvenci, potrebno je izvršiti preuređenje slika pri slanju i rekonstrukciji, što povećava kašnjenje u formiranju video sekvence, što je nepovoljno za interaktivni rad. Takođe, zahteva se veći memorijski prostor u dekoderu jer se slike moraju memorisati a zatim rekonstruisati u ispravnom redosledu. Video sekvenca koja koristi samo I-slike dopušta najveću slobodu u pristupu i najbolje pretraživanje, ali ima i najniži stepen kompresije. Sekvenca koja je kodovana tako da ima I- i P-slike se odlikuje visokim srednjim nivoom kompresije i ima izvesnu slobodu slučajnog pristupa i pretraživanja video materijala. Sekvence koje sadrže sva tri tipa slika (I B P) imaju visok stepen kompresije, ali i znatno povećano kašnjenje, što se može javiti kao ozbiljan nedostatak za interaktivni rad (kao u videotelefoniji ili videokonferenciji), i zahtevaju veći memorijski prostor u dekoderu [3-6].



Slika 2. Multifraktalni spektar filma "The Firm" kodovanog sa H.261 za različite protoke bita.

U standardu H.263 se dodatno povećava efikasnost kompresije time što se, umesto B-slika, koriste PB koje su u suštini kombinacija dve slike: P- i B-kodovanih kao jedna jedinica, slika 1.a. S obzirom da B-slika dopušta veliku kompresiju, ovakvo kodovanje omogućava povećanje brzine prenosa slika bez značajnijeg povećanja protoka bita. Rad sa PB-slikama je naročito pogodan za sporo promenljive scene, dok je neefikasan za brzo promenljive kadrove ili za rad sa veoma malim brojem slika u sekundi.

Za predikciju pokreta se koristi lumentni sadržaj iz makro-blokova (MB), tako što se vrši procena najboljeg slaganja blokova (*block matching*). Svaki makro-blok iz trenutno procesirane slike se poredi sa makro-blokovima koji leže unutar prozora za pretraživanje u prethodno rekonstruisanoj slici. Kriterijum slaganja blokova može biti bilo koja mera greške: srednja kvadratna greška ili suma apsolutnih grešaka. Vektor pomeraja (MV = *motion vector*) između trenutnog MB i najbližeg MB iz prethodne slike se opisuje vertikalnim i horizontalnim komponentama koje se šalju dekoderu radi korekcije pri rekonstrukciji odgovarajućeg MB. S obzirom da se koduju samo vektori pomeraja, time se obezbeđuje visok stepen kompresije. Za MV se koriste lumentne komponente iz makro-blokova, koji su sadržani u matricama 16x16 piksela (4 lumentna bloka), a hrominantne komponente unutar makro-blokova su u matricama dva puta manjih dimenzija (8x8 – dva hrominantna bloka). Poboljšanje efikasnosti kompresije se dodatno ostvaruje predikcijom vektora pomeraja. U H.261 se vrši predikcija na osnovu vektora pomeraja od najbližeg MB sa leve strane. U H.263 se dopušta predikcija vektora pomeraja sa tačnošću pola piksela. Kada se dobija necelobrojna vrednost vektora pomeraja, koristi se bilinearna interpolacija za pronalaženje odgovarajućeg položaja piksela. Predikcija se ovde vrši na osnovu vektora pomeraja prema tri susedna makro-bloka: prema levom horizontalnom (d1), gornjem vertikalnom (d2) i gornjem desnom (dijagonalnom) makrobloku (d3), slika 1.b – gore levo. Za svaku komponentu koduje se samo razlika vektora pomeraja posmatranog bloka u odnosu na vektore pomeraja susednih blokova, ali se posmatra razlika u odnosu na median vrednosti vektora pomeraja susednih blokova po horizontalnom i vertikalnom pravcu. Blokovi koji se nalaze van granica slike računavaju se sa nulom, ako je samo jedan u pitanju, odnosno kao levi (d1), ako su dva u pitanju, slika 1.b.

