

Metodologija statističke analize muzičkih signala u oktavnim opsezima

Aleksandar M. Dikanović

Sadržaj — U ovom radu opisan je postupak analize statističkih karakteristika muzičkih signala u oktavnim opsezima (raspodele verovatnoće nivoa, srednji nivo, kumulativna funkcija). S posebnom pažnjom razmatrani su problemi normalizacije, vremenskog usrednjavanja signala, kao i izbora odgovarajućeg formata muzičkog signala. Rezultati analize nekoliko muzičkih signala prikazani su grafički i tabelarno. Analiza je sprovedena u programskom paketu Matlab®.

Ključne reči — muzički signal, oktavni opsezi, dinamika signala, raspodela nivoa.

I. UVOD

Audio signale karakterišu brze, vremenski zavisne promene svih njihovih parametara (amplitude, pitch-a, trajanja, harmonijskog sadržaja itd). Sve ove varijacije se ne mogu opisati analitičkim putem. Ipak u poslednje vreme razvijeni su brojni algoritmi za analizu audio signala koji su našli svoju inženjersku primenu. Većina ih je zasnovana na dva osnovna moda manipulacije audio signala (a) obrada u vremenskom domenu (b) obrada u frekvencijskom domenu. Na ovaj način dobija se uvid u dinamičke osobine audio signala [1]. Dinamičke osobine nekog audio signala mogu se najčešće predstaviti dvema numeričkim vrednostima. Prva predstavlja potreban headroom, dok drugi numerički podatak je dinamički opseg [2]. Podaci o širini dinamičkog opsega neke muzičke sekvence dobijaju se ispitivanjem raspodele verovatnoće nivoa signala. Ona ustvari predstavlja frekvenciju pojavljivanja različitih nivoa u analiziranoj muzičkoj sekvenci. Dinamičke osobine audio signala dominantno zavise od vrste zvukova koje predstavljaju. Potreba poznavanja dinamičkih karakteristika signala je od posebnog značaja kada su u pitanju muzički signali. Raspodela nivoa signala zavisi značajno od muzičkog žanra, ali u izvesnoj meri zavisi i od načina na koji je muzički signal zabeležen, odnosno od formata muzičkog signala.

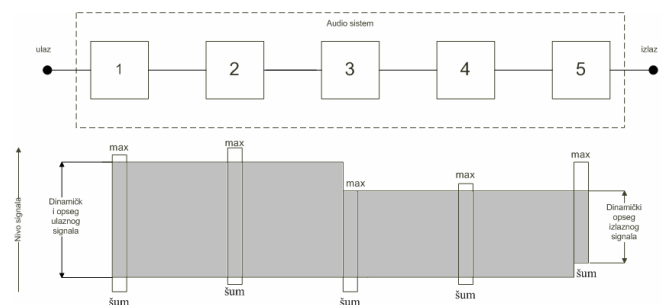
Rad je organizovan u pet poglavlja. Prenos audio signala kroz audio sistem i problem očuvanja dinamike obrađen je u narednom poglavlju. U trećem poglavlju objašnjen je postupak analize dinamičkih karakteristika i definisani su posmatrani parametri na primerima muzičkih signala filmske muzike. U četvrtom poglavlju analiziran je uticaj stepena

kompresije muzičkih signala na njegove dinamičke karakteristike. U petom poglavlju dat je zaključak.

II. PRENOS SIGNALA KROZ AUDIO SISTEM

Ono što se nameće kao praktičan problem proizilazi iz činjenice da su zahtevani dinamički gabariti signala bliski granicama tehničkih mogućnosti upotrebljenih uređaja [3]. U audio sistemu, kao i u svim njegovim elementima ponaosob, postoje dve tehničke granice koje određuju maksimalni mogući dinamički opseg signala. Sa njegove gornje strane to je najveća moguća amplituda signala koja može da prođe kroz sistem bez bitne degradacije, a sa donje strane to je najmanji održivi nivo signala koji se još uvek može prepoznati u šumu.

