

PROBABILITY DENSITY FUNCTION OF AVERAGE POWER OF REAL SPEECH SIGNALS

Goran Petković

The Academy of Applied Technical and Preschool Studies, Serbia, goran.petkovic@atvssnis.edu.rs

Abstract: It is known that many real signals, such as speech signals, are non-stationary processes that express their feature through changes in parameters over time. By observing speech in shorter time intervals, the property of stationarity can be noticed. This characteristic enables the application of techniques for adaptation to local signal characteristics in signal processing algorithms. Many of these algorithms are described by standards, and in addition to intensive development in this area in last decades, there is a constant need for the development of new solutions and standards. One of the most used parameters of the speech signal for adaptation is the mean (average) signal power. The change of speech in time results in a wide dynamic range of average power. In addition to the predicted dynamics of average power, in the design of systems with adaptive techniques it is important to include other parameters, among which is the Probability Density Function (PDF) of the average power. The goal of the research presented in this paper is the analysis of the probability distribution of the average power of speech signals, based on which the adaptability in the development of algorithms in digital processing would be improved, which would ensure higher quality and less requirements in data transmission and storage. In addition to the theoretical consideration, the research was conducted on real speech signals of different speakers. In modern technical systems, where Internet technologies are prominent, processing, transmission and memorization of speech is executed frame by frame. Therefore, an analysis of the probability density of the average power for different frame lengths was carried out in the experiment. In the experimental part, for each of the speech signals of the speech corpus, the Probability Density Function that best describes the average power values per frame was determined. Experimental research results indicate that the function that best describes the average power is different for different speakers. In addition, when observing one speaker, the Probability Density Function is different for different frame lengths. It can be concluded that when it comes to adaptive techniques in the digital processing of the speech signal, it is important to consider the characteristics of the average power, among which is the Probability Density Function of the average power.

Keywords: Speech signals, Statistical parameters of the signal, Signal processing, Adaptation techniques

FUNKCIJA GUSTINE RASPODELE VEROVATNOĆE SREDNJE SNAGE REALNIH GOVORNIH SIGNALA

Goran Petković

Akademija tehničko-vaspitačkih strukovnih studija, Srbija, goran.petkovic@atvssnis.edu.rs

Rezime: Poznato je da mnogi realni signali, poput govornog signala, predstavljaju nestacionarne procese koji tu svoju osobenost izražavaju kroz promene parametara u vremenu. Posmatranjem govora u kraćim vremenskim intervalima može se zapaziti osobina stacionarnosti. Ova karakteristika omogućuje da se u algoritmima za obradu signala primene tehnike kojima se vrši adaptacija na lokalne karakteristike signala. Mnogi od ovih algoritama su opisani standardima, a pored višedecenijskog intenzivnog razvoja u ovoj oblasti postoji stalna potreba za razvojem novih rešenja i standarda. Jedan od najkorišćenijih parametara govornog signala koji se koristi u adaptaciji je srednja (prosečna) snaga signala. Promena govora u vremenu ima za posledicu širok dinamički opseg srednje snage. Pored predviđene dinamike srednje snage, u projektovanju sistema sa adaptabilnim tehnikama značajno je uključiti i druge parametre među kojima je i funkcija gustine raspodele verovatnoće srednje snage. Cilj sprovedenih istraživanja koja su prezentovana u ovom radu je analiza gustine raspodele verovatnoće prosečne snage govornih signala, na osnovu koje bi se poboljšala adaptabilnost u razvoju algoritama u digitalnog obradi, koji bi obezbedili veći kvalitet i manje zahteve u prenosu i memorisanju podataka. Pored teorijskog razmatranja, istraživanje je sprovedeno na realnim govornim signalima različitih govornika. S obzirom na to da se u savremenim tehničkim sistemima, među kojima su sve izraženije internet tehnologije, za obradu, prenos i memorisanje govora koristi obrada po frejmovima, u sprovedenom eksperimentu je izvršena analiza gustine raspodele srednje snage za različite dužine frejma. U eksperimentalnom delu je za svaki od govornih signala govornog korpusa određena funkcija gustine raspodele verovatnoće koja najbolje opisuje vrednosti prosečne snage po frejmovima. Eksperimentalni rezultati istraživanja ukazuju na to da je funkcija koja najbolje opisuje prosečnu snagu različita za različite govornike a pored toga, kada

se posmatra jedan govornik, funkcija gustine raspodele se razlikuje za različite dužine frejma. Može se zaključiti da je u adaptibilnim tehnikama u digitalnoj obradi govornog signala važno uzeti u obzir karakteristike prosečne snage među kojima je i funkciju gustine raspodele prosečne snage.

