

ANALIZA MPEG-4 VIDEO SEKVENCI

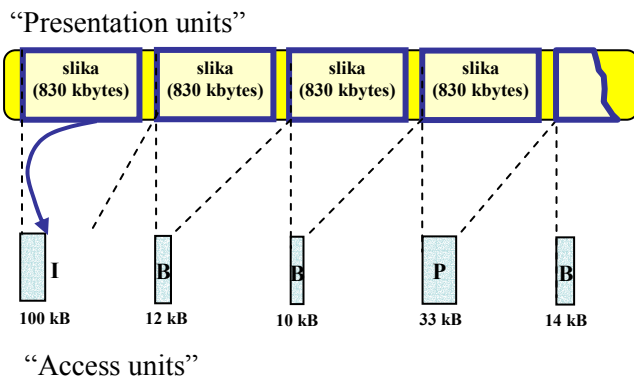
Irina Reljin^{(1),(2)}, ireljin@ptt.yu, Branimir Reljin⁽²⁾, reljinb@etf.bg.ac.yu, Marija Zajeganović-Ivančić⁽¹⁾, zmarija@eunet.yu
⁽¹⁾Viša škola za informacione i komunikacione tehnologije, Beograd, ⁽²⁾Elektrotehnički fakultet, Beograd

Sadržaj – U radu su analizirane video sekvence kodovane prema standardu MPEG-4. Analiziran je broj bajtova po slici, kao i broj bita po makrobloku, u video sekvenci, nelinearnim metodama: fraktalnom i multifraktalnom analizom.

1. UVOD

Standardni digitalni video signal sadrži veliku količinu informacija: jedna slika (*frame*) standardne TV rezolucije zahteva 830 kB, što predstavlja problem pri prenosu i arhiviranju video materijala. Stoga se velika pažnja posvećuje izvornom kodovanju, posebno, različitim tehnikama kompresije video sadržaja. Pri kompresiji video sekvenci koriste se dva osnovna moda kodovanja:

- *Intraframe* (unutar *frame*-a) kada se uklanja prostorna i statistička redundansa unutar jedne slike (*frame*-a), najčešće primenom diskretne kosinusne transformacije, pri čemu se formiraju tzv. I *frame*-ovi, i
- *Interframe* (između *frame*-ova) kada se formiraju tzv. P *frame*-ovi (sa predikcijom unapred) ili B (bi-direkcionni, sa predikcijom unapred i unazad), ili neki složeniji tipovi *frame*-ova (zavisno od kompresije).



Slika 1. Ilustracija veličine frame-ova pri MPEG-x i H.26x kompresiji

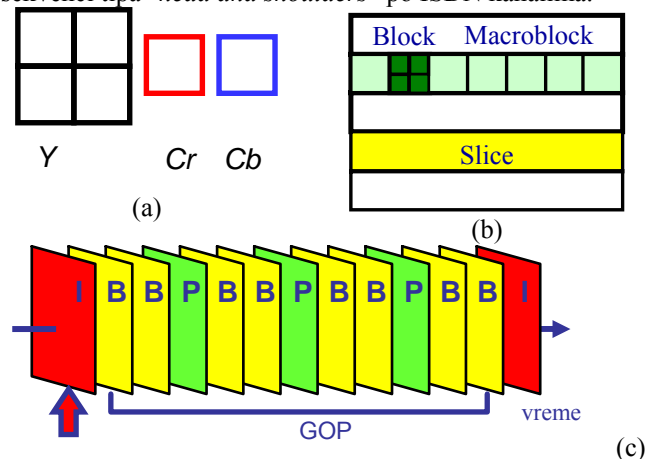
Nakon kompresije, signal sadrži znatno manje podataka u odnosu na izvorni, nekomprimovan signal, što je ilustrovano na slici 1. Nekomprimovani video signal (*stream*) se sastoji od *frame*-ova fiksne veličine, "Presentation units", a odgovarajući komprimovani *frame*-ovi, "Access units", su promenljive veličine, zavisno od sadržaja video materijala i od tipa kodovanja.