Problem očuvanja dinamike u audio sistemu može se simbolički ilustrovati na primeru jednostavne blok šeme koja je prikazana na Sl. 1[3]. Sistem je predstavljen nekim nizom elemenata koji su redno vezani. Ulazni signal ima neki dinamički opseg koji je na slici simbolično označen šrafiranom površinom. On je određen njegovom maksimalnom i minimalnom mogućom amplitudom koja se pojavljuje na ulazu u audio sistem. Istovremeno svaki od audio elemenata kroz koji prolazi signal ima svoj „dinamički prozor“. Prema tome svaki element u audio sistemu može biti mesto gde se degradira dinamički opseg signala.



Sl. 1. Šematski prikaz problema očuvanja dinamičkog opsega signala pri prolasku kroz audio sistem

Da bi se to izbeglo neophodno je zadovoljiti dva uslova. Prvo, da dinamičke mogućnosti svih pojedinih elemenata povezanih u sistem budu veće od dinamičkog opsega ulaznog signala. Drugo, da ispravnom manipulacijom sa signalom, njegovim pojačanjima ili slabljenjima, njegov dinamički opseg duž čitavog sistema nalazi između dinamičkih granica elemenata kroz koji prolazi.

Aleksandar Dikanović, Elektrotehnički fakultet u Beogradu, Srbija; (telefon: +381-64-3229775; e-mail: dikanovic@beotel.net).

III. METODOLOGIJA ANALIZE

Analiza signala koja se sprovodi podrazumeva određivanje sledećih statističkih parametara:

1. Raspodela nivoa verovatnoće signala u sedam oktavnih opsega sa centralnim učestanostima 125Hz, 250Hz, 500Hz, 1000Hz, 2000Hz, 4000Hz, 8000Hz respektivno.
2. Srednji nivo širokopojasnog signala, kao i srednji nivo signala u oktavnim opsezima.
3. Vrednosti signala premašenih u 1%, 10%, 99% vremena (L_{-1} , L_{10} , L_{99}) respektivno. Kako u dinamičkom opsegu koji se posmatra signali zauzimaju manji ili veći deo opsega, L_{99} i L_{-1} mogu se posmatrati kao indikatori širine tog opsega. Upravo rastojanje između L_{-1} i L_{99} može se označiti kao dinamički opseg muzičkog signala.

Na početku analize ulazni signal se normalizuje svojom maksimalnom apsolutnom vrednošću. Signal se zatim filtrira korišćenjem banke filtera [4]. Prilikom formiranja banke filtera treba voditi računa da filteri imaju što veću selektivnost i budu što manjeg reda kako bi se smanjio broj računskih operacija.

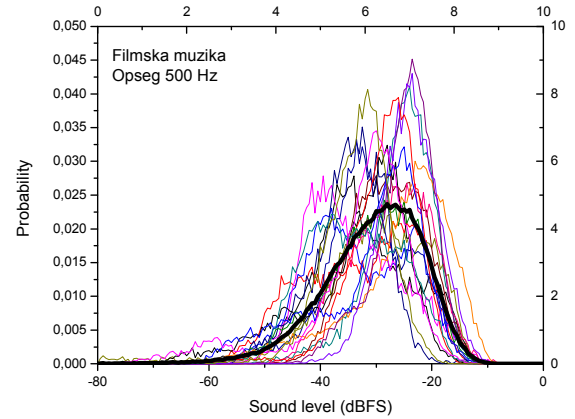
Nakon filtriranja vrši se usrednjavanje signala korišćenjem prozorske funkcije trajanja T , bez preklapanja. Kako konvolucija u vremenskom domenu odgovara množenju Furijeovih transformacija odgovarajućih funkcija u frekvencijskom domenu[5] u slučaju pravouganog prozora imali bismo množenje sa funkcijom $|\sin x/x|$, pri čemu su nule na celobrojnim umnošcima $1/T$. Otuda proces usrednjavanja signala u frekvencijskom domenu predstavlja njegovo filtriranje niskopropusnim filtrom:

$$f_{gr} = \frac{1}{2T} \quad (1)$$

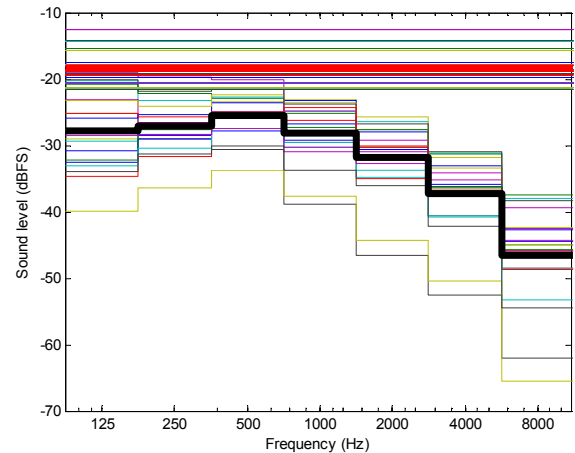
Iz prethodnog izraza može se zaključiti da je izbor vremenske konstante T pri merenju efektivne vrednosti zvučnog pritiska u praksi relativno složen. Zbog nestacionarnosti zvučnih signala u audio sistemima koriste se relativno male vrednosti za T [6]. U praksi se često koriste i prozori sa preklapanjem u cilju smanjenja gubitaka po ivicama prozora.

Verovatnoća raspodele definiše se na opsegu od -80dBFS (vrednost koja je definisana kao prag čujnosti širokopojasnog signala[2]) do 0dBFS. Raspodela se računa sa korakom od 0,5 dB. Iz svake raspodele moguće je odrediti kumulativna funkcija. Ipak u cilju dobijanja što preciznijih rezultata potrebno je izvršiti interpolaciju krive čime se dobija rezolucija od 0,1dB. Tako dobijena funkcija koristi se za dobijanje vrednosti nivoa premašenih u 1%, 10%, 99% vremena (L_{-1} , L_{10} , L_{99}) respektivno.

Raspodela verovatnoće nivoa 21 numere koje pripadaju filmskoj muzici dati su na Sl.2, dok je na Sl.3 data raspodela nivoa signala po oktavama (srednji nivo označen crnom bojom), kao i srednji nivo signala na čitavom opsegu za svaku numeru (srednji nivo označen crvenom bojom). Vremensko usrednjavanje vršeno je pravougaonom prozorskom funkcijom trajanja 10ms.

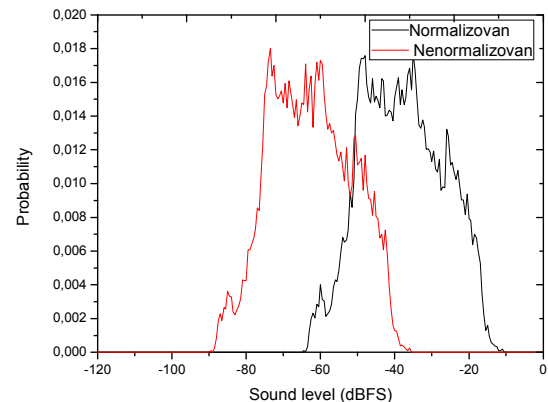


Sl. 2. Raspodela verovatnoće nivoa signala iz žanra filmske muzike u opsegu sa centralnom učestanošću 500Hz



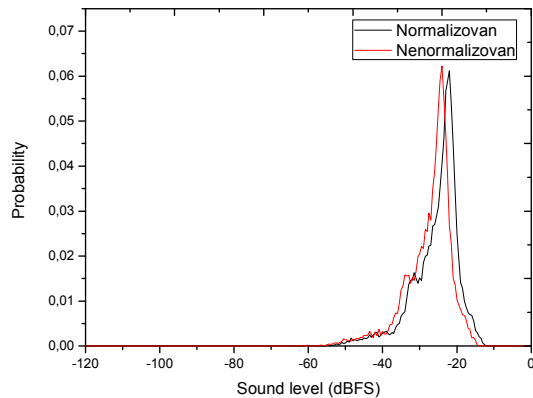
Sl. 3. Raspodela verovatnoće nivoa signala po oktavama i srednji nivoi signala na čitavom opsegu

