

Simulacija MPEG kodovanja audio signala

Jovana Pavlović

Sadržaj — U ovom radu je opisan postupak i način kodovanja audio signala za model 1 nivoa I. Ovo je osnovni postupak kodovanja koji je suštinski isti i za više nivoa modela 1. Rezultati simulacije i analize obuhvataju poređenje stepena i uštede kompresije za pet različitih audio signala. Na kraju je dat grafički prikaz uštede kompresije za svaki podopseg, predstavljen na primeru jednog od pet analiziranih signala.

Ključne reči — Kompresija po podopsezima, maskiranje, prag čujnosti, prag maskiranja, stepen kompresije, ušteda kompresije.

I. UVOD

RADI ostvarenja kvalitetnijeg i efikasnijeg prenosa digitalnog audio signala javila se potreba redukcije veličine audio fajlova [1]. Ovaj postupak je zasnovan na percepciji čovekovog auditornog sistema i odnosi se na otklanjanje komponenata koje su irelevantne za čovekovo uho. Čovekovo uho ne registruje komponente koje su maskirane ili ispod praga čujnosti, pa te delove izvornog signala nije potrebno prenositi [2].

Simultano maskiranje u frekvencijskom domenu je pojava u kojoj se signal niskog nivoa (maskirani) može učiniti nečujnim (maskiranim) od strane istovremeno postojećeg jačeg signala (maskera) ukoliko su dugo i masker i maskirani signal dovoljno bliski frekvencijski [1] i [2]. Ukoliko signal sadrži puno simultanih maskera, svaki od njih ima svoj prag maskiranja, te se može odrediti globalni prag maskiranja [1]. Na povećanje simultanog maskiranja utiče vremensko maskiranje koje se javlja kada se dva, po intenzitetu različita zvuka, pojave u bliskim vremenskim intervalima.

U radu [3] su prikazani rezultati maskiranja signala koji se sastoji od 8 prostoperiodičnih komponenata. Takav signal je odabiran sa učestanošću od 44.1KHz pri protoku od 128Kb/s. U okviru [3] je odrađena simulacija kodovanja na osnovnom nivou, a za potrebe ovog rada, preuređen je kod za kompresiju čime je omogućena analiza složenih, realnih audio sekvenci. Izračunat je ukupan broj bita odabranih sekvenci pre i nakon kompresije sa uklonjenim pragom čujnosti, kao i nakon totalne kompresije, što je dalje iskorišćeno za računanje stepena i uštede. Najzad, izračunata je ušteda za svaki podopseg jednog od signala i dat njen grafički prikaz.

U poglavlju II, tretira se problem kompresije audio signala i data je gruba ilustracija postupka obrade. U poglavlju III, grafički su prikazane etape signala tokom odbacivanja audio komponenata za koje se očekuje da

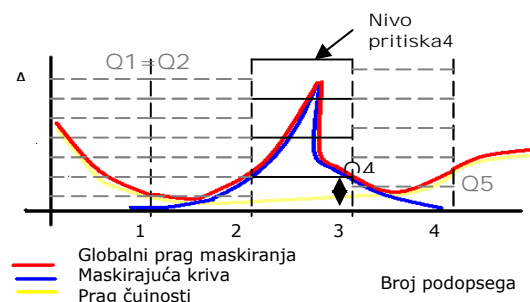
budu maskirane. Na osnovu metoda [3], komprimovani su signali pet raspoloživih audio klipova na kojima su, po obavljenim simulacijama, izvršeni dalji proračuni.

II. POSTAVKA PROBLEMA

Umesto signala generisanog u radu [3], može se učitati zvučni signal *wave* tipa i izvršiti uklanjanje suvišnih spektralnih komponenata uz grafički prikaz obrade signala [4]. U radu je obrađeno pet zvučnih signala koji su prethodno odabirani sa učestanošću od 44.1KHz pri bitskoj brzini od 128Kb/s, podešeni na mono signal i moraju biti *wave* tipa. Paralelno je svaki signal prebačen u MP3 verziju da bi se poredile odgovarajuće veličine fajlova i radi utvrđivanja razlike zvučnih predstava dve verzije istog signala.