Kompresije video sekvenci su blok tipa, pri čemu su blokovi najčešće veličine 8x8 piksela (u H.264 kompresiji ta veličina varira). Vektori pokreta se određuju na nivou makrobloka. Kvalitet komprimovanog sadržaja zavisi u velikoj meri i od strukture makrobloka. U sekvencama čije će sadržaj biti analiziran u ovom radu, makroblokovima su

formata 4:2:0, što znači da na svaka 4 lumentna dolaze ukupno dva (Cr i Cb) hrominentna bloka, slika 2.a. Struktura MPEG sloja (*layer*), kao i grupe slika (GOP – *Group of Pictures*) u tim sekvencama, prikazani su na slikama 2.b. i 2.c, respektivno.

MPEG Forum je bio osnovni pokretač razvoja standarda za kompresiju pokretne slike. Prvi standard, MPEG-1, koji je specificiran od strane međunarodne organizacije za standardizaciju (ISO), bio je namenjen snimanju kvalitetnog video materijala na CD. Protoci kojima odgovara ovakav zapis su reda veličine 1.5MB/s. MPEG-2 je namenjen komprimovanju televizijskih signala kako u standardnoj rezoluciji tako i u televiziji visoke rezolucije (HDTV), te ima, do sada, najširu primenu. MPEG-4 je standard namenjen multimedijalnim aplikacijama i doživeo je deset verzija. Naredni MPEG standardi (MPEG-7 i MPEG-21) nisu klasični kompresioni, već definišu semantiku za opisivanje komprimovanih sadržaja, odnosno definišu načine pretraživanja i upravljanja njima.

Međunarodna unija za telekomunikacije (ITU) je prvo definisala način kompresije videokonferencijskih sadržaja, uobličeni u preporuci H.261. Ovaj se standard u žargonu naziva $p \times 64 \text{ kb/s}$, jer mu je osnovna namena bila prenos video sekvenci tipa "head and shoulders" po ISDN kanalima.



Slika 2. (a) Format odmeravanja; (b) Struktura layer-a u MPEG4 kodovanju; (c) Struktura GOP u analiziranim sekvencama.

Razvoj mobilnih telekomunikacionih sistema je nametnuo potrebu za definisanjem poboljšanog standarda visoke kompresije, koji bi omogućio prenos video signala pri veoma niskim protocima bita. Tako je, preporukom H.263, definisana kompresija sa ciljem što manji protok (čak ispod 32kb/s). Ova je preporuka dugo dopunjavana, te je doživela, pored osnovne, još dve verzije, H.263+ i H.263++.

U toku razvoja dve paralelne grane standarda (MPEG-x i H.26x), zajedničkim naporima ISO i ITU organizacija formirana su dva standarda: MPEG-2 (H.262) i MPEG-4 v.10 (H.264 AVC, *Advanced Video Coding*). Poslednji od ova dva

imao je zadatak da obezbedi što bolji kvalitet komprimovanog video sadržaja pri što manjem protoku.

Može se konstatovati da su u razvoju tehnika kompresije video sekvenci uočena tri pravca sa ciljevima: televizijski kvalitet (MPEG-2), "video konferencijski" kvalitet (H.261, H.263) i usavršeni multimedijalni (MPEG-4 v.10 ≡ H.264 AVC) koji ima za cilj da eventualno potisne MPEG-2, bar u onim aplikacijama koje su namenjene prenosu televizijskog signala po kanalima malog protoka. Za sada se predviđa primena H.264 u prenosu TV signala preko Interneta, odnosno u digitalnom TV prenosu za mobilne prijemnike (DVB-H, *Digital Video Broadcast – Handheld*).

