

Jedno rešenje dinamičkog poboljšanja perceptualnog kvaliteta širokopojasnog audio signala u prisustvu buke

Petar Petrović, Robert Pečkai-Kovač, Stanislav Očovaj, Jelena Kovačević

Sadržaj — U ovom radu je prikazano jedno rešenje dinamičkog poboljšanja perceptualnog kvaliteta širokopojasnog audio signala u prisustvu buke na jednoj klasi procesora za digitalnu obradu signala uz korišćenje minimalnih resursa. Realizovani algoritam je zasnovan na adaptivnom širokopojasnom kompresoru, prilagođenom za rad u realnom vremenu na platformi sa ograničenim resursima. Opisana je struktura kompresora, kao i realizacija na DSP procesoru firme Cirrus Logic.

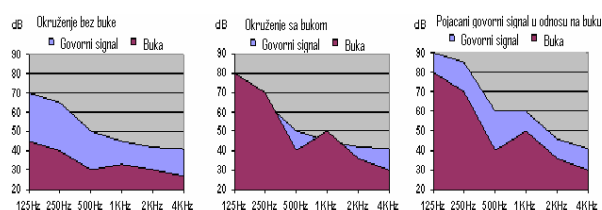
Ključne reči — DFT (Discrete Fourier Transformation), DSP (Digital Signal Processor), FFT (Fast Fourier Transformation), NMR (Noise Masking Ratio), PCM (Pulse Code Modulation), SNR (Signal To Noise Ratio)

I. UVOD

ALGORITAM za dinamičko poboljšanje perceptualnog kvaliteta širokopojasnog audio signala u okruženju sa bukom je namenjen prenosivim audio uređajima, kojih na tržištu ima sve više. Kada se prilikom korišćenja tih uređaja korisnik nađe u okruženju sa ambijentalnom bukom neki delovi audio signala mogu biti maskirani bukom [1]. U tom slučaju korisnik najčešće reaguje pojačavanjem muzičkog signala. Ako je intenzitet signala smetnje vremenski promenljiv, intenzitet zvuka će biti pojačan i kada to nije potrebno, što može dovesti do kraćeg trajanja baterija ili čak oštećenja sluha korisnika. Odnosno problem kod ručnog podešavanja intenziteta je vremenski promenljiv intenzitet signala smetnje, pa je potrebno podešavati intenzitet audio signala (muzika, film) tako da prati intenzitet signala smetnje. Na slici 1 je prikazano maskiranje govora, ambijentalnom bukom.

U ovom radu je dat opis rešenja za dinamičko podešavanje intenziteta audio signala koji analizira karakteristike signala smetnje i rezultate analize primenjuje na sadržaj ulaznog audio signala, tako da i

najtiši delovi ulaznog signala ne budu maskirani bukom. Ovo se postiže kombinacijom nekoliko psihoakustičkih efekata, karakteristika ljudskog slušnog sistema i karakteristika slušalica.



Sli 1. Govor maskiran šumom

II. STRUKTURA I OPIS ALGORITMA

Korisnički scenario predstavlja upotrebu uređaja za digitalnu reprodukciju audio signala u okruženju sa bukom. Prisustvo pozadinskog šuma otkriva se mikrofonom ugrađenim u uređaj.

Poboljšanje perceptualnog kvaliteta audio signala [2] je zasnovano na frekventno zavisnom kompresoru, optimizovanom da obezbedi SNR (odnos signala i šuma) u onom frekventnom opsegu u kom je signal smetnje najdominantniji.

Prilikom obrade potrebno je uzeti u obzir i nekoliko psihoakustičkih efekata [3] [4]:

- Efekat maskiranja na istom audio kanalu
- Efekat maskiranja u susjednim kanalima
- Efekat vremenskog maskiranja
- Model audio kanala raspodeljen po bark skali

Opseg pojačanja je ograničen u nekoliko slučajeva, tako da pojačanje ne bi trebalo da naruši subjektivni osećaj tonskog balansa između visokih i niskih frekvencija, a takođe je potrebno izbeći prekomernu kompresiju. Korisnik ne bi trebao da oseti značajnu razliku u boji tonova kada se uspostavi frekventna ravnoteža.

