



**UNIVERZITET U NIŠU**  
**ELEKTRONSKI FAKULTET**



**Stefan S. Tomić**

**RAZVOJ ALGORITAMA ZA KODOVANJE  
GOVORNOG I AUDIO SIGNALA**

**DOKTORSKA DISERTACIJA**

**Niš, 2018.**



UNIVERSITY OF NIŠ

FACULTY OF ELECTRONIC ENGINEERING



Stefan S. Tomić

**DEVELOPMENT OF SPEECH AND AUDIO  
SIGNAL CODING ALGORITHMS**

**DOCTORAL DISSERTATION**

**Niš, 2018.**

# ZAHVALNICA

*Želeo bih da zahvalim svima koji su me podržavali u različitim periodima života i rada tokom doktorskih studija.*

*Veliku zahvalnost dugujem svom mentoru prof. Zoranu Periću, koji je bio uz mene od samog početka, nesebično delio svoje znanje i ideje i sa velikom posvećenošću pratio moj rad. Takođe, posebno se zahvaljujem dr Jeleni Nikolić, koja je svojim savetima i sugestijama u velikoj meri uticala na moj rad i doprinela mom razvoju kao naučnog radnika.*

*Zahvaljujem se svojoj supruzi i porodici, koji su verovali u mene i nesebično me podržavali da završim doktorsku disertaciju.*

Autor,

Avgust 2018.

## Podaci o doktorskoj disertaciji

Mentor:

dr. Zoran Perić, redovni profesor, Elektronski fakultet, Univerzitet u Nišu

Naslov:

Razvoj algoritama za kodovanje govornog i audio signala

Rezime:

Govor predstavlja jednu od najznačajnijih osobina ljudske vrste, koja ga izdvaja od ostalih bića. Kao takav, on predstavlja i veoma bitan signal u digitalnim sistemima. Od začetka prenosa govornog signala u komunikacionim sistemima, tehnologija kojom se vršila njegova digitalizacija je zauzimala značajno mesto u digitalnoj obradi signala. Ekspanzija multimedijalnih sistema, koja i dalje traje, dovodi do toga da se govor i audio zapis može prenositi lakše i raznovrsnije nego ikad pre. Nove mogućnosti primene zahtevaju i napredak u tehnikama kodovanja, što čini ovu oblast aktuelnom i veoma važnom, i podstiče istraživače da rade na razvoju novih algoritama za kodovanje govornog i audio signala.

U disertaciji će biti predstavljeni različiti tipovi algoritama za kodovanje govornog i audio signala, zvanani na kodovanju talasnog oblika signala. Biće procenjene prednosti i nedostaci primene različitih tehnika kodovanja, kao i različitih tipova kvantizera. Kvalitet izlaznog signala će biti procenjen objektivnim i subjektivnim merama kvaliteta.

Najpre će biti predstavljena primena tehnika podopsežnog kodovanja u kodovanju govornog i audio signala. Podopsežno kodovanje vrši podelu ulaznog signala na podopsege, koji se mogu kodovati nezavisno i time ostvariti ušteda u korišćenoj bitskoj brzini. Biće predstavljena nekoliko tipa algoritama za podopsežno kodovanje, različiti složenosti i uz primenu različitih tehnika kvantizacije.

U disertaciji će biti predložen i modifikovani BTC (Block Truncation Coding) algoritam i njegova primena u kodovanju govornog i audio signala. BTC algoritam je originalno projektovan za kodovanje crno-bele slike, dok

modifikovani algoritam iskorišćava neke zajedničke osobine ovih tipova signala.

Biće predstavljen i jednostavan algoritam zasnovan na diferencijalnom kodovanju i adaptaciji unazad, , koji će prikazati prednost primene ovih tehnika, kojima se uz nisku kompleksnot izračunavanja mogu ostvariti dobre performanse izlaznog signala.

Pored procene performansi pojedinačnih algoritama, u disertaciji će biti prikazana i analiza korelisanosti rezultata subjektivnih i objektivnih mera kvaliteta kodovanog signala. Pored toga biće predstavljena ograničenja upotrebe pojedinih metoda, kao i mogućnost primena objektivnih mera kvaliteta koje imaju zadatak da simuliraju subjektivnu meru kvaliteta koja uključuje učešće ljudskih ispitanika.

Performanse predloženih algoritama će biti poređene sa performansama algoritama slične kompleksnosti, kako bi se stekao bolji uvid u kvalitet kodovanog signala, kao i u potencijalne mogućnosti primene algoritama koji će biti predloženi u disertaciji.

Naučna  
oblast:  
Naučna  
disciplina:

Elektrotehnika i računarstvo
Telekomunikacije

Ključne  
reči:

algoritam, audio signal, govorni signal, kvantizacija, kodovanje, podopsežno kodovanje, BTC algoritam.
--

UDK:

621.391:004(075.8)(076)
-------------------------

CERIF  
klaisfikacija:

T 180 Telekomunikacioni inženjering
-------------------------------------

Tip licence  
kreativne  
zajednice:

CC BY-NC-ND
-------------

## Data on Doctoral Dissertation

Doctoral  
Supervisor:

dr. Zoran Perić, full professor, Faculty of Electronic Engineering,  
University of Nis

Title:

Development of speech and audio signal coding algorithms

Abstract:

Speech represents one of the most important characteristics of the human species, which sets humans apart from other living beings. Thus, speech represents highly important signal in digital communications. Since the transmission of the speech signal in digital communications began, the technology dedicated to the digitalization of speech took a significant place in digital signal processing. Still lasting growth in the usage of multimedia systems enables us to send and distribute digital speech and audio signals easier than ever. New implementation possibilities demand progress in coding techniques, which inspires researchers to develop new speech and audio signal coding algorithms.

This dissertation will present different speech and audio signal coding algorithms, based on waveform signal coding. Also, the advantages and disabilities of implementing different coding techniques and quantizer types will be evaluated. The output signal quality will be evaluated by subjective and objective quality measures.

Firstly, the implementation of subband coding techniques in speech and audio signal coding will be presented. Subband coding performs decomposing of the input signal into subbands, which can be coded independently, by which we can reduce the required bit rate. Four different subband coding algorithms will be presented, determined by different complexity levels and quantization techniques.

The dissertation will present modified BTC (Block Truncation Coding) algorithm and its implementation in speech and audio signal coding. BTC

algorithm was originally developed for application in black and white image coding, while the modified algorithm exploits some of the common characteristics of black and white images and speech and audio signals. Also, a simple differential and backward adaptive coding algorithm will be presented. This algorithm will show the advantages of implementing aforementioned techniques, by which we can obtain high output signal quality, while preserving low computational complexity.

Along with the performance evaluation of the individual coding algorithms, in the dissertation will be analyzed the correlation between subjective and objective assessment of the coded signal quality. Limitations of individual methods will be denoted as well as the limitations and benefits of their implementation in the scientific research. The performances of the algorithms to be proposed in the dissertation will be compared with the performances of similar complexity coding algorithms, thus gaining a better insight into the encoded signal quality and possible practical application.

Scientific Field:	Electrical engineering
Scientific Discipline:	Telecommunications
Key Words:	adaptation, algorithm, audio signal, BTC, DPCM, speech signal, quantization, coding, subband coding
UDC:	621.391:004(075.8)(076)
CERIF Classification:	T 180 Telecommunication engineering
License Type:	CC BY-NC-ND

# SADRŽAJ

1. UVOD .....	1
2. TEHNIKE KODOVANJA GOVORNOG I AUDIO SIGNALA .....	9
2.1 Uvod .....	9
2.2 Koderi govornog signala .....	10
2.2.1 Koderi talasnog oblika govornog signala .....	11
2.2.2 Parametarski koderi .....	12
2.2.3 Hibridni koderi govornog signala .....	12
2.3 Koderi audio signala .....	14
3. OSNOVE TEHNIKA KVANTIZACIJE .....	17
3.1 Osnovni pojmovi .....	17
3.2 Tipovi kvantizera .....	20
3.3 Kompanding tehnika .....	21
3.3.1 Optimalna kompresorska funkcija .....	21
3.3.2 Logaritamski A zakon kompresije .....	22
3.3.3 Logaritamski $\mu$ zakon kompresije .....	23
4. PODOPSEŽNO KODOVANJE GOVORNOG I AUDIO SIGNALA .....	24
4.1 Osnove podopsežnog kodovanja .....	25
4.2 Algoritam za podopsežno kodovanje govornog i audio signala sa dva podopsega .....	28
4.2.1 Projektovanje i razvoj algoritama .....	28
4.2.2 Analiza eksperimentalnih rezultata .....	29
4.3 Adaptivni algoritam za podopsežno kodovanje govornog i audio signala sa dva podopsega .....	31
4.3.1 Tehnike adaptacije .....	32
4.3.2 Projektovanje algoritma .....	33
4.3.3 Analiza eksperimentalnih rezultata .....	35
4.4 Algoritam za podopsežno kodovanje audio signala sa tri podopsega i adaptacijom unazad .....	41
4.4.1 Projektovanje algoritma i primena podopsežnog kodovanja .....	41
4.4.2 Projektovanje kvantizera i određivanje parametra kodovanja .....	43



