

# Realizacija dva širokopojasna modela za izračunavanje glasnosti

Sanja Rakić i Ivana Čpajak

**Sadržaj** — Predikcija i kontrola glasnosti su bitne u televizijskom i radio emitovanju, naročito od kada je među emiterima došlo do takozvanog „rata glasnosti“. U ovom radu je dat pregled modela koji se koriste za estimaciju glasnosti. Na osnovu postojećih algoritama realizovana su dva širokopojasna modela za izračunavanje glasnosti nestacionarnih audio signala čiji je opis prikazan u radu.

**Ključne reči** — glasnost, modeli za izračunavanje glasnosti, uskopojasni i širokopojasni modeli.

## I. UVOD

Percipirana glasnost je funkcija svih anatomskih i fizioloških osobina čula sluha kao i psiholoških faktora koji deluju u procesu stvaranja zvučne slike. Problem pri kvantifikovanju glasnosti predstavlja činjenica da ona zavisi od brojnih faktora, kao što su dinamičke osobine zvučnog stimulusa, njegov frekvencijski sadržaj, trajanje i vremenska anvelopa [1]. Glasnost je subjektivni doživljaj, tako da neće svi slušaoci doživeti isti stimulus jednako glasnim, mogu čak postojati razlike i u percepciji glasnosti kod istog slušaoca u zavisnosti od brojnih psiholoških faktora. Kvantifikacija zvučne pobude koja bi bila u korelaciji sa subjektivnom glasnošću je složen problem koji treba da objedini brojne zahteve.

Priroda i sadržaj programskog materijala koji se emituje putem radija ili televizije se neprekidno menjaju, a ovakve promene rezultuju značajnim promenama glasnosti, koje slušaocu mogu biti neprijatne [2]. Na percipiranu glasnost utiču i različite forme dinamičkog procesiranja koje se uobičajeno primenjuju na signale. Da ne bi bilo naglih skokova, tj. da bi materijal koje se emituje bio ujednačen po glasnosti, neophodno je merenje i kontrola izlaznog nivoa glasnosti [3]. Pitanje ovakve kontrole je od velikog značaja i u muzičkoj industriji gde se često primenjuje kompresija u cilju maksimiziranja percipirane glasnosti.

Merači glasnosti čiji se rad bazira na praćenju vršnog nivoa signala pokazali su se neefikasnim, jer je glasnost u većoj meri korelisana sa ukupnom energijom koju signal nosi nego sa vršnom vrednošću [1]. Zato su počeli da se razvijaju modeli sa drugačijim pristupima. Neki od postojećih načina za izračunavanje glasnosti su ukratko opisani u ovom radu, i na osnovu postojećih algoritma u Matlab-u su realizovana dva modela za predikciju subjektivne glasnosti.

Modeli za izračunavanje glasnosti se mogu značajno pojednostaviti ako se razmatraju samo neke vrste ulaznih signala ali takvi modeli će se pokazati kao loši ako se koriste van svojih okvira. Prvi modeli su dizajnirani za rad isključivo sa stacionarnim signalima dok noviji prediktuju glasnost i vremenski promenljivih signala [3]. Takođe, od modela se zahteva da ima mogućnost operisanja sa mono, stereo i višekanalnim formatima audio signala. Cilj je da izlazna vrednost koju daje model bude što bolja kvantifikacija onoga što slušalac čuje, da bi se te vrednosti mogle koristiti pri kontroli nivoa glasnosti.

