

TEHNIKE FILTRIRANJA MUZIČKIH SIGNALA ZASNOVANE NA VREMENSKO-FREKVENCIJSKOJ REPREZENTACIJI

Darko Nikčević, Muzička Akademija Cetinje
Irena Orović, Srdjan Stanković, Elektrotehnički fakultet Podgorica
dn@cg.yu, irenao@cg.ac.yu, srdjan@cg.ac.yu

Abstract – U radu će biti predstavljene dvije tehnike filtriranja muzičkih signala: vremensko-promjenljivo filtriranje i soft-thresholding tehnika. Tehnike su zasnovane na korišćenju S-metoda kao vremensko-frekvencijskoj distribucije koja predstavlja pogodnu reprezentaciju za analizu ove vrste signala. Uspješnost primjene navedenih tehnika biće ilustrovana kroz primjer.

Abstract – The two procedures for musical signal filtering are introduced: time-varying filtering and soft-thresholding technique. Those techniques are based on S-method as time-frequency representation that represents the common representation for musical signal analysis. The efficiency of proposed procedures will be illustrated through the example.

Ključne riječi: vremensko-frekvencijsko filtriranje, vremenski-promjenljivo filtriranje, soft-thresholding

Keywords: time-frequency analysis, time-varying filtering, soft-thresholding

1. UVOD

Od samih početaka realizacije muzičkih snimanja, operatori se suočavaju sa problemima vezanim za pojavu slučajnog šuma koji vrši degradaciju korisnog signala. Pri analognom snimanju koriste se magnetne trake, koje imaju odnos signal-šum oko 60dB, što je nedovoljno za ostvarivanje snimka sa visokom vjernošću. Energija audio signala je uglavnom koncentrisana na niskim frekvencijama, ispod 2kHz. Međutim, visoke frekvencije, koje nose značajno manju energiju u poređenju sa niskim, su veoma značajne kako za razumljivost tako i za sveopšti utisak o kvalitetu datog audio signala. Visoke frekvencije su, zbog svoje male energije, podložne šumu tako da često veoma niski nivoi šuma trake mogu dovesti do značajne degradacije snimka. Da bi se poboljšale karakteristike analognih sistema za snimanje muzike, korišćene su razne vrste reduktora šuma, od kojih je najznačajniji Dolby. Ovaj sistem, koderom kojeg čini kombinacija kola preemfazisa i dinamičkog kompresora, pojačava energiju na visokim frekvencijama prije snimanja na magnetnu traku. Niske frekvencije uvijek nose dovoljno energije i na njima dolazi do maskiranja šuma, usljed čega ga uho gotovo i ne registruje. Pri reprodukciji, signal se obnavlja koristeći dekoder koji predstavlja kombinaciju odgovarajućih kola deemfazisa i dinamičkog ekspandera. Koder i dekoder moraju biti veoma pažljivo podešeni kako ne bi došlo do izobličenja. U praksi se koristilo nekoliko vrsta Dolby sistema koji su se, zahvaljujući popularnosti uređaja za analogno snimanje sa magnetnim trakama (koji su naročito dobri za snimanje bubnjeva), zadržali i dan danas. Najznačajniji su sistemi Dolby A, Dolby B i Dolby C [1].

Dolby A sistem je razvijen za profesionalne primjene. On dijeli spektar audio signala na četiri podopsega: prvi opseg obuhvata frekvencije ispod 80Hz, drugi je između

80Hz i 3kHz, treći opseg je iznad 3kHz, dok je četvrti iznad 9kHz. Ovaj sistem obezbjeđuje najveće pojačanje od 10 do 15dB u svakom podopsegu, u slučaju da nivo signala opadne za više od 45dB ispod maksimalnog nivoa snimanja. Dolby B i Dolby C sistemi su za svakodnevnu upotrebu i dijele spektar audio signala na samo dva podopsega. Pojačanja koja ostvaruju su 10dB i 20dB, respektivno.

Koristeći Dolby reduktore šuma vrhunski uređaji za analogno snimanje postizali su odnos signal-šum i do 90dB. Pojavom uređaja za digitalno audio snimanje, značajno su poboljšane performanse audio sistema. Na primjer, 20-bitni uređaji ostvaruju odnos signal-šum veći od 120dB. Napomenimo da je glavni problem digitalnih uređaja šum kvantizacije. Teorijski se odnos signal-šum kvantizacije računa po sljedećoj formuli [2]:

$$SNR = 6,02n - 10 \log_{10} \left(\frac{P_F^2}{3} \right) \quad (1)$$

gdje su: n – broj bita, dok je $P_F^2 = \frac{2^{n-1} Q}{\sigma_X}$ i predstavlja

vršni faktor analognog signala. Q je korak kvantizacije, a σ_X^2 predstavlja varijansu analognog signala.