4.4.3	Analiza eksperimentalnih rezultata .....	45
4.5	Podopsežno kodovanje audio signala sa četiri podopsega .....	48
4.5.1	Projektovanje algoritma .....	48
4.5.2	Konstrukcija kvantizera .....	50
4.5.3	Analiza eksperimentalnih rezultata .....	51
4.6	Poređenje rezultata predloženih algoritama za podopsežno kodovanje govornog i audio signala .....	53
5.	KODOVANJE GOVORNOG I AUDIO SIGNALA PRIMENOM MODIFIKOVANOG BTC ALGORITMA I DIFERENCIJALNOG KODOVANJA .....	54
5.1	Originalni BTC algoritam.....	55
5.2	Modifikovani BCT algoritam za kodovanje audio signala.....	56
5.2.1	Projektovanje modifikovanog BTC algoritma za kodovanje audio signala.....	57
5.2.2	Projektovanje kvantizera.....	60
5.2.3	Analiza eksperimentalnih rezultata .....	63
5.3	Modifikovani BTC algoritam za kodovanje širokopojasnog govornog signala visokog kvaliteta .....	71
5.3.1	Projektovanje modifikovanog BTC algoritma za kodovanje širokopojasnog govornog signala .....	71
5.3.2	Analiza eksperimentalnih rezultata .....	74
5.4	Algoritam za jednostavno diferencijalno kodovanje govornog signala zasnovano na adaptaciji unazad.....	76
5.4.1	Tok algoritma.....	79
5.4.2	Analiza numeričkih rezultata .....	81
6.	ANALIZA KORELISANOSTI REZULTATA SUBJEKTIVNIH I OBJEKTIVNIH MERA KVALITETA KODOVANOG SIGNALA .....	85
6.1	Objektivne metode procene kvaliteta kodovanog signala .....	85
6.2	Subjektivna mera kvaliteta kodovanog signala .....	91
6.2.1	MOS- Mean Opinion Score .....	91
6.2.2	Tok eksperimenta.....	92
6.2.3	Rezultati eksperimenta subjektivne procene kvaliteta kodovanog signala .....	93
7.	ZAKLJUČAK .....	101

8.	LITERATURA.....	104
9.	SPISAK REFERENCI AUTORA.....	112

# 1. UVOD

Sposobnost govora je jedna od glavnih karakteristika ljudske vrste, koja omogućava neposredan prenos informacija između ljudi. Čovjek je od davnina težio da informacije sačuva u obliku zapisa kako bi ostale nadolazećim generacijama. Začetkom digitalnog doba, velika energija je usmerena ka digitalizaciji i prenosu govora na daljinu putem telekomunikacione infrastrukture. Prenos, a kasnije i skladištenje govornog i audio signala u izvornom obliku, zahtevaju veliku količinu memorijskog prostora i propusnog opsega sistema kojim se vrši prenos. To je inspirisalo istraživače da veliku pažnju posvete razvoju algoritama za kompresiju, kojim će se smanjiti upotreba resursa a održati zahtevani kvalitet signala.

Moderno doba i ekspanzija multimedijalnih tehnologija doprineli su tome da se telekomunikacionom infrastrukturom trenutno prenosi ogromna količina audio i video signala svakodnevno. Iako smo svedoci da širina propusnog opsega telekomunikacionih sistema raste, veliki porast ostvarenog saobraćaja i dalje čini primenu efikasnih algoritama za kompresiju signala neophodnim i izuzetno važnim. Rasprostranjenost upotrebe mobilnih telefona i razvoj mobilnih aplikacija doprineli su tome da se multimedijalni sadržaj može deliti između korisnika na veoma veliki broj načina. Prisutna je i česta upotreba servisa slanja glasovnih poruka putem mobilnih aplikacija, uz standardne visoko kvalitetne video i audio pozivi koji su prisutni već duže vreme. Uspešnost primene ovih tehnologija i servisa je u velikoj meri baziran na brzini izvršavanja i lagodnosti korišćenja istih. Svedoci smo da korisnik telekomunikacionih servisa u XXI veku očekuje od svake usluge da se izvrši trenutno, a da pritom kvalitet istog servisa bude na visokom nivou. U takvoj situaciji postoji stalna potreba za unapređenjem postojećih i razvojem novih algoritama za kodovanje audio i govornog signala, kako bi se ostvario još bolji kvalitet komprimovanog signala uz primenu što manje bitske brzine. Stalni porast upotrebe digitalnih servisa i širok spektar potencijalne primene projektovanih algoritama čini ovu oblast aktuelnom i privlači veliki broj autora da se njome bavi. To je inspirisalo i razvoj algoritama koji će biti predstavljeni u ovoj doktorskoj disertaciji.

Kako telekomunikacioni sistemi ne prepoznaju govor i audio signale u izvornom obliku, potrebno je izvršiti njihovu digitalizaciju. Digitalizacija se sastoji iz tri koraka: odmeravanje, kvantovanje i kodovanje [1] - [3]. Svaki od koraka je izuzetno važan i ima veliki uticaj na kvalitet digitalizovanog signala. U disertaciji je posebna pažnja posvećena postupku kvantizacije, koji je veoma bitan u ostvarivanju kompresije izlaznog signala. Kompresija signala podrazumeva postizanje što boljeg kvaliteta izlaznog signala uz korišćenje što manjeg broja bita po odmerku. Disertacija se bavi pravilnim projektovanjem kvantizera i njihovom primenom u algoritmima za kodovanje govornog i audio signala, sa ciljem ostvarenja što višeg nivoa kompresije signala.

Kvantizer se može definisati kao redna veza kodera i dekodera [3]. Uloga kvantizera je da izvrši preslikavanje amplituda ulaznog signala u grupu dozvoljenih amplituda. Na taj način se vrši kompresija signala koji se predstavlja značajno manjim brojem reprezentacionih nivoa u odnosu na realan ulazni signal u kojem, teorijski, svaki odmerak može da ima različitu vrednost amplitude. Tipovi kvantizera se mogu definisati po više kriterijuma, a dva osnovna tipa kvantizera su skalarni i vektorski. Osnovna razlika između ova dva tipa kvantizera je u tome što se kod skalarnih kvantizera kvantuje svaki odmerak pojedinačno, dok se kod vektorskog vrši kvantizacija više odmeraka istovremeno, koji su grupisani u vektore [1] - [3]. U disertaciji se razmatra projektovanje i implementacija skalarnih kvantizera u algoritmima za kodovanje govornog i audio signala.

U postupku kvantizacije se praktično vrši zaokruživanje vrednosti amplitude ulaznog signala na dozvoljenu vrednost amplitude određenu reprezentacionim nivoima kvantizera. Kom reprezentacionom nivou (kvantu) će trenutni odmerak biti dodeljen zavisi od granica odlučivanja, koje se nalaze između 2 reprezentaciona nivoa. Kako broj reprezentacionih nivoa direktno zavisi od korišćene bitske brzine, veća bitska brzina povećava rezoluciju kvantizera i samim tim će razlika između originalne i kvantovane vrednosti odmerka biti manja. Takođe, prilikom projektovanja kvantizera, bitan parametar predstavlja i amplituda maksimalnog opeterećenja, koja određuje granice opsega u kojem će kvantizacioni intervali biti raspoređeni [3].

Disertacija se bavi primenom uniformnih i neuniformnih skalarnih kvantizera u projektovanju algoritama za kodovanje govornog i audio signala. Pored toga, bitna podela kvantizera je i na adaptivne i neadaptivne, u zavisnosti od toga da li se vrši prilagođenje

reprezentacionih nivoa i granica odlučivanja na statističke parametre ulaznog signala. Adaptacija se može implementirati kao adaptacija unapred, koja se vrši na osnovu odmeraka ulaznog signala koji treba da se kvantuju i adaptacija unazad, koja se vrši na osnovu već kvantovanih odmeraka ulaznog signala [3]. Kako svaki od navedenih tipova kvantizera ima svoje prednosti i nedostatke, u predloženim algoritmima se razmatra implementacija različitih tipova kvantizera, u zavisnosti od toga koji je najpogodniji za određenu namenu.

Algoritmi za kodovanje govornog i audio signala se generalno mogu podeliti na one koji vrše kodovanje bez gubitaka i one koji primenjuju tehnike kodovanja sa gubicima [1], [4], [5]. Kodovanje bez gubitaka podrazumeva da je moguće izvršiti idealnu rekonstrukciju kodovanog signala i povratiti izvorni oblik, dok kod kodovanja sa gubicima nije moguće idealno rekonstruisati originalni ulazni signal. Algoritmi predloženi i razmatrani u disertaciji su zasnovani na kodovanju sa gubicima, uz primenu kodovanja talasnog oblika signala. Dalje, kodovanje se može podeliti na kodovanje talasnog oblika signala, parametarsko kodovanje i hibridno, koje predstavlja kombinaciju prvih dveju tehnika [1], [3], [4]. Kod kodovanje talasnog oblika signala, talasni oblik kodovanog je približan talasnom obliku originalnog signala [1], [3], [4], [6]. Ovaj tip kodera je široko zastupljen u različitim standardima, među kojima je prvi i verovatno najrasprostranjeniji PCM (*Pulse Code Modulation*) G.711 standard [7]. Ovaj standard predstavlja osnovu telefonskih sistema, a prisutan je i danas. Širokopoljaska verzija ovog standarda je usvojena 2008. godine i poznata je pod nazivom G.711.1. Pogodna je za kodovanje širokopoljaskog govornog i audio signala, sa frekvencijom odmeravanja od 16 kHz [8]. To čini G.711.1 standard pogodnim za primenu u VoIP (*Voice over IP*) protokolu, široko zastupljenim u internet komunikaciji.

U kodovanju govornog i audio signala, izuzetno je bitna i preporuka međunarodne telekomunikacione unije (ITU) G.712, koja definiše donji prag kvaliteta kodovanog govornog signala visokog kvaliteta izraženog odnosom signal šum kvantizacije [9]. Poređenjem prerformansi novih algoritama za kodovanje sa ovim standardom se može utvrditi da li je zadovoljen minimalni prag potreban za visok kvalitet kvantovanog signala.