Obrada signala radi kompresije zahteva određivanje individualnih pragova maskiranja, dok se globalni prag maskiranja dobija objedinjavanjem individualnih. Odgovarajuće komponente signala koje su značajno ispod praga globalnog maskiranja podopsega ne moraju biti prenešene. Broj bita po podopsegu kvantizera bira se tako da se šum kvantizacije zadržava značajno ispod praga globalnog maskiranja. Za svaki podopseg određuje se SMR [dB] (razlika između nivoa maskera i minimalnog praga maskiranja), jer raspodela bita koje bi trebalo preneti po svakom podopsegu grubo prate SMR.

Za svaki ulazni signal vrši se kvantizacija, tj. komponente signala se nakon odabiranja smeštaju na odgovarajući kvantizacioni nivo pri čemu je svaki nivo determinisan jedinstvenom bitskom predstavom. Takvi signali se mogu međusobno porediti, pa je moguće i poređenje pragova maskiranja i čujnosti signala.



Sl.1. Formiranje globalnog praga maskiranja uzimanjem veće od dve poredene vrednosti.

Na sl. 1. je slikovito prikazan način dobijanja globalnog praga maskiranja od praga čujnosti i praga maskiranja. Za svaki podopseg, na odgovarajućim frekvencijama duž opsega, se poredi nivo praga čujnosti i praga maskiranja i za vrednost globalnog praga uzima veća od dve poredene vrednosti. Globalni prag maskiranja direktno utiče na veličine kvantizacionih koraka. Radi dodatne redukcije

broja bita potrebnih za kodovanje signala, određuje se SGMR (odnos signala i globalnog praga maskiranja):

$$SGMR = L_{sb} - L_{Tmin} \text{ [dB]} \quad (1)$$

gde je L_{sb} maksimalni nivo zvučnog pritiska za dati podopseg, dok je L_{Tmin} minimum globalnog praga maskiranja za dati podopseg. Broj bita se računa prema relaciji:

$$N_{bits} = \frac{SGMR * \log_2(10)}{20} \quad (2)$$

III. ANALIZA

A. Postupak obrade signala

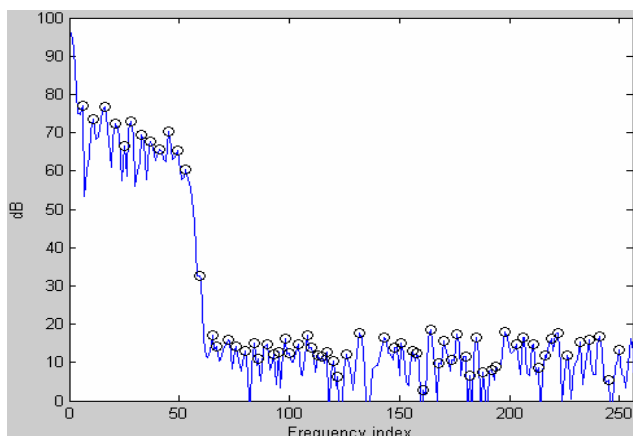
Nakon učitavanja audio signala i odabiranja na učestanosti od $f_s = 44100$ Hz, obrada signala se obavlja u više koraka.

Prvo se prema standardizovanoj tabeli postavljaju frekvencije, određuju se veličine kritičnih opsega i apsolutni pragovi. Ovaj deo se zapravo odnosi na podešavanje analize signala prema percepciji čula sluha, jer je utvrđeno da ono „raspoređuje“ signale prema kritičnim opsezima, a ne prema opsezima jednakih širina. Takođe se određuje apsolutni prag čujnosti ispod koga čovekovo čulo sluha ne detektuje signale.

Zatim sledi postavljanje indeksa za granice kritičnih opsega u tabele apsolutnih pragova za odgovarajući nivo (nivo I u ovom slučaju).

Svaki od podopsega se prekriva prozorom od 512 $C[i]$ koeficijenata i u njemu se obavlja analiza signala. Ovaj, kao i prethodna dva koraka, koriste tabelarne vrednosti definisane prirodom čula sluha i načinom obrade signala.