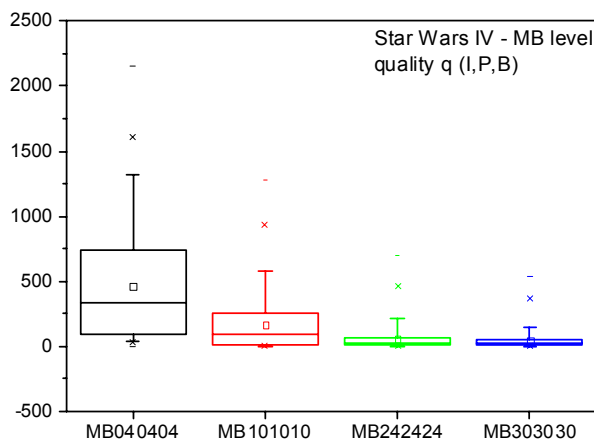
Očigledno je da je široka oblast kompresionih tehnika proizvela veliki broj standarda različitih namena (ne treba zaboraviti da su proizvođači softvera ovaj broj dodatno uvećali raznim nestandardnim rešenjima), a da je u centru pažnje, ipak, *protok proizvedenih saobraćajnih sekvenci* koje će opterećivati telekomunikacionu mrežu. Komprimovane video sekvence prate, takođe komprimovane, audio sekvence, kao i različite servisne informacije. Ipak, najveće opterećenje potiče od video sadržaja, te je jasno zašto postoji veliki broj radova koji se bave analizom video sekvenci sa stanovišta saobraćaja.

U ovom radu su analizirane video sekvence komprimovane prema MPEG-4 standardu. U drugom poglavlju su opisane osnovne karakteristike upotrebljenih sekvenci. Glava 3 daje rezultate simulacije u analizi na nivou makroblokova, za različite kvalitete kodovanja, kao i poređenje sa sekvencama iste kompresije na nivou frame-ova. Analiza je izvršena i za H.263 sekvence odgovarajućeg kvaliteta. Neke zaključne napomene su date u poglavlju 4.

2. MPEG-4 VIDEO SEKVENCE

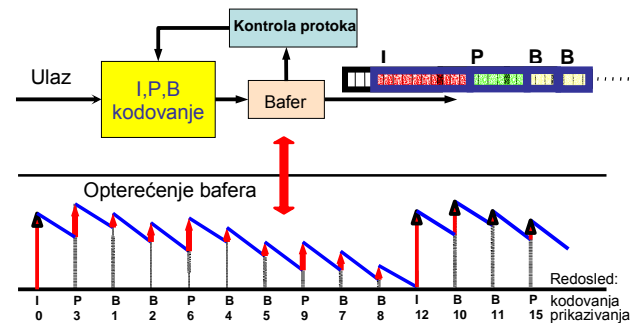
Za standard MPEG-4, koji je namenjen, pre svega, multimedijalnim sistemima, se kaže da je "defokusiran". On sadrži niz poboljšanja koja nisu usmerena ka jednom cilju, međutim, širok dijapazon novina koje uvodi ovaj standard nudi rešenja neophodna različitim aplikacijama. Na primer, kodovanje po ravnima objekata je osnov svih virtuelnih setova, kodovanje zasnovano na modelu je neizbežan alat u animacijama, itd. Standard H.264 AVC je, takođe, sastavljen od niza malih poboljšanja, a proizašao je iz MPEG-4.

Prostorna redundansa u slici se, kao što je poznato, umanjuje prema karakteristici ljudskog vizuelnog sistema (HVS). U tu svrhu se koristi matrica čiji sadržaj odražava karakteristike HVS. Koder video sekvenci, za zadati oblik