U procesu kvantizacije se zaokruživanjem ulaznog odmerka na vrednost reprezentacionog nivoa unosi nepovratna greška, koja se može sumirati u veličinu nazvanu distorzija signala [1], [3]. Kvalitet kodovanog signala se može proceniti subjektivnim i objektivnim merama kvaliteta. Najrasprostranjenija subjektivna mera kvaliteta, korišćena i u izradi disertacije je MOS (*Mean Opinion Score*). U ovoj metodi se kao konačna mera kvaliteta koristi srednja vrednost rezultata dobijenih od različitih ispitanika [4], [6]. Postoji više objektivnih metoda procene kvaliteta kodovanog signala, ali se benefiti pravilnog projektovanja i implementacije odgovarajućeg kvantizera najlakše uočava analiziranjem odnosa signal-šum kvantizacije, odnosno SQNR-a (*Signal to Quantization Noise Ratio*). SQNR je definisan kao odnos snage ulaznog signala i šuma kvantizacije [1] - [3]. Pored SQNR-a u disertaciji je korišćen i PESQ (*Perceptual Evaluation of Speech Quality*), koji predstavlja objektivnu metodu procene kvaliteta sa zadatkom da modeluje subjektivnu metodu procene kvaliteta za čiju primenu je potrebno više vremena i resursa [10]. Uz korišćenje samih metoda za procenu performansi predloženih algoritama, razmatra se i dobitak ostvaren poređenjem dobijenih rezultata sa rezultatima primene drugih metoda slične kompleksnosti.

Kako je oblast kodovanja govornog i audio signala izuzetno široka, postoji veliki broj algoritama koji se bave istim. Algoritmi često imaju komercijalnu primenu, bilo da je to standard koji se primenjuje u telekomunikacionim sistemima širom sveta, poput PCM-a i G.711 standarda, ili u definisanju formata za prenos i skladištenje audio signala, poput AAC (*Advanced Audio Coding*), MP3 (*MPEG-2 Audio Layer III*) ili Dolby digital standarda koji su široko rasprostranjeni i koriste se svakodnevno od strane velikog broja korisnika [11]. Pored komercijalne primene samog algoritma u definisanju standarda i formata digitalnih fajlova, algoritmi se često primenjuju kao delovi većih sistema sa dosta kompleksnijom primenom. Time je doprinos disertacije značajniji, jer je oblast istraživanja aktuelna, a potencijalna oblast primene predloženih algoritama izuzetno široka.

Cilj disertacije je razvoj novih algoritama za kodovanje govornog i audio signala. U disertaciji je predloženo nekoliko različitih tipova algoritama i razmatrana njihova primena u kodovanju govornih i audio signala. Uz projektovanje samih algoritama akcenat je stavljan na postupku projektovanja kvantizera koji se implementira u predloženi algoritam.

Jedan od algoritama predloženih u disertaciji je algoritam zasnovan na principima podopsežnog kodovanja, sa nekoliko različitih nivoa složenosti i različitim tehnikama kvantizacije. Ovaj algoritam koristi činjenicu da se kod govornog i audio signala različita količina informacija nalazi u različitim frekvencijskim podopsezima. Algoritam predlaže razlaganje signala na podopsege primenom idealnih digitalnih filtara i njihovu pojedinačnu obradu. Primenjuju se idealni digitalni FIR (*Finite Impulse Response*) niskopropusni (NF) i visokopropusni (VF) filtri. U disertaciji su projektovani su algoritmi koji vrše razlaganje ulaznog signala na dva i tri podopsega, sa ciljem postizanja optimalne kompleksnosti i zadovoljavajuće kompresije signala [12] - [14]. Prednost podele signala na podopsege je ta što se signali podopsega mogu kvantovati posebno i iskoristiti njihove karakteristike kako bi se primenom različitih tehnika kodovanja i bitskih brzina ostvarila bolja kompresija signala. Podopsezima koji sadrže manju količinu korisne informacije se mogu dodeliti niže bitske brzine i time smanjiti ukupna bitska brzina potrebna za kodovanje ulaznog signala. Doprinos takođe leži i u projektovanju kvantizera odgovarajućih za primenu u podopsežnom kodovanju, kao i u predlogu optimalne bit-alokacije i poređenju teorijski i eksperimentalno dobijene raspodele bitske brzine.

U disertaciji je takođe razmatrana primena algoritma za kodovanje crno-bele slike u kodovanju govornog i audio signala. Zapravo, izvršena je modifikacija BTC (*Block Truncation Coding*) algoritma i predstavljena njegova primena u kodovanju govornog i audio signala. Kako se prirode audio signala i crno-bele slike u velikoj meri razlikuju, bilo je potrebno primeniti različite tehnike kodovanja i konstruisati nove kvantizere, koji odgovaraju prirodi signala na koji se algoritam primenjuje. Na taj način je dobijen novi algoritam, koji zadržava osnovne principe BTC algoritma uz njegovo prilagođenje prirodi govornog i audio signala. Govorni i audio signali poseduju slične karakteristike, pogotovo u slučaju širokopojasnog govornog signala, sa frekvencijom odmeravanja od 16 kHz. Oba tipa signala se odlikuju postojanjem korelacije između susednih odmeraka, dok je audio signal okarakterisan nešto višom amplitudskom dinamikom. Pored toga, audio signala se najčešće koduje primenom više bitske brzine, usled čestog zahteva za reprodukcijom visokog kvaliteta.

Originalni BTC algoritam je zasnovan na podeli ulaznog signala na blokove podataka, odnosno frejmove, dok se kodovanje vrši frejm po frejm [15]. Pored frejmova odmeraka ulaznih signala, koriste se i statistički parametri koji opisuju frejm, kao što je standardna

devijacija i srednja vrednost frejma. Originalni algoritam koristi postojanje korelacije između odmeraka crno-bele slike kako bi se ostvarila ušteda u korišćenoj bitskoj brzini. Principe BTC algoritma je moguće primeniti na govorni i audio signal, jer kod njih takođe postoji korelacija između susednih odmeraka. Predloženi algoritam zadržava samo osnovne principe BTC algoritma, dok su modifikacije primenjene u odabiru tehnika kodovanja, odabiru bitske brzine i veličini frejma. Pored toga, u disertaciji je predložen BTC algoritam sa primenom diferencijalnog kodovanja, kao i prikaz prednosti primene adaptivnih tehnika kodovanja u projektovanju modifikovanog BTC algoritma [16], [17].

Diferencijalno kodovanje predstavlja jednostavnu, ali veoma efikasnu kodnu tehniku u kojoj se umesto originalnog signala vrši kodovanje signala razlike (diferencije). Ova tehnika je prvi put primenjena u DPCM (*Differential Pulse Code Modulation*) koderu, koji nadgrađuje PCM koder [4], [18] - [20]. Signal razlike je dobijen oduzimanjem predviđene vrednosti odmerka od njegove stvarne vrednosti. Ovaj postupak predstavlja i primenu predikcionog kodovanja, kojim se primenom jednostavnih transformacija smanjuje dinamički opseg singla, i dobija ulazni signal pogodniji za primenu različitih tehnika kodovanja. U disertaciji je prikazana primena diferencijalnog kodovanja u predloženim algoritmima, čime se dobija nova, efikasnija verzija algoritma i postiže bolja kompresija signala, dok se kompleksnost i vreme izvršenja algoritma neznatno povećavaju.

Najvažniji doprinos disertacije su predloženi algoritmi za kodovanje govornog i audio signala, koji su proizvod primene više različitih tehnika, a između ostalog i zasnovani na kvantizerima projektovanim i predloženim u disertaciji. Projektovani su kvantizeri različitih tipova, uz primenu tehnika adaptacije i diferencijalnog kodovanja. Pored toga, pruža se uvid u performanse predloženih algoritama na osnovu analize eksperimentalnih rezultata. Analiza eksperimentalno dobijenih rezultata primene predloženih algoritama pokazuje nivo kvaliteta koji se primenom predloženih algoritama ostvaruje, kao i dobitak u poređenju sa standardizovanim koderima slične kompleksnosti.

Doprinos disertacije je i uporedna analiza rezultata i zaključak o postojanju korelacije između rezultata ostvarenih primenom različitih subjektivnih i objektivnih metoda procene kvaliteta kodovanog signala.



Disertacija obuhvata teorijsku postavku i opis razmatranih algoritama, kao i eksperimente koji ukazuju na mogućnosti praktične primene predloženih algoritama u kodovanju govornog i audio signala. Oblast istraživanja je aktuelna i obuhvata više različitih tehnika čije se pojedinačne prednosti i nedostaci sisematično obrađuju za svaku od tehnika pojedinačno.

Disertacija je organizovana u pet poglavlja, podeljenih u sekcije. Nakon prvog, uvodnog poglavlja, u drugom poglavlju je dat pregled široko rasprostranjenih tehnika i algoritama korišćenih u kodovanju govornog i audio signala, kao i klasifikacija kodera i tehnika kodovanja.

U trećem poglavlju su predstavljene osnove tehnika kvantizacije, klasifikacija kvantizera, kao i bliži opis tehnika kvantizacije korišćenih u izradi disertacije.

U četvrtom poglavlju su predloženi algoritmi za kodovanje govornog i audio signala zasnovani na podopsežnom kodovanju. Razmatrani su algoritmi različitih nivoa složenosti, u pogledu broja podopsega i primenjenih tehnika kvantizacije. Projektovani su i primenjeni novi adaptivni i neadaptivni kvazilogaritamski kvantizeri, zasnovani na  $\mu$  zakonu kompresije. Izvedena je i optimalna bit-alokacija za predložene algoritama, odnosno raspodela raspoložive bitske brzine po različitim frekvencijskim podopsezima. Izvršena je analiza pojedinačnih frekvencijskih podopsega i izvedena formula za procenu varijansi istih, koja se koristi u postupku adaptivne kvantizacije. Predloženi algoritmi su primenjeni u kodovanju realnih govornih i audio signala, sa ciljem procene performansi algoritama.

U petom poglavlju je predstavljen i opisan algoritam za kodovanje govornog i audio signala zasnovan na BTC algoritmu, originalno korišćenom u kodovanje crno-bele slike. Prikazane su primenjene modifikacije i razlike između originalnog i predloženog algoritma, kao i njihova zajednička svojstva. Kako originalni BTC algoritam koristi jednostavnu uniformnu kvantizaciju sa dva nivoa, projektovani su i primenjeni složeniji kvantizeri, pogodni za primenu u kodovanju govornog i audio signala. Takođe, primenjene su i tehnike diferencijalnog i adaptivnog kodovanja, kako bi se poboljšao kvalitet kodovanog signala. Razmotrena je i primena različitih dužina frejma, kako bi statistički parametri celog frejma na najbolji način predstavljali svaki odmerak tog frejma. Uvedena je i nadgradnja BTC algoritma podelom frejmova na podfrejmove različitih

dužina, kako bi se selektivnom procenom statističkih parametara za glavni frejm i podfrejm ostvarila bolja kompresija ulaznog signala.