Algoritam treba da ispuni dva zadatka. Prvi zadatak je da se analizira spektar ulaznog signala i proračunaju pojačanja u frekventnim podopsezima. Drugi zadatak je filtriranje ulaznog audio signala, filtrom čiji su koeficijenti izračunati koristeći prethodni proračun pojačanja.

Oba zadatka su zasnovana na diskretnoj Furijeovoj transformaciji (DFT). Blok dijagram algoritma je prikazan na slici 2. Ulazni signal $x(t)$ može da bude muzika, govor ili njihova kombinacija pri čemu je frekventni opseg ulaznog signala od 20Hz do 20000Hz a frekvencija

Ovaj rad je delimično finansiran od ministarstva za nauku Republike Srbije, projekat 161003, od 2008 god.

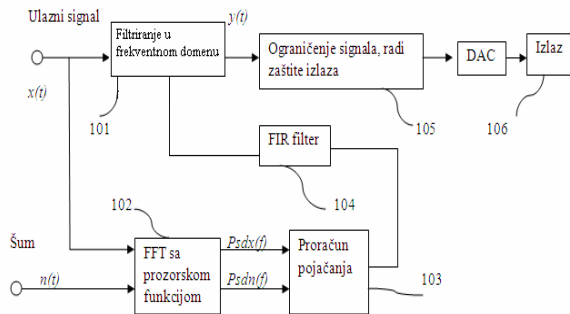
Petar Petrović, Fakultet tehničkih nauka u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija, (e-mail: petar.petrovic@rt-sp.com)

Robert Pečkai-Kovač, Fakultet tehničkih nauka u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija, (e-mail: robert.peckai-kovac@rt-sp.com)

Stanislav Očovaj, Fakultet tehničkih nauka u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija, (e-mail: stanislav.ocovaj@rt-sp.com)

Jelena Kovačević, Fakultet tehničkih nauka u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija, (e-mail: jelena.kovacevic@rt-sp.com)

odabiranja 48000Hz. Signal šuma $n(t)$ može biti otkriven posebnim mikrofonom ili postojećim mikrofonom celularnog telefona. Signal šuma ima isti frekventni opseg kao i signal $x(t)$.



Sl. 2 Blok dijagram algoritma

Spektralna analiza oba kanala je urađena u modulu 102. Signali su podeljeni u blokove veličine 1024 odbiraka, blokovi se preklapaju polovinom svoje veličine. Signal smetnje kao i muzički signal pomnožen je prozorskom funkcijom pre brze Furijeove transformacije (FFT). Izlaz modula 102 je spektar energije oba signala. Najznačajniji modul je modul 103, koji koristi prethodne rezultate za proračun pojačanja audio kanala.

U skladu sa psihoakustičkim modelom primenjenim u ovom rešenju [2], algoritam koristi 50 audio kanala raspoređenih prema bark skali. Modul 104 transformiše pojačanja dobijena u prethodnom modulu u željeni frekventni odziv. FIR filter sa željenim frekventnim odzivom je dobijen tehnikom prozoriranja. Koeficijenti filtera su osveženi za svaki novi blok podataka koji se sastoji od 1024 odbiraka.

Ulazni signal $x(t)$ je obrađen u modulu 101, FIR filtrom dobijenim u modulu 103. FFT obrada je izvršena nezavisno od modula 102. Nivo obrađenog signala $y(t)$ se ograničava u modulu 105 kako ne bi došlo do preopterećenja izlaznog zvučnog sistema.