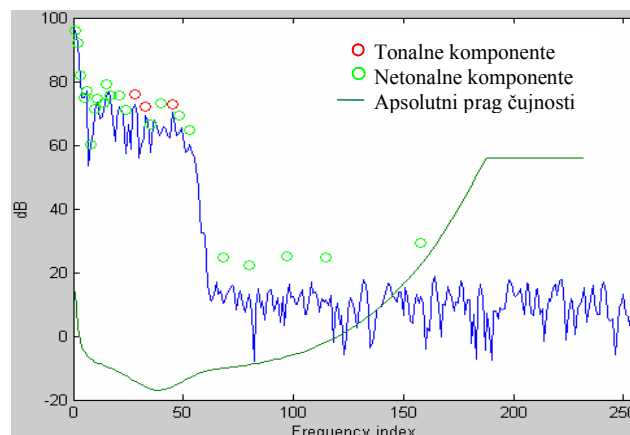
Analiza u okviru samog podopsega vrši se za svaki podopseg posebno (postupak opisan u [2]). Za obradu je potrebno 512 elemenata, zbog čega je prethodno definisano 512 $C[i]$ koeficijenata.



Sl.2. Lokalni maksimumi signala za jedan podopseg.

Dalja obrada obuhvata skaliranje odbiraka podopsega. To znači da se određuje maksimum apsolutnih vrednosti svih 12 odbiraka i prema najvećoj vrednosti se za taj kritični opseg bira kvantizacija sa odgovarajućim korakom. Informaciju o izboru nosi faktor skaliranja. (dinamička raspodela bita [1] i [2].) Postoji se u [3] obrađuje nivo I, onda svaka grupa od 12 odbiraka po podopsegu ima svoj faktor skaliranja.

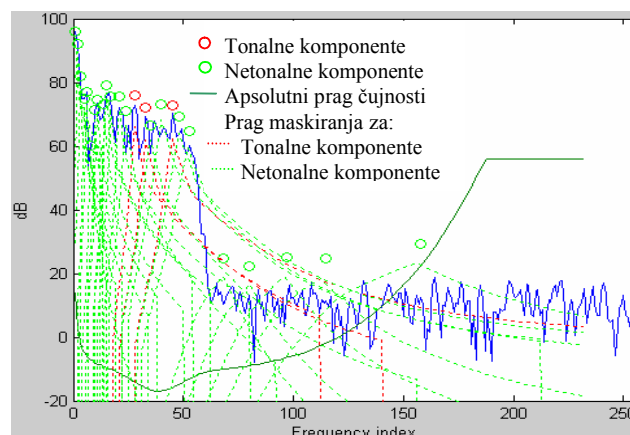
Naredni korak se odnosi na proces obrade u bankama filtera, dok se paralelno otklanjaju komponente ispod praga čujnosti i maskirane komponente (gornja i donja grana šeme strukture enkodera [1]). Stoga je potrebno izvršiti FFT (Fast Fourier Transform) analizu, koja daje spektar audio signala. Sledi centriranje 384 odbiraka od 512, jer je za analizu signala u bankama filtera potrebno 512 elemenata obrade, dok opseg zahteva $12 \times 32 = 384$ odbiraka.



Sl. 3. Tonalne komponente i apsolutni prag čujnosti.

Korak koji sledi u 32 podopsega određuje nivo zvučnog pritiska (sl. 2.) sa uklanjanim pragom čujnosti. Ovako dobijeni signal se može tretirati kao signal obrađen kompresijom bez gubitaka (zadržane su sve komponente u opsegu čujnosti 20-20000Hz).

Dalje se detektuju i obeležavaju lokalni maksimumi (sl. 2.). Neki od njih su tonalne komponente (tonovi), dok su drugi netonalne komponente (ponašaju se kao šum). Ovo je, zajedno sa apsolutnim pragom čujnosti, prikazano na sl. 3.