U šestom poglavlju je predstavljena analiza performansi predloženih algoritama primenom različitih subjektivnih i objektivnih metoda procene kvaliteta kodovanog signala. Data je veza između rezultata subjektivnih testova slušanja i objektivnih mera kvaliteta dobijenih na osnovu egzaktnih analitičkih formula, korišćenjem softverskog paketa Matlab. Takođe, u analizi je pored predloženih algoritama razmatran i PCM koder, sa ciljem poređenja rezultata dobijenih različitim metodama procene kvaliteta i isticanja očekivanih dobitaka u kvalitetu primenom predloženih algoritama za kodovanje.

Sedmo poglavlje predstavlja zaključak gde je sumiran doprinos disertacije, kao i najvažniji ostvareni rezultati. Dati su i mogući dalji pravci istraživanja, a na kraju i spisak korišćene literature.

## 2. TEHNIKE KODOVANJA GOVORNOG I AUDIO SIGNALA

### 2.1 Uvod

Važnost govornog i audio signala i rasprostranjenost korišćenja njihovih digitalnih reprezentacija dovela je do razvoja velikog broja tehnika za njihovo kodovanje i obradu u digitalnom obliku. Iako su počeci obrade i šireg korišćenja govornog i audio signala bili različiti, svedoci smo da se sa razvojem multimedije sve češće mogu primeniti slični sistemi za obradu signala različitih karakteristika.

Govorni signal prenosi veliku količinu podataka svakodnevno i kao takav predstavlja veoma važan signal u telekomunikacijama, a obrada govornog signala bitnu oblast istraživanja. Osnovne razlike u parametrima korišćenim prilikom obrade govornog i audio signala su frekvencija odmeravanja i bitska brzina, koji se definišu u zavisnosti od primene. Prva najrasprostranjenija primena digitalnog govornog signala bila je u fiksnoj telefoniji. Osnovna frekvencija odmeravanja govornog signala kod fiksne telefonije je 8 kHz [1] - [4], [21], [22], što prema Nikvistovoj teoremi znači da je digitalni signal koji se prenosi ograničen na opseg do 4 kHz [23]. Time je sačuvana razumljivost govora koji se najčešće nalazi u opsegu od 300 Hz do 3.4 kHz, a sva informacija koja se nalazi van ovog opsega će biti izobličena ili izostavljena. Koder koji je korišćen u fiksnoj telefoniji je standardizovan od strane ITU kao G.711 i baziran je na PCM koderu talasnog oblika signala [7]. Interesantno je to da se PCM koristi i kod prvog široko rasprostranjenog digitalnog audio formata, kompakt diska (CD – Compact Disc), s tim što je frekvencija odmeravanja kod kompakt diska 44.1 kHz ili 48 kHz [24]. Iako je PCM sistem pružao visok nivo kvaliteta, njegov nedostatak je korišćenje visoke bitske brzine koja je ograničavala upotrebu. Ova činjenica je motivisala veliki broj istraživača da se bave razvojem algoritama za kompresiju signala, koji bi omogućio lakše i jeftinije (sa stanovišta resursa) korišćenje, prenos i skladištenje digitalnih signala. Vremenom je definisan veliki broj standarda, sa specifičnim oblastima primene i implementiranim tehnikama kodovanja. U nastavku će biti navedeni i kratko opisani najznačajniji koderi govornog i audio signala, sa najvećom oblasti primene.

## 2.2 Koderi govornog signala

Dok se sa stanovišta nauke posmatraju koderi i dekoderi, sa stanovišta komercijalne primene potrebno je ove tehnike definisati kao standarde, kako bi se omogućila primena istih na brojnim uređajima širom sveta. Standardizacijom se bave velike grupe naučnika objedinjene pod okriljem velike međunarodne organizacije, kao što je slučaj sa ITU, a to mogu biti i veliki kompanije koje razvijaju standarde za primenu u svojim proizvodima i uslugama. U nastavku je prikazana Tabela sa nekim od primena kodera govornog signala, koja sadrži dodatne parametre poput bitske brzine, naziva standarda, primenjenog algoritama za kodovanje, kao i godine kada je standard definisan [25].

Tabela 2.2.1. Deo standarda za kodovanje govornog signala i njihova primena

Primena	Bitska brzina	Standard	Algoritam kodovanja	Godina
Fiksna telefonija	64	G.711	PCM ( $\mu$ ili A zakon)	1988
	16-40	G.726	ADPCM	1990
	16-40	G.727	ADPCM	1990
Tele-konferencija	48-64	G.722	ADPCM (split band)	1988
	16	G.728	CELP	1992
	13	Full-rate	RPE-LTP	1992
Mobilna telefonija	12.2	EFR	ACELP	1997
	6.5	Half-rate	VSELP	1995
	8	G.729	ACELP	1996
	4.75-12.2	AMR	ACELP	1998
Multimedija	5.3-6.3	G.723.1	MPLPC, CELP	1996
	2-18.2	MPEG-4	HVXC, CELP	1998
Satelitska telefonija	4.15	M	IMBE	1991
	3.6	Mini-M	AMBE	1995

Koderi govornog signala se generalno mogu podeliti u tri grupe i to na: kodere talasnog oblika signala, parametarske kodere i hibridne kodere. U disertaciji se razmatra projektovanje i primena kodera talasnog oblika signala, pa je stoga ova vrsta kodera detaljnije opisana u nastavku. Pred toga, navedeni su i najznačajniji predstavnici hibridnih kodera govornog signala, koji imaju široku praktičnu upotrebu u mobilnim komunikacijama.

### 2.2.1 Koderi talasnog oblika govornog signala

Prvi najrasprostranjeniji koder talasnog oblika govornog signala je PCM, definisan G.711 standardom, čija je primena počela 1988. godine. Ovaj koder je osnovni standard u digitalnim telefonskim servisima, a često se koristi i prilikom digitalizacije govornog signala kod novijih koder. G.711 standard definiše primenu logaritamskih  $\mu$  i  $A$  zakona kompresije, prilikom čega se  $A$  zakon koristi u Evropi, dok se  $\mu$  zakon koristi u Japanu i Severnoj Americi. Najčešće korišćene standardizovane vrednosti faktora kompresije su  $A = 87.56$  i  $\mu = 255$ . Kod oba zakona se vrši kodovanje sa po 8 bita po odmerku, odnosno 64 kbit/s [3], [4], [7]. Iako PCM predstavlja standard koje je star više decenija, on i dalje ima veliku praktičnu primenu, a sistemi za kodovanje govornog signala zasnovani na PCM-u predstavljaju aktuelnu oblast istraživanja [26], [27].

PCM je nadgrađen primenom tehnike diferencijalnog kodovanja, što je za rezultat dalo diferencijalni PCM, tj. DPCM. Razlika leži u tome što se kod DPCM-a ne vrši direktno kodovanje odmeraka, već se koduje signal razlike, drugačije nazvan signal greške predikcije. DPCM se oslanja na činjenicu da se kod govornog (ali i audio signala uopšte) signala susedni odmerci često ne razlikuju u velikoj meri, odnosno da postoji prisustvo korelacije između odmeraka. Time se vrši predikcija vrednosti sledećeg odmerka na osnovu prethodnog, a razlika između predviđene i stvarne vrednosti odmerka se stoga naziva signal greške predikcije. Signal razlike je okarakterisan nižom varijansom u odnosu na sam ulazni signal i na taj način može se ostvariti viši kvalitet kodovanog signala u odnosu na PCM, za istu bitsku brzinu [4], [18].

Kako DPCM primenjuje fiksni kvantizer i prediktor, poboljšanje je bilo moguće ostvariti primenom tehnika adaptacije. Kao rezultat je nastao adaptivni DPCM, ADPCM, prvi put standardizovan kao G.721, sa brzinom od 32 kbit/s [28]. Adaptaciju je moguće primeniti na kvantizer, prediktor ili na oba elementa i to kao adaptaciju unapred i adaptacija unazad. Adaptacija unapred se zasniva na korišćenju odmeraka ulaznog signala koji tek treba da se kvantuju, dok se adaptacija unazad zasniva na već kvantovanim odmercima signala. Prednost adaptacije unapred je to što pruža nešto veći SQNR u odnosu na adaptaciju unazad, od oko 1 dB [3], ali zahteva prenos dodatne informacije o ulaznim odmercima, što povećava bitsku brzinu, ali i kompleksnost realizacije. Sa druge strane, adaptacija unazad ne zahteva prenos dodatne informacije, ali

se njenom primenom prenosi i greška kvantizacije, koja je neminovna i utiče na nešto niži kvalitet kodovanog signala.

Nakon toga je uvedena serija standarda zasnovanih na ADPCM-u, koji podržavaju rad na različitim brzinama, kao što su standardi G.726 i G.727, koji podržavaju rad na brzinama od 16 kbit/s , 24 kbit/s , 32 kbit/s i 40 kbit/s. G.727 je namenjen primeni u paketskom prenosu govora [29], [30]. Definisana je i koder namenjen kodovanju širokopojasnog govornog signala, koji je standardizovan kao ITU-T standard G.722, sa brzinama od 48 kbit/s , 56 kbit/s i 64 kbit/s [31]. Takođe, definisana su još dva koder za širokopojasno kodovanje govornog signala, standardizovana kao G.722.1 i G.722.2. Koder definisan ITU-T G.722.1 standardom je zasnovan na transformacionom kodovanju i projektovan je za rad na brzinama od 24 kbit/s i 32 kbit/s [32], dok je G.722.2 predstavljen 2003. godine i namenjen je primeni u multimedijalnim komunikacijama [33].