Efekat šuma kvantizacije koji uho doživljava kao neprijatnu pojavu može se umanjiti dodavanjem tzv. dithera koji umanjuje odnos signal-šum sistema [3].

Na današnjem stepenu razvoja audio tehnologije možemo reći da su performanse sistema takve da ne zahtijevaju dodatna kola za redukciju šuma. Međutim, njihova primjena nije nimalo zapostavljena imajući u vidu potrebu za njihovim korišćenjem zbog velikog broja

snimaka koji se realizuju u standardnim studijima za snimanje muzike. Veoma značajna primjena reduktora šuma je i u restauraciji starih magnetofonskih analognih audio snimaka [4].

Reduktori šuma koji se danas gotovo isključivo koriste u audio-tehnici, rade na principima vremenski promjenljivih Wiener-ovih filtara i tzv. spektralnog oduzimanja, kod koga postoje dva osnovna tipa, oduzimač snage i oduzimač magnitude [2]. Njihove prenosne funkcije djeluju na spektar svakog segmenta kratkotrajne Furijeove transformacije signala. U praksi se koriste uglavnom Hammingov i Hanningov prozor, pri čemu se za realizovanje kratkotrajne Furijeove transformacije koristi 1024 ili 2048 bita. Uobičajeni problem koji se javlja pri njihovoj primjeni je pojava tzv. muzičkog šuma, koja značajno doprinosi ograničenju njihovog korišćenja [4], [5], [6], a u nekim slučajevima, kad je odnos signal-šum snimka relativno mali, restaurisani audio snimak je slabijeg kvaliteta od originalnog.

U ovom radu predlažemo novi pristup redukcije šuma zasnovan na vremensko-frekvencijskoj analizi i obradi signala [7]. Model šuma koji ćemo koristiti je, kao i uobičajeno u audio-tehnici, aditivni stacionarni bijeli šum čiji statistički parametri podliježu normalnoj raspodjeli [8].

2. TEORIJSKA OSNOVA

U ovom poglavlju biće dat osvrt na vremensko-frekvencijske reprezentacije korišćene u radu. Naime, filtriranje muzičkih signala biće izvedeno u vremensko-frekvencijskom domenu, korišćenjem S-metoda. Imajući u vidu da pruža kompletnu predstavu o pružanju svih harmonika različitih tonova simultano po vremenu i frekvenciji, združena vremensko-frekvencijska reprezentacija signala pokazuje se kao najbolje sredstvo za analizu muzičkih signala, a time i za konkretnu primjenu u filtriranju.

Podimo od Short Time Fourier-ove Transformacije (STFT) kao osnove mnogih vremensko-frekvencijskih distribucija. Korišćenjem funkcije prozora STFT se može definisati u diskretnom obliku kao:

$$STFT(n, k) = \sum_{k=-N/2}^{N/2} w(m)x(n+m)e^{-j2\pi mk/N} \quad (2)$$

gdje n i k predstavljaju diskretne promjenljive vremena i frekvencije, dok je $w(m)$ funkcija prozora širine m . Na osnovu STFT definiše se spektrogram kao najzastupljenija vremensko-frekvencijska reprezentacija signala:

$$SPEC(n, k) = |STFT(n, k)|^2 \quad (3)$$

Međutim, imajući u vidu da se muzički signali sastoje od većeg broja tonova koji dalje sadrže u sebi harmonike kao komponente na cjelobrojnim umnošcima osnovne

frekvencije, oni se tretiraju kao multikomponentni signali. S toga, ukoliko se komponente poklapaju u vremensko-frekvencijskom planu spektrogram postaje:

$$SPEC(n, k) = \sum_{m=1}^M |STFT_m(n, k)|^2 + \sum_{\substack{m=1 \\ m \neq n}}^M \sum_{n=1}^M STFT_m(n, k)STFT_n^*(n, k) \quad (4)$$

gdje je M broj komponenti. Drugi član u prethodnoj relaciji predstavlja kros članove. Pored pojave kros članova za slučaj multikomponentnih signala, nedostatak spektrograma je i nemogućnost istovremenog postizanja dobre vremenske i frekvencijske rezolucije. U cilju prevazilaženja pomenutih nedostataka uveden je S-metod kao vremensko-frekvencijska distribucija oslobođena kros članova i time pogodna za reprezentaciju multikomponentnih signala. S-metod se definiše kao [9]:

$$SM(n, k) = SPEC(n, k) + 2 \operatorname{Re} \left[\sum_{i=1}^L P(i) STFT(n, k+i) STFT^*(n, k-i) \right] \quad (5)$$

gdje članovi u okviru sume poboljšavaju kvalitet spektrograma.

3. FILTRIRANJE MUZIČKIH SIGNALA

U nastavku će biti objašnjeni principi filtriranja muzičkih signala zasnovani na S-metodu.