## II. MODELI ZA IZRAČUNAVANJE GLASNOSTI

Modeli za izračunavanje glasnosti se dele na uskopojasne i širokopojasne [4]. Širokopojasni modeli, za razliku od uskopojasnih, ne vrše frekvencijsku analizu signala. Kako podela po frekvencijskim opsezima ne postoji nemoguće je modelovati efekte spektralnog sumiranja glasnosti, pa je upotreba ovih modela ograničena na širokopojasne signale čiji spektralni sadržaj ne varira puno i čiji je opseg nivoa relativno uzak. U okviru širokopojasne metode postoje dva principa: računa se ili  $L_{eq}$  (ekvivalentni nivo zvuka) ili PPM (Peak Programme Meter) nivo [4]. U energetskom domenu  $L_{eq}$  odgovara usrednjavanju nivoa signala čija se glasnost procenjuje u nekom vremenskom intervalu. Pre usrednjavanja signal se obično propusti kroz filter čiji frekvencijski odziv simulira jednu od krivih jednakih glasnosti (obično A, B, C, D, M ili RLB). Smatra se da zajedno sa odgovarajućim frekvencijskim skaliranjem  $L_{eq}$  merenje predstavlja način za izračunavanje glasnosti. Za razliku od  $L_{eq}$  merenja koje računa jednu prosečnu vrednost PPM merenje daje signal u vremenskom domenu – anvelopu u dB; anvelopa se može detektovati korišćenjem filtra propusnika niskih učestanosti i opisuje se raspodelom decibelskih vrednosti. Jedan od načina da se na osnovu takve raspodele proceni glasnost audio signala jeste da se pronađe vrednost koja je premašena u određenom procentu vremena i da se ona smatra dugovremenom glasnošću audio signala. TC Electronic je razvio model nazvan LARM koji kombinuje dve poznate merne tehnike: kombinuje se kvazi-vršni detektor sa frekvencijski skaliranim RMS merenjem, i predstavlja značajno poboljšanje u odnosu na postojeće merne tehnike.

Uskopojasne metode se baziraju na podeli spektra signala po frekvencijskim opsezima. Svaki frekvencijski opseg se dalje procesira nezavisno [4]. Na ovaj način se modeluju efekti frekvencijskog maskiranja i spektralnog sumiranja glasnosti. Neki modeli za izdvajanje

Sanja Rakić, student Elektrotehničkog fakulteta u Beogradu, Srbija (telefon: 381-64-1998759, e-mail: [sanja.rakic@yahoo.com](mailto:sanja.rakic@yahoo.com))

Ivana Čpajak, student Elektrotehničkog fakulteta u Beogradu, Srbija (telefon: 381-64-0217200, e-mail: [ivanacpajak@yahoo.com](mailto:ivanacpajak@yahoo.com))

frekvencijskog sadržaja koriste filtre koji se poklapaju sa kritičnim opsezima dok drugi radi smanjenja kompleksnosti koriste manje filtera koji obuhvataju šire opsege. Jedan od najrasprostranjenijih modela je Cvikerov metod, ali on je prilično neefikasan za predikciju glasnosti vremenski promenljivih signala [4]. TC Electronic je razvio novi model baziran na uskopojasnom pristupu nazvan HEIMDAL koji se pokazao najbolji pri evaluaciji u odnosu na subjektivne baze podataka [4]. Prednost modela je postojanje mogućnosti da se određeni parametri menjaju da bi model dao što tačniju estimaciju glasnosti za određenu vrstu zvučnog stimulusa. Vrednosti parametara se određuju za vreme optimizacione faze i ne moraju se menjati tokom upotrebe modela.

U [4] je opisan eksperimentalni metod za evaluaciju performansi različitih modela, kao i način za formiranje dva referentna seta podataka koji će biti korišćeni pri evaluaciji. Merenjem performansi modela i poređenjem došlo se do zaključka da bi se modeli mogli svrstati u četiri klase. U klasu 1 se nalaze modeli koji su pokazali najbolje ukupne performanse, dok su u klasi 4 najlošiji modeli. Noviji modeli napravljeni specijalno za predikciju nivoa glasnosti programskog materijala, LARM i HEIMDAL, pokazali su se najboljim i kao takvi se nalaze u klasi 1. Dva široko rasprostranjena modela, Cvikerov i  $L_{eq}(A)$ , su neočekivano svrstani u klase 3 i 4, respektivno. Cvikerov model je prvobitno razvijen za sintetičke i stacionarne signale i zato ne može da obezbedi tačnu predikciju glasnosti za govor i muziku. Najlošije performanse u odnosu na ova dva seta podataka pokazali su  $L_{eq}(A)$  i  $L_{eq}(M)$  modeli, koji su implementirani u analizatorima signala i meračima nivoa zvuka u bioskopima [4].

### III. REALIZOVANI MODELI

U okviru ovoga rada realizovana su dva modela za izračunavanje glasnosti zvučnih segmenata: jedan prema algoritmu iz ITU-R BS. 1770 preporuke a drugi prema LARM algoritmu (Model I i Model II respektivno).

