

# Prekidačka skalarna kvantizacija sa adaptacijom na snagu i raspodelu govornog signala

Goran M. Petković, Zoran H. Perić, Leonid V. Stoimenov

**Sadržaj** — U ovom radu su analizirani modeli prekidačke skalarnе kvantizacije za izvor sa Laplasovom i Gausovom raspodelom. Analizirani su rezultati na realnom telefonskom govoru i predložen je model prekidačke skalarnе kvantizacije koji, pored adaptacije na snagu govora, sadrži i adaptaciju na raspodelu signala (Gausovu i Laplasovu), što kao rezultat daje bolji kvalitet kvantizovanog govornog signala izražen odnosom signal-šum kvantizacije.

**Ključne reči** — Algoritmi kvantizacije, Obrada govornog signala, Skalarna prekidačka kvantizacija

## I. UVOD

KVANTIZERI imaju važnu ulogu u teoriji i praksi današnje moderne obrade signala [1]-[5]. Postupkom kvantizacije trenutna vrednost pobudnog signala, koja u opštem slučaju pripada neizbrojivom skupu vrednosti iz kontinualnog amplitudskog opsega, transformiše se u najbližu dozvoljenu vrednost iz konačnog, diskretnog skupa amplituda. Kvantizacija, nije samo jednostavna operacija koja prethodi kodiranju signala već predstavlja efikasnu tehniku za kompresiju podataka [1]-[5].

Poznato je da mnogi realni signali, poput govornog signala, predstavljaju nestacionarne procese koji tu svoju osobenost izražavaju kroz promene srednje snage u vremenu, što za posledicu ima širok dinamički opseg. Kako je kod mnogih realnih signala promena srednje snage spora, u većini slučajeva se u kratkim vremenskim intervalima proces može smatrati stacionarnim.

Da bi se postigao maksimalan kvalitet primljenog govornog signala, za datu brzinu prenosa, pri njegovom kodovanju preporučuje se upotreba adaptivnih koda, odnosno kvantizera koji se stalno u toku kvantizacije adaptiraju na lokalne statističke karakteristike izvora [6].

Pored često korišćene adaptacije na snagu govornog signala, u predloženom modelu, vrši se i adaptacija na raspodelu govornog signala što poboljšava kvalitet kvantovanja izražen odnosom signal-šum kvantizacije (SQNR).

U poglavlju II date su teorijske osnovе skalarnе

Goran M. Petković, Visoka škola primenjenih strukovnih studija u Vranju, Filipa Filipovića 20, 17500 Vranje, Srbija (telefon: 381-17-421589, e-mail: [goranpetko@yahoo.com](mailto:goranpetko@yahoo.com))

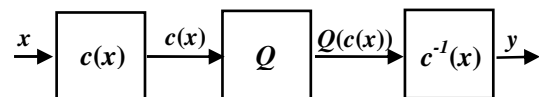
Zoran H. Perić, Elektronski fakultet u Nišu, Aleksandra Medvedeva 14, 18000 Niš, Srbija (telefon: 381-63-691961, e-mail: [zoran.peric@elfak.ni.ac.rs](mailto:zoran.peric@elfak.ni.ac.rs))

Leonid V. Stoimenov, Elektronski fakultet u Nišu, Aleksandra Medvedeva 14, 18000 Niš, Srbija (telefon: 381-18-642235, e-mail: [leni@elfak.ni.ac.rs](mailto:leni@elfak.ni.ac.rs))

kvantizacije za izvor sa Laplasovom i izvor sa Gausovom raspodelom. U poglavlju III prezentovan je novi model adaptivne skalarnе prekidačke kvantizacije govornog signala. Eksperimentalni rezultati prikazani su u poglavlju IV.

## II. MODEL KOMPANDOVANJA

Jedan od načina realizacije neuniformne kvantizacije predložio je Bennet [1]-[3]. On je uveo tehniku kompondovanja kojom se neuniformna kvantizacija postiže kompresijom ulaznog signala  $x$  korišćenjem kompresora sa nelinearnom karakteristikom  $c(\cdot)$ , zatim kvantizacijom komprimovanog signala  $c(x)$  primenom uniformnog kvantizera i na kraju ekspanovanjem kvantovane vrednosti komprimovanog signala, korišćenjem nelinearne inverzne karakteristike kompresije  $c^{-1}(\cdot)$  [1]-[3]. Struktura opisanog neuniformnog kvantizera sastavljena od redne veze kompresora, uniformnog kvantizera i ekspandora se naziva kompondor i prikazana je na Sl.1.



