



Univerzitet u Nišu
Elektronski fakultet



MILAN Ž. TANČIĆ

**KONSTRUKCIJA LOGARITAMSKIH
KVANTIZERA ZA VISOKO KVALITETNO
ADAPTIVNO TRANSFORMACIONO
KODOVANJE GOVORNOG SIGNALA**

Doktorska disertacija

Niš, 2019.



University of Niš
Faculty of Electronic Engineering



MILAN Ž. TANČIĆ

**DESIGN OF LOGARITHMIC
QUANTIZERS FOR HIGH QUALITY
ADAPTIVE TRANSFORM CODING OF
SPEECH SIGNAL**

Doctoral dissertation

Niš, 2019.

Podaci o mentoru i doktorskoj disertaciji

Mentor: dr Zoran Perić, redovni profesor, Univerzitet u Nišu, Elektronski fakultet

Naslov: Konstrukcija logaritamskih kvantizera za visoko kvalitetno adaptivno transformaciono kodovanje govornog signala

Rezime: U ovoj disertaciji razmatra se projektovanje kvantizera, uređaja koji se u okviru telekomunikacionih sistema koriste za kodovanje signala. Glavni cilj ove disertacije je konstrukcija logaritamskih kvantizera za kodovanje govornog signala koji će obezbediti visok kvalitet rekonstruisanog signala na prijemnoj strani. Upotrebom transformacionih i tehnika adaptacije poboljšava se kvalitet kodovanja u odnosu na klasične telekomunikacione sisteme predložene od strane Međunarodne Telekomunikacione Unije.

Kvantizacija, jedan od najznačajnijih procesa pri obradi signala, predstavlja diskretizaciju po amplitudi kao osnovni korak u postupku digitalizacije analognih signala. Karakteristika održanja približno konstantnog kvaliteta kodovanja u širokom opsegu varijansi naziva se robustnost, te je pogodno koristiti logaritamske kvantizere koji zadovoljavaju ovu karakteristiku. Adaptacijom logaritamskih kvantizera povećava se kvalitet kodovanja jer adaptivni kvantizeri prate promenu snage signala i prilagođavaju kvantizacione nivoe čime obezbeđuju konstantno visok kvalitet u širokom opsegu varijansi.

Primenom metoda transformacionog kodovanja dodatno raste kvalitet kodovanja govornog signala. Transformaciono kodovanje podrazumeva udruživanje uzastopnih odmeraka signala i primenu određene transformacije nad njima čime se vrši preraspodela energije (informacije), što predstavlja dekorelaciju signala. U ovoj

disertaciji za projektovanje kodnih šema iskorišćene su neke od najpoznatijih transformacija kao što su Hadamarova, diskretna kosinusna i diskretna *wavelet* transformacija.

Dovođenjem diskretnog signala na ulaz sistema za kodovanje dodatno se povećava kompresija jer je diskretni signal amplitudski ograničen i nema distorzije prekoračenja što i na nižim bitskim brzinama obezbeđuje visok kvalitet kodovanja. Iz ovog razloga predložene su šeme za kodovanje diskretnog ulaznog signala.

Doprinos predloženih rešenja u okviru ove disertacije jasno se pokazuje kombinacijom navedenih tehnika kvantizacije i transformacionog kodovanja, jer se postiže visok kvalitet rekonstrukcije govornog signala na prijemu i na nižim bitskim brzinama.

Predložena tema je aktuelna sa naučnog stanovišta jer sa razvojem komunikaciono-informacionih tehnologija i telekomunikacionih sistema primena predloženih rešenja telekomunikacionih sistema u prenosu govornog signala postaje sve zastupljenija. Rezultati izloženi u ovoj disertaciji mogu imati i praktičnu primenu.

Ključne reči:

Kvantizacija, Logaritamski kvantizeri, Adaptivni kvantizeri, Transformaciono kodovanje, Kodovanje govornog signala, Diskretni govorni signal

Naučna oblast:

Elektrotehničko i računarsko inženjerstvo

**Naučna
disciplina:**

Telekomunikacije

UDK:

((621.391.:004)+519.7):621.394.14

**CERIF
klasifikacija:**

T 121

**Tip licence
kreativne
zajednice:**

CC BY-NC-ND

Mentor and Doctoral dissertation informations

Doctoral Supervisor: dr Zoran Perić, full professor, University of Niš, Faculty of Electronic Engineering

Title of Doctoral Dissertation: Design of logarithmic quantizer for high quality adaptive transform coding of speech signal

Abstract: In this dissertation, the quantizer design is being considered, devices which within telecommunication systems are used for the signal coding. The main goal of this dissertation is logarithmic quantizers design for speech signal coding that will provide high signal quality at the receiving end. The usage of transform coding and adaptation techniques increases coding quality comparing to classical telecommunicational systems suggested by International Telecommunication Union.

Quantization, as one of the most important procedures in signal processing, presents discretization of the signal by amplitude as a basic step in analogue signal digitization. The characteristics of maintaining approximately constant coding quality in a wide range of input signal variances is called robustness, so it is suitable to use the logarithmic quantizers due their ability to satisfy this feature.

The adaptation of the logarithmic quantizers provides higher reconstructed signal quality at the receiving end, in the manner that adaptive quantizers track the changes in signal strength, change the quantization levels to adapt to input signal, thereby ensuring consistently high quality in the wide range of the input signal variances.

The usage of transform coding methods in coding scheme for speech processing additionally increases quality of the reconstructed signal at the receiver. Transform coding involves association of adjacent signal samples, and the application of certain transformation over them, which indicated redistributing energy

(information). In this dissertation, in the order of coding schemes design, will be used some of the best known transformations like Hadamard transform, discrete cosine transform and discrete wavelet transform.

By leading the discrete signal to the input of a coding system, the compression is further enhanced, since the discrete signal is limited by amplitude and there is no overload distortion, which ensures high quality coding with lower bit rate. For this reason, dissertation proposes coding schemes for discrete input signal. The gain of the solutions that are proposed in this dissertaiton is clearly demonstrated by combining these quantization techinques and transform coding which results in high coding quality of speech signal at the receiving end even on the lower bit rates.

The subject of this dissertation is a very popular in current scientific research directions, because with the development of communication-informational technologies and telecommunication systems, the application of the proposed solutions in the speech signal coding and transmission systems significantly increases. Results presented in this thesis may have many practical applications.

Key Words: Quantization, Logarithmic quantizers, Adaptive quantzers, Transform coding, Speech signal coding, Discrete speech signal

Scientific Field: Electrical and Computer Engineering

Scientific Telecommunications

Discipline:

Key Words: Quantization, Logarithmic quantizers, Adaptive quantzers, Transform coding, Speech signal coding, Discrete speech signal

UDC: ((621.391.:004)+519.7):621.394.14

CERIF
Classification:

T 121

Creative
Commons
Licence Type:

CC BY-NC-ND

Sadržaj

Spisak slika.....	i
1 Uvod.....	13
2 Tehnike kodovanja govornog signala	19
2.1 Kodovanje govornog signala	19
2.2 Kvantizacija i tipovi kvantizera	23
2.2.1 Uniformni skalarni kvantizer.....	26
2.2.2 Skalarni neuniformni kvantizer	29
2.2.3 Kompanding tehnika	30
2.2.4 Logaritamski kvantizeri.....	31
2.3 Tehnike adaptacije kvantizera na ulazni signal	36
2.3.1 Adaptivni kvantizeri sa adaptacijom unazad	37
2.3.2 Adaptivni kvantizeri sa adaptacijom unapred	38
2.4 Optimizacija faktora kompresije kvazi-logaritamskog kvantizera	40
3 Tehnike transformacionog kodovanja govornog signala.....	46
3.1 Diskretna Hadamarova transformacija	47
3.2 Diskretna kosinusna transformacija - DCT.....	51
3.2.1 DCT-I	53
3.2.2 DCT-II	54
3.2.3 DCT-III.....	54
3.2.4 DCT-IV.....	54
3.2.5 Inverzne transformacije	55

3.3	Diskretna wavelet transformacija (DWT).....	55
3.3.1	Implementacija wavelet-a korišćenjem banke filtera	58
4	Konstrukcija logaritamskih kvantizera za adaptivno transformaciono kodovanje govornog signala.....	62
4.1	Šeme za kodovanje kontinualnog govornog signala konstruisane koristeći modifikovanu Hadamarovu transformaciju	63
4.2	Šeme za kodovanje kontinualnog govornog signala konstruisane koristeći diskretnu kosinusnu transformaciju – DCT	84
4.3	Šeme za kodovanje diskretnog govornog signala.....	93
5	Zaključak.....	107
6	Literatura.....	110
7	Biografija Autora	118

Spisak slika

SLIKA 2.1.1. GAUSOVA RASPODELA GUSTINE VEROVATNOĆE.....	20
SLIKA 2.1.2. LAPLASOVA RASPODELA GUSTINE VEROVATNOĆE.....	20
SLIKA 2.1.3. ODNOS GAUSOVE I LAPLASOVE FUNKCIJE RASPODELE GUSTINE VEROVATNOĆE..	21
SLIKA 2.1.4. PROCES DIGITALIZACIJE SIGNALA.....	22
SLIKA 2.2.1. STRUKTURA KVANTIZERA KAO REDNE VEZE ENKODERA I DEKODERA	23
SLIKA 2.2.2. PRIKAZ REGIONA (OBLASTI) KVANTIZERA	24
SLIKA 2.2.3. SKALARNI UNIFORMNI KVANTIZER	27
SLIKA 2.2.4. SKALARNI UNIFORMNI KVANTIZER PROJEKTOVAN ZA RAZNE BITSKE BRZINE	28
SLIKA 2.2.5. SKALARNI UNIFORMNI KVANTIZER PROJEKTOVAN ZA RAZNE MAKSIMALNE AMPLITUDE OPRETEĆENJA.....	28
SLIKA 2.2.6. SKALARNI NEUNIFORMNI KVANTIZER.....	29
SLIKA 2.2.7. KOMPANDING TEHNIKA	30
SLIKA 2.2.8. SEMILOGARITAMSKA KOMPRESORSKA FUNKCIJA	33
SLIKA 2.2.9. SQNR [dB] ZA SIGNAL MODELOVAN LAPLASOVOM RASPODELOM U ŠIROKOM OPSEGU ULAZNOG SIGNALA ZA RAZNE VREDNOSTI FAKTORA KOMPRESIJE μ	35
SLIKA 2.3.1. ADAPTIVNI KVANTIZER SA ADAPTACIJOM UNAZAD	37
SLIKA 2.3.2. ADAPTIVNI KVANTIZER SA ADAPTACIJOM UNAPRED (KODER (GORE) I DEKODER (DOLE))	39
SLIKA 2.4.1. SQNR U USKOM OPSEGU ULAZNIH VARIJANSI [-3 dB, 3 dB] ZA $N=256$, $x_{MAX}=9.12$ I RAZNE VREDNOSTI FAKTORA KOMPRESIJE μ	45
SLIKA 2.4.2. SQNR U USKOM OPSEGU ULAZNIH VARIJANSI [-4 dB, 2 dB] ZA $N=256$, $x_{MAX}=9.12$ I RAZNE VREDNOSTI FAKTORA KOMPRESIJE μ	45
SLIKA 3.1.1. MODIFIKACIJA HADAMAROVE TRANSFORMACIJE H_2 U DVOSTRUKU TRANSFORMACIJU H_1	50
SLIKA 3.2.1. ULAZNI I IZLAZNI PARAMETRI DCT TRANSFORMACIJE	52
SLIKA 3.2.2. SLIKOVIT PRIKAZ BAZNH FUNKCIJA DCT TRANSFORMACIJE	52
SLIKA 3.3.1. A) PRIMER JEDNOG WAVELET-A B) PRIMER EKG SIGNALA SA DVA WAVELET-A	55
SLIKA 3.3.2. PROSTA FUNCKIJA $F(x)$	56
SLIKA 3.3.3. SKALIRANA FUNCKIJA $F(x/0.5)$	57
SLIKA 3.3.4. TRANSLIRANA SKALIRANA FUNCKIJA $F((x-1)/0.5)$	58

SLIKA 3.3.5. IDEALNI A) NISKO FREKVENCIJSKI FILTAR B) VISOKO FREKVENCIJSKI FILTAR.....	61
SLIKA 4.1.1. ŠEMA ZA KODOVANJE GOVORNOG SIGNALA IMPLEMENTIRANA KORISTEĆI PROSTU TRANSFORMACIJU I ADAPTACIJU UNAPRED	64
SLIKA 4.1.2. TEHNIKA ADAPTACIJE UNAPRED U PREDAJNIKU	65
SLIKA 4.1.3. ŠEMA ZA KODOVANJE GOVORNOG SIGNALA IMPLEMENTIRANA KORISTEĆI DVOSTRUKU PROSTU TRANSFORMACIJU I ADAPTACIJU UNAPRED	70
SLIKA 4.1.4. ŠEMA ZA KODOVANJE GOVORNOG SIGNALA IMPLEMENTIRANA KORISTEĆI DVOSTRUKU PROSTU TRANSFORMACIJU NA SAMO JEDNOJ GRANI I ADAPTACIJU UNAPRED 74	
SLIKA 4.1.5. ŠEMA ZA KODOVANJE GOVORNOG SIGNALA IMPLEMENTIRANA KORISTEĆI PROSTU TRANSFORMACIJU I G.711 KODERE U GRANAMA	77
SLIKA 4.1.6. SQNR ZA G.711 STANDARD ($R=6$ BITA PO ODMERKU) I PREDLOŽENE KODNE ŠEME SA SLIKE 4.1.5 ($R_1=8$ BITA PO ODMERKU, $R_2=4$ BITA PO ODMERKU)	80
SLIKA 4.1.7. SQNR ZA G.711 STANDARD ($R=7$ BITA PO ODMERKU) I PREDLOŽENE KODNE ŠEME SA SLIKE 4.1.5 ($R_1=8$ BITA PO ODMERKU, $R_2=6$ BITA PO ODMERKU)	80
SLIKA 4.2.1. ŠEMA ZA KODOVANJE GOVORNOG SIGNALA IMPLEMENTIRANA KORISTEĆI DISKRETNU KOSINUSNU TRANSFORMACIJU 3×3 I ADAPTACIJU UNAPRED	84
SLIKA 4.2.2. ŠEMA ZA KODOVANJE GOVORNOG SIGNALA IMPLEMENTIRANA KORISTEĆI DISKRETNU KOSINUSNU TRANSFORMACIJU 4×4 I ADAPTACIJU UNAPRED	89
SLIKA 4.3.1. ŠEMA ZA KODOVANJE DISKRETNOG GOVORNOG SIGNALA IMPLEMNTRANA KORIŠĆENJEM JEDNOG KVAZI-LOGARITAMSKOG KVANTIZERA	94
SLIKA 4.3.2. ALGORITAM OBRADJE SIGNALA ZA PREDLOŽENU KODNU ŠEMU SA SLIKE 4.3.1	96
SLIKA 4.3.3. SQNR U ŠIROKOM OPSEGU ULAZNIH VARIJANSI ZA $\mu=20$ A) EKSPERIMENTALNI REZULTATI B) TEORIJSKI REZULTATI.....	98
SLIKA 4.3.4. SQNR U ŠIROKOM OPSEGU ULAZNIH VARIJANSI ZA $\mu=255$ A) EKSPERIMENTALNI REZULTATI B) TEORIJSKI REZULTATI.....	99
SLIKA 4.3.5. ŠEMA ZA KODOVANJE DISKRETNOG GOVORNOG SIGNALA KOJA KORISTI TRANSFORMACIONO KODOVANJE I TEHNIKU ADAPTACIJE UNAPRED.....	101
SLIKA 4.3.6. SQNR U ŠIROKOM OPSEGU ULAZNIH VARIJANSI ZA $R_{SR}=4,5,6$ BITA PO ODMERKU, RAZNE VREDNOSTI μ I R_3 ($R_3=4$ - CRVENA LINIJA, $R_3=5$ - ZELENA LINIJA).....	103
SLIKA 4.3.7. SQNR U ŠIROKOM OPSEGU ULAZNIH VARIJANSI ZA $R_{SR} \approx 4$ BITA PO ODMERKU I RAZNE VREDNOSTI μ	105
SLIKA 4.3.8. SQNR U ŠIROKOM OPSEGU ULAZNIH VARIJANSI ZA $R_{SR} \approx 5$ BITA PO ODMERKU I RAZNE VREDNOSTI μ	106

1 Uvod

U ovoj disertaciji razmatrane su performanse logaritamskih kvantizera uključenih u šeme za kodovanje govornog signala. Cilj analize performansi logaritamskih kvantizera (adaptivnih i neadaptivnih) u kodnim šemama sa implementiranim transformacionim kodovanjem je poboljšanje pouzdanosti sistema i povećanje kvaliteta rekonstruisanog signala na prijemnoj strani bez povećanja broja bita po odmerku potrebnih za kodovanje.

Kodovanje govornog signala je proces predstavljanja odmeraka govora u pogodnoj formi sa što manje kodnih simbola ali bez degradacije kvaliteta u cilju efikasnijeg prenosa digitalnim putem kroz žičane i bežične kanale, kao i njegovo skladištenje [1] - [11]. U savremenom društvu kodovanje govora ima veoma značajnu ulogu u telekomunikacionim i multimedijalnim sistemima. Komercijalni sistemi bazirani na efikasnom kodovanju govornog signala uključuju javnu digitalnu i mobilnu telefoniju, satelitske i vojne komunikacione sisteme, audiotehniku, sisteme za prenos signala preko interneta (*Voice over Internet Protocol* - VoIP), videokonferencije, kao i razne multimedijalne aplikacije. Navedene oblasti primene predstavljaju razlog zbog koga se značajna pažnja pridaje istraživanjima u ovoj oblasti [1] - [4], [6] - [8], [9], [12] - [17].

Jedan od osnovnih koraka pri kodovanju govornog signala je kvantizacija. U obradi digitalnih signala kvantizacija se odnosi na proces predstavljanja kontinualne vrednosti amplitude ograničenim setom diskretnih vrednosti [1] - [3], [5], [7], [8], [18] - [28]. Kontinualna amplituda uzoraka mapirana je na n reprezentacionih nivoa [1] - [3], [5], [7], [8], [18] - [28]. Uređaj ili algoritamska funkcija koji izvodi kvantizaciju predstavlja kvantizer [1] - [8]. Kvantovanjem n bitova, raspon amplitude signala je podeljen u 2^n diskretnih nivoa, a svaki uzorak je kvantovan (zaokružen) na najbliži reprezentacioni nivo. Ovo zaokruživanje na

najbliži nivo dovodi do greške koja predstavlja najveći izvor smetnje u digitalnom procesuiranju signala [2], [3], [8], [11], [17], [20], [24], [29] – [32]. U pretvaranju analognih signala u digitalne, razlika između stvarne analogne vrednosti i kvantovane digitalne vrednosti naziva se greškom kvantizacije. Signal koji je nastao zbog ove greške je ponekad modelovan kao dodatni slučajni signal zbog stohastičkog ponašanja. Kvalitet kodovanja raste povećanjem bitske brzine a samim tim i broja reprezentacionih nivoa na koje se analogni odmerci zaokružuju. Međutim, cilj ove disertacije predstavlja projektovanje kvantizera koji će sa što manjom bitskom brzinom obezbediti visok kvalitet kodovanja govornog signala. Iz tog razloga ova disertacija predlaže različite kodne šeme koje za cilj imaju obezbeđivanje što višeg kvaliteta upotrebom što manje bitova po odmerku potrebnih za kodovanje [33] - [34].

Osnovna podela kvantizera vrši se prema broju odmeraka koji se kvantuju u jednom trenutku vremena. Na ovaj način razlikuju se skalarni i vektorski kvantizeri. Skalarni kvantizeri kvantuju svaki odmerak ulaznog signala odvojeno, dok se kod vektorskih m uzastopnih odmeraka ulaznog signala udružuje (formira se m -dimenzionalni vektor) i nad njima se vrši združena kvantizacija [5]. U ovoj disertaciji akcenat je stavljen na konstrukciju skalarnih kvantizera i njihovu primenu u kodnim šemama koje će biti predstavljene, tako da će u nastavku biti više reči o skalarnoj kvantizaciji i skalarnim kvantizerima [2], [3], [7], [8], [11], [17], [22] – [23], [30] – [32], [35].

Skalarni kvantizer sa N kvantizacionih nivoa i ulazni signal, opisan odgovarajućom gustinom raspodele verovatnoće, mogu se predstaviti matematičkim modelom na sledeći način. Amplitudski opseg kvantizera deli se na N ćelija (kvantizacionih intervala) unutar kojih se bira jedna reprezentaciona tačka, čime se definiše kodna knjiga kvantizera. Pragovi odluke (granice između kvantizacionih intervala) i reprezentacioni nivoi definišu skalarni kvantizer [1] – [4], [7], [8], [11]. Skalarna kvantizacija vrši se određivanjem kvantizacionog intervala kome pripada ulazni odmerak (jedan za skalarnu kvantizaciju), nakon čega se taj odmerak zaokružuje na reprezentacioni nivo tog intervala. Mapiranjem originalne (realne) vrednosti ulaznog odmerka na vrednost reprezentacionog nivoa pravi se greška koju nije moguće ispraviti nekim dodatnim procesom i ta greška naziva se greškom kvantizacije (šum kvantizacije) ili distorzijom [1] – [4], [7], [8], [11]. Da bi se ta greška mogla kvantitativno proceniti potrebno je definisati objektivnu meru na osnovu koje se može odrediti ukupna degradacija kvaliteta signala. Kao jedna od brojnih mera distorzije najčešće se koristi kvadrat razlike realnog odmerka i zaokružene (mapirane) vrednosti na reprezentacioni nivo kvantizacionog intervala kome odmerak pripada [1] - [11], [17] - [21], [23] - [24], [27] - [35]. Iz gore navedenog može se zaključiti da je cilj prilikom projektovanja kvantizera za kodovanje ulaznog signala,

opisanog određenom gustinom raspodele verovatnoće, određivanje optimalnih vrednosti pragova odluke i reprezentacionih nivoa kako bi se obezbedilo da distorzija bude minimalna, odnosno kvalitet kvantizacije maksimalan. Kao objektivna mera kvaliteta kvantizacije u ovoj disertaciji biće korišćen odnos signala i kvantizacionog šuma (*Signal to Quantization Noise Ratio*) SQNR [1] - [11], [17] - [21], [23] - [25], [27] - [37]. SQNR je mera koja predstavlja logaritamski odnos snage signala i snage kvantizacionog šuma i data je u decibelima.

Skalarni kvantizeri mogu biti uniformni i neuniformni. Uniformni kvantizeri su takvi da su njihovi pragovi odlučivanja i reprezentacioni nivoi na međusobno jednakim rastojanjima, pri čemu su reprezentacioni nivoi na sredini kvantizacionog intervala. Uniformni kvantizeri su jednostavniji za realizaciju i pogodni su za korišćenje kada imamo ulazni signal sa uniformnom gustinom raspodele verovatnoće [1] – [3], [5] – [11]. Međutim, za signale sa neuniformnom raspodelom (između ostalih i govorni signal) uniformni kvantizeri nisu optimalni i ne obezbeđuju dobar kvalitet kvantizacije. Za takve signale pogodno je koristiti neuniformne kvantizere. Kod neuniformnih kvantizera kvantizacioni intervali i reprezentacioni nivoi nisu na međusobno jednakim rastojanjima, dok reprezentacioni nivoi nisu na sredini kvantizacionih intervala. Rastojanje između nivoa kod ovih kvantizera malo je na mestima gde funkcija gustine raspodele verovatnoće ima veliku vrednost, dok u oblastima gde gustina raspodele uzima male vrednosti, rastojanje je veliko [1] – [3], [5] – [11]. Takav je slučaj i sa govornim signalom. U oblastima malo verovatnih signala, kvantizacioni nivoi su retko raspoređeni, samim tim su i veće širine (odmerci ulaznog signala su grubo kvantovani), dok su u oblastima visoko verovatnih signala nivoi gusti raspoređeni, dosta uski, i tu su odmerci finije kvantovani. Na ovaj način se obezbeđuje viši kvalitet kodovanja, odnosno manja distorzija i veći SQNR. Za projektovanje neuniformnih kvantizera koriste se dve tehnike: Iterativni Lojd-Maks algoritam i kompanding tehnika [1] – [3], [5] – [11], [37], [38] – [43]. Iterativni Lojd-Maks algoritam se, usled svoje velike kompleksnosti zasnovane na ponovnom proračunu pragova odluka i reprezentacionih nivoa kroz svaku iteraciju algoritma, najčešće koristi pri konstrukciji kvantizera za male bitske brzine, dok se za veće brzine koristi kompanding tehnika kod koje se neuniformni kvantizer realizuje kao redna veza kompresora, uniformnog kvantizera i ekspandora o kojoj će više reči biti u narednom poglavlju disertacije [1] – [3], [5] – [11], [18] – [24], [27], [38], [44] – [52].

Mnogi realni signali, poput govornog signala, su nestacionarni, što znači da se njihova snaga menja u vremenu, te je za kvantizaciju takvih signala potrebno konstruisati kvantizer koji će održavati približno konstantan kvalitet za ceo opseg. U cilju ispunjenja ovog zahteva pogodno je koristiti logaritamske kompanding kvantizere koji poseduju određeni stepen

robustnosti i time delimično obezbeđuju konstantan kvalitet kvantizacije. Kako je zahtev savremenih tehnologija da se kvalitet kodovanja održava konstantnim u velikoj meri u vremenu potrebno je projektovanje kvantizera koji bi se prilagođavao snazi signala. Kvantizeri koji se mogu prilagođavati promenama snage signala nazivaju se adaptivnim kvantizerima. Njihova konstrukcija je takva da prati promene signala u vremenu i menja (adaptira) svoje pragove odlučivanja i reprezentacione nivoe prilagođavajući se signalu čime se obezbeđuje konstantan kvalitet kvantizacije [1] – [3], [5] – [11], [53] - [59]. Tehnike adaptacije kvantizera biće detaljnije obrađene u narednom poglavlju disertacije.

Zahtev za smanjenjem broja bita potrebnih za kvantizaciju nalaže upotrebu tehnika koje mapiraju informacije iz velikog broja odmeraka ulaznog signala na manji broj odmeraka i time pružaju mogućnost uštede u bitskoj brzini. Ovakvo mapiranje korisnih informacija sadržanih u odmercima signala naziva se transformacijom, a kodovanje koje se vrši primenom transformacija naziva se transformacionim kodovanjem [1] – [3], [5] – [11], [59] – [64]. Kod signala sa velikim stepenom korelacije između odmeraka moguće je izvršiti grupisanje uzastopnih odmeraka sa približnim vrednostima u kojima je koncentrisana energija, a samim tim i korisna informacija, i nad njima primeniti neku vrstu transformacije. Time se obezbeđuje da se energija (korisna informacija koju želimo da prenesemo) signala pregrupiše tako da određeni odmerci nose najveću količinu informacija, dok ostali nose malu količinu. Odmerke sa najvećom količinom informacija kodujemo i prenosimo kroz komunikacioni kanal do prijemnika dok odmerke sa malom količinom informacija možemo zanemariti. Na ovaj način smanjujemo broj odmeraka za kvantizaciju a time i broj bitova potrebnih za njihovo kodovanje [1] – [3], [5] – [11], [59] – [64].

Disertacija je podeljena na šest poglavlja koja su dalje podeljena na odeljke. Označavanje formula, slika i tabela u okviru jednog poglavlja izvršeno je tako da poslednji broj ukazuje na redni broj formule, slike ili tabele u odeljku definisanom preostalim brojevima.

U drugom poglavlju disertacije akcenat je na teorijskoj obradi tehnika kodovanja govornog signala, kao i na proučavanju standardizovanih kodnih tehnika predloženih od strane Međunarodne Telekomunikacione Unije (ITU-T). Drugo poglavlje podeljeno je na tri podpoglavlja u kojima je više pažnje posvećeno samom govornom signalu i njegovom kodovanju, kao i detaljnoj obradi skalarne kvantizacije i njene performanse, zatim detaljnijoj obradi komandirajućih tehnika sa njenim pogodnostima i primenom, i na kraju teorijska obrada tehnika transformacionog kodovanja i njegova primena na kodovanje govornog signala.

U trećem poglavlju dat je teorijski osvrt na tehnike transformacionog kodovanja govornog signala. Tačnije, u ovom poglavlju obrađene su tri tehnike transformacionog kodovanja, i predstavljena je njihova primena u kodnim šemama. Obradene transformacione tehnike su: Hadamarova transformacija sa matricama transformacionih koeficijenata 2×2 i 4×4 [1] – [3], [5] – [8], [59], diskretna kosinusna transformacija (*Discrete Cosine Transform* - DCT) sa matricama transformacionih koeficijenata 3×3 i 4×4 i diskretna wavelet transformacija (*discrete Wavelet transform* - DWT) sa matricom 2×2 koja se u stvari svodi na Hadamarovu transformaciju sa istom veličinom matrice. Kvantizeri korišćeni u predloženim kodnim šemama su logaritamski kompanding kvantizeri bazirani na μ logaritamskom zakonu i optimizovanim parametrima za region rada kvantizera i faktor kompresije [1] – [3], [5] – [11], [65] – [71].

Glavni doprinos disertacije predstavljen je u poglavlju četiri. Samo poglavlje podeljeno je na tri podpoglavlja od kojih su u prvom prikazane kodne šeme bazirane na upotrebi Hadamarove transformacije sa matricama koeficijenata 2×2 i 4×4 , u drugom kodne šeme za čiju se implementaciju koristi diskretna kosinusna transformacija 3×3 i 4×4 , dok je u trećem podpoglavlju akcentat stavljen na projektovanje kodnih šema za diskretni ulazni signal.

Prvo podpoglavlje predstavlja četiri kodne šeme implementirane primenom proste (modifikovane) Hadamarove transformacije pogodne za kodovanje govornog signala. Prva šema implementirana je jednostrukom upotrebom H_1 i tehnike adaptacije unapred [59]. Druga predstavljena kodna šema strukturalno je slična standardu G.711.1 s razlikom u upotrebi dva G.711 koda u granama nakon upotrebe transformacija u odnosu na G.711.1 standard gde se u jednoj grani koristi vektorska kvantizacija, čime se u velikoj meri uprošćava kompleksnost implementacije kao i smanjuje kašnjenje pri kodovanju. Takođe je opisan koncept segmentne uniformne kvantizacije primenjene u predloženoj kodnoj šemi [71]. Prikazani su benefiti korišćenja ove kodne šeme kroz objektivnu meru kvaliteta SQNR i uporedivu analizu sa standardom G.711.1 koji uključuje vektorsku kvantizaciju. Za potrebe eksperimenta i projektovanja kodne šeme korišćeni su govorni signali snimljeni u laboratoriji Elektronskog fakulteta u Nišu na katedri za Telekomunikacije, odmereni na 8 kHz i 16 kHz (širokopojasna ekstenzija). Još jedan važan doprinos disertacije koji je objašnjen i prikazan u ovom poglavlju jeste novi model objektivnog merenja kvaliteta kodovanja govornog signala [45], [46], [61] – [64], [71] – [73]. Ostale dve predložene šeme za kodovanje govornog signala bazirane su na dvostrukoj upotrebi proste (modifikovane) Hadamarove transformacije H_1 . Upotrebom dvostruke transformacije dobijaju se četiri grane transformisanog signala koje se koduju nezavisno. U jednoj od prikazanih šema koduju se signali iz svih grana, dok se u drugoj ne

primenjuje još jedna transformacija nad signalom sa malom korelacijom dobijenom posle prve transformacije, već se on koduje takav kakav je jer nosi malu količinu informacija u sebi zbog male korelacije pa dodatna transformacija ne donosi dobitak u kodovanju.

Drugo podpoglavlje četvrtog poglavlja predlaže kodne šeme bazirane na upotrebi diskretne kosinusne transformacije DCT i tehnike adaptacije unapred. Prva predložena kodna šema bazirana je na upotrebi DCT-a sa matricom koeficijenata veličine 3×3 dok druga koristi matricu DCT-a veličine 4×4 . Rezultati kvaliteta kodovanja dobijeni upotrebom ovih šema pokazuju pogodnost korišćenja za kodovanje govornog signala.

Treće podpoglavlje četvrtog poglavlja predlaže kodne šeme projektovane za diskretni ulazni signal [74]. Data je teorijska i eksperimentalna obrada kodnih šema projektovanih za diskretni ulazni signal. Predložene kodne šeme predstavljaju model za kodovanje govornog signala koji se sastoji iz dva nivoa. Prvi nivo predstavlja konverziju iz analognog domena u diskretni (za ovu svrhu je korišćen logaritamski kvantizer sa velikim brojem nivoa), a drugi nivo predstavlja sistem sa kodovanje tako diskretizovanog govornog signala. Prva od predloženih kodnih šema u ovom poglavlju predstavlja drugi nivo sistema za kodovanje diskretizovanog signala kao jedan logaritamski kvantizer [74] – [76], dok druga predložena kodna šema drugi nivo predstavlja kao sistem za kodovanje diskretnog govornog signala koji uključuje transformaciono kodovanje i adaptacione tehnike. U oba slučaja optimizovani su parametri kvantizera uključenih u kodne šeme i to region rada kvantizera i faktor kompresije. Kroz objektivnu meru kvaliteta, SQNR, prikazani su rezultati i obrazložena je pogodnost korišćenja predloženih kodnih šema za kodovanje diskretnog govornog signala.

Peto poglavlje, zaključak, sumira rezultate i doprinose disertacije. Ovde su eksplicitno navedeni svi zaključci disertacije i na osnovu analize ovih zaključaka prikazani su dalji pravci istraživanja u ovoj oblasti. Na kraju je prikazan spisak literature korišćene pri izradi ove disertacije. Veliki deo dobijenih rezultata već je prikazan u naučnim radovima objavljenim u naučnim časopisima i zbornicima radova sa domaćih i međunarodnih konferencija.

2 Tehnike kodovanja govornog signala

Porast u korišćenju komunikaciono-informacionih tehnologija poslednjih godina uslovio je ekspanzivni razvoj tehnika kodovanja signala koje obezbeđuju pouzdan i kvalitetan prenos signala sa predajne na prijemnu stranu. Sve do sada razvijane tehnike bazirane su na ideji da se sa što manjim bojem bita potrebnih za kodovanje signala obezbedi što viši kvalitet rekonstruisanog signala na prijemu [1] – [3], [5] – [11], [50] – [59], [66] – [70], [77] – [83]. U ovoj disertaciji pažnja se polaže na kompanding tehniku za kodovanje, kao i na razne tehnike transformacionog kodovanja o kojima će više reči biti u narednim odeljcima ovog poglavlja. Teorijskim sagledavanjem kodnih tehnika primenjenim u kodnim šemama koje su predložene u ovoj disertaciji izvode se zaključci o pogodnosti njihove primene u kodovanju govornog signala. U prvom odeljku ovog poglavlja objašnjena je priroda govornog signala, tipovi funkcija raspodele gustine verovatnoće, dok je detaljno opisan proces kodovanja govornog signala. Drugi odeljak predstavlja opšte pojmove vezane za kvantizaciju i klasifikaciju kvantizera. Ovde su detaljno predstavljeni uniformni kvantizer, neuniformni, kompanding tehnika i logaritamski kvantizeri koji se koriste u ovoj disertaciji .

2.1 Kodovanje govornog signala

Govorni signal, kao jedan od najvažnijih signala u prirodi, predstavlja zvučnu interpretaciju načina komunikacije između ljudi, a u poslednje vreme puno pažnje posvećuje se i komunikaciji ljudi i mašina putem govornog signala. Upravo je ovo uzrok istraživanja u polju kodovanja i kompresije govornog signala koja imaju za cilj obezbeđivanje bezbednog prenosa i skladištenja govornog signala [1] – [11].

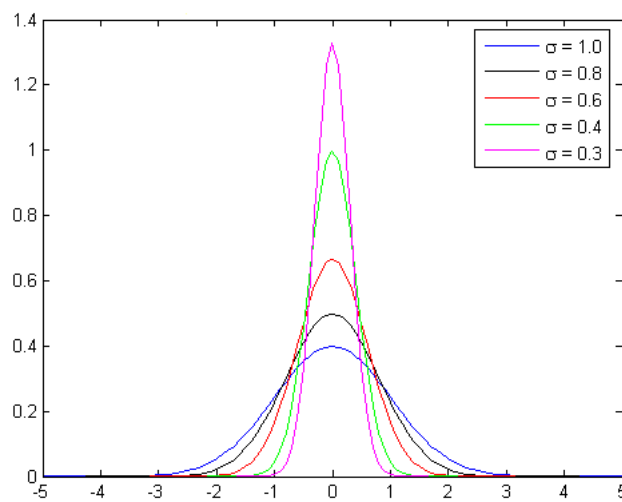
Kako bi bio predstavljen matematički, govorni signal je neophodno modelovati nekom funkcijom raspodele gustine verovatnoće $p(x)$, gde x predstavlja trenutnu vrednost signala. U praksi su najkorišćenije Gausova i Laplasova funkcija raspodele gustine verovatnoće. Gausova funkcija se matematički može predstaviti na sledeći način [1] – [3], [5] – [11]:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \quad (2.1.1)$$

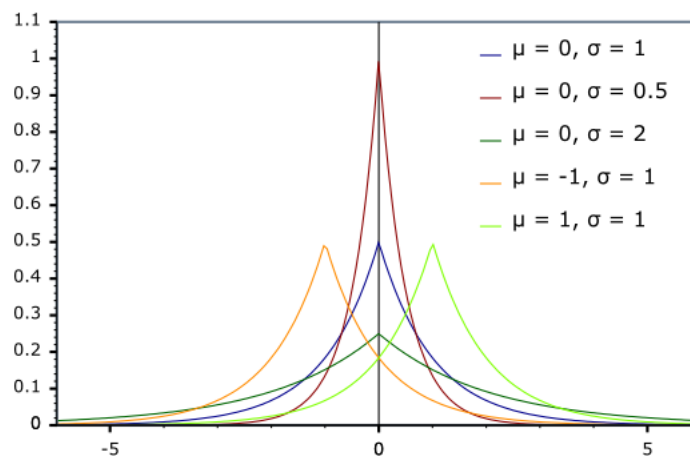
dok se Laplasova raspodela matematički opisuje:

$$p(x) = \frac{1}{\sqrt{2}\sigma} e^{-\frac{|x|\sqrt{2}}{\sigma}} \quad (2.1.2)$$

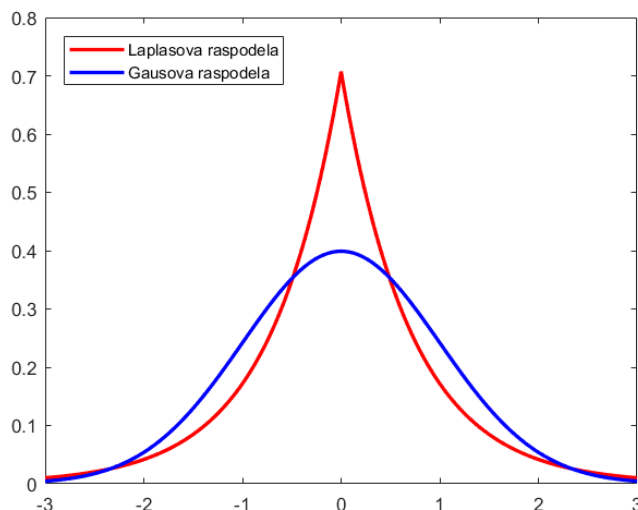
gde parametar σ predstavlja standardnu devijaciju signala, dok σ^2 predstavlja varijansu signala. Izrazi 2.1 i 2.2 opisuju funkcije raspodele verovatnoće sa nultom srednjom vrednošću i za signale sa takvom srednjom vrednošću varijansa σ^2 ujedno predstavlja i snagu signala. Grafički prikaz Gausove funkcije sa nultom srednjom vrednošću i raznim vrednostima varijanse data je na slici 2.1.1 dok je Laplasova raspodela verovatnoće sa raznom srednjom vrednošću i varijansom prikazana na slici 2.1.2. Slika 2.1.3 prikazuje odnos Laplasove i Gausove funkcije raspodele gustine verovatnoće.



Slika 2.1.1. Gausova raspodela gustine verovatnoće



Slika 2.1.2. Laplasova raspodela gustine verovatnoće

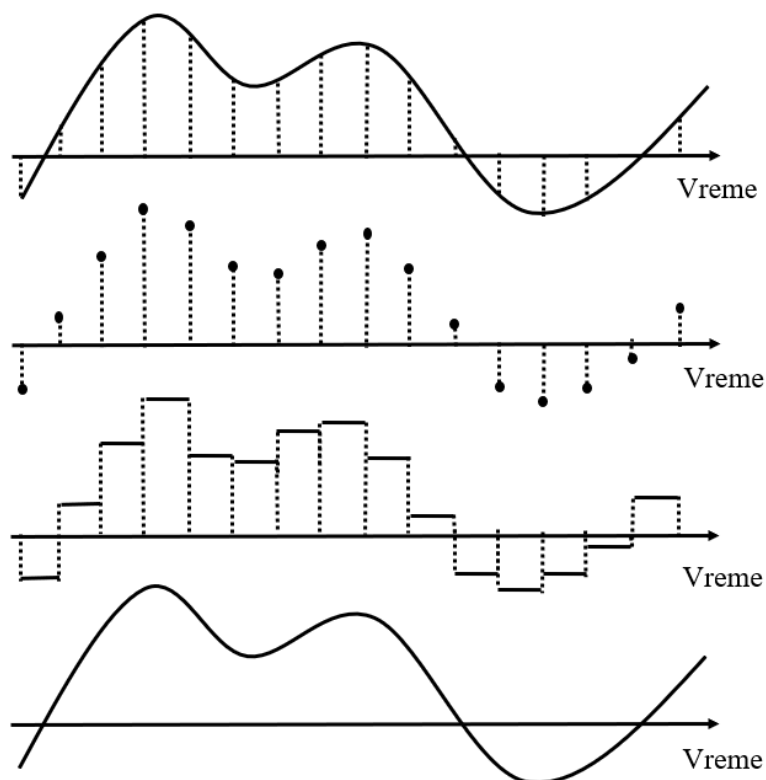


Slika 2.1.3. Odnos Gausove i Laplasove funkcije raspodele gustine verovatnoće

Prilikom modelovanja signala, ukoliko se posmatra kratkoročna statistika signala, pogodno je koristiti Gausovu funkciju raspodele verovatnoće (modelovanje signala sa malim brojem odmeraka), dok u slučaju dugoročne statistike (signal koji se modeluje sadrži veliki broj odmeraka) Laplasova funkcija raspodele verovatnoće pružiće bolji uspeh u modelovanju. Preciznije rečeno, Gausova funkcija raspodele verovatnoće se koristi ukoliko je frekvencija odmeravanja signala f_s do 20 Hz (perioda odmeravanja do 5 ms), dok se u ostalim slučajevima najčešće koristi Laplasova funkcija [77]. U disertaciji u svrhu izvođenja eksperimenata korišćeni su govorni signali snimljeni u laboratoriji Elektronskog fakulteta u Nišu, koji sadrže više miliona odmeraka te je za njihovo modelovanje korišćena Laplasova funkcija.