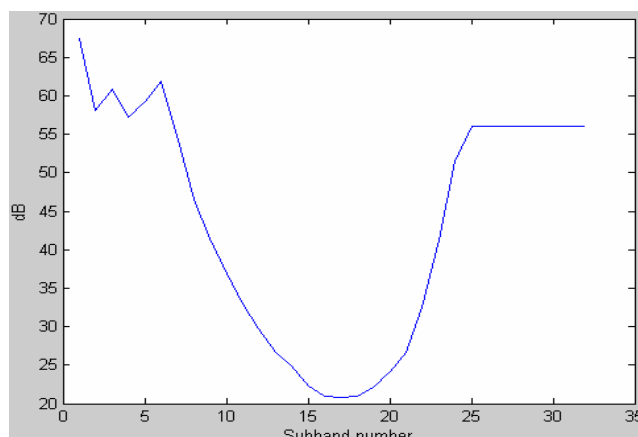
Sl. 4. Eliminisanje bliskih komponenta ispod praga maskiranja tonalnih i netonalnih komponenta

U ovom delu analize se odstranjuju po intenzitetu slabije komponente, ukoliko su one udaljene od susednih jačih komponenta za ne više od jedne polovine širine kritičnog opsega. Na taj način se u velikoj meri smanjuje količina podataka koji se dalje obrađuju. Uklonjene su samo komponente koje su sigurno irelevantne za čovekovo uho, ali ne i sve maskirane.

TABELA 1: REZULTATI ANALIZE.
FAJL 1-DRWHO, FAJL 2-BANDIDO, FAJL 3-ESENA MONO, FAJL 4-MEXIKO, FAJL 5-KALAJDIZIJSKO KOLO

Talasni fajl	Simultani rezultati za 5 različitih talasnih fajlova								MP3 Kompresija			
	Odbacivanje praga čujnosti (Sound pressure level)				Globalno maskiranje (Global masking)							
Ukupni broj bita	Nakon kompresije	Pre kompresije	Ušteda kompresije (%)	Stepen kompresije	Nakon kompresije	Pre kompresije	Ušteda kompresije (%)	Stepen kompresije	Nakon kompresije (MB)	Pre kompresije (MB)	Ušteda kompresije (%)	Stepen kompresije
Fajl 1	4417140	2331036	81.05	5.2772	898380	2331036	96.15	25.9471	0.518	2.77	81.3	5.3475
Fajl 2	11856684	21190656	44.05	1.7872	1417752	21190656	93.31	14.9467	0.47	2.52	81.35	5.3617
Fajl 3	5709384	12730368	55.15	2.2297	835284	12730368	93.44	15.2408	0.283	1.51	81.26	5.3357
Fajl 4	13626336	26130432	47.85	1.9176	1738536	26130432	93.35	15.0301	0.579	3.11	81.38	5.3713
Fajl 5	4687620	9873408	52.52	2.1063	709428	9873408	92.81	13.9174	0.22	1.17	81.19	5.3182

Radi dobijanja globalnog praga maskiranja, potrebno je odrediti individualne pragove maskiranja za tonalne i za netonalne komponente (sl. 4.), kao i njihov uticaj na susedne spektralne frekvencije. Poznato je iz [2] da tonalne komponente u odnosu na netonalne imaju „uži“ prag maskiranja, što ukazuje da je šum (ili netonalna komponenta) bolji masker od tona.



Sl. 5. Minimalni prag maskiranja.

Dobijanje globalnog praga maskiranja je opisano u II i slikovito prikazano na sl. 1. Ukoliko se u svakom podopsegu odredi minimum globalnog praga maskiranja, onda minimumi svih 32 podopsega zajedno čine minimalni prag maskiranja za dati opseg (sl. 5.).

Razlika lokalnih maksimuma (nivoa zvučnog pritiska) i minimalnog praga maskiranja (jednačina (1)) su komponente koje čine komprimovani zvučni signal (kompresija sa gubicima). Pomoću SGMR-a se prema jednačini (2) može odrediti potreban broj bita za kodovanje signala.

