

Kodovanje govornog signala unazad adaptivnim optimalnim skalarnim komparatorom

Zoran Perić, Jelena Nikolić, Aleksandar Jocić

Sadržaj — Ovaj rad predlaže novi model adaptivnog skalarnog komparatora koji se projektuje prema optimalnom zakonu kompresije, pri čemu se adaptacija amplitude maksimalnog opterećenja vrši korišćenjem memorije od jedne reči i tehnike adaptacije unazad. Značaj predloženog modela kvantizera pored njegove jednostavnosti ogleda se i u dobitku odnosa signal-šum kvantizacije koji se pri obradi govornog signala postiže u odnosu na slučaj kada adaptacija nije uključena. Dodatno, u radu je sagledan i značaj izbora težinskog faktora kao i njegov uticaj na performanse predloženog modela kvantizera.

Ključne reči — Optimalni zakon kompresije, skalarni komparator, tehnika adaptacija unazad, težinski faktor.

I. UVOD

PREDSTAVLJANJE odmeraka govornih signala upotrebom po mogućnosti što manjeg broja bita uz zadovoljavajući nivo kvaliteta govora definiše postupak poznat kao kvantizacija govora ili kodovanje govora [1]-[6]. Poznato je da se u cilju maksimizacije kvaliteta primljenog govora za datu brzinu prenosa, usled nestacionarnosti govornih signala, tj. promenljivosti njihove srednje snage u širokom opsegu, pri kodovanju ovih signala preporučuje upotreba adaptivnih koda, odnosno kvantizera koji se stalno u toku kvantizacije adaptiraju na lokalne statističke karakteristike izvora [1]-[6]. Konceptualna razlika adaptivnih tehnika ogleda se u tome da li se adaptacija vrši na osnovu niza odmeraka koji treba da se koduju tek posle posmatanog odmerka ili se pak adaptacija vrši na osnovu niza odmeraka koji su prethodili posmatranom odmerku i čije je kodovanje u trenutku nailaska posmatranog odmerka završeno [2], [3], [7]-[9]. Tako razlikujemo adaptaciju unapred i adaptaciju unazad. Opredelili smo se za tehniku adaptacije unazad koja za razliku od tehnike adaptacije unapred ne zahteva prenos tzv. dodatne informacije koja zapravo predstavlja informaciju o adaptaciji [2], [3], [8]. Pored toga, opredelili smo se za model adaptivnog kvantizera sa memorijom od jedne reči što znači da ćemo adaptaciju tekućeg odmerka vršiti na osnovu samo jednog prethodno kodovanog odmerka [3]. Važno je reći da smo se zbog jednostavnosti postupka skalarnog kvantizacije, kao i znatno manje

kompleksnosti projektovanja skalarnih kvantizera u odnosu na vektorske kvantizere, opredelili upravo za skalarnu kvantizaciju. Konačno, obzirom da smo u [10] pokazali da se optimalnim komparatorom, projektovanim za proces jedinične snage i Laplasov izvor mogu postići performanse bliske teorijski optimalnim performansama Lloyd-Max-ovog kvantizera, u ovom radu načinimo korak dalje i razmatraćemo implementaciju pomenutog modela skalarnog komparatora sa optimalnim zakonom kompresije u unazad adaptivnoj kodnoj šemi sa memorijom od jedne reči. Zapravo, cilj je istražiti da li predložena kodna šema, tj. model kvantizera, predstavlja efikasno rešenje pri obradi govornih signala za koje je poznato da se srednja snaga menja u širokom dinamičkom opsegu.