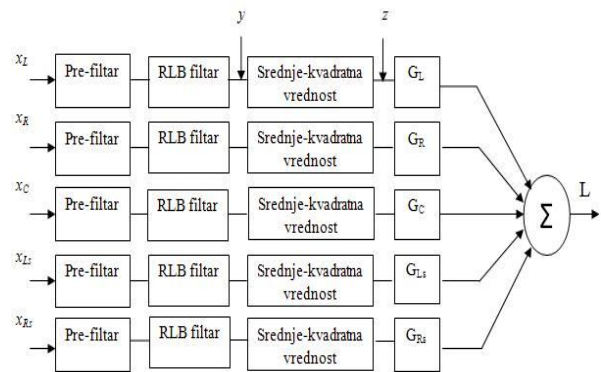
#### A. Model 1

U preporuci ITU-R BS. 1770 je predstavljen algoritam za merenje glasnosti višekanalnog audia [4]. Na Sl.1 je prikazana blok-šema modela. Dijagram prikazuje ulaze za 5 glavnih kanala (levi, desni, centrali, levi *surround* ( $L_s$ ) i desni *surround* ( $R_s$ )) što omogućava monitoring programa koji sadrže od jednog do pet kanala. Za programe koji imaju manje od pet kanala neki ulazi se neće koristiti. Kanal sa nisko-frekvencijskim efektima (LFE) nije uključen u merenje.

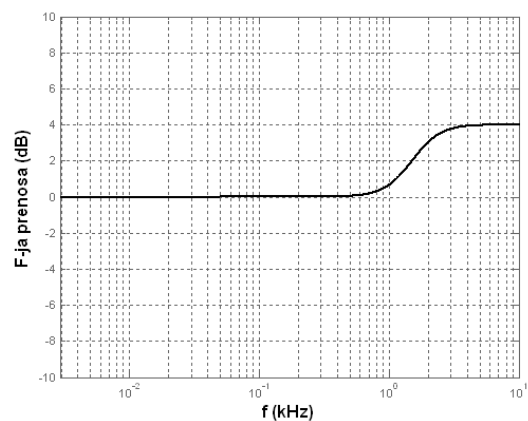
U prvoj fazi algoritma vrši se pre-filtriranje signala; signal se propušta kroz filter čija je funkcija prenosa prikazana na Sl. 2. Pre-filtriranje uračunava uticaj glave na percepciju signala; glava je modelovana kao gruba sfera. Pre-filtar je definisan kao IIR filter drugog reda sa funkcijom prenosa

$$H_1(z) = \frac{b_{01} + b_{11} \cdot z^{-1} + b_{21} \cdot z^{-2}}{1 + a_{11} \cdot z^{-1} + a_{21} \cdot z^{-2}} \quad (1)$$

Koeficijenti filtra su dobijeni iz preporuke [5] i prikazani su u Tabeli 1.



Sl 1. Blok šema algoritma po kojem je realizovan model I.



Sl. 2 Funkcija prenosa pre-filtra

TABELA 1: KOEFICIJENTI PRE-FILTRA.

		$b_{01}$	1.53512485958697
$a_{11}$	-1.69065929318241	$b_{11}$	-2.69169618940638
$a_{21}$	0.73248077421585	$b_{21}$	1.19839281085285

U drugoj fazi algoritma se primenjuje RLB težinska kriva, koja predstavlja jednostavan filter propusnik visokih frekvencija čija je funkcija prenosa prikazana na Sl 3. Koeficijenti za oba filtera su definisani za frekvenciju odabiranja od 48 kHz. Implementacije za druge frekvencije odabiranja zahtevaju druge vrednosti koeficijenata, koje treba izabrati tako da filter ima istu funkciju prenosa odziv kao za 48 kHz [5].

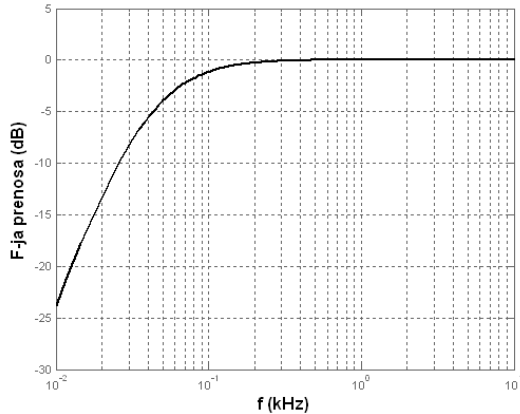
RLB filter je specificiran kao filter drugog reda, čiji je frekvencijski odziv prikazan na Sl.3. Funkcija prenosa ovog filtera je definisana kao

$$H_2(z) = \frac{b_{02} + b_{12} \cdot z^{-1} + b_{22} \cdot z^{-2}}{1 + a_{12} \cdot z^{-1} + a_{22} \cdot z^{-2}} \quad (2)$$

Koeficijenti filtra koji su dobijeni na osnovu preporuke [5] prikazani su u Tabeli 2.