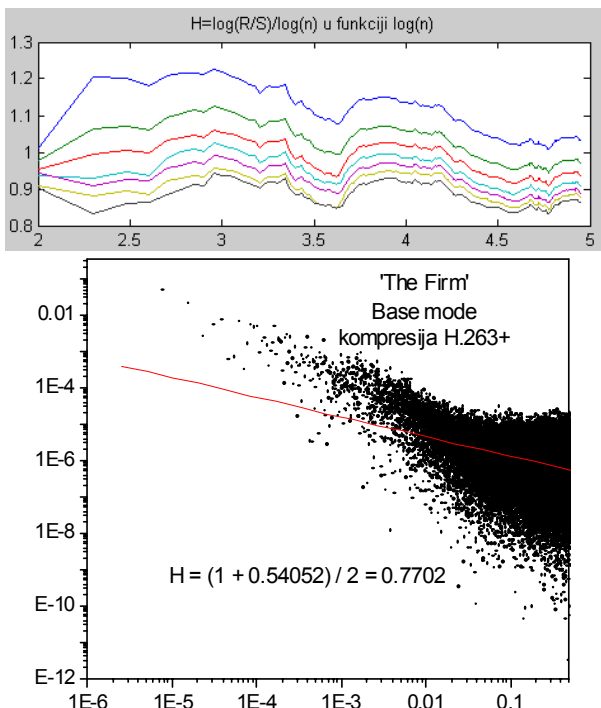


Slika 3. Multifraktalni spektar filma "The Firm" kodovanog sa H.263 za različite protoke bita.

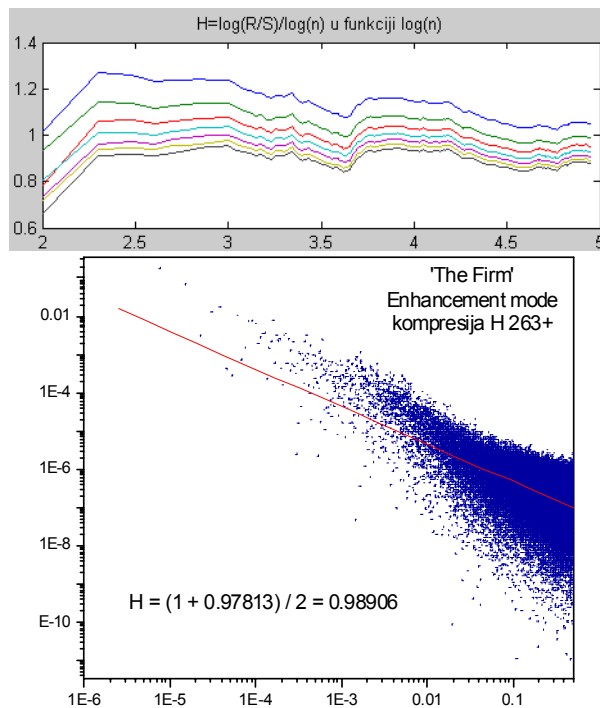
U prenosu kodovanog video signala postoji potreba da se zadovolje zahtevi korisnika različitih prijemnih uređaja, a da se pri tome prenese isti sadržaj video materijala. To implicira prijem sekvenci različitih protoka. Stoga je u standardne kompresije uvedena skalabilnost koja može biti: kvalitativna (SNR = *Signal-to-Noise Ratio*), prostorna, vremenska, frekvencijska ili neka od njihovih kombinacija, slika 1.c. i 1.d. Koder, u slučaju H.263+ kompresije, na svom izlazu formira dve sekvence koje se šalju jednovremeno slika-po-

slika. Prva sekvenca pripada osnovnom sloju (BL=Base Layer) i sadrži samo I i PB slike. U drugoj sekvenci, obogaćenom (EL=Enhancement Layer) vrši se korekcija komprimovanog video signala, slika 1.c1, i formira sekvenca slika EI i EP. Već kvantovan signal se koristi za korekciju greške i na taj način dobija sekvenca obogaćenog sloja.

Koder H.263+ na izlazu daje dve sekvence čije se slike (I EI P EP...) sukcesivno šalju. Ako se dozvoljava lošiji kvalitet prijema koristi se samo osnovni sloj, dakle samo jedna sekvenca. Za kvalitetni prijem se koriste obe sekvence, osnovna i obogaćena, a dekodovanje se vrši prema slici 1.c2.



Slika 4. RS statistika i periodogram osnovnog sloja filma "The Firm" kodovanog sa H.263+.

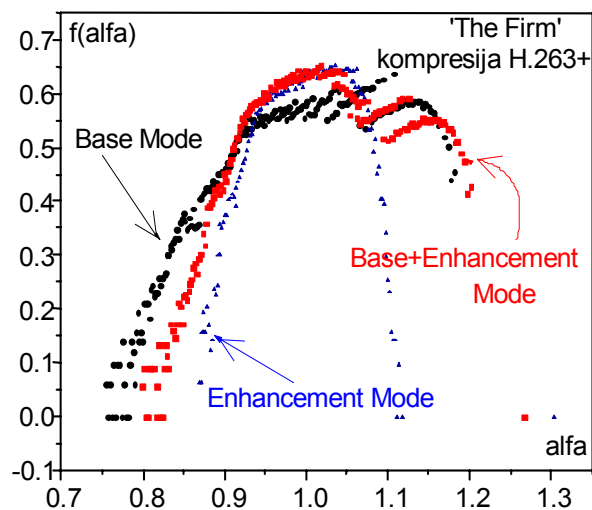


Slika 5. RS statistika i periodogram obogaćenog sloja filma "The Firm" kodovanog sa H.263+.