A. KONCEPT VREMENSKI-PROMJENLJIVOG FILTRIRANJA

Podimo od pretpostavke da se zašumljeni muzički signal može zapisati kao: $x(t) = f(t) + n(t)$, gdje je $f(t)$ čist muzički signal a $n(t)$ označava šum. Izlaz vremenski-promjenljivog filtra dobija se na osnovu relacije [10]:

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} h\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}\right) w(\tau) x(t + \tau) d\tau \quad (6)$$

Ukoliko se definiše funkcija oslonca $L(t, \omega)$ kao:

$$L(t, \omega) = \int_{-\infty}^{\infty} h\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}\right) e^{-j\omega\tau} d\tau \quad (7)$$

tada se korišćenjem Parservalove teoreme signal na izlazu vremenski-promjenljivog filtra može zapisati na sljedeći način [7], [10]:

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} L(t, \omega) STFT(t, \omega) d\omega \quad (8)$$

Odnosno u diskretnom obliku važi da je[10]:

$$y(n) = \frac{1}{N} \sum_{k=-N/2}^{N/2} L(n,k) STFT(n,k) \quad (9)$$

Imajući u vidu S-metod kao distribuciju oslobođenu članova interferencije, funkcija oslonca L se definiše kao[7]:

$$L(n,k) = 1 - \frac{SM_n(n,k)}{SM_x(n,k)} \quad (10)$$

gdje je sa SM_n označen S-metod šuma, dok je sa SM_x označen S-metod zašumljenog signala.

B. FILTRIRANJE SIGNALA ZASNOVANO NA SOFT-TRESHOLDING TEHNICI

Pretpostavimo da šum $n(t)$ predstavlja Gauss-ov šum, nulte srednje vrijednosti i varijanse σ_n . Pokazuje se da se gotovo 99% vrijednosti ovakve šumne sekvence nalazi u granicama [11]:

$$-3\sigma \leq n \leq 3\sigma \quad (11)$$

Sa druge strane, imajući u vidu da se pri numeričkoj realizaciji spektograma, a time i S-metoda, koristi algoritam brze Fourier-ove transformacije, za neki proizvoljan signal $s(n)$ sa odgovarajućim FFT koeficijentima $S(k)$, na osnovu Parservalove teoreme važi da je:

$$\sum_{n=0}^{N-1} s(n) = \frac{1}{N} \sum_{k=0}^{N-1} S(k) \quad (12)$$

Na osnovu prethodne relacije može se zaključiti da je FFT \sqrt{N} puta jedinična transformacija [11]. Dakle, može se očekivati da 99% amplituda spektograma šuma a na osnovu (4) i 99% amplituda S-metoda, biće manje od $3\sigma\sqrt{N}$. Proglašavajući ovu vrijednost za prag ε , soft-thresholding funkcija se može definisati kao:

$$St(SM_x) = \begin{cases} SM_x - \varepsilon, & \text{for } x > \varepsilon \\ 0, & \text{for } x \leq \varepsilon \end{cases} \quad (13)$$

IV PRIMJERI I ILUSTRACIJE

Posmatrajmo primjer muzičkog signala odabranog sa frekvencijom 8kHz. Prilikom računanja STFT korišćen je prozor širine 256 odbiraka. Signalu je dodat Gaussov šum nulte srednje vrijednosti i varijanse 0.05. Najprije je testirana procedura vremenski-promjenljivog filtriranja. U praktičnoj realizaciji, relacija (9) kojom se definiše funkcija

oslonca L je donekle modifikovana korišćenjem spektralnog praga α :

$$L(n,k) = \max\left(1 - \frac{SM_n(n,k)}{SM_x(n,k)}, \alpha\right)$$

Vrijednost spektralnog praga postavljena je eksperimentalno na 0.04. Vremensko-frekvencijaska reprezentacija signala sa šumom i filtriranog signala prikazana je u Fig.1.a i Fig.1.b respektivno.

Prilikom testiranja procedure zasnovane na soft-thresholding tehnici, izračunata vrijednost praga je $\varepsilon = 10.3$. Pri proračunu STFT kao i SM korišćeni su iste vrijednosti parametara, kao u slučaju tehnike vremenski-promjenljivog filtriranja. Signalu je i u ovom slučaju dodat Gaussov šum istih parametara, kao u prvom slučaju. U Fig.2. ilustrovana je vremensko frekvencijska reprezentacija signala filtriranog korišćenjem soft-thresholding tehnike.