TABELA 2: KOEFICIJENTI RLB FILTRA.

		$b_{02}$	1
$a_{12}$	-1.99004745483398	$b_{12}$	-2
$a_{22}$	0.99007225036621	$b_{22}$	1



Sl. 3 Funkcija prenosa RLB filtra

Nakon primene ova dva filtriranja vrši se izračunavanje srednje-kvadratne energije signala prema formuli:

$$z_i = \frac{1}{T} \sum_0^T y_i^2 \quad (3)$$

U prethodnoj formuli T je interval vremena u kom se vrši sumiranje, a  $y_i$  je signal na izlazu iz RLB filtra ( $i = L, D, C, Ls, Rs$  je oznaka odgovarajućeg kanala). Postoji mogućnost izbora vrednosti T, tako da se ovim algoritmom može dobiti ukupna glasnost određenog programskog materijala, a mogu se i pratiti kratkovremene promene glasnosti. Kada se skalirana srednje-kvadratna energija,  $z_i$ , izračuna za svaki kanal, poslednji korak u računanju nivoa glasnosti, L, je sumiranje nivoa iz svih kanala prema formuli:

$$L = -0.8979 + 10 \log_{10} \sum_i^N G_i \quad (4)$$

U jednačini (4), N predstavlja broj kanala na ulazu, a  $G_i$  su težinski koeficijenti za različite kanale, koji su dati u Tabeli 3, i koji imaju funkciju da uračunaju uticaj pravca nailaska zvuka na auditorni sistem [5].

Jedinica kojom se izražava glasnost proračunata na osnovu modela I je LUFS (Loudness Unit relative to nominal Full Scale) [5]. Povećanje nivoa signala na ulazu za 1 dB rezultuje povećanjem glasnosti za 1 LUFS. Ako na ulaz levog, centralnog ili desnog kanala dovedemo sinusoidu na frekvenciji od 1 kHz sa nivoom amplitude od 0 dBFS, indikovani nivo glasnosti će biti -3 LUFS. Da bi ovo bilo zadovoljeno uvedena je kalibraciona konstanta -0.8979.

Kako je ovaj model pravljen prema [5] on je pogodan za procenu glasnosti audio programa koji su tipični za emitovanje, ali generalno nije dobar za estimaciju subjektivne glasnosti čistih tonova. Prednost ovog višekanalnog merača glasnosti je njegova jednostavnost. Ceo algoritam je napravljen od jednostavnih blokova za procesiranje signala koji se lako mogu implementirati. Druga ključna prednost je skalabilnost algoritma. Pošto se identično pocesiranje primenjuje na svaki kanal, može se raditi sa bilo kojim brojem kanala, od 1 do N. Šta više, pošto se doprinosi pojedinačnih kanala sumiraju kao vrednosti glasnosti, a ne kao nivoi signala, algoritam je nezavisan od faza ili korelacije između kanala.

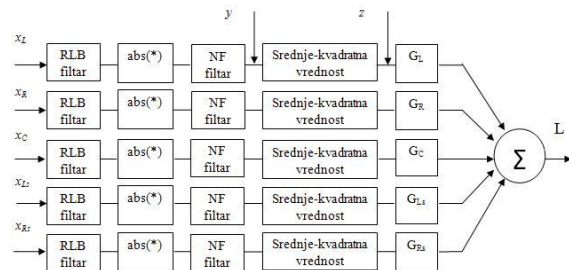
TABELA 3: TEŽINSKI KOEFICIJENTI ZA AUDIO KANALE.

kanal	$G_i$
Levi ( $G_L$ )	1.0 (0 dB)
Desni ( $G_D$ )	1.0 (0 dB)
Centralni ( $G_C$ )	1.0 (0 dB)
Levi surround ( $G_{LS}$ )	1.41 (~1.5 dB)
Desni surround ( $G_{DS}$ )	1.41 (~1.5 dB)