II. OPTIMALNI ZAKON KOMPRESIJE

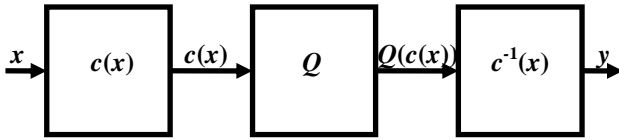
Opis modela unazad adaptivnog optimalnog skalarnog komparatora započemo u ovoj sekciji definisanjem postupka skalarnog kvantizacije i optimalnog zakona kompresije. Naime, funkcija skalarnog kvantizera je da trenutnu vrednost pobudnog signala, koja u opštem slučaju pripada neizbrojivom skupu vrednosti iz kontinualnog amplitudskog opsega, transformiše u najbližu dozvoljenu vrednost iz diskretnog skupa od N amplituda [2], [3]. Poznato je da je skalarni kvantizer jednoznačno određen izlaznim vrednostima y_i , koje se nazivaju još i reprezentacionim nivoima, i podelom ulaznog opsega vrednosti u N ćelija ili intervala kvantovanja $\Delta_i = (t_{i-1}, t_i]$, $i=1,2,\dots,N$ koji su definisani granicama odluke $\{t_0, t_1, \dots, t_N\}$ [2], [3]. Takođe, poznato je da konstantna veličina ćelija odgovara modelu uniformnog kvantizera dok je ne jednakim veličinama ćelija definisan model neuniformnog kvantizera.

Model neuniformnog skalarnog kvantizera predstavljen na Sl. 1. kaskadnom vezom kompresora, uniformnog kvantizera i ekspandora predstavlja model skalarnog komparatora [2], [3]. Prema ovom modelu neuniformno kvantovanje se može postići kompresijom ulaznog signala x korišćenjem kompresora sa nelinearnom karakteristikom $c(\cdot)$, zatim kvantovanjem komprimovanog signala $c(x)$ primenom uniformnog kvantizera i na kraju ekspandovanjem kvantovane vrednosti komprimovanog signala, korišćenjem nelinearne inverzne karakteristike kompresije $c^{-1}(\cdot)$. Poznato je da je karakteristika kompresije monotono rastuća neparno simetrična funkcija, čiji je oblik u prvom kvadrantu ilustrovan na Sl. 2.

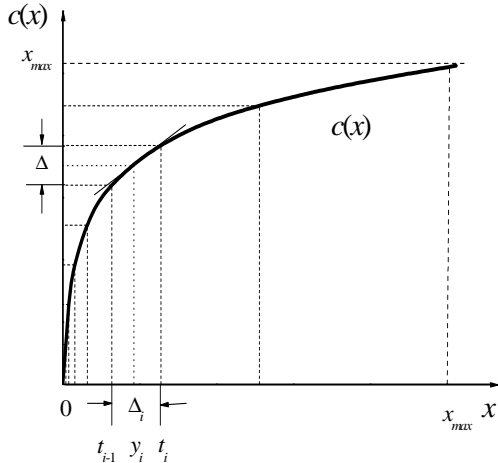
Z. Perić, Elektronski fakultet u Nišu, Srbija (telefon: 381-18-529101; faks: 381-18-588399; e-mail: zoran.peric@elfak.ni.ac.yu).

J. Nikolić, Elektronski fakultet u Nišu, Aleksandra Medvedeva 14, 18000 Niš, Srbija; (e-mail: jelena.nikolic@elfak.ni.ac.yu).

A. Jocić, Elektronski fakultet u Nišu, Aleksandra Medvedeva 14, 18000 Niš, Srbija; (e-mail: aleksandar.jocic@elfak.ni.ac.yu).



Sl. 1. Blok dijagram modela komandora



Sl. 2. Karakteristika kompresije

Podelom komprimovanog signala na jednake kvante označene sa Δ , moguće je korišćenjem karakteristike kompresije podeliti amplitudski opseg ulaznog signala x , na nejednake intervale kvantovanja Δ_i [2], [3]. Kako smo se opredelili za model skalarnog komandora sa optimalnim zakonom kompresije, potrebno je najpre definisati optimalni zakon kompresije koji određuje pomenuti model kvantizera [2], [3], [10]:

$$c(x) = x_{\max} \frac{\int_0^x p^{1/3}(x) dx}{\int_0^{x_{\max}} p^{1/3}(x) dx}, \quad (1)$$

pri čemu je sa x_{\max} označena amplituda maksimalnog opterećenja kvantizera koja definiše amplitudski opseg skalarnog komandora u granicama $[-x_{\max}, x_{\max}]$, a $p(x)$ je funkcija gustine verovatnoće trenutnih vrednosti signala na ulazu skalarnog komandora. Pretpostavljanjem Laplasove funkcije gustine verovatnoće odmeraka govornog signala:

$$p(x) = \frac{1}{\sqrt{2}\sigma} \exp\left(-\frac{|x|\sqrt{2}}{\sigma}\right), \quad (2)$$

jednostavno izvodimo sledeće relacije koje važe za reprezentacione nivoe y_i , $N/2 < i \leq N$ i granice odluke x_i , $N/2 < i < N$ koji se nalaze na pozitivnom delu ose:

$$x_i = \frac{3\sigma}{\sqrt{2}} \ln \left(\frac{N}{2N - 2i + (2i - N) \exp\left(-\frac{\sqrt{2}x_{\max}}{3\sigma}\right)} \right), \quad (3)$$

$$y_i = \frac{3\sigma}{\sqrt{2}} \ln \left(\frac{N}{2N - 2i + 1 + (2i - 1 - N) \exp\left(-\frac{\sqrt{2}x_{\max}}{3\sigma}\right)} \right), \quad (4)$$

na osnovu kojih se jednostavno određuju i simetrične granice i reprezentacioni nivoe na negativnom delu ose. Očigledno je da veličina reprezentacionih nivoe i granica odluke zavisi između ostalog i od veličine amplitude maksimalnog opterećenja. Kako je u radu [10] izveden eksplicitan izraz na osnovu kojeg se određuje veličina amplitude maksimalnog opterećenja komandora koji je optimalno projektovan za Laplasov izvor jedinične snage:

$$x_{\max}^f = \frac{3}{\sqrt{2}} \ln\left(\frac{N}{2}\right) + \sqrt{2}, \quad (5)$$

određićemo njegov faktor relativnog opsega C_{opt} :

$$C_{opt} = \frac{x_{\max}^f}{\sigma} = x_{\max}^f, \quad (6)$$

koji ćemo koristiti u narednoj sekciji pri projektovanju unazad adaptivnog skalarnog komandora.

III. TEORIJSKA POSTAVKA MODELA

U ovoj sekciji predložićemo veoma jednostavan model adaptivnog kvantizera sa memorijom od jedne reči. Zapravo opredelili smo se da razmatramo implementaciju modela skalarnog komandora sa optimalnim zakonom kompresije u unazad adaptivnoj kodnoj šemi sa memorijom od jedne reči. Napomenimo da je u literaturi do sada razmatrana jedino implementacija uniformnog modela skalarnog kvantizera u unazad adaptivnoj kodnoj šemi sa memorijom od jedne reči, a pomenuti model kvantizera sreće se pod imenom *Jayant quantizer* [3], [9].

Slično kao kod Jayant-ovog kvantizera, adaptacija amplitude maksimalnog opterećenja pri kvantizaciji n -tog odmerka se kod predloženog modela unazad adaptivnog komandora vrši na sledeći način:

$$x_{\max} = C_{opt} \hat{\sigma}_{y(n)}, \quad (7)$$

pri čemu važi rekurentna relacija:

$$\hat{\sigma}_{y(n)}^2 = \alpha \hat{\sigma}_{y(n-1)}^2 + (1 - \alpha) y^2(n-1), \quad (8)$$

gde je α težinski koeficijent koji može da uzima vrednosti u intervalu $0 < \alpha < 1$. Osim toga, $\hat{\sigma}_{y(n)}^2$ označava procenjenju varijansu adaptacije n -tog odmerka čija je kvantovana vrednost označena sa $y(n)$. Očigledno je da se adaptacija maksimalne amplitude pri kvantovanju tekućeg odmerka vrši na osnovu samo jednog prethodno kvantovanog odmerka. Takođe važno je istaći da izbor vrednosti težinskog koeficijenta α utiče na procenu varijanse a time i maksimalne amplitude opterećenja kvantizera. Stoga smo se opredelili da za raspoloživi govorni signal utvrdimo kako se izbor težinskog koeficijenta odražava na performanse predloženog modela kvantizera.