III. OPIS FUNKCIONALNIH BLOKOVA ALGORITMA

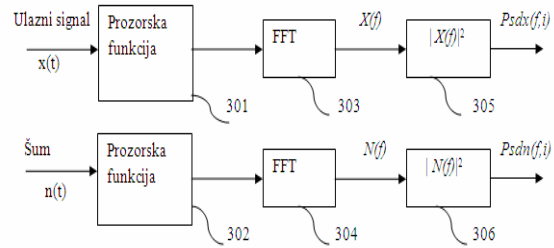
A. FFT filtriranje

Standardni pristup rešavanju problema je korišćenje banke filtera, kod koga problem dostizanja odgovarajućeg frekventnog odziva zahteva puno proračuna. Ovdje je primenjeno rešenje sa jednim FIR filtrom, čiji se koeficijenti računaju sa svakim novim blokom podataka prema željenom frekventnom odzivu. Za okruženja gde je šum niskog nivoa modul 103 kreira FIR filter sa ravnim frekventnim odzivom koji neće izmeniti originalni muzički signal. Realizovani filter obezbeđuje strm frekventni odziv kako bi smanjio uticaj filtera na susedne kanale [2]. Kako bi se smanjila izobličenja izazvana naglim promenama između susednih blokova primenjuje se dvostruko filtriranje sa prethodnim i novim koeficijentima. Izlazni signali ova dva filtriranja množe se prozorskom funkcijom, a zatim sabiraju.

B. FFT analiza

Ovaj modul računa gustinu spektra energije muzičkog

signala $x(t)$ i signala šuma $n(t)$. Pre FFT obrade nad ovim signalima primenjena je prozorska funkcija (Hann prozorom) slika 3, kako bi se potisnule komponente nastale curenjem spektra.



Sl. 3 FFT analiza ulaznog signala i signala šuma

Izlaz ovog modula su kvadrati koeficijenata diskretne Furijeove transformacije (DFT) koji predstavljaju gustinu spektra energije ulaznog signala i signala šuma.

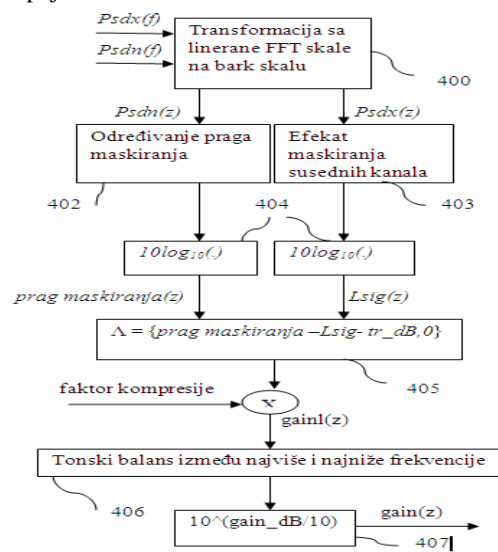
C. Određivanje pojačanja

Na osnovu modela ljudskog slušnog sistema, osnovne frekvencije audio signala su raspoređene po bark skali. Frekventni raspon od 20Hz do 20000Hz je podeljen na 25 opsega (barka). U ovom rešenju koristi se 50 audio kanala sa širinom od 0,5 barka. Svaki frekventni blok je preslikan u odgovarajući audio kanal. Snaga muzičkog signala i šuma se računa na samom početku algoritma.

Pored praga maskiranja šuma uzet je u obzir i efekat maskiranja susednih kanala u modulu 403, slika 4. Za ovu svrhu izračunata je ovojnica $P_{sd_x}(z)$.

Prilikom računanja prve aproksimacije pojačanja, primenjene su dodatna konstanta za pojačanje, dobijena na osnovu apsolutnog praga čujnosti, dodatni prag za pojačanje signala (tr_dB) i frekventno zavisna faktor kompresije.

Tonski balans primenjen u modulu 406 predstavlja poboljšanje zasnovano na ukupnom nagibu krive $gain(z)$. Ako je nagib negativan to znači da će niske frekvencije biti pojačane, ali kako bi se održala originalna boja ulaznog muzičkog signala i visoke frekvencije se moraju dodatno pojačati.