### **2.2.2 Parametarski koderi**

Parametarski koderi vrše rekonstrukciju govornog signala na osnovu određenih parametara ulaznog signala, za razliku od koder talasnog oblika koji koriste konkretne odmerke signala. Talasni oblik signala kodovanog parametarskim koderom ne prati talasni oblik originalnog signala i ovaj tip koder se često koristi na nižim bitskim brzinama, kada je akcenat na uštedi memorijskog prostora. Kako prijemnik rekonstruiše signal na osnovu kvantovanih vrednosti parametara, taj signal se naziva sintetizovani govor. Kvalitet sintetizovanog govora se uglavnom određuje subjektivnim metodama, kako SQNR često nije korelisan sa stvarnim kvalitetom sintetizovanog govora. Najčešće korišćene grupe parametarskih koder su LP (*Linear Prediction*) koderi i Harmonijski koderi.

### **2.2.3 Hibridni koderi govornog signala**

Hibridni koderi praktično kombinuju tehnike koder talasnog oblika signala i parametarskih koder. Oni se oslanjaju na model sinteze govora, kao i parametarski koderi, dok optimizovanje parametra vrši sa ciljem postizanja sličnosti u talasnom obliku signala, kao kod koder talasnog oblika. Hibridni koderi su bazirani na tehnici nazvanoj analiza nakon sinteze (Analysis by Synthesis – Abs). U ovoj tehnici se ne vrši direktno

kodovanje parametara, već se iz njih vrši sinteza jednog segmenta signala, koji se poredi sa odgovarajućim segmentom originalnog signala. Poređenjem se utvrđuje sličnost između ova dva segmenta i parametri kodova se podešavaju sve dok se ne ostvari unapred definisana i prihvatljiva sličnost. Tako dobijeni parametri se porede sa pre-definisanim parametrima sadržanim u kodnoj knjizi, kako se ne bi vršio prenos samog parametra, već rednog broja u kodnoj knjizi onog parametra sa kojim je najbliži prilikom poređenja.

Ova tehnika je često korišćena u projektovanju koda za rad na nižim bitskim brzinama, što je odgovaralo razvoju mobilne telefonije i paketskog prenosa putem interneta, čiji je propusni opseg inicijalno bio veoma ograničen. To je slučaj sa CELP (*Code Excited Linear Prediction*) koderom, čija je prva generacija (1986–1992) projektovana za rad na brzinama od 5.3 kbit/s do 16 kbit/s [3], [6], [31]. Nedostatak ovog koda je bila visoka kompleksnost u pretraživanju kodne knjige, što je uvodilo određeno kašnjenje. Pored toga, CELP je bio osnova za mnoge algoritme i standarde za kodovanje govornog signala na niskim bitskim brzinama, korišćenim i u mobilnoj telefoniji druge generacije. To je slučaj sa Japanskim GSM standardom na 8 kbit/s [35], half rate GSM-om, na 5.6 kbit/s [36] i ostalim standardima zasnovanim na VSELP algoritmu, koji koristi visoko strukturane kodne knjige kako bi smanjio kompleksnost izračunavanja i popravio otpornost sistema na greške.

Druga generacija CELP koda na čelu sa Algebarskim CELP-om (ACELP) ima izuzetno veliku primenu u mobilnim i paketskim komunikacijama. Tu se posebno ističe AMR (*Adaptive Multirate*) kodek (koder i dekodek), sa svojom uskopojasnom (AMR-NB) i širokopojasnom verzijom (AMR-WB), primenjene u mobilnim komunikacijama. Uskopojasni AMR radni na brzinama od 4.75, 5.15, 5.9, 6.7, 7.4, 7.95, 10.2, i 12.2 kbits/s, [37], dok je AMR-WB je projektovan za rad na brzinama od 6.6, 8.85, 12.65, 14.25, 15.85, 18.25, 19.85, 23.05, and 23.85 kbits/s [38]. Odlika AMR koda je da bitska brzina u različitim frejmovima varira u zavisnosti od raspoloživog propusnog opsega u mreži. Oba standarda imaju široku primenu u mobilnim komunikacijama treće i četvrte generacije, omogućavajući glasovne pozive visokog kvaliteta.

Nakon AMR koda, usledio je razvoj VMR (*Variable Multirate*) kodeka, kod koga je bitska brzina promenjiva u zavisnosti od raspoloživog propusnog opsega, ali i od karakteristika ulaznog signala [39]. Kako se generacije mobilne telefonije smenjuju, to

podrazumeva i razvoj novih koderera koji će odgovarati infrastrukturi sistema. Tako je na osnovama ACELP i AMR-WB koderera nastao EVS (*Enhanced Voice Services*) koder, projektovan za četvrtu generaciju mobilne telefonije (VoLTE) [39]. VMR omogućava kodovanje govornog signala u punom opsegu, na brzinama od 16.4 kbit/s.

Gore navedeni su samo neki od mnogobrojnih koderera govornog signala, koji se ističu rasprostranjenošću svoje primene ili uticajem na razvoj oblasti uopšte. Sa stalnim razvojem sistema za prenos podataka javlja se i potreba za novim algoritmima za kodovanje signala, što čini ovu oblast istraživanja aktuelnom.

### **2.3 Koderi audio signala**

Upotreba digitalnog audio signala je započeta pojavom kompaktnog diska, koji je doneo prednosti korišćenja digitalnih signala u domen audija. Kompaktni disk pruža značajno bolji kvalitet od svojih prethodnika i imao je veoma široku upotrebu, koja u poslednjoj deceniji polako jenjava. Razlog za to je između ostalog taj što je „cena“ kvaliteta bila korišćenje visokih bitskih brzina za kodovanje audio signala. Drugi razlog sve manje upotrebe kompaktnog diska i sličnih medijuma za skladištenje audio signala je i sve rasprostranjenija upotreba internet servisa koji digitalne audio podatke skladište na svojim serverima (*cloud computing*) dok korisnik samo privremeno preuzima sadržaj na svoj uređaj (*deezer, spotify, youtube*). Korišćenje ovakvih servisa otvara prostor za razvoj algoritama za kodovanje audio signala, koji će ostvariti bolju kompresiju signala i omogućiti brže preuzimanje sadržaja sa servera, a istovremeno se zahteva visok kvalitet signala.

Kompaktni disk je zasnovan na PCM koderu, koji koristi 16 bita po odmerku, dok je frekvencija odmeravanja 44.1 kHz ili 48 kHz [24], [34]. Iako je kvalitet bio na zadovoljavajućem nivou i komercijalna upotreba je bila izuzetno velika, ograničenje primenjene tehnike kodovanja je da jedan kompaktni disk može da uskladišti samo 70 minuta digitalnog audio signala. Pored toga, zajednička upotreba digitalnog video i audio signala je motivisala dalji razvoj i primenu algoritama za kompresiju. Tako je oformljena i MPEG (*Moving Picture Experts Group*) koja je imala za cilj definisanje standarda koji bi omogućio nesmetanu zajedničku upotrebu digitalnog video i audio signala. MPEG grupa je definisala set međunarodnih standarda za kodovanje audio i video signala, koji



su promenili način distribucije istih. Prvi veći rezultat grupe bio je MPEG-1 standard, koji je imao cilj da zadrži subjektivni kvalitet audio signala CD-a, uz primenu algoritama za kompresiju. MPEG-1 koder je omogućio korišćenje audio signala visokog kvaliteta u digitalnom emitovanju i u različitim internet aplikacijama [11], [41]. Definisan je kao standard za sinhrono kodovanje audio i video signala na ukupnom bitskim brzinom od 1.5 Mb/s. Naslednik ovog standarda, MPEG-2 je bio projektovan za rad na brzinom od 10 Mb/s, sa velikom primenom u digitalnoj satelitskoj i kablovskoj televiziji, DVD videu, itd. Format koji se ističe je MPEG-2 AAC (*Advanced Audio Codec*), široko primenjeni audio standard.

MPEG grupa je predstavila još jedan standard sa još širim spektrom primene od svojih prethodnika, a to je MPEG-4. Delovi ovog standarda su u masovnoj upotrebi i danas, pre svega MPEG-4 AAC, kao i MPEG-4 odeljak 10, odnosno H.264 standard. MPEG-4 AAC je aktuelan i široko rasprostranjeni audio standard, naslednik izuzetno uticajnog i široko korišćenog MP3 (*MPEG Layer 3*) standarda. MP3 je bio nezaobilazan standard za skladištenje i prenos audia putem interneta, da bi ga sa razvojem „*cloud*“ servisa sve više zamenjivao AAC standard. Ovaj standard se koristi kao osnova audio uređaja kao što je *Apple Ipod* i njegove verzije internet audio prodavnice *Itunes*, kao i kod najpopularnijeg internet video servisa *Youtube*. Sa druge strane, H.264 je izuzetno uticajan i široko primenjen u domenu video kodovanja, korišćen kod digitalne televizije visokog kvaliteta, kao i za skladištenje i distribuciju video datoteka putem interneta. U disertaciji nije data detaljnija analiza tehnika kodovanja primenjenih u MPEG standardima kako one nisu obuhvaćene i korišćene u algoritmima predloženih u disertaciji. Navedeni MPEG standardi spadaju u kodere sa gubicima, a pored njih se ističu i *Dolby* standardi, *Dolby AC-2* i *AC-3*. *Dolby AC-3* je standard široko korišćen u bioskopskoj video distribuciji, kao i kod DVD i *Blu-ray* diskova [34].

Pored kodovanja sa gubicima, kodovanje audio signala može biti i bez gubitaka (*lossless*), kod koga je moguće izvršiti idealnu rekonstrukciju kodovanog signala i povratiti njegov prvobitni oblik. Komercijalna primena audio signala kodovanih bez gubitaka je pojava medijuma za skladištenje audio signala visokog kvaliteta kao što je DVD-audio i super-audio CD (SACD), koji su pružali veći prostor za skladištenje [34]. Pored toga, ovaj format je pogodan i u procesu snimanja audio signala u muzičkoj industriji i sličnim profesionalnim primenama. Ovako kodovan signal nije pogodan za

korišćenje u sistemima koji rade u realnom vremenu ili za prenos datoteka putem interneta, ali se može koristiti u sistemima bogatim prostorom za skladištenje.