Sl.1. Blok diajagram tehnike kompondovanja

U situacijama kao što je kodovanje govora, tačna vrednost ulazne varijanse nije poznata unapred, a pored toga, ona se menja u vremenu. U takvoj situaciji, konstantan odnos signal-šum kvantizacije u širokom opsegu ulazne varijanse, može se dobiti primenom logaritamskog zakona kompresije:

$$c(x) = x_{\max} \frac{\ln(1 + \mu \frac{|x|}{x_{\max}})}{\ln(1 + \mu)} \operatorname{sgn} x, \quad (1)$$

pri čemu je  $\mu$  bezdimenzioni parametar koji se zove faktor kompresije, a  $x_{\max}$  je amplituda maksimalnog opterećenja kvantizera koja definiše amplitudski opseg skalarnog kompondora u granicama  $[-x_{\max}, x_{\max}]$ .

N-to stepeni neuniformni skalarni kvantizer izvora okarakterisan kao kontinualna slučajna promenljiva sa raspodelom verovatnoće  $p(x)$  ima distorziju definisanu kao očekivanu srednje kvadratnu grešku između originalnog i kvantovanog signala. Ukupna distorzija se sastoji od dve komponente, granularne distorzije i distorzije prekoračenja. Simbolički,

$$D_t = D_g + D_o \quad (2)$$

gde su granularna i distorzija prekoračenja definisane izrazima:

$$D_g = \sum_{i=2}^{N-1} \int_{t_{i-1}}^{t_i} (x - y_i)^2 p(x) dx \quad (3)$$

$$D_o = 2 \int_{t_{N-1}}^{\infty} (x - y_N)^2 p(x) dx \quad (4)$$

Jednakost (3) može biti predstavljena izrazom:

$$D_g = \frac{\ln^2(1+\mu)}{3N^2} \sigma^2 \left[ \frac{1}{\mu^2} \frac{x_{\max}^2}{\sigma^2} + \frac{2}{\mu\sigma^2} x_{\max} \overline{|x|} + \frac{\overline{x^2}}{\sigma^2} \right] \quad (5)$$

gde su  $\overline{|x|}$  i  $\overline{x^2}$  definisani izrazima:

$$\overline{|x|} = 2 \int_0^{x_{\max}} xp(x) dx \quad (6)$$

$$\overline{x^2} = 2 \int_0^{x_{\max}} x^2 p(x) dx \quad (7)$$

#### A. Distorzija za Laplasov izvor

Neka je govorni signal predstavljen Laplasovom funkcijom raspodele. Laplasova funkcija gustine raspodele originalne slučajne promenljive  $x$  sa jediničnom varijansom može biti predstavljena izrazom:

$$p(x, \sigma) = \frac{\sqrt{2}}{2\sigma} e^{-\frac{|x|\sqrt{2}}{\sigma}} \quad (8)$$

Smenom (8), (7), (6) u (5) i nakon rešavanja ova dva integrala i određene aproksimacije za distorziju dobijamo[7]:

$$D_i = \frac{\ln^2(1+\mu)}{3N^2} \sigma^2 \left[ \frac{1}{\mu^2} \frac{x_{\max}^2}{\sigma^2} + \frac{x_{\max}}{\sigma} \frac{\sqrt{2}}{\mu} + 1 \right] + \sigma^2 e^{-\frac{\sqrt{2}x_{\max}}{\sigma}} \quad (9)$$

Može se izračunati odnos signal-šum kvantizacije (SQNR) za kvantizaciju Laplasovog izvora za snagu u širokom dinamičkom opsegu.

$$SQNR = 10 \lg \frac{\sigma^2}{D_i} \quad (10)$$

Obično se u cilju pojednostavljenja analize performansi kvantizera uvodi pretpostavka da je srednja vrednost ulaznog signala jednaka nuli, tako da je u poslednjem izrazu snaga ulaznog signala zamenjena njegovom varijansom. Može se pokazati da uvođenjem ove pretpostavke nije izgubljena opštost.