Performanse kvantizera najčešće su određene odnosom signal-šum kvantizacije $SQNR$, (signal to quantization noise ratio) koji se definiše na sledeći način [2], [3]:

$$SQNR = 10 \log \left(\frac{\sigma^2}{D} \right), \quad (9)$$

i izražava u dB pri čemu σ^2 predstavlja varijansu ulaznog signala x , dok je D distorzija uneta kvantizacijom. Postoje različite mere izobličenja (distorzije) ali je najčešće korišćena mera kvadrat razlike vrednosti ulaznih odmeraka i odgovarajućih vrednosti koje se javljaju na izlazu iz kvantizera [2], [3], [6], [9]. Zapravo, distorzija je cena koja se plaća diskretizacijom prostora R na N reprezentacionih nivoa, odnosno zamenom bilo koje ulazne amplitude x iz ćelije R_i $i=1, \dots, N$, reprezentacionim nivoom y_i $i=1, \dots, N$. U opštem slučaju distorzija uneta kvantizacijom predstavlja zbir granularne distorzije D_g i distorzije prekoračenja D_p :

$$D = D_g + D_p, \quad (10)$$

koje se definišu na sledeći način [2], [3], [6], [9]:

$$D_g = \sum_{j=1}^N \int_{t_{j-1}}^{t_j} (x - y_j)^2 p(x) dx, \quad (11)$$

$$D_p = 2 \int_{t_N}^{\infty} (x - y_N)^2 p(x) dx. \quad (12)$$

Kako performanse predloženog modela kvantizera određujemo na osnovu eksperimenta koji sprovodimo na raspoloživom realnom govornom signalu, potrebno je zapravo definisati trenutnu vrednost odnosa signal-šum kvantizacije recimo n -tog odmerka na sledeći način:

$$SQNR_n = 10 \log \left(\frac{(\sigma_n^{ex})^2}{D_n^{ex}} \right), \quad (13)$$

pri čemu je $(\sigma_n^{ex})^2$ varijansa n -tog odmerka, dok je D_n^{ex} distorzija tj. izobličenje n -tog odmerka koje je određeno srednje-kvadratnom greškom. Dodatno, definišemo i usrednjenu vrednost odnosa signal-šum kvantizacije:

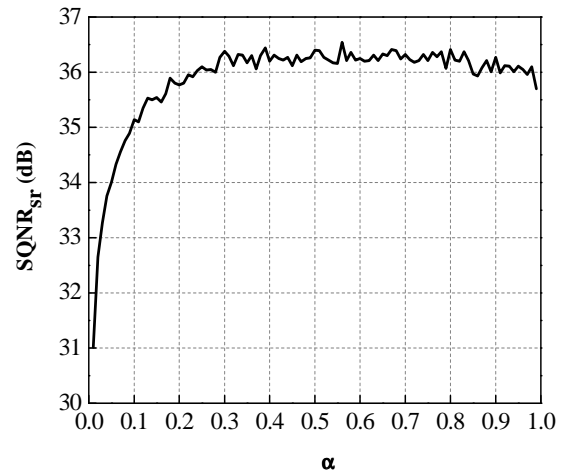
$$SQNR_{sr} = 10 \log \left(\frac{\sum_{k=1}^L (\sigma_k^{ex})^2}{\sum_{k=1}^L D_k^{ex}} \right), \quad (14)$$

radi efikasnijeg poređenja performansi predloženog modela kvantizera za različite vrednosti težinskog koeficijenta α .

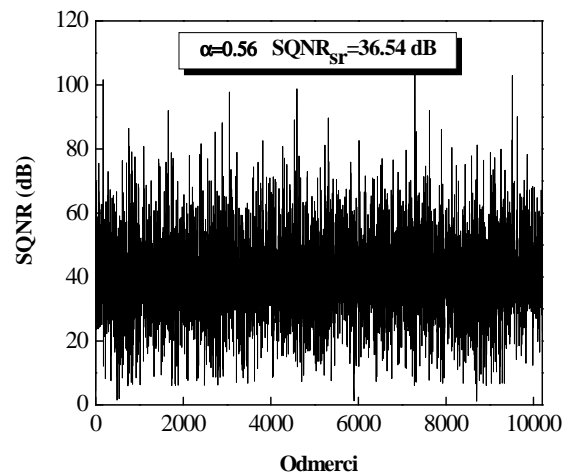
IV. NUMERIČKI REZULTATI

U cilju analize uticaja izbora težinskog koeficijenta, označenog sa α , na performanse predloženog modela kvantizera izvršili smo eksperiment na realnom govornom signalu dužine $L=10200$ odmeraka. Korišćenjem poslednjeg izraza za raspoloživi realni govorni signal odredili smo zavisnost usrednjene vrednosti odnosa signal-šum kvantizacije od težinskog koeficijenta α . Na osnovu pomenute zavisnosti koja je prikazana na Sl. 3., utvrdili smo da se izborom vrednosti koeficijenta $\alpha=0.56$ postiže maksimalna usrednjena vrednost odnosa signal-šum kvantizacije $SQNR_{sr}^{max}=36.54$ dB. Za tako izabranu