Algoritmi za kodovanje audio signala bez gubitaka vrše tačnu kompresiju svakog bita ulaznog signala eliminišući statističke zavisnosti signala i redundantnost u talasnom obliku. Mogu se podeliti na dve velike grupe: algoritmi koji primenjuju prediktivno kodovanje i one koji primenjuju transformaciono kodovanje signala. Kod prve grupe algoritama, na ulazni signal se primenjuje linearna predikcija, koja daje signal ostatka na koji se primenjuje entropijsko kodovanje. Teži se da signal ostatka bude što manji, odnosno da predikcija bude što bliža originalnom ulaznom signalu. U drugom pristupu, audio signali se najpre konvertuju u frekvencijski domen putem celobrojnog mapiranja i transformacija, dok su spektralni podaci entropijski kodovani [34]. Kako kodovanje bez gubitaka zahteva veći memorijski prostor za skladištenje, ovi standardi nisu doživeli masovnu upotrebu, kao što je slučaj sa nekim od standarda koji primenjuju kodovanje sa gubicima. Neki od rasprostranjenijih standarda za kodovanje audio signala bez gubitaka je MPEG-4 ALS, *Apple Lossless*, *WMA Lossless*, *FLAC*, itd. *FLAC* standard predstavlja standard otvorenog koda (*open source*), besplatan je za upotrebu i naišao je na široku primenu u skladištenju audio signala. Pored toga, zastupljen je i MPEG-4 ALS, iza kojeg stoji MPEG grupa, koja je definisala veliki broj standarda za kodovanje govornog i audio signala. MPEG-4 ALS primenjuje linearno prediktivno kodovanje sa adaptacijom unapred, sa ciljem postizanja proširenja mogućnosti primene standarda na više različitih uređaja [42], [43]. Najveća oblast primene kodovanja bez gubitaka je u studijskoj upotrebi, prilikom snimanja audio signala, kao i skladištenja signala gde se zahteva najviši mogući kvalitet sa mogućnošću rekonstrukcije signala u prvobitni oblik.

### 3. OSNOVE TEHNIKA KVANTIZACIJE

Signali u prirodi mogu imati proizvoljnu vrednost amplituda, koje se međusobno mogu razlikovati za takođe proizvoljno malu vrednost. Sistem koji bi beležio i razdvajao beskonačno male razlike u amplitudama signala morao bi da ima beskonačno veliku rezoluciju, odnosno raspoloživu bitsku brzinu, a digitalne verzije fajlova bi zahtevale previše memorijskog prostora za skladištenje. Zbog toga je potreban proces koji vrši preslikavanje skupa beskonačno različitih amplituda u skup unapred definisanih vrednosti, koje su bliske originalnoj vrednosti amplitude odmerka. Ovaj proces se naziva kvantizacija, i nezaobilazan je korak prilikom digitalizacije analognih signala iz prirode [3].

#### 3.1 Osnovni pojmovi

Prilikom projektovanja kvantizera ulazni signal je potrebno modelovati određenom funkcijom gustine verovatnoće, koja odgovara tipu signala kome je kvantizer namenjen. U slučaju govornog i audio signala najčešće se koriste Laplasova i Gausova funkcija gustine verovatnoće. Drugim rečima, prilikom projektovanja kvantizera, govorni i audio signal najčešće modelujemo Laplasovim ili Gausovim izvorom. Laplasov izvor bez memorije, sa srednjom vrednošću jednakoj nuli i varijansom  $\sigma^2$  definisan je kao [1], [3], [4]:

$$p(x) = \frac{1}{\sqrt{2\sigma^2}} \exp\left\{-\frac{|x|\sqrt{2}}{\sigma}\right\}, \quad (3.1)$$

dok je Gausov izvor sa srednjom vrednošću jednakoj nuli i varijansom  $\sigma^2$  definisan kao:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{x^2}{2\sigma^2}\right\}. \quad (3.2)$$

Izrazi (3.1) i (3.2) su dati za signal sa nultom srednjom vrednošću jer se takvi signali najčešće sreću u praksi. Kod takvih signala varijansa  $\sigma^2$  predstavlja i snagu signala. U disertaciji se razmatraju oba modela ulaznog signala sa nultom srednjom vrednošću. Ovim postupkom se ne gubi na opštosti, jer se za slučaj srednje vrednosti različite od nule

funkcija gustine verovatnoće zadržava isti oblik, uz pomeranje po  $x$  osi za srednju vrednost signala.

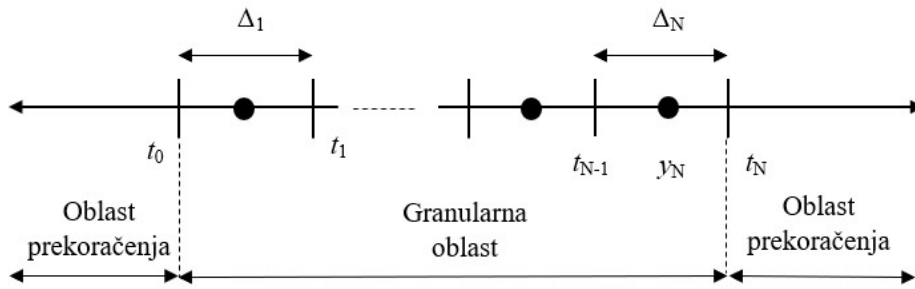
Pored statističkog modela izvora ulaznog signala, osnovni parametri i gradivni elementi kvantizera su:

- Bitska brzina,  $r$ ;
- Amplituda maksimalnog opterećenja,  $t_{\max}$ ;
- Granice odlučivanja;
- Kompresorska funkcija.

Kvantizer  $Q$ , sa  $N$  reprezentacionih nivoa ( $N=2^r$ ) je definisan mapiranjem  $Q: R \rightarrow Y$ , gde  $R$  predstavlja skup realnih brojeva, dok je  $Y \equiv (y_1, y_2, y_3, \dots, y_N) \subset R$ , skup reprezentacionih nivoa koji čine kodnu knjigu veličine  $|Y| = N$  [1]. Skalarni kvantizer vrši podelu skupa realnih brojeva u  $N$  ćelija,  $R_i = (t_{i-1}, t_i]$ ,  $i = 1, \dots, N$ , gde  $t_i$ ,  $i = 0, 1, \dots, N$  predstavljaju granice odlučivanja kvantizera, definisanog kao as  $Q(x) = y_i$ ,  $x \in R_i$ . U postupku kvantizacije, realna osa se deli na kvantizacione intervale (ćelije širine  $\Delta_i$ ), koji mogu biti uniformno ili neuniformno raspodeljeni, odnosno biti istih ili različitih dužina. Ova osobina svrstava kvantizer u grupu uniformnih ili neuniformnih. Kvantizacioni intervale su definisani pragovima odluke, u odnosu na koje je moguće odrediti kojem intervalu pripada odmerak ulaznog signala koji treba kvantovati. Nakon što se odredi kojem intervalu odmerak ulaznog signala pripada, on se zaokružuje na vrednost reprezentacionog nivoa, koji se nalazi unutar tog intervala.

Dozvoljene vrednosti amplitude signala se nalaze unutar amplitudskog opsega kvantizera, koja je definisana amplitudom maksimalnog opterećenja. U slučaju simetričnih kvantizera, koji se razmatraju u disertaciji, amplitudski opseg kvantizera je definisan opsegom  $[-t_{\max}, t_{\max}]$ , prilikom čega važi da je  $t_{\max} = t_N = -t_0$ .

Ovako definisan opseg predstavlja granularnu oblast kvantizera. Kako odmerak može imati proizvoljnu vrednost amplitude, ona može biti veća od maksimalne predviđene vrednosti, odnosno amplitude maksimalnog opterećenja  $t_{\max}$  ili manja od minimalne predviđene vrednosti  $-t_{\max}$ . U ovom slučaju trenutna vrednost odmerka pripada oblasti prekoračenja, prilikom čega se vrši odsecanje amplitude ulaznog signala i zaokruživanje na vrednost definisanu granicama granularne oblasti.



Slika 3.1. Granularna oblast i oblast prekoračenja skalarnog kvantizera

Primenom kvantizacije se unosi nepovratna greška prilikom zaokruživanja odmerka ulaznog signala na vrednost reprezentacionog nivoa. Ova greška se sumira i prikazuje kroz veličinu koja se naziva distorzija. Kako bismo odredili distorziju potrebno je naći razliku između kvantovane i originalne vrednosti ulaznog odmerka. Ova razlika se određuje za svaki odmerak ulaznog signala i sumira u srednjekvadratnu grešku kvantizacije kao:

$$D = \sum_{i=1}^N \int_{x_{i-1}}^{x_i} (x - y_i)^2 p(x) dx. \quad (3.1)$$

Ukupna distorzija se računa kao zbir granularne i distorzije prekoračenja,  $D = D_g + D_p$ . Granularna distorzija se definiše kao srednjekvadratna greška kvantizacije u granularnoj oblasti kvantizera:

$$D_g = \sum_{i=2}^{N-1} \int_{x_{i-1}}^{x_i} (x - y_i)^2 p(x) dx. \quad (3.2)$$

Distorzija prekoračenja obuhvata dva neograničena segmenta koja čine oblast prekoračenja i definiše se kao:

$$D_p = 2 \int_{x_{N-1}}^{\infty} (x - y_N)^2 p(x) dx \quad (3.3)$$

Neizbežno postojanje greške kvantizacije nameće potrebu za određenom veličinom kojom bi se ta greška kvantitativno izrazila. U postupku kvantizacije se najčešće koristi SQNR, koji se izračunava uz pomoć distozije i smage ulaznog signala kao:

$$\text{SQNR}[\text{dB}] = 10 \log_{10} \left( \frac{\sigma^2}{D} \right). \quad (3.4)$$

SQNR je veličina koja se izražava u decibelima [dB] i praktično predstavlja logaritamski odnos snage ulaznog signala i snage šuma kvantizacije. Nakon što se odredi SQNR performanse projektovanog kvantizera postaju uporedive sa ostalim, kako bi se procenio potencijalni dobitak u objektivnom kvalitetu kodovanog signala.