Cljučne reči: Govorni signali, Statistički parametri signala, Obrada signala, Adaptibilne tehnike

1. UVOD

Govor je jedan od najvažnijih signala u prirodi, koji omogućava neposredan prenos informacija između ljudi. U poslednje vreme značajna pažnja posvećuje se i komunikaciji ljudi i mašina putem govornog signala.

Početak digitalnog doba, velika energija je usmerena ka digitalizaciji i prenosu govora na daljinu putem telekomunikacione infrastrukture. Sa neslučenim razvojem interneta i porastom broja korisničkih uređaja, dodatno su povećani zahtevi za racionalno korišćenje propusnog opsega komunikacionih sistema i efikasno skladištenje podataka. Iako je razvijen veliki broj tehnika u obradi govornog signala još uvek postoji potreba za razvojem novih i unapređenjem postojećih tehnika koje obezbeđuju pouzdaniji prenos i vernu reprodukciju govornog signala.

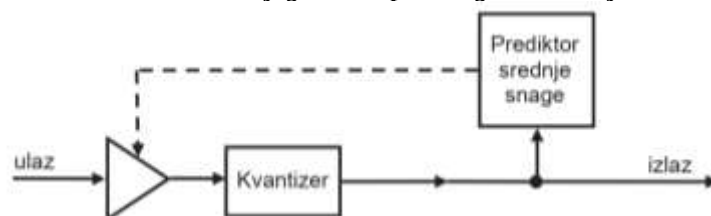
Posebno mesto u sistemima za digitalnu obradu signala od digitalizacije do kompresije pripada adaptibilnim tehnikama među kojima je najkorišćeniji parametar za adaptaciju srednja snaga signala (Chu, 2003; Sayood, 2005; Perić, & Nikolić, 2012; Perić i dr., 2020; Petković, 2016).

Ako je govorni signal uzorkovan (odmeren) i sadrži N odmeraka srednja (prosečna) snaga P se može odrediti korišćenjem sledećeg izraza:

$$P = \frac{1}{N-1} \sum_{i=1}^N x_i^2$$

Na slici 1 je prikazan primer primene adaptacije u kvantovanju korišćenjem srednje snage.

Slika 1. Blok dijagram adaptabilnog kvantovanja

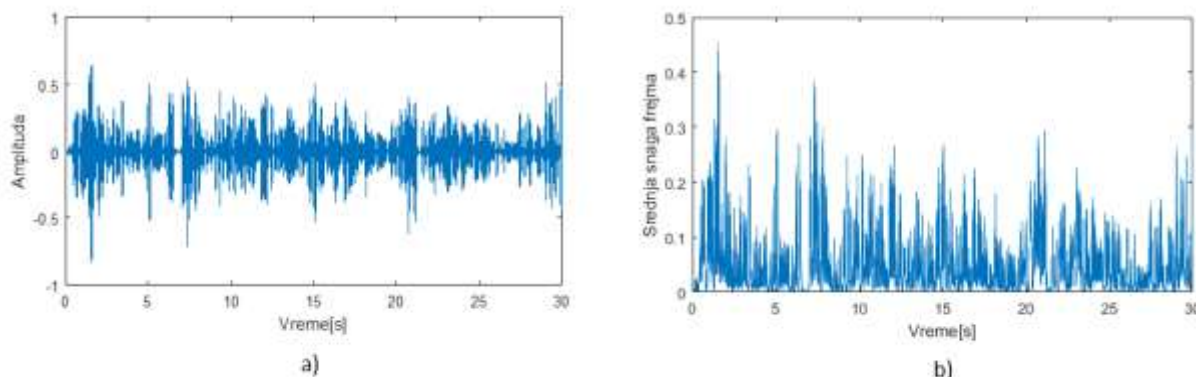


Uspešna adaptacija na srednju snagu uključuje poznavanje funkcije gustine raspodele srednje snage.

2. FUNKCIJA GUSTINE RASPODELE SREDNJE SNAGE GOVORNOG SIGNALA

Govorni signal predstavlja nestacionarni proces koji tu svoju osobenost izražava kroz promene srednje snage u vremenu, što za posledicu ima širok dinamički opseg signala (Gazor & Zhang, 2003; Petković, 2015).