Završni korak obrade daje ukupan broj bita nekomprimovanog signala, kao i ukupni broj bita komprimovanog signala, i to posebno za kompresiju sa i bez gubitaka. Ukupan broj bita se računa za svaki podopseg posebno. Sumiraju se sve dobijene vrednosti, pri čemu se određuje i broj prozora analize. Broj bita po svakom podopsegu zavisi od SGMR-a (jednačina (2)). Ukoliko se dobiju negativne vrednosti one se izjednače sa

nulom, a ostale se zaokružuju na najbližu celobrojnu vrednost. Svaki podopseg komprimovanog signala se koduje brojem bita koji je određen odnosom signal/maska. Za dobijanje ukupnog broja bita nekomprimovanog signala, uzima se da je za jedan odbirak potrebno 16 bita. Broj bita za svaki podopseg posebno moguće je odrediti na isti način kao i za ukupan broj bita.

$$\text{stepenkompresije} = \frac{\text{nekomprimovanavel}}{\text{komprimovanavel}} \quad (3)$$

$$\text{uštedakompresije} = 1 - \frac{\text{komprimovanavel}}{\text{nekomprimovanavel}} \quad (4)$$

Komprimovane i nekomprimovane veličine su veličine fajlova MP3 i wave signala (ukupan broj bita), odnosno ukupan broj bita sračunatih analizom za komprimovani i nekomprimovani signal.

B. Rezultati analize

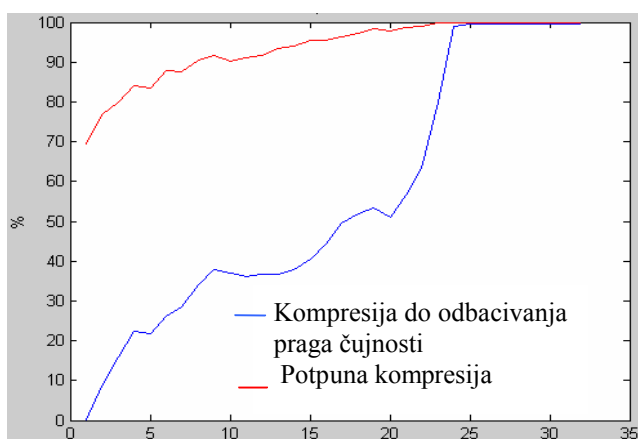
Prvi deo rezultata analize, Tabela 1, odnosi se na kompletne signale. Za svaki signal je sračunat ukupan broj bita pre i nakon kompresije signala, stepen kompresije i ušteda na čitavom opsegu signala. Rezultati su dati u Tabeli 1 za pet obrađenih signala i poređeni sa rezultatima savremene MP3 kompresije. Za MP3 kompresiju vrednosti se dobijaju na osnovu broja bita MP3 i wave fajlova prema obrascima (3) i (4).

Prema Tabeli 1, kompresija bez gubitaka ima uštedu kompresije 40-55% za sve fajlove osim za prvi (oko 80% što se može obrazložiti prirodom samog signala kome se već otklanjanjem praga čujnosti otklanja značani broj bita). Analiza koja je obavljena za potpunu kompresiju signala (ušteda kompresije više od 93%) pokazuje slične rezultate sa MP3 kompresijom (ušteda oko 81.3%). Mala razlika rezultata uslovljena je drugačijom organizacijom fajlova. MP3 sadrži dodatne bite koji u korišćenoj zvučnoj realizaciji nisu uzeti u obzir, te stoga MP3 ima veći koeficijent kompresije. Uočimo da efekat maskiranja značajno povećava uštedu kompresije, a broj kodovanih bita smanjuje, i više od 5 puta u odnosu na kompresiju bez gubitaka.

TABELA 2: REZULTATI ANALIZE ZA FAJL 3-ESENA MONO ZA NEKE OD 32 PODOPSEGA SIGNALA (RSB - ODNOS BROJA BITA NAKON KOMPRESIJE I UKUPNOG BROJA BITA U %, SSB - UŠTEDA KOMPRESIJE TAKOĐE U %, CRSB - STEPEN KOMPRESIJE ZA SVAKI PODOPSEG).