Govorni signal predstavlja analogni nestacionarni prirodni signal čija se snaga menja kroz vreme tako da je potrebno dobro projektovati kvantizere koji će taj analogni signal promenljive snage predstaviti u digitalnom obliku, pogodnom za skladištenje, kao i prenos kroz komunikacioni kanal. Proces prestavljanja analognih signala u digitalnom domenu naziva se digitalizacija signala ili A/D konverzija (*Analogue/Digital conversion*) [1] – [3], [5] – [11]. Digitalizacija je proces predstavljanja analognih signala nizom brojeva. Ovi brojevi predstavljaju izmerene vrednosti tog signala u sukcesivnim vremenskim intervalima. To podrazumeva da se ne beleže sve vrednosti signala, već samo vrednosti u određenim trenucima u vremenu. Dakle, skup tačaka u kojima se posmatra vrednost signala je diskretan, te se signal dobijen na ovaj način zove diskretni signal. Kako digitalizacija podrazumeva kompjutersku obradu signala koja se svodi na rad sa brojevima koji predstavljaju vrednosti tog signala, time i te vrednosti moraju biti iz diskretnog skupa kako bi računar mogao da ih koristi. Takav signal predstavlja digitalni signal [1] – [3], [5] – [11]. Svođenje neprekidnog, tj. analognog signala na

diskretni signal, odnosno diskretizacija po vremenu naziva se smplovanje (biranje uzoraka ili uzorkovanje), dok se svođenje diskretnog signala na digitalni naziva diskretizacija po amplitudi, tj. kvantovanje ili kvantizacija, koja će detaljno biti opisana u narednom odeljku ovog poglavlja. Dakle, obavlja se proces dvostruke diskretizacije čiji je grafički prikaz dat na slici 2.1.4. Vrednost signala u jednoj tački predstavlja uzorak (sample) [1] – [3], [5], [7], [8] – [11], [18] – [28].



Slika 2.1.4. Proces digitalizacije signala

Nakon procesa diskretizacije signala po amplitudi (kvantizacije), vrši se kodovanje signala. Uređaji koji obavljaju proces kodovanja nazivaju se koderima. Koderi se mogu podeliti u tri grupe u odnosu na način rada i to na:

- Kodere talasnog oblika
- Parametarske kodere
- Hibridne kodere

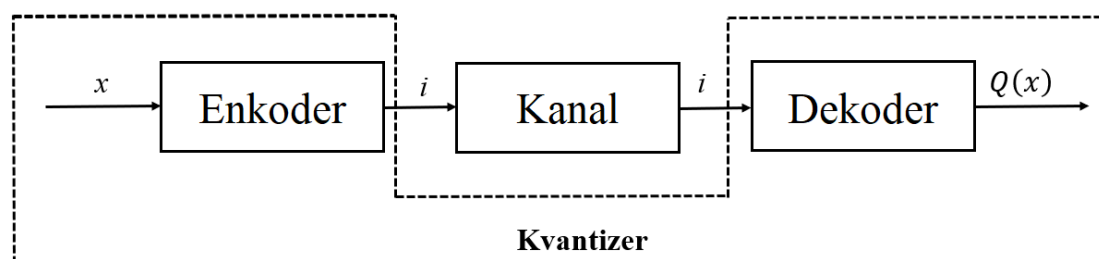
Koderi talasnog oblika rade na taj način što prate i zadržavaju talasni oblik signala na ulazu kodera i u ovu grupu kodera spadaju PCM (*Pulse Code Modulation*) koderi i ADPCM (*Adaptive Differential Pulse Code Modulation*) koderi [1] – [3], [5], [7], [8], [18] – [28], [37], [45], [46]. Parametarski koderi imaju za cilj da prilikom rekonstrukcije kodovanog signala na prijemnoj strani proizvedu razumljiv govorni signal na osnovu određenih parametara koji zavise od samog modela kvantizera. Koderi koji spadaju u grupu parametarskih kodera su LP

(*Linear Prediction*) koderi i harmonijski koderi koji rade na veoma niskim bitskim brzinama [1] – [3], [47], [78] – [81].

Hibridni koderi predstavljaju kombinaciju prednosti koderata talasnog oblika i hibridnih koderata jer rade na taj način što pokušavaju da prate talasni oblik signala koji se koduje ali i parametre modela [1] – [3], [7] – [11], [47], [78] – [81]. Tipičan predstavnik hibridnih koderata je CELP (*Code Excited Linear Prediction*) koder [1] – [3], [7] – [11], [47], [78] – [81].

2.2 Kvantizacija i tipovi kvantizera

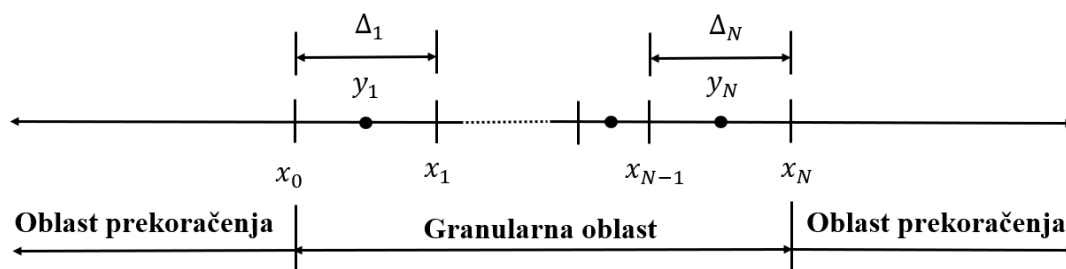
Kao što je pomenuto u prethodnom poglavlju, kvantizacija predstavlja diskretizaciju signala po amplitudi, dok se uređaji koji vrše kvantizaciju nazivaju kvantizerima. Njihova struktura je takva da se sastoje od enkodera na predajnoj i dekodera na prijemnoj strani sistema za kodovanje. Struktura kvantizera prikazana je na slici 2.2.1.



Slika 2.2.1. Struktura kvantizera kao redne veze enkodera i dekodera

Formalno, kvantizacija predstavlja mapiranje Q skupa od R realnih brojeva u skup dozvoljenih izlaznih vrednosti $Y \subset R$, koje predstavljaju skup reprezentacionih nivoa kvantizera, odnosno kodnu knjigu kvantizera [1] – [3], [5] – [11], [18] – [28], [35], [36], [44], [48], [49], [82] – [86]. Detaljnije opisano, kvantizacija predstavlja preslikavanje $Q([x_{i-1}, x_i]) \rightarrow y_j, j=1, \dots, N$, gde je N broj nivoa kvantizera, sa $x_i, i=0, \dots, N$, označeni su pragovi odlučivanja kvantizera, dok su sa y_j označeni reprezentacioni nivoi. Kvantizacioni interval je ograničen sa dva praga odluke x_{i-1} i x_i , i označava se sa Δ_i . Unutar njega se nalazi jedan reprezentacioni nivo na koji se mapiraju svi odmerci koji prilikom procesa kvantizacije upadnu u taj kvantizacioni interval. Amplitudski neograničeni signali imaju odmerke koji ne upadaju u region rada kvantizera, što znači da se može uzeti da je $x_0 \rightarrow -\infty$ a $x_N \rightarrow \infty$, vrednosti koje predstavljaju minimalnu i maksimalnu amplitudu kvantizera pa se zato mogu nazivati i amplitudama maksimalnog opterećenja kvantizera x_{\max} [1] – [3], [5] – [11], [18] – [28], [66] – [69]. Region rada kvantizera čine nivoi koji se nalaze između x_1 i x_{N-1} i taj region se naziva granularni region

(granularna oblast). Nivoi koji se nalaze izvan krajeva regiona rada kvantizera, manji od x_1 i veći od x_N spadaju u region prekoračenja (oblast prekoračenja, *overflow* region) [1] – [3], [5] – [11], [18] – [28], [66] – [69]. Na slici 2.2.2 prikazan je odnos regiona kvantizera.



Slika 2.2.2. Prikaz regiona (oblasti) kvantizera

Iz navedenog se može zaključiti da je izbor maksimalne amplitude kvantizera od ključne važnosti za samu konstrukciju kvantizera. Ukoliko se prilikom projektovanja loše izabere vrednost x_{\max} -a može doći do slučaja da veliki broj odmeraka ispadne iz regiona rada kvantizera u oblast prekoračenja i mapira na poslednji reprezentacioni nivo regiona (najveći ili najmanji), čime se pravi veća greška kvantizacije usled prekoračenja. Sa druge strane, izbor velikih vrednosti x_{\max} -a proširuje granularni region, a time i kvantizacione intervale što mapiranje odmeraka čini netačnim i dovodi se do veće greške kvantizacije u granularnoj oblasti. U procesu kvantizacije signala prilikom mapiranja (zaokruživanja) odmeraka na reprezentacione nivoe kvantizera pravi se nepovratna greška koja se naziva greškom kvantizacije ili šumom kvantizacije. Greška kvantizacije se ne može izbeći u potpunosti, već se samo dobrim izborom maksimalne amplitude kvantizera može smanjiti na prihvatljivu vrednost. Srednja vrednost greške kvantizacije izražava se kroz veličinu koja se naziva distorzija D (ukupna distorzija), koja predstavlja i objektivnu meru kvaliteta kvantizacije. Ukupna distorzija D predstavlja zbir distorzije granularnog regiona i distorzije prekoračenja (*overflow* distorzije) i data je izrazom [1] – [3], [5] – [11], [18] – [28], [66] – [71], [74] – [76], [78], [79]:

$$D = D_g + D_{ov} \quad (2.2.1)$$

Za računanje distorzije potrebno je odrediti rastojanja između realnog odmerka i vrednosti na koju je taj odmerak mapiran (zaokružen) i uvesti određena pravila po kojima se distorzija računa. Standardno se koristi Euklidsko kvadratno rastojanje prema kome distorzija predstavlja srednjekvadratnu grešku kvantizacije (predstavlja snagu šuma kvantizacije) i definisana je opštim izrazom [1] – [3], [5] – [11], [18] – [28]:

$$D = \sum_{i=1}^N \int_{x_{i+1}}^{x_i} (x - y_i)^2 p(x) dx, \quad (2.2.2)$$

gde je N broj reprezentacionih nivoa kvantizera, dok x_i predstavlja pragove odluke kvantizacionih intervala.

Na osnovu izraza 2.1.3, distorzija D predstavlja zbir D_g i D_{ov} koje su definisane izrazima:

$$D_g = \sum_{i=2}^{N-1} \int_{x_{i+1}}^{x_i} (x - y_i)^2 p(x) dx, \quad (2.2.3)$$

$$D_{ov} = 2 \int_{x_{N-1}}^{\infty} (x - y_N)^2 p(x) dx. \quad (2.2.4)$$

Ukoliko je funkcija $p(x)$ parna, dobija se simetrični kvantizer što podrazumeva da su mu pozitivni i negativni reprezentacioni nivoi međusobno simetrični. Sa druge strane, ukoliko je funkcija $p(x)$ ograničena, dobija se slučaj u kome nema distorzije prekoračenja pa se optimalnim izborom maksimalne amplitude kvantizera (njegovim izjednačavanjem sa maksimalnom vrednošću amplitude signala) obezbeđuje maksimalni kvalitet kodovanja. Za neograničene funkcije $p(x)$ uvek će postojati distorzija prekoračenja koja se ne može neutralisati već se samo smanjuje optimalnim izborom amplitude maksimalnog opterećenja kvantizera [1] – [3], [5] – [11], [18] – [28]:

Distorzija je samo jedna od mera kvaliteta jer pored nje performanse primenjene tehnike kvantizacije mogu se izraziti i objektivnom veličinom koja se naziva SQNR (*Signal-to-Quantization-Noise-Ratio*). Ova veličina predstavlja logaritamski odnos snage signala i kvantizacionog šuma i izražava se u decibelima (dB) [1] – [3], [5] – [11], [18] – [28], [66] – [71], [74] – [76], [78], [79], [82] – [86]:

$$SQNR = 10 \log_{10} \left(\frac{\sigma^2}{D} \right), \quad (2.2.5)$$

gde je σ^2 varijansa signala.

Postoje dva načina za klasifikaciju kvantizera: prema načinu projektovanja i prema fizičkim karakteristikama. Prema načinu projektovanja kvantizeri se dele na kvantizere projektovane minimizacijom distorzije i kvantizere ograničene entropijom. Projektovanje kvantizera minimizacijom distorzije ima za cilj postizanje maksimalnog kvaliteta kodovanja - SQNR-a (tj. minimalne distorzije) za zadati broj kvantizacionih nivoa N kodovanjem izlaznih nivoa kvantizera kodnim rečima fiksne dužine. Kod entropijom ograničenih kvantizera, izlazni nivoi kvantizera se koduju kodnim rečima promenljive dužine i cilj projektovanja nije samo postizanje maksimalnog SQNR-a, već i minimizacija bitske brzine R , odnosno broja nivoa

kvantizera potrebnih za kvantovanje [1] – [3], [5] – [11], [18] – [28], [66] – [71], [74] – [76], [78], [79], [82] – [86].

Fizičke karakteristike kvantizere dele na vektorske i skalarne. Kod vektorskih kvantizera grupa odmeraka signala formira niz (vektor) koji se ceo propušta kroz kvantizer (kvantuje se), dok se kod skalarnih svaki odmerak ulaznog signala kvantuje pojedinačno. Tabela 2.2.1 prikazuje dalju podelu kvantizera prema fizičkim karakteristikama [1], [7].

Tabela 2.2.1 Podela kvantizera prema fizičkim karakteristikama

Kvantizeri				
Vektorski	Skalarni			
	Uniformni	Neuniformni		
		Lojd-Maks	Kompadding	
			Optimalni	Logaritamski
				A-zakon μ -zakon

Kako je u ovoj disertaciji akcenat stavljen na projektovanje logaritamskih kvantizera baziranih na μ logaritamskom zakonu koji spadaju u skalarne, u nastavku će biti predstavljen detaljan opis određenih grupa skalarnih kvantizera.

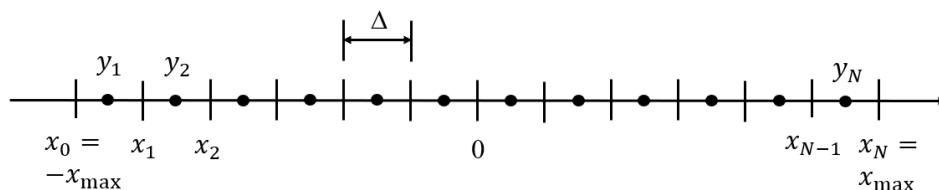
2.2.1 Uniformni skalarni kvantizer

Skalarni uniformni kvantizeri predstavljaju najprostiju grupu skalarnih kvantizera zato što se kod njih pragovi odluke i reprezentacioni nivoi nalaze na jednakom međusobnom rastojanju, koje se naziva korakom kvantizacije Δ . Još jedna specifičnost za uniformne kvantizere je to da se reprezentacioni nivoi nalaze na sredini kvantizacionih intervala, osim u dva spoljna intervala za koje važi $x_0 \rightarrow -\infty, x_N \rightarrow \infty$. Za projektovanje uniformnih kvantizera potrebno je znati samo maksimalnu amplitudu, x_{\max} , i broj nivoa kvantizera N . Samim tim imamo i informaciju o koraku kvantizacije, koji se računa kao $\Delta = (2 \cdot x_{\max})/N$. Pragovi odluke i reprezentacioni nivoi se određuju korišćenjem sledećih izraza:

$$x_i = -x_{\max} + i\Delta, \quad i = 0, 1, \dots, N \quad (2.2.6)$$

$$y_i = -x_{\max} + (i - 0.5)\Delta, \quad i = 1, \dots, N \quad (2.2.7)$$

Na slici 2.2.3 prikazan je uniformni skalarni kvantizer.



Slika 2.2.3. Skalarni uniformni kvantizer

Uniformni kvantizer je pogodno koristiti za signale sa uniformnom raspodelom, međutim, vrlo često se koristi i za signale sa neuniformnom raspodelom (ograničenom ili neograničenom). U slučaju amplitudski ograničenih signala sa neuniformnom raspodelom, za konstrukciju uniformnog kvantizera potrebno je postaviti vrednost maksimalne amplitude signala za maksimalnu amplitudu kvantizera čime se obezbeđuje da distorzija prekoračenja ne postoji, tj bude jednaka nuli. Granularna distorzija u ovom slučaju predstavlja ukupnu distorziju prilikom procesa kvantizacije i dobija se pomoću izraza:

$$D = D_g = \sum_{i=1}^N \int_{x_{i-1}}^{x_i} (x - y_i)^2 p(x) dx \quad (2.2.8)$$

U slučaju konstrukcije kvantizera za veliki broj kvantizacionih nivoa ($N \geq 64$) može se uvesti aproksimacija $p(x) \approx p(y_i)$, gde važi da $x \in [x_{i-1}, x_i)$, jer je za veliki broj nivoa kvantizacioni interval male širine i unutar njega je funkcija gustine verovatnoće signala približno jednaka, pa se jednačina (2.2.8) može dalje uprostiti:

$$D = \sum_{i=1}^N \int_{-x_{max} + (i-1)\Delta}^{-x_{max} + i\Delta} \left(x - \left(-x_{max} + \left(i - \frac{1}{2} \right) \Delta \right) \right)^2 p(x) dx, \quad (2.2.9)$$

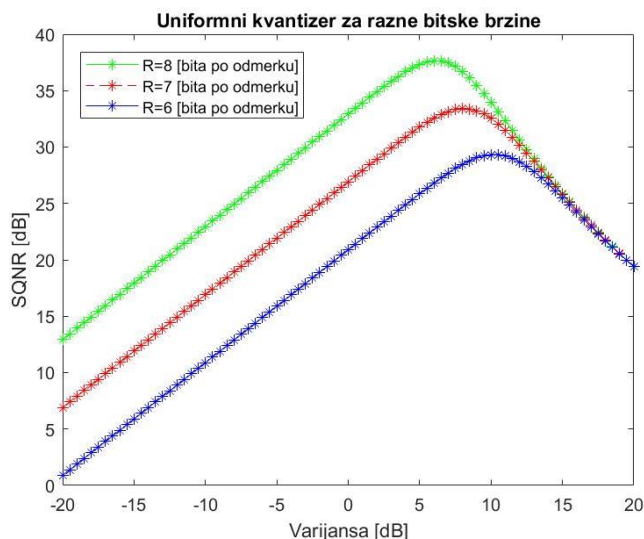
$$D = \sum_{i=1}^N p(y_i) \int_{-\Delta/2}^{\Delta/2} t^2 dt = \frac{\Delta^2}{12} \sum_{i=1}^N P_i, \quad (2.2.10)$$

gde je P_i verovatnoća i -tog intervala, a suma verovatnoća svih intervala jednaka je jedinici. Dalje se dobija izraz za ukupnu distorziju uniformnog kvantizera za ograničene neuniformne signale na sledeći način:

$$D = \frac{\Delta^2}{12} = \frac{x_{max}^2}{3N^2} \quad (2.2.11)$$

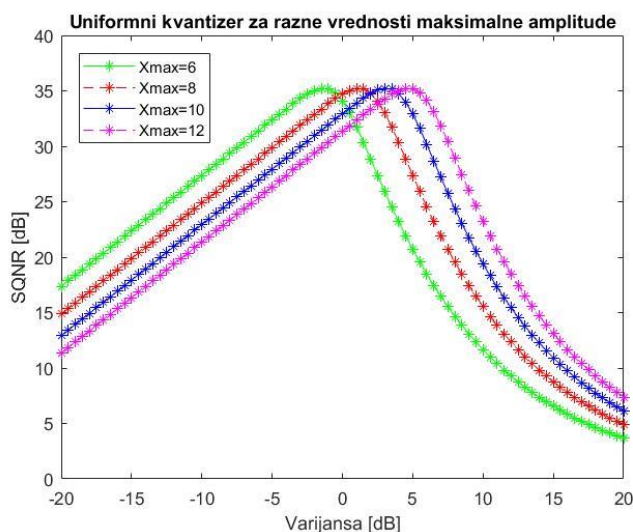
U slučaju neuniformnih signala koji nisu amplitudski ograničeni, distorzija prekoračenja nije jednaka nuli i ukupna distorzija tada predstavlja zbir granularne i distorzije prekoračenja. Najteži deo posla pri konstrukciji kvantizera je određivanje maksimalne amplitude kvantizera x_{max} . Optimalna vrednost x_{max} -a određuje se minimizacijom distorzije, tj. izjednačavanjem

prvog izvoda distorzije po maksimalnoj amplitudi sa nulom, $\frac{dD}{dx_{max}} = 0$. Na ovaj način obezbeđuje se minimalna distorzija odnosno maksimalni SQNR pri kvantizaciji [1] – [3], [5] – [11], [18] – [28]. Na slici 2.2.4 prikazan je SQNR u širokom opsegu ulaznih varijansi za bitske brzine uniformnog kvantizera R=6,7,8 bita po odmerku.



Slika 2.2.4. Skalarni uniformni kvantizer projektovan za razne bitske brzine

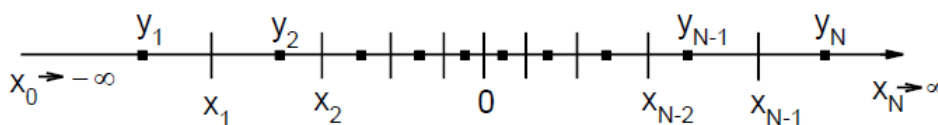
Izbor maksimalne amplitude opterećenja kvantizera je od velike važnosti pri projektovanju (naročito ako je na njegov ulaz doveden signal neuniformne raspodele) jer se smanjivanjem, tj. povećavanjem, x_{max} -a grafik SQNR-a u širokom opsegu ulaznih varijansi pomera levo-desno i time se direktno utiče na srednju vrednost SQNR-a u tom opsegu. Slika 2.2.5 prikazuje slučajeve projektovanja uniformnog kvantizera za Laplasov izvor i razne vrednosti maksimalne amplitude opterećenja.



Slika 2.2.5. Skalarni uniformni kvantizer projektovan za razne maksimalne amplitude opterećenja

2.2.2 Skalarni neuniformni kvantizer

Za signale koji nemaju uniformnu raspodelu gustine verovatnoće (kao što je govorni signal i većina signala u prirodi) pogodnije je koristiti skalarne neuniformne kvantizere. Neuniformni signali imaju karakteristiku da je verovatnoća da signal uzme male vrednosti po amplitudi mnogo veća od verovatnoće da uzme velike vrednosti po amplitudi i upravo zato je potrebno finije kodovanje male amplitudske vrednosti signala dok se kodovanje velikih vrednosti može vršiti dosta grublje. Upravo ovakvo kodovanje omogućavaju neuniformni kvantizeri kod kojih se pragovi odluke i reprezentacioni nivoi ne nalaze na međusobno jednakom rastojanju i reprezentacioni nivoi nisu na sredini kvantizacionog intervala. Neuniformni kvantizeri imaju karakteristiku da su nivoi gušće raspoređeni u oblastima malih signala po amplitudi (na manjem međusobnom rastojanju) dok su u oblastima signala sa velikim amplitudskim vrednostima dosta ređe raspoređeni (veće međusobno rastojanje). Na ovaj način se vrši fina kvantizacija malih signala koji su mnogo verovatniji od velikih signala čija je kvantizacija znatno grublja [1] – [3], [5] – [11], [18] – [28]. Na slici 2.2.4 prikazan je skalarni neuniformni kvantizer.



Slika 2.2.6. Skalarni neuniformni kvantizer

Kao što je prikazano u tabeli podele kvantizera po fizičkim karakteristikama, neuniformni kvantizeri mogu biti realizovani na dva načina:

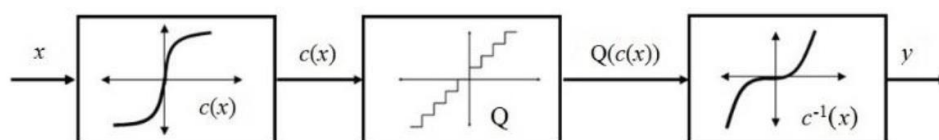
- Primenom iterativnog Lojd-Maks algoritma,
- Primenom kompanding tehnike.

Primenom Lojd-Maks iterativnog algoritma konstruišu se optimalni kvantizeri, koji za zadatu funkciju gustine verovatnoće imaju minimalnu distorziju [39] – [42]. Da bi se obezbedila minimalna distorzija pri procesu kvantizacije potrebno je da pragovi odluke x_i i reprezentacioni nivoi y_i budu optimalno raspoređeni, što se obezbeđuje postavljanjem uslova da su prvi izvodi distorzije po x_i i y_i jednaki nuli. Prolaskom kroz iteracije Lojd-Maks algoritma dobijaju se optimalne vrednosti pragova odluke x_i i reprezentacionih nivoa y_i , jedna po jedna i tako za sve kvantizacione nivoe. Sam algoritam je jako kompleksan te je samim tim pogodan za konstruisanje neuniformnih kvantizera sa malim i srednjim brojem kvantizacionih nivoa ($N \leq 128$). Ukoliko postoji potreba da se obezbedi viši kvalitet kvantizacije, tj da se kvantizeri projektuju sa više kvantizacionih nivoa ($N \geq 128$), kompleksnost primene Lojd-Maks iterativnog

algoritma se značajno povećava i nije pogodno koristiti ga, pa se tada za konstrukciju neuniformnih kvantizera primenjuje kompanding tehnika [1] – [3], [5] – [11], [18] – [28], [66] – [71], [74] – [76], [78], [79], [82] – [86].

2.2.3 Kompanding tehnika

U slučaju kada postoji potreba za konstruisanjem neuniformnog kvantizera sa velikim brojem kvantizacionih nivoa i obezbeđivanjem visokog kvaliteta kvantizacije pogodno je primeniti kompanding tehniku. Njenom primenom neuniformni kvantizeri se konstruišu iz tri koraka, kao redna veza kompresora, uniformnog kvantizera i ekspanzora, što je prikazano na slici 2.2.5 [1] – [3], [5] – [11], [18] – [28], [66] – [71], [74] – [76], [78], [79], [82] – [86].



Slika 2.2.7. Kompanding tehnika

Ovom metodom najpre se vrši kompresija signala neuniformnom kompresorskom funkcijom $c(x)$, zatim se takav signal kvantuje uniformnim kvantizerom Q i na kraju ekspanduje inverznom kompresorskom funkcijom $c^{-1}(x)$. Na izlazu se dobija signal sa funkcijom $y = c^{-1}(Q[c(x)])$. Kompresorska funkcija $c(x)$ ima takav oblik da male signale (signale sa velikom verovatnoćom) pojačava mnogo više od velikih signala, koji imaju malu verovatnoću. Na ovaj način smanjuje se razlika verovatnoća malih i velikih signala, dok funkcija gustine raspodele verovatnoće ima približno uniforman oblik na izlazu kompresora. Takav signal se kvantuje uniformnim kvantizerom Q što predstavlja drugi korak kompanding tehnike. Treći korak kompanding tehnike jeste ekspanzovanje, koje podrazumeva da se primenom inverzne kompresorske funkcije dobijaju reprezentacioni nivoi neuniformnog kvantizera [1] – [3], [5] – [11], [18] – [28]. Dve osnovne kompresorske funkcije koje se koriste prilikom konstrukcije neuniformnih kvantizera primenom kompanding tehnike su optimalna kompresorska funkcija i logaritamska kompresorska funkcija. Optimalna kompresorska funkcija je takvog oblika da se njome minimizuje distorzija D , i to granularna distorzija D_g zato što se za zadatu vrednost maksimalne amplitude kvantizera distorzija prekoračenja D_p ne može menjati jer ne zavisi od kompresorske funkcije i ima konstantnu vrednost. Granularna distorzija izražena kroz Benetov integral se minimizuje primenom Holderove nejednakosti [1] – [3], [5] – [11], [18] – [28].

Optimalna kompresorska funkcija ima karakteristiku da obezbeđuje maksimalni kvalitet kvantizacije SQNR (odnosno minimalnu distorziju) kada je ulazna snaga signala σ^2 jednaka

referentnoj snazi σ_0^2 za koju se projektuje kvantizer. Udaljavanjem ulazne snage signala σ^2 od referentne snage σ_0^2 za koju se projektuje kvantizer SQNR sve više opada i u širokom opsegu ulazne snage signala optimalni kvantizer ne obezbeđuje konstantno dobar kvalitet kvantizacije. Ovo je slučaj kod nestacionarnih signala kakav je i govorni signal, jer takvi signali znatno menjaju snagu u vremenu. Ovim se pokazuje da neuniformni kvantizatori sa optimalnom kompresorskom funkcijom nisu najbolje rešenje za kvantovanje nestacionarnih signala. Za takve signale potrebno je konstruisati kvantizer koji će obezbediti približno konstantan SQNR za ceo opseg ulaznih varijansi σ^2 . Da bi se ovo ostvarilo koriste se neuniformni kvantizatori sa logaritamskom kompresorskom funkcijom [1] – [3], [5] – [11], [18] – [28].

2.2.4 Logaritamski kvantizatori

Glavni zadatak neuniformnih kvantizatora sa logaritamskom kompresorskom funkcijom je da obezbede približno konstantan kvalitet kvantizacije, SQNR, u širokom opsegu ulaznih varijansi signala σ^2 , odnosno da kvalitet kvantizacije ne zavisi od ulazne snage signala σ . Da bi se dokazalo da logaritamski kvantizatori ne zavise od ulazne snage signala, a samim tim i od funkcije raspodele gustine verovatnoće $p(x)$, polazi se od pretpostavke da maksimalna amplituda opterećenja kvantizatora ima dovoljno veliku vrednost da se može zanemariti distorzija prekoračenja D_p . Shodno tome, izraz za dobijanje SQNR-a postaje [1] – [3], [5] – [11], [18] – [28]:

$$SQNR = \frac{\sigma^2}{D} = \frac{3N^2}{x_{\max}^2} \frac{\int_{-\infty}^{+\infty} x^2 p(x) dx}{\int_{-\infty}^{+\infty} p(x) \left(\frac{dc}{dx}\right)^{-2} dx}. \quad (2.2.12)$$

Uvođenjem smene

$$\frac{dc}{dx} = (kx)^{-1}, \quad (2.2.13)$$

gde k predstavlja konstantu izraz za računanje SQNR-a koji ne zavisi od ulazne snage signala ni od funkcije raspodele gustine verovatnoće. Da bi SQNR imao približno konstantnu vrednost u širokom opsegu snaga ulaznog signala, uslov iz izraza (2.2.13) mora biti zadovoljen. Integraljenjem leve i desne strane izraza (2.2.13) dobija se:

$$c(x) = \frac{1}{k} \ln(x) + C_1. \quad (2.2.14)$$

Konstanta integracije C_1 dobija se iz uslova:

$$c(x_{\max}) = x_{\max}. \quad (2.2.15)$$

Izjednačavanjem leve i desne strane izraza (2.2.15) dobija se konačni izraz za računanje konstante C_1 :

$$C_1 = x_{\max} - \frac{1}{k} \ln(x_{\max}). \quad (2.2.16)$$

Konačno, uvođenjem konstante k iz izraza (2.2.13) u izraz (2.2.12) dobija se jednačina za računanje SQNR-a neuniformnog kvantizera koja ne zavisi ni od ulazne varijanse signala σ^2 , niti od funkcije raspodele gustine verovatnoće $p(x)$ [1] – [3], [5] – [11], [18] – [28]:

$$SQNR = \frac{3N^2 \int_{-\infty}^{+\infty} x^2 p(x) dx}{x_{\max}^2 k^2 \int_{-\infty}^{+\infty} x^2 p(x) dx} = \frac{3N^2}{k^2 x_{\max}^2}. \quad (2.2.17)$$

Logaritamska kompresorska funkcija $C_{\log}(x)$ ima oblik:

$$C_{\log}(x) = \left(\frac{1}{k} \ln \frac{|x|}{x_{\max}} + x_{\max} \right) \operatorname{sgn}(x) = x_{\max} \left(1 + \frac{1}{K x_{\max}} \ln \frac{|x|}{x_{\max}} \right) \operatorname{sgn}(x), \quad (2.2.18)$$

gde $\operatorname{sgn}(x)$ predstavlja funkciju znaka.

Logaritamska kompresorska funkcija je takvog oblika da ne prolazi kroz koordinatni početak i da za male signale menja znak u nuli. Ovo dovodi do pojave velike greške prilikom kvantizacije jer upravo mali signali u okolini nule moraju biti najfinije kvantovani zato što imaju najveću verovatnoću. Matematički se ovaj nedostatak logaritamske kompresorske funkcije dokazuje iz uslova:

$$C_{\log}(x_0) = 0. \quad (2.2.19)$$

Rešavanjem jednačine (2.2.19) dobija se vrednost logaritamske kompresorske funkcije u nuli i ona iznosi:

$$x_0 = x_{\max} \cdot e^{-k x_{\max}}. \quad (2.2.20)$$

Problem neprolaska logaritamske kompresorske funkcije kroz koordinatni početak što uzrokuje veliku grešku pri kvantovanju malih signala prevaziđen je na način što su predložene dve modifikacije logaritamske kompresorske funkcije, a to su A i μ logaritamski zakoni kompresije [1] – [3], [5] – [11], [18] – [28].

2.2.4.1 Logaritamski A zakon kompresije

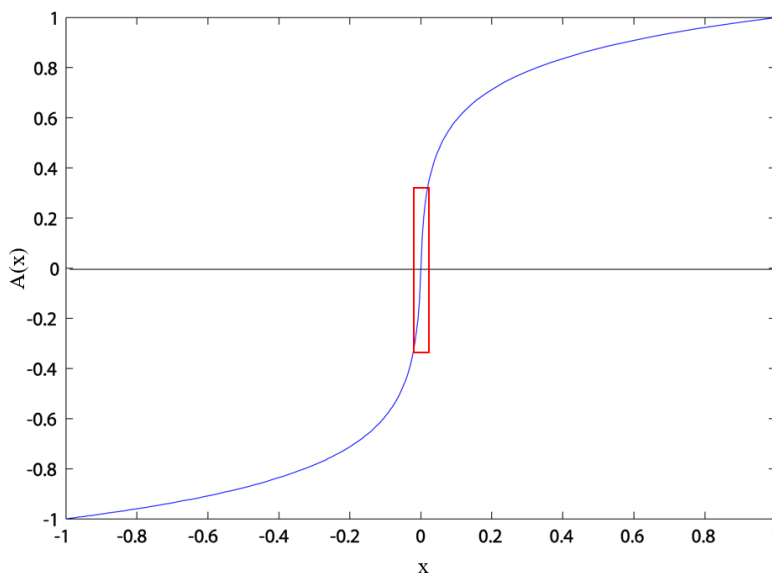
Logaritamski A zakon (semilogaritamski zakon) kompresije prevazilazi problem “neprolaska” kroz koordinatni početak na taj način što se funkcija $C_A(x)$ sastoji iz dva dela [1] – [3], [5] – [11], [18] – [28]:

- 1) Linearnog dela u okolini nule (koordinatnog početka) za male signale
- 2) Logaritamskog dela definisanog za srednje i velike signale.

Mora postojati jasno definisana granica između linearnog i logaritamskog dela funkcije i ta vrednost se označava sa x_{\min} i izračunava na osnovu izraza:

$$x_{\min} = \frac{x_{\max}}{A}. \quad (2.2.21)$$

gde je A parametar čija je vrednost uvek veća od 1 i definiše se u procesu projektovanja [1] – [3], [5] – [11]. Na slici 2.2.8 prikazana je semilogaritamska kompresorska funkcija gde se jasno može videti razlika između linearnog i logaritamskog dela.



Slika 2.2.8. Semilogaritamska kompresorska funkcija

Deo uokviren pravougaonikom predstavlja linearni deo ograničen vrednostima $-x_{\min}$ i x_{\min} , dok se nagib funkcije menja u zavisnosti od parametra A koji se definiše prilikom projektovanja kvantizera. Kompresorska funkcija $C_A(x)$ definisana A zakonom kompresije može se predstaviti kao:

$$C_A(x) = \begin{cases} \alpha|x|\operatorname{sgn}(x), & 0 \leq |x| \leq x_{\min} \\ x_{\max} \left(1 + \frac{1}{kx_{\max}} \ln \left(\frac{|x|}{x_{\max}} \right) \right), & x_{\min} \leq |x| \leq x_{\max} \end{cases} \quad (2.2.22)$$

Mora se dodati i uslov neprekidnosti u tački x_{\min} , što znači da funkcija $C_A(x)$ i njen prvi izvod $C'_A(x)$ budu neprekidni u tački x_{\min} :

$$C_A(x_{\min}^-) = C_A(x_{\min}^+), \quad (2.2.23)$$

$$C'_A(x_{\min}^-) = C'_A(x_{\min}^+), \quad (2.2.24)$$

Zamenom izraza iz jednačine (2.2.22) u (2.2.23) i (2.2.24) dobija se sledeće:

$$C_A(x_{\min}^-) = \alpha x_{\min} \quad (2.2.25)$$

$$C'_A(x_{\min}^+) = x_{\max} \left(1 + \frac{1}{kx_{\max}} \ln \left(\frac{x_{\min}}{x_{\max}} \right) \right), \quad (2.2.26)$$

pa se iz uslova (2.2.23) dobija:

$$K = \frac{1}{x_{\max}} \left(1 - \ln \left(\frac{x_{\min}}{x_{\max}} \right) \right) = \frac{1 + \ln(A)}{x_{\max}}. \quad (2.2.27)$$

Iz uslova (2.2.24) sledi:

$$\alpha = \frac{1}{kx_{\min}} = \frac{A}{1 + \ln(A)}, \quad (2.2.28)$$

pa se zamenom k i α u izraz (2.2.22) dobija konačni izraz za $C_A(x)$:

$$C_A(x) = \left\{ \begin{array}{ll} \frac{A|x|}{1 + \ln(A)} \operatorname{sgn}(x), & 0 \leq |x| \leq x_{\min} \\ \frac{x_{\max}}{1 + \ln(A)} \left(1 + \ln \left(\frac{A|x|}{x_{\max}} \right) \right) \operatorname{sgn}(x), & x_{\min} \leq |x| \leq x_{\max} \end{array} \right\} \quad (2.2.28)$$

Evropski PCM (*European Pulse Code Modulation*) standard definiše vrednost parametra $A=87.56$ i ova vrednost se najčešće koristi u praksi [1] – [3], [5] – [11], [18] – [28].

2.2.4.2 Logaritamski μ zakon kompresije

Druga modifikacija logaritamске kompresorske funkcije $C_{\log}(x)$, koja je čini primenljivom i za male signale, predstavljena je kao μ logaritamski zakon – kvazilogaritamski zakon. Kako bi se dobio eksplicitan izraz za logaritamsku kompresorsku funkciju definisanu μ zakonom $C_{\mu}(x)$, mora se početi od transliranja funkcije $C_{\log}(x)$ za x_0 (x_0 je definisano izrazom (2.2.20)) u levo pri čemu se dobija [1] – [3], [5] – [11], [18] – [28]:

$$\begin{aligned} C_{\mu}(x) &= C_{\log}(x + x_0) = x_{\max} \left[1 + \frac{1}{kx_{\max}} \ln \left(\frac{x + x_0}{x_{\max}} \right) \right] = \\ &= x_{\max} \left(1 + \frac{1}{kx_{\max}} \ln \left(\frac{x + x_{\max} e^{-kx_{\max}}}{x_{\max}} \right) \right) = \\ &= x_{\max} + \frac{1}{k} \ln \left(\frac{x}{x_{\max}} + e^{-kx_{\max}} \right) \end{aligned} \quad (2.2.29)$$

Kako se x_{\max} može napisati u obliku:

$$x_{\max} = \frac{1}{k} \ln \cdot e^{kx_{\max}}, \quad (2.2.30)$$

dobija se:

$$C_{\mu}(x) = \frac{1}{k} \ln \left(1 + \frac{x}{x_{\max}} e^{kx_{\max}} \right). \quad (2.2.31)$$

Ukoliko k zamenimo izrazom:

$$k = \frac{\ln(1 + \mu)}{x_{\max}}, \quad (2.2.32)$$

izraz za $C_{\mu}(x)$ postaje:

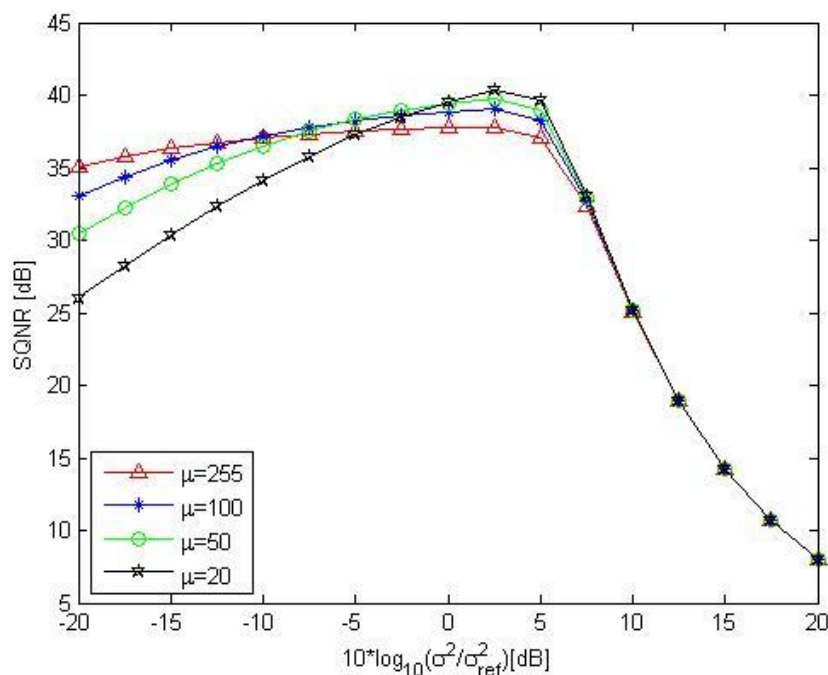
$$C_{\mu}(x) = \frac{x_{\max}}{\ln(1 + \mu)} \ln \left(1 + \frac{(1 + \mu)x}{x_{\max}} \right) \operatorname{sgn}(x). \quad (2.2.33)$$

Ukoliko je zadovoljen uslov $C_{\mu}(x_{\max}) = x_{\max}$ može se uvesti aproksimacija $1 + \mu = \mu$ pa konačan izraz za kvazilogaritamsku kompresorsku funkciju postaje:

$$C_{\mu}(x) = \frac{x_{\max}}{\ln(1 + \mu)} \ln \left(1 + \frac{\mu x}{x_{\max}} \right) \operatorname{sgn}(x), \quad -x_{\max} \leq x \leq x_{\max}. \quad (2.2.34)$$

U praksi se najčešće koristi vrednost faktora kompresije $\mu = 255$, koja je propisana standardom PCM za Severnu Ameriku [1] – [3], [5] – [11], [18] – [28].

Slika 2.2.9 prikazuje zavisnost SQNR-a od ulazne snage signala σ^2 u širokom opsegu snaga ulaznog signala (-20 dB, +20 dB), u slučaju projektovanja kvantizera za referentnu snagu σ_0^2 , koristeći kvazilogaritamsku kompresorsku funkciju za razne vrednosti faktora kompresije μ , kada je ulazni signal modelovan Laplasovom funkcijom raspodele gustine verovatnoće.



Slika 2.2.9. SQNR [dB] za signal modelovan Laplasovom raspodelom u širokom opsegu ulaznog signala za razne vrednosti faktora kompresije μ

Slika 2.2.9 daje jasan zaključak da se izborom vrednosti faktora kompresije μ direktno utiče na konstantnost kvaliteta kodovanja u širokom opsegu ulaznih varijansi. Vrednost $\mu = 20$ pruža najviši pik (vrh) SQNR-a za vrednost ulazne varijanse od 2.5 dB dok za ostale vrednosti ulazne varijanse naglo opada. Sa druge strane, vrednost $\mu = 255$ (standardna vrednost faktora kompresije u PCM standardu za Severnu Ameriku) obezbeđuje najkonstantniji SQNR u širokom opsegu ulaznih varijansi. Za vrednosti od -20 dB do 5 dB približno je konstantan, dok

za veće varijanse opada usled greške kvantizacije. Ovaj fenomen opadanja SQNR-a za veće varijanse prevazilazi se primenom neke od tehnika adaptacije kvantizera na ulazni signal, čijom se primenom postiže konstantan kvalitet kodovanja u celom opsegu ulaznih varijansi.