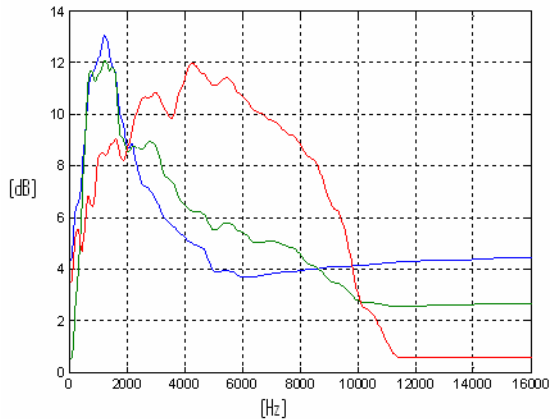
Sl. 4 Proračun pojačanja

D. FIR filter

Nad željenim frekventnim odzivom se primenjuje:

- inverzna FFT obrada,
- prozorska funkcija, i
- FFT obrada.

Amplitudska karakteristika FIR filtra u prisustvu različitih tipova šuma su prikazane na slici 5.



Sl. 5 Amplitudska karakteristika FIR filtra u prisustvu tri različita tipa pozadinskog šuma

IV. OPIS REALIZACIJE

Algoritam je realizovan u tri faze, pri čemu svaka faza ima odgovarajući model [5]. Model 1 predstavlja C kod sa aritmetikom sa pokretnim zarezom, prilagođen prelasku na aritmetiku sa fiksnim zarezom i daje iste rezultate kao referentni Matlab kod. Model 2 predstavlja C kod sa aritmetikom sa fiksnim zarezom koja unosi zanemarljivu grešku prilikom obrade u odnosu na model 1, dok model 3 predstavlja asemblerski kod koji daje identične rezultate kao model 2.

Opisani algoritam je realizovan i ispitivan na 32-bitnom DSP procesoru (Cirrus Logic CS47048) sa jednim jezgrom koji je prvenstveno namenjen prenosnim uređajima za reprodukciju digitalnog audio signala [6]. Ciljni procesor obezbeđuje 150 miliona instrukcija u sekundi (MIPS). Memorija je podeljena na programsku memoriju P i dve odvojene memorije X i Y za podatke. Programaska memorija i memorija za podatke Y dele isti memorijski blok pa se ukupna memorija procesora može konfigurisati na četiri načina, kao što je prikazano u tabeli 1.

Tabela 1: Memorijska mapa DSP-a

P-RAM	X-RAM	Y-RAM
14K words	10K words	8K words
12K words	10K words	10K words
10K words	10K words	12K words
8K words	10K words	14K words

Realizacija algoritma na procesoru namenjenom prenosivim uređajima zahteva: malu potrošnju, odnosno što manje procesorskog vremena (MIPS-a), i malu

memoriju odnosno što manju cenu krajnjeg uređaja. Proces obrade signala se sastoji od sedam funkcija, tabela 2. Najzahtevnije funkcije po potrošnji MIPS-a su filtriranje FIR filtrom sa novim koeficijentima (*ApplyNewFIR*) i sam proračun novih koeficijenta za FIR filter (*CalculateFIRCoefs*). U prvoj verziji koda ove funkcije su zajedno trošile 11,8 MIPS-a.

Tabela 2: Broj potrebnih MIPS-a za svaku funkciju

Funkcije	Potrošnja MIPS-a
GetSignalSpectrum	2.927
ApplyOldFIR	3.313
GetSignalPsd	2.975
GetNoisePsd	2.591
CalculateGains	2.590
CalculateFIRCoefs	4.466
ApplyNewFIR	3.514

Analiza koda je pokazala da su ove funkcije realizovane od kratkih programskih petlji, koje su se ponavljale veliki broj puta (npr. za svaki odbirak svakog od kanala) pa je svako skraćivanje ovih programskih petlji imalo značajan uticaj na brzinu izvršavanja.