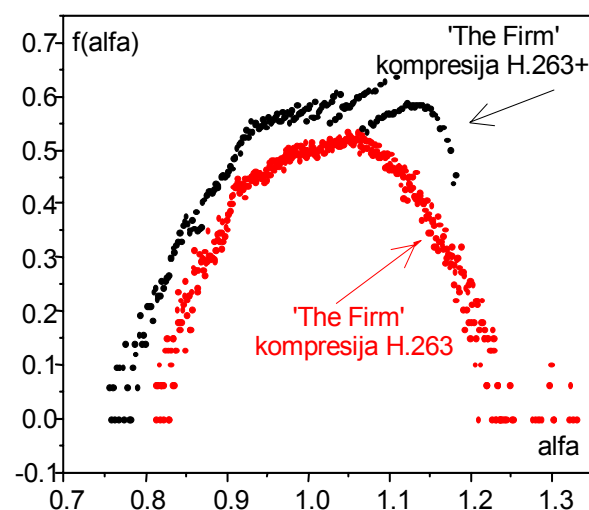
3. ANALIZA KOMPRIMOVANIH VIDEO SEKVENCI

Video sekvenca filma «The Firm» je, najpre, formirana u QCIF formatu (176x144 piksela lumentnog signala i 88x72 piksela za svaku od hrominentnih komponenata) sa 8 bita po pikselu i 25 slika u sekundi. Ovako formirana digitalne sekvenca je dalje kodovana prema H.261 i H.263 standardima za četiri protoka bita: 16 kb/s, 64 kb/s, 256 kb/s i

za VBR (Variable Bit Rate), dakle, bez definisane gornje granice protoka. U H.263+ kompresiji su formirani osnovni i obogaćeni sloj. Komprimovane sekvence iskazuju izrazitu eksplozivnost, a time i svojstvo dugotrajne zavisnosti, odnosno, fraktalnosti. To je potvrđeno izračunavanjem Hurstovog indeksa (vrednosti između 0.5 i 1.0). Ovakvo svojstvo iskazuje i savremeni telekomunikacioni saobraćaj [6-7].



Slika 6. Multifraktalni spektar osnovnog sloja, obogaćenog sloja i celog filma "The Firm" kodovanog sa H.263+.



Slika 7. Multifraktalni spektar osnovnog sloja filma "The Firm" kodovanog sa H.263+ i sa H.263.

Velike varijacije vrednosti Hurst-ovog indeksa ukazuju da video sekvence, pored fraktalnih karakteristika, poseduju izrazitu multifraktalnu komponentu. Stoga su izračunati multifraktalni spektri metodom histograma. Rezultati su u saglasnosti sa nekim nedavno dobijenim u [6-7]. Prema slikama 2 i 3, široka parabola multifraktalnih spektara se dobija za manje iznose kompresije. Što je kompresija veća, to su izrazitije neravnomernosti u spektrima i uočavaju se strukture sekvenci, u kojima će se isticati oštri maksimumi koji, prema [6-7], potiču od slika različitog tipa (I i P u slučaju H.261, odnosno I i PB u slučaju H.263 kompresije).

Ispitivanja sekvenci filma u osnovnom i obogaćenom sloju najpre su vršena izračunavanjem Hurst-ovog indeksa (H). Korišćen je RS metod razvijen u [6]. Pokazuje se, slike 4 i 5, da H-indeksi veoma variraju, i to u okolini vrednosti 1. Metodom linearne regresije određen je nagib RS krive, a odatle je izračunat Hurst-ov koeficijent. Dobijene vrednosti ukazuju na izrazito fraktalno svojstvo procesa veličine slika pri H.26x kompresijama. Metodom periodograma, koja se inače može smatrati lošijom, jer je zasnovana na linearnoj operaciji izračunavanja Furijeovog spektra, dobijene su,

takođe, velike vrednosti ($H=0.7702$ za osnovni sloj i 0.989 za obogaćeni sloj).

Varijacije H-indeksa su bile razlog za ispitivanje multifraktalnih svojstava. Spektri prikazani na slikama 6 i 7, iskazuju globalni maksimum oko vrednosti Holder-ovog eksponenta $\alpha=1.0$ i lokalni maksimum pri $\alpha=1.15$, što ukazuje na aditivne komponente vremenske serije. Detaljnijom analizom, posmatranjem pojedinačnih MF spektara samo za P-slike, odnosno samo za PB-slike, mogu se uočiti uticaji tih komponentata na ukupan oblik MF spektara. Globalni MF spektri u slučaju H.263 i osnovnog sloja H.263+ su neosporno slični – oblika parabole. Razlike u detaljima su posledica modifikacija koje H.263+ preporuka unosi u kodovanje: usavršeno kodovanje blokova u intra modu, filter za uklanjanje blok efekata u slici, ubacivanje informacije koje spoljašnjem korisniku pružaju poboljšanje u prikazu (zamrzavanje slike, zum ili *chroma-key* efekat), poboljšanje PB moda (regularno formiranje bidirekciono predikcije slike B, unutar PB strukture koja se koduje kao celina).