2.3 Tehnike adaptacije kvantizera na ulazni signal

Kao što je ranije napomenuto, u prirodi se nalazi veliki broj nestacionarnih signala čija se snaga, srednja vrednost, pa čak i funkcija raspodele gustine verovatnoće menjaju kroz vreme. Postoje različiti načini za prevazilaženje promena ovih parametara koji mogu biti kombinovani ukoliko se više parametara menja kroz vreme [1] – [3], [5] – [11], [18] – [28], [42], [53], [54], [56] – [59], [84]. Kvantizeri koji obezbeđuju konstantan SQNR u širokom opsegu ulaznih varijansi nazivaju se robustnim kvantizerima. U ovu grupu spadaju logaritamski neuniformni kvantizeri (A i μ zakon) i adaptivni kvantizeri. Logaritamski neuniformni kvantizeri imaju karakteristiku da obezbeđuju približno konstantan kvalitet kvantizacije u širokom opsegu ulaznih snaga (varijansi) pri velikim bitskim brzinama, dok se na srednjim i niskim bitskim brzinama njihovom upotrebom ne održava konstantni SQNR, već drastično opada. Adaptivni kvantizeri spadaju u drugu grupu robustih kvantizera koji pružaju konstantan kvalitet kvantizacije i na srednjim i na niskim bitskim brzinama, zato što prate promene snage ulaznog signala u vremenu, vrše procenu trenutne snage i prilagođavaju se njoj. Adaptacija se može vršiti kako za uniformni kvantizer, tako i za neuniformne kvantizere. Adaptacija kvantizera može biti urađena na dva načina [1] – [3], [5] – [11], [18] – [28], [56], [87]:

- 1) Adaptacijom unazad (*Backward adaptation*)
- 2) Adaptacijom unapred (*Forward adaptation*)

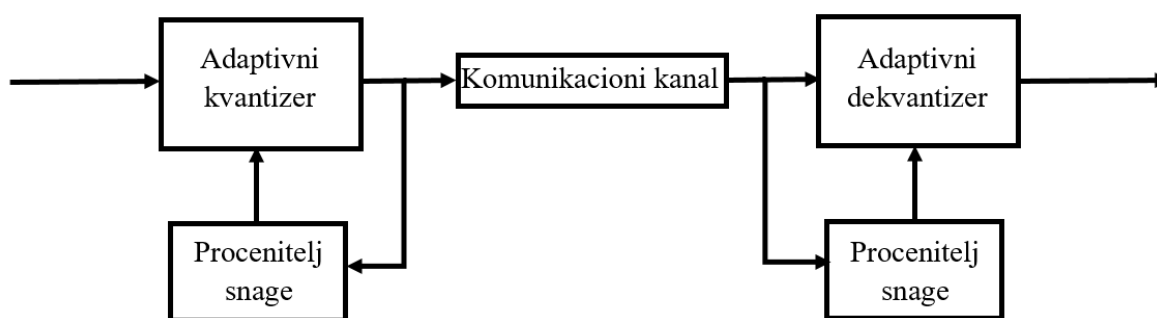
Prilikom adaptacije kvantizera adaptacijom unazad koristi se fiksni kvantizer projektovan za referentnu snagu σ_0^2 , a odmerci signala pri ulasku u kvantizer se dele sa $\frac{\hat{\sigma}_x(n)}{\sigma_0}$ kako bi se izvršila njihova kvantizacija na fiksni kvantizer. Ukoliko je trenutna snaga signala $\hat{\sigma}_x^2(n)$ i ona se подели sa $\frac{\hat{\sigma}_x(n)}{\sigma_0}$ trenutna snaga signala postaje σ_0^2 , što je upravo referentna snaga fiksnog kvantizera. Shodno tome u prijemu, nakon dekvantizacije, potrebno je pomnožiti snagu signala istom vrednošću $\frac{\hat{\sigma}_x(n)}{\sigma_0}$ kako bi se snaga odmerka signala vratila u realno stanje [1] – [3], [5] – [11], [18] – [28], [56], [87].

Tehnika adaptacije unapred je nesto složenija od adaptacije unazad jer se zahteva da se prijemu prosledi informacija o varijansi koja se kvantuje kao dodatna informacija. Naime,

kod adaptacije unapred koristi se kvantizer koji menja vrednosti pragova odluke i reprezentacionih nivoa na osnovu trenutne snage signala $\hat{\sigma}_x^2(n)$, dok je prilikom procesa dekvantizacije potrebno prijemniku menjati pragove odluke i reprezentacione nivoe na isti način na osnovu procenjene trenutne snage. Zbog toga je potrebno postojanje bloka procenitelja snage [1] – [3], [5] – [11], [18] – [28], [53], [54], [59], [71].

2.3.1 Adaptivni kvantizeri sa adaptacijom unazad

Tehnika adaptacije unazad vrši se na osnovu izlaznih (već kvantovanih) odmeraka [1] – [3], [5] – [11], [18] – [28], [56], [87]. S obzirom na to da odmerci koji su već kvantovani u sebi sadrže i grešku kvantizacije, ova tehnika adaptacije ima nešto slabije performanse u odnosu na tehniku adaptacije unapred, što predstavlja nedostatak ove tehnike. Sa druge strane, prednost adaptacije unazad u odnosu na adaptaciju unapred jeste činjenica da su izlazni nivoi kvantizera dostupni prijemniku i nije potrebno prosleđivati dodatne informacije o njima. Još jedna prednost ove tehnike je nepostojanje ulaznog bafera u kome se ulazni signal deli na frejmove te se samim tim ne unosi dodatno kašnjenje. Na slici 2.3.1 je prikazan adaptivni kvantizer sa adaptacijom unazad.



Slika 2.3.1. Adaptivni kvantizer sa adaptacijom unazad

S obzirom da se adaptacija vrši na osnovu već kvantovanih odmeraka može se zaključiti da je trenutna snaga ulaznog i izlaznog signala približno jednaka $\hat{\sigma}_x^2(n) \approx \hat{\sigma}_y^2(n)$. Procena snage se vrši na osnovu N već kvantovanih odmeraka kao:

$$\hat{\sigma}_y^2(n) = \frac{1}{N} \sum_{i=1}^N y^2(n-i). \quad (2.3.1)$$

U ovom slučaju svi izlazni odmerci imaju istu težinu, što ne mora uvek biti tako. Trenutna snaga može se izraziti i na način da već kvantovani odmerci imaju eksponencijalno opadajuće težine $1, \alpha^2, \alpha^3, \alpha^4, \alpha^5, \dots$ gde α predstavlja težinski faktor i može imati vrednost između nula i jedan. Tada se procena snage trenutnog odmerka računa kao:

$$\begin{aligned}\hat{\sigma}_y^2(n) &= C(y^2(n-1) + \alpha y^2(n-2) + \alpha^2 y^2(n-3) + \dots) \\ &= C \sum_{i=1}^{\infty} \alpha^{i-1} y^2(n-i)\end{aligned}\quad (2.3.2)$$

Iz jednačine (2.3.2) se vidi da najveću težinu ima poslednji izlazni odmerak signala, dok težina ostalih, koji su kvantovani pre njega (gledajući unazad), eksponencijalno opada [1] – [3], [5] – [11], [18] – [28]. Konstanta C predstavlja faktor normalizacije, pa je suma svih težinskih koeficijenata jednaka jedinici:

$$C(1 + \alpha^2 + \alpha^3 + \alpha^4 + \dots) = C \sum_{i=1}^{\infty} \alpha^{(i-1)} = C \frac{1}{1-\alpha} = 1. \quad (2.3.3)$$

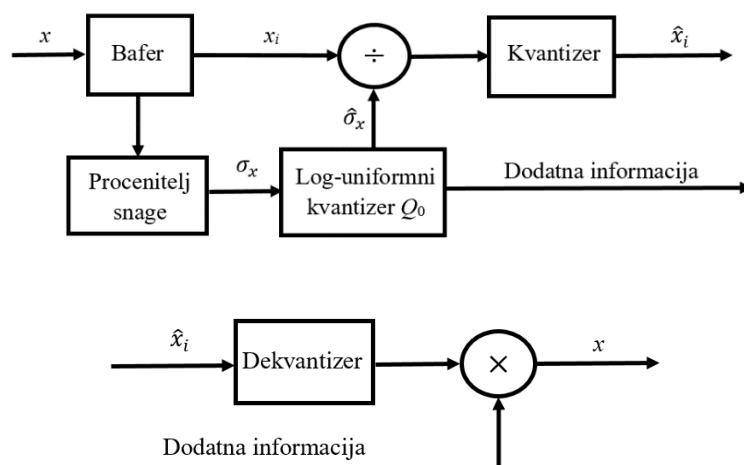
odakle se dobija da je $C = 1 - \alpha$, na osnovu čega sledi da je procenjena snaga trenutnog odmerka signala:

$$\hat{\sigma}_y^2(n) = (1 - \alpha)y^2(n-1) + \alpha\hat{\sigma}_y^2(n-1). \quad (2.3.4)$$

Na osnovu jednačine (2.3.4) zaključuje se da se procena snage trenutnog odmerka signala može izvršiti koristeći samo poslednji kvantovani odmerak jer on sadrži informacije o svim prethodnim odmercima koji su već kvantovani [1] – [3], [5] – [11], [18] – [28]. Ovo dokazuje činjenicu da nema potrebe za postojanjem bafera koji bi unosi dodatno kašnjenje jer je kašnjenje unešeno tehnikom adaptacije unazad samo zbog jednog odmerka (prvog koji se kvantuje jer on nema prethodni kvantovani odmerak).

2.3.2 Adaptivni kvantizeri sa adaptacijom unapred

Adaptivni kvantizeri koji za adaptaciju koriste tehniku adaptacije unapred nešto su drugačiji od adaptivnih kvantizera sa adaptacijom unazad. Adaptivni kvantizeri sa adaptacijom unapred sastoje se od bafera (u koji se učitavaju odmerci signala deleći ga na frejmove), procenitelja snage (gde se vrši procena snage na osnovu ulaznih, nekvantovanih odmeraka), kvantizera za kvantovanje varijanse, odnosno standardne devijacije signala u frejmu (u ovoj disertaciji koristi se log-uniformni kvantizer) i fiksnog kvantizera, što je prikazano na slici 2.3.2 [1] – [3], [5] – [11], [18] – [28], [53], [54], [59], [71].



Slika 2.3.2. Adaptivni kvantizer sa adaptacijom unapred (koder (gore) i dekoder (dole))

Na ulaz bafera dovodi se signal i u bafer se smešta frejm od M odmeraka signala ($x(n)$, $x(n-1), \dots, x(n-M+1)$) za koji procenitelj snage računa snagu (varijansu) na osnovu formule:

$$\sigma_x^2 = \frac{1}{M} \sum_{i=0}^{M-1} x^2(n-i). \quad (2.3.5)$$

Standardna devijacija signala u frejmu σ_x se dobija kao kvadratni koren varijanse frejma na osnovu koje se dobija pojačanje g :

$$g = \frac{\sigma_x}{\sigma_0}, \quad (2.3.6)$$

gde je σ_0^2 referentna snaga fiksnog kvantizera. Svaki odmerak u frejmu se deli pojačanjem g kako bi imao snagu σ_0^2 i prilagodio se fiksnom kvantizeru. Da bi se u prijemniku signal frejma rekonstruisao u prvobitni, potrebno je svaki odmerak pomnožiti pojačanjem g . S obzirom da je pojačanje g izračunato na osnovu ulaznih odmeraka signala, koji nisu dostupni prijemniku, potrebno je do prijemnika proslediti informaciju o pojačanju g kao dodatnu informaciju. Da bi se kroz komunikacioni kanal prosledila ta informacija, pojačanje g mora biti u digitalnom obliku, tj. mora da se kvantuje. Za ovu potrebu koristi se log-uniformni kvantizer [1] – [3], [5] – [11], [18] – [28], [53], [54], [59], [71]. Log-uniformni kvantizer je uniformni kvantizer čiji su pragovi odluke i reprezentacioni nivoi u logaritamskom domenu (u disertaciji u opsegu -20 dB, +20 dB) i projektuje se sa Ng kvantizacionih nivoa. Na izlazu iz log-uniformnog kvantizera nalazi se kvantovano pojačanje \hat{g} kojim se dele ulazni odmerci signala u frejmu i koje se prosleđuje prijemniku kao dodatna informacija na osnovu koje se, množenjem tom vrednošću, rekonstruiše kodovani signal [1] – [3], [5] – [11], [18] – [28], [53], [54], [59], [71].

2.4 Optimizacija faktora kompresije kvazi-logaritamskog kvantizera

Faktor kompresije μ kvazi-logaritamskog kvantizera je, pored maksimalne amplitude opterećenja kvantizera, od velike važnosti pri projektovanju, te je potrebno odrediti optimalnu vrednost ovog parametra. U ovoj disertaciji predložen je novi metod u dva koraka za određivanje optimalne vrednosti ovog parametra, čije su performanse upoređene sa Milerovim metodom u radu [88]. Glavni cilj ove metode je dobijanje eksplicitnog izraza za optimalnu vrednost faktora kompresije. Oba metoda, predloženi i Milerov metod, kreću izjednačavanjem prvog izvoda Bentovog integrala sa nulom. Nakon ovoga, primenjeni su Milerov i predloženi metod u dva koraka za dobijanje optimalne vrednosti i upoređene vrednosti.

Milerov iterativni metod radi na principu aproksimiranja funkcije u blizini korena kvadratnog polinoma. Za rešenje dobijenih transedentalnih jednačina dobijenih ovim aproksimacijama primenom predloženog metoda u dva koraka primećuje se da je moguće uvesti smenu koja bi kompleksnost tih jednačina značajno smanjila na rešavanje kvadratne jednačine. Važan parametar u oba metoda je maksimalna amplituda kvantizera x_{\max} i iz tog razloga ona mora biti određena optimalno [30], [31], [66], [68], [69]. Eksplicitni izraz za dobijanje optimalne vrednosti maksimalne amplitude izveden je u radu [66] i korišćen u predloženom metodu. Kao što je napomenuto, oba metoda počinu Benetovim integralom koji predstavlja granularnu distorziju kvazi-logaritamskog kvantizera:

$$D_g = \frac{x_{\max}^2}{3N^2} \int_{-x_{\max}}^{x_{\max}} \frac{p(x)}{[c'(x)]^2} dx, \quad (2.4.1)$$

gde $c'(x)$ predstavlja prvi izvod kompresorske funkcije kvazi-logaritamskog kvantizera, prikazane formulom 2.2.34, $p(x)$ je Laplasova funkcija raspodele gustine verovatnoće, izražena formulom 2.1.2, N broj kvantizacionih nivoa kvantizera gde važi $x_N \approx x_{\max}$, a x_{\max} je maksimalna amplituda opterećenja kvantizera. Zamenom jednačina 2.1.2 i 2.2.34 u jednačinu 2.4.1, uz aproksimaciju $y_N \approx x_{\max}$, dobija se izraz za ukupnu distorziju koji zavisi od varijanse:

$$D_g = \frac{\ln^2(1 + \mu)\sigma^2}{3N^2} \left(\frac{x_{\max}^2}{\mu^2\sigma^2} + \frac{\sqrt{2}x_{\max}}{\mu\sigma} + 1 \right) + \sigma^2 \cdot \exp\left(-\frac{\sqrt{2}x_{\max}}{\sigma}\right). \quad (2.4.2)$$

Prvi izvod izraza prikazanog u formuli 2.4.2 (za jediničnu varijansu $\sigma^2=1$) po faktoru kompresije μ se u oba metoda izjednačava sa nulom čime se dobija polazna tačka za optimizaciju faktora kompresije predloženim i Milerovim metodom.

Osnova Milerovog metoda optimizacije je aproksimacija funkcije u blizini kvadratnog polinoma [69], što praktično znači pronalazak tih korena koristeći kvadratnu interpolaciju u tri tačke. Korišćenjem tih tačaka konstruiše se parabola čiji presek sa x osom predstavlja sledeću aproksimaciju. Predloženi metod optimizacije faktora kompresije je znatno prostiji. Izjednačavanjem prvog izvoda ukupne distorzije sa nulom dobija se sledeći izraz:

$$\mu^2 + \sqrt{2}x_{\max} \left(1 - \frac{\mu + 1}{\mu} \frac{\ln(\mu + 1)}{2}\right) \mu + x_{\max}^2 \left(1 - \frac{\mu + 1}{\mu} \ln(\mu + 1)\right) = 0. \quad (2.4.3)$$

Primećuje se da je moguće uvesti smenu:

$$t_0 = \frac{\mu + 1}{\mu} \ln(\mu + 1). \quad (2.4.4)$$

Uvođenjem ove smene, jednačina 2.4.3 postaje prosta kvadratna jednačina:

$$\mu^2 + \sqrt{2}x_{\max} \left(1 - \frac{t_0}{2}\right) \mu + x_{\max}^2(1 - t_0) = 0, \quad (2.4.5)$$

čijim se rešavanjem dobijaju dva rešenja:

$$\mu = \frac{\sqrt{2}x_{\max}}{4} \left[(t_0 - 2) \pm \sqrt{t_0^2 + t_0 + 1} \right], \quad (2.4.6)$$

od kojih se uzima samo vrednost koja ima fizički smisao a to je pozitivna vrednost faktora kompresije. Optimalna vrednost faktora kompresije dobija se u slucaju minimalne distorzije, odnosno maksimalnog SQNR-a, što se postoze u drugom koraku predloženog metoda. Vraćanjem smene 2.4.4 u jednačinu 2.4.6 dobija se formula u zatvorenom obliku koja se koristi u oba koraka predloženog metoda. U prvom koraku dobija se μ_1 , koji zavisi od početne vrednosti faktora kompresije μ_0 :

$$\begin{aligned} \mu_1 = \frac{\sqrt{2}x_{\max}}{4} & \left[\left(\left(\frac{\mu_0 + 1}{\mu_0} \ln(\mu_0 + 1) \right) - 2 \right) \right. \\ & \left. \pm \sqrt{\left(\frac{\mu_0 + 1}{\mu_0} \ln(\mu_0 + 1) \right)^2 + \left(\frac{\mu_0 + 1}{\mu_0} \ln(\mu_0 + 1) \right) + 1} \right], \end{aligned} \quad (2.4.7)$$

dok se u drugom koraku dobija optimalna vrednost faktora kompresije, koji zavisi od μ_1 :

$$\begin{aligned} \mu_2 = \frac{\sqrt{2}x_{\max}}{4} & \left[\left(\left(\frac{\mu_1 + 1}{\mu_1} \ln(\mu_1 + 1) \right) - 2 \right) \right. \\ & \left. \pm \sqrt{\left(\frac{\mu_1 + 1}{\mu_1} \ln(\mu_1 + 1) \right)^2 + \left(\frac{\mu_1 + 1}{\mu_1} \ln(\mu_1 + 1) \right) + 1} \right]. \end{aligned} \quad (2.4.8)$$

Maksimalna amplituda x_{\max} kao važan parametar mora se računati iznova za svaki korak na osnovu jednačine iz rada [66]. Procedura dobijanja optimalnog faktora kompresije je sledeća: na osnovu početne vrednosti faktora kompresije μ_0 računaju se maksimalna amplituda opterećenja kvantizera, distorzija, SQNR i nova vrednost faktora kompresije μ_1 . Zatim se u drugom koraku na osnovu μ_1 računa sve ispočetka i dobija μ_2 koja predstavlja optimalnu vrednost zato što obezbeđuje minimalnu distorziju. Početna vrednost faktora kompresije μ_0 iznosi 128, što je otprilike polovina vrednosti μ u standardu G.711. Polazeći od ove vrednosti, primenom predloženog metoda u dva koraka dobija se optimalna vrednost za μ . Proces optimizacije je kroz rezultate prikazan u tabelama 2.4.1 – 2.4.4.

Tabela 2.4.1. Faktor kompresije μ , maksimalna amplituda x_{\max} , distorzija D , i SQNR za broj kvantizacionih nivoa $N=32$

Korak	μ	x_{\max}	D	SQNR [dB]
μ_0	128	6.872	0.00835415	20.781
μ_1	19.881	6.218	0.00477807	23.2075
μ_2	10.331	6.073	0.00477807	23.6032

Tabela 2.4.2. Faktor kompresije μ , maksimalna amplituda x_{\max} , distorzija D , i SQNR za broj kvantizacionih nivoa $N=64$

Korak	μ	x_{\max}	D	SQNR [dB]
μ_0	128	7.8622	0.00211105	26.775
μ_1	22.6374	7.2347	0.00130122	28.8565
μ_2	12.5906	7.0919	0.00121541	29.1528

Tabela 2.4.3. Faktor kompresije μ , maksimalna amplituda x_{\max} , distorzija D , i SQNR za broj kvantizacionih nivoa $N=128$

Korak	μ	x_{\max}	D	SQNR [dB]
μ_0	128	8.83234	0.000533447	32.7291
μ_1	25.4629	8.24843	0.000349829	34.5614
μ_2	14.9603	8.1095	0.00033213	34.7869

Tabela 2.4.4. Faktor kompresije μ , maksimalna amplituda x_{\max} , distorzija D , i SQNR za broj kvantizacionih nivoa $N=256$

Korak	μ	x_{\max}	D	SQNR [dB]
μ_0	128	9.81245	0.000134797	38.7032
μ_1	28.2885	9.25981	0.000093138	40.3087
μ_2	17.4093	9.12575	0.000089491	40.4822

Za primenu Milerovog iterativnog metoda potrebno je izabrati početne vrednosti 3 tačke kojima je označena parabola i u radu [88] postavljeno je $(x_1, x_2, x_3) = (1, 255, 128)$. Nakon inicijalne postavke parabole primenjuje se Milerov algoritam i za svaku iteraciju računa se relativna greška δ , na osnovu koje se algoritam zaustavlja kada ona padne ispod praga $\varepsilon=1\%$. Postavljanjem graničnih tačaka $(x_1, x_2) = (1, 255)$ osigurava se da algoritam konvergira ka optimalnom rešenju. Optimizacija faktora kompresije Milerovim iterativnim metodom urađena je za iste bitske brzine kao i predloženim metodom i rezultati su prikazani u tabelama 2.4.5 – 2.4.8.

Tabela 2.4.5. Parametri za analizu tačnosti Milerovog iterativnog metoda za broj kvantizacionih nivoa $N=32$

I	x_1	x_2	x_3	x_4	δ [%]
1	1	255	128	88.028	45.4
2	1	128	88.028	54.8561	60.47
3	1	88.028	54.8561	36.2045	51.52
4	1	54.8561	36.2045	24.1396	49.98
5	1	6.2045	24.1396	16.8432	43.31
6	1	24.1396	16.8432	12.4921	34.83
7	1	16.8432	12.4921	10.1286	23.34
8	1	12.4921	10.1286	9.0669	11.62
9	1	10.1286	9.0669	8.7428	3.71
10	1	9.0669	8.7428	8.6904	0.6

Tabela 2.4.6. Parametri za analizu tačnosti Milerovog iterativnog metoda za broj kvantizacionih nivoa $N=64$

I	x_1	x_2	x_3	x_4	δ [%]
1	1	255	128	88.568	44.52
2	1	128	88.568	55.6518	59.14
3	1	88.568	55.6518	37.2201	49.52
4	1	55.6518	37.2201	25.3594	46.77
5	1	37.2201	25.3594	18.3005	38.57
6	1	25.3594	18.3005	14.2467	28.45
7	1	18.3005	14.2467	12.2252	16.53
8	1	14.2467	12.2252	11.463	6.64
9	1	12.2252	11.463	11.2913	1.52
10	1	11.463	11.2913	11.2726	0.16

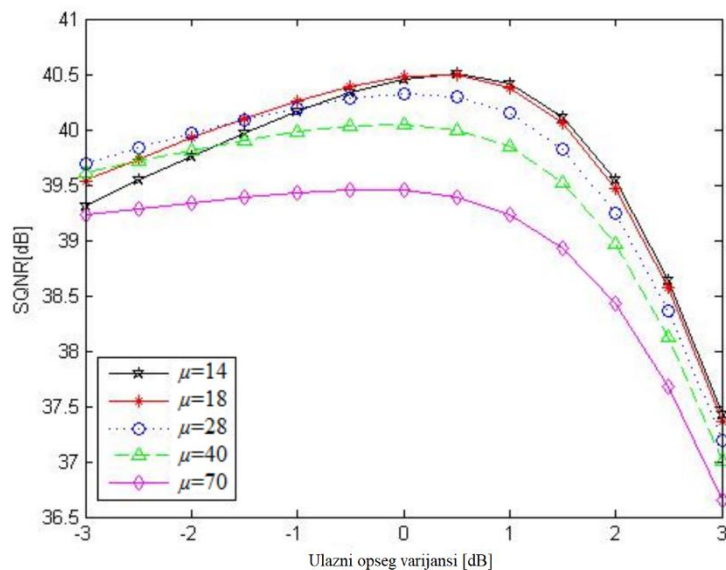
Tabela 2.4.7. Parametri za analizu tačnosti Milerovog iterativnog metoda za broj kvantizacionih nivoa $N=128$

I	x_1	x_2	x_3	x_4	δ [%]
1	1	255	128	89.1369	43.59
2	1	128	89.1369	56.507	57.74
3	1	89.1369	56.507	38.3309	47.42
4	1	56.507	38.3309	26.721	43.44
5	1	38.3309	26.721	19.9559	33.9
6	1	26.721	19.9559	16.2559	22.76
7	1	19.9559	16.2559	14.5981	11.35
8	1	16.2559	14.5981	14.0895	3.61
9	1	14.5981	14.0895	14.0065	0.59
10	1	14.0895	14.0065	14.0001	0.045

Tabela 2.4.8. Parametri za analizu tačnosti Milerovog iterativnog metoda za broj kvantizacionih nivoa $N=256$

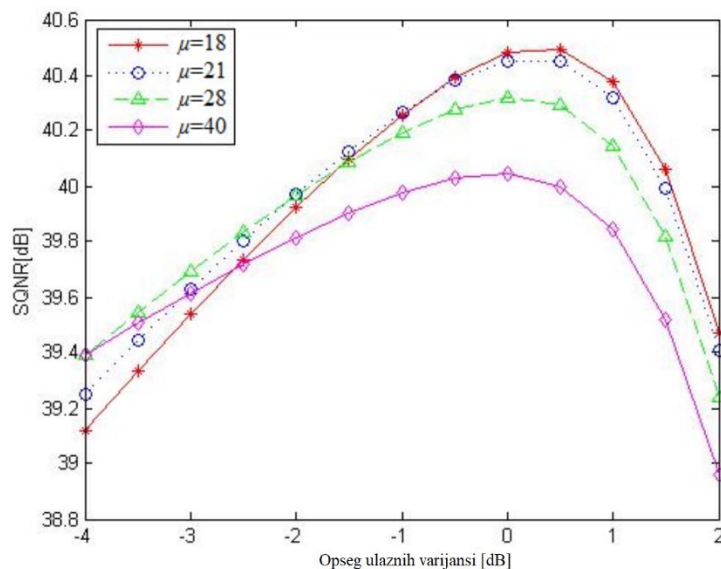
I	x_1	x_2	x_3	x_4	δ [%]
1	1	255	128	89.7358	45.69
2	1	128	89.7358	57.424	56.27
3	1	89.7358	57.424	39.5407	45.22
4	1	57.424	39.5407	28.2291	40.07
5	1	39.5407	28.2291	21.8116	29.42
6	1	28.2291	21.8116	18.5076	17.85
7	1	21.8116	18.5076	17.2046	7.57
8	1	18.5076	17.2046	16.8868	1.88
9	1	17.2046	16.8868	16.8491	0.22
10	1	16.8868	16.8491	16.8469	0.01

Nakon određivanja optimalne vrednosti faktora kompresije μ u slučaju Laplasovog izvora jedinične varijanse, optimalna vrednost je potražena u uskom opsegu ulaznih varijansi $\sigma^2 \in [-3 \text{ dB}, 3 \text{ dB}]$. Na slici 2.4.1 prikazana je analiza optimizacije faktora kompresije u slučaju $N=256$ i $x_{\max}=9.12$. Analiza je urađena za vrednosti faktora kompresije μ , $\mu \in (14, 18, 28, 40, 70)$, što čini širok opseg vrednosti faktora kompresije i iz koga se može zaključiti koja je vrednost μ optimalna za dati opseg ulaznih varijansi.



Slika 2.4.1. SQNR u uskom opsegu ulaznih varijansi [-3 dB, 3 dB] za $N=256$, $x_{\max}=9.12$ i razne vrednosti faktora kompresije μ

Da bi se zadovoljio uslov da SQNR treba da ima približno iste vrednosti na obe strane posmatranog opsega, sagledan je i opseg ulaznih varijansi $\sigma^2 \in [-4 \text{ dB}, 2 \text{ dB}]$, gde su vrednosti faktora kompresije $\mu \in (18, 21, 28, 40)$ i isti broj kvantizacionih nivoa i vrednost maksimalne amplitude. Ovaj slučaj prikazan je na slici 2.4.2.



Slika 2.4.2. SQNR u uskom opsegu ulaznih varijansi [-4 dB, 2 dB] za $N=256$, $x_{\max}=9.12$ i razne vrednosti faktora kompresije μ

Sa slike 2.4.1 može se zaključiti da u opsegu $\sigma^2 \in [-3 \text{ dB}, 3 \text{ dB}]$ optimalna vrednost faktora kompresije iznosi $\mu=18$, zato što obezbeđuje maksimalni SQNR, a ujedno je i najbliža

optimalnoj vrednosti za jediničnu varijansu. Slika 2.4.2 pokazuje da u posmatranom opsegu ulaznih varijansi optimalna vrednost faktora kompresije iznosi $\mu=21$ jer se za nju dobija najviši SQNR.

Kao što je već naglašeno, optimizacija faktora kompresije μ je od velike važnosti pri projektovanju kvazi-logaritamskog kvantizera. Zbog toga ova disertacija predlaže metod optimizacije ovog parametra u 2 koraka, što predstavlja veliko umanjeње kompleksnosti poredeći sa Milerovim iterativnim metodom koji do optimalne vrednosti dolazi u 9-oj ili 10-oj iteraciji u zavisnosti od broja kvantizacionih nivoa.

3 Tehnike transformacionog kodovanja govornog signala

U ovom poglavlju opisane su tehnike prema kojima se ulazni signal razlaže (transformiše) na komponente prema određenim pravilima, a zatim te komponente koduju pojedinačno prema svojim individualnim karakteristikama. Detaljno su obrađene tehnike transformacionog kodovanja koje uključuju transformacije kao što su diskretna Hadamardova transformacija DHT (*Discrete Hadamard Transform*), diskretna kosinusna transformacija DCT (*Discrete Cosine Transform*), kao i diskretna wavelet transformacija DWT (*Discrete Wavelet Transform*) [1] – [3], [5] – [11], [18] – [28], [59], [60], [89] – [111].

Transformaciono kodovanje predstavlja skup tehnika koje originalni ulazni signal u vremenskom domenu transformišu u drugi reprezentacioni domen (npr. frekvencijski) u cilju obezbeđivanja efikasnijih karakteristika signala za kodovanje [1] – [3], [5] – [11], [18] – [28], [53], [54], [59], [71]. Cilj transformacionog kodovanja je pregrupisanje odmeraka ulaznog signala kako bi se dobile sekvence signala koje nose veću količinu informacije. Ovakvo pregrupisanje pružilo bi mogućnost korišćenja manjeg broja bita po odmerku potrebnih za kodovanje poredeći sa tehnikama koje ne uključuju transformacije s tim da kvalitet

rekonstruisanog signala ostane isti. Da bi signal bio uspešno rekonstruisan na prijemnoj strani, koeficijenti transformacije takođe moraju biti kvantovani i poslani ka prijemniku zajedno sa odmercima signala. Na prijemnoj strani primenjuju se inverzne transformacije primenjenim u predajniku i na taj način se rekonstruiše signal. Transformacione tehnike su deo nekoliko ITU-T standarda za visoko kvalitetno kodovanje govornog i audio signala.

Transformacione tehnike koje su obrađene u disertaciji imaju matrice linearnih koeficijenata, tj. linearno transformišu ulazni signal x u izlazni signal y ,

$$y = A \cdot x, \quad (3.0.1)$$

dok se ulazni signal x dobija primenom inverzne transformacije nad signalom y :

$$x = B \cdot y, \quad (3.0.2)$$

gde je matrica B inverzna matrica matrici A i za njih važi:

$$A \cdot B = I. \quad (3.0.3)$$

Odmerci izlaznih signala, transformisani odmerci, dobijaju se koristeći jednačinu:

$$y_n = \sum_{i=0}^{N-1} x_i a_{i,n}, \quad (3.0.4)$$

dok se rekonstruišući odmerci na prijemnoj strani dobijaju:

$$x_n = \sum_{i=0}^{N-1} y_i b_{n,i}, \quad (3.0.5)$$

gde su A i B kvadratne matrice veličine $N \times N$ a i -ti i j -ti elementi matrica $a_{i,j}$ i $b_{i,j}$ dati su sa:

$$\begin{aligned} [A]_{i,j} &= a_{i,j}, \\ [B]_{i,j} &= b_{i,j}. \end{aligned} \quad (3.0.6)$$

Na osnovu jednačina (3.0.4) i (3.0.5) dobijaju se jednodimenzionalne sekvence signala, što je pogodno za govorni i audio signal. Međutim, transformaciono kodovanje je jedna od najčešćih metoda kodovanja slike, ali se onda zbog prirode slike dobijaju dvodimenzionalne matrice izlaznih sekvenci. Pošto je glavni cilj disertacije projektovanje kvantizera za kodovanje govornog signala, u nastavku će više reči biti o transformacionim tehnikama sa jednodimenzionalnim matricama.

3.1 Diskretna Hadamarova transformacija

Diskretna Hadamarova transformacija predstavlja generalizovanu klasu Furije-ove transformacije koja uz pomoć koeficijenata transformacije vrši simetrične i linearne operacije

nad odmercima ulaznog signala [1] – [3], [5] – [11], [18] – [28], [59] – [62], [92] – [98]. Hadamarova transformacija se vrši na osnovu matrice koeficijenata koja može biti različitih veličina, te se tako razlikuju matrice veličine 2×2 , 3×3 , 4×4 , ... Na osnovu ovih matrica koeficijenata Hadamarova transformacija deli ulazni signal na više sekvenci (grana) koje se posebno koduju i tako prenose kroz komunikacioni kanal do prijemnika. Hadamarova transformacija je dosta korišćena prilikom enkripcije podataka, kao i u raznim algoritmima za obradu i kompresiju signala. Za kompresiju video signala obično se koristi u formi sume apsolutnih transformacionih razlika. Takođe je ključni deo Groverovog i Šorovog (*Grover* i *Shor*) algoritma u kvantnom računarstvu. Hadamarova transformacija je takođe primenjena u naučnim metodama spektroskopije i kristalografije. Pogodna je za upotrebu prilikom analize i poređenja velikih govornih signala u smislu kompresije zbog svojih prednosti, kao što su brzo računanje transformacionih koeficijenata, manje zahtevanog prostora za čuvanje podataka, kao i brze rekonstrukcije signala [1] – [3], [5] – [11], [18] – [28].

Hadamarova transformacija H_m je predstavljena matricom koeficijenata veličine $2^m \times 2^m$ (skalirana faktorom normalizacije), koji transformišu 2^m odmeraka ulaznog signala u 2^m transformisanih odmeraka ulaznog signala. Hadamarova matrica koeficijenata može biti definisana rekurzivno ili koristeći binarnu prezentaciju (sa osnovom 2) indeksa n i k .

Rekurzivno, matrica Hadamarovih koeficijenata veličine 1×1 (Hadamarova transformacija H_0) ima samo jedan element i definiše se kao:

$$H_0 = 1, \quad (3.1.1)$$

a na osnovu nje se mogu definisati matrice H_m koeficijenata veličine $m \times m$, gde je $m > 0$:

$$H_m = \frac{1}{\sqrt{2}} \begin{bmatrix} H_{m-1} & H_{m-1} \\ H_{m-1} & -H_{m-1} \end{bmatrix}, \quad (3.1.2)$$

gde je $\frac{1}{\sqrt{2}}$ faktor normalizacije koji se ponekad izostavlja. Za $m > 1$, Hadamarova matrica se može definisati i kao:

$$H_m = H_1 \otimes H_{m-1}. \quad (3.1.3)$$

Ekvivalentno, koristeći binarnu prezentaciju koeficijenata n i k (sa osnovom 2), Hadamarova matrica se definiše preko:

$$k = \sum_{i=0}^{m-1} k_i 2^i = k_{m-1} 2^{m-1} + k_{m-2} 2^{m-2} + \dots + k_1 2 + k_0, \quad (3.1.4)$$

$$n = \sum_{i=0}^{m-1} n_i 2^i = n_{m-1} 2^{m-1} + n_{m-2} 2^{m-2} + \dots + n_1 2 + n_0, \quad (3.1.5)$$

gde su k_j i n_j binarna prezentacija (nula ili jedan) indeksa n i k .

Matrice koeficijenata Hadamarove transformacije H_m gde je $0 \leq m \leq 3$ su:

$$H_0 = 1, \quad m = 0 \quad (3.1.6)$$

$$H_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad m = 1, \quad (3.1.7)$$

$$H_2 = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \quad m = 2, \quad (3.1.8)$$

$$H_3 = \frac{1}{2^{\frac{3}{2}}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix} \quad m = 3, \quad (3.1.9)$$

Elementi matrica za bilo koje m se računaju po formuli:

$$(H_n)_{i,j} = \frac{1}{2^{\frac{n}{2}}} (-1)^{i \cdot j}, \quad (3.1.10)$$

gde su i i j vrsta i kolona matrice koeficijenata Hadamarove transformacije [1] – [3], [5] – [11], [18] – [28].

Kao što je već napomenuto, Hadamarova transformacija dekomponuje signal od M odmeraka, na osnovu matricnih koeficijenata, u posebne sekvence koje zbirno i dalje imaju M odmeraka ali u sebi nose veću količinu informacija (prealocira se energija signala). Ovo podrazumeva da je informacija koju je potrebno preneti od predajnika ka prijemniku sada sadržana u manjem broju odmeraka koji se fino kvantuju, dok se na kvantovanju odmeraka koji ne nose veliku količinu informacija može smanjiti broj bitova potrebnih za kodovanje. Na ovaj način Hadamarova transformacija održava kvalitet kodovanja dok se ono vrši manjim brojem bitova po odmerku.

U ovoj disertaciji praktično je primenjena modifikacija Hadamarove transformacije H_m za $m=1,2$ (matrica koeficijenata 2×2 i 4×4). Modifikacija Hadamarove transformacije H_m za $m=1$ je u vidu uprošćenja faktora normalizacije sa $\frac{1}{\sqrt{2}}$ na $\frac{1}{2}$ čime nije poremećen koncept transformacije [59], [71]. Modifikacija Hadamarove transformacije H_m za $m=2$ je u tome što za transformaciju ulaznog signala nije korišćena klasična Hadamarova transformacija H_2 već je dva puta iskorišćena Hadamarova transformacija H_1 . Time se bolje prealocira energija signala (informacija koju prenosimo) u jednu sekvencu u odnosu na slučaj korišćenja Hadamarove transformacije H_2 .

Modifikovana matrica koeficijentata Hadamarove transformacije H_1 može se prikazati kao:

$$H_1 = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad m = 1, \quad (3.1.11)$$

pa se odmerci ulaznog signala transformišu u paru (dva po dva) na sledeći način:

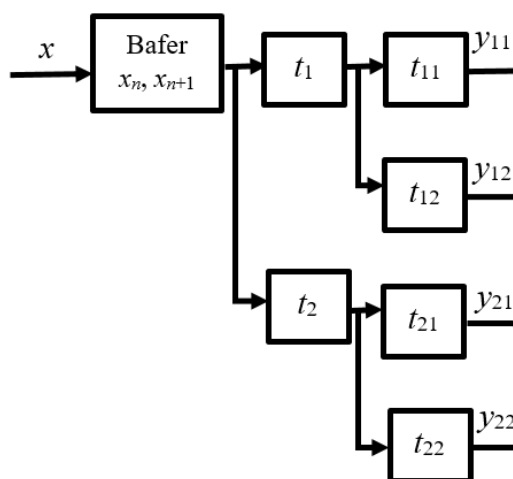
$$\begin{bmatrix} x_n^T \\ x_{n+1}^T \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_n \\ x_{n+1} \end{bmatrix}, \quad (3.1.12)$$

gde x_n i x_{n+1} predstavljaju dva susedna odmerka ulaznog signala dok x_n^T i x_{n+1}^T predstavljaju dva odmerka signala posle primene transformacije. Množenjem ovih matrica dobijaju se izrazi preko koji se formiraju dve posebne sekvence transformisanog signala:

$$t_1 = \frac{x_n + x_{n+1}}{2}, \quad (3.1.13)$$

$$t_2 = \frac{x_n - x_{n+1}}{2}. \quad (3.1.14)$$

Šema koja bi grafički prikazala drugu modifikaciju Hadamarove transformacije data je na slici 3.1.1.



Slika 3.1.1. Modifikacija Hadamarove transformacije H_2 u dvostruku transformaciju H_1

Sa slike 3.1.1 može se videti da se nakon primene dvostruke Hadamarove transformacije H_1 dobijaju četiri sekvence odmeraka signala, kao i korišćenjem Hadamarove transformacije H_2 . Kodovanje se vrši pojedinačno s tim da je energija signala (korisna informacija koja se prenosi do prijemnika) bolje raspoređena.

3.2 Diskretna kosinusna transformacija - DCT

Kao i Hadamarova, diskretna kosinusna transformacija (DCT) dekomponuje (razdvaja) ulazni signal na posebne sekvence odmeraka koristeći matrice koeficijente i time prealocira (transformiše) energiju signala na pogodniji način za kodovanje manjim brojem bita po odmerku. Diskretna kosinusna transformacija generiše konačni broj izlaznih sekvenci odmeraka signala kao sumu kosinusnih funkcija sa različitim frekvencijama, te je matrica koeficijenata data kao funkcija kosinusa [1] – [3], [5] – [11], [18], [99] – [104]:

$$[C]_{i,j} = \left\{ \begin{array}{l} \sqrt{\frac{1}{N}} \cos \frac{(2j+1)i\pi}{2N}, i = 0, j = 0, 1, \dots, N-1 \\ \sqrt{\frac{2}{N}} \cos \frac{(2j+1)i\pi}{2N}, i = 1, 2, \dots, N-1, j = 0, 1, \dots, N-1 \end{array} \right\} \quad (3.2.1)$$

Diskretna kosinusna transformacija je ključni deo mnogih segmenata nauke i inženjeringa, od audio kompresije sa gubicima (MP3) i slike (JPEG), do spektralnih metoda za numeričko rešavanje parcijalnih diferencijalnih jednačina. Korišćenje kosinusne umesto sinusne funkcije u transformaciji je jako bitno pri kompresiji signala zato što je potrebno manje kosinusnih funkcija kako bi se aproksimirao signal [1] – [3], [5] – [11], [18], [99] – [104].