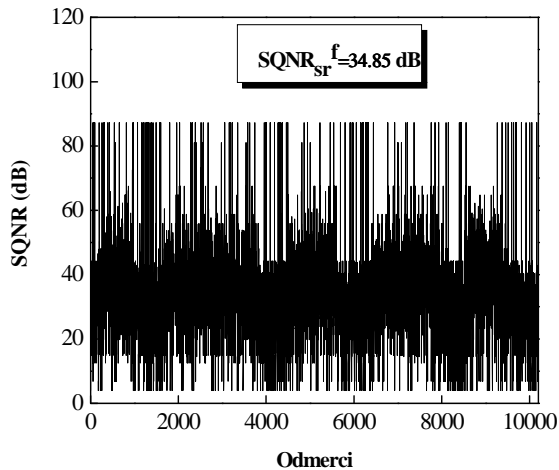
vrednost težinskog koeficijenta ($\alpha=0.56$) odredili smo promenu trenutne vrednosti odnosa signal-šum kvantizacije koja je prikazana na Sl. 4. Radi sagledavanja značaja izbora težinskog faktora kao i prednosti korišćenja adaptivne tehnike unazad, na Sl. 5. ilustrovana je promena trenutne vrednosti odnosa signal-šum kvantizacije za slučaj kada je raspoloživi govorni signal kvantovan fiksnim skalarnim komparatorom koji je optimalno projektovan za Laplasov izvor jedinične snage. Obzirom da smo utvrdili da varijansa raspoloživog govornog signala dužine $L=10200$ odmeraka iznosi $\sigma^2=0.9172$, (bliska jediničnoj varijansi), za varijanse različite od jedinične pretpostavljamo da će dobitak ostvaren korišćenjem predloženog modela kvantizera u odnosu na slučaj kada adaptacija nije uključena biti veći od 1.69 dB. Da bismo to i pokazali rešili smo da od raspoloživog govornog signala analiziramo samo prvih L_d odmeraka i



Sl. 3. Ilustracija uticaja parametra α na usrednjenu vrednost odnosa signal-šum kvantizacije $SQNR_{sr}$.



Sl. 4. Prikaz promene trenutne vrednosti odnosa signal-šum kvantizacije za slučaj kada je težinski faktor α izabran tako da se za raspoloživi realni govorni signal postiže maksimum usrednjene vrednosti odnosa signal-šum kvantizacije.



Sl. 5. Prikaz promene trenutne vrednosti odnosa signal-šum kvantizacije za slučaj kada je raspoloživi govorni signal kvantovan fiksnim skalarnim komparatorom koji je optimalno projektovan za Laplasov izvor jedinične snage.

TABELA 1: ILUSTRACIJA ZNAČAJA ADAPTACIJE

L_d	σ^2	$SQNR_{sr}^f$ [dB]	$SQNR_{sr}$ [dB]	G [dB]
640	1.30	35.29	36.56	1.27
1280	2.10	33.41	36.70	3.29
2550	1.27	33.81	36.58	2.77
5100	1.06	34.49	36.60	2.11
10200	0.92	34.85	36.54	1.69