#### B. Distorzija za Gausov izvor

Slično prethodno sprovedenoj analizi može se analizirati govorni signal predstavljen Gausovom funkcijom raspodele:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \quad (11)$$

Zamenom (11), (7), (6) u (5), nakon rešavanja integrala, uz aproksimaciju, dobija se izraz za distorziju i odnos signal-šum kvantizacije:

$$D_i = \sigma^2 \left( \frac{\ln^2(1+\mu)}{3N^2} \left[ \frac{1}{\mu^2} c^2 + c \frac{2}{\mu} \sqrt{\frac{2}{\pi}} + 1 \right] + \left[ (1+c^2) \left( 1 - \operatorname{erf} \frac{c}{\sqrt{2}} \right) - c e^{-\frac{c^2}{2}} \sqrt{\frac{2}{\pi}} \right] \right) \quad (12)$$

$$SQNR = 10 \lg \frac{\sigma^2}{D_i} \quad (13)$$

### III. MODEL PREKIDAČKE KVANTIZACIJE SA ADAPTACIJOM NA PROSEČNU SNAGU I RASPODELU GOVORNOG SIGNALA

Jedan od načina za postizanje konstantnog odnosa signal-šum kvantizacije je prekidačka adaptivna skalarna kvantizacija sa kodnom knjigom kod koje se na osnovu ulaznog frejma odlučuje da li blok uzoraka pripada određenoj statističkoj klasi iz skupa od  $K$  mogućih klasa. Do prijemnika se kao dodatna informacija prenosi indeks koji identifikuje klasu odnosno kodnu knjigu.

Razvijeno je više modela iz ove klase koji se međusobno razlikuju u načinu adaptacije. U predloženom modelu, adaptacija se vrši na osnovu srednje snage i raspodele govornog signala.

#### A. Adaptacija na prosečnu snagu

Kvantizaciona tehnika u predloženom modelu koristi frejm po frejm princip obrade ulaznih odmeraka. Postupak obrade signala započinje baferovanjem. Nakon baferovanja  $j$ -tog frejma dužine  $M$  koji sadrži odmerke označene sa  $x_{j+i}$ ,  $i=0,1,\dots,M-1$ , vrši se najpre procena varijanse baferovanog frejma  $\sigma_j^2$  a zatim i njena log-uniformna kvantizacija  $\hat{\sigma}_j^2$  [3], [4], čime se od  $k$  raspoloživih kvantizera koji su projektovani za varijanse  $\hat{\sigma}_p^2$ ,  $p=1,\dots,k$  koje su log-uniformno raspodeljene u dinamičkom opsegu varijansi  $B=20 \log(\sigma_{\max}/\sigma_{\min})$ :

$20 \log \hat{\sigma}_p = 20 \log \sigma_{\min} + (2p-1)B/(2k)$ ,  $p=1,\dots,k$  primenom prekidačke tehnike vrši izbor jednog  $s$ -tog.

Pri projektovanju svakog od  $k$  raspoloživih kvantizera u svakom od  $k$  opsega na koji je čitav dinamički opseg varijansi  $B$  podeljen, optimalna vrednost odgovarajuće amplitude maksimalnog opterećenja određena je tako da je za  $\sigma = \hat{\sigma}_p$  ukupna distorzija minimalna. Optimizacija amplitude maksimalnog opterećenja odnosno ukupne distorzije izvršena je za fiksiranu standardnu vrednost parametra  $\mu=255$  tako što je optimizovan parametar  $c$ :

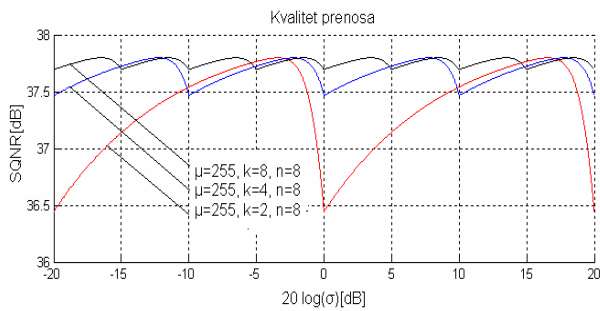
$$\frac{\partial D_i}{\partial c} = 0 \Rightarrow c = c_{opt} \quad (14)$$

koji predstavlja odnos  $x_{\max}/\hat{\sigma}_p$  i naziva se faktorom relativnog opsega kvantizera.