Dvodimenzionalna diskretna kosinusna transformacija najčešće korišćena za kompresiju slike, podeljene na blokove, data je jednačinom [99]:

$$F(u, v) = 2 \cdot \frac{C(u) \cdot C(v)}{\sqrt{M \cdot N}} \left[\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \cdot \cos \left(\frac{(2x+1)u\pi}{2M} \right) \cdot \cos \left(\frac{(2x+1)v\pi}{2N} \right) \right], \quad (3.2.2)$$

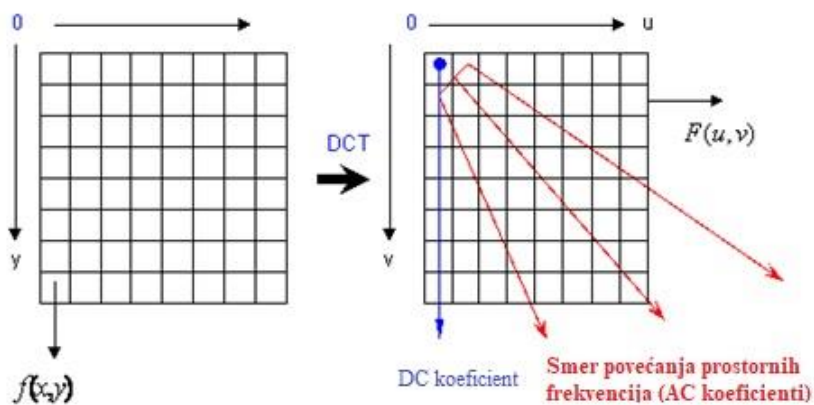
gde je M broj redova bloka, N je broj kolona bloka, $f(x,y)$ je diskretna vrednost pojedinačne tačke u bloku sa koordinatama (x, y) , a $C(u)$ i $C(v)$ su definisani izrazom:

$$C(u) = C(v) = \begin{cases} \frac{1}{\sqrt{2}}, & \text{za } u, v = 0 \\ 1, & \text{za } u, v \neq 0 \end{cases} \quad (3.2.3)$$

Inverzna diskretna kosinusna transformacija data je izrazom:

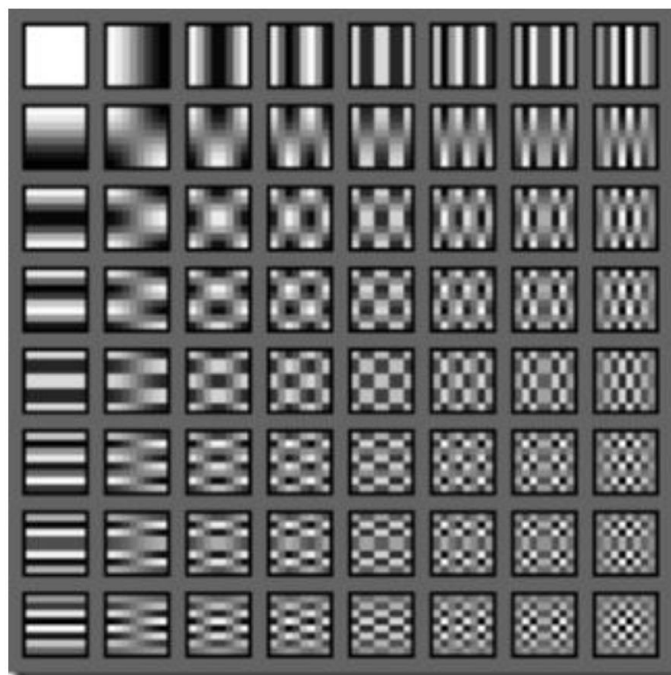
$$f(u, v) = 2 \cdot \frac{C(u) \cdot C(v)}{\sqrt{M \cdot N}} \left[\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} F(x, y) \cdot \cos\left(\frac{(2x+1)u\pi}{2M}\right) \cdot \cos\left(\frac{(2y+1)v\pi}{2N}\right) \right] \quad (3.2.4)$$

Slika 3.2.1 pokazuje ulazne i izlazne parametre DCT transformacije.



Slika 3.2.1. Ulazni i izlazni parametri DCT transformacije

Bazne funkcije DCT transformacije prikazane su na slici 3.2.2, gde su belom bojom označene pozitivne, a crnom negativne vrednosti bazne funkcije.



Slika 3.2.2. Slikovit prikaz baznih funkcija DCT transformacije

Bazna funkcija nulte frekvencije (DC vrednost) nalazi se u gornjem levom uglu slike 3.2.1, a koordinate te funkcije su (0,0). Svi ostali koeficijenti su koeficijenti kosinusa frekvencije koja odgovara toj tački prema izrazu 3.2.2. Samim tim, što je koordinata funkcije veća, to je frekvencija kosinusa koja je opisuje veća. Ovo je slučaj kretanja ka donjem desno uglu na slici 3.2.1.

Diskretna kosinusna transformacija je, kao i Hadamarova, jako bliska diskretnoj Furijeovoj transformaciji, s tim što koristi isključivo koeficijente koji su realni brojevi. Koeficijenti diskretne kosinusne transformacije su ekvivalentni koeficijentima diskretne Furijeove transformacije po periodi, s tim što imaju skoro dvostruko veću vrednost. Postoji osam standardnih DCT varijanti od kojih su četiti najčešće u upotrebi, dok je najpoznatija DCT-II, poznata i kao simetrična DCT. Transformacije bliske diskretnoj kosinusnoj transformaciji su diskretna sinusna transformacija (DST) i modifikovana diskretna kosinusna transformacija (MDCT), bazirana na DCT-u sa preklapanjem podataka [100].

Diskretna kosinusna transformacija je najčešće korišćena pri obradi signala i slike, naročito u kompresiji sa gubicima, zbog jake kompaktnosti energije koju održava. DCT generiše sekvence odmeraka koje najveću količinu informacija čuvaju u komponentama sa malom frekvencijom, što omogućava kodovanje manjim brojem bita po odmerku.

Kao što je već napomenuto, diskretna kosinusna transformacija predstavlja linearnu invertibilnu funkciju koja $f: R^N \rightarrow R^N$ (gde se sa R označava skup realnih brojeva), što znači da se skup N realnih brojeva $\{x_0, x_1, x_2, \dots, x_{N-1}\}$ mapira (transformiše) u skup N realnih brojeva $\{x_0, x_1, x_2, \dots, x_{N-1}\}$ koristeći neku od varijanti DCT-a (najčešće se koriste 4 varijante):

3.2.1 DCT-I

$$x_k = \frac{1}{2}(x_0 + (-1)^k \cdot x_{N-1}) + \sum_{n=1}^{N-2} x_n \cdot \cos\left(\frac{\pi}{N-1}nk\right), k = 0, \dots, N-1 \quad (3.2.5)$$

DCT-I (varijanta I diskretne kosinusne transformacije) je ekvivalentna diskretnoj Furijeovoj transformaciji sa $2N-2$ koeficijenata (realni brojevi) sa parnom simetrijom. Primera radi, DCT-I sa 5 koeficijenata ($N = 5, \{1,2,3,4,5\}$) je ekvivalentna DFT transformaciji sa 8 koeficijenata podeljenih sa 2 ($N = 85, \{1,2,3,4,5,4,3,2\}$). Da bi se kreirala ortogonalna matrica koeficijenata DCT-a prvi i poslednji (x_0 i x_{N-1}) koeficijent se moraju pomnožiti i podeliti sa $\sqrt{2}$. Varijanta I DCT transformacije nije definisana za $N < 2$ koeficijenata, za razliku od ostalih varijanti koje su definisane za sve pozitivne vrednosti N [101].

3.2.2 DCT-II

$$x_k = \sum_{n=0}^{N-1} x_n \cdot \cos\left(\frac{\pi}{N}\left(n + \frac{1}{2}\right)k\right), k = 0, \dots, N - 1 \quad (3.2.6)$$

Varijanta II diskretne kosinusne transformacije (DCT-II) je najčešće korišćena varijanta u praksi. DCT-II je ekvivalentna DFT transformaciji sa $4N$ odmeraka parne simetrije gde je parna simetrija definisana u odnosu na nulu. Ovo znači da DCT-II predstavlja DFT od $4N$ odmeraka y_n , gde je $y_{2n} = 0$, $y_{2n+1} = x_n$, $0 \leq n \leq N$, $y_{2N} = 0$, $y_{4N-n} = y_n$, $0 < n < 2N$. Da bi se dobila ortogonalna matrica koeficijenata DCT-II koeficijent x_0 se mora podeliti sa $\sqrt{2}$, a rezultujuća matrica pomnožiti sa ukupnim skalirajućim faktorom $\sqrt{\frac{2}{N}}$. Pri upotrebi DCT-II u kompresiji slike (JPEG) skaliranje može biti proizvoljno zato što se skalirajući faktori kombinuju sa izračunavajućim koracima podsekvenci, te se skaliranje može izabrati tako da se DCT dobija sa manje multipliciranja [100], [101].

3.2.3 DCT-III

$$x_k = \frac{1}{2}x_0 + \sum_{n=1}^{N-1} x_n \cdot \cos\left(\frac{\pi}{N}n\left(k + \frac{1}{2}\right)\right), k = 0, \dots, N - 1 \quad (3.2.7)$$

Matrica koeficijenata varijante III diskretne kosinusne transformacije je inverzna matrici koeficijenata varijante II DCT-a. Deljenjem prvog koeficijenta x_0 sa $\sqrt{2}$ i množenjem rezultujuće matrice skalirajućim faktorom $\sqrt{\frac{2}{N}}$, matrica DCT-III se transponuje u matricu DCT-II [100], [101].

3.2.4 DCT-IV

$$x_k = \sum_{n=0}^{N-1} x_n \cdot \cos\left(\frac{\pi}{N}\left(n + \frac{1}{2}\right)\left(k + \frac{1}{2}\right)\right), k = 0, \dots, N - 1 \quad (3.2.8)$$

Matrica koeficijenata DCT-IV postaje ortogonalna njenim daljim množenjem skalirajućim faktorom $\sqrt{\frac{2}{N}}$. Četvrta varijanta DCT transformacije naziva se još i modifikovanom diskretnom kosinusnom transformacijom (MDCT) [100], [101].

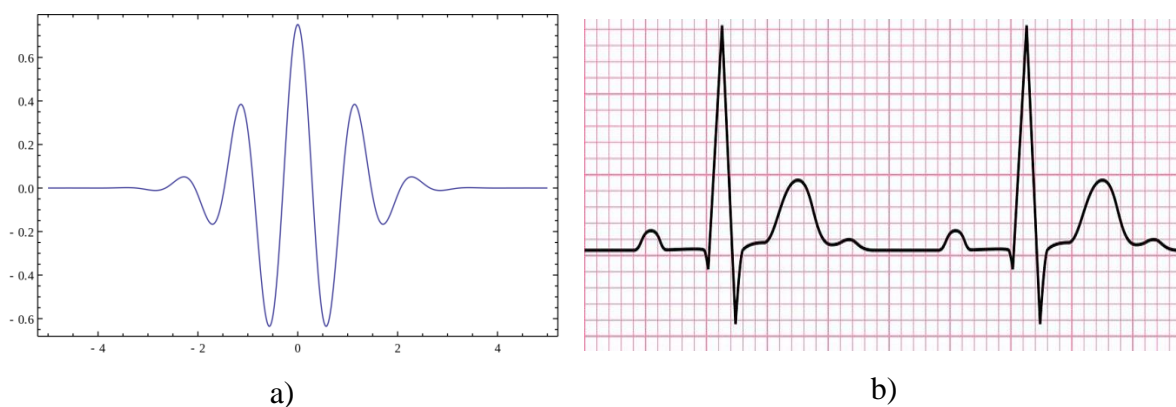
3.2.5 Inverzne transformacije

Inverzne matrice koeficijenata diskretne kosinusne transformacije dobijaju se na sledeći način:

- Inverzna matrica DCT-I dobija se množenjem matrice koeficijenata DCT-I skalarnom vrednošću normalizacionog faktora koja iznosi $\frac{2}{N-1}$,
- Inverzna matrica DCT-II dobija se množenjem matrice koeficijenata DCT-III skalarnom vrednošću normalizacionog faktora koja iznosi $\frac{2}{N}$,
- Inverzna matrica DCT-III dobija se množenjem matrice koeficijenata DCT-II skalarnom vrednošću normalizacionog faktora koja iznosi $\frac{2}{N}$,
- Inverzna matrica DCT-IV dobija se množenjem matrice koeficijenata DCT-IV skalarnom vrednošću normalizacionog faktora koja iznosi $\frac{2}{N}$.

3.3 Diskretna wavelet transformacija (DWT)

Wavelet (*mali talas*) predstavlja oscilaciju čija amplituda počinje u nuli, zatim raste i na kraju opada do nule. DWT se najlaške predstavlja (vizualizuje) kao snimljeni seizmograf ili monitoring otkucaja srca (jedan otkucaj srca na monitoru tretira se kao wavelet – *mali talas*), što je prikazano na slici 3.3.1 [1] – [3], [5] – [11], [18] – [28], [107] – [111].



Slika 3.3.1. a) primer jednog wavelet-a b) primer EKG signala sa dva wavelet-a

Wavelet-i su napravljeni tako da imaju specifična svojstva pogodna pri obradi signala, najpre zato što korišćenjem konvolucionih tehnika wavelet-i u kombinaciji sa poznatim delovima oštećenog signala mogu “izvući” informacije iz nepoznatih delova. Komplementarni wavelet-i se koriste za dekompoziciju (deljenje) podataka bez preklapanja, što proces dekompozicije čini matematički reverzibilnim (povratnim u početno stanje). Naime, setovi komplementarnih

wavelet-a su korisni u algoritmima za kompresiju-dekompresiju u kojima je cilj povratak celokupne korisne informacije iz kompresovanog signala sa minimalnim gubicima [1] – [3], [5] – [11], [18] – [28], [107] – [112].

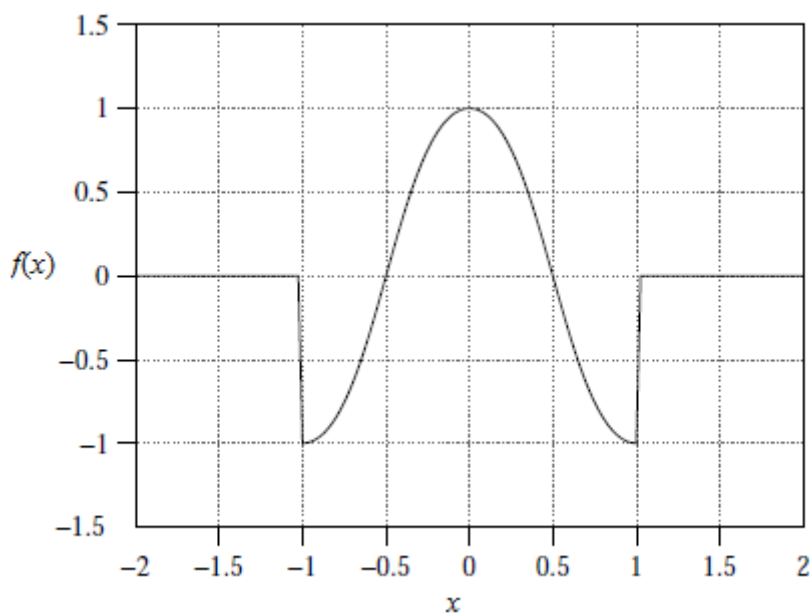
Praktična primena wavelet kompresionih šema je jako slična šemama podopsežnog kodovanja jer se koriste banke filtara za dekompoziciju signala. Izlazi filtara se odmeravaju (*downsample-uju*), kvantuju, koduju i tako šalju kroz komunikacioni kanal, dok se na prijemnoj strani signali odmeravaju (*upsample-uju*), dekvantuju i dekoduju.

Wavelet-i se mogu prikazati kao proste (*single*) funkcije, dok se sve ostale funkcije wavelet-a dobijaju menjanjem veličine prostih funkcija, odnosno skaliranjem prostih funkcija. Skaliranje prostih funkcija se naziva majkom wavelet-a (*mother wavelet*). Skaliranje proste funkcije

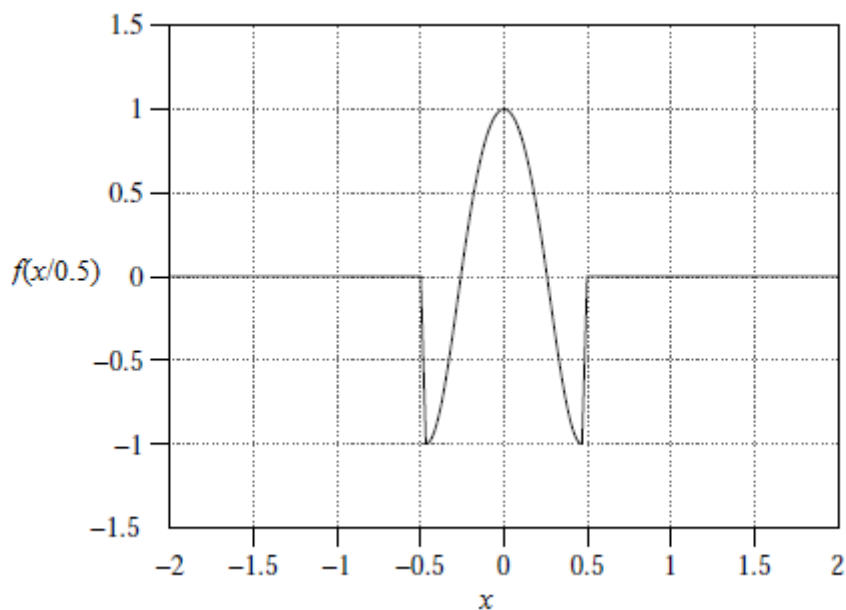
$$f(x) = \begin{cases} \cos(\pi x), & -1 \leq x \leq 1 \\ 0, & \text{drugde} \end{cases} \quad (3.3.1)$$

prikazane na slici 3.3.2, može se ostvariti menjanjem parametra x sa x/a , gde je a parametar skaliranja, čime se za $a = 0.5$ dobija skalirana funkcija $f\left(\frac{x}{0.5}\right)$ prikazana na slici 3.3.3:

$$f\left(\frac{x}{0.5}\right) = \begin{cases} \cos\left(\pi \frac{x}{0.5}\right), & -1 \leq \frac{x}{0.5} \leq 1 \\ 0, & \text{drugde} \end{cases} = \begin{cases} \cos(\pi 2x), & -\frac{1}{2} \leq x \leq \frac{1}{2} \\ 0, & \text{drugde} \end{cases}. \quad (3.3.2)$$



Slika 3.3.2. Prosta funkcija $f(x)$

Slika 3.3.3. Skalirana funkcija $f(x/0.5)$

Skaliranje će, kao i prostu funkciju $f(x)$, promeniti i normu funkcije $f(x)$:

$$\|f(x)\|^2 = \int_{-\infty}^{\infty} f^2(x) dx, \quad (3.3.3)$$

i to kao:

$$\left\| f\left(\frac{x}{a}\right) \right\|^2 = \int_{-\infty}^{\infty} f^2\left(\frac{x}{a}\right) dx = a \int_{-\infty}^{\infty} f^2(t) dt, \quad (3.3.4)$$

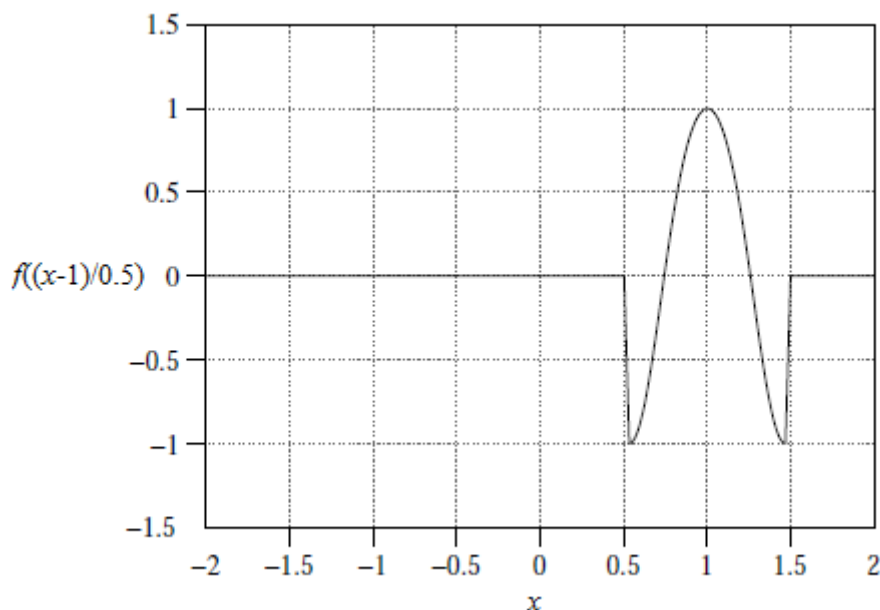
pri čemu je uvedena smena $t = x/a$ tako da se dobija:

$$\left\| f\left(\frac{x}{a}\right) \right\|^2 = a \cdot \|f(x)\|^2. \quad (3.3.5)$$

Kako bi norme proste i skalirane funkcije bile identične, norma skalirane funkcije množi se sa $\frac{1}{\sqrt{a}}$

Skalirana funkcija $f\left(\frac{x}{0.5}\right)$ može biti translirana u levu ili desnu stranu po x osi (slika 3.3.4) za vrednost b zamenom x sa $x-b$ ili $x+b$, te se dobija:

$$\begin{aligned} f\left(\frac{x-1}{0.5}\right) &= \begin{cases} \cos(\pi 2(x-1)), & -\frac{1}{2} \leq x-1 \leq \frac{1}{2} \\ 0, & \text{drugde} \end{cases} \\ &= \begin{cases} \cos(\pi 2(x-1)), & \frac{1}{2} \leq x \leq \frac{3}{2} \\ 0, & \text{drugde} \end{cases} \end{aligned} \quad (3.3.6)$$



Slika 3.3.4. Translirana skalirana funkcija $f((x-1)/0.5)$

3.3.1 Implementacija wavelet-a korišćenjem banke filtara

Filteri predstavljaju jednu od osnovnih grupa alata pri obradi signala. Služe za izdvajanje frekvencijskih grupa iz signala, odnosno za izdvajanje svih komponenti čije frekvencije pripadaju unapred zadatom opsegu. Filteri su definisani koeficijentima koji predstavljaju koeficijente diletacione jednačine, tako da se analizom digitalnih filtara dolazi do uslova koje treba da zadovolje koeficijenti diletacione jednačine da bi wavelet-i imali željena svojstva [1] – [3], [5] – [11], [18] – [28], [107] – [112].

Banke filtara predstavljaju skup filtara, pri čemu se razlikuju banke filtara analize i banke filtara sinteze. Banke filtara analize razlažu ulazni signal na frekvencijske grupe, dok banke filtara za sintezu od tih razloženih frekvencijskih grupa rekonstruišu ulazni signal. Najjednostavnija banka filtara sastoji se od dva filtra (jedan niskofrekvencijski i jedan visokofrekvencijski) i naziva se dvokanalnom bankom. Dvokanalnom bankom analize ulazni signal se razlaže u dve sekvence sa različitim frekvencijama, što ima za posledicu da se te sekvence mnogo efikasnije mogu kompresovati, kodovati i čuvati u odnosu na ulazni signal. Na prijemnoj strani se dvokanalnom bankom sinteze te sekvence sjedinjuju i time se rekonstruiše ulazni signal, pri čemu za rekonstrukciju ulaznog signala nije potrebno čuvati sve komponente sekvenci već samo parne komponente izlaza filtara analize. Analogno, kod banki filtara koje se sastoje od M filtara potrebno je, za rekonstrukciju, čuvati samo M -te komponente svake od sekvenci dobijenih analizom. Time se dobija da je ukupna dužina izlaznog signala jednaka dužini ulaznog signala [112].

Posmatrajući vremenski domen, filter predstavlja linearni operator invarijantan u vremenu, određen signalom h , čiji elementi $h(0)$, $h(1)$, $h(2)$, ... predstavljaju koeficijente filtra. Konvolucijom ulaznog signala $x(n)$ i signala h dobija se izlazni signal $y(n)$ [1] – [3], [5] – [11], [18] – [28], [107] – [112]:

$$y = h \cdot x, \quad y(n) = \sum_k h(k) \cdot x(n - k). \quad (3.3.7)$$

Matematički, elementi konvolucije predstavljaju koeficijente polinoma dobijenog množenjem dva polinoma, čiji su koeficijenti elementi signala h i x (za primer dva kvadratna polinoma):

$$\begin{aligned} H(t) \cdot X(t) &= (h(0) + h(1)t + h(2)t^2) \cdot (x(0) + x(1)t + x(2)t^2) = \\ &= h(0)x(0) + (h(0)x(1) + h(1)x(0))t \\ &\quad + (h(2)x(0) + h(1)x(1) + h(0)x(2))t^2 \\ &\quad + (h(2)x(1) + h(1)x(2))t^3 + h(2)x(2)t^4. \end{aligned} \quad (3.3.8)$$

U opštem slučaju, ulazni signal x je beskonačan dok je signal h dimenzije $(N+1)$ te se konvolucija ova dva signala predstavlja kao signal određen koeficijentima proizvoda polinoma:

$$\begin{aligned} H(t) \cdot X(t) &= (h(0) + h(1)t + h(2)t^2 + \dots + h(N)t^N) \\ &\quad \cdot (\dots + x(-1)t^{-1} + x(0) + x(1)t + x(2)t^2 + \dots). \end{aligned} \quad (3.3.9)$$

Konvolucija dva signala se može izračunati i množenjem signala x (vektora koji je definisan signalom x) i cikličnom matricom $F = \{h_{i-j}\}$ (Toeplitzovom matricom) [111]:

$$F = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & h(0) & 0 & 0 & \cdot \\ \cdot & h(1) & h(0) & 0 & \cdot \\ \cdot & h(2) & h(1) & h(0) & \cdot \\ \cdot & h(3) & h(2) & h(1) & h(0) \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}, \quad (3.3.10)$$

ili postmatrano preko polinoma:

$$\begin{aligned} y(0) &= h(N)x(-N) + \dots + h(1)x(-1) + h(0)x(0) \\ &\quad \vdots \\ y(n) &= h(N)x(n - N) + \dots + h(1)x(n - 1) + h(0)x(n) \\ &\quad \vdots \end{aligned} \quad (3.3.11)$$

Wavelet-i se mogu predstaviti i korišćenjem kauzalnih filtera (uzročni filtri) za koje važi da je $h(k) = 0, k < 0$. Ovaj uslov znači da izlazni signal ne može zavisiti od ulaznog jer bi se u izlaznom $y(n)$ signalu pojavio sabirak za koji važi:

$$h(k)x(n + |k|), k < 0, \quad (3.3.12)$$

pa se matrica ovog F kauzalnog filtra naziva i donje trougaonom matricom [111]:

$$F = \begin{bmatrix} h(0) & h(N-1) & h(N-2) & \cdots & \cdot & h(1) \\ h(1) & h(0) & h(N-1) & \cdots & 0 & h(2) \\ h(2) & h(1) & h(0) & \cdots & 0 & h(3) \\ \vdots & \vdots & \vdots & \ddots & 0 & \vdots \\ \vdots & \vdots & \vdots & \vdots & h(0) & \vdots \\ h(N-1) & h(N-2) & h(N-3) & \cdots & \cdot & h(0) \end{bmatrix}, \quad (3.3.13)$$

Dva najjednostavnija filtra su filtar za usrednjavanje i filtar za razliku. Filtar za usrednjavanje određuje odmerak izlaznog signala na taj način što usrednjuje vrednosti dva uzastopna odmerka ulaznog signala:

$$y(n) = \frac{1}{2}x(n) + \frac{1}{2}x(n-1), \quad n = \cdots, -1, 0, 1, \dots \quad (3.3.14)$$

pa se vidi da su koeficijenti filtra $h(0) = h(1) = \frac{1}{2}$. Ovaj filtar predstavlja kombinaciju identičkog operatora (koji daje izlaz jednak ulazu) i operatora kašnjenja, čiji je izlaz jednak ulazu u prethodnom trenutku. Jednačina (3.3.14) može biti predstavljena u matricnom obliku:

$$y = F_0 x, \quad \begin{pmatrix} y(-1) \\ y(0) \\ y(1) \\ \cdot \end{pmatrix} = \begin{pmatrix} \cdot & & & 0 \\ 1/2 & 1/2 & & \\ & 1/2 & 1/2 & \\ & & 1/2 & 1/2 \\ 0 & & & \cdot \end{pmatrix} \begin{pmatrix} x(-1) \\ x(0) \\ x(1) \\ \cdot \end{pmatrix} \quad (3.3.15)$$

Koeficijenti dati po glavnoj dijagonali matrice F_0 predstavljaju koeficijente filtra identičnosti, dok koeficijenti na dijagonali ispod glavne dijagonale predstavljaju koeficijente filtra kašnjenja.

Odziv filtra za usrednjavanje, u frekvencijskom domenu, dat je sa:

$$h_0(\omega) = \frac{1}{2}e^0 + \frac{1}{2}e^{-i\omega} = \frac{1}{2}(1 + e^{-i\omega}) = \cos \frac{\omega}{2} e^{-i\omega/2}. \quad (3.3.16)$$

Ukoliko ulazni signal sadrži samo jednu frekvenciju, za izbor $\omega=0$, ulazni signal je konstantan a na osnovu formule (3.3.16) vidi se da filtar za usrednjavanje ne menja konstantni signal. U slučaju izbora $\omega=\pi$, ulazni signal će oscilovati u vremenu, dok će filtar za usrednjavanje potpuno prigušiti njegove maksimalne frekvencije (svi izlazni odmerci su nula). Na osnovu navedenog može se zaključiti da filtar za usrednjavanje niske frekvencije ne menja uopšte, dok visoke frekvencije potpuno ili delimično prigušuje. Ovaj filtar spada u grupu niskofrekvencijskih filtara kojima se iz signala izdvajaju harmonici niskih frekvencija [107] – [112].

Filtar za razliku određuje odmerak izlaznog signala tako što usrednjuje razliku dva uzastopna odmerka ulaznog signala [112]:

$$y(n) = \frac{1}{2}x(n) - \frac{1}{2}x(n-1), \quad n = \dots, -1, 0, 1, \dots \quad (3.3.17)$$

pa se vidi da su koeficijenti filtra $h(0) = \frac{1}{2}$, $h(1) = -\frac{1}{2}$. Kao i filter za usrednjavanje, i filter razlike predstavlja kombinaciju operatora identičnosti i operatora kašnjenja:

$$\text{filter za razliku} = \frac{1}{2}(\text{identičnost}) - \frac{1}{2}(\text{kašnjenje}). \quad (3.3.18)$$

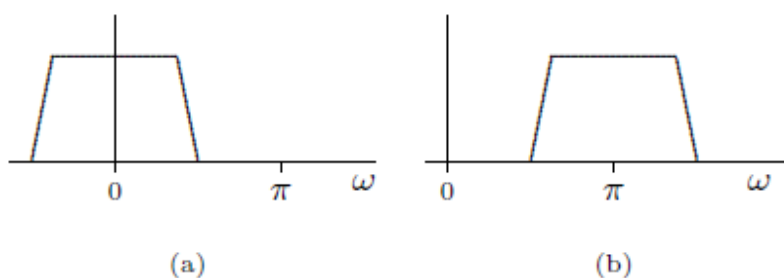
Jednačina (3.3.17) se može prikazati i u matricnom obliku:

$$y = F_1 x, \quad \begin{pmatrix} y(-1) \\ y(0) \\ y(1) \\ \vdots \end{pmatrix} = \begin{pmatrix} -1/2 & 1/2 & & & 0 \\ & -1/2 & 1/2 & & \\ & & -1/2 & 1/2 & \\ & & & -1/2 & 1/2 \\ 0 & & & & \end{pmatrix} \begin{pmatrix} x(-1) \\ x(0) \\ x(1) \\ \vdots \end{pmatrix}, \quad (3.3.19)$$

gde je F_1 matrica koeficijenata filtra za razliku. Predstavljanjem odziva filtra za razliku u frekvencijskom domenu dobija se :

$$h_0(\omega) = \frac{1}{2}e^0 - \frac{1}{2}e^{-i\omega} = \frac{1}{2}(1 - e^{-i\omega}) = \sin \frac{\omega}{2} e^{-i\omega/2}. \quad (3.3.20)$$

Analizom slučaja kada ulazni signal sadrži samo jednu frekvenciju, za izbor $\omega=0$, filter za razliku poništava niskofrekvencijske odmerke ulaznog signala, dok za $\omega=\pi$, visokofrekvencijske odmerke ne menja, tj oni ostaju isti. Na osnovu date analize zaključuje se da filter za razliku odmerke sa niskom frekvencijom potpuno ili delimično prigušuje dok odmerke visoke frekvencije ne menja uopšte ili menja vrlo malo, te se može reći da filter za razliku spada u grupu visokofrekvencijskih filtera. Na slici 3.3.5 prikazani su idealni niskofrekvencijski i visokofrekvencijski filteri i na osnovu nje se jasno vidi koje frekvencije se prigušuju a koje ne menjaju.



Slika 3.3.5. Idealni a) niskofrekvencijski filter b) visokofrekvencijski filter

Niskofrekvencijski filter za usrednjavanje i visokofrekvencijski filter za razliku nisu invertibilni svaki za sebe jer oba filtera transformišu određeni signal u nulu te se zbog toga ne može izvršiti rekonstrukcija ulaznog signala na osnovu izlaznog. Ipak, ukoliko su filteri

posmatrani u paru (niskofrekvencijski filter za usrednjavanje i visokofrekvencijski filter za razliku), bilo koji ulazni signal može biti rekonstruisan. Za kombinaciju ovih filtera kaže se da su oni filteri sa ogledalskom simetrijom (*QMR – quadrature mirror filters*) jer su jedan drugome lik u ogledalu [1] – [3], [5], [7], [11], [18], [107] – [112]. Rad ovih filtera može se opisati na sledeći način:

Niskofrekvencijski filter usrednjuje signal i time izravjava njegove neravnine koje predstavljaju visokofrekvencijske komponente pa ih ovaj filter prigušuje za sve frekvencije bliske $\omega=\pi$. Visokofrekvencijski filter pojačava razlike u signalu (ističe ih) dok prigušuje (uklanja) niskofrekvencijske komponente za sve frekvencije bliske $\omega=0$.

Posmatrajući wavelet kao kombinaciju niskofrekvencijskog i visokofrekvencijskog filtra, koeficijenti takvog wavelet-a se mogu tretirati i kao koeficijenti matrice modifikovane Hadamarove transformacije H_2 , koja se praktično primenjuje u ovoj disertaciji.

$$y = (F_0 \otimes F_1)x, \quad y(n) = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} x(n) = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} x(n), \quad (3.3.21)$$

gde kombinacija matrica F_0 i F_1 predstavlja matricu modifikovane Hadamarove transformacije [1] – [3], [5], [7], [11], [18], [107] – [112].

4 Konstrukcija logaritamskih kvantizera za adaptivno transformaciono kodovanje govornog signala

Fokus ovog poglavlja predstavlja praktična konstrukcija logaritamskih kvantizera primenom tehnika adaptacije i transformacionog kodovanja u slučaju da je na ulaz sistema doveden govorni signal. Postavljen je cilj da se projektuje kodna šema koja obezbeđuje visok

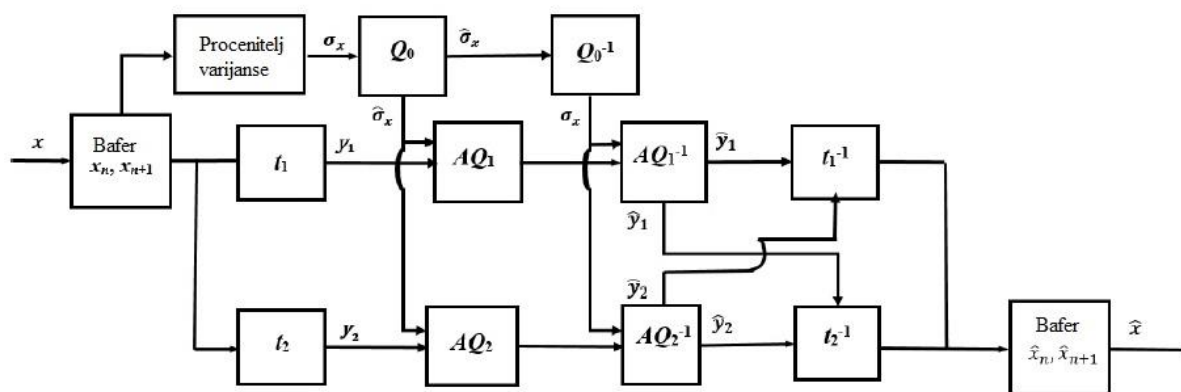
kvalitet kodovanja, odnosno da se sa što manje bita po odmerku korišćenih za kodovanje obezbedi što viši SQNR (*Signal-to-Quantization-Noise-Ratio*). Kao ulazni test signali za eksperimente prikazane u ovoj disertaciji korišćeni su govorni signali odmereni na 8 kHz i 16 kHz (širokopojasna ekstenzija), snimljeni u laboratoriji za Akustiku na Elektronskom fakultetu u Nišu. U ovom poglavlju prikazane su kodne šeme projektovane za slučajeve kada je na ulaz sistema doveden kontinualni govorni signal i kada je taj kontinualni signal prvo diskretizovan [74]. Sve prikazane kodne šeme implementiraju neuniformne kvantizere sa μ logaritamskim zakonom kompresije. Kodne šeme prikazane u ovom poglavlju uključuju različite tipove transformacionog kodovanja koji su detaljno opisani u prethodnom poglavlju, uključujući Hadamarovu transformaciju H_1 i H_2 (modifikovane matrice transformacionih koeficijenata 2×2 i 4×4) [59], [71], diskretnu kosinusnu transformaciju matrice koeficijenata 3×3 i 4×4 i diskretnu wavelet transformaciju matrice koeficijenata 2×2 koja se svodi na Hadamarovu transformaciju H_2 pa za nju nije izvođen eksperiment. Adaptacija kvantizera na signal u svim kodnim šemama prikazanim u ovoj disertaciji izvršena je tehnikom adaptacije unapred, detaljno opisana u poglavlju 2.3. Kodne šeme implementirane koristeći modifikovanu Hadamarovu transformaciju prikazane su u radovima publikovanim u naučnim časopisima sa impakt faktorom [59], [71].

4.1 Šeme za kodovanje kontinualnog govornog signala konstruisane koristeći modifikovanu Hadamarovu transformaciju

U ovoj sekciji prikazane su kodne šeme za kodovanje kontinualnog govornog signala implementirane koristeći modifikovanu Hadamarovu transformaciju H_1 i H_2 matricama koeficijenata veličine 2×2 i 4×4 , respektivno.

Jedna od prikazanih šema za kodovanje implementirana je koristeći prostu (modifikovanu) Hadamarovu transformaciju čija je matrica koeficijenata 2×2 i tehniku adaptacije unapred predložene u radu [59]. Prikazana šema se prilagođava (adaptira) varijansi ulaznog signala čime se obezbeđuje visoka efikasnost iskorišćenja propusnog opsega dok se implementacijom modifikovane Hadamarove transformacije signal dekomponuje (deli) u podsekvence čije su karakteristike signala predvidljivije, pa se upotreba predložene kodne šeme pokazuje pogodnom za procesiranje signala. Kodna šema prikazana na slici 4.1.1 koncipirana je tako da se ulazni signal najpre deli na podsekvence primenom transformacije obezbeđujući uži propusni opseg, nakon čega se vrši adaptacija kvantizera na varijansu signala u sekvencama,

što rezultuje korišćenjem manje bita po odmerku za kodovanje signala. U radu [59] kao adaptaciona tehnika korišćena je adaptacija unapred zbog veće tolerancije na greške prilikom prenosa signala kroz komunikacioni kanal, što je česta pojava u VoIP sistemima. Takođe, kroz literaturu je pokazano da adaptacija unapred obezbeđuje do 1 dB viši SQNR u odnosu na adaptaciju unazad.



Slika 4.1.1. Šema za kodovanje govornog signala implementirana koristeći prostu transformaciju i adaptaciju unapred

Sa slike 4.1.1 se vidi da ulazni signal (kontinualni) dolazi do bafera u kome se deli na frejmove veličine M odmeraka. Nad signalom u frejmu na izlazu iz bafera primenjena je prosta transformacija koja ga deli na dve grane koje se nezavisno koduju. Pošto se vrši adaptacija kvantizera na signal tehnikom adaptacije unapred, varijansa signala u frejmu mora biti izračunata, kodovana i prosleđena prijemniku kao dodatna informacija, za šta je iskorišćen log-uniformni kvantizer detaljno opisan u sekciji 2.3.2. Signali u granama kvantovani su logaritamskim neuniformnim kvantizerima konstruisanim korišćenjem μ zakona kompresije [1] – [3], [5], [7], [11], [18] – [28], [59], [71].

Prosta transformacija primenjena na kodnu šemu sa slike 4.1.1 predstavlja modifikovanu Hadamarovu transformaciju H_1 , matrice koeficijenata veličine 2×2 . Hadamarova transformacija detaljno je opisana u prethodnom poglavlju a njena modifikacija primenjena na ovu kodnu šemu ogleda se u tome što se kao faktor normalizacije ne uzima standardno $\frac{1}{\sqrt{2}}$ već $\frac{1}{2}$. Prosta transformacija deli ulazni signal x na signale y_1 i y_2 i definisana je sledećim izrazima:

$$\begin{aligned} y_1 &= \frac{x_n + x_{n+1}}{2}, \\ y_2 &= \frac{x_n - x_{n+1}}{2}, \end{aligned} \quad (4.1.1)$$

gde su x_n i x_{n+1} dva susedna odmerka signala u frejmu, dok y_1 i y_2 predstavljaju transformisane signale. Signali y_1 i y_2 dobijeni primenom proste transformacije imaju varijanse σ_1^2 i σ_2^2 , respektivno, koji zavise od varijanse ulaznog signala u frejmu σ_x^2 i koeficijenta korelacije signala u frejmu ρ [1] – [3], [5] – [8], definisane sledećim jednačinama:

$$\begin{aligned}\sigma_1^2 &= \frac{\sigma_x^2}{2}(1 + \rho), \\ \sigma_2^2 &= \frac{\sigma_x^2}{2}(1 - \rho).\end{aligned}\tag{4.1.2}$$

Transformisane sekvence signala y_1 i y_2 predstavljaju nezavisne signale i koduju se posebno koristeći kvazilogaritamske kvantizere koji se prilagođavaju signalu primenom tehnike adaptacije unared. Kvantizeri AQ_1 i AQ_2 sa slike 4.1.1 konstruisani su koristeći μ logaritamski zakon kompresije koji je definisan sledećom jednačinom [1], [7], [42], [66], [67]:

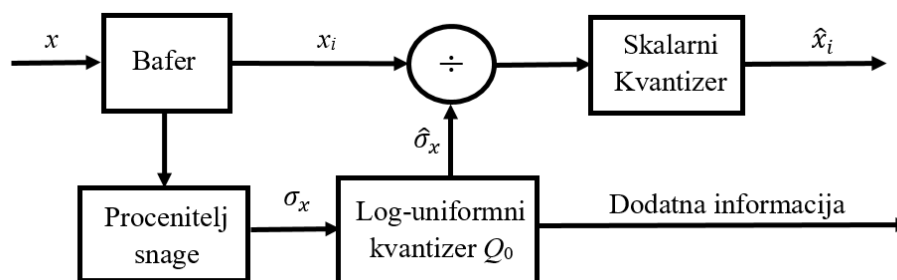
$$c_\mu(x) = \frac{x_{\max}}{\ln(1 + \mu)} \ln \left(1 + \mu \frac{|x|}{x_{\max}} \right) \text{sgn}(x),\tag{4.1.3}$$

gde je $|x| \leq x_{\max}$, dok x_{\max} predstavlja maksimalnu amplitudu opterećenja kvantizera a μ je faktor kompresije. Pragovi odluka x'_i i reprezentacioni nivoi kvantizera y'_i se, prema logaritamskom zakonu, dobijaju koristeći sledeće jednačine:

$$\begin{aligned}x'_i &= \frac{x_{\max}}{\mu} \left(\exp \left(\frac{2i}{N} \ln(1 + \mu) \right) - 1 \right), \\ y'_i &= \frac{x_{\max}}{\mu} \left(\exp \left(\frac{2(i - 0.5)}{N} \ln(1 + \mu) \right) - 1 \right),\end{aligned}\tag{4.1.4}$$

gde i uzima vrednosti $i = 1, 2, \dots, N$ (N predstavlja broj nivoa kvantizera).

Kao što je već napomenuto, da bi se ostvario veći kvalitet rekonstrukcije signala na prijemu, kvantizeri se adaptiraju signalu koristeći tehniku adaptacije unapred [1] – [3], [5] – [11], [18] – [28], [42], [56] – [59]. Tehnika adaptacije unapred, njen deo implementiran u predajniku, objašnjena detaljno u prethodnim poglavljima, prikazana je na slici 4.1.2.



Slika 4.1.2. Tehnika adaptacije unapred u predajniku

Log-uniformni kvantizer, kao što je objašnjeno ranije, konstruisan je kao uniformni kvantizer čiji su reprezentacioni nivoi i pragovi odluke u logaritamskom domenu, pa zbog toga i varijansa koju izračunava procenitelj varijansi i prosleđuje log-uniformnom kvantizeru mora biti u logaritamskom domenu [42]:

$$20\log_{10}(\hat{\sigma} = \hat{\sigma}_k) = 20\log_{10}(\sigma_{\min}) + \frac{2k-1}{2}\Delta^{\text{lu}}, \quad (4.1.5)$$

gde je Δ^{lu} širina kvanta definisana sa:

$$\Delta^{\text{lu}} = \frac{20\log_{10}\left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)}{N_g}. \quad (4.1.6)$$

Brojač k uzima vrednosti $1, 2, \dots, N_g$, gde je N_g broj nivoa log-uniformnog kvantizera [67].