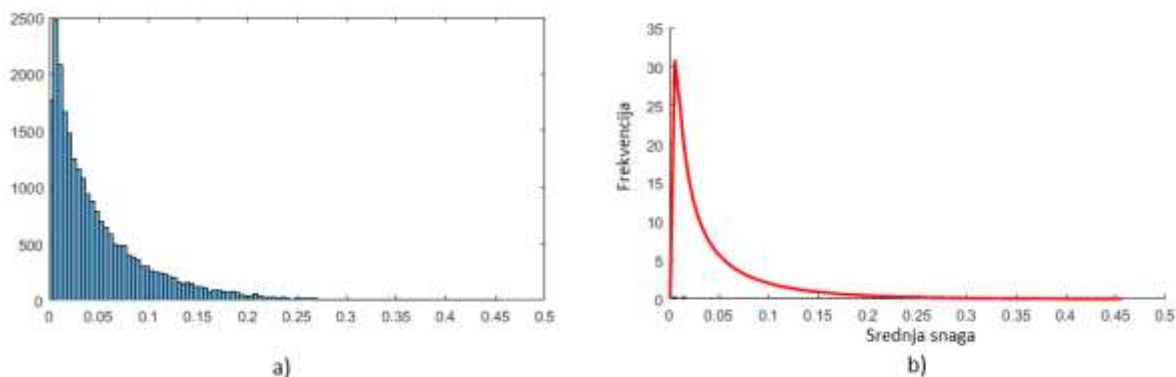
Slika 2. Govorni signal u vremenskom domenu (a) i promena vrednosti srednje snage u vremenu (b)



Na slici 2 je prikazan primer govornog signala u vremenskom domenu (a) i promene srednje snage istog govornog signala u vremenu (b) za frejm dužine 10 odmeraka. Može se primetiti da su vrednosti srednje snage u korelaciji sa promenama amplitude govornog signala (Petković, 2021; Petković, Perić, & Stoimenov, 2016).

Obradom podataka dobija se histogram vrednosti srednje snage, koji prikazuje raspodelu vrednosti po intervalima (slika 3.a). Na slici 3.b je prikazana funkcija gustine raspodele srednje snage. Cilj je izabrati funkciju koja će po određenom kriterijumu najviše odgovarati vrednostima srednje snage signala. Jedan od načina u izboru je metod procene maksimalne verovatnoće (engl. maximum likelihood estimation - MLE) tako da su u pretpostavljenom statističkom modelu uočeni podaci najverovatniji.

Slika 3. Histogram srednje snage (a) i funkcija gustine raspodele srednje snage



Očekuje se da se korišćenjem funkcije gustine raspodele srednje snage u adaptaciji može postići veći kvalitet u obradi signala (Petković, Perić, & Despotović, 2017).

3. EKSPERIMENTALNI REZULTATI

Analiza funkcije gustine raspodele srednje snage govornog signala izvršena je na dostupnim govornim signalima. Analizirano je sedam signala snimljenih korišćenjem PC računara koji su označeni sa GS01, GS02...GS07. Signali su filtrirani filtrom opsega 400Hz – 3400Hz.

Cilj eksperimentalnog dela istraživanja je bio odrediti funkciju gustine raspodele verovatnoće koja će najbolje opisivati srednju snagu realnih govornih signala.

Za analizu su korišćene funkcije *fitdist* i *mle* softverskog paketa Matlab. Za izabranu dužinu frejma koja je izražena brojem odmeraka govorni signal je podeljen na frejmova i za svaki frejm je određena srednja snaga. Izabrana je funkcija gustine raspodele iz liste dostupnih funkcija za gustinu raspodele u Matlabu (tabela 1), a funkcija *fitdist* je iskorišćena za podešavanje parametara izabrane funkcije dobijenim vrednostima srednje snage za signal. Postupak je ponovljen za sve funkcije gustine raspodele dostupne u Matlabu (tabela 1). Na osnovu rezultata funkcije *mle* određena je funkcija gustine raspodele koja najbolje opisuje srednju snagu govornog signala. Rezultati opisanog postupka („najbolja“ funkcija) prikazani su u tabeli 2 za različite dužine frejma.

Tabela 1. Funkcije gustine raspodele verovatnoće

Funkcija gustine raspodele	Skraćeni naziv	Funkcija gustine raspodele	Skraćeni naziv
Normalna	norm	Generalizovana ekst. vr.	gev
Lognormalna	logn	Generalizovana pareto	gp
Eksponecijalna	eksp	Inverzna Gausova	invga
Gama	gama	Birnbaum-Saundersova	birn
Logistička	logi	Beta	beta
Log-logistička	logl	Nakagami	nkg
Uniformna	uni	Rajlijeva	rajl
Vejbulova	vejb	Ricianova	ric
Ekstremnih vrednosti	ev		

U tabeli 2 su prikazane funkcije gustine raspodele prosečne snage koje najbolje opisuju govorni signal za izabranu dužinu frejma. Zbog preglednosti, u tabeli 2. su upisani skraćeni nazivi funkcija, koji su prethodno dati u tabeli 1. Rezultati su prikazani za različite dužine frejma (10 do 1000 odmeraka).