Tabela 1. Karakteristični podaci komprimovanih sekvenci.

Standard	H.261			H.263				H.263+
	64 kb/s	256 kb/s	vbr	16 kb/s	64 kb/s	256 kb/s	vbr	vbr
Stepen kompresije	118.95	29.71	22.82	476.20	118.96	29.73	59.67	133.73
Eksploziv. slika	6.22	4.97	4.41	55.00	14.17	7.68	14.30	4.54
Eksploziv. protoka	6.2	5	4.4	5.9	5.3	5.4	13	4.5
H-indeks	0.25-0.78	0 -0.996	0.66-0.84	0-0.97	0-0.923	0-0.923	0.48-0.88	0.735-1

Prema podacima navedenim u Tabeli 1, uočava se da stepen kompresije veoma varira u zavisnosti od zadatog (ciljnog) protoka bita. Za odabranu vrednost protoka od 16 kb/s (H.263) ostvaren je izuzetno visok stepen kompresije od 476:1. Kodovanje prema preorukama H.261 i H.263 za iste ciljne protoke, daje slične vrednosti odnosa kompresije. Sama H.263+ sekvenca, međutim, za vbr varijantu (dakle, bez ograničenja protoka) postiže visok stepen kompresije (133.73).

Eksplozivnost sekvenci slika (definisana maksimalnim brojem prema srednjem broju bajtova u slici) kodovanih prema H.261 preporuc, je bliska stvarnoj eksplozivnosti saobraćaja (maksimalan broj bita prema srednjem broju bita). U slučaju H.263 sekvence, međutim, ova dva broja se veoma razlikuju. Razlog tome je postojanje PB-slika koje u principu nose veliku količinu informacija, pa time povećavaju ukupnu eksplozivnost slika.

4. ZAKLJUČAK

S obzirom da je priroda video sekvenci nelinearna, analiza saobraćaja koji se generiše na izlazu video kodera zahteva primenu neke od nelinearnih metoda, na primer fraktalne, odnosno, multifraktalne. Ispitivanja fraktalnih svojstava sekvenci, izračunavanjem Hurst-ovog indeksa ukazala su na postojanje multifraktalne prirode u komprimovanim video sekvencama. Izračunati su RS dijagrami i periodogrami, kao potvrda prethodnih rezultata.

Multifraktalni spektri potvrđuju rezultate dobijene u prethodnim radovima [6-7] za druge tipove kompresije. Ovde je dodatno, prvi put analiziran MF spektar jedne H.263+ sekvence, i to posebno za osnovni i obogaćeni sloj, kao i za

čitav film sastavljen od dve, primarno generisane, sekvence (BL+EL).

LITERATURA

- [1] K. Rao, Z. Bojkovic, D. Milovanovic, *Multimedia Communication Systems*, Prentice Hall, NJ, 2002.
- [2] Y. Wang, J. Osterman, Y.Q. Zhang, *Video processing and Communications*, Prentice Hall, NJ, 2002.
- [3] ITU-T, H.263 Recommendation, ITU-T, Geneva, 2000.
- [4] F. Fitzek, M. Reisslein, "MPEG-4 and H.263 video traces for network performance evaluation", *TKN Technical Report TKN-00-06*, Technical University Berlin, 2000.
- [5] C. Evertsz, B. Mandelbrot: "Multifractal Measures", Appendix B in H. Peitgen, H. Jurgens, P. Andrews *Chaos and Fractals*, Springer, 1992.
- [6] I. Reljin, B. Reljin, "Fractal and multifractal analyses of compressed video sequences" (invited paper), in *Proc. X Conf. TELFOR-02*, pp. 471-476, Belgrade, 26-28 Nov. 2002.
- [7] I. Reljin, B. Reljin, "Statističke i multifraktalne karakteristike video signala komprimovanog prema H.263 standardu", *Zbornik radova 20. simp. o novim tehnologijama u poštanskom i telekomunikacionom saobraćaju*, str. 193-205, Saobraćajni fakultet, Beograd, 10-11. dec. 2002.

Abstract – *The paper analyses the video sequences coded according to the H.261, H.263 and H.263+ recommendations. The number of bytes-per-frame in video sequences was analysed through non-linear methods: the fractal and multifractal analyses.*

ANALYSIS OF VIDEO SEQUENCES AT LOW BIT RATES

Irina Reljin, Marija Zajeganović-Ivančić