U radu [21] gde je predložena kodna šema sa slike 4.1.1 log-uniformni kvantizer konstruisan je za male i srednje bitske brzine, tako da broj kvantizacionih nivoa uzima vrednosti 2, 4, 8 i 16. Procenitelj varijanse ka log-uniformnom kvantizeru šalje vrednost standardne devijacije signala u frejmu, log-uniformni kvantizer je kvantuje i prosleđuje ka adaptivnim kvantizerima AQ_1 i AQ_2 , kao i ka prijemniku kao dodatnu informaciju. Kvantovana vrednost standardne devijacije se koristi da bi se dobila maksimalna amplituda opterećenja kvantizera $x_{\max1}$ i $x_{\max2}$. Ovim se kvantizeri u granama prilagođavaju signalu za svaki frejm. Vrednosti maksimalne amplitude kvantizera AQ_1 i AQ_2 , $x_{\max1}$ i $x_{\max2}$ respektivno, izračunavaju se koristeći sledeće formule [59]:

$$\begin{aligned} x_{\max1} &= x_{\text{ml}} \cdot \hat{\sigma} \cdot \sqrt{\left(\frac{1+\rho}{2}\right)}, \\ x_{\max2} &= x_{\text{ml}} \cdot \hat{\sigma} \cdot \sqrt{\left(\frac{1-\rho}{2}\right)}, \end{aligned} \quad (4.1.7)$$

gde je $\hat{\sigma}$ kvantovana vrednost standardne devijacije signala u frejmu, ρ je koeficijent korelacije signala u frejmu, dok x_{ml} predstavlja vrednost maksimalne amplitude opterećenja kvantizera ukoliko se na njegov ulaz dovede signal modelovan Laplasovom funkcijom raspodele gustine verovatnoće za jediničnu varijansu, koji se dobija iz formule [66], [68], [69]:

$$x_{\text{ml}} = \frac{1}{\sqrt{2}} \ln\left(\frac{3\mu N_{sr}^2}{\ln^2(1+\mu)}\right), \quad (4.1.8)$$

gde N_{sr}^2 predstavlja prosečni broj kvantizacionih nivoa kvantizera AQ_1 i AQ_2 , dok je μ faktor kompresije.

Objektivna mera kvaliteta kodovanja prikazana je, kao što je već napomenuto u prethodnim poglavljima, kroz odnos snage signala i kvantizacionog šuma (SQNR) koji se dobija usrednjavanjem SQNR-a svih frejmova:

$$\text{SQNR}_{sr}[\text{dB}] = \frac{1}{N_f} \sum_{j=1}^{N_f} \text{SQNR}_j, \quad (4.1.9)$$

gde se SQNR po frejmovima dobija kao:

$$\text{SQNR}_j[\text{dB}] = 10 \cdot \log_{10} \left(\frac{\sigma^2}{D_j} \right). \quad (4.1.10)$$

U jednačini (4.1.10) D_j predstavlja distorziju j -tog frejma, koja se dobija iz izraza:

$$D_j = \sum_{i=1}^M (x_i - \hat{x}_i)^2, \quad (4.1.11)$$

gde su \hat{x}_i označene kvantovane vrednosti odmeraka poslate kao prijemniku [1], [7].

Prosečna bitska brzina kodne šeme data je izrazom:

$$R_{av} = R + R_{log}, \quad (4.1.12)$$

gde je R srednja vrednost bitskih brzina adaptivnih kvantizera AQ_1 i AQ_2 , a R_{log} broj bitova po odmerku potrebnih za kodovanje varijanse u log-uniformnom kvantizeru koji se izračunava:

$$R_{log} = \frac{R_0}{M}. \quad (4.1.13)$$

Za potrebe eksperimenta korišćen je govorni test signal odmeren na 8 kHz trajanja oko 10 sekundi čija varijansa iznosi $\sigma^2=0.9086$. Kako bi dobit korišćenja transformacionog kodovanja pri procesiranju govornog signala bila jasnija, eksperimentalni rezultati dobijeni korišćenjem kodne šeme sa slike 4.1.1 upoređeni su sa odgovarajućim rezultatima kodne šeme koja ne implementira transformaciono kodovanje pre adaptacije unapred, kao i sa standardnim PCM (*Pulse Code Modulation*) sistemom. Eksperiment je urađen za različite vrednosti veličine frejma M , $M \in \{40, 80, 160, 240\}$, male i srednje bitske brzine log-uniformnog kvantizera Q_0 , $N_0 \in \{2, 3, 8, 16\}$, faktor kompresije $\mu=16.9227$ (što je optimalna vrednost faktora kompresije za Laplasov izvor jedinične varijanse) preuzet iz [113], i različite bitske brzine adaptivnih kvantizera u granama AQ_1 i AQ_2 [59].

U tabeli 4.1.1 prikazane su vrednosti SQNR-a dobijene koristeći kodnu šemu sa slike 4.1.1 za različite vrednosti prosečne bitske brzine: $R = 5.5$ bita po odmerku ($R_1(AQ_1) = 6$ bita po odmerku, $R_2(AQ_2) = 5$ bita po odmerku), $R = 6$ bita po odmerku ($R_1(AQ_1) = 7$ bita po odmerku, $R_2(AQ_2) = 5$ bita po odmerku), $R = 7$ bita po odmerku ($R_1(AQ_1) = 8$ bita po odmerku, $R_2(AQ_2) = 6$ bita po odmerku), i sve vrednosti bitskih brzina log-uniformnih kvantizera ($R_0 = \{1, 2, 3, 4\}$)

bita po odmerku). Dobijeni rezultati su upoređeni sa rezultatima koristeći PCM sa istim bitskim brzinama i faktorom kompresije.

Tabela 4.1.1. Eksperimentalni rezultati (SQNR) dobijeni koristeći kodnu šemu sa slike 4.1.1 i adekvatno poređenje sa PCM sistemom za razne parametre kodovanja i $\mu = 16.9227$

M	N_0	SQNR [dB]	R_{sr} [bitova /odmerku]	SQNR _{PCM} [dB]	SQNR [dB]	R_{sr} [bitova /odmerku]	SQNR _{PCM} [dB]	SQNR [dB]	R_{sr} [bitova /odmerku]	SQNR _{PCM} [dB]
40	2	27.914	5.525	25.5040	33.855	6.525	32.0016	38.315	7.025	35.1832
	4	28.562	5.55		34.466	6.55		38.971	7.05	
	8	28.816	5.575		34.461	6.575		39.074	7.075	
	16	28.859	5.6		34.466	6.6		39.089	7.1	
80	2	27.925	5.5125		33.830	6.5125		38.262	7.0125	
	4	28.720	5.525		34.461	6.525		39.086	7.025	
	8	28.788	5.5375		34.460	5.5375		39.042	7.0375	
	16	28.852	5.55		34.510	6.55		39.257	7.05	
160	2	27.585	5.5062		33.594	6.5062		38.044	7.0062	
	4	28.781	5.5125		34.569	6.5125		39.231	7.0125	
	8	28.841	5.5187		34.456	6.5187		39.223	7.0187	
	16	28.887	5.525		34.574	6.525		39.281	7.025	
240	2	27.120	5.5042		32.927	6.5042		37.345	7.0042	
	4	28.660	5.5083		34.271	6.5083		38.978	7.0083	
	8	28.809	5.5125		34.288	6.5125		39.206	7.0125	
	16	28.897	5.5167		34.603	6.5167		39.223	7.0167	

Posmatrajući rezultate iz tabele 4.1.1 može se zaključiti da predložena kodna šema sa slike 4.1.1 pruža bolje performanse kodovanja u odnosu na PCM sistem sa istim parametrima (kombinacije veličine frejma, bitske brzine log-uniformnog kvantizera Q_0), tj. SQNR je čak 4.0983 dB veći u odnosu na PCM (u slučaju najkompleksnijeg dizajna iz tabele). SQNR za PCM sistem je konstantna vrednost za određenu bitsku brzinu (ne zavisi od veličine) pa za $R_{sr} = 5.5$ bitova po odmerku SQNR ima vrednost 25.5040 dB, za $R_{sr} = 6.5$ bitova po odmerku SQNR ima vrednost 32.0016 dB, dok za $R_{sr} = 7$ bitova po odmerku SQNR ima vrednost 35.1832 dB. Vidi se da kodna šema sa slike 4.1.1 obezbeđuje bolji kvalitet kodovanja u odnosu na PCM za sve kombinacije parametara kodovanja i može se zaključiti da je pogodna za kodovanje govornog signala.

Tabela 4.1.2 prikazuje poređenje SQNR-a dobijenog korišćenjem kodne šeme koja ne uključuje transformaciono kodovanje već samo tehniku adaptacije unapred i SQNR-a dobijenog korišćenjem PCM sistema za iste parametre kodovanja kao i tabela 4.1.1.

Tabela 4.1.2. Eksperimentalni rezultati (SQNR) dobijeni koristeći kodnu šemu koja ne uključuje transformaciono kodovanje i adekvatno poređenje sa PCM sistemom za razne parametre kodovanja i $\mu = 16.9227$

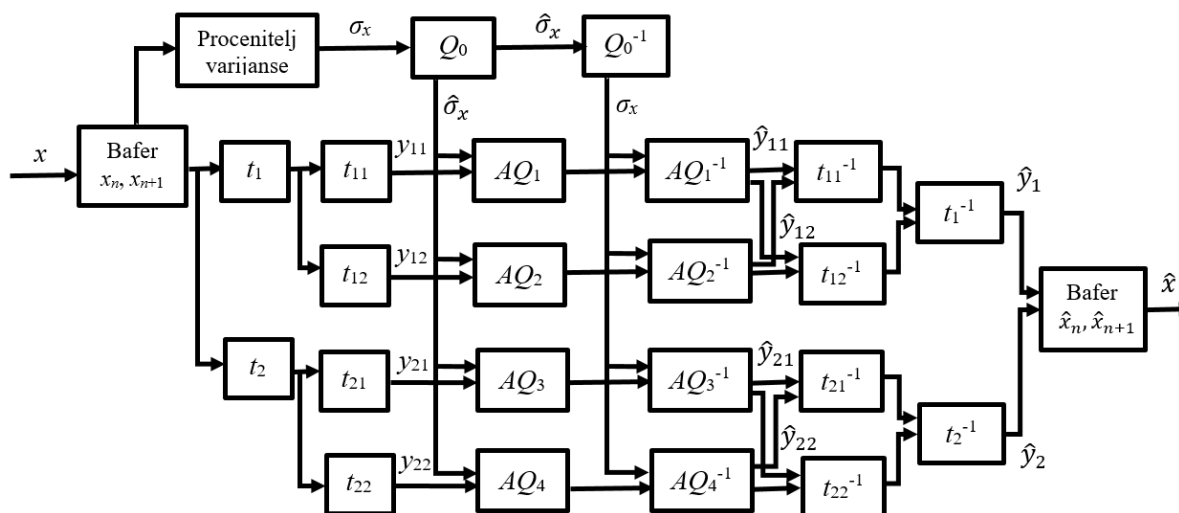
M	N_0	SQNR [dB]	R_{sr} [bitova /odmerku]	SQNR _{PCM} [dB]	SQNR [dB]	R_{sr} [bitova /odmerku]	SQNR _{PCM} [dB]
40	2	28.313	6.025	28.8200	33.920	7.025	35.1832
	4	28.952	6.05		34.800	7.05	
	8	28.953	6.075		34.808	7.075	
	16	28.978	6.1		34.827	7.1	
80	2	28.322	6.125		33.959	7.0125	
	4	29.042	6.025		34.649	7.025	
	8	29.061	6.0375		34.662	7.0375	
	16	29.086	6.05		34.816	7.05	
160	2	28.139	6.0062		33.921	7.0062	
	4	29.007	6.0125		34.557	7.0125	
	8	29.056	6.0187		34.669	7.0187	
	16	29.092	6.025		34.815	7.025	
240	2	27.885	6.0042		33.732	7.0042	
	4	29.019	6.0083		34.304	7.0083	
	8	29.096	6.0125		34.647	7.0125	
	16	29.106	6.0167		34.905	7.0167	

Iz tabele 4.1.2 može se videti da za srednju vrednost bitske brzine od $R_{sr} = 6$ bita po odmerku kodna šema bez transformacionog kodovanja pokazuje dobitak tek pri projektovanju log-uniformnog kvantizera za bitske brzine od 2, 3, 4 bita po odmerku, dok za $R_0 = 1$ bit po odmerku PCM sistem obezbeđuje bolje performanse kodovanja. Primećuje se da je dobitak u odnosu na PCM mnogo manji nego kada se u kodnu šemu implementira i transformaciono kodovanje, jer iznosi manje od 0.3 dB u odnosu na skoro 5 dB uvršćavanjem transformacionog kodovanja u kodnu šemu. Takođe, za srednju bitsku brzinu od 7 bita po odmerku PCM sistem pruža bolje performanse kodovanja za sve posmatrane slučajeve.

Posmatrajući tabele 4.1.1 i 4.1.2 zaključuje se da implementacijom transformacionog kodovanja u kodnu šemu obezbeđuju se daleko bolje performanse kodovanja govornog signala, čijim se uključenjem dobija preko 4 dB viši SQNR, dok ukoliko se upoređi najbolji slučaj iz tabele 4.1.1 sa log-PCM sistemom, baziranim na G.711 standardu, dobitak kodovanja je čak i do 10 dB.

Sledeća predložena kodna šema za kodovanje govornog signala predstavlja još jednu modifikaciju Hadamarove transformacije. Naime, u ovoj kodnoj šemi je, posle primene proste (modifikovane) Hadamarove transformacije H_1 i dekompozicije ulaznog signala na dve posebne grane, prosta (modifikovana) Hadamarova transformacija primenjena još jednom, tj.

svaka od dobijene dve grane je dekomponovana još na dve. Na slici 4.1.3 prikazana je kodna šema u kojoj je primenjena dvostruka modifikovana Hadamarova transformacija H_1 , što zapravo predstavlja modifikaciju Hadamarove transformacije H_2 .



Slika 4.1.3. Šema za kodovanje govornog signala implementirana koristeći dvostruku prostu transformaciju i adaptaciju unapred

Sa slike 4.1.3 se vidi da ulazni signal (kontinualni) dolazi do bafera u kome se deli na frejmove veličine M odmeraka. Nad signalom u frejmu na izlazu iz bafera primenjena je prosta transformacija koja ga deli na dve grane y_1 i y_2 . Nad signalima y_1 i y_2 se ponovo primenjuje prosta Hadamarova transformacija H_1 i time se svaki od njih dekomponuje na još po dva signala. Posle primene dvostruke proste transformacije ulazni signal x je dekomponovan na četiri signala sa predvidljivijim karakteristikama koji se zasebno koduju, i to su signali y_{11} , y_{12} , y_{21} i y_{22} . S obzirom da signali u granama imaju različite karakteristike, moraju se kvantovati kvantizerima različitih bitskih brzina da bi se ostvario maksimalni kvalitet kodovanja. Kako se vrši adaptacija kvantizera na signal tehnikom adaptacije unapred, varijansa signala u frejmu mora biti izračunata, kodovana i prosleđena prijemniku kao dodatna informacija, za šta je iskorišćen log-uniformni kvantizer detaljno opisan u sekciji 2.3.2. Signali u granama kvantovani su, kao i u prethodnoj šemi, logaritamskim neuniformnim kvantizerima konstruisanim korišćenjem μ logaritamskog zakona kompresije, čija je kompresorska funkcija prikazana u formuli (4.1.3) [1] – [3], [5] – [11], [18] – [28], [59], [65]. Vrednosti pragova odluke i reprezentacionih nivoa adaptivnih kvantizera u granama izračunavaju se prema formuli (4.1.4), kao i u prethodnoj šemi. Vrednosti odmeraka signala u granama y_{11} , y_{12} , y_{21} i

y_{22} mogu se izračunati dvostukom primenom formule (4.1.1) dok se njihove varijanse dobijaju iz sledećih izraza [1], [7]:

$$\begin{aligned}\sigma_{11}^2 &= \frac{\sigma_x^2}{4}(1 + \rho)(1 + \rho_1), \\ \sigma_{12}^2 &= \frac{\sigma_x^2}{4}(1 + \rho)(1 - \rho_1), \\ \sigma_{21}^2 &= \frac{\sigma_x^2}{4}(1 - \rho)(1 + \rho_2), \\ \sigma_{22}^2 &= \frac{\sigma_x^2}{4}(1 - \rho)(1 - \rho_2),\end{aligned}\tag{4.1.14}$$

gde je σ_x^2 varijansa signala u frejmu, ρ koeficijent korelacije signala u frejmu, a ρ_1 i ρ_2 koeficijenti korelacije signala y_1 i y_2 , respektivno.

Samim tim što se vrši dvostruko transformaciono kodovanje, odmerci ulaznog signala se transformišu (mapiraju) u odmerke koji nose veću količinu informacija i odmerke koji nose manju količinu. Kodna šema poseduje četiri adaptivna kvantizera za kvantovanje signala u granama i jedan log-uniformni za kvantovanje varijanse signala u frejmu. Jako je bitno voditi računa od bit alokaciji, odnosno o broju bitova po odmerku potrebnih za kodovanje svakog kvantizera. Bit alokacija se izračunava na osnovu formule [7]:

$$R_k = R + \frac{1}{2} \log_2 \left(\frac{\sigma_k^2}{\prod_{k=1}^L (\sigma_k^2)^{\frac{1}{L}}} \right),\tag{4.1.15}$$

gde je R_k bitska brzina kvantizera u grani K , R je prosečna bitska brzina svih kvantizera u šemi, σ_k^2 predstavlja varijansu signala u grani K , dok je u imeniocu proizvod svih varijansi po granama na stepenu $1/L$, dok je L broj grana kodne šeme. Log-uniformni kvantizer za kodovanje varijanse signala (standardne devijacije) u frejmu objašnjen je ranije i prikazan na slici 4.1.2 i kroz formule (4.1.5) i (4.1.6). Kvantovana vrednost standardne devijacije koristi se kako bi se dobila maksimalna amplituda opterećenja kvantizera $x_{\max 1}$, $x_{\max 2}$, $x_{\max 3}$ i $x_{\max 4}$. Na ovaj način se kvantizeri u granama prilagođavaju signalu za svaki frejm. Distorzija i SQNR po frejmu, kao i ukupan prosečni SQNR prikazani su kroz formule (4.1.9) – (4.1.11) u prethodnoj kodnoj šemi. Vrednosti maksimalne amplitude kvantizera $AQ_1 - AQ_4$, ($x_{\max 1} - x_{\max 4}$) respektivno, izračunavaju se koristeći sledeće formule [1] – [3], [5] – [11], [18] – [28], [59]:

$$\begin{aligned}x_{\max 1} &= x_{\text{ml}} \cdot \hat{\sigma}, \\ x_{\max 2} &= x_{\text{ml}} \cdot \hat{\sigma}, \\ x_{\max 3} &= x_{\text{ml}} \cdot \hat{\sigma}, \\ x_{\max 4} &= x_{\text{ml}} \cdot \hat{\sigma},\end{aligned}\tag{4.1.16}$$

gde je x_{ml} vrednost maksimalne amplitude opterećenja kvantizera za Laplasov izvor jedinične varijanse koja se dobija na osnovu formule (4.1.8).

Za potrebe eksperimenta korišćen je govorni test signal odmeren na 16 kHz (širokopojasna ekstenzija) trajanja oko 1 minut čija varijansa iznosi $\sigma^2=0.0021$. Eksperiment je urađen za veličinu frejma $M=240$, bitsku brzinu log-uniformnog kvantizera Q_0 , $N_0=16$, vrednosti faktor kompresije $\mu=20$ i $\mu=255$ (PCM standard) [1] – [3], [5] – [11], [18] – [29], [37], [45], [46], i različite srednje bitske brzine adaptivnih kvantizera u granama AQ_1 , AQ_2 , AQ_3 i AQ_4 , $R_{sr}=5,6,7,8$ bita po odmerku. Na osnovu formule (4.1.15) određene su bitske brzine adaptivnih kvantizera AQ_1 , AQ_2 , AQ_3 i AQ_4 , i to na sledeći način: $R_1=R_{sr}+3$, $R_2=R_{sr}-1$, $R_3=R_{sr}-1$, $R_4=R_{sr}-1$.

U tabeli 4.1.3 prikazane su vrednosti SQNR-a dobijene koristeći kodnu šemu sa slike 4.1.3 za različite vrednosti prosečne bitske brzine: $R_{sr}=5$ bita po odmerku ($R_1(AQ_1)=8$ bita po odmerku, $R_2(AQ_2)=4$ bita po odmerku, $R_3(AQ_3)=4$ bita po odmerku, $R_4(AQ_4)=4$ bita po odmerku); $R=6$ bita po odmerku ($R_1(AQ_1)=9$ bita po odmerku, $R_2(AQ_2)=5$ bita po odmerku, $R_3(AQ_3)=5$ bita po odmerku, $R_4(AQ_4)=5$ bita po odmerku); $R=7$ bita po odmerku ($R_1(AQ_1)=10$ bita po odmerku, $R_2(AQ_2)=6$ bita po odmerku, $R_3(AQ_3)=6$ bita po odmerku, $R_4(AQ_4)=6$ bita po odmerku); $R=8$ bita po odmerku ($R_1(AQ_1)=11$ bita po odmerku, $R_2(AQ_2)=7$ bita po odmerku, $R_3(AQ_3)=7$ bita po odmerku, $R_4(AQ_4)=7$ bita po odmerku), i sve vrednosti bitskih brzina log-uniformnih kvantizera ($R_0=4$ bita po odmerku), i vrednosti faktora kompresije $\mu=20$ i $\mu=255$.

Tabela 4.1.3. Eksperimentalni rezultati (SQNR) dobijeni koristeći kodnu šemu sa slike 4.1.3 za razne bitske brzine i vrednosti faktora kompresije

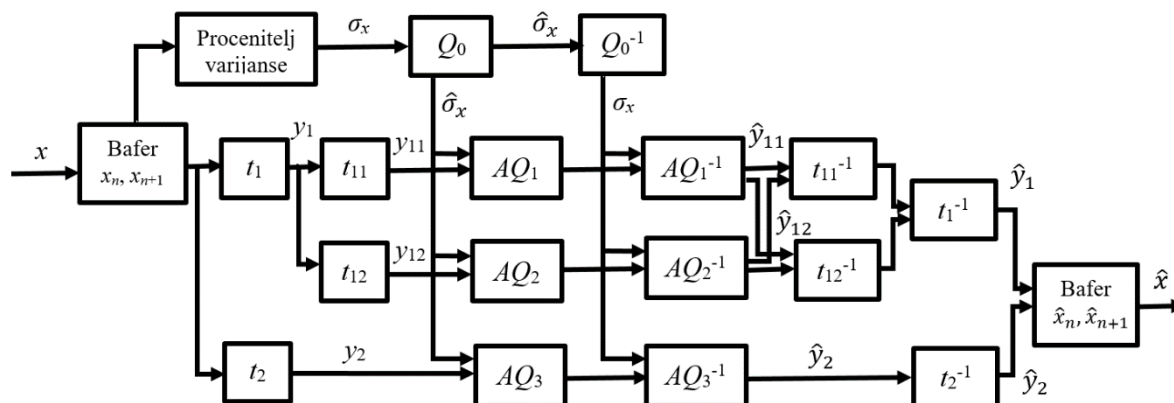
R_{sr} [bita po odmerku]	μ	SQNR[dB]	SQNR _{2 grane} ($\mu=16.9227$)[dB]	SQNR _{PCM} [dB]
5	20	30.1968	27.9027	23.4112
	255	27.1547		
6	20	36.0036	33.4459	29.2385
	255	33.1343		
7	20	41.5876	39.2283	35.1832
	255	39.0757		
8	20	47.7315	45.1963	41.0952
	255	45.3552		

U tabeli 4.1.3 prikazani su eksperimentalni rezultati kvaliteta kodovanja koristeći kodnu šemu sa slike 4.1.3. SQNR dobijen primenom pomenute kodne šeme upoređen je sa SQNR-om dobijenim u radu [59], gde je nad ulaznim signalom primenjena samo jedna prosta (modifikovana) Hadamarova transformacija H_1 i time ulazni signal dekomponovan na dve podsekvence. Ovo poređenje urađeno je u cilju prikazivanja dobitka u kvalitetu usled primene

još jedne transformacije čime se, već predvidivim karakteristikama signala y_1 i y_2 , dodaje na predvidivosti jer transformacije povećavaju korelaciju signala što podiže kvalitet kodovanja. U radu [59], kvantizeri su projektovani za optimalnu vrednost faktora kompresije koja iznosi $\mu=16.9227$, dok je eksperiment primenom kodne šeme sa slike 4.1.3 urađen za $\mu=20$, 255. Faktor kompresije $\mu=255$ preuzet je iz standarda G.711 i kvazi-logaritamske kvantizere čini robustnim zato što “širi” krivu kompresorske funkcije dok $\mu=16.9227$, izračunat u radu [68], predstavlja optimalnu vrednost i utiče na kodovanje tako što “skuplja” krivu kompresorske funkcije kvazi-logaritamskog zakona i, za optimalan izbor regiona rada kvantizera obezbeđuje maksimalni SQNR. Ovaj faktor približno je jednak faktoru $\mu=20$ tako da se poređenjem rezultata za ove dve vrednosti mogu izvući određeni zaključci.

Za $\mu=20$ kodna šema pokazuje dobitak od 2.2 do 2.5 dB u odnosu na šemu iz rada [59], gde su kvantizeri projektovani za $\mu=16.9227$, dok u odnosu na PCM sistem ($\mu=16.9227$) dobitak kodovanja iznosi 6.5 – 6.8 dB, u zavisnosti od bitske brzine. Faktor kompresije $\mu=255$ pruža niži SQNR u odnosu na $\mu=20$ jer su kvazi-logaritamski kvantizeri robustniji za ovu vrednost. Naime, cilj dvostruke primene proste Hadamarove transformacije H_1 nad ulaznim signalom bio je obezbeđivanje signala sa predvidljivijim karakteristikama za kodovanje u odnosu na samo jednu primenu transformacije. Posle primena transformacija, signal y_{11} nosi najviše korisnih informacija (ima najveću korelaciju) te se zbog toga on mora kodovati najfinije tj. kvantizerom sa najviše kvantizacionih nivoa (najvećom bitskom brzinom) kako ne bi došlo do gubitka informacija. Korelacije (a samim tim i varijanse) signala y_{12} , y_{21} i y_{22} opadaju u odnosu na y_{11} , ti signali nose sve manje količine informacije tako da se za njihovo kodovanje mogu projektovati kvantizeri sa manjim brojem kvantizacionih nivoa (manjih bitskih brzina), što potvrđuje i formula (4.1.15) [7].

Kao što se vidi na šemi sa slike 4.1.3 ulazni signal se, primenom dvostruke modifikovane Hadamarove transformacije H_1 , dekomponuje na 4 sekvence (4 grane) koje se posebno koduju. Međutim, kako signal y_{22} ima jako malu korelaciju postavlja se pitanje da li je uopšte potrebno nad signalom y_2 primenjivati još jednu transformaciju i time ga deliti na signale y_{21} i y_{22} , s obzirom na to da se njegovim transformisanjem ne dobija na kvalitetu kodovanja a angažuje se dodatni adaptivni kvantizer. Zbog toga je izveden još jedan eksperiment gde je signal y_1 (koji ima veliku korelaciju) dekomponovan na signale y_{11} i y_{12} , dok je signal y_2 kodovan takav kakav jeste bez dodatne transformacije. Kodna šema je prikazana na slici 4.1.4.



Slika 4.1.4. Šema za kodovanje govornog signala implementirana koristeći dvostruku prostu transformaciju na samo jednoj grani i adaptaciju unapred

Osnovna razlika šeme sa slike 4.1.4 u odnosu na šemu 4.1.3 jeste u uklanjanju jednog adaptivnog kvantizera iz upotrebe zato što nad signalom y_2 nije primenjena još jedna transformacija već je kodovan ceo. Time se štedi na bitovima po odmerku potrebnim za kodovanje gledajući ceo sistem. Varijanse signala y_{11} i y_{12} (σ_{11}^2 , σ_{12}^2 , respektivno) računaju se prema izrazu (4.1.14) dok je varijansa signala y_2 , σ_2^2 izračunata na osnovu izraza (4.1.2). Maksimalne amplitude opterećenja kvantizera AQ_1 i AQ_2 ($x_{\max 1}$ i $x_{\max 2}$, respektivno) računaju se na osnovu izraza (4.1.16) dok se maksimalna amplituda kvantizera AQ_3 , $x_{\max 3}$ izračunava koristeći izraz (4.1.7). Bitske brzine adaptivnih kvantizera dobijene su kao i u prethodno predloženim šemama koristeći izraz (4.1.15). Kao i u prethodnim šemama, za kodovanje varijanse angažovan je log-uniformni kvantizer detaljno objašnjen u odeljku 2.3.2. Tabela 4.1.4 prikazuje kvalitet kodovanja signala kroz SQNR koristeći kodnu šemu sa slike 4.1.4.

Tabela 4.1.4. Eksperimentalni rezultati (SQNR) dobijeni koristeći kodnu šemu sa slike 4.1.4 za razne bitske brzine i vrednosti faktora kompresije i rezultati prethodnih kodnih šema

R_{sr} [bita po odmerku]	μ	SQNR[dB]	SQNR _{4 grane} [dB]	SQNR _{2 grane} ($\mu=16.9227$) [dB]	SQNR _{PCM} [dB]
5	20	30.4192	30.1968	27.9027	23.4112
	255	27.2343	27.1547		
6	20	36.1893	36.0036	33.7459	29.2385
	255	33.4370	33.1343		
7	20	41.6938	41.5876	39.2063	35.1832
	255	39.0851	39.0757		
8	20	48.0339	47.7315	45.1963	41.0952
	255	45.4819	45.3552		

U tabeli 4.1.4 prikazani su eksperimentalno rezultati kvaliteta kodovanja koristeći kodnu šemu sa slike 4.1.4. SQNR dobijen primenom pomenute kodne šeme upoređen je sa SQNR-om dobijenim primenom šeme sa slike 4.1.3 (ulazni signal x se deli na 4 grane primenom

dvostruke transformacije), sa SQNR-om dobijenim u radu [59], gde je nad ulaznim signalom primenjena samo jedna prosta (modifikovana) Hadamarova transformacija H_1 i time ulazni signal dekomponovan na dve podsekvence, kao i sa SQNR-om dobijenim promenom PCM sistema. Ovo poređenje urađeno je kako bi se prikazao dobitak u kvalitetu kodovanja usled izbacivanja transformacije signala y_2 na signale y_{21} i y_{22} , zato što signal y_2 nema veliku korelaciju i ne nosi veliku količinu informacija u sebi pa dodatna transformacija ne doprinosi povećanju kvaliteta kodovanja. Eksperiment je urađen za vrednosti faktora kompresije $\mu=20$, 255, kao i za prethodnu kodnu šemu. Rezultati iz tabele 4.1.4 pokazuju da se izbacivanjem adaptivnog kvantizera AQ_4 iz kodne šeme ne gubi na kvalitetu već dobija dok se smanjuje kompleksnost cele šeme.

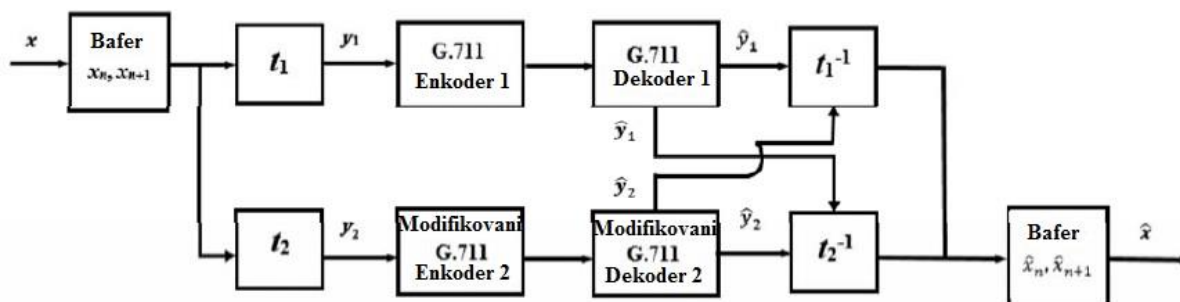
Za vrednost faktora kompresije $\mu=20$ kodna šema pokazuje dobitak od 0.1 do 0.3 dB u odnosu na šemu sa slike 4.1.3, 2.5 – 2.9 dB u odnosu na šemu iz rada [59], gde su kvantizeri projektovani za $\mu=16.9227$, dok u odnosu na PCM sistem ($\mu=16.9227$) dobitak kodovanja iznosi 6.7 – 7 dB, u zavisnosti od bitske brzine. Faktor kompresije $\mu=255$ pruža niži SQNR u odnosu na $\mu=20$ jer su kvazi-logaritamski kvantizeri robustniji za ovu vrednost. Ipak, poredeći rezultate sa onim dobijenim primenom šeme sa slike 4.1.3, vidi se dobitak kodovanja 0.01 – 0.3 dB za razne bitske brzine. Naime, cilj izbacivanja još jedne transformacije signala y_2 je smanjivanje kompleksnosti kodne šeme, a ujedno i povećanje kvaliteta kodovanja jer se dodatnom transformacijom ne dobija već, što se vidi iz prikazanih rezultata, gubi na kvalitetu kodovanja. Posle primena transformacija, signal y_{11} nosi najviše korisnih informacija (ima najveću korelaciju) te se zbog toga on mora kodovati najfinije tj. kvantizerom sa najviše kvantizacionih nivoa (najvećom bitskom brzinom) da ne bi došlo do gubitka informacija. Korelacije (a samim tim i varijanse) signala y_{12} i y_2 opadaju u odnosu na y_{11} i ti signali nose sve manje količine informacije, tako da se za njihovo kodovanje mogu projektovati kvantizeri sa manjim brojem kvantizacionih nivoa (manjih bitskih brzina), što potvrđuje i formula (4.1.15) [7]. Primenom jednačine (4.1.5) dolazi se do bitskih brzina za koje treba projektovati kvantizere i iznose $R_1 = R_{sr} + 3$, $R_2 = R_{sr} - 1$ i $R_3 = R_{sr} - 1$ bit po odmerku [7].

Sledeća šema za kodovanje govornog signala predložena u disertaciji a bazirana na modifikovanoj Hadamarovoj transformaciji H_1 , prvi put je prikazana u radu [71]. Šema je strukturalno slična G.711.1 sistemu za kodovanje signala, međutim za razliku od standarda G.711.1 ima dosta manju kompleksnost implementacije zato što se, nakon transformacionog kodovanja, obe dobijene sekvence koduju G.711 kvantizerima. Na ovaj način, ulazni signal se ne deli u frejmove, već je kašnjenje uzrokovano obradom samo prva dva odmerka signala koji

se procesiraju umesto celog frejma, što doprinosi značajnom smanjenju kašnjenja pri obradi signala.

Kao što je prikazano u literaturi [46], [62], [64], osnovni metod za kodovanje govornog signala standardizovan od strane ITU-T je preporuka G.711 [1], [45]. Ovim standardom je pokriveno kodovanje govornog signala bitskom brzinom od 8 bita po odmerku, odmerenog na 8 kHz, što obezbeđuje bitsku brzinu prenosa signala od 64 kb/s koristeći impulsnu kodnu modulaciju. G.711 predstavlja standard za kodovanje uskopojasnog govornog signala. Što se tiče obrade i prenosa širokopojasnih signala (odmerenih na 16 kHz) predložen je standard G.711.1 od strane ITU-T [46], [113]. Glavna karakteristika G.711.1 standarda je dekompozicija ulaznih signala koristeći frekvencijske filtere (niskopropusni filter i visokopropusni filter). Signal u grani posle niskopropusnog filtera kodovan je G.711 kvantizerom, dok je za kodovanje signala u grani iza visokopropusnog filtera korišćena vektorska kvantizacija. U radu [71] predložena je kodna šema koja umesto frekvencijskih filtara za dekompoziciju signala u sekvence koristi Hadamarovu transformaciju H_1 i obe nastale sekvence koduje G.711 koderima. Samim tim što vektorska kvantizacija nije iskorišćena u predloženoj šemi, kompleksnost same implementacije je redukovana u velikoj meri. Izbacivanjem vektorske kvantizacije iz šeme dodatno se smanjuje kašnjenje pri kodovanju u poređenju sa G.711.1 standardom i time se dobija realniji prenos u vremenu (*“real-time”*), kao i bolja kompresija. Prema G.711.1 standardu ulazni signal je podeljen u frejmove od 40 odmeraka od kojih svaki unosi po 0.125 ms kašnjenja pri kodovanju, dok predložena šema smanjuje to kašnjenje na 0.25 ms koje je uzrokovano kodovanjem samo prva dva odmerka (po jedan u svakoj grani). Dakle, ulazni govorni signal odmeren na 16 kHz transformacionim kodovanjem deli se na dve zasebne sekvence koje se obe koduju G.711 koderima. Prema G.711 standardu, ulazni signal se koduje bitskom brzinom od 8 bita po odmerku, dok vrednost faktora kompresije μ iznosi 255. Koder za kodovanje signala u grani iza transformacije t_1 je dizajniran upravo po G.711 standardu, dok je koder u grani iza transformacije t_2 , u cilju sniženja srednje bitske brzine, dizajniran za nižu bitsku brzinu (modifikovani G.711 koder za $R_2(Q_2)=4$ i $R_2(Q_2)=6$ bita po odmerku). Najveći problem pri implemetaciji kodne šeme predstavlja pronalazak optimalne vrednosti maksimalne amplitude opterećenja kvantizera u granama ($x_{\max 1}$ za Q_1 i $x_{\max 2}$ za Q_2) tako da je glavni akcenat stavljen upravo na to. Da bi bila urađena detaljna analiza uticaja maksimalne amplitude kvantizera na kvalitet kodovanja, eksperimenti su urađeni za muški i ženski govorni signal (odmereni na 16 kHz), zbog razlike u dinamici signala. Vrednosti maksimalne amplitude opterećenja oba kvantizera dobijene su kao proizvodi standardne devijacije svake grane i optimalne vrednosti maksimalne amplitude opterećenja kvazi-logaritamskog kvantizera kada

se na njegov ulaz dovede signal Laplasove gustine raspodele verovatnoće jedinične varijanse. Varijanse σ_1^2 i σ_2^2 signala y_1 i y_2 , odnosno standardne devijacije σ_1 i σ_2 , zavise od varijanse ulaznog signala σ_x^2 i njegovog koeficijenta korelacije ρ [7]. Ova dva parametra ulaznog signala jako variraju u zavisnosti od toga da li je u pitanju muški ili ženski govorni signal na ulazu te je od velike važnosti određivanje optimalnog regiona rada kvantizera. Šema za kodovanje govornog signala koja uključuje transformaciono kodovanje i G.711 kvantizere prikazana je na slici 4.1.5.



Slika 4.1.5. Šema za kodovanje govornog signala implementirana koristeći prostu transformaciju i G.711 kodere u granama

Ova kodna šema je predložena sa idejom da se, poredeći sa G.711.1 standardom, smanji kašnjenje prilikom obrade signala a kvalitet rekonstruisanog signala ostane nepromenjen. Kao što je već rečeno, ulazni signal x je primenom proste (modifikovane) transformacije, prikazanom u izrazu (4.1.1), dekomponovan u dve nezavisne sekvence y_1 i y_2 , sa većom korelacijom i užim dinamičkim opsegom, koje se dalje koduju upotrebom G.711 koda. Urađena je i modifikacija G.711 koda u grani signala y_2 smanjivanjem bitske brzine sa 8 na 4 bita po odmerku u odnosu na G.711 standard [1], [7], [73], [114]. G.711 koder u predajniku i G.711 dekodekter u prijemniku predstavljaju G.711 kvantizer koji je implementiran koristeći segmentnu uniformnu kvantizaciju, tako da se N kvantizacionih nivoa kvantizera deli u L segmenata, od kojih svaki segment sadrži n nivoa [71]. Eksperiment je urađen za $L=16$ nivoa, dok broj kvantizacionih nivoa unutar svakog segmenta $n=N/L$ zavisi od ukupnog broja kvantizacionih nivoa, odnosno ukupne prosečne bitske brzine ($R_1 = 2^{N(Q_1)}, R_2 = 2^{N(Q_2)}$). Signal y_1 je kodovan kvantizerom Q_1 (G.711 Koder 1/Dekoder 1) dok je signal y_2 kodovan kvantizerom Q_2 (modifikovani G.711 Koder 2 / Dekoder 2), što je prikazano i na slici 4.1.5. Bitska brzina kvantizera Q_1 je $R_1=8$ bita po odmerku, shodno G.711 standardu, dok se modifikacija ogleda u smanjenju bitske brzine kvantizera Q_2 , i iznosi $R_2=4$ bita po odmerku i $R_2=6$ bita po odmerku.

Granice između segmenata definisane su izrazom:

$$x_i = \frac{(2^i - 1)}{255} \cdot x_{\max}, i = 1, 2, \dots, L, \quad (4.1.17)$$

gde je x_{\max} maksimalna amplituda opterećenja kvantizera, i je brojač segmenata kvantizera koji ide od 1 do L . U svakom segmentu L_i nalazi se n kvantizacionih (reprezentacionih) nivoa, definisanih izrazom:

$$\Delta_i = \frac{x_{i+1} - x_i}{n}. \quad (4.1.18)$$

Pragovi odluke x'_j i reprezentacioni nivoi y'_j za kvantizacione nivoe sadržane u svakom segmentu ($j=1, 2, \dots, n$) definisani su na sledeći način:

$$\begin{aligned} x'_j &= j \cdot \Delta, & j &= 1, 2, \dots, n, \\ y'_j &= (j - 0.5) \cdot \Delta, & j &= 1, 2, \dots, n, \end{aligned} \quad (4.1.19)$$

gde Δ predstavlja širinu kvantizacionog nivoa, dok je j brojač reprezentacionih nivoa unutar svakog L_i segmenta. Na prijemnoj strani kvantovane sekvence signala \hat{y}_1 i \hat{y}_2 formiraju kvantovani izlazni signal \hat{x} koristeći inverzne transformacije t_1^{-1} i t_2^{-1} , definisane sa:

$$\begin{aligned} t_1^{-1} &= \hat{y}_1 + \hat{y}_2 = \hat{x}_k, \\ t_2^{-1} &= \hat{y}_1 - \hat{y}_2 = \hat{x}_{k+1} \end{aligned} \quad (4.1.20)$$

gde su \hat{x}_k i \hat{x}_{k+1} dva susedna odmerka kvantovanog signala \hat{x} . Kao što je već napomenuto, glavni izazov prilikom projektovanja ove kodne šeme je određivanje optimalnih vrednosti maksimalnih amplituda opterećenja kvantizera Q_1 i Q_2 , s obzirom da ove vrednosti imaju značajan uticaj na kvalitet rekonstruisanog signala, posmatrajući SQNR. Vrednosti regiona rada (maksimalnih amplituda opterećenja) $x_{\max 1}$ i $x_{\max 2}$, kvantizera Q_1 i Q_2 respektivno, direktno zavise od varijansi signala y_1 i y_2 (σ_1^2 i σ_2^2 , respektivno), koje se dobijaju koristeći jednačinu (4.1.2) [70]. Regioni rada kvantizera $x_{\max 1}$ i $x_{\max 2}$ dobijaju se na sledeći način:

$$\begin{aligned} x_{\max 1} &= l \cdot \sigma_1 \cdot x_{\max}^{\sigma=1}, \\ x_{\max 2} &= l \cdot \sigma_2 \cdot x_{\max}^{\sigma=1}. \end{aligned} \quad (4.1.21)$$

Vrednosti σ_1 i σ_2 iz jednačine (4.1.21) predstavljaju standardne devijacije signala y_1 i y_2 , respektivno, l je koeficijent multiplikacije, dok je $x_{\max}^{\sigma=1}$ optimalna vrednost regiona rada kvantizera kada je na njegov ulaz doveden signal Laplasove funkcije raspodele gustine verovatnoće dobijen iz izraza (4.1.8).