Najveće uštede donelo je korišćenje sledećih optimizacionih tehnika [7]:

- izdvajanje provera pojedinih uslova i uslovnih skokova izvan programskih petlji, kao i
- skraćivanje programskih petlji većim korišćenjem paralelnog izvršavanja instrukcija

Nakon optimizacije koda, spomenute dve kritične funkcije su smanjene za oko 4 MIPS-a. Ukupno procesorsko vreme izvršavanja obrade je smanjeno na približno 20 MIPS-a što predstavlja 13.33 % iskorišćenosti procesorskog vremena, uz povećanje utroška memorijskih resursa.

Iskorišćenost memorije DSP procesora je prikazano u tabeli 3.

Tabela 3: Iskorišćenost memorije DSP-a

Memorija	Zauzeta memorija
X	6,189k
Y	7,265k
P	1,304k

V. ISPITIVANJE I VERIFIKACIJA

Predloženi algoritam ispitivan je u okviru nekoliko faza koje odgovaraju fazama razvoja.

Model 1 je ispitivan subjektivnom metodom merjenja kvaliteta obrađenog audio signala. Za potrebe ispitivanja korišćeno je nekoliko stacionarnih i nestacionarnih tipova šuma, kao što su:

- Buka u avionskoj kabini
- Pneumatski čekić
- Veći broj istovremenih razgovora
- Buka ulice

Jedna od najčešće korišćenih subjektivnih metoda merjenja kvaliteta obrađenog audio signala je metoda koja se zasniva na ispitivanju slušanjem, kod koje veći broj slušalaca ocenjuje kvalitet prikazanih ispitnih signala, koristeći ocene na skali od 1 do 5, gde 5 označava odličan kvalitet, dok 1 označava nezadovoljavajući, odnosno loš kvalitet ispitnog signala (tabela 4). Krajnja ocena kvaliteta dobija se računanjem srednje vrednosti ocene svih slušalaca. Ovako dobijena ocena obično se naziva „srednja ocena mišljenja“ (*Mean Opinion Score, MOS*). Preporuke za sprovođenje MOS ispitivanja definisane su u okviru standarda ITU-R BS 1284-1. Prosečan odnos šuma i maske (NMR) ispitnih signala je za 21,1 % bolji kod obrađenog audio signala u odnosu na ulazni signal.

Tabela 4: Skala subjektivne ocene kvaliteta govornog signala

Ocena	Nivo kvaliteta
5	Šum je potpuno neprimetan
4	Šum je delimično primetan
3	Šum je primetan ali nije ometajući
2	Šum je primetan i delimično ometajući
1	Šum je veoma primetan i ometajući

U drugoj fazi razvoja algoritma, ispitivanjem modela 2, u odnosu na model 1, eksperimentalnim putem je utvrđena bitska razlika sa maksimalnom greškom od jednog bita, pošto ljudsko uho u toku reprodukcije izlaznog signala ovu razliku ne može da oseti, je definisana dozvoljena razlika od 1 bita.

Ispitivanje modela 3 je izvršeno u realnom vremenu. Ovaj način ispitivanja korišćen je u toku razvoja kompresora i služio je za otkrivanje grešaka kao što su gubitak sinhronizacije, beskonačne petlje i služio je kao indikator da su svi kanali prisutni i obrađeni na izlazu kompresora. Ispitivanje u realnom vremenu u toku razvoja projekta izvršeno je posle svakog koraka optimizacije ili bilo koje druge promene u kodu. Izlazni audio signal sa ciljane platforme je sniman audio karticom „*Echo Digital Audio Fireworks*“ kako bi se uporedio sa izlaznim signalom referentnog C koda modela 2.