Analza sprovedena na muzičkim signalima pokazala je da je za neke signale potrebna analiza u znatno širem opsegu od 80dB, kako bi bilo moguće odrediti vrednost širine



Sl. 4. Raspodela verovatnoća signala iz žanra klasične muzike normalizovanog i nenormalizovanog u opsega sa centralnom učestanošću 8000Hz

dinamičkog opsega. Informaciju o širini dinamičkog opsega moguće je zadržati normalizacijom signala nakon filtriranja svojom maksimalnom apsolutnom vrednošću u datom oktavnom opsegu. U ovom trenutku osvrnućemo se na pitanje normalizacije signala nakon njihovog filtriranja. S tim u vezi sprovedena je analiza kojom su obuhvaćeni po jedno delo iz klasične i pop muzike. Obrada svakog signala vršena je sa i bez normalizacije nakon filtriranja. Analiza je sprovedena na opsegu širine 120dB. Raspodela verovatnoća nivoa audio signala u najvišem oktavnom opsegu dati su na Sl. 4. i Sl. 5.:



Sl. 5. Raspodela verovatnoća signala iz žanra pop muzike normalizovanog i nenormalizovanog u opseg sa centralnom učestanošću 8000Hz

Kao što se može sa slika videti u slučaju klasične muzike raspodela verovatnoća se znatno pomerila ka nižim nivoima, i ispod 80dB. To se može objasniti činjenicom da kod klasične muzike zvuk ne prolazi nikakvu predobradu, a kako je snaga zvuka koju proizvode muzički instrumenti uglavnom koncentrisana u središnjem opsegu (250-4000Hz) to na visokim učestanostima uglavnom preovlađuju tiši tonovi[7]. Iz tog razloga postavlja se potreba za proširenjem opsega posmatranja raspodele. Za razliku od klasične muzike pop muzika predstavlja muziku novog doba koja se uglavnom pušta po klubovima. Iz tog razloga diskografske kuće i izvođači stvaraju muziku koja bi se svojom glasnoćom izdigla iznad buke koju stvaraju ljudi međusobnom konverzijom[8]. Ta izuzetna glasnoća se postiže kompresijom dinamičkog opsega signala po opsezima, počev od veoma niskih koji daju zvuku bas, pa do visokih koji zvuku daju potrebnu oštrinu.

Svi prethodni grafici predstavljaju rezultat obrade signala zapisanih u mp3 formatu koji su za potrebe obrade u softverskom paketu Matlab® konvertovani u wav format. U cilju što boljeg sagledavanja rezultata u narednom poglavlju biće ispitan uticaj kompresije muzičkog signala na njene dinamičke osobine.

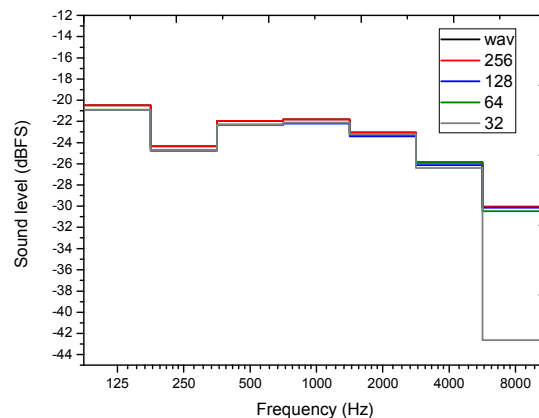
IV. ANALIZA RAZLIČITIH FORMATA

Paralelno sa uvodjenjem digitalnog načina zapisivanja podataka, počelo se sa uvodjenjem formata njihovog čuvanja na kompjuteru. Jedan od formata koji pruža najbolji kvalitet

jeste wav. Wav predstavlja standardni format za zapisivanje nekomprimovanog audio signala na CD-u. Signali čuvani na ovaj način imaju učestonost odabiranja od 44,1 KHz i kodiraju se sa 16 bita. Međutim veliki nedostatak ovakvog načina zapisivanja audio signala predstavlja velika glomaznost podataka. Na primer, veličina muzičkog signala, učestonosti odabiranja od 44,1 kHz sa 16 bitnom konverzijom, prosečnog trajanja iznosi 30-ak MB. Kako bi se ovaj nedostatak ukolnio javila se potreba za kompresijom audio signala.