Bit alokacija, važan parametar pri projektovanju kodne šeme, dobijena je korišćenjem izraza (4.1.15) koji se za dve grane može svesti na:

$$R_1 = R + \frac{1}{4} \log_2 \left(\frac{1 + \rho}{1 - \rho} \right), \quad (4.1.22)$$

$$R_2 = R + \frac{1}{4} \log_2 \left(\frac{1 - \rho}{1 + \rho} \right),$$

gde R predstavlja srednju bitsku brzinu, dok je ρ koeficijent korelacije ulaznog signala [7].

Urađeni eksperiment ima svoju teoretsku pozadinu prikazanu kroz SQNR i dobitak SQNR-a za slučaj kada je ulazni signal modelovan Laplasovom funkcijom raspodele gustine verovatnoće. Laplasova raspodela za određeni kvantizacioni nivo definisana je na sledeći način:

$$P_i(x_{\max}, \sigma) = \int_{x_i}^{x_{i+1}} \frac{1}{\sqrt{2}\sigma} \exp\left(-\frac{|x|\sqrt{2}}{\sigma}\right) dx. \quad (4.1.23)$$

Granularna distorzija zavisi od broja reprezentacionih nivoa, regiona rada kvantizera i standardne devijacije ulaznog signala i definisana je:

$$D_g(x_{\max}, \sigma) = \frac{1}{6} \sum_{i=0}^{L-1} \Delta_i^2 \cdot P_i(x_{\max}, \sigma). \quad (4.1.24)$$

Distorzija prekoračenja zavisi od regiona rada kvantizera i varijanse ulaznog signala:

$$D_{ov}(x_{\max}, \sigma) = \sigma^2 \cdot \exp\left(-\frac{\sqrt{2}x_{\max}}{\sigma}\right), \quad (4.1.25)$$

dok je ukupna distorzija suma granularne i distorzije prekoračenja, kao što je već napomenuto u prethodnim poglavljima:

$$D = D_g + D_{ov}. \quad (4.1.26)$$

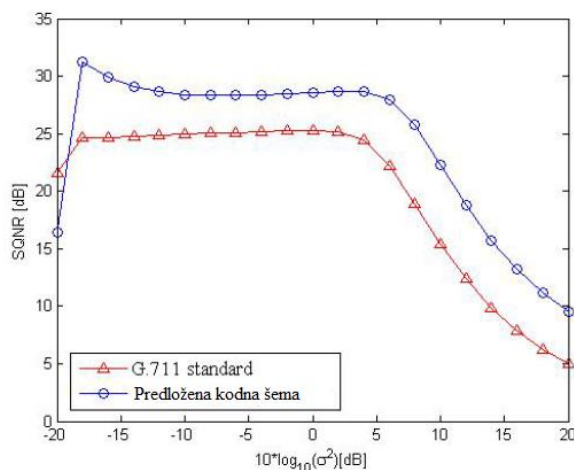
Iz jednačina (4.1.24) i (4.1.25) vidi se velika zavisnost distorzije, odnosno SQNR-a, od maksimalne amplitude opterećenja kvantizera te je zbog toga veoma važan optimalan odabir ove vrednosti pri projektovanju kvantizera. Kvalitet rekonstrukcije signala definisan je kroz SQNR sa:

$$\text{SQNR[dB]} = 10 \cdot \log_{10} \left(\frac{\sigma^2}{D} \right). \quad (4.1.27)$$

Dobitak u kvalitetu kodovanja kodne šeme sa slike 4.1.5 u odnosu na G.711 standard, definisan je sa:

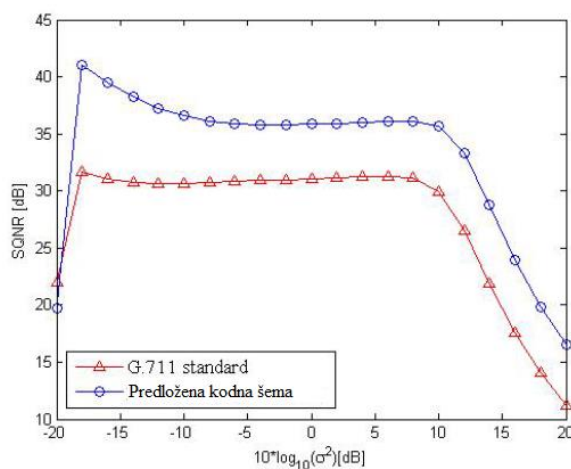
$$G[\text{dB}] = 10 \cdot \log_{10} \left(\frac{D_{G.711}(R, \sigma)}{D(R_1, R_2, \sigma)} \right). \quad (4.1.28)$$

Teorijski rezultati prikazani su na slikama 4.1.6 i 4.1.7 [66], [69].



Slika 4.1.6. SQNR za G.711 standard ($R=6$ bita po odmerku) i predložene kodne šeme sa slike 4.1.5 ($R_1=8$ bita po odmerku, $R_2=4$ bita po odmerku)

Slika 4.1.6 pokazuje SQNR dobijen korišćenjem predložene kodne šeme sa slike 4.1.5 i SQNR dobijen korišćenjem G.711 standarda. Bitske brzine kvantizera Q_1 i Q_2 su 8 i 4 bita po odmerku, respektivno, dok je bitska brzina G.711 standarda 6 bita po odmerku u cilju adekvatnog poređenja. Posmatrajući sliku 4.1.6, u opsegu ulaznih varijansi [-20 dB, 5 dB] može se videti da predložena kodna šema dostiže do 30 dB, što je oko 5.5 dB više u odnosu na G.711 standard. Posmatrajući opseg [-15 dB, 20 dB] primećuje se da kodna šema obezbeđuje između 3.5 i 7 dB viši SQNR u odnosu na G.711 standard. Kako bi bila potvrđena pogodnost kodne šeme za kodovanje govornog signala teoretske kalkulacije su urađene za još jedan set bitskih brzina, Q_1 i Q_2 projektovani su za 8 i 6 bita po odmerku, respektivno, dok je bitska brzina G.711 standarda 7 bita po odmerku.



Slika 4.1.7. SQNR za G.711 standard ($R=7$ bita po odmerku) i predložene kodne šeme sa slike 4.1.5 ($R_1=8$ bita po odmerku, $R_2=6$ bita po odmerku)

Posmatrajući sliku 4.1.7, u opsegu ulaznih varijansi [-20 dB, 10 dB] može se videti da predložena kodna šema dostiže do 35 dB, sa jednim pikom od 41 dB na -17.5 dB dok približno linearno opada za više ulazne varijanse. Dobitak kodovanja predložene kodne šeme u odnosu na G.711 za pomenute bitske brzine varira od -2 dB do 9 dB. Negativni dobitak kodovanja se javlja samo u oblasti jako niskih ulaznih varijansi i u toj oblasti za pomenute brzine predložena kodna šema ne obezbeđuje bolje performanse u odnosu na G.711. U oblastima viših ulaznih varijansi predložena kodna šema pruža oko 6 dB viši SQNR od G.711 standarda. Na osnovu teorijskih proračuna može se zaključiti da je predložena kodna šema pogodna za kodovanje širokopojasnog govornog signala [71].

Teorijski rezultati potvrđeni su i eksperimentalnim putem. Kao što je navedeno, eksperiment je urađen za muški i ženski govorni signal odmeren na 16 kHz. Muški i ženski govorni signal se razlikuju u dinamici što uzrokuje razliku u kvalitetu kodovanja signala. Kako bi bio prikazan uticaj vrednosti regiona rada kvantizera na kvalitet kodovanja, eksperiment je urađen za dva seta vrednosti $x_{\max 1}$ i $x_{\max 2}$. Prvi set vrednosti dobijen je koristeći formulu (4.1.21) za $l=1$ dok u drugom setu $x_{\max 1}$ i $x_{\max 2}$ uzimaju dvostruko veće vrednosti, tako da faktor multipliciranja ima vrednost $l=2$ [71]. Performanse kodovanja dobijene eksperimentom prikazane su, kroz SQNR, u tabelama 4.1.5 – 4.1.8.

Tabela 4.1.5. Eksperimentalni rezultati (SQNR) dobijeni koristeći kodnu šemu sa slike 4.1.5 i G.711 standarda za muški govorni signal i prosečnu bitsku brzinu $R=6$ bita po odmerku

		Prvi set vrednosti $x_{\max 1}$ i $x_{\max 2}$		Drugi set vrednosti $x_{\max 1}$ i $x_{\max 2}$	
R [bita po odmerku]	μ	SQNR [dB]	SQNR _{G.711} [dB]	SQNR [dB]	SQNR _{G.711} [dB]
$R_1=8$ $R_2=4$	255	28.6658	24.7077	29.5098	25.1288

Tabela 4.1.5 pokazuje SQNR dobijen korišćenjem predložene kodne šeme i G.711 standarda za muški govorni signal. Koeficijent korelacije ulaznog signala iznosi $\rho=0.9672$, dok je varijansa $\sigma^2=0.0021$. S obzirom da se koriste G.711 kvantizeri (standardni i modifikovani) faktor kompresije kvantizera iznosi $\mu=255$. Kao što je prethodno pomenuto, što se i vidi u tabeli 4.1.5, eksperiment je urađen za dva seta vrednosti $x_{\max 1}$ i $x_{\max 2}$, korišćenjem fomule (4.1.21) za $l=1$, pa je $x_{\max 1}=0.3733$ i $x_{\max 2}=0.0482$, i $l=2$ gde vrednosti regiona rada kvantizera Q_1 i Q_2 iznose $x_{\max 1}=0.7466$ i $x_{\max 2}=0.0964$. Maksimalna amplituda opterećenja kvantizera (region rada) u G.711 standardu, radi adekvatnog poređenja, ima istu vrednost kao i region rada kvantizera u grani posle transformacije t_1 . Bitske brzine kvantizera Q_1 i Q_2 iznose $R_1=8$ bita po

odmerku i $R_2=4$ bita po odmerku, dok bitska brzina kvantizera u G.711 standardu iznosi 6 bita po odmerku (modifikovani G.711), što je prosečna bitska brzina kvantizera Q_1 i Q_2 sa slike 4.1.5, tako da se može izvršiti adekvatno poređenje performansi. Vidi se da predložena kodna šema pruža 3.8951 dB viši SQNR u odnosu na G.711 standard za prvi set vrednosti $x_{\max 1}$ i $x_{\max 2}$, i 4.383 dB za drugi set. Eksperimentom je utvrđeno da se za $l=2$ dobijaju optimalne vrednosti regiona rada kvantizera jer dalje povećavanje faktora multipliciranja obara performanse kodovanja [71].

Tabela 4.1.6. Eksperimentalni rezultati (SQNR) dobijeni koristeći kodnu šemu sa slike 4.1.5 i G.711 standarda za ženski govorni signal i prosečnu bitsku brzinu $R=6$ bita po odmerku

R [bita po odmerku]	μ	Prvi set vrednosti $x_{\max 1}$ i $x_{\max 2}$		Drugi set vrednosti $x_{\max 1}$ i $x_{\max 2}$	
		SQNR [dB]	SQNR _{G.711} [dB]	SQNR [dB]	SQNR _{G.711} [dB]
$R_1=8$ $R_2=4$	255	25.9416	23.5089	28.6113	25.0890

Tabela 4.1.6 pokazuje SQNR dobijen korišćenjem predložene kodne šeme i G.711 standarda za ženski govorni signal za iste bitske brzine kao i tabela 4.1.5. Koeficijent korelacije ulaznog signal iznosi $\rho = 0.9586$, dok je varijansa $\sigma^2=0.0062$. I u tabeli 4.1.6, eksperiment je urađen za dva seta vrednosti $x_{\max 1}$ i $x_{\max 2}$, korišćenjem fomule (4.1.21) za $l=1$, pa je $x_{\max 1}=0.6376$ i $x_{\max 2}=0.0925$, i $l=2$ gde vrednosti regiona rada kvantizera Q_1 i Q_2 iznose $x_{\max 1}=1.2734$ i $x_{\max 2}=0.185$. Veća dinamika ženskog signala u odnosu na muški uzrokuje 2.7242 dB niži SQNR za predloženu kodnu šemu i 1.1988 dB niži SQNR za G.711 standard. Međutim, predložena kodna šema i dalje ostvaruje dobitak u SQNR-u u odnosu na klasični G.711, i to 2.4327 dB za prvi set regiona rada kvantizera i 3.5223 dB za drugi [71].

U cilju potvrde pogodnosti korišćenja predložene kodne šeme za kodovanje širokopojasnog govornog signala (odmerenog na 16 kHz), eksperiment je urađen za još jedan set bitskih brzina, za muški i ženski govorni signal. Bitska brzina kvantizera Q_2 u grani posle transformacije t_2 iznosi $R_2=6$ bita po odmerku, tako da je vrednost srednje bitske brzine $R=7$ bita po odmerku.

Tabela 4.1.7. Eksperimentalni rezultati (SQNR) dobijeni koristeći kodnu šemu sa slike 4.1.5 i G.711 standarda za muški govorni signal i prosečnu bitsku brzinu $R=7$ bita po odmerku

R [bita po odmerku]	μ	Prvi set vrednosti $x_{\max 1}$ i $x_{\max 2}$		Drugi set vrednosti $x_{\max 1}$ i $x_{\max 2}$	
		SQNR [dB]	SQNR _{G.711} [dB]	SQNR [dB]	SQNR _{G.711} [dB]
$R_1=8$ $R_2=6$	255	34.7762	31.1869	36.1886	31.1594

Maksimalne amplitude opterećenja kvantizera zavise od bitske brzine i za slučaj prikazan u tabeli 4.1.7 (muški govorni signal) imaju vrednosti $x_{\max 1}=0.4128$ i $x_{\max 2}=0.0540$ kada faktor multiplikacije l ima vrednost 1. Bitske brzine kvantizera Q_1 i Q_2 iznose $R_1=8$ i $R_2=6$ bita po odmerku, dok je bitska brzina kvantizera u G.711 standardu $R=7$ bita po odmerku. Iz tabele se vidi da predložena kodna šema obezbeđuje 3.5893 dB dobitka u odnosu na klasičan G.711 za prvi set vrednosti $x_{\max 1}$ i $x_{\max 2}$, i 5.0292 dB dobitka u SQNR-u za drugi set vrednosti regiona rada kvantizera ($l=2$) koje iznose $x_{\max 1}=0.8355$ i $x_{\max 2}=0.1081$.

U slučaju ženskog govornog signala maksimalne amplitude opterećenja kvantizera imaju vrednosti $x_{\max 1}=0.7132$ i $x_{\max 2}=0.1036$ za vrednost faktora multiplikacije $l=1$, i $x_{\max 1}=1.4264$ i $x_{\max 2}=0.2072$ za $l=2$.

Tabela 4.1.8. Eksperimentalni rezultati (SQNR) dobijeni koristeći kodnu šemu sa slike 4.1.5 i G.711 standarda za muški govorni signal i prosečnu bitsku brzinu $R=7$ bita po odmerku

R [bita po odmerku]	μ	Prvi set vrednosti $x_{\max 1}$ i $x_{\max 2}$		Drugi set vrednosti $x_{\max 1}$ i $x_{\max 2}$	
		SQNR [dB]	SQNR _{G.711} [dB]	SQNR [dB]	SQNR _{G.711} [dB]
$R_1=8$ $R_2=6$	255	30.9070	28.8145	36.0050	31.0751

Iz tabele se može videti da predložena kodna šema ostvaruje 2.0925 dB dobitka u SQNR-u, u odnosu na klasičan G.711 standard, za prvi set vrednosti regiona rada kvantizera i 4.9335 dB dobitka za drugi set. Poredeći SQNR za prvi set vrednosti regiona rada, dobijen kodovanjem muškog i ženskog govornog signala, vidi se da ženski govorni signal ostvaruje 3.8692 dB niži SQNR, dok za drugi set vrednosti regiona rada (optimalni) ostvaruje 0.183 dB niži SQNR.

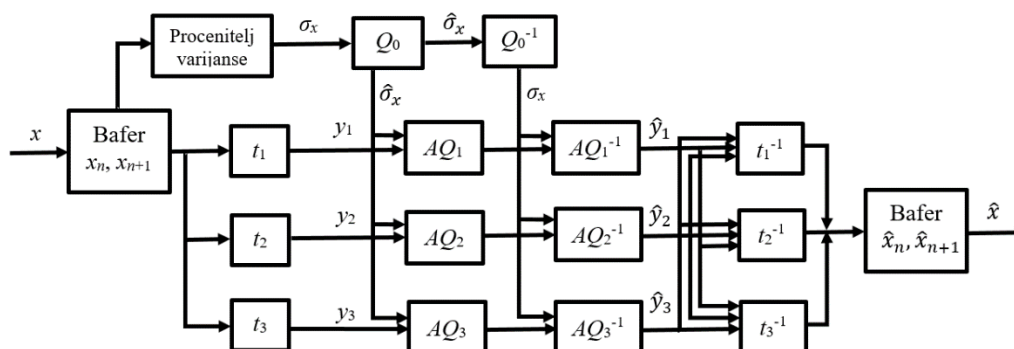
Ideja je bila da se predloži kodna šema koja pruža realnije vreme kodovanja (što manje kašnjenje) u odnosu na G.711.1 standard, dok istovremeno obezbeđuje što viši kvalitet kodovanja poredeći sa G.711 standardom. Konačno, u poređenju sa G.711 standardom,

predložena kodna šema pruža viši SQNR i za muški i za ženski govorni signal. Takođe, kompleksnost kodne šeme je takva da obezbeđuje 20 puta manje kašnjenje u odnosu na G.711.1 standard i 2.5 – 5 dB viši SQNR u odnosu na G.711 standard [71].

4.2 Šeme za kodovanje kontinualnog govornog signala konstruisane koristeći diskretnu kosinusnu transformaciju – DCT

U ovoj sekciji prikazane su kodne šeme za kodovanje kontinualnog govornog signala implementirane koristeći diskretnu kosinusnu transformaciju (DCT) sa matricama koeficijenata 3×3 i 4×4 .

Prva prikazana kodna šema implementirana je koristeći diskretnu kosinusnu transformaciju sa matricom koeficijenata 3×3 i tehniku adaptacije unapred [1] – [3], [5] – [8], [42], [56] – [59], [71]. Prikazana šema se prilagođava (adaptira) varijansi ulaznog signala čime se obezbeđuje visoka efikasnost iskorišćenja propusnog opsega, dok se implementacijom diskretne kosinusne transformacije signal dekomponuje (deli) u podsekvence čije su karakteristike signala predvidljivije, pa se upotreba predložene kodne šeme pokazuje pogodnom za procesiranje signala. Na slici 4.2.1 prikazana je kodna šema koja je koncipirana tako da se ulazni signal najpre deli na podsekvence primenom DCT transformacije obezbeđujući uži propusni opseg, nakon čega se vrši adaptacija kvantizera na varijansu signala u sekvencama, što rezultuje time da se koristi manje bita po odmerku za kodovanje signala.



Slika 4.2.1. Šema za kodovanje govornog signala implementirana koristeći diskretnu kosinusnu transformaciju 3×3 i adaptaciju unapred

Ulazni govorni signal x , odmeren na 8 kHz, dovodi se do bafera u kome se formiraju frejmovi veličine M odmeraka. Nad signalom u frejmu primenjuje se DCT transformacija 3×3 koja ga transformiše u tri nezavisne sekvence (y_1, y_2, y_3) koje se dalje koduju adaptivnim

kvantizerima zasnovanim na μ -logaritamskom zakonu. S obzirom da se koriste adaptivni kvantizeri sa tehnikom adaptacije unapred, potrebno je do prijemnika preneti informaciju o varijansi ulaznog signala kao dodatnu informaciju. Procenitelj varijansi računa varijansu signala u frejmu, log-uniformni kvantizer je koduje i tako kodovana varijansa (standardna devijacija) prosleđuje se do prijemnika. U prijemniku se koriste inverzne transformacije DCT-a 3×3 koje formiraju izlazni signal y .

Matrica koeficijenata A diskretne kosinusne transformacije veličine 3×3 iskorišćena u predloženoj kodnoj šemi data je sa:

$$A = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 1 \\ \frac{-1}{\sqrt{2}} & \sqrt{2} & \frac{-1}{\sqrt{2}} \\ -\sqrt{\frac{3}{2}} & 0 & \sqrt{\frac{3}{2}} \end{bmatrix}. \quad (4.2.1)$$

Primenom ove matrice koeficijenata nad ulaznim signalom x , dobijaju se tri izlazne sekvence y_1, y_2, y_3 i to na sledeći način:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = A \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 1 \\ \frac{-1}{\sqrt{2}} & \sqrt{2} & \frac{-1}{\sqrt{2}} \\ -\sqrt{\frac{3}{2}} & 0 & \sqrt{\frac{3}{2}} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad (4.2.2)$$

što se može prikazati i pojedinačno:

$$\begin{aligned} y_1 &= \frac{1}{\sqrt{3}} (x_1 + x_2 + x_3), \\ y_2 &= \frac{1}{\sqrt{3}} \left(-\frac{1}{\sqrt{2}} x_1 + \sqrt{2} x_2 - \frac{1}{\sqrt{2}} x_3 \right), \\ y_3 &= \frac{1}{\sqrt{3}} \left(-\sqrt{\frac{3}{2}} x_1 + \sqrt{2} x_2 - \frac{1}{\sqrt{2}} x_3 \right). \end{aligned} \quad (4.2.3)$$

Varijanse signala y_1, y_2 i y_3 (σ_1^2, σ_2^2 i σ_3^2 , respektivno) zavise od varijanse ulaznog signala u frejmu σ_x^2 , kao i od koeficijenta korelacije između dva susedna odmerka ρ_1 i koeficijenta korelacije između dva nesusedna odmerka ρ_2 , i definisane su sledećim jednačinama:

$$\begin{aligned} \sigma_1^2 &= \sigma_x^2 \left(1 + \frac{4}{3} \rho_1 + \frac{2}{3} \rho_2 \right), \\ \sigma_2^2 &= \sigma_x^2 \left(1 - \frac{4}{3} \rho_1 + \frac{1}{3} \rho_2 \right), \\ \sigma_3^2 &= \sigma_x^2 (1 - \rho_2). \end{aligned} \quad (4.2.4)$$

S obzirom da su u granama za kodovanje sekvenci y_1 , y_2 i y_3 angažovani adaptivni kvantizeri, procenitelj varijansi računa varijansu (standardnu devijaciju) signala u frejmu, čiju vrednost šalje na kodovanje log-uniformnom kvantizeru, detaljno opisanom u odeljku 2.3.2, te se ta kodovana varijansa signala u frejmu šalje ka prijemniku kao dodatna informacija. Adaptivni kvantizeri AQ_1 , AQ_2 i AQ_3 , bazirani na μ -logaritamskom zakonu čija je kompresorska funkcija data jednačinom (4.1.3), prilagođavaju se signalima y_1 , y_2 i y_3 , respektivno, tako da se njihovi regioni rada (maksimalne amplitude opterećenja) dobijaju kao proizvod maksimalne amplitude opterećenja kvantizera kada se na njegov ulaz dovede signal modelovan Laplasovom funkcijom raspodele gustine verovatnoće jedinične varijanse i kodovane standardne devijacije signala po granama [1], [7]:

$$\begin{aligned}x_{\max 1} &= x_{\text{ml}} \cdot \hat{\sigma}_1, \\x_{\max 2} &= x_{\text{ml}} \cdot \hat{\sigma}_2, \\x_{\max 3} &= x_{\text{ml}} \cdot \hat{\sigma}_3,\end{aligned}\tag{4.2.5}$$

gde su $\hat{\sigma}_1$, $\hat{\sigma}_2$ i $\hat{\sigma}_3$ kvantovane vrednosti standardne devijacije, dobijene od log-uniformnog kvantizera a x_{ml} već pomenuta maksimalna amplituda kvantizera za signal sa Laplasovom raspodelom dobijena iz fomule (4.1.8). Na ovaj način dobijene su kvantovane sekvence \hat{y}_1 , \hat{y}_2 i \hat{y}_3 koje primenom inverzne diskretne kosinusne transformacije formiraju izlazni signal \hat{x} na način:

$$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{bmatrix} = A^{-1} \cdot \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix}.\tag{4.2.6}$$

Inverzna matrica koeficijenta DCT-a 3×3 ima oblik [60], [97] – [103]:

$$A^{-1} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & -\frac{1}{\sqrt{2}} & -\sqrt{\frac{3}{2}} \\ 1 & \sqrt{2} & 0 \\ 1 & \frac{1}{\sqrt{2}} & \sqrt{\frac{3}{2}} \end{bmatrix},\tag{4.2.7}$$

pa se odmerci izlaznog signala \hat{x} dobijaju:

$$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{bmatrix} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & -\frac{1}{\sqrt{2}} & -\sqrt{\frac{3}{2}} \\ 1 & \sqrt{2} & 0 \\ 1 & \frac{1}{\sqrt{2}} & \sqrt{\frac{3}{2}} \end{bmatrix} \cdot \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix}\tag{4.2.8}$$

ili pojedinačno:

$$\begin{aligned}\hat{x}_1 &= \frac{1}{\sqrt{3}} \left(\hat{y}_1 - \frac{1}{\sqrt{2}} \hat{y}_2 - \sqrt{\frac{3}{2}} \hat{y}_3 \right), \\ \hat{x}_2 &= \frac{1}{\sqrt{3}} (\hat{y}_1 + \sqrt{2} \hat{y}_2), \\ \hat{x}_3 &= \frac{1}{\sqrt{3}} \left(\hat{y}_1 - \frac{1}{\sqrt{2}} \hat{y}_2 + \sqrt{\frac{3}{2}} \hat{y}_3 \right).\end{aligned}\quad (4.2.9)$$

Kao mera kvaliteta kodovanja koristi se SQNR, koji se računa za svaki frejm korišćenjem formula (4.1.10) i (4.1.11) i na kraju prosečni SQNR svih frejmova za ceo signal.

Kao i u prethodno prikazanim šemama, od velikog značaja je izbor bit alokacije, tako da se i ovde primenjuje jednačina (4.1.15) [1], [7].

Eksperiment je urađen za vrednosti faktora kompresije $\mu=20$ i $\mu=255$ i bitske brzine kvantizera u granama dobijenih iz formule (4.1.15) koje zavise od srednje bitske brzine:

$$\begin{aligned}R_1 &= R_{sr} + 2, \\ R_2 &= R_{sr} - 2, \\ R_3 &= R_{sr}.\end{aligned}\quad (4.2.10)$$

Da bi se pokazala pogodnost korišćenja kodne šeme sa slike 4.2.1, dobijeni SQNR upoređen je sa SQNR-om G.711 standarda. Rezultati su prikazani za srednju bitsku brzinu $R_{sr}=5, 6, 7, 8$ bita po odmerku.

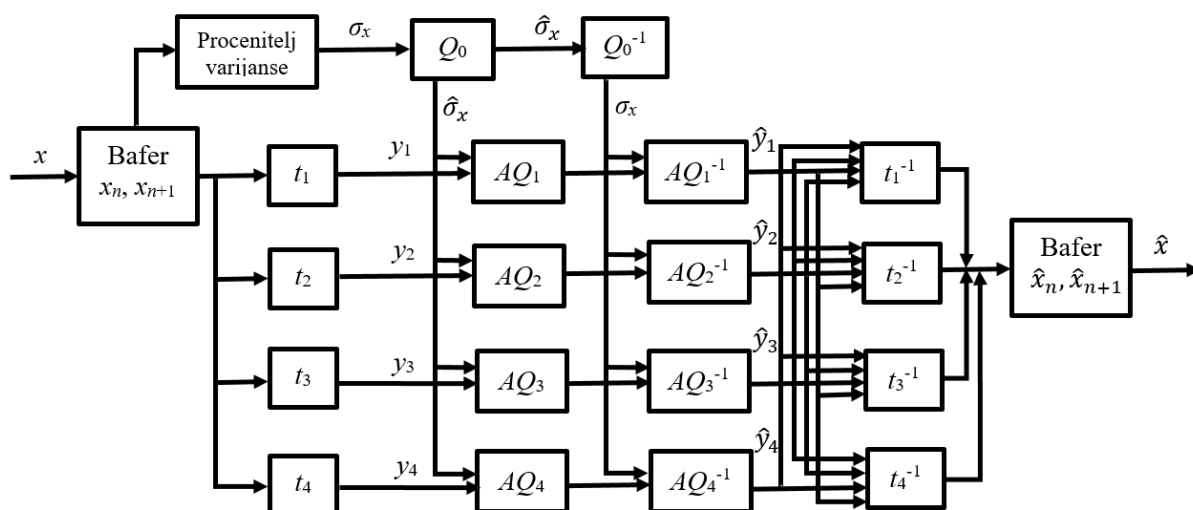
Tabela 4.2.1. Eksperimentalni rezultati (SQNR) dobijeni koristeći kodnu šemu sa slike 4.2.1, kodnih šema sa Hadamarovom transformacijom i G.711 standarda i različite bitske brzine

R_{sr} [bita po odmerku]	μ	SQNR[dB]	SQNR _{4 grane} [dB]	SQNR _{2 grane} ($\mu=16.9227$)[dB]	SQNR _{PCM} [dB]
5	20	28.7478	30.1968	25.9027	23.4112
	255	27.0165	27.1547		
6	20	34.2541	36.0036	31.7459	29.2385
	255	33.0727	33.1343		
7	20	40.0320	41.5876	39.2063	35.1832
	255	38.9444	39.0757		
8	20	45.6249	47.7315	45.1963	41.0952
	255	44.9313	45.3552		

Tabela 4.2.1 prikazuje eksperimentalne rezultate u vidu SQNR-a dobijenih korišćenjem šeme sa slike 4.2.1 za kodovanje govornog signala, SQNR prethodno prikazanih kodnih šema koje koriste prostu (modifikovanu) Hadamarovu transformaciju za transformaciono kodovanje, kao i SQNR dobijen kodovanjem govornog signala impulsnom kodnom modulacijom. Ulazni test govorni signal snimljen je u laboratoriji za Akustiku na Elektronskom fakultetu u Nišu,

odmeren na 8 kHz isti je kao i test signal nad kojim je urađen eksperiment u kodnoj šemi sa slike 4.1.1, radi adekvatnog poređenja. Trajanje ulaznog test signala je oko jedan minut, dok varijansa ima vrednost 0.9068. Iz tabele se vidi da primena diskretne kosinusne transformacije sa matricom koeficijenata veličine 3×3 obezbeđuje približno isti kvalitet kodovanja za posmatrani faktor kompresije $\mu=20$ kao i kodna šema sa slike 4.1.1 (jednostruka Hadamarova transformacija H_1), dok pruža oko 2 dB niži SQNR u odnosu na šemu sa slike 4.1.3 (dvostruka prosta Hadamarova transformacija H_1). Za vrednost faktora kompresije $\mu=255$, kodna šema pruža rezultate približno jednake kao i kodna šema implementirana dvostrukom primenom Hadamarove transformacije, 0.1 - 0.3 dB niži SQNR. Iz navedenih rezultata može se zaključiti da ukoliko je zahtev sistema za kodovanje takav da kvantizeri budu implementirani za faktor kompresije $\mu=20$, veći kvalitet kodovanja će se ostvariti primenom transformacionog kodovanja baziranog na dvostrukoj primeni proste Hadamarove transformacije H_1 , dok ukoliko je zahtev da se kvantizeri projektuju za $\mu=255$, što je standardna vrednost faktora kompresije u PCM sistemima, bolje je koristiti diskretnu kosinusnu transformaciju 3×3 koja obezbeđuje približno jednak kvalitet kodovanja dok zahteva implementaciju jednog kvantizera manje.

Sledeća prikazana kodna šema je implementirana koristeći diskretnu kosinusnu transformaciju sa matricom koeficijenata 4×4 i tehniku adaptacije unapred. Adaptacija kvantizera na ulazni signal obezbeđuje visoku efikasnost iskorišćenja propusnog opsega dok se implementacijom diskretne kosinusne transformacije signal dekomponuje (deli) u podsekvence signala čije su karakteristike predvidljivije, pa se upotreba predložene kodne šeme pokazuje pogodnom za procesiranje signala. Slika 4.2.2 prikazuje kodnu šemu koja je koncipirana tako da se ulazni signal najpre deli na četiri zasebne podsekvence primenom DCT transformacije obezbeđujući uži propusni opseg, nakon čega se vrši adaptacija kvantizera na varijansu signala u sekvencama, što rezultuje time da se koristi manje bita po odmerku za kodovanje signala.



Slika 4.2.2. Šema za kodovanje govornog signala implementirana koristeći diskretnu kosinusnu transformaciju 4×4 i adaptaciju unapred

Kao i u prethodnoj šemi, ulazni govorni signal x , odmeren na 8 kHz, dovodi se do bafera u kome se formiraju frejmovi veličine M odmeraka nad kojima se primenjuje DCT transformacija 4×4 koja ih transformiše u četiri nezavisne sekvence signala (y_1, y_2, y_3, y_4). Dobljene sekvence koduju se adaptivnim kvantizerima zasnovanim na μ -logaritamskom zakonu. S obzirom da se koriste adaptivni kvantizeri sa tehnikom adaptacije unapred, potrebno je do prijemnika preneti informaciju o varijansi ulaznog signala kao dodatnu informaciju. Procenitelj varijansi računa varijansu signala u frejmu, log-uniformni kvantizer je koduje i tako kodovana varijansa (standardna devijacija) se prosleđuje do prijemnika. U prijemniku se koriste inverzne transformacije DCT-a 4×4 koje formiraju izlazni signal y .

Matrica koeficijenata A diskretne kosinusne transformacije veličine 4×4 iskorišćena u kodnoj šemi sa slike 4.2.2 data je sa:

$$A = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ a & b & -b & -a \\ 1 & -1 & -1 & 1 \\ b & -a & a & -b \end{bmatrix}, \quad (4.2.11)$$

gde su koeficijenti a i b dati izrazima:

$$\begin{aligned} a &= \sqrt{2} \cos\left(\frac{\pi}{8}\right), \\ b &= \sqrt{2} \cos\left(\frac{3\pi}{8}\right). \end{aligned} \quad (4.2.12)$$

Primenom ove matrice koeficijenata nad ulaznim signalom x dobijaju se četiri izlazne sekvence y_1, y_2, y_3, y_4 i to na sledeći način:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = A \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ a & b & -b & -a \\ 1 & -1 & -1 & 1 \\ b & -a & a & -b \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}. \quad (4.2.13)$$

EksPLICITNO prikazano, sekvence signala posle transformacije imaju oblik:

$$\begin{aligned} y_1 &= \frac{1}{2}(x_1 + x_2 + x_3 + x_4), \\ y_2 &= \frac{1}{2}(ax_1 + bx_2 - bx_3 - ax_4), \\ y_3 &= \frac{1}{2}(x_1 - x_2 - x_3 + x_4), \\ y_4 &= \frac{1}{2}(bx_1 - ax_2 + ax_3 - bx_4). \end{aligned} \quad (4.2.14)$$

Varijanse signala y_1, y_2, y_3 i y_4 ($\sigma_1^2, \sigma_2^2, \sigma_3^2$ i σ_4^2 , respektivno) zavise od varijanse ulaznog signala u frejmu σ_x^2 , kao i od korelacije između dva susedna odmerka ρ_1 (korelacija između x_1 i x_2), korelacije između dva nesusedna odmerka ρ_2 (korelacija između x_1 i x_3) i korelacije između dva nesusedna odmerka između kojih postoje još dva odmerka ρ_3 (korelacija između x_1 i x_4) i definisane su sledećim jednačinama:

$$\begin{aligned} \sigma_1^2 &= \sigma_x^2 \left(1 + \frac{3}{2}\rho_1 + \rho_2 + \frac{1}{2}\rho_3 \right), \\ \sigma_2^2 &= \frac{1}{4}\sigma_x^2 (2a^2 + 2b^2 + (4ab - 2b^2)\rho_1 - 4ab\rho_2 - 2a^2\rho_3), \\ \sigma_3^2 &= \sigma_x^2 \left(1 - \frac{1}{2}\rho_1 - \rho_2 + \frac{1}{2}\rho_3 \right), \\ \sigma_4^2 &= \frac{1}{4}\sigma_x^2 (2a^2 + 2b^2 - (4ab + 2a^2)\rho_1 + 4ab\rho_2 - 2b^2\rho_3). \end{aligned} \quad (4.2.15)$$

S obzirom da su u granama za kodovanje sekvenci y_1, y_2, y_3 i y_4 angažovani adaptivni kvantizeri, procenitelj varijansi računa varijansu (standardnu devijaciju) signala u frejmu, čiju vrednost šalje na kodovanje log-uniformnom kvantizeru. Dalje se ta kodovana varijansa signala u frejmu šalje ka prijemniku kao dodatna informacija. Adaptivni kvantizeri AQ_1, AQ_2, AQ_3 i AQ_4 bazirani na μ -logaritamskom zakonu čija je kompresorska funkcija data jednačinom (4.1.3), prilagođavaju se signalima y_1, y_2, y_3 i y_4 respektivno, tako da se njihovi regioni rada (maksimalne amplitude opterećenja) dobijaju kao proizvod maksimalne amplitude opterećenja kvantizera kada se na njegov ulaz dovede signal modelovan Laplasovom funkcijom raspodele gustine verovatnoće jedinične varijanse i kodovane standardne devijacije signala po granama [1], [7]:

$$\begin{aligned}
x_{\max 1} &= x_{\text{ml}} \cdot \hat{\sigma}_1, \\
x_{\max 2} &= x_{\text{ml}} \cdot \hat{\sigma}_2, \\
x_{\max 3} &= x_{\text{ml}} \cdot \hat{\sigma}_3, \\
x_{\max 4} &= x_{\text{ml}} \cdot \hat{\sigma}_4,
\end{aligned} \tag{4.2.16}$$

gde su $\hat{\sigma}_1$, $\hat{\sigma}_2$, $\hat{\sigma}_3$ i $\hat{\sigma}_4$ kvantovane vrednosti standardne devijacije, dobijene od log-uniformnog kvantizera, a x_{ml} već pomenuta maksimalna amplituda kvantizera za signal sa Laplasovom raspodelom dobijena iz fomule (4.1.8). Na ovaj način dobijene su kvantovane sekvence \hat{y}_1 , \hat{y}_2 , \hat{y}_3 i \hat{y}_4 koje primenom inverzne diskretne kosinusne transformacije formiraju izlazni signal \hat{x} na način:

$$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \end{bmatrix} = A^{-1} \cdot \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \\ \hat{y}_4 \end{bmatrix}. \tag{4.2.17}$$

Inverzna matrica koeficijenata DCT-a 4×4 ima oblik [60], [97] – [103]:

$$A^{-1} = \frac{1}{2} \begin{bmatrix} 1 & a & 1 & b \\ 1 & b & -1 & -a \\ 1 & -b & -1 & a \\ 1 & -a & 1 & -b \end{bmatrix}, \tag{4.2.18}$$

pa se odmerci izlaznog signala \hat{x} dobijaju:

$$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & a & 1 & b \\ 1 & b & -1 & -a \\ 1 & -b & -1 & a \\ 1 & -a & 1 & -b \end{bmatrix} \cdot \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \\ \hat{y}_4 \end{bmatrix}. \tag{4.2.19}$$

Eksplisitni prikazi odmeraka izlaznog signala \hat{x} dati su sledećim jednačinama:

$$\begin{aligned}
\hat{x}_1 &= \frac{1}{2} (\hat{y}_1 + a\hat{y}_2 + \hat{y}_3 + \hat{y}_4), \\
\hat{x}_2 &= \frac{1}{2} (\hat{y}_1 + b\hat{y}_2 - \hat{y}_3 - a\hat{y}_4), \\
\hat{x}_3 &= \frac{1}{2} (\hat{y}_1 - b\hat{y}_2 - \hat{y}_3 + a\hat{y}_4), \\
\hat{x}_4 &= \frac{1}{2} (\hat{y}_1 - a\hat{y}_2 + \hat{y}_3 - b\hat{y}_4).
\end{aligned} \tag{4.20}$$

Kao i u svim prethodno prikazanim šemama, mera kvaliteta kodovanja je SQNR, koji se računa za svaki frejm korišćenjem formula (4.1.10) i (4.1.11) i na kraju prosečni SQNR svih frejmova za ceo signal. Izbor optimalne bit alokacije je od velikog značaja za postizanje maksimalnog kvaliteta kodovanja, te se i ovde koristi jednačina (4.1.15) za računanje optimalnih bitskih brzina adaptivnih kvantizera u granama [1], [7].

Eksperiment je urađen za vrednosti faktora kompresije $\mu=20$ i $\mu=255$ i bitske brzine kvantizera u granama dobijenih iz formule (4.1.15) koje zavise od srednje bitske brzine:

$$\begin{aligned} R_1 &= R_{sr} + 2, \\ R_2 &= R_{sr} - 1, \\ R_3 &= R_{sr} - 1, \\ R_4 &= R_{sr} - 2 \end{aligned} \quad (4.2.21)$$

Kako bi se pokazala pogodnost korišćenja kodne šeme sa slike 4.2.2, dobijeni SQNR upoređen je sa SQNR-om dobijenim korišćenjem kodne šeme sa slike 4.2.1, kao i SQNR-om G.711 standarda. Rezultati su prikazani za srednju bitsku brzinu $R_{sr}=5, 6, 7, 8$ bita po odmerku.

Tabela 4.2.2. Eksperimentalni rezultati (SQNR) dobijeni koristeći kodnu šemu sa slike 4.2.2, šemu baziranu na DCT 3×3 , šemu sa dvostrukim H_1 transformacionim kodovanjem, šemu sa jednostrukim H_1 transformacionim kodovanjem i impulsnu kodnu modulaciju

R_{sr} [bita po odmerku]	μ	SQNR[dB]	SQNR $_{3 \times 3}$ [dB]	SQNR $_{H_2}$ [dB]	SQNR $_{H_1}$ [dB]	SQNR $_{PCM}$ [dB]
5	20	30.5984	28.7478	30.1968	27.9027	23.4112
	255	27.3692	27.0165	27.1547		
6	20	36.0576	34.2541	36.0036	33.7459	29.2385
	255	33.5828	33.0727	33.1343		
7	20	41.5176	40.0320	41.5876	39.2063	35.1832
	255	39.0174	38.9444	39.0757		
8	20	47.2364	45.6249	47.7315	45.1963	41.0952
	255	45.0508	44.9313	45.3552		

U tabeli 4.2.2 prikazani su eksperimentalni rezultati u vidu SQNR-a dobijeni korišćenjem šeme sa slike 4.2.2 za kodovanje govornog signala, šeme 4.2.1 za kodovanje govornog signala, kao i SQNR prethodno prikazanih kodnih šema koje koriste prostu (modifikovanu) Hadamarovu transformaciju za transformaciono kodovanje i SQNR dobijen kodovanjem govornog signala impulsnom kodnom modulacijom. Ulazni test govorni signal snimljen je u laboratoriji za Akustiku na Elektronskom fakultetu u Nišu, odmeren na 8 kHz isti je kao i test signali nad kojima su urađeni eksperimenti u prethodnoj kodnoj šemi radi adekvatnog poređenja. Trajanje ulaznog test signala je oko jedan minut, dok varijansa ima vrednost 0.9068. Iz tabele se vidi da primena diskretne kosinusne transformacije sa matricom koeficijenata veličine 4×4 obezbeđuje približno isti kvalitet kodovanja za posmatrani faktor kompresije $\mu=20$ kao i kodna šema sa slike 4.1.3 (dvostruka Hadamarova transformacija H_1), dok pruža oko 2 dB viši SQNR u odnosu na šemu sa slike 4.1.1 (jednostruka prosta Hadamarova transformacija H_1). Za vrednost faktora kompresije $\mu=255$, kodna šema pruža rezultate približno jednake kao i kodna šema implementirana dvostrukom primenom Hadamarove

transformacije i 0.1 - 0.3 dB niži SQNR. Poređenjem rezultata šema implementiranih DCT-em 4×4 i 3×3 vidi se da šema koja ulazni signal deli na 4 podsekvence koje se koduju nezavisno pruža do 2 dB dobitka u kodovanju ali je dosta kompleksnija za implementaciju, što zbog samog transformacionog kodovanja što zbog dodavanja još jednog kvantizera. Iz ovoga se može zaključiti da ukoliko je zahtev sistema za kodovanje takav da se obezbedi visok kvalitet bez pitanja za cenu kompleksnosti implementacije kodne šeme onda je pogodno koristiti kodnu šemu sa slike 4.2.2 za kodovanje govornog signala, a ako je slučaj da su zahtevi takvi da kompleksnost implementacije kodne šeme bude niska, što bi značilo neznatno smanjenje kvaliteta, za kodovanje govornog signala trebalo bi koristiti šemu sa slike 4.2.1. Kvalitet kodovanja naravno varira, za istu srednju bitsku brzinu, u zavisnosti od vrednosti faktora kompresije jer za faktor $\mu=20$ šema pruža oko 2 dB viši kvalitet kodovanja u odnosu na $\mu=255$, što je normalno zato što snižavanjem ove vrednosti sužava se karakteristika kompresorske funkcije, dok se njenim povećanjem dobija na robustnosti kvantizera.