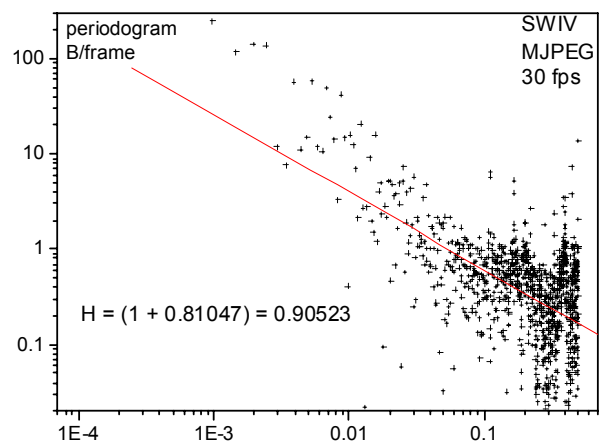
Slika 4. Box dijagram veličina makroblokova u MPEG-4 sekvencama filma «Star Wars IV» za različite faktore q .

GOP-a (najčešće korišćen je IBBPBBPBBPBB), na izlazu ima bafer pomoću kojeg se vrši kontrola protoka, slika 3. Za unapred definisanu veličinu protoka, kontroliše se ispunjenost bafera i, zavisno od toga, povratnom granom šalje informacija na osnovu koje se smanjuje ili povećava rezolucija komprimovanog sadržaja. Kvalitet kompresije se kontroliše korišćenjem faktora kvaliteta q (celobrojne vrednosti manje od 32). Za veće q , stepen kompresije je veći, odnosno, kvalitet komprimovanog videa je lošiji.

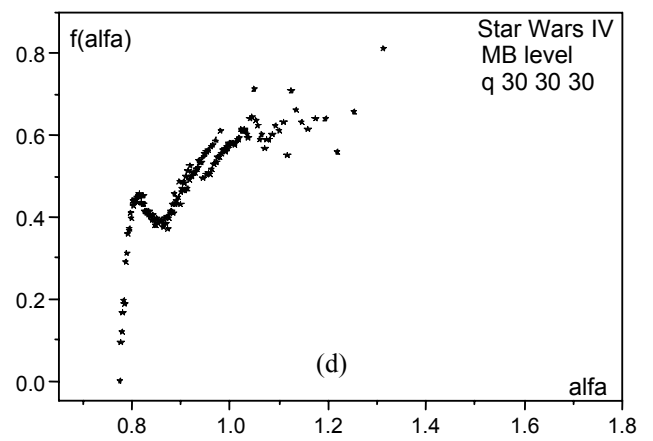
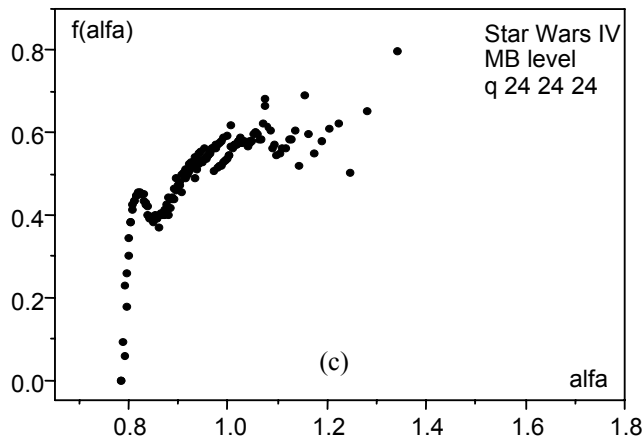
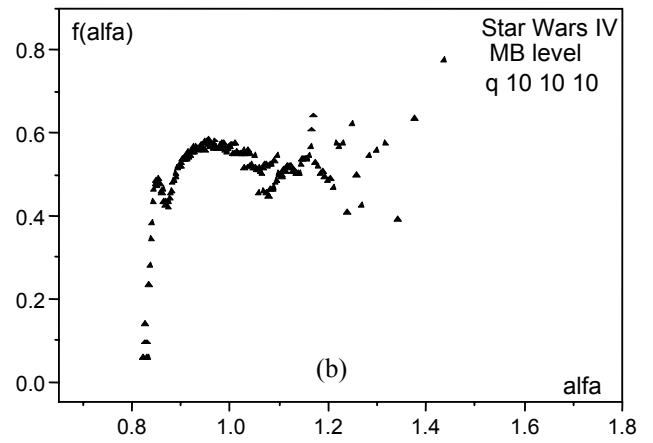
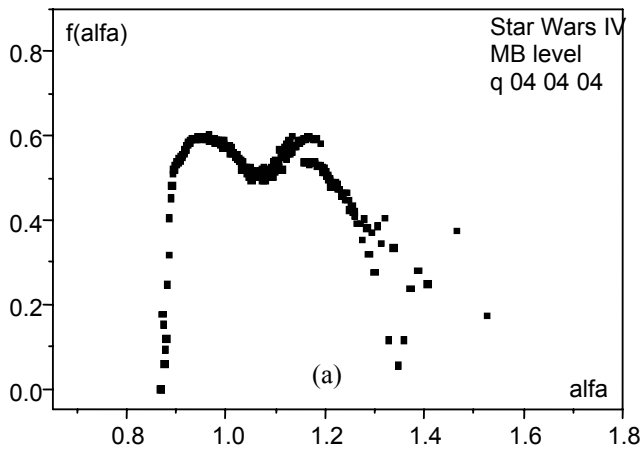