podo pseg	Totalna kompresija za svaki podopseg			Kompresija otklanjanjem praga čujnosti za svaki podopseg		
	Rsb (%)	Ssb (%)	CRsb	Rsb (%)	Ssb (%)	CRsb
1	30.5532	69.4468	3.2729	99.9729	0.0271	1.0003
4	15.8663	84.1337	6.3027	77.6394	22.3606	1.2880
8	9.6193	90.3807	10.3958	65.9568	34.0432	1.5161
12	8.2951	91.7049	12.0553	63.3416	36.6584	1.5784
16	4.5216	95.4784	22.1161	55.7010	44.2990	1.7953
20	2.1175	97.8825	47.2255	48.8688	51.1312	2.0463
24	0	100.0000	-	0.9291	99.0709	107.63
28	0	100.0000	-	0.5460	99.4540	183.15
32	0	100.0000	-	0.5430	99.4570	184.16

Rezultati dobijeni potpunom kompresijom signala daju ne baš realne vrednosti, a razlog su izvršene aproksimacije prilikom obračunavanja broja bita i činjenice da je izvršena gruba analiza signala. Razlika u broju bita za isti fajl koje očitava računar i ukupnog broja bita sračunatih prilikom obrade signala potiče takođe od aproksimacija prilikom vršenja proračuna.



Sl. 6. Ušteda kompresije za 32 podopsega za Fajl 3-Esena mono.

Drugi deo rezultata analize je prikaz uštede za svaki podopseg posebno i dat je grafički prikaz krivih za jedan od obrađenih signala. Na sl. 6. vidi se značajna razlika uštede kompresije za dva posmatrana slučaja. Za neke od 32 podopsega rezultati su dati tabelarno. Oni pokazuju veću uštedu na višim nego na nižim frekvencijama. Razlog je to što je maskiranje ka višim frekvencijama veće pa se samim tim u tom delu otklanja više maskiranih komponenata ([1]-prag maskiranja ka nižim frekvencijama je strmiji nego ka višim). Stepen kompresije za pojedine podopsege je beskonačan (tabela 2), što je takođe posledica aproksimacije negativnih vrednosti broja bita nulom u samom procesu proračuna. Realno je moguće da neki podopsezi na najvišim frekvencijama budu kompletno maskirani i da im se mogu ukloniti svi odbirci, ali bi ovu analizu više trebalo shvatiti aproksimativno.

III. ZAKLJUČAK

Rezultati obavljene analize pokazuju da se obradom zvuka zasnovanoj na audio percepciji značajno redukuje količina informacija koje se prenose (Tabela 1). Odbacivanjem maskiranih komponenata ukupan broj bita se može smanjiti više puta u odnosu na nekomprimovani signal. Ka višim frekvencijama maskiranje je izraženije, pa u toj oblasti značajno raste ušteda (sl. 6.), tj. stepen kompresije sa porastom frekvencije takođe raste.

Postupak obrade signala korišćen u ovom radu je naknadno kroz više nivoe i model 2, modifikovan drugačijim grupisanjem odbiraka po podopsegu i drugačijom dinamičkom raspodelom bita. Time je ostvaren adekvatan stepen kompresije zvuka uz zadržavanje kvalitetnog audio doživljaja.

LITERATURA

- [1] Peter Noll, *MPEG Digital Audio Coding*, 2000 CRC Press LLC, <http://www.csie.ncnu.edu.tw/~hychen/multimedia/mpeg>.
- [2] Davis Pan, „A Tutorial on MPEG/Audio Compression“, *IEEE Multimedia*, <http://www.cs.ubc.ca/~krasic/cpsc538a-2005/papers/tutorial.pdf>, Summer 1995.
- [3] Fabien Petitcolas, program za simulaciju MPEG kompresije nivoa I, 1997-2008, www.cl.cam.ac.uk/~fapp2/software/mpeg
- [4] J. Pavlović, *Simulacija MPEG kodovanja audio signala*, diplomski rad, ETF, Beograd, sept. 2008.

ABSTRACT

This paper describes procedure and method of audio coding for first level of model 1. This is a basic coding process that is essentially the same for higher levels of model 1. The results of simulation and analysis include the comparison of the compression ratio and space savings for five different audio signals. The last figure shows graphic display of compression for each of the subbands, presented in the example of one of five analysed signals.

MPEG AUDIO CODING SIMULATION

Jovana Pavlović