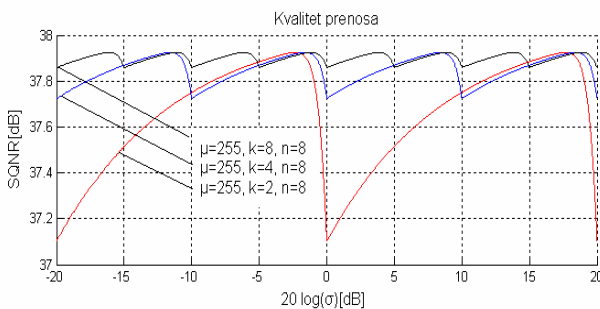
Veličina kodne knjige  $N$  zavisi od broja bitova korišćenih za kodovanje  $n$ . Relacija između  $N$  i  $n$  je  $N=2^n$ , gde je  $n$  broj bitova po odmerku [7].

Ako ne postoje ograničenja u pogledu veličine memorije i broja bitova po odmerku za prenosni sistem, tada postoji mogućnost izbora optimalnog broja kvantizera u modelu, za sistem u kome se može dobiti visok kvalitet meren SQNR-om, uz robusnost u širokom opsegu na ulazu. Ako se poveća broj kvantizera  $k$  dobija se odnos signal-šum

kvantizacije koji nema veće varijacije tokom promena ulazne snage (SI.2 i SI.3).



SI.2. Kvalitet prenosa izražen SQNR-om, za model sa dva, četiri i osam kvantizera za Laplasov izvor.



SI.3. Kvalitet prenosa izražen SQNR-om, za model sa dva, četiri i osam kvantizera za Gausov izvor.

### B. Adaptacija na raspodelu signala

Laplasova raspodela je dobra aproksimacija stvarne funkcije gustine verovatnoće trenutnih vrednosti kvalitetno snimljenog govornog signala. Za govorni signal snimljen mikrofonom slabijeg kvaliteta odstupanja od ove raspodele su znatno veća [1].

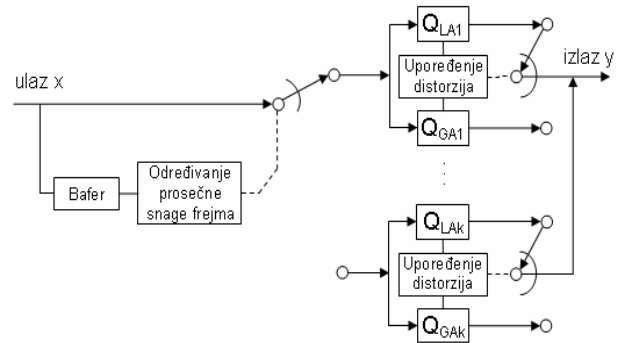
Ako posmatramo vremenski oblik govornog signala, može se primetiti da zvučni glasovi (vokali, poluvokali) imaju velike trenutne vrednosti. Relativno male trenutne vrednosti predstavljaju bezvučne glasove (frikative, afrikate itd.) Pošto su talasni oblici bezvučnih glasova stohastičke prirode (turbulencija strujanja vazduha u vokalnom traktu govornika) kao što je to i šum sistema za snimanje i prenos govora, to može uticati da dominantna raspodela bude Gausova.

Iz ovih razloga nameće se zaključak da nekada Laplasova a nekada Gausova raspodela bolje opisuje govorni signal, i to obzirom na prirodu govora i tehničke faktore u obradi i prenosu i to može biti izraženo kako na celom signalu tako i po segmentima govornog signala.

U predloženom modelu, nakon izbora kvantizera po kriterijumu prosečne snage frejma, vrši se kvantizacija i kvantizerom projektovanim za Gausovu raspodelu i kvantizerom za Laplasovu raspodelu, izračunava se distorzija za tako kvantovane signale i bira kvantizer koji obezbeđuje bolji kvalitet (SI.4).

Do prijemnika se kao dodatna informacija prenosi indeks koji identifikuje klasu odnosno kodnu knjigu određenu srednjom snagom frejma i informacija dužine jednog bita koja određuje da li će biti korišćen kvantizer projektovan za Laplasovu ili Gausovu raspodelu.

Informacije o izboru kvantizera se obično šalju na početku ili na kraju bloka.