Slika 3. Kontrola protoka u postupku kompresije video sekvenci.

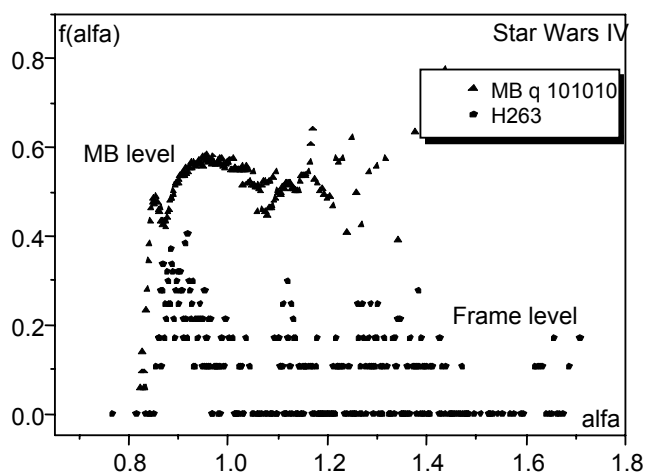
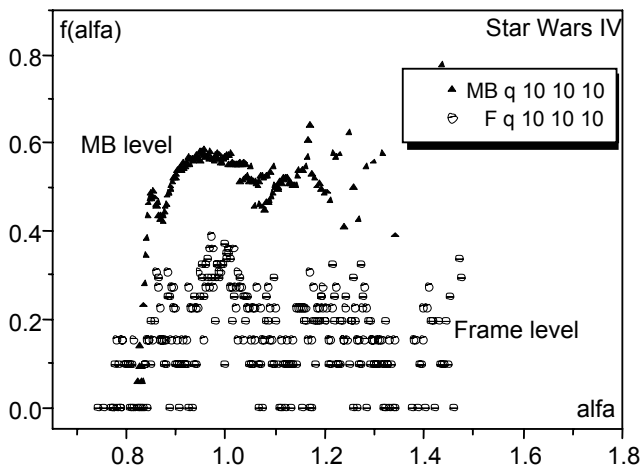
Video sekvenca filma «Star Wars IV» je, najpre, formirana u QCIF formatu (176x144 piksela lumnentnog signala i 88x72 piksela za svaku od hrominentnih komponenta) sa 8 bita po pikselu i 25 slika u sekundi. Ovako formirana digitalna sekvenca je, zatim, kodovana prema MPEG-4 standardu za četiri različita faktora kvaliteta kodovanja I,P,B frame-ova [1]. U tabeli 1 su navedeni statistički podaci ispitivanih sekvenci. U prethodnim radovima [2,3] su analizirane sekvence opisane brojem bajtova po *frame*-u, a ovde, u boljoj rezoluciji, brojem bita po makrobloku. Time se dobija precizniji podatak o saobraćajnim karakteristikama sekvenci, što je posebno važno sa stanovišta fraktalnih i multifraktalnih parametara. Prema podacima iz table 1, eksplozivnost sekvenci za $q=30$ je nekoliko puta veća od iste za $q=4$. Eksplozivnost upotrebljenih sekvenci se uočava i na *box* dijagramu, na slici 4. Naime, srednja vrednost broja bita po makrobloku je znatno veća od medijane, što ukazuje na postojanje izrazito velikih makroblokova, odnosno, upućuje na to da sekvence moraju imati izrazita fraktalna svojstva, odnosno, svojstva dugotrajne zavisnosti (*long-range dependency*) [4].