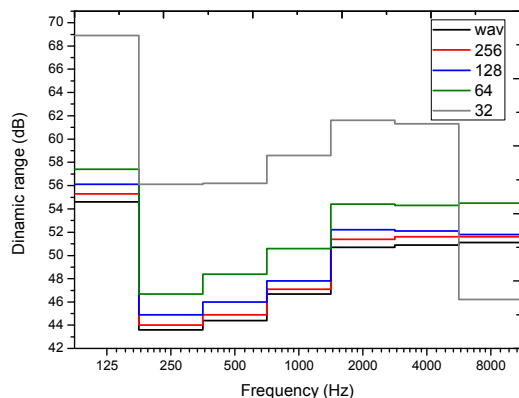
Kompresija audio signala razvijena je nešto pre formiranja MPEG-a [9]. Jedan od prvih načina kompresije zvao se MUSICAM (Masking pattern adapted Universal Sub-band Integrated Coding and Multiplexing). Paralelno sa MUSICAM-om pristupilo se razvoju još jednog sistema pod nazivom ASPEC (Adaptive Spectral Perceptual Entropy Coding). Nešto kasnije MPEG je kombinacijom ova dva sistema stvorio standard sa tri nivoa (layers), poznatija kao layer I, II i III. Ova tri nivo razlikuju se u kompleksnosti kodiranja kao i u stepenu smanjenja bitskog protoka. Ipak zajedničko za sve MPEG nivoe jeste da se kompresija signala postiže primenom (1) maskiranja i (2) skaliranja i kvantizacije. Maskiranje predstavlja postupak eliminacije tonova koje naše uvo ne bi moglo da registruje usled efekta frekvencijskog i vremenskog maskiranja, kao i usled činjenice da ljudko uvo nema podjednaku osetljivost na svim frekvencijama. Kvantizacijom i skaliranjem određuje se minimalan broj bita neophodan za kodiranje, koji je dovoljan da nivo šuma bude ispod nivoa maskiranja. Najpoznatiji među njima MPEG Layer III, ili mp3 format, predstavlja najčešći način zapisivanja muzičkih file-ova, i pruža visok stepen kompresije.

Treba imati u vidu da svaka digitalana kompresija signala može dovesti do izvesnih promena u dinamičkim osobinama signala. Te promene mogu se ogledati u smanjenju nivoa signala, ili sužavanju dinamičkog opsega. Neka nezvanična istraživanja pokazala su da su ova izobličenja primetna u slučaju promena dinamičkog opsega većih od $\pm 2\text{dB}$ [8].



Sl. 6. Raspodela nivoa signala po oktavama za različite stepene kompresije

Kako bi se utvrdio uticaj kompresije signala na dinamiku i srednji nivo sprovedena je analiza. Eksperimentom je bila obuhvaćena jedna pesma, zapisana u wav i mp3 formatu sa različitim stepenim kompresije (različitim bitskim protokom). Raspodela nivoa signala data je na Sl.6. Analogno, na Sl.7. prikazana je zavisnost širine dinamičkog opsega od stepena kompresije.



Sl. 7. Raspodela dinamičkog opsega signala po oktavama za različite stepene kompresije

Sa slika se može uočiti da sa povećanjem stepena kompresije povećava i stepen odsutpanja dinamičkih osobina komprimovanog u odnosu na nekomprimovan signal. Kako bi se preciznije uvideo utcaj kompresije na dinamičke osobine muzičkog signala tabelama 1 i 2 data je širina dinamičkog opsega signala za različite stepene kompresije. Vrednosti su date u decibelima.