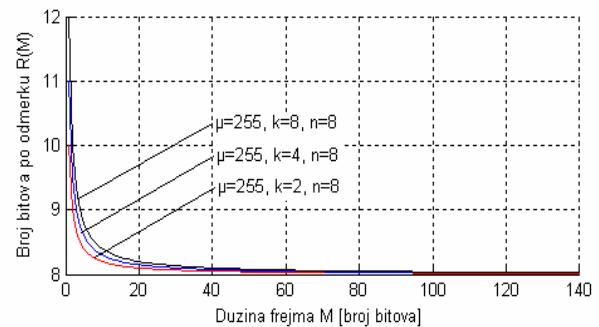


SI.4. Prekidačka adaptivna skalarna kvantizacija sa adaptacijom na prosečnu snagu i raspodelu signala govora

Ako svaka kodna knjiga od njih  $k$  ima veličinu  $N$ , broj bitova po odmerku je:

$$R = \log_2 N + \frac{\log_2 k}{M} + \frac{1}{M} = \log_2 N + \frac{\log_2 k + 1}{M} \quad (15)$$

Prenos informacije od jednog bita, koji opisuje raspodelu, povećava bitsku brzinu za sabirak  $1/M$  (15), i izražen je samo kod vrlo kratkih frejmova (SI.5).



SI.5. Bitska brzina za kvantizer sa adaptacijom na prosečnu snagu i raspodelu signala govora

## IV. EKSPERIMENTALNI REZULTATI

Analizirana su tri raspoloživa telefonska govorna signala označena sa GS 01, GS 02 i GS 03 čija je frekvencija odmeravanja 8kHz, a odmerci su kodirani sa 16 bitova. Signali su filtrirani filtrom opsega 300Hz-3400Hz. Kvantizeri su projektovani za dinamički opseg signala (-20dB 20dB). Analizirana je kvantizacija uzoraka sa 8 bitova. Analiza je izvršena za različite dužine frejmova.

Tabelama 1-3 predstavljene su vrednosti odnosa signal-šum kvantizacije za različit broj kodnih knjiga i za različite dužine frejmova. Sa La i Ga označeni su rezultati prekidačke kvantizacije za projektovane kvantizere sa Gausovom i Laplasovom raspodelom i adaptacijom samo na prosečnu snagu frejma, a sa LaGa rezultati dobijeni predloženim modelom sa adaptacijom i na srednju snagu i na raspodelu.

Može se zaključiti da su vrednosti dobijene eksperimentom nad realnim signalima, a koje se odnose na grešku kvantizacije odnosno odnos signal-šum kvantizacije u granicama predviđenim teorijskim razmatranjem (SI.2 i SI.3). Iz tabela se može potvrditi i teorijski zaključak da se

povećanjem broja kodnih knjiga postiže bolji odnos signal-šum kvantizacije.

TABELA 1: SQNR ZA MODEL SA DVA KVANTIZERA

2 kvantizera		Veličina frejma (broj odmeraka)						
		5	10	20	40	80	160	320
GS 01	La	37,15	37,19	37,25	37,32	37,31	37,32	37,33
	Ga	37,57	37,60	37,64	37,67	37,67	37,66	37,67
	LaGa	38,79	38,53	38,34	38,20	38,01	37,89	37,82
GS 02	La	37,12	37,11	37,10	37,08	37,08	37,08	37,09
	Ga	37,57	37,56	37,56	37,55	37,55	37,54	37,40
	LaGa	38,71	38,33	38,07	37,88	37,74	37,65	37,61
GS 03	La	37,09	37,12	37,13	37,14	37,15	37,17	37,19
	Ga	37,49	37,51	37,52	37,53	37,53	37,54	37,53
	LaGa	38,71	38,36	38,11	37,93	37,78	37,69	37,64

TABELA 2: SQNR ZA MODEL SA ČETIRI KVANTIZERA

4 kvantizera		Veličina frejma (broj odmeraka)						
		5	10	20	40	80	160	320
GS 01	La	37,56	37,56	37,57	37,60	37,60	37,60	37,61
	Ga	37,77	37,78	37,79	37,80	37,80	37,71	37,71
	LaGa	39,26	38,93	38,65	38,47	38,25	38,09	38,00
GS 02	La	37,57	37,60	37,60	37,62	37,61	37,62	37,63
	Ga	37,80	37,82	37,81	37,82	37,82	37,82	37,69
	LaGa	39,32	38,97	38,64	38,42	38,21	38,08	37,99
GS 03	La	37,55	37,57	37,58	37,59	37,59	37,61	37,61
	Ga	37,78	37,79	37,79	37,79	37,79	37,80	37,77
	LaGa	39,25	38,88	38,57	38,35	38,16	38,04	37,95