Slika 5. Periodogram intrakodovane video sekvence, digitalizovane sa 30 fr/s.



Slika 6. Multifraktalni spektri filma «Star Wars IV» za različite faktore q : (a) $q(4,4,4)$; (b) $q(10,10,10)$; (c) $q(24,24,24)$; (d) $q(30,30,30)$



Slika 7. Multifraktalni spektri MPEG-4 kodovanog filma «Star Wars IV» na nivou makroblokova, odnosno frame-a.

Slika 8. Multifraktalni spektar MPEG-4 kodovanog filma «Star Wars» na nivou makrobloka, odnosno H.263 na nivou frame-a.

U standardu H.263 se povećava efikasnost kompresije time što se, umesto B-slika, koriste PB koje su kombinacija dve slike: P- i B-kodovanih kao jedna jedinica. S obzirom da B-slika dopušta veliku kompresiju, ovakvo kodovanje omogućava povećanje brzine prenosa 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.

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 standarde kompresije uvedena skalabilnost koja može biti: kvalitativna (SNR = *Signal-to-Noise Ratio*), prostorna, vremenska, frekvencijska ili neka od njihovih kombinacija. U ovom radu su, međutim, korišćene sekvence sa jednim slojem.

3. ANALIZA KOMPRIMOVANIH VIDEO SEKVENCI

Sekvence koje su u ovom radu analizirane su dobijene sa mreže [1] i odnose se na komprimovani sadržaj QCIF formata, što znači da jedan *frame* sadrži 99 makroblokova (11x9). U analizi je korišćeno 100 GOP-ova, odnosno 1200 *frame*-ova (ili 118800 makroblokova) sekvenci koje su uzimane sa 25 fr/s. Poređenje je izvršeno i sa video sekvencom MJPEG tipa (*Motion* JPEG, bez kompenzacije pokreta) digitalizovanom sa 29.97 fr/s. Da bi se uporedio isti video sadržaj (vremensko trajanje video materijala), poslednja sekvenca je uzeta u dužini 1438 *frame*-ova.

Ispitivanjem vrednosti Hurst-ovog indeksa, prema [1-4], pokazuje se da analizirane video sekvence, pored fraktalnih karakteristika, poseduju izrazitu multifraktalnu komponentu. Multifraktalni spektri sekvenci različitih kvaliteta kompresije, prikazani na slici 6, to potvrđuju. Široka "parabola" multifraktalnih spektara se dobija za manje iznose kompresije (manje q). Pri tome je globalni maksimum oko vrednosti $\alpha=0.95$ Holder-ovog eksponenta, što ukazuje na postojanje singulariteta. Pri većim vrednostima q se uočavaju izrazitije

neravnomernosti u spektrima. Tu se izdvaja uzan lokalni maksimum pri $\alpha=0.85$ za koji se može pretpostaviti da odgovara I *frame*-ovima ili *frame*-ovima bilo kog tipa koji odgovaraju početku nove scene. Uočavaju se maksimumi koji, prema [2,3], potiču od slika različitog tipa (I,P,B u slučaju MPEG-4, odnosno I i PB u slučaju H.263 kompresije). Na slici 6.b. se uočava preraspodela maksimuma koji, u slučaju velikog stepena kompresije ($q=24,30$) počinju da se međusobno prepliću, slike 6.c i d.