Pošto je merenje glasnosti po ITU preporuci merenje sa integraljenjem, jako tihi segmenti utiču na smanjenje izmerene vrednosti. Da bi se ovo izbeglo, trebalo bi da se integracija pauzira tokom tiših segmenata, tako da pronalaženje načina za isključivanje izuzetno tihih segmenata iz proračuna predstavlja prostor za unapređenje modela. Druga mana ovog načina računanja glasnosti je što ignoriše uticaj LFE kanala [6]. Bilo je pokušaja da se algoritam proširi tako da obuhvati LFE kanal sumiranjem K-skalirane snage LFE kanala. Ovakvo modifikovan algoritam se ne slaže sa subjektivnim procenama osim ako se primeni slabljenje od 10 dB za LFE kanal pre sumiranja njegove snage sa ostalim kanalima [6]. Leq(R2LB) merenje, na kojem se zasniva algoritam iz Modela I, ne estimira glasnost na optimalan način kada programski materijal sadrži znatnu energiju na niskim frekvencijama. Filmska i popularna muzika imaju značajne komponente na niskim frekvencijama, tako da Leq(R2LB) nije optimalan izbor za ove tipove signala [7].

### B. Model II

Model II je napravljen na osnovu algoritma koji se zove LARM i koji je razvijen u TC Electronic-u [4]. Dijagram po kom je realizovan ovaj model prikazan je na Sl. 4. Algoritam je razvijen sa ciljem da pruži tačnu i robusnu estimaciju glasnosti zvučnih segmenata koji sadrže govor i/ili muziku. Predviđen je za estimaciju dugovremene glasnosti, ali se može prilagoditi i računanju kratkovremene [8]. Za detekciju vrhova se koristi asimetrični filter propusnik niskih frekvencija koji ima zadatak da izdiže više nivoe anvelope. Glasnost se dalje estimira analogno Leq modelu sa frekvencijskim skaliranjem, ali se umesto klasičnog računanja RMS vrednosti izračunava neka vrsta srednje snage (*power mean*),  $z_i$ .



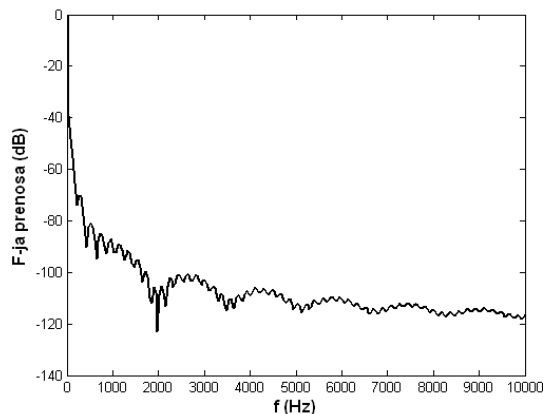
Sl.4 Blok šema algoritma po kojem je realizovan model II

Posle prve faze koja je ekvivalentna drugoj fazi modela I (Sl. 3), izračunava se apsolutna vrednost signala. U trećoj fazi primenjuje se asimetrični filtar propusnik niskih frekvencija koji ima brz linearan uspon i sporo ekponencijalno opadanje, čija je funkcija prenosa prikazana na Sl. 5. Pri realizaciji je omogućeno da se promenom parametara menja izgled filtra, tj. izgled uzlazne i silazne ivice filtra.

U četvrtoj fazi vrši se izračunavanje srednje snage prema sledećoj jednačini:

$$z_i = \sqrt[p]{\frac{1}{T} \sum_{i=0}^T |y(i)|^p} \quad (5)$$

U jednačini (5)  $p$  je parametar modela,  $T$  je veličina prozora za usrednjavanje (podložna je promeni, tako da se pomoću modela mogu pratiti i kratkovremena i dugovremena glasnost), a  $y_i$  je ulazni signal u tački procesiranja kao što je naznačeno na Sl. 4 ( $i = L, R, C, Ls, Rs$  je oznaka odgovarajućeg kanala). Izračunavanje vrednosti  $z_i$ , za  $p > 1$ , predstavlja nelinearno skaliranje vrednosti anvelope koja je dobijena na izlazu iz asimetričnog filtra propusnika niskih učestanosti. Važno je primetiti da je za  $p=2$  vrednost  $z_i$  ekvivalentna RMS vrednosti, dok je za  $p=1$   $z_i$  zapravo srednja apsolutna vrednost. Za estimaciju glasnosti optimalna vrednost parametra  $p$  je oko 1.5 [4].