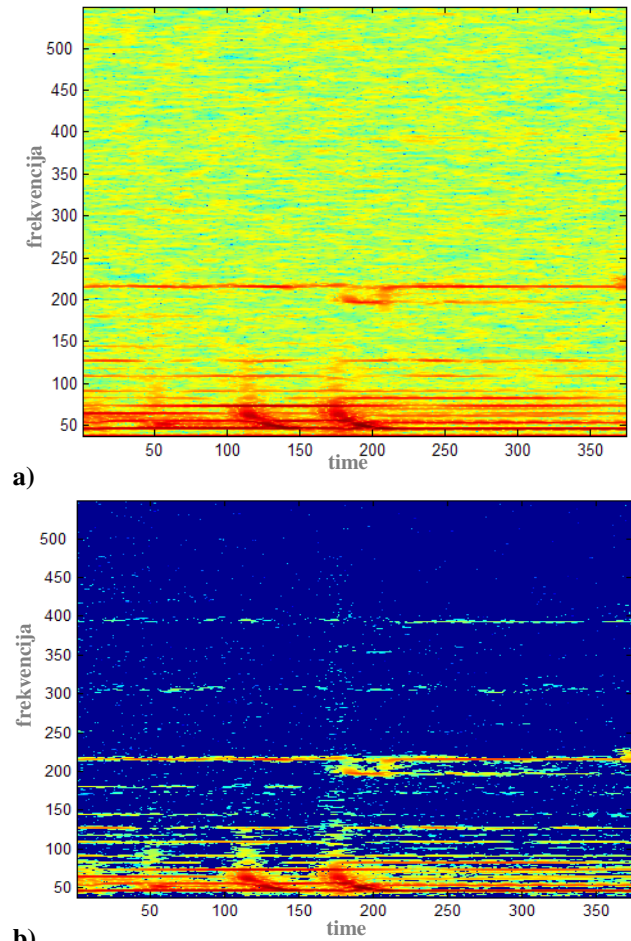


Fig.1. a) S-metod muzičkog signala sa šumom, b) S-metod filtriranog signala

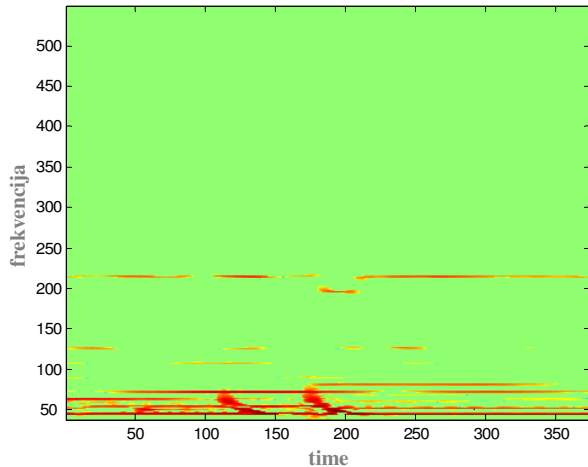


Fig.2. S-metod muzičkog signala filtriranog korišćenjem soft-thresholding tehnike

ZAKLJUČAK

U radu su predložene procedure redukcije šuma, zasnovane na novom konceptu filtriranja u vremensko-frekvencijskom domenu. Primjenom ovih procedura nije došlo do pojave tzv. muzičkog šuma, karakterističnog za metode koje se danas uobičajeno koriste u praksi, a koje su navedene u uvodu. Takođe, predloženi postupci pokazuju dobre rezultate u pogledu perceptualne neosjetljivosti. Kao nastavak budućih istraživanja veća pažnja biće posvećana izbor optimalnog praga za različite frekventne opsege muzičkog signala.

REFERENCE

- [1] Ballou G., ed (2005.), *Handbook for Sound Engineers*, Focal Press
- [2] Vaseghi S. (2000), *Advanced Digital Signal Processing and Noise Reduction*, Wiley
- [3] Zolzer U (1997), *Digital Audio Signal Processing*, Wiley
- [4] Godsill S. J & Rayner P. J. W. (1998), *Digital Audio Restoration*, Springer
- [5] Huang Y. & Benesty J. (2004), *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, Kluwer Academic Publishers
- [6] Kahrs M. & Brandenburg K. (2002), *Applications of Digital Signal Processing to Audio and Acoustics*, Kluwer Academic Publishers
- [7] S.Stanković, "About Time-Variant Filtering of Speech Signals with Time-Frequency Distributions for Hands-Free Telephone Systems", *Signal processing*, Vol.80, No.9, 2000.
- [8] Peebles P. Z. (1987), *Probability, Random Variables and Random Signal Principles*, McGraw-Hill
- [9] LJ.Stanković, "A method for Time-Frequency Signal Analysis", *IEEE Transaction on Signal Processing*, Vol.42, No.1, January 1994.
- [10] LJ.Stanković, "On the Time-Frequency Analysis Based Filtering", *Annales des telecommunications*, Vol.54, No.5/6, May/June, 2000.
- [11] G. W. Don, J. S. Walker, "Time-frequency analysis of music"