Poređenjem dijagrama na slikama 7 i 8, uočava se da, u slučaju sekvenci izraženih brojem bajtova po *frame*-u, postoje istaknutiji maksimumi za koje je pokazano [2,3] da su posledica aditivnosti procesa, postojanja pojedinačnih *frame*-ova I,P,B, odnosno I,P,PB (u slučaju H.263 kompresije).

Eksplozivnost sekvenci slika (definisana maksimalnim brojem prema srednjem broju bita po makrobloku odnosno, bajtova po slici) izražena je pri velikim stepenima kompresije i u slučaju *intraframe* kodovanog video materijala.

Tabela 1. Statistički podaci ispitivanih sekvenci.

Kompresija	Oznaka	Skala faktor za kvantizaciju (q)			Hurstov indeks H	Sred. veličina <i>frame</i> -a ili makrobloka	Max veličina <i>frame</i> -a ili makrobloka	Eksplozivnost
		I <i>frame</i>	P <i>frame</i>	B <i>frame</i>				
MPEG-4 Makroblok nivo	(04,04,04)	4	4	4	0.91613	459	2158	4.701
	(10,10,10)	10	10	10	0.90586	163	1282	7.865
	(24,24,24)	24	24	24	0.84739	53	697	13.15
	(30,30,30)	30	30	30	0.83587	39	540	13.846
MPEG-4, <i>frame</i> nivo		10	10	10	0.94557	747	4690	6.278
H.263, <i>frame</i> nivo		var	var	PB	>> 1	1987	8969	4.514
Intrakodovan MPEG-4		10	-	-	0.90523	2032	23064	11.35

4. ZAKLJUČAK

Priroda video sekvenci je nelinearna tako da je za analizu saobraćaja na izlazu video kodera pogodna primena neke od nelinearnih metoda, na primer fraktalna, odnosno, multifraktalna analiza. Doista, pokazuje se da Hurst-ov indeks ima vrednosti u opsegu od 0.83-0.95 što ukazuje na fraktalni proces [4]. Detaljnija ispitivanja fraktalnih svojstava sekvenci, iskazanih promenama vrednosti Hurst-ovog indeksa na različitim vremenskim skalama [1-3], ukazala su na postojanje multifraktalne prirode u komprimovanim video sekvencama.

Multifraktalni spektri potvrđuju rezultate dobijene u prethodnim radovima [2-3] za druge tipove kompresije. Ovdje je, dodatno, prvi put analiziran MF spektar jedne MPEG-4 kodovane video sekvence na nivou makroblokova.

LITERATURA

[1] 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.

[2] I. Reljin, B. Reljin, "Fractal and multifractal analyses of compressed video sequences", *FACTA UNIVERSITATIS* (Niš), Series: Electr. Eng., Vol. 16, Dec. 2003, pp. 401-414.

[3] I. Reljin, B. Reljin, "Multimedia traffic performances – Different compression techniques" (invited paper), in *Proc. 6th Int. Conf. on Telecomm. in Modern Satellite, Cable and Broadcasting Services* (TELSIKS-03), Niš (Serbia and Montenegro), Oct. 1-3, 2003, pp. 369-376.

[4] C. Evertsz, B. Mandelbrot: "Multifractal Measures", Appendix B in H. Peitgen, H. Jurgens, P. Andrews *Chaos and Fractals*, Springer, 1992.

Abstract – The paper analysed the video sequences coded according to MPEG-4 standard. The number of bytes-per-frame, as well as bytes-per-macroblock, in video sequences was analysed through non-linear methods: the fractal and multifractal analyses.

ANALYSIS OF MPEG-4 VIDEO SEQUENCES

Irina Reljin, Branimir Reljin, Marija Zajeganović-Ivančić