Slika 5. Funkcija prenosa asimetričnog filtra

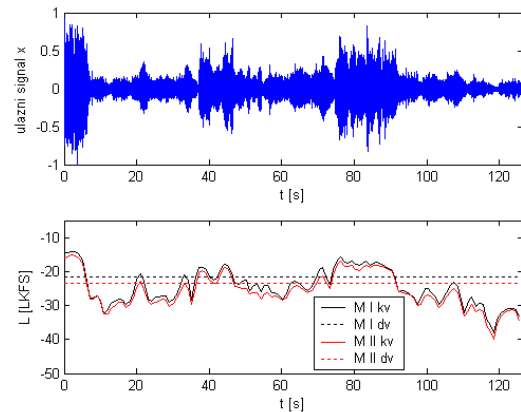
Nakon izračunavanja srednje snage,  $z_i$ , za svaki kanal, na kraju se računa nivo glasnosti na sledeći način:

$$L = -0.9062 + 20 \log_{10} \sum_i^N G_i \cdot z_i \quad (6)$$

Ukoliko računamo glasnost za različite formate audia, ugao nailaska signala i u ovom modelu uzimamo u obzir preko težinskih koeficijenata  $G_i$ , koji su dati u Tabeli 3. Kalibraciona konstanta je +0.9062, a jedinica kojom se izražava glasnost je takođe LUFS.

U okviru ovog rada u programskom paketu MATLAB realizovana su dva širokopolasna modela za estimaciju glasnosti. Ovi modeli omogućavaju procenu dugovremene i kratkovremene glasnosti kao i dinamičkih karakteristika ovih funkcija, za različite vrednosti kratkovremenog usrednjavanja. Na Sl. 6. prikazan je rezultat analize

dugovremene i kratkovremene glasnosti za jedan segment opere „Karmen“ Žorža Bizeta. Na gornjem grafiku data je vremenska promena analiziranog signala, a na donjem promena glasnosti po dva realizovana modela. Osim grafičkog prikaza program daje i podatke o parametrima koji kvantifikuju varijacije glasnosti na zadatom signalu. Program je realizovan sa ciljem da posluži kao alatka u kreiranju signala za subjektivna testiranja.



Slika 6. Izlaz koji daju realizovani modeli, kv je kratkovremena, dv dugovremena glasnost

#### LITERATURA

- [1] E. Skovborg, R. Quesnel, and S.H. Nielsen, "Loudness Assessment of Music and Speech", in *Proceedings of the AES 116<sup>th</sup> Convention*, Berlin, Preprint 6143, 2000.
- [2] M. Mijić, D. Mašović, M. Petrović, D. Šumarac Pavlović, "Statistical properties of musical signals", 126<sup>th</sup> AES Convention, Preprint No. , 2009.
- [3] E. Skovborg and T. Lund "Loudness Descriptors to Characterize Programs and Music Tracks", in *Proceedings of the AES 125<sup>th</sup> Convention*, San Francisco, 2008.
- [4] E. Skovborg, and S.H. Nielsen, "Evaluation of different loudness Models with Music and Speech Material", In proceeding of the AES 117<sup>th</sup> Convention, San Francisco, Preprint 6234, 2004.
- [5] ITU-R Rec. BS.1770 "Algorithms to Measure Audio Programme Loudness and True-Peak Level", Recommendation BS.1770-1, 2006.
- [6] Norcross, Scott G, Lavoie, Michel C. "Investigations on the Inclusion of the LFE Channel in the ITU-R BS.1770-1 Loudness Algorithm," AES Convention Paper 7829, 127<sup>th</sup> AES Convention, New York 2009
- [7] Cabrera, Densil, Dash, Ian, Miranda, Luis, "Multichannel loudness istening Tests", AES Convention Paper 7451, 124<sup>th</sup> AES Convention, Amsterdam 2008
- [8] Soulodre, G. & Lavoie, M.C. "Development and Evaluation of Short-Term Loudness Meters", in Proceedings of the 121<sup>th</sup> AES Convention, 2006.

#### ABSTRACT

Loudness prediction and control are important in television and radio broadcasting, especially since so called 'loudness-war' between emitters has started. This paper presents an overview of models which are used for loudness estimation. Using the known algorithms, the two single-band models for loudness estimation of non-stationary signals have been realized and described in this paper.

#### REALIZATION OF TWO SINGLE-BAND LOUDNESS MODELS

Sanja Rakić and Ivana Čpajak