### 3.2 Tipovi kvantizera

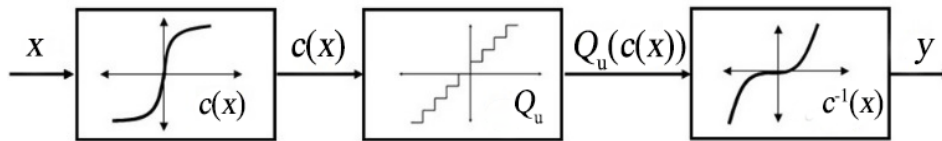
Osnovna podela kvantizera je na skalarne i vektorske [3]. Skalarni kvantizeri vrše kvantovanje odmerak po odmerak, dok se kod vektorskih vrši istovremeno kvantovanje više grupisanih odmeraka. Disertacija se bavi projektovanjem i implementacijom skalarnih kvantizera, pa će o njima biti više reči u nastavku.

Skalarni kvantizeri se dele na uniformne i neuniformne [3]. Ova podela se odnosi na raspored kvantizacionih intervala, odnosno njihovu širinu. Kod uniformnih kvantizera, svi reprezentacioni nivoi i pragovi odlučivanja su na jednakim udaljenostima i kvantizacioni intervali su istih širina. Kod neuniformnih kvantizera ovo nije slučaj i kvantizacioni intervali su različitih širina, koje se određuju na osnovu pravila uspostavljenih u postupku projektovanja kvantizera. Takođe, reprezentacioni nivoi se ne moraju nalaziti na sredini kvantizacionog intervala. U praktičnoj primeni, kod govornog signala visokog kvaliteta i audio signala se uglavnom primenjuju neuniformni kvantizeri, koji bolje odgovaraju prirodi signala. Neuniformni kvantizeri se mogu projektovati primenom Lojd-Maksovog algoritma i kompanding tehnike.

Lojd-Maksov algoritam je iterativni algoritam, čijom se primenom vrši minimizacija distorzije za pretpostavljenu funkciju gustine verovatnoće ulaznog signala. Kako bi se dobio optimalni algoritam, odnosno izvršila minimizacija distorzije, moraju se zadovoljiti uslov centroida i uslov najbližeg suseda [1]- [3], [49]. Kako se Lojd-Maksov algoritam najčešće koristi za projektovanje kvantizera sa manjim brojem nivoa ( $N \leq 128$ ) u disertaciji se ne razmatra njegova primena u kodovanju govornog i audio signala, pa ovaj algoritam neće biti detaljnije opisan.

### 3.3 Kompanding tehnika

U disertaciji se razmatra projektovanje neuniformnih kvantizera primenom kompadding tehnike. Kompanding tehnikom se može projektovati optimalni komparator ili logaritamski kompadding kvantizer, koji se može realizovati primenom  $A$  ili  $\mu$  zakona kompresije. Ova tehnika predlaže realizaciju neuniformnog kvantizera kao rednu vezu kompresora, uniformnog kvantizera i ekspandora.



Slika 3.3.1. Model kompaddinga

Na ulazni signal se najpre primenjuje kompresorska funkcija  $c(x)$ . Kompresorska funkcija može biti zasnovana na  $A$  ili  $\mu$  zakonu kompresije, dok se u slučaju optimalnog kompaddinga primenjuje optimalna kompresorska funkcija. Tako kompresovani signal se dovodi na ulaz uniformnog kvantizera  $Q_u$ . Kvantovani signal se dalje vodi na ulaz ekspandora koji primenjuje inverznu kompresorsku funkciju  $c^{-1}(x)$  i daje izlazni signal  $y$ .

Granularna distorzija optimalnog kompaddinga se može analitički izraziti preko Benetovog integrala [1]- [3], [50] :

$$D_g^B = \frac{x_{\max}^2}{3N^2} \int_{-x_{\max}}^{x_{\max}} \frac{p(x)}{(c'(x))^2} dx \cdot \quad (3.3.1)$$

Ako pretpostavimo da amplituda maksimalnog opterećenja ima dovoljno veliku vrednost i da se distorzija prekoračenja može zanemariti, moguće je izvesti opšti izraz za SQNR kompadding kvantizera, koji je definisan kao:

$$\text{SQNR}[\text{dB}] = 10 \log_{10} \left( \frac{1}{D} \right) \cdot \quad (3.3.2)$$

#### 3.3.1 Optimalna kompresorska funkcija

Optimalna kompresorska funkcija ima za cilj da ostvari maksimalni SQNR za referentnu varijansu ulaznog signala. Izvođenje optimalne kompresorske funkcije je moguće izvršiti na više načina, a najčešće korišćeni i opšte prihvaćeni metod je

minimizacija Benetovog integrala. Ovo je prvi uradio Gersho, koji je pokazao da je nagib optimalne kompresorske funkcije proporcijalan kubnom korenu funkcije gustine verovatnoće [51] :

$$c^{opt}(x) = C_1 (p(x))^{\frac{1}{3}}, \quad (3.3.1.1)$$

gde  $C_1$  predstavlja konstantu dobijenu izračunavanjem Benetovog integrala za poznatu vrednost kompresorske funkcije i granice koje odgovaraju amplitudskom opsegu komandora. Izvođenje optimalne kompresorske funkcije je poznato i prikazano u literaturi, pa stoga neće biti prikazano u disertaciji. Konačan izraz za optimalnu kompresorsku funkciju je definisan kao [3], [6]:

$$c^{opt}(x) = x_{\max} \frac{\int_0^x (p(x))^{\frac{1}{3}} dx}{\int_0^{x_{\max}} (p(x))^{\frac{1}{3}} dx}, \quad x \geq 0. \quad (3.3.1.2)$$

Za vrednosti  $x < 0$  optimalna kompresorska funkcija je definisana kao:

$$c^{opt}(x) = -c^{opt}(-x). \quad (3.3.1.3)$$

### 3.3.2 Logaritamski $A$ zakon kompresije

Kao što je već rečeno, komandring kvantizer može biti zasnovan i na logaritamskim zakonima kompresije,  $A$  i  $\mu$  zakonu. Primenom logaritamskih zakona kompresije projektuju se robusniji kvantizeri, kojima je zadatak da ostvare što konstantniji SQNR u širem opsegu varijansi ulaznog signala. Kako ova osobina pogoduje primeni u kodovanju audio signala i govornog signala visokog kvaliteta, kvantizeri zasnovani na logaritamskim zakonima kompresije se primenjuju u algoritmima za kodovanje koji će biti predstavljeni u disertaciji.

Komandor koji implementira  $A$  zakon predstavlja jednu praktičnu primenu logaritamskih komandora.  $A$  zakon kompresije je definisan kompresorskom funkcijom,



koja se sastoji iz dva dela: linearnog, koji se nalazi u okolini nule i logaritamskog, definisanog logaritamskom kompresorskom funkcijom [3], [6], [52]:

$$c_A(x) = \begin{cases} \frac{A|x|}{1 + \ln A} \operatorname{sgn}(x), & 0 \leq \frac{|x|}{x_{\max}} \leq \frac{1}{A} \\ \frac{x_{\max}}{1 + \ln A} \left( 1 + \ln \frac{A|x|}{x_{\max}} \right) \operatorname{sgn}(x), & \frac{1}{A} \leq \frac{|x|}{x_{\max}} \leq 1 \end{cases}, \quad (3.3.2.1)$$

gde  $A$  predstavlja parametar kompresije koji se definiše prilikom projektovanja kvantizera, dok  $\operatorname{sgn}(x)$  predstavlja znak trenutne vrednosti ulaznog signala. Tipična vrednost faktora  $A$  iznosi 87.56 i ona se koristi u Evropskom PCM sistemu za prenos govora [3], [6].

SQNR za  $A$  zakon se može dobiti zamenom konstante  $K = \frac{1 + \ln A}{x_{\max}}$  u opšti izraz za

SQNR kompadding kvantizera (3.3.2) i on se može odrediti kao [3], [6]:

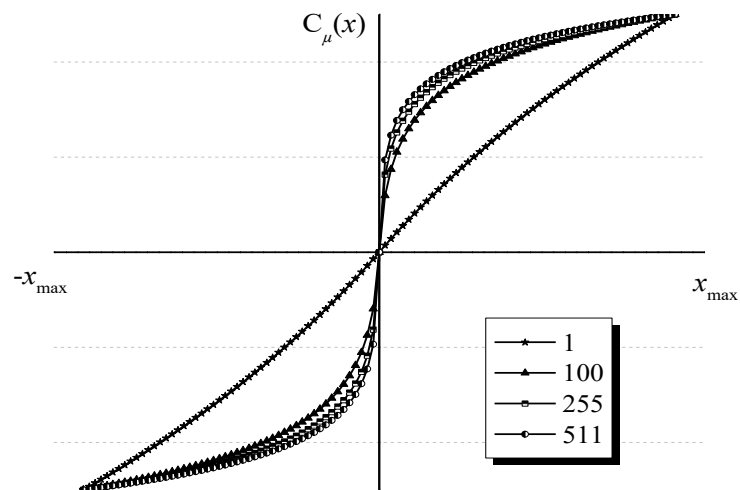
$$\text{SQNR}^A = 6.02R + 4.77 - 20 \log_{10} (1 + \ln A), \quad (3.3.2.2)$$

### 3.3.3 Logaritamski $\mu$ zakon kompresije

Logaritamski komparator zasnovan na  $\mu$  zakonu kompresije predstavlja još jednu varijantu primene kompadding tehnike. Pored bitske brzine i amplitude maksimalnog opterećenja, veličina koja definiše ovakav tzv. kvazilogaritamski komparator je faktor kompresije  $\mu$ . Kompresorska funkcija za  $\mu$  zakon kompresije je definisana kao [3], [6]:

$$c_\mu(x) = \frac{x_{\max}}{\ln(1 + \mu)} \ln \left( 1 + \mu \frac{|x|}{x_{\max}} \right) \operatorname{sgn}(x), \quad -x_{\max} \leq x \leq x_{\max}. \quad (3.3.3.1)$$

Kompresorska funkcija  $\mu$  zakona kompresije prikazana je na sledećem grafiku, za tipične vrednosti faktora kompresije  $\mu$ .