TABELA 1 : DINAMIČKI OPSEG SIGNALA ZA RAZLIČITE NIVOE KOMPRESIJE U I, II, III I IV OKTAVNOM OPSEGU

	I opseg (125Hz)	II opseg (250Hz)	III opseg (500Hz)	IV opseg (1000Hz)
wav	54,6	43,6	44,4	46,7
256kbps	55,3	44	44,9	47,1
128kbps	56,1	44,9	46	47,8
64kbps	57,4	46,7	48,4	50,6
32kbps	68,9	56,1	56,2	58,6

TABELA 2 : DINAMIČKI OPSEG SIGNALA ZA RAZLIČITE NIVOE KOMPRESIJE U V, VI I VII OKTAVNOM OPSEGU

	V opseg (2000Hz)	VI opseg (4000Hz)	VII opseg (8000Hz)
wav	50,7	50,9	51,1
256kbps	51,4	51,6	51,6
128kbps	52,2	52,1	51,8
64kbps	54,4	54,3	54,5
32kbps	61,6	61,3	46,2

Sa tabela 1 i 2 može se zaključiti da najveće odstupanje ima signal sa najmanjim protokom (14,3dB), dok najveća odstupanja signala sa protocima 256kbps i 128kbps iznose 0,7dB i 1,6dB respektivno.

V. ZAKLJUČAK

U ovom radu grafički su predstavljeni rezultati analize više muzičkih signala koji pripadaju žanru filmske muzike. Rezultati analize signala sa i bez normalizacije u oktavnim opsezima pokazali su da je za dobijanje preciznijih rezultata potrebno proširenje opsega posmatranja raspodele. Posebno su analizirane razlike u frekvencijskoj zavisnosti dinamičkih parametara kod različitih kompresija signala. Analizom uticaja kompresije dobijeno je da se signali protoka već od 128kbps, po svojim karakteristikama bitnije ne razlikuju od nekomprimovanog signala i mogu biti upotrebljeni u analizi. Nakon definisanja metodologije analize dinamičkih karakteristika dalji koraci u istraživanju podrazumevaju analizu muzičkih signala iz različitih muzičkih žanrova.

LITERATURA

- [1] M.Paraskevas, J.Mourjopoulos, *A statistical study of the variability and features of audio signals: some preliminary results*, 100th AES Convention (1996), Preprint No. 4256
- [2] Miomir Mijić, Draško Mašović, Milan Petrović, Dragana Šumarac-Pavlović, *Statistical Properties of Music Signals*, AES 126th Convention, Munich, Germany, 2009
- [3] Dr Miomir Mijić, tema *Dinamičke osobine audio signala*, predavanja iz Audio sistema, Beograd, 2009
- [4] IEC 1260 Octave filter bank
- [5] Dr Miomir Mijić, tema *Zvučni pritisak kao signal*, predavanja iz Elektroakustike, Beograd, 2007
- [6] Sanjit K. Mitra, *Digital signal processing: A computer based approach*, Mc Graw Hill, 2006
- [7] F. Alton Everest, *The master handbook of acoustics*, Mc Graw Hill, 2001
- [8] N.Thiele, *Some thoughts on the dynamics of reproduced sound*, Journal AES, VOL. 53, No. 1/2, 2005, pp 130-132
- [9] K.F. Ibrahim, *Newnes Guide to Television and Video Technology*, 2007

ABSTRACT

In this paper, the methodology of analysis of frequency dependence of dynamical characteristics of musical signals is shown. To approach such an analysis, it was necessary to readjust the method to the level changes that resulted from filtering of signals, also to define the principles of normalization of statistical distribution and, moreover, to determine the ways of comparing the different octave band results. As an illustration the results of analysis of several statistical parameters of different musical genres are shown. The difference between dynamical characteristics of compressed signals, written with different levels of compression, is analyzed separately.

THE METHODOLOGY OF OCTAVE BAND STATISTICAL ANALYSIS OF MUSICAL SIGNALS

Aleksandar M. Dikanović