Tabela 2. Funkcije gustine raspodele koje najbolje opisuju prosečnu snagu govornog signala

	Dužina frejma [broj odmeraka]												
	10	20	40	60	80	100	150	300	400	500	600	800	1000
GS01	gp	gama	gp	gp	gp	vejb	vejb	vejb	vejb	vejb	vejb	beta	gev
GS02	birn	birn	birn	birn	birn	birn	birn	birn	birn	birn	logn	logn	llogl
GS03	vejb	vejb	beta	gp	gp	beta	beta	beta	beta	vejb	vejb	vejb	vejb
GS04	birn	gama	beta	beta	beta	beta	beta	beta	beta	beta	beta	beta	beta
GS05	birn	gama	gama	gama	gama	gama	gama	gama	gama	gama	beta	gama	gama
GS06	beta	beta	nkg	nkg	nkg	nkg	nkg	nkg	nkg	nkg	vejb	vejb	vejb
GS07	gama	gama	gama	gama	gama	gama	gama	gama	gama	beta	gama	gama	gama

4. DISKUSIJA

Iz prikazanih podataka se može konstatovati da govornim signalima različitih govornika odgovaraju različite funkcije gustine raspodele srednje snage. Posmatrajući jednog govornika pojedinačno, zapaža se da se razlikuje funkcija gustine raspodele za različite dužine frejma.

Dobijeni rezultati se mogu primeniti u adaptibilnim tehnikama u algoritmima za digitalnu obradu signala.

U daljim istraživanjima, u cilju potvrde dobijenih rezultata i izvedenih zaključaka, nameće se potreba analize signala govornog korpusa koji obuhvata veći broj govornika, govornike različitih jezika, pola i uzrasta.

REFERENCE

- Chu, W. C. (2003). *Speech Coding Algorithms*. New Jersey: Foundation and Evolution of Standardized Coders. John Wiley & Sons, Chapter 5, 143-158.
- Gazor S., & Zhang, W. (2003). *Speech probability distribution*. IEEE Signal Process. Letters, 10(7), 204-207. doi: 10.1109/LSP.2003.813679
- Perić, Z., & Nikolić, J. (2012). An adaptive waveform coding algorithm and its application in speech coding. *Digital Signal Processing*, 22(1), 199-209. doi: 10.1016/j.dsp.2011.09.001
- Peric Z., Petkovic G., Denic, B., Stanimirovic, A., Despotovic, V., & Stoimenov, L. (2020). Gaussian Source Coding using a Simple Switched Quantization Algorithm and Variable Length Codewords. *Advances in Electrical and Computer Engineering*, 20(4), 11-18. doi: 10.4316/AECE.2020.04002
- Petković, G. (2015). The Characteristics of the Speech Signal and Application in Signal Compression. *Sixth International Scientific Conference "Knowledge-The Power of Knowledge", Agia Triada, Greece, Institute of knowledge management - Skopje*, 10(1), 545-548.
- Petković G. (2016). Average power and variance in the processing of real voice signals. *Tenth International Scientific Conference "Knowledge-The Power of Knowledge", Agia Triada, Greece, 7-9.10.2016, Institute of knowledge management - Skopje*, 14(2), 740-744.
- Petković, G., Perić, Z., & Despotović, V. (2017). Switched uniform scalar quantization adapted to mean and variance for speech coding. *Facta universitatis - Series: Automatic Control and Robotics*, 16(3), 263-274. doi: 10.22190/FUACR1703263P
- Petković G. (2021). Autokorelacija odmeraka realnih govornih signala. Niš: *Zbornik radova, Akademija tehničko-vaspitačkih strukovnih studija*.
- Petković, G., Perić, Z., & Stoimenov, L. (2016). Switched Scalar Optimal μ -law Quantization with Adaptation Performed to both the Variance and the Distribution of Speech Signal. *Elektronika Ir Elektrotehnika, Kaunas University of Technology*, 22(1), 64-67. doi: 10.5755/j01.eee.22.1.14111
- Sayood, K. (2005). *Introduction to Data Compression*. San Francisco: Elsevier Science, Chapter 9, 227-270.
- Vucic, N., Peric, Z., & Petkovic, G. (2018). Design of Switched Quantizers and SpeechCoding Based on Quasi-Logarithmic Compandor. *Elektronika Ir Elektrotehnika, Kaunas University of Technology*, 24(6), 82-86. doi: 10.5755/j01.eie.24.6.22295