4.3 Šeme za kodovanje diskretnog govornog signala

U ovoj sekciji prikazane su šeme za kodovanje govornog signala koje pokazuju značaj diskretizacije ulaznog signala pre kodovanja, implementirane tako da predstavljaju “*two-stage*” sistem za kodovanje (sistem dva stanja) jer se mogu razdvojiti deo za diskretizaciju ulaznog signala i ostatak šeme za kodovanje.

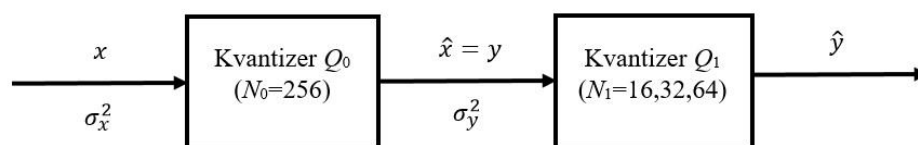
Najznačajniji deo šema za kodovanje diskretnog signala predstavlja projektovanje kvantizera kojim se vrši kodovanje, te je shodno tome, u radu [74] akcenat stavljen upravo na to. U radu [74] urađena je analiza performansi kvantizera projektovanog za kodovanje korelisanog diskretnog signala. Teorijska istraživanja su urađena za slučaj da se na ulaz kvantizera dovede signal modelovan Laplasovom funkcijom raspodele gustine verovatnoće, dok je za eksperiment korišćen test govorni signal odmeren na 16 kHz [74]. Sam kvantizer predstavlja drugi korak dvofaznog (“*two-stage*”) sistema u kome je u prvom koraku urađena diskretizacija kontinualnog ulaznog signala, dok drugi korak obezbeđuje dodatnu kompresiju.

“*Mismatch*” kvantizacija je poznata tehnika koja se koristi za obradu različitih tipova signala [76], [114] – [116]. Sastoji se iz dva koraka od kojih se u prvom koraku određuje region rada kvantizera iz drugog koraka, dok su kvantizacioni nivoi, smešteni u tom regionu, određeni u drugom koraku. Osnovni model “*mismatch*” dvofaznog kvantizacionog sistema predložen je u radu [75], gde se za diskretizaciju koristi uniformni kvantizer, dok se za dodatnu kompresiju

koristi neuniformni (optimalni komparator). Prvi korak u takvom sistemu je diskretizacija kontinualnog signala kvantizerom sa velikim brojem nivoa, čiji su izlaz diskretni odmerci signala. Zatim, ti diskretni odmerci kvantuju se kvantizerom sa malim brojem nivoa u cilju dodatne kompresije. Dovođenjem diskretnih odmeraka signala na ulaz drugog kvantizera, koji je amplitudski ograničen, gubi se distorzija prekoračenja. Predložena kodna šema, za razliku od one u radu [75], projektovana je tako da se u oba koraka dvofaznog sistema koriste logaritamski kvantizeri, bazirani na μ -logaritamskom zakonu, kako za diskretizaciju kontinualnog govornog signala, tako i za dodatnu kompresiju u drugom koraku sistema [74]. Upotreba kvazi-logaritamskih kvantizera dolazi iz činjenice da oni pružaju veću robustnost u odnosu na optimalni komparator iz rada [75] zbog čega su pogodniji za kodovanje diskretnih odmeraka.

Performanse se obično u ovakvim sistemima daju na osnovu procene varijansi ulaznog kontinualnog signala dok se promene u varijansi posle prvog koraka odbacuju. Međutim, česta je pojava u sistemima da informacija o kontinualnom ulazu ne postoji, već je prisutna samo informacija o diskretizovanim odmercima. U radu [74] pokazano je da je informacija o varijansi diskretnog signala posle prve faze od velikog značaja za dalju kompresiju. Performanse su prikazane kroz SQNR dobijen na način detaljno objašnjen u odeljku 4.1. Predložena kodna šema za kodovanje diskretnog govornog signala prikazana je na slici 4.3.1.

Kao što je već napomenuto, diskretizacija kontinualnog signala vrši se kvantizerom sa velikim brojem nivoa. Za ovu namenu korišćen je kvazilogaritamski kvantizer Q_0 sa $N_0=256$ kvantizacionih nivoa. U teorijskoj obradi, signal je modelovan Laplasovom funkcijom raspodele gustine verovatnoće čiji je izraz dat kroz jednačinu (2.1.2).



Slika 4.3.1. Šema za kodovanje diskretnog govornog signala implementirana korišćenjem jednog kvazi-logaritamskog kvantizera

Kompresorska funkcija μ -logaritamskog zakona na kome su zasnovana oba kvantizera prikazana je jednačinom (2.2.34) [1], [7], [66]. Prema μ -logaritamskoj kompresorskoj funkciji reprezentacioni nivoi x_i i pragovi odluke y_i kvantizera Q_0 dobijeni su na sledeći način:

$$x_i = \frac{x_{\max}}{\mu} \left(\exp \left(\frac{2i}{N_0} \ln(1 + \mu) \right) - 1 \right), i = 1, 2, \dots, N_0,$$

$$y_i = \frac{x_{\max}}{\mu} \left(\exp \left(\frac{2(i - 0.5)}{N_0} \ln(1 + \mu) \right) - 1 \right), i = 1, 2, \dots, N_0.$$
(4.3.1)

gde je x_{\max} maksimalna amplituda opterećenja kvantizera Q_0 . Varijansa kontinualnog signal x , σ_x^2 uzima vrednosti iz opsega (-30 dB, 30 dB) koje se konvertuju u realne vrednosti formulom:

$$\sigma_x^2 = \sigma_{\text{reff}}^2 \cdot 10^{\frac{\sigma_x^2[\text{dB}]}{10}},$$
(4.3.2)

dok se varijansa diskretnog signala dobija iz formule:

$$\sigma_y^2 = 2 \cdot \sum_{i=1}^{N_0/2} y_i^2 P_i.$$
(4.3.3)

U jednačini (4.3.2) σ_{reff}^2 označava referentnu varijansu, dok u jednačini (4.3.3) P_i označava verovatnoću diskretnog ulaznog nivoa kvantizera Q_1 . Verovatnoće P_i dobijene su iz jednačine [75], [117]:

$$P_i = \int_{x_i}^{x_{i+1}} p(x) dx = \frac{1}{2} \left(\exp \left(-\frac{\sqrt{2}x_i}{\sigma_x} \right) - \exp \left(-\frac{\sqrt{2}x_{i+1}}{\sigma_x} \right) \right).$$
(4.3.4)

Cilj šeme jeste da druga faza sistema sa slike 4.3.1 obezbedi dodatnu kompresiju sa što manje potrebnih bita po odmerku. Diskretizovani izlaz kvantizera Q_0 predstavlja ulaz u kvantizer Q_1 , koji je takođe kvazi-logaritamski jer je diskretni ulazni signal nestacionaran kao i kontinualni, te je pogodno njegovo korišćenje zbog osobine velike robustnosti. Kvantizer u drugoj fazi, Q_1 , treba da se projektuje za manju bitsku brzinu i na osnovu rada [11] najbolja razlika u bitskim brzinama između kvantizera Q_0 i Q_1 je 4 bita po odmerku, što znači da bi kvantizer Q_1 trebalo projektovati za $R_1 = 4$ bita po odmerku. Pragovi odluke t_i i reprezentacioni nivoi w_i kvantizera Q_1 dobijeni su iz jednačine:

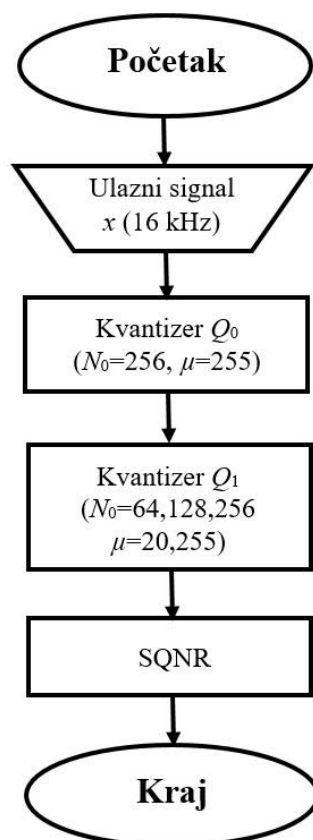
$$t_i = \frac{x_{\max 1}}{\mu_1} \left(\exp \left(\frac{2i}{N_1} \ln(1 + \mu_1) \right) - 1 \right), i = 1, 2, \dots, N_1,$$

$$w_i = \frac{x_{\max 1}}{\mu_1} \left(\exp \left(\frac{2(i - 0.5)}{N_1} \ln(1 + \mu_1) \right) - 1 \right), i = 1, 2, \dots, N_1.$$
(4.3.5)

gde je $x_{\max 1}$ maksimalna amplituda opterećenja kvantizera Q_1 , dok je N_1 broj kvantizacionih nivoa i važi ($N_0 > N_1$).

Performanse sistema dobijene su poređenjem tradicionalnog načina modelovanja kada se koristi samo varijansa ulaznog signal σ_x^2 i predložene “two-stage” kodne šeme kada se koristi

i varijansa diskretnog signala σ_y^2 za procenu performansi. Ovi teorijski dobijeni rezultati poređeni su sa eksperimentalnim koji su dobijeni obradom test govornog signala odmerenog na 16 kHz. Takođe, rezultati su upoređeni i sa rezultatima iz rada [75] gde su uniformni kvantizer i optimalni komparator upotrebljeni u prvoj i drugoj fazi, respektivno. Algoritam koji opisuje obradu ulaznog govornog signala prikazan je na slici 4.3.2.



Slika 4.3.2. Algoritam obrade signala za predloženu kodnu šemu sa slike 4.3.1

Kao što je napomenuto, kao mera kvaliteta kodovanja koriste se distorzija i SQNR. Za posmatrani sistem sa slike 4.3.1 ukupna distorzija predstavlja granularnu distorziju zbog amplitudskog ograničenja diskretnih odmeraka, odnosno odsustva distorzije prekoračenja. Distorzija se izračunava na osnovu izraza:

$$D_{uk}(\sigma_y^2(j)) = 2 \cdot \sum_{i=N_1/2}^{N_1} \sum_{j=1}^{r_i} (y_{ij} - w_i)^2 \cdot P(y_{ij}), \quad (4.3.6)$$

gde parameter r_i označava broj ulaznih nivoa mapiranih sa w_i , dok y_{ij} predstavlja izlazne nivoe kvantizera Q_0 , a N_1 broj kvantizacionih nivoa kvantizera Q_1 . Kvalitet rekonstruisanog signala izražen je kroz meru SQNR-a kao:

$$SQNR[\text{dB}] = 10 \log_{10} \left(\frac{\sigma_y^2}{D_{uk}} \right). \quad (4.3.7)$$

Unapređenje konstrukcije i određivanje sistemskih parametara bazirano je na poređenju teorijskih i eksperimentalnih rezultata za sve varijanse.

Dinamika ulaznog signala B određena je količnikom najviše i najniže varijanse frejma na koje je podeljen ulazni signal (svaki frejm $M=240$ odmeraka):

$$B[\text{dB}] = 10 \log \left(\frac{\sigma_{\max}^2(M)}{\sigma_{\min}^2(M)} \right). \quad (4.3.8)$$

Referentna varijansa postavljena je na 0 dB i to na sredini dinamičkog opsega, dok maksimalna i minimalna varijansa određuju opseg:

$$\begin{aligned} \sigma_{\max}^2[\text{dB}] &= \frac{B}{2}, \\ \sigma_{\min}^2[\text{dB}] &= -\frac{B}{2}. \end{aligned} \quad (4.3.9)$$

Maksimalna amplituda kvantizera Q_0 adaptirana je signalu na način:

$$x_{\max 0} = x_{\max}^{(\sigma^2=1)} \cdot \sigma_{\text{reff}}, \quad (4.3.10)$$

gde $x_{\max}^{(\sigma^2=1)}$ predstavlja regiona rada kvantizera kada se na njegov ulaz dovede signal modelovan Laplasovom funkcijom raspodele gustine verovatnoće jedinične varijanse čija je vrednost data jednačinom (4.1.8) [66].

Kao što je već pomenuto, za eksperiment je korišćen govorni test signal odmeren na 16 kHz, snimljen u laboratoriji za Akustiku na Elektronskom fakultetu u Nišu, čija varijansa ima vrednost $\sigma_x^2 = 0.0021$. Referentna varijansa iznosi $\sigma_{\text{reff}}^2 = 5.84 \cdot 10^{-5}$, dok je dinamika signala $B=54$ dB. Optimalna vrednost regiona rada kvantizera Q_0 iznosi $x_{\max 0}=0.0775$, dok optimalne vrednosti regiona rada kvantizera Q_1 zavise od sistemskih parametara (faktora kompresije μ i bitske brzine R_1).

S obzirom da referentna varijansa ima veoma nisku vrednost, opseg varijansi od interesa je (-30 dB, 30 dB). U svrhu poređenja ovaj opseg je podeljen u segmente veličine 2 dB, dok je ulazni signal podeljen na frejmove veličine $M=240$ odmeraka, zbog određivanja varijanse (standardne devijacije). Za svaki segment izračunat je nivo na sledeći način:

$$L_i[\text{dB}] = 10 \cdot \log_{10} \left(\frac{\hat{\sigma}_{y_i}^2}{\sigma_{\text{reff}}^2} \right), \quad (4.3.11)$$

gde je $\hat{\sigma}_{y_i}^2$ nivo varijanse i -tog segmenta. Nakon određivanja vrednosti L_i , urađena je kvantizacija na oba načina – tradicionalnim sistemom kvantovanja samo kontinualnog signala kada je bitska brzina $R=4,5,6$ bita po odmerku i predloženim algoritmom sa parametrima $R_0=8$

bita po odmerku i $R_1=4,5,6$ bita po odmerku. Za svaki segment u opsegu ulaznih varijansi izbrojano je koliko se frejmova javlja unutar svakog opsega za posmatranu varijansu, nakon čega je dobijen srednji SQNR svih frejmova u okviru segmenta, za kontinualni i diskretni signal. Srednje vrednosti SQNR-a po segmentu označene su sa $SQNR(L_i)$ i za i -ti segment se dobijaju:

$$SQNR(L_i)[dB] = \frac{1}{M_i} \sum_{k=1}^{M_i} SQNR(\sigma_{y_{ik}}^2)[dB], \quad (4.3.12)$$

dok je ukupni srednji SQNR celog opsega ulaznih varijansi (za sve segmente):

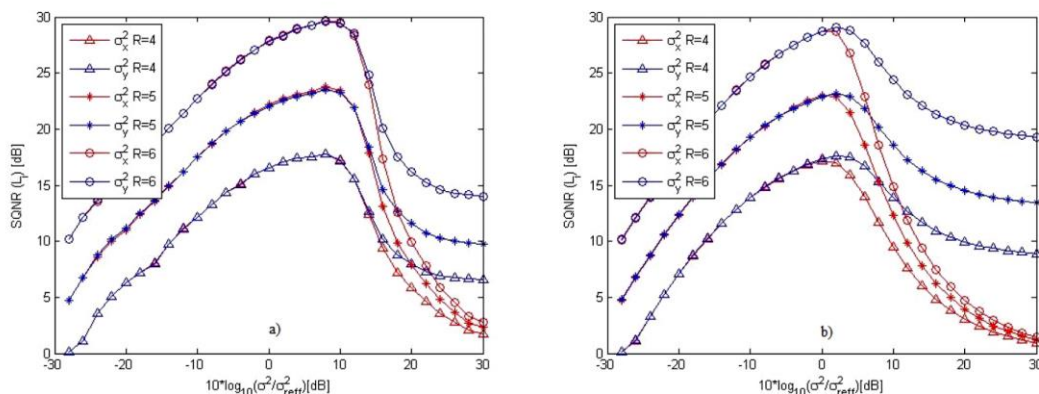
$$SQNR_{sr}[dB] = \frac{1}{L} \sum_{k=1}^L SQNR(\sigma_{y_k}^2)[dB]. \quad (4.3.13)$$

Ovim načinom merenja performansi kodovanja dobija se adekvatno poređenje teorijskih i eksperimentalnih rezultata. Standardni srednji SQNR za ceo dinamički opseg, bez podele na segmente, dobija se:

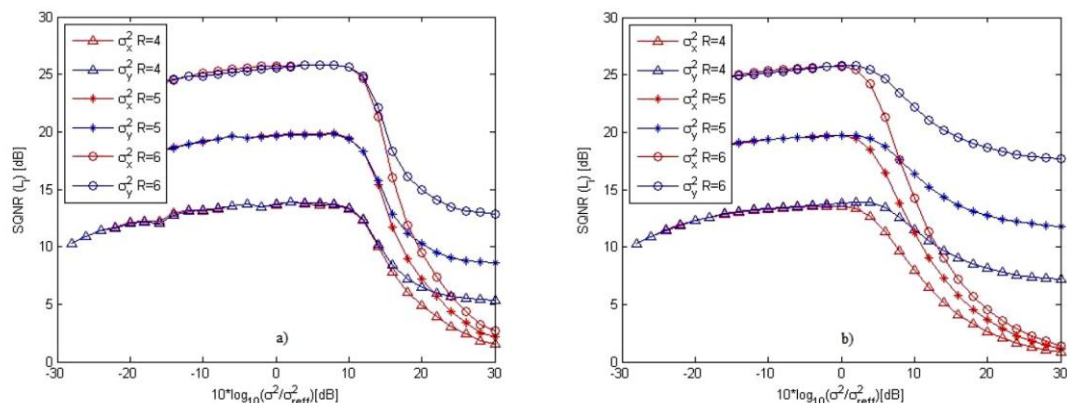
$$SQNR_{sr}[dB] = \frac{1}{F} \sum_{k=1}^F SQNR(k)[dB]. \quad (4.3.14)$$

gde je F ukupni broj frejmova na koje je podeljen ulazni signal.

Na slikama 4.3.3 i 4.3.4 prikazane su performanse teorijskih istraživanja i eksperimentalnih rezultata u širokom opsegu ulaznih varijansi za razne vrednosti faktora kompresije μ .



Slika 4.3.3. SQNR u širokom opsegu ulaznih varijansi za $\mu=20$ a) Eksperimentalni rezultati b) Teorijski rezultati



Slika 4.3.4. SQNR u širokom opsegu ulaznih varijansi za $\mu=255$ a) Eksperimentalni rezultati b) Teorijski rezultati

Sa slika 4.3.3 i 4.3.4 vidi se odlično podudaranje eksperimentalnih i teorijskih rezultata. Razlika između tradicionalnog kodovanja samo kontinualnog signala i korišćenja predloženog algoritma vidi se pri većim varijansama ulaznog signala i upravo u ovim oblastima je dobitak u kvalitetu kodovanja korišćenjem predloženog algoritma.

Tabela 4.3.1. Srednji SQNR u širokom opsegu varijansi ulaznog signala (-30 dB, 30 dB)

R [bita po odmerku]	μ	Teorijski rezultati		Eksperimentalni rezultati	
		$\text{SQNR}_{sr}^{\sigma_x^2}$ [dB]	$\text{SQNR}_{sr}^{\sigma_y^2}$ [dB]	$\text{SQNR}_{sr}^{\sigma_x^2}$ [dB]	$\text{SQNR}_{sr}^{\sigma_y^2}$ [dB]
4	20	8.5325	11.2632	9.7321	10.5400
	255	9.0984	11.2908	10.3218	10.9580
5	20	12.3055	16.3280	14.1703	15.4675
	255	13.2782	16.6781	15.1109	16.2019
6	20	16.0018	21.9298	18.6834	20.7717
	255	17.3339	22.5549	19.9489	21.7933

Iz tabele 4.3.1 može se zaključiti da su performanse predloženog teorijskog algoritma u vidu prosečnog SQNR-a veoma bliske eksperimentalnim rezultatima kao i da eksperimentalni i teorijski rezultati prate promene parametara na isti način. Može se videti da je prosečna razlika u performansama eksperimentalnih i teorijskih rezultata 1.3 dB, kao i da je prosečni SQNR diskretnog ulaznog signala 2.2 – 5.2 dB viši u odnosu na kontinualni za teorijsku analizu i 0.6 – 2.1 dB za eksperiment, u zavisnosti od bitske brzine i faktora kompresije [74]. U cilju adekvatne kompresije sa rezultatima prikazanim u radu [75], gde se umesto kvazi-logaritamskog kvantizera Q_0 koristi uniformni a umesto kvazi-logaritamskog kvantizera Q_1 koristi optimalni komandor, analiza je urađena u opsegu ulaznih varijansi (-20 dB, 20 dB). U

radu [75], bitske brzine za koje je projektovan Q_1 su $R_1=4,5$ bita po odmerku, te su u tabeli 4.3.2 rezultati poređenja prikazani za ove bitske brzine.

Tabela 4.3.2. Srednji SQNR u širokom opsegu varijansi ulaznog signala (-20 dB, 20 dB)

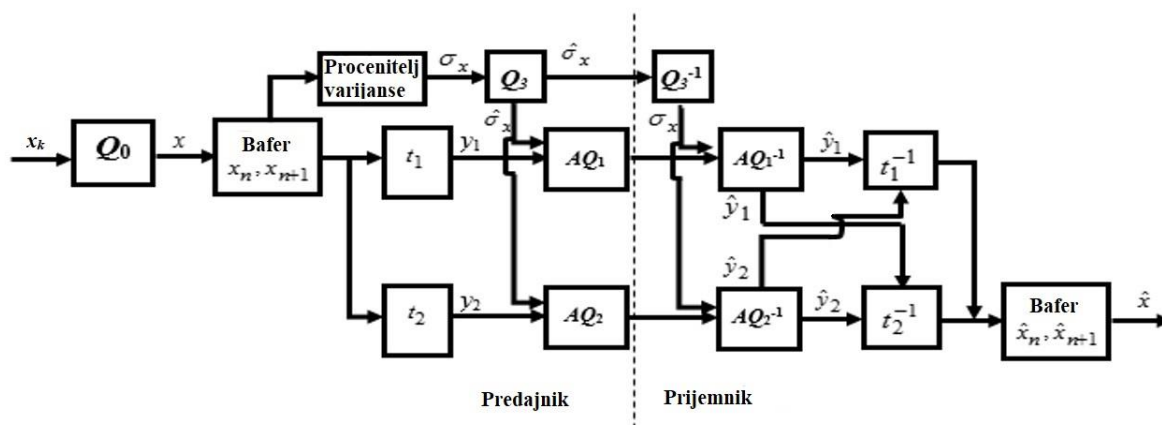
R [bita po odmerku]	μ	Teorijski rezultati [dB]		Eksperimentalni rezultati [dB]		Rezultati iz rada [75]	
		$SQNR_{sr}^{\sigma_x^2}$	$SQNR_{sr}^{\sigma_y^2}$	$SQNR_{sr}^{\sigma_x^2}$ [dB]	$SQNR_{sr}^{\sigma_y^2}$ [dB]	$SQNR_{sr}^{\sigma_x^2}$	$SQNR_{sr}^{\sigma_y^2}$
4	20	11.7611	13.6013	12.7228	12.9509	11.0023	12.6012
	255	10.9441	12.8421	12.0344	12.2179		
5	20	16.2082	18.7768	17.8673	18.2416	15.6085	17.6616
	255	15.8721	17.9857	17.5601	17.8639		

Rezultati prikazani u tabeli 4.3.2 potvrđuju pogodnost korišćenja kvazi-logaritamskih kvantizera u drugoj fazi “*two-stage*” sistema, upravo zbog robustnosti u širokom opsegu ulaznih varijansi. Korišćenje kvazi-logaritamskih kvantizera umesto optimalnog komandora, kao što je prikazano u radu [75], pruža oko 1 dB viši SQNR, poredeći teorijski dobijene rezultate sa onim u radu [75], za obe bitske brzine i vrednost faktora kompresije $\mu=20$, i oko 0.2-0.6 dB za vrednost faktora kompresije $\mu=255$.

U ovoj disertaciji prikazana je i šema za kodovanje diskretnog govornog signala bazirana na primeni proste (modifikovane) Hadamarove transformacije H_1 i tehnike adaptacije unapred. Kao i u prethodno predloženoj šemi (“*two-stage*” sistem za kodovanje), kvazi-logaritamski kvantizer Q_0 iskorišćen je za diskretizaciju ulaznog govornog signala, odnosno pripremu diskretizovanog signala, dok je druga faza sistema implementirana koristeći kodnu šemu sa slike 4.1.1. Kao i u šemi sa slike 4.1.1 upotreba tehnike adaptacije unapred obezbeđuje efikasniju upotrebu propusnog opsega dok transformaciono kodovanje dekomponuje signal u podsekvence sa predviljivijim karakterisikama, što omogućuje viši kvalitet rekonstrukcije signala u prijemniku. I u ovoj kodnoj šemi performanse su analizirane za dva tipična slučaja – kada postoji informacija o varijansi kontinualnog signala na ulazu i kada je dostupna informacija samo o varijansi diskretnog signala, što znači da postoji gubitak informacije ulaznog signala. Pored digitalizacije i kompresije govornog signala, skalarna kvantizacija je pogodna za upotrebu kod algoritama za prepoznavanje govora korišćenjem dubokih konvolucionih neuronskih mreža [118] – [124].

Diskretni signal, dobijen kao izlaz kvantizera Q_0 , amplitudski je ograničen što znači da se optimalnim projektovanjem kvantizera u ostatku kodne šeme može zanemariti distorzija prekoračenja [1], [7]. Tako diskretni signal se, upotrebom transformacionog kodovanja (prosta

Hadamarova transformacija H_1), dekomponuje na podsekvence sa užim propusnim opsegom nad kojima se primenjuje tehnika adaptacije unapred. Cilj upotrebe transformacionog kodovanja je obezbeđivanje dodatne kompresije pre adaptacije unapred [1], [59], [87]. Diskretni signal se u baferu deli na frejmove veličine $M=240$ odmeraka nad kojima se računa varijansa (standardna devijacija) koja se dalje koduje log-uniformnim kvantizerom i prenosi do prijemnika kao dodatna informacija zbog upotrebe adaptacije unapred [1], [7], [66]. Šema za kodovanje diskretnog govornog signala koja koristi transformaciono kodovanje i tehniku adaptacije unapred prikazana je na slici 4.3.5.



Slika 4.3.5. Šema za kodovanje diskretnog govornog signala koja koristi transformaciono kodovanje i tehniku adaptacije unapred

Kao i u prethodno prikazanoj šemi za kodovanje diskretnog govornog signala, koristi se kvantizer Q_0 za diskretizaciju kontinualnog govornog signala zasnovan na μ -logaritamskom zakonu projektovan za bitsku brzinu $R_0=8$ bita po odmerku i faktor kompresije $\mu=255$. Izlazni odmerci kvantizera Q_0 vode se u bafer gde se signal deli na frejmove veličine $M=240$ odmeraka. Iako se koristi robustni kvazi-logaritamski kvantizer može doći do velike greške za visoke vrednosti varijanse, što rezultuje razlikom u performansama kada je informacija o kontinualnom signalu prisutna i kada postoji informacija samo o diskretnim odmercima. Zbog toga je poželjno koristiti dva kvantizera u prvoj fazi “two-stage” sistema, jedan za niske i jedan za visoke varijanse. Međutim, eksperiment je urađen za slučaj jednog kvazi-logaritamskog kvantizera u prvoj fazi.

Nakon diskretizacije i podele diskretnog signala na frejmove u baferu, primenjuje se prosta modifikovana Hadamarova transformacija H_1 deleći signal na podsekvence y_1 i y_2 koristeći jednačinu (4.1.1). Transformisani signali y_1 i y_2 imaju varijanse σ_1^2 i σ_2^2 koje zavise od varijanse ulaznog diskretnog signala x i koeficijenta korelacije ρ , koje se dobijaju koristeći formulu

(4.1.2) [59]. Tako dobijene podsekvence se dalje vode na adaptivne kvantizere AQ_1 i AQ_2 . Kako se koriste adaptivni kvantizeri bazirani na tehnici adaptacije unapred, varijansa signala u frejmu mora biti kodovana i za to se koristi log-uniformni kvantizer Q_3 , detaljno opisan u odeljku 2.3.2 [1], [7]. Na osnovu varijanse (standardne devijacije) signala u frejmu dobijaju se vrednosti maksimalnih amplituda adaptivnih kvantizera u granama, $x_{\max 1}$ i $x_{\max 2}$, koristeći jednačinu (4.1.7). Adaptivni kvantizeri u granama zasnovani su na μ -logaritamskom zakonu čija kompresorska funkcija je data jednačinom (4.1.3) [1], [59], [69], [87]. Pragovi odluke x_i i reprezentacioni nivoi y_i adaptivnih kvantizera dobijeni su jednačinom (4.1.4) [1], [59], [69], [87]. Kao i u svim prethodno prikazanim šemama koje koriste transformaciono kodovanje cilj je postizanje maksimalnog kvaliteta kodovanja, tako da je od velike važnosti projektovati kvantizere u granama za optimalne bitske brzine. Tako se i ovde za optimalnu bit alokaciju koristi jednačina (4.1.15) [7], dok su bitske brzine adaptivnih kvantizera AQ_1 i AQ_2 , R_1 i R_2 , respektivno:

$$\begin{aligned} R_1 &= R_{sr} + \frac{1}{4} \log_2 \left(\frac{1 + \rho}{1 - \rho} \right), \\ R_2 &= R_{sr} + \frac{1}{4} \log_2 \left(\frac{1 - \rho}{1 + \rho} \right). \end{aligned} \quad (4.3.15)$$

Iz fomule (4.3.15) se može zaključiti da je optimalna bitska brzina kvantizera AQ_1 za 1 bit po odmerku veća od srednje bitske brzine, dok je bitska brzina kvantizera AQ_2 za 1 bit po odmerku manja od srednje bitske brzine. Opseg ulaznih varijansi (-30 dB, 30 dB) je podeljen na segmente od po 2 dB kao u prethodnoj šemi, dok je na osnovu najviše i najniže varijanse po frejmovima koristeći formulu (4.3.8) izračunat dinamički opseg B . Za svaki frejm ulaznog signala izračunat je nivo L_i koristeći formulu (4.3.11) kako bi se odredilo kom segmentu dinamičkog opsega pripada i -ti frejm. Za svaki segment iz dinamičkog opsega izbrojano je koliko frejmova ima nivo koji pripada datom segmentu i izračunat je srednji SQNR za svaki frejm u i -tom segmentu na osnovu formule:

$$\text{SQNR}(\sigma_{y^{(i)}}^2)[\text{dB}] = 10 \cdot \log_{10} \left(\frac{\sigma_{x^{(i)}}^2}{D(\sigma_{y^{(i)}}^2)} \right) [\text{dB}], \quad (4.3.16)$$

gde je distorzija:

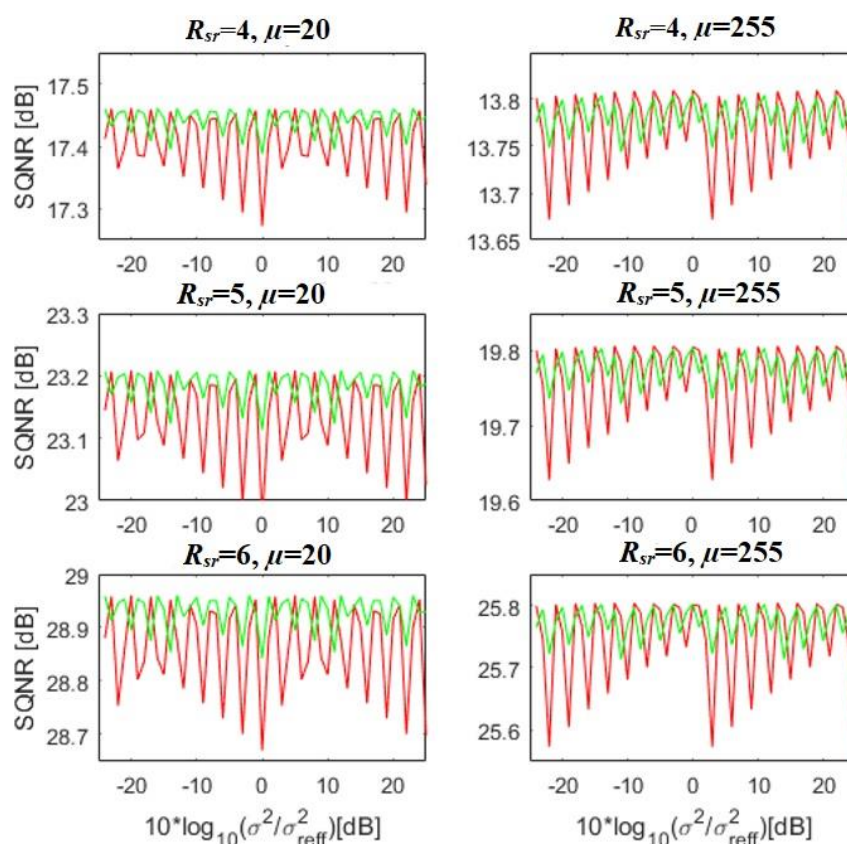
$$D(\sigma_{y^{(i)}}^2) = \sum_{k=1}^M (x_k - \hat{x}_k)^2. \quad (4.3.17)$$

SQNR i -tog segmenta je prosečni SQNR svih frejmova alociranih u tom segmentu:

$$\text{SQNR}(L_i)[\text{dB}] = \frac{1}{M_i} \sum_{m=1}^{M_i} \text{SQNR}(\sigma_{y(m)}^2) [\text{dB}], \quad (4.3.18)$$

dok je ukupni srednji SQNR dobijen koristeći formulu (4.3.13) [7], [59], [66], [86].

Eksperiment je urađen za $R_1=5,6,7$ bita po odmerku i $R_2=3,4,5$ bita po odmerku i faktor kompresije $\mu=20, 255$, dok je bitska brzina log-uniformnog kvantizera $R_3=4,5$ bita po odmerku. Na slici 4.3.6 prikazan je teoretski dobijen SQNR u širokom opsegu ulaznih varijansi kvazi-logaritamskog kvantizera, bez implementacije transformacionog kodovanja na kojoj se vidi dobitak kodovanja za razne vrednosti srednje bitske brzine.



Slika 4.3.6. SQNR u širokom opsegu ulaznih varijansi za $R_{sr}=4,5,6$ bita po odmerku, razne vrednosti μ i R_3 ($R_3=4$ - crvena linija, $R_3=5$ - zelena linija)

Sa slike 4.3.6 se vidi da povećanjem bitske brzine log-uniformnog kvantizera Q_3 za 1 bit po odmerku performanse sistema postaju bolje i grafički gledano linearnije. 1 bit po odmerku log-uniformnog kvantizera Q_3 ne povećava kompleksnost šeme u velikoj meri te je zbog konstantnijih performansi kodovanja pogodnije projektovati log-uniformni kvantizer za višu bitsku brzinu. U tabeli 4.3.3 prikazan je teoretski dobijen SQNR za sistem sa slike 4.3.5, sa dodatnom kolonom neadaptivne kvantizacije u cilju utvrđivanja pogodnosti korišćenja predložene kodne šeme.

Tabela 4.3.3. Teoretski dobijen SQNR predložene kodne šeme sa i bez primene adaptacije unapred

R [bita po odmerku]	R ₃ = 4 bita po odmerku		R ₃ = 5 bita po odmerku		Neadaptivna kvantizacija	
	SQNR _{μ=255} [dB]	SQNR _{μ=20} [dB]	SQNR _{μ=255} [dB]	SQNR _{μ=20} [dB]	SQNR _{μ=255} [dB]	SQNR _{μ=20} [dB]
4	13.8025	17.4037	13.8080	17.4406	11.2086	11.7467
5	19.7970	23.1321	19.8045	23.1819	15.7951	15.7195
6	25.7909	28.8026	25.8009	28.9296	20.2153	19.6486

Iz tabele 4.3.3 može se videti da uključivanjem tehnike adaptacije pružaju 2.55 – 9.2 dB dobitka u vidu SQNR-a, u zavisnosti od bitske brzine i faktora kompresije, u odnosu na neadaptivnu kvantizaciju što se smatra značajnim dobitkom. Poređenjem teoretskih rezultata sa onim dobijenim u radu [74] dokazuje se pogodnost korišćenja predložene kodne šeme koja pruža preko 3 dB viši srednji SQNR za istu bitsku brzinu.

U tabelama 4.3.4 i 4.3.5 prikazani su rezultati eksperimenta za različite vrednosti srednje brzine $R_{sr} = R + R_v$, gde je $R = (R_1 + R_2)/2$, a $R_v = R_3/M$. Tabela 4.3.4 prikazuje vrednosti SQNR-a dobijenog korišćenjem šeme sa slike 4.3.5, kada je kvantizer AQ_1 projektovan za brzine $R_1 = 4, 5, 6$ bita po odmerku, kvantizer AQ_2 za $R_2 = 2, 3, 4$ bita po odmerku a log-uniformni kvantizer Q_3 za $R_3 = 4$ bita po odmerku. Glavni razlog izbora ovih bitskih brzina je adekvatno poređenje sa rezultatima dobijenim u radu [74], gde kodna šema ne uključuje transformaciono kodovanje i adaptaciju unapred.

Tabela 4.3.4. Eksperimentalni rezultati predložene kodne šeme za $R_3 = 4$ bita po odmerku

μ	Ulazni signal	SQNR [dB]	R _{sr} [bita po odmerku]		
			4.01667	5.01667	6.01667
20	Kontinualni	SQNR _{sr}	20.0062	25.6692	31.1689
	Diskretni	SQNR _{sr}	20.7713	27.3206	32.0262
	Kontinualni	SQNR _{sr(st)}	20.7398	26.4888	32.2254
	Diskretni	SQNR _{sr(st)}	21.5754	27.7954	33.1885
255	Kontinualni	SQNR _{sr}	16.9248	22.9970	28.9011
	Diskretni	SQNR _{sr}	18.3151	24.5445	30.4459
	Kontinualni	SQNR _{sr(st)}	17.3997	23.3876	29.4358
	Diskretni	SQNR _{sr(st)}	19.2437	24.6927	30.8136

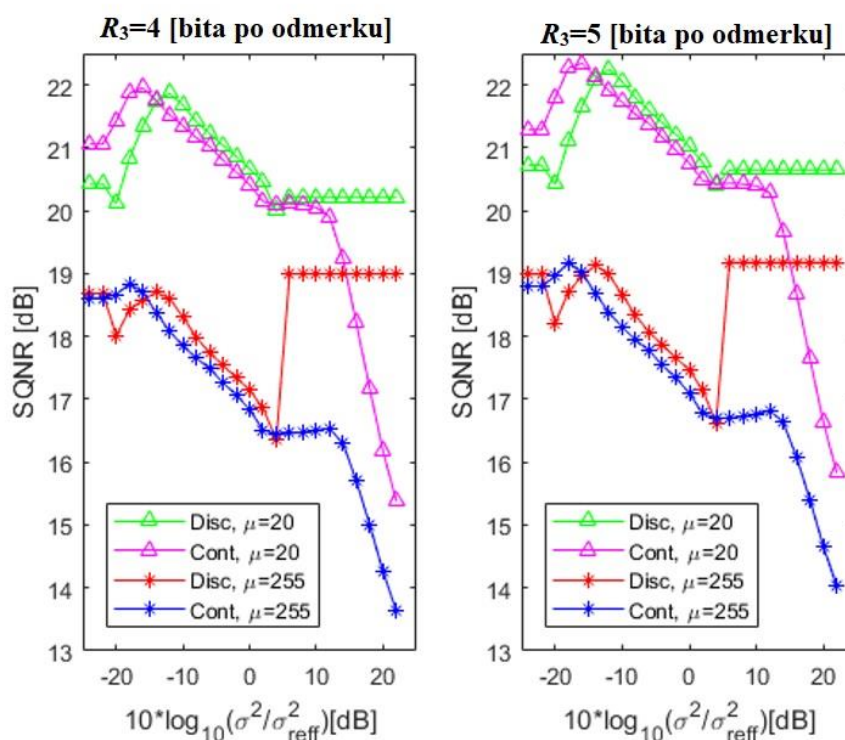
Iz tabele 4.3.4 se vidi da predložena kodna šema sa slike 4.3.5 pruža 4.9 – 5.9 dB viši SQNR u slučaju vrednosti faktora kompresije $μ = 255$ i 5.8 – 7.8 dB u slučaju $μ = 20$ za razne bitske brzine. Vrednosti SQNR_{sr(st)} su dobijene standardnim proračunom performansi sistema, koristeći jednačinu 4.3.14 dok su vrednosti SQNR_{sr} dobijene jednačinom 4.3.13.

Tabela 4.3.5. Eksperimentalni rezultati predložene kodne šeme za $R_3=5$ bita po odmerku

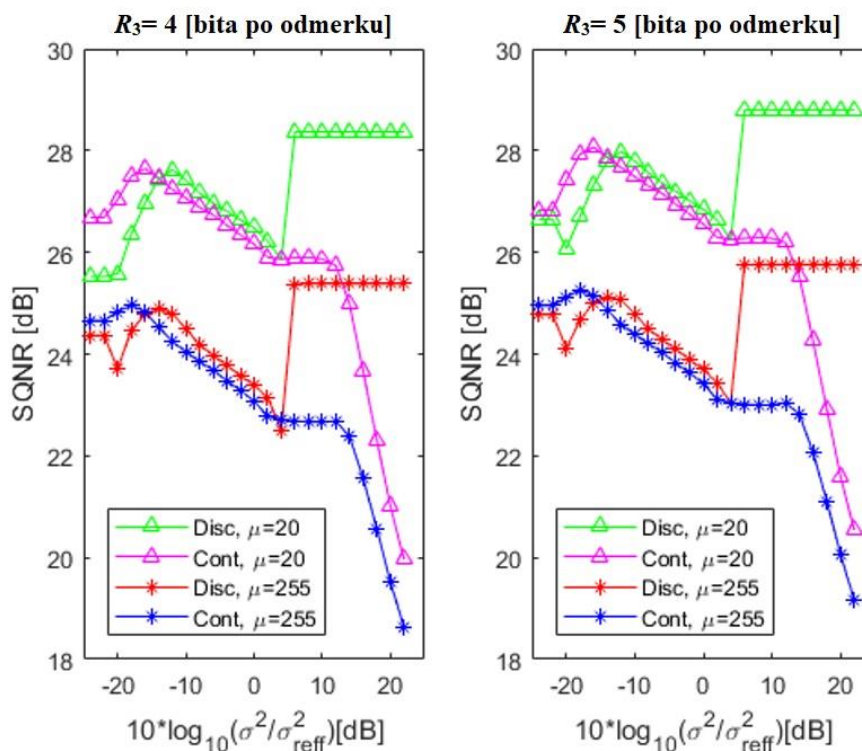
μ	Ulazni signal	SQNR [dB]	R_{sr} [bita po odmerku]		
			4.02083	5.02083	6.02083
20	Kontinualni	$SQNR_{sr}$	20.1395	25.7187	31.4030
	Diskretni	$SQNR_{sr}$	20.9472	27.7113	32.9634
	Kontinualni	$SQNR_{sr(st)}$	20.8485	26.6353	32.3656
	Diskretni	$SQNR_{sr(st)}$	21.6146	28.0653	33.9725
255	Kontinualni	$SQNR_{sr}$	16.9693	23.0142	29.2311
	Diskretni	$SQNR_{sr}$	18.4096	24.8170	30.9516
	Kontinualni	$SQNR_{sr(st)}$	17.4584	23.5475	29.5184
	Diskretni	$SQNR_{sr(st)}$	19.3170	25.3108	31.3613

Tabela 4.3.5 pokazuje rezultate u obliku SQNR-a za slučaj $R_3=5$ bita po odmerku. Poredeći rezultate iz tabela 4.3.5 i 4.3.4 vidi se da povećanjem bitske brzine log-uniformnog kvantizera za 1 bit po odmerku (ukupno 0.00416 bita po odmerku) kvalitet kodovanja poraste za 0.9 dB, u zavisnosti od faktora kompresije i bitske brzine. Takođe se vidi da predložena kodna šema obezbeđuje 1 – 3 dB viši SQNR kada je poznata informacija samo o diskretnom signalu, u odnosu na slučaj kada je dostupna informacija i o kontinualnom.