Pošto je utvrđeno da model 3 daje identične obrađene signale kao i model 2, zbog velikog broja ispitnih signala obavljena je automatska verifikacija tačnosti snimljenih podataka. Automatsko ispitivanje rešenja u realnom vremenu izvršeno je primenom ispitnog okruženja na principu crne kutije (Black Box Testing – BBT [8]). Operacije koje se izvršavaju tokom automatskog ispitivanja su:

- Učitavanje audio kompresora na razvojnu ploču
- Pripremanje referentnog audio signala
- Puštanje ulaznog i signala šuma na razvojnu ploču
- Snimanje audio signala sa DSP-a
- Poređenje audio signala obrađenog referentnim kodom i obrađenog signala sa razvojne ploče

Na osnovu rezultata automatskog ispitivanja, pri čemu je utvrđeno da rešenje na ciljnoj platformi daje identične rezultate kao referentni C kod u modelu 2 i MOS ispitivanja kod koga je prosečna ocena neobrađenog signala 1,2 dok je ocena obrađenog audio signala 3,3, ustanovljeno je da rešenje zadovoljava zadate kriterijume referentnog modela kao i ograničenja ciljne platforme.

VI. ZAKLJUČAK

U okviru ovog rada prikazana je jedna realizacija širokopolasnog audio kompresora na platformi sa ograničenim resursima (Cirrus Logic CS47048). Kao referenca korišćen je MATLAB kod koji obezbeđuje sve funkcionalnosti kompresora zadate specifikacijom.

Osnovni cilj bio je realizacija algoritma kojim se postiže prijatnije slušanje muzike u bučnom okruženju, namenjenog prenosnim audio uređajima.

Iako ovakva rešenja još nisu zastupljena na tržištu u velikoj meri, subjektivnim ispitivanjima, što potvrđuje i bolja NMR vrednost, utvrđeno je poboljšanje perceptualnog kvaliteta muzičkog signala.

LITERATURA

- [1] Meir Tzur (Zibulski), Alexander A. Goldin, "Sound Equalization in a Noisy Environment", Audio Engineering Society, 2001 May 12–15 Amsterdam, The Netherlands
- [2] Marina Bosi, "Perceptual Audio Coding", IEEE Signal Processing Magazine Sept. 1997
- [3] E. Zwicker and H. Fastl, "Psychoacoustics Facts and Models" Berlin: Springer Verlag 1990.
- [4] J. Thiemann „Acoustic Noise Suppression for Speech Signals Using Auditory Masking Effects“, Master thesis, Department of Electrical & Computer Engineering, McGill University, Montreal, Canada, 2001
- [5] Djukic Miodrag, Četic Nenad, Kovačević Jelena, Popovic Miroslav, "A C Compiler Based Methodology For Implementing Audio DSP Applications on a Class of Embedded Systems", ISCE, IEEE, ISBN: 978-1-4244-2422-1, 2008
- [6] Perovic Srdjan, Kovacevic Jelena, "Jedno rešenje radnog okruženja audio kodeka na platformi sa ogranicenim resursima", ISBN: 987-86-80509-62-4, Etran 2007.
- [7] Dejan Mišković, Miroslav Popović, Zoran Jovanović: Jedno rešenje optimizacije koda pomeranjem instrukcija bezuslovnih skokova i arhitekturom podržanih petlji, XLIVIII konferencija ETRAN, Čačak, 2004,
- [8] Nikola Teslic, Vukota Pekovic, Miodrag Temerinac, Uwe Strohbeck "On the Effectiveness of TV System Validation Based on Black Box Testing Technique," MicronasNIT, Novi Sad Institute of Information Technology Fruskogorska 11, SRB-21000 Novi Sad, Serbia

Abstract – This paper presents a short overview of one realization of wide bandwidth audio compressor based on DSP processor.

ONE SOLUTION OF THE DINAMIC IMPROVEMENT OF PERCEPTUAL QUALITY OF BROADBAND AUDIO SIGNALS IN THE PRESENCE OF NOISE

Petar Petrović, Robert Pečkai-Kovač, Stanislav Očovaj, Jelena Kovačević