da zapravo izvršimo poređenje kako varijansi tako i usrednjenih vrednosti odnosa signal-šum kvantizacije, pri čemu smo sa G označili *gain*, tj. dobitak unet adaptacijom a izražen razlikom odgovarajućih srednjih vrednosti odnosa signal-šum kvantizacije koje su prikazane u Tabeli 1. Na osnovu utvrđenih vrednosti ovog dobitka, može se uočiti da njegova maksimalna vrednost odgovara slučaju kada varijansa govornog signala, koji ima dužinu $L_d=1280$, ima vrednost $\sigma^2=2.10$, što je potvrdilo naša očekivanja. Napomenimo da smo rezultate prikazane u Tabeli 1 dobili za slučaj kada težinski koeficijent ima vrednost $\alpha=0.56$, što predstavlja optimalnu vrednost za signal dužine $L_d=10200$. Primitimo i to da su vrednosti $SQNR_{sr}$ prikazane u Tabeli 1 bliske, tako da možemo izvesti zaključak da predloženi model kvantizera predstavlja efikasno rešenje pri kvantovanju govornog signala, pri čemu je efekat poboljšanja u odnosu na slučaj kada ne postoji adaptacija veći ukoliko se varijansa signala više razlikuje od jedinične.

V. ZAKLJUČAK

U ovom radu predložen je jednostavan model unazad adaptivnog skalarnog komparatora koji je projektovan na osnovu optimalnog zakona kompresije, pri čemu je adaptacija maksimalne amplitude opterećenja skalarnog kvantizera izvršena na osnovu memorije od samo

jedne reči. Pokazano je da se predloženim modelom kvantizera može ostvariti značajan dobitak u odnosu na slučaj kada adaptacija nije uključena pri čemu je efekat poboljšanja veći ukoliko se varijansa signala više razlikuje od jedinične. Takođe, ustanovljeno je da se predloženim modelom kvantizera za različite vrednosti varijansi govornog signala postižu bliske usrednjene vrednosti odnosa signal-šum kvantizacije, na osnovu čega možemo opravdano zaključiti da predloženi model kvantizera predstavlja efikasno rešenje pri obradi govornih signala.

LITERATURA

- [1] W. C. Chu, *Speech coding algorithms, Foundation and Evolution of Standardized Coders*. New Jersey: John Wiley & Sons, 2003, ch. 5-6, pp. 143-183.
- [2] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Boston, Dordrecht, London: Kluwer Academic Publishers, 1992, ch. 5, pp. 133-172.
- [3] N. S. Jayant and P. Noll, *Digital Coding Of Waveforms, Principles and Applications to Speech and Video*. New Jersey: Prentice Hall Secondary Education Division, 1984, ch. 4-5, pp. 115-251.
- [4] M. H. Johnson and A. Alwan, *Speech coding: Fundamentals and applications*. In J. Proakis (Ed.), *Wiley Encyclopedia of Telecommunications 5*. New York: John Wiley & Sons, 2002, pp. 2340-2359.
- [5] A. Moffat and A. Turpin, *Compression and coding algorithms*. Norwell, MA, USA: Kluwer Academic Publishers, 2002.
- [6] 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.
- [7] J. Vass, Y. Zhao and X. Zhuang, "Adaptive Forward-Backward Quantizer for Low Bit Rate High Quality Speech Coding", *IEEE Trans. on Speech and Audio Processing*, vol. 5, issue 6, pp. 552-557, Nov. 1997.
- [8] A. Ortega and M. Vetterli, "Adaptive scalar quantization without side information", *IEEE Trans. on Image Processing*, vol. 6, no. 5, pp. 665-676, May 1997.
- [9] K. Sayood, *Introduction to Data Compression*. San Francisco, CA, USA: Morgan Kaufmann, 2000, ch. 8, pp.232-236.
- [10] Z. Peric and J. Nikolic, "Analysis of compressor functions for Laplacian source's scalar compandor construction", *Data Recording, Storage and Processing*, vol. 8, no. 2, pp. 15-24, June 2006.

ABSTRACT

This paper proposes a novel adaptive model of scalar compandor designed according to the optimal compression low. In order to adapt the support region of the quantizer, the proposed model requires one word memory and uses the backward adaptive technique. It is shown that the importance of the proposed model lies in its simplicity, but also can be viewed through the gain in signal to quantization noise ratio that can be obtained over the non adaptive model. Additionally, an analysis is performed in order to provide dependance of the proposed model performances on the weighted factor as well as in order to find the best solution for coding the available real speech signal.

SPEECH CODING USING BACKWARD ADAPTIVE OPTIMAL SCALAR COMPANDOR

Zoran Peric, Jelena Nikolic, Aleksandar Jovic