Na slikama 4.3.7 i 4.3.8 prikazani su numerički rezultati dobijeni korišćenjem jednačine (4.3.13) za razne vrednosti bitskih brzina i faktora kompresije, u slučajevima kodovanja diskretnog i kontinualnog signala.



Slika 4.3.7. SQNR u širokom opsegu ulaznih varijansi za $R_{sr} \approx 4$ bita po odmerku i razne vrednosti μ



Slika 4.3.8. SQNR u širokom opsegu ulaznih varijansi za $R_{sr} \approx 5$ bita po odmerku i razne vrednosti μ

Na slici 4.3.7 prikazan je SQNR u slučajevima vrednosti srednje bitske brzine kodne šeme od $R_{sr} = 4.01667$ bita po odmerku ($R_1=5$, $R_2=3$, $R_3=4$), $R_{sr} = 4.02083$ bita po odmerku ($R_1=5$, $R_2=3$, $R_3=5$), dok slika 4.3.8 prikazuje SQNR za slučajeve srednje bitske brzine $R_{sr} = 5.01667$ bita po odmerku ($R_1=6$, $R_2=4$, $R_3=4$), $R_{sr} = 5.02083$ bita po odmerku ($R_1=6$, $R_2=4$, $R_3=5$). Posmatrajući prikazane grafike sa slika 4.3.7 i 4.3.8 primećuje se da predložena kodna šema, u slučaju kada je dostupna samo informacija o diskretnom signalu, za male varijanse obezbeđuje približno isti SQNR kao i kada je dostupna informacija o kontinualnom signalu. Međutim, za više vrednosti varijanse iz dinamičkog opsega (veće od 5 dB), pruža značajni dobitak u kvalitetu kodovanja koji nije očekivan, ali se javlja usled lošeg odabira vrednosti maksimalne amplitude opterećenja kvantizera Q_0 [75].

5 Zaključak

U ovoj disertaciji razmatrane su performanse logaritamskih kvantizera uključenih u šeme za kodovanje govornog signala. Ciljevi u ovoj disertaciji su nove kodne šeme sa transformacionim kodovanjem. Najkraće rečeno, cilj je poboljšanje kvaliteta rekonstruisanog signala na prijemnoj strani bez povećanja broja bita po odmerku potrebnih za kodovanje ili smanjivanje bitske brzine uz zadržavanje kvaliteta prijemnog signala.

Sumirani najvažniji doprinosi ove disertacije su:

1. Predložena je šema za kodovanje govornog signala, implementirana korišćenjem modifikovane Hadamarove transformacije H_1 i tehnike adaptacije unapred. Ulazni signal doveden je do bafera u kome se deli na frejmove veličine M odmeraka i za koje procenitelj varijanse računa varijansu, odnosno standardnu devijaciju koju šalje na kodovanje log-uniformnom kvantizeru. Signal u frejmu se dekomponuje na dve podsekvence (y_1 i y_2) primenom transformacija koje se koduju pojedinačno upotrebom adaptivnih kvazi-logaritamskih kvantizera AQ_1 i AQ_2 . Ovako kvantovani signali \hat{y}_1 i \hat{y}_2 , kao i standardna devijacija $\hat{\sigma}_x^2$ prenose se na prijemnu stranu gde se vrši rekonstrukcija signala. Eksperimentalno je pokazano da se upotrebom ove kodne šeme postiže 2 – 4 dB dobitka u SQNR-u u odnosu na adaptivnu impulsnu kodnu modulaciju, dok je u odnosu na standard G.711 dobitak značajno veći.
2. Ova disertacija predlaže i kodnu šemu implementiranu primenom dvostruke Hadamarove transformacije H_1 i tehnike adaptivne kvantizacije unapred. Dvostruka primena transformacije podrazumeva da se nad signalima y_1 i y_2 (već primenjena H_1) primenjuje još jednom modifikovana transformacija H_1 , čime se dobijaju podsekvence y_{11} , y_{12} , y_{21} i y_{22} . Dvostrukom primenom Hadamarove transformacije dodatno se preraspoređuju korisne informacije iz ulaznog signala i dobija se veća korelacija odmeraka u nekim sekvencama što omogućava viši kvalitet rekonstrukcije u prijemniku. Korišćeni ulazni test signal isti je kao u prethodnom eksperimentu. Za vrednost faktora kompresije $\mu=20$ predloženo rešenje obezbeđuje 2 – 2.5 dB viši SQNR u odnosu na kodnu šemu sa jednostrukom primenom Hadamarove transformacije H_1 u zavisnosti od bitske brzine i 4 – 4.5 dB u odnosu na PCM.
3. Prilikom izvođenja eksperimenta uočeno je da signal y_{22} ima jako nisku korelaciju i da kodovanjem tog signala nema dobitka u SQNR-u celokupne kodne šeme. Predložena je

kodna šema sa tri signala (podsekvence) koja se koduju zasebno (y_{11} , y_{12} i y_2). Izbacivanjem dodatne transformacije signala y_2 smanjuje se kompleksnost projektovanja kodne šeme, dok kvalitet kodovanja ostaje približno isti kao sa četiri sekvence (četiri grane). Eksperiment je urađen, kao i prethodni, za vrednosti faktora kompresije $\mu=20$ i $\mu=255$. Rezultati pokazuju da predložena kodna šema ostvaruje dobitak 2.1 – 2.2 dB dobitka u SQNR-u u odnosu na šemu sa jednostrukom Hadamarovom transformacijom H_1 .

4. U disertaciji je predložena i šema za kodovanje govornog signala strukturalno slična G.711.1 standardu, mada za razliku od G.711.1 standarda, kompleksnost predložene kodne šeme je znatno niža. Umanjenje kompleksnosti posledica je izbacivanja vektorske kvantizacije iz grane posle transformacije t_2 i primene G.711 koda za kodovanje signala u obe grane. Na ovaj način signal se ne deli u frejmove koji se koduju odjednom (vektorska kvantizacija) već se svaki odmerak ulaznog signala koduje zasebno (skalarna kvantizacija) čime se kašnjenje pri kodovanju smanjuje na samo 0.25 ms (prva dva odmerka ulaznog signala po 0.125 ms) sa 5 ms (u zavisnosti od veličine frejma, 40 odmeraka). S obzirom da je G.711.1 standard za kodovanje širokopojasnih signala, odmerenih na 16 kHz, a predložena kodna šema je strukturalno slična pomenutom standardu, eksperiment je urađen za ulazni test signal odmeren na 16 kHz. Kako je u disertaciji predložen novi način računanja objektivne mere kvaliteta, kvalitet kodovanja predložene šeme dobijen je njegovom primenom. Novi način računanja SQNR-a podrazumeva podelu dinamičkog opsega ulaznog signala na segmente određene širine (u ovom slučaju 2 dB), računanje srednjeg SQNR-a za svaki segment i na kraju prosečnog SQNR-a svih segmenata dinamičkog opsega. Rezultati pokazuju da predložena kodna šema obezbeđuje visok kvalitet kodovanja i da je pogodna za kodovanje govornog signala odmerenog na 16 kHz jer obezbeđuje 2.5 – 5 dB viši SQNR u odnosu na G.711 standard i 20 puta manje kašnjenje pri kodovanju u odnosu na G.711.1 standard, što je približno kodovanju u realnom vremenu.
5. U disertaciji predložena je šema za kodovanje govornog signala, implementirana upotrebom DCT transformacije. Struktura šeme slična je šemama sa Hadamarovom transformacijom (bafer, procenitelj varijanse, log-uniformni kvantizer i adaptivni kvazi-logaritamski kvantizeri u granama). Za razliku od Hadamarove transformacije, diskretna kosinusna transformacija dekomponuje signal na tri podsekvence odjednom koje se koduju pojedinačno. Rezultati dobijeni eksperimentom pokazuju dobitak predložene

- kodne šeme za kodovanje govornog signala od oko 3 dB u odnosu na kodnu šemu sa jednostrukom primenom Hadamarove transformacije H_1 , dok ostvaruje 1 – 2 dB niži SQNR u odnosu na šemu sa dvostrukom Hadamarovom transformacijom H_1 .
6. Predložena je i kodna šema bazirana na DCT transformaciji sa transformacionom matricom veličine 4×4 . Eksperimentom je pokazana pogodnost upotrebe predložene kodne šeme za kodovanje govornog signala jer obezbeđuje 1 – 2 dB viši SQNR u odnosu na šemu sa DCT transformacijom 3×3 i približno isti kvalitet kodovanja kao i šema sa dvostukom Hadamarovom transformacijom H_1 , s tim što se ulazni signal ne dekomponuje 2 puta, već se ulazni signal odjednom deli na 4 podsekvence, te je kompleksnost implementacije predložene šeme dosta niža.
 7. Ova disertacija predlaže i šeme za kodovanje diskretnog govornog signala. Šeme za kodovanje diskretnog signala mogu se predstaviti kao dvofazni sistemi. U prvoj fazi nalazi se kvantizer zadužen za diskretizaciju ulaznog signala, dok se u drugoj fazi nalazi kvantizer ili čitava šema za kodovanje tako diskretnog signala. U obe predložene šeme u prvoj fazi upotrebljen je kvazi-logaritamski kvantizer projektovan za bitsku brzinu $R_0=8$ bita po odmerku i faktor kompresije $\mu=255$. Predložena je šema za kodovanje diskretnog signala gde je u drugoj fazi sistema primenjen kvazi-logaritamski kvantizer za razne bitske brzine. Teorijskim proračunima potvrđen je eksperiment i pokazana pogodnost predložene šeme za kodovanje diskretnog signala. Iz eksperimentalnih rezultata može se zaključiti da ukoliko se govorni signal prvo diskretizuje pa onda koduje kvazi-logaritamskim kvantizerom, dobija 1 – 2 dB dobitka u odnosu na slučaj kodovanja kontinualnog signala, dok se obezbeđuje 0.3 – 0.6 dB viši SQNR u odnosu na šemu iz rada [75], gde je u prvoj fazi upotrebljen uniformni kvantizer a u drugoj optimalni komparator. Viši SQNR kodovanja diskretnog u odnosu na kontinualni signal objašnjava se time što diskretizacijom signal postaje amplitudski ograničen (nema distorzije prekoračenja) te se optimalnim izborom maksimalne amplitude opterećenja kvantizera u drugoj fazi postiže viši SQNR ovom tehnikom kodovanja.
 8. Predložena je i šema za kodovanje diskretnog govornog signala, koja je razlikuje od prethodne šeme po tome što se u drugoj fazi primenjuje transformaciono kodovanje i tehnika adaptacije unapred, što podiže kvalitet kodovanja. Eksperimentalnim rezultatima se pokazuje da je prodložena kodna šema obezbeđuje 0.75 – 2 dB dobitka u SQNR-u u odnosu na kodovanje kontinualnog signala, u zavisnosti od bitske brzine.

Rezultati prikazani u ovoj disertaciji pokazuju da je predložene kodne šeme moguće praktično primeniti u kodovanju govornog signala. Primenom novih i kombinacijom postojećih tehnika, predložene šeme je moguće nadograditi čime se mogu povećati vrednosti u SQNR-u, što može biti buduće istraživanje u ovoj oblasti.

6 Literatura

- [1] **N.S. Jayant, P. Noll.** *Digital Coding of Waveforms*. Prentice-Hall, 1984.
- [2] **Allen Gersho**, “Quantization”, *IEEE Communications Society Magazine*, pp. 20 - 29, September, 1977.
- [3] **Allen Gersho**, “Principles of Quantization”, *IEEE Transactions on Circuits and Systems*, Vol. 25, No. 7, pp. 427 - 436, July, 1978.
- [4] **Andreas Spanias, Ted Painter, Venkatraman Atti**, *Audio Signal Processing and Coding*, Wiley-Interscience, February 9, 2007.
- [5] **A. Gersho, R.M. Gray.** *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, Massachusetts, 1992.
- [6] **D. Hankerson, G.A. Harris, P.D.Jr. Johnson.** *Introduction to information theory and data compression*. CHAPMAN & HALL/CRC, 2nd edition, 004.
- [7] **K. Sayood.** *Introduction to Data Compression*. Elsevier Inc., 3rd edition, 2006.
- [8] **D. Salomon.** *Data Compression – The Complete Reference*, 4th edition, Springer, 2007.
- [9] **R. Gallager.** *Information Theory and Reliable Communications*, John Wiley & Sons, 1968.
- [10] **R. Gallager**, *Principles of Digital Communication*, Cambridge University Press, March 24, 2008.
- [11] **Robert Gray, David Neuhoff**, “Quantization”, *IEEE Transactions on Information Theory*, Vol. 44, No. 6, pp. 2325 - 2383, October, 1998
- [12] **Bhumip Khasnabish**, *Implementing Voice over IP*, John Wiley & Sons, 1st edition, May 23, 2003.
- [13] **O. Hersent, D. Gurle, J. Petit.** *IP Telephony – Packet-based multimedia communication systems*, Addison Wesley, 2000.
- [14] **Jonathan Davidson, James Peters, Manoj Bhatia, Satish Kalidindi, Sudipto Mukherjee**, *Voice over IP Fundamentals*, Cisco Press, 2nd edition, July 27, 2006.

- [15] **Olivier Hersent, Jean-Pierre Petit, David Gurle**, *Beyond VoIP Protocols: Understanding Voice Technology and Networking Techniques for IP Telephony*, John Wiley & Sons, 1st edition, March 11, 2005.
- [16] **Daniel Minoli**, *Voice Over MPLS: Planning and Designing Networks*, McGraw-Hill, 1st edition, May 2, 2002.
- [17] **Floriano De Rango, Mauro Tropea, Peppino Fazio, Salvatore Marano**, "Overview on VoIP: Subjective and Objective Measurement Methods", *IJCSNS International Journal of Computer Science and Network Security*, Vol. 6, No. 1B, pp. 140 - 153, January, 2006.
- [18] **R. Gray**, *Quantization and Data Compression, Lecture Notes*, Stanford University, 2004.
- [19] **C. Chu**, *Speech coding algorithms*, John Wiley & Sons, New Jersey, 2003.
- [20] **Joel Max**, "Quantizing for Minimum Distortion", *IRE Transactions on Information Theory*, Vol. 6, No. 1, pp. 7 - 12, March, 1960.
- [21] **Lawrence Rabiner, Ronald Schafer**, *Introduction to Digital Speech Processing*, November 30, 2007.
- [22] **Tadeusz Wysocki, Bahram Honary, Beata Wysocki**, *Signal Processing for Telecommunications and Multimedia*, Springer, 1st edition, October 1, 2004
- [23] **Lawrence Rabiner, Ronald Schafer**, *Digital Processing of Speech Signals*, Prentice Hall, 1st edition, September 15, 1978.
- [24] **John Proakis, Masoud Salehi**, *Digital Communications*, McGraw-Hill, 5th edition, November 6, 2007.
- [25] **Marina Bosi, Richard Goldberg**, *Introduction to Digital Audio Coding and Standards*, Springer, 1st edition, December 31, 2002.
- [26] **Mark Hasegawa-Johnson, Abeer Alwan**, "Speech Coding: Fundamentals and Applications", *Wiley Encyclopedia of Telecommunications*, pp. 1-20, John Wiley & Sons, 2003
- [27] **G. Lukatela, D. Drajić, G. Petrović, R. Petrović**, *Digitalne telekomunikacije*, Građevinska knjiga, Beograd, 1984.
- [28] **S. Lloyd**, "Least Squares Quantization in PCM", *IEEE Transactions on Information Theory*, Vol. 28, No. 2, pp. 129-136, 1982.
- [29] **Paez, Glisson**, "Minimum Mean-squared-error Quantization in Speech PCM and DPCM Systems", *IEEE Transactions on Communications*, Vol. 20, No. 2, pp. 225 - 230, April, 1972.
- [30] **Sangsin Na**, "On the Support of Fixed-Rate Minimum Mean-Squared Error Scalar Quantizers for a Laplacian Source", *IEEE Transactions on Information Theory*, Vol. 50, No. 5, pp. 937 - 944, May, 2004.
- [31] **Sangsin Na, David Neuhoff**, "On the Support of MSE-Optimal, Fixed-Rate, Scalar Quantizers", *IEEE Transactions on Information Theory*, Vol. 47, No. 7, pp. 2972 - 2982, November, 2001.
- [32] **Fleischer**, "Sufficient Conditions for Achieving Minimum Distortion in a Quantizer", *IEEE International Convention Record*, pp. 104 - 111, 1964.
- [33] **James Bucklew, Neal Gallagher**, "A Note on the Computation of Optimal Minimum Mean-Square Error Quantizers", *IEEE Transactions on Communications*, Vol. 30, No. 1, pp. 298 - 301, January, 1982.
- [34] **Roger Wood**, "On Optimum Quantization", *IEEE Transactions on Information Theory*, Vol. 15, No. 2, pp. 248 - 252, March, 1969.
- [35] **Siegfried Graf, Harald Luschgy**, *Foundations of Quantization for Probability Distributions*, Springer, 1st edition, June 16, 2000.
- [36] **Ahmet Kondoz**, *Digital Speech: Coding for Low Bit Rate Communication Systems*,

- John Wiley & Sons, 2nd edition, October 20, 2004.
- [37] **Oliver Bernard, John Pierce, Claude Shannon**, “The Philosophy of PCM”, *Proceedings of the Institute of Radio Engineers*, Vol. 36, No. 11, pp. 1324 - 1331, November, 1948.
 - [38] **Demetrios Kazakos, Kami Makki**, “Robust Companders”, *Proceedings of the 6th WSEAS International Conference on Telecommunications and Informatics*, pp. 32 - 35, Dallas, Texas, 2007.
 - [39] **Neil Judell, Louis Scarf**, “A Simple Derivation of Lloyd's Classical Result for the Optimum Scalar Quantizer”, *IEEE Transactions on Information Theory*, Vol. 32, No. 2, pp. 326 - 328, March, 1986.
 - [40] **Zoran Perić, Jelena Nikolić**, “An Effective Method for Initialization of Lloyd-Max's Algorithm of Optimal Scalar Quantization for Laplacian Source”, *Informatica*, Vol. 18, No. 2, pp. 279 - 288, I O S Press, 2007.
 - [41] **Zoran Perić, Jelena Nikolić**, “Inicijalizacija Lloyd-Max-ovog algoritma optimalne skalarne kvantizacije za Laplasov izvor”, *Zbornik radova 50. konferencije ETRAN*, Sveska 2, str. 23 - 26, Beograd, 6 - 8 Jun, 2006.
 - [42] **Jelena Nikolić, Zoran Perić**, “Lloyd-Max's Algorithm Implementation in Speech Coding Algorithm Based on Forward Adaptive Technique”, *Informatica*, Vol. 19, No. 2, pp. 255 - 270, I O S Press, 2008.
 - [43] **Fu-Sheng Lu, Gary Wise**, “A Further Investigation of Max's Algorithm for Optimum Quantization”, *IEEE Transactions on Communications*, Vol.33., No.7, pp. 746-750, Jul, 1985.
 - [44] **Ida Mengyi Pu**, *Fundamental Data Compression*, Elsevier, January 11, 2006.
 - [45] ITU-T, Recommendation G.711, *Pulse Code Modulation (PCM) of Voice Frequencies*, International Telecommunication Union, 1972.
 - [46] ITU-T, Recommendation G.711.1, *Wideband Embedded Extension for G.711 Pulse Code Modulation*, International Telecommunication Union, 2008.
 - [47] **Jacob Benesty, Mohan Sondhi, Yiteng Huang**, *Springer Handbook of Speech Processing*, Springer, 1st edition, December 13, 2007.
 - [48] **Jerry Gibson**, *Multimedia Communications: Directions and Innovations*, Academic Press, 1st edition, November 2, 2000.
 - [49] **Jerry Gibson**, “Speech Coding Methods, Standards and Applications”, *IEEE Circuits and Systems Magazine*, Vol. 5, No. 4, pp. 30 - 49, 2005.
 - [50] **Z. Perić, J. Nikolić, D. Pokrajac**. “Estimation of the support region for Laplacian source scalar quantizers”, *Journal of Electrical Engineering*, Vol. 58, No. 1, pp. 47–51, 2007.
 - [51] **Z. Perić, A. Mosić, S. Panić**. “Coding Algorithm based on Loss Compression using Scalar Quantization Switching Technique and Logarithmic Companding”, *Journal of Information Science and Engineering*, Vol. 26, No. 3, pp. 967-976, 2010.
 - [52] **Z. Perić, J. Nikolić, Z. Eškić, S. Krstić, N. Marković**. “Design of Novel Scalar Quantizer Model for Gaussian Source”, *Information Technology and Control*, Vol. 37, No. 4, pp. 321-325, 2008.
 - [53] **Antonio Ortega, Martin Vetterli**, “Adaptive Scalar Quantization Without Side Information”, *IEEE Transactions on Image Processing*, Vol. 6, No. 5, pp. 665 - 676, May, 1997.
 - [54] **David Goodman, Allen Gersho**, “Theory of an Adaptive Quantizer”, *IEEE Transactions on Communications*, Vol. 22, No. 8, pp. 1037 - 1045, August, 1974.
 - [55] **David Salomon**, *Data Compression: The Complete reference*, Springer, 4th edition, December 19, 2006.

- [56] Jozsef Vass, Yunxin Zhao, Xinhua Zhuang, “Adaptive Forward-Backward Quantizer for Low Bit Rate High Quality Speech Coding”, *IEEE Transactions on Speech and Audio Processing*, Vol. 5, No. 6, pp. 552 - 557, November, 1997.
- [57] Z. Perić, M. Dinčić, D. Denić, A. Jocić. “Forward Adaptive Logarithmic Quantizer with New Lossless Coding Method for Laplacian Source”, *Wireless Personal Communications*, Vol. 59, No. 4, pp. 625-641, 2011.
- [58] M. Savić, Z. Perić, M. Dinčić. “Design of Forward Adaptive Uniform Quantizer for Discrete Input Samples for Laplacian Source”, *Electronics and Electrical Engineering*, Vol. 9(105), pp. 73-76, 2010.
- [59] M. Tancic, Z. Peric, S. Tomic, N. Simic, “Speech Signal Coding Using Forward Adaptive Quantization and Simple Transform Coding”, *ELEKTRONIKA IR ELEKTROTEHNIKA*, ISSN 1392-1215, Vol. 22, No. 3, pp. 74-77, 2016, DOI: <http://dx.doi.org/10.5755/j01.eie.22.3.15318>
- [60] R. Clarke. *Transform Coding of Images*, Academic Press, London, 1985.
- [61] Y. Hiwasaki, H. Ohmuro, T. Mori, S. Kurihara, A. Kataoka. “A G.711 Embedded Wideband Speech Coding for VoIP Conferences”, *IEICE Trans. Inf. & Syst.*, Vol. 89, No. 9, pp. 2542, 2006.
- [62] Yusuke Hiwasaki, Shigeaki Sasaki, Hitoshi Ohmuro, Takeshi Mori, Jongmo Seong, Mi Suk Lee, Balázs Kövesi, Stéphane Ragot, Jean-Luc Garcia, Claude Marro, Lei Miao, Jianfeng Xu, Vladimir Malenovsky, Jimmy Lapierre, Roch Lefebvre, “G.711.1: A Wideband Extension to ITU-T G.711”, *Proceedings of the European Signal Processing Conference EUSIPCO '08*, Lausanne, Switzerland, August 25 - 29, 2008.
- [63] Y. Hiwasaki, T. Mori, S. Sasaki, H. Ohmuro, A. Kataoka, „A wideband speech and audio coding candidate for ITU-T G.711 WBE standardization“, *Proc. of the IEEE ICASSP*, 2008, pp. 4017–4020, DOI: 10.1109/ICASSP.2008.4518535
- [64] S. Sasaki, T. Mori, Y. Hiwasaki, H. Ohmuro. “Global Standard for Wideband Speech Coding: ITU-T G.711.1 (G.711 wideband extension)”, *NTT Technical Review*, Vol. 6, No. 8, 2008.
- [65] A. D. Lyon. “The μ -law CODEC”, *Journal of Object Technology*, Vol. 7, No. 8, pp. 17- 31, 2008.
- [66] Z. Peric, D. Aleksic, M. Stefanovic, J. Nikolic, “New approach to support region determination of the μ -law quantizer”, *Elektronika Ir Elektrotehnika*, vol. 19, no. 8, pp. 111–114, 2013, DOI: <http://dx.doi.org/10.5755/j01.eee.19.8.5406>
- [67] G. Petkovic, Z. Peric, L. Stoimenov, “Switched scalar optimal μ -law quantization with adaptation performed on both the variance and the distribution of speech signal”, *Elektronika Ir Elektrotehnika*, Vol. 22, No. 1, pp. 64–67, 2016. DOI: <http://dx.doi.org/10.5755/j01.eee.22.1.14111>
- [68] J. Nikolic, Z. Peric, D. Aleksic, “Optimization of μ -law companding quantizer for Laplacian source using Muller’s method”, *Przegląd Elektrotechniczny*, Vol. 89, No. 3A, pp. 206–208, 2013.
- [69] D. Aleksic, Z. Peric, J. Nikolic, “Support region determination of quasi-logarithmic quantizer for Laplacian source”, *Przegląd Elektrotechniczny*, vol. 88, no. 7A, pp. 130–132, 2012.
- [70] Z. Peric, M. Tancic, S. Tomic, D. Ciric, “Subband coding of audio signal with logarithmic companders”, in *Proc. Telsiks 2015*, Serbia, Nis, 2015, pp. 19–22. DOI: <http://dx.doi.org/10.1109/telsks.2015.7357728>
- [71] M. Tancic, Z. Peric, N. Simic, “Modified Wideband Speech Coding System with Embedded G.711 Coders”, *Revue Roumanie Des Sciences Techniques, serie*

- Electrotechnique et Energetique*, Vol. 62, No. 3, 2017, pp. 299-304. DOI: <http://dx.doi.org/10.5755/j01.eie.22.3.15318>
- [72] **J. Nikolic, Z. Peric, J. Lukic**, “Konstrukcija dvostepenog kvantizera sa ugrađenim G.711 kodekom”, *Zbornik radova 53. konferencije ETRAN*, T.E.3.5.1-4, Vrnjačka Banja, 15 - 18 Jun, 2009.
- [73] **Z. Peric, J. Nikolic, J. Lukic, D. Denic**, “Analysis of Two-stage Quantizer with Embedded G.711 Quantizer and Segmental Uniform Quantizer”, *Electronics & Electrical Engineering*, Vol.19, No.2, pp. 88–91, 2013.
- [74] **M.Tancic, Z. Peric, N. Simic, S. Tomic**, “Performance of quasi-logarithmic quantizer for discrete input signal”, *Journal of Information Technology and Control*, Vol. 46, No. 3, pp. 395-402, 2017, DOI: <http://dx.doi.org/10.5755/j01.itc.46.3.16197>
- [75] **Peric Z., Simic N., Savic M.**, “Analysis and Design of Two Stage Mismatch Quantizer for Laplacian Source”, *Elektronika Ir Elektrotehnika*, Vol. 21 (3), pp. 49-53, 2015, DOI: <http://dx.doi.org/10.5755/j01.eee.21.3.10380>
- [76] **Nikolic J., Peric Z., Jovanovic A.**, “Variance Mismatch Analysis of Unrestricted Polar Quantization for Gaussian Source”, *IEEE Signal Processing Letters*, Vol. 21(5), pp. 540-544, 2014, DOI: <http://dx.doi.org/10.1109/LSP.2014.2309093>
- [77] **Gazor S., Zhang W.**, „Speech probability distribution“, *IEEE Speech Processing Letters*, Vol. 10, No.7, pp. 204-207, 2003, DOI: 10.1109/LSP.2003.813679
- [78] **Lajos Hanzo, Clare Somerville, Jason Woodard**, *Voice and Audio Compression for Wireless Communications*, John Wiley & Sons - IEEE Press, 2nd edition, October 8, 2007.
- [79] **Ian McLoughlin**, *Applied Speech and Audio Processing: With Matlab Examples*, Cambridge University Press, 1st edition, March 23, 2009.
- [80] **Bishnu Atal**, “The History of Linear Prediction”, *IEEE Signal Processing Magazine*, Vol. 23, No. 2, pp. 154 - 161, March, 2006.
- [81] **Wai Chu**, *Speech Coding Algorithms: Foundation and Evolution of Standardized Coders*, John Wiley & Sons, 1st edition, March 7, 2003.
- [82] **James Bucklew, Neal Gallagher**, “A Note on Optimal Quantization”, *IEEE Transactions on Information Theory*, Vol. 25, No. 3, pp. 365 - 366, May, 1979.
- [83] **Peter Moo, David Neuhoff**, “Optimal Compressor Functions for Multidimensional Companding”, *Proceedings of the IEEE International Symposium on Information Theory*, pp. 515, Ulm, Germany, July, 1997.
- [84] **Jelena Nikolić, Zoran Perić, Aleksandar Jocić**, “Speech Coding Using a Novel Forward Adaptive Companding Technique”, *Zbornik radova VII konferencije Digitalna obrada govora i slike DOGS 2008*, str. 102 - 105, Kelebija, 2 - 3 Oktobar, 2008.
- [85] **Randy Goldberg**, *A Practical Handbook of Speech Coders*, CRC Press LLC, 1st edition, April 21, 2000.
- [86] **A. M. Aldajani**. “Logarithmic quantization in the least mean squares algorithm”, *Digital Signal Processing*, Vol. 18, pp. 321-333, 2008.
- [87] **S. Tomic, Z. Perić, M. Tančić, J. Nikolić**, “Backward adaptive and quasi-logarithmic quantizer for sub-band coding of audio”, *Journal of Information Technology and Control*, Vol. 47, No. 1, pp. 131-139, 2018, DOI: <http://dx.doi.org/10.5755/j01.itc.47.1.16190>
- [88] **M. Tancic, Z.Peric, A. Jovanovic, S. Tomic**, “Determining Compression Factor of Quasi-logarithmic Quantizers for Laplacian Source in Narrow Dynamic Variance Range”, *Facta Universitatis Series Automatic Control And Robotics*, Vol. 15, No. 3, pp. 217-226, 2016, DOI: 10.22190/FUACR1603217T
- [89] **E. O. Brigham, R. E. Morrow**, "The fast Fourier transform", *IEEE Spectrum*, vol. 4,

- pp. 63-70, December 1967.
- [90] **H. C. Andrews**, "A high speed algorithm for the computer generation of Fourier transforms", *IEEE Trans. Computers (Short Notes)*, vol. C-17, pp. 373-375, April 1968.
 - [91] **H. C. Andrews, W. K. Pratt**, "Fourier transform coding of images", *Hawaii Internat'l Conf. on System Sciences*, pp. 677-679, 1968-January.
 - [92] **Kunz, H.O.**, "On the Equivalence Between One-Dimensional Discrete Walsh-Hadamard and Multidimensional Discrete Fourier Transforms". *IEEE Transactions on Computers*. Vol.28, No.3, 1979.
 - [93] **Yarlagadda, Hershey**, "Hadamard Matrix Analysis and Synthesis", 1997 (Springer).
 - [94] **B. Fino, V. Algazi**, "Unified Matrix Treatment of the Fast Walsh-Hadamard Transform," in *IEEE Transactions on Computers*, vol. 25, no. , pp. 1142-1146, 1976.
 - [95] **W. K. Pratt, J. Kane, H. C. Andrews**, "Hadamard transform image coding", *Proceedings of the IEEE*, Vol. 57, No.1, 1969.
 - [96] **Song Wang, Guang Deng, Jiankun Hu**, "A Partial Hadamard Transform approach to the design of cancelable fingerprint templates containing binary biometric representations", *Pattern recognition*, Vol. 61, pp. 447-458, 2017, DOI: <https://doi.org/10.1016/j.patcog.2016.08.017>
 - [97] **Robin Scheibler, Saeid Haghhighatshoar, Martin Vetterli**, "A Fast Hadamard Transform for Signals with Sublinear Sparsity in the Transform Domain", *IEEE Transactions on Information Theory*, Vol. 61, No.4, 2015, pp. 2115-2132, DOI: 10.1109/TIT.2015.2404441
 - [98] **D. Jabeen, G. Monir, F. Azim**, "Sequency Domain Signal Processing Using Complex Hadamard Transform", *Circuits, Systems, and Signal Processing*, Vol. 35, No.5, pp. 1783-1793, 2016, DOI: <https://doi.org/10.1007/s00034-015-0138-x>
 - [99] **S. Liu, A. C. Bovik**. "Foveation embedded DCT domain video transcoding", *J. Vis. Commun. Image R.*, Vol. 16, pp. 643-667, 2005.
 - [100] **N. Ahmed, T. Natarajan, K.R. Rao**, "Discrete Cosine Transform", *IEEE Transactions on Computers*, Vol. 23, No.1, 1974, DOI: 10.1109/T-C.1974.223784
 - [101] **Byeong Lee**, "A new algorithm to compute the discrete cosine transform", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol.32, No.6, 1984. DOI: 10.1109/TASSP.1984.1164443.
 - [102] **Wen-Hsiung Chen, C. Smith, S. Fralick**, "A Fast Discrete Cosine Transform algorithm for the Discrete Cosine Transform", *IEEE Transactions on Communications*, Vol. 25, No. 9, 1977, DOI: 10.1109/TCOM.1977.1093941.
 - [103] **Hsieh Hou**, "A fast recursive algorithm for computing the discrete cosine transform", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 25, No. 10, 1987, DOI: 10.1109/TASSP.1987.1165060.
 - [104] **M. Narasimha, A. Peterson**, "On the computation of the discrete cosine transform", *IEEE Transactions on Communications*, Vol. 26, No. 6, 1978, DOI: 10.1109/TCOM.1978.1094144
 - [105] **Ravi K Sheth, V. V. Nath**, "Secured digital image watermarking with discrete cosine transform and discrete wavelet transform method", *International Conference on Advances in Computing, Communication, & Automation (ICACCA) (Spring)*, 2016, DOI: 10.1109/ICACCA.2016.7578861
 - [106] **Tasweer Ahmad, Junaid Rafique, Hassam Muazzam, Tahir Rizvi**, "Using Discrete Cosine Transform Based Features for Human Action Recognition", *Journal of Image and Graphics*, Vol. 3, No. 2, December 2015, DOI: 10.18178/joig.3.2.96-101
 - [107] **Ramesh Kumar Lama, Soekjoo Shin, Moonsoo Kang, Goo-Rak Kwon, Moo-Rak**

- Choi**, “Interpolation using wavelet transform and discrete cosine transform for high resolution display”, *IEEE International Conference on Consumer Electronics (ICCE)*, 2016, DOI: 10.1109/ICCE.2016.7430573
- [108] **Dipalee Gupta, Siddhartha Choubey**, “Discrete Wavelet Transform for Image Processing”, *International Journal of Emerging Technology and Advanced Engineering*, Vol. 4, No.3, 2015
- [109] **Nasrin M. Makbol, Bee Ee Khoo, Taha H. Rassem**, “Block-based discrete wavelet transform – singular value decomposition image watermarking scheme using human visual system characteristics”, *IET Image Processing*, Vol. 10, No.1, 2016, DOI: 10.1049/iet-ipr.2014.0965
- [110] **Shradhananda Beura, Banshidhar Majhi, Ratnakar Dash**, “Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer”, *Neurocomputing*, Vol. 154, pp. 1-14, 2015, DOI: <https://doi.org/10.1016/j.neucom.2014.12.032>
- [111] **S. A. Kumar, K. Basant, D. Mayank, M. Anand**, “Multiple watermarking on medical images using selective discrete wavelet transform coefficients”, *Journal of Medical Imaging and Health Informatics*, Vol. 5, No. 3, pp. 607-614, 2015, DOI: <https://doi.org/10.1166/jmihi.2015.1432>
- [112] **D.P. Radunovic**, “Talasići”, *Akademski Misao*, Beograd, 2005
- [113] **J. E. Garcia, A. Ortega, A. Miguel, E. Lleida**, “Low bit rate compression methods of feature vectors for distributed speech recognition”, *Speech Communication*, vol. 58, pp. 111–123, 2014. DOI: <http://dx.doi.org/10.1016/j.specom.2013.11.007>
- [114] **Hiwasaki Yusuke, Hitoshi Ohmuro**, ITU-T G. 711.1: *Extending G. 711 to higher-quality wideband speech*, *Communications Magazine*, IEEE 47.10 (2009): 110–116
- [115] **Na S.**, “Asymptotic Formulas for Mismatched Fixed Rate Minimum MSE Laplacian Quantizers”, *IEEE Signal Processing Letters*, Vol.15, pp. 13-16, 2008, DOI: <http://dx.doi.org/10.1109/LSP.2007.910240>
- [116] **Na S.**, “Asymptotic Formulas for Variance-Mismatched Fixed-Rate Scalar Quantization of a Gaussian Source”, *IEEE Transactions on Signal Processing*, Vol.59(5), pp.2437-2441, 2011, DOI: <http://dx.doi.org/10.1109/TSP.2011.2112354>
- [117] **Na S., Neuhoﬀ D. L.**, “Asymptotic MSE Distortion of Mismatched Uniform Scalar Quantization”, *IEEE Transactions on Information Theory*, 2012, Vol.58, No.5, 3169-3181. DOI: <http://dx.doi.org/10.1109/TIT.2011.2179843>
- [118] **Wang Y., Li J., Gong Y.**, “Small-footprint high-performance deep neural network-based speech recognition using split-VQ”, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2015 ISBN: 978-1-4673-6997-8, DOI: 10.1109/ICASSP.2015.7178919
- [119] **Zhang Y., Chan W., Jaitly N.**, “Very deep convolutional networks for end-to-end speech recognition”, *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 5-9 March 2017, New Orleans, LA, USA, pp. 4845-4849
- [120] **Polap D.**, “Model of Identity Verification Support System Based on Voice and Image Samples”, *Journal of Universal Computer Science*, 2018, Vol. 24 No.4, pp. 460-474.
- [121] **Polap D.**, “Neuro-heuristic voice recognition”, *Proceedings of the Federated Conference on Computer Science and Information Systems*, 11-14 September 2016, Gdansk, Poland, pp. 487-490
- [122] **Xiao X., Watanabe S., Erdogan H., Lu L., Hershey J., Seltzer M. L., Chen G., Zhang Y., Mandel M., Yu D.**, “Deep beamforming networks for multi-channel speech recognition”, *Proceedings of IEEE International Conference on Acoustics,*

Speech and Signal Processing (ICASSP), 20-25 March 2016, Shanghai, China, pp. 5745–5749

- [123] **Fayek H. M., Lech M., Cavedon L.**, “Evaluating deep learning architectures for Speech Emotion Recognition”, *Neural Networks*, Vol. 92, pp. 60–68, 2017
- [124] **Li L., Wang D., Chen Y., Shi Y., Tang Z., Zheng T. F.**, “Deep Factorization for Speech Signal”, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, Canada, 2018, Electronic ISSN: 2379-190X, DOI: 10.1109/ICASSP.2018.8462169

7 Biografija Autora

Milan Ž. Tančić rođen u Pirotu, Srbija, 22.11.1989. Osnovnu školu “8. Septembar” i “Gimnaziju” završio je u Pirotu. Zvanje diplomirani inženjer elektrotehnike i računarstva dobio je 2013. godine na Elektronskom Fakultetu u Nišu. U novembru 2013. godine upisuje doktorske studije na istom fakultetu na smeru Telekomunikacije pod mentorstvom prof. dr Zorana Perića, čiji je i trenutno student. Od oktobra 2016. do decembra 2018. radi kao programer informacionih sistema u kompaniji Tigar Tyres (*Michelin group*), dok od tada radi u firmi SORS d.o.o čiji je su-osnivač. Njegova oblast istraživanja je obrada signala, sa akcentom na kodovanje govornog signala primenom tehnika adaptacije i transformacionog kodovanja. Do sada je objavio preko 15 naučnih radova, od kojih 5 u časopisima koji se nalaze na SCI listi.

IZJAVA O AUTORSTVU

Izjavljujem da je doktorska disertacija, pod naslovom

Konstrukcija logaritamskih kvantizera za visoko kvalitetno adaptivno transformaciono kodovanje govornog signala

koja je odbranjena na Elektronskom Fakultetu u Nišu:

- rezultat sopstvenog istraživačkog rada;
- da ovu disertaciju, ni u celini, niti u delovima, nisam prijavljivao na drugim fakultetima, niti univerzitetima;
- da nisam povredio autorska prava, niti zloupotrebio intelektualnu svojinu drugih lica.

Dozvoljavam da se objave moji lični podaci, koji su u vezi sa autorstvom i dobijanjem akademskog zvanja doktora nauka, kao što su ime i prezime, godina i mesto rođenja i datum odbrane rada, i to u katalogu Biblioteke, Digitalnom repozitorijumu Univerziteta u Nišu, kao i u publikacijama Univerziteta u Nišu.

U Nišu, _____

Potpis autora disertacije:

dr Milan Ž. Tančić

IZJAVA O ISTOVETNOSTI ŠTAMPANOG I ELEKTRONSKOG OBLIKA
DOKTORSKE DISERTACIJE

Naslov disertacije:

Konstrukcija logaritamskih kvantizera za visoko
kvalitetno adaptivno transformaciono kodovanje
govornog signala

Izjavljujem da je elektronski oblik moje doktorske disertacije, koju sam predao na unošenje u **Digitalni repozitorijum Univerziteta u Nišu**, istovetan štampanom obliku.

U Nišu, _____

Potpis autora disertacije:

dr Milan Ž. Tančić

IZJAVA O KORIŠĆENJU

Ovlašćujem Univerzitetsku biblioteku “Nikola Tesla” da u Digitalni repozitorijum Univerziteta u Nišu unese moju doktorsku disertaciju, pod naslovom:

Konstrukcija logaritamskih kvantizera za visoko kvalitetno adaptivno transformaciono kodovanje govornog signala

Disertaciju sa svim priložima predao sam u elektronskom obliku, pogodnom za trajno arhiviranje.

Moju doktorsku disertaciju, unetu u Digitalni repozitorijum Univerziteta u Nišu, mogu koristiti svi koji poštuju odredbe sadržane u odabranom tipu licence Kreativne zajednice (Creative Commons), za koju sam se odlučio.

1. Autorstvo (CC BY)
2. Autorstvo – nekomercijalno (CC BY-NC)
- 3. Autorstvo – nekomercijalno – bez prerađe (CC BY-NC-ND)**
4. Autorstvo – nekomercijalno – deliti pod istim uslovima (CC BY-NC-SA)
5. Autorstvo – bez prerađe (CC BY-ND)
6. Autorstvo – deliti pod istim uslovima (CC BY-SA)

U Nišu, _____

Potpis autora disertacije:

dr Milan Ž. Tančić