TABELA 3: SQNR ZA MODEL SA OSAM KVANTIZERA

8 kvantizera		Veličina frejma (broj odmeraka)						
		5	10	20	40	80	160	320
GS 01	La	37,66	37,70	37,70	37,70	37,70	37,70	37,71
	Ga	37,84	37,85	37,86	37,85	37,86	37,75	37,43
	LaGa	38,66	38,58	38,47	38,37	38,21	38,11	38,04
GS 02	La	37,66	37,69	37,70	37,71	37,71	37,72	37,72
	Ga	37,85	37,86	37,86	37,87	37,87	37,86	37,59
	LaGa	38,71	38,57	38,41	38,30	38,17	38,07	38,02
GS 03	La	37,66	37,67	37,69	37,69	37,69	37,70	37,70
	Ga	37,82	37,83	37,83	37,84	37,84	37,83	37,81
	LaGa	38,64	38,50	38,35	38,22	38,10	38,03	37,97

Upoređenjem modela sa adaptacijom samo na srednju snagu frejma zaključuje se da bolje rezultate daje model projektovan za Gausovu raspodelu u odnosu na model projektovan za Laplasovu raspodelu, a razlike su izraženije kod kraćih frejmova. Ovo se može objasniti time da je

snimljeni govor bolje opisan Gausovom raspodelom obzirom na to da u procesu snimanja nisu korišćene kvalitetne tehničke komponente. Veće razlike u kvalitetu su kod modela sa 2 kvantizera u odnosu na modele sa 4 i 8 kvantizera.

Predloženi model sa adaptacijom i na prosečnu snagu i na raspodelu govornog signala daje bolje rezultate za sve dužine frejmova, a poboljšanje je posebno izraženo kod kraćih frejmova kada je razlika u odnosu na prethodne modele i veća od 1dB.

Kod predloženog modela povećani su zahtevi u pogledu procesorske snage obzirom na to da je potrebno kvantizovati signal i za slučaj Gausove i za slučaj Laplasove raspodele, izračunati distorzije, a onda odabrati jedan od ova dva kvantizera. Razvojem savremenih hardverskih komponenata ovaj nedostatak je sve manje izražen kao faktor ograničenja u realizaciji.

## V. ZAKLJUČAK

Predloženim modelom prekidačke kvantizacije postiže se bolji kvalitet u prenosu govornog telefonskog signala u odnosu na model prekidačke skalarne kvantizacije gde se adaptacija vrši samo na snagu signala. Ovo poboljšanje postiže se na račun neznatnog povećanja prenosa informacija i povećanih zahteva u pogledu procesorske snage.

## LITERATURA

- [1] G. Lukatela, D. Drajić, G. Petrović, R. Petrović, *Digitalne telekomunikacije*, Građevinska knjiga, Beograd 1981.
- [2] W. C. Chu, *Speech Coding Algorithms, Foundation and Evolution of Standardized Coders*, New Jersey: John Wiley & Sons, 2003.
- [3] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Boston, Dordrecht, London: Kluwer Academic Publishers, 1992.
- [4] N. S. Jayant and P. Noll, *Digital Coding Of Waveforms, Principles and Applications to Speech and Video*, New Jersey: Prentice Hall, 1984.
- [5] L. R. Rabiner and R. W. Schafer, "Introduction to Digital Speech Processing", *Foundations and Trends in Signal Processing*, vol. 1, pp. 1-194, Jan. 2007.
- [6] Z. Perić, G. Petković, J. Nikolić, "Prekidačka skalarna kvantizacija govornog signala u širokom dinamičkom opsegu varijansi", *ETRA 2010*, Jun 2010.
- [7] J. Nikolic and Z. Peric, "Lloyd-Max's Algorithm Implementation in Speech Coding Algorithm Based on Forward Adaptive Technique", *Informatica*, vol. 19, No. 2, pp. 255-270, 2008.

## ABSTRACT

This paper analyzes the models for switching scalar quantization of a source with Laplace and Gaussian distribution. We analyzed the results of real telephone speech and proposed a model of switching scalar quantization, which, in addition to adaptation on the power of speech, includes the adaptation on the distribution of signals (Gaussian and Laplace), which resulted in a better quality of voice signal pronounced with Signal-to-Quantization-Noise Ratio.

## SWITCHED SCALAR QUANTIZATION WITH ADAPTATION PERFORMED ON BOTH THE POWER AND THE DISTRIBUTION OF SPEECH SIGNAL

Goran M. Petković, Zoran H. Perić, Leonid V. Stoimenov