Slika 3.3.3.1. Kompresorska funkcija  $\mu$  zakona kompresije

Najčešće korišćena vrednost faktora kompresije je 255, koja se koristi kod PCM-a definisanog G.711 standardom, dok se kod projektovanja adaptivnih kvantizera primenjuju niže vrednosti faktora kompresije. SQNR za logaritamski komparator sa primenom  $\mu$  zakona kompresije se može izvesti po istom principu kao i kod onog sa  $A$  zakonom i može se izračunati kao:

$$\text{SQNR}^{\mu} = 6.02R + 4.77 - 20 \log_{10}(\ln(1 + \mu)) \quad (3.3.3.2)$$

## 4. PODOPSEŽNO KODOVANJE GOVORNOG I AUDIO SIGNALA

Algoritmi za kodovanje govornog i audio signala se često zasnivaju na iskorišćavanju pogodnih osobina i statističkih karakteristika signala, kako bi se postigao što bolji kvalitet kodovanog signala. Kako su u pitanju nestacionarni signali, njihovi statistički parametri se značajno menjaju u vremenu, što čini zadatak projektovanja algoritma za kodovanje složenijim. Iako se ne može sa sigurnošću predvideti promena signala u vremenu, moguće je odrediti frekvencijski podopseg signala u kome je sadržan najveći deo korisne informacije. Nakon određivanja podopsega sa najvećom koncentracijom korisne informacije, poželjno je taj podopseg izdvojiti i kodovati nezavisno od ostatka ulaznog

signala. Razdvajanje (dekompozicija) signala na frekvencijske podopsege i njihovo nezavisno kodovanje je moguće izvršiti primenom tehnike podopsežnog kodovanja. Iz literature je poznato da se tehnike podopsežnog kodovanja mogu uspešno primeniti u kodovanju govornog i audio signala [44]- [46], kao i u kodovanju slike i video signala [46]- [48]. U ovom odeljku se pored projektovanja novih algoritama i same primene podopsežnog kodovanja u kodovanju govornog i audio signala, akcentat stavlja i na projektovanje novih kvantizera koji su primenjeni u razvijenim algoritmima.

## 4.1 Osnove podopsežnog kodovanja

Podopsežno kodovanje se zasniva na podeli ulaznog signala na frekvencijske podopsege, sa ciljem njihove nezavisne obrade i kodovanja. Prednost primene ove tehnike kodovanja leži u tome što se frekvencijski podopsezi mogu tretirati kao nezavisni signali, čije je karakteristike moguće predvideti u većoj meri nego što je slučaj sa kompletnim ulaznim signalom. Prilikom projektovanja algoritma zasnovanog na podopsežnom kodovanju, potrebno je definisati nekoliko ključnih komponentata:

- Digitalni filter;
- Bit alokaciju;
- Kvantizer.

Svaka od navedenih komponenti u velikoj meri utiče na kvalitet kodovanog signala, a za uspešno projektovanje algoritma zasnovanog na podopsežnom kodovanju neophodan je i skladan odabir metoda korišćenih u definisanju navedenih komponenti.

Namena filtra je da izdvoji predstavlja uređaj ili proces koji izdvaja određeni frekvencijski podopseg signala, odnosno eliminiše neželjene komponente signala. Svaki filter se projektuje da vrši propuštanje određenog frekvencijskog podopsega ulaznog signala, dok se ostatak frekvencijskog opsega signala blokira, odnosno ne propušta se na izlazu. Filtri se generalno mogu podeliti na tri velike grupe [1]:

1. Niskopropusni (NF) filtri;
2. Visokopropusni (VF) filtri;
3. Filtri propusnici opsega (PO).

NF filtri propuštaju komponente ulaznog signala koje se nalaze ispod unapred definisane frekvencije  $f_0$ , VF filtri propuštaju komponente signala koje su iznad definisane frekvencije  $f_0$ , dok filtri propusnici opsega propuštaju frekvencijske komponente signala koje se nalaze iznad frekvencije  $f_1$ , ali ispod frekvencije  $f_2$ . U algoritmima predstavljenim u disertaciji, dekompozicija signala na podopsege vrši se primenom idealnih digitalnih NF i VF filtara.

Pored odabira i projektovanja odgovarajućih digitalnih filtara i dobijanja podopsega, veoma je važno pravilno rasporediti raspoloživi broj bita na podopsege. Da bi se tako nešto i postiglo, potrebno je iskoristiti činjenicu da određeni podopsezi nose veću količinu informacije i njima se može dodeliti veći broj bita. Pored toga što se bit alokacija može izvršiti na osnovu prethodnog iskustva i podataka iz literature, bit alokaciju je moguće odrediti teorijski na osnovu raspoloživih parametara sistema. Ukoliko je ukupan broj bita jednak  $R$ , a broj podopsega iznosi  $L$ , bit alokacija  $k$ -tog podopsega se može odrediti kao [1]:

$$R_k = R + \frac{1}{2} \log_2 \frac{\sigma_k^2}{\prod_{k=1}^L (\sigma_k^2)^{\frac{1}{L}}}, \quad (4.1.1)$$

gde je sa  $\sigma_k$  označena varijansa  $k$ -tog podopsega. U najjednostavnijem slučaju kada imamo 2 podopsega, teorijsku bit alokaciju možemo odrediti na osnovu prethodne jednačine kao:

$$R_1 = R + \frac{1}{4} \log_2 \left( \frac{1+\rho}{1-\rho} \right), \quad (4.1.2)$$

$$R_2 = R + \frac{1}{4} \log_2 \left( \frac{1-\rho}{1+\rho} \right), \quad (4.1.3)$$

gde je sa  $\rho$  označen koeficijent korelacije ulaznog signala. U idealnom slučaju za očekivati je da su teorijski i eksperimentalni rezultati optimalne bit-alokacije jednaki i o tome će biti govora u nastavku, prilikom opisivanja projektovanih algoritama. Definisana bit alokacija se nakon toga dodeljuje kvantizerima koji su zaduženi za određene podopsege ulaznog signala. Kvantizacija ima presudni uticaj na kvalitet kodovanog signala, dok ostali parametri imaju zadatak da omoguće što efikasniju primenu

kvantizera. Projektovanje kvantizera za primenu u podopsežnom kodovanju će biti opisano u nastavku poglavlja.

Algoritmi koji će biti predstavljeni u disertaciji primenjuju idealne digitalne filtre koji vrše dekompoziciju signala: digitalni FIR niskopropusni filter (NF), označen sa  $H_1(z)$  i digitalni FIR visokopropusni (VF) filter, označenom sa  $H_2(z)$ . Prenosne funkcije filtra određene sledećim relacijama [53]:

$$F_1(z) = H_2(-z) \text{ i } F_2(z) = -H_1(-z), \quad (4.1.4)$$

$$H_1(z) = H_2(-z), \quad (4.1.5)$$

$$H_2(z) = 1 - z^{-1}. \quad (4.1.6)$$

Jedna od najjednostavnijih veoma tačnih filter banaka za rekonstrukciju koristi Haar-ovu osnovu i definisana je kao [54]:

$$h_1(n) = (1/2, 1/2), \quad H_1(f) = e^{-j\pi f} \cos(\pi f), \quad (4.1.7)$$

$$h_2(n) = (1/2, -1/2), \quad H_2(f) = je^{-j\pi f} \sin(\pi f). \quad (4.1.8)$$

Na osnovu prethodnih dveju formula, filtri se u vremenskom domenu mogu modelovati uz pomoć sledećih aproksimacija:

$$y_1(n) = \frac{x(n) + x(n+1)}{2}, \quad n = 1, 2, \dots, M, \quad (4.1.9)$$

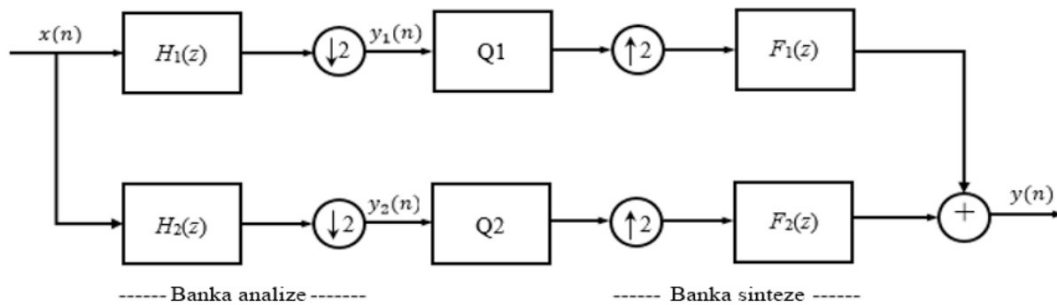
$$y_2(n) = \frac{x(n) - x(n+1)}{2}, \quad n = 1, 2, \dots, M, \quad (4.1.10)$$

gde  $x(n)$  i  $x(n+1)$  predstavljaju odmerke ulaznog signala,  $y_1(n)$  i  $y_2(n)$  označavaju signale u odgovarajućim podopsezima dobijene primenom digitalnih filtera, dok  $M$  označava ukupan broj ulaznih odmeraka. Dakle izrazi (4.1.9) i (4.1.10) predstavljaju aproksimacije idealnih digitalnih NF i VF filtera, respektivno. Ovako definisani filtri će biti primenjeni u podopsežnom kodovanju govornog i audio signala što će biti opisano u nastavku.

## 4.2 Algoritam za podopsežno kodovanje govornog i audio signala sa dva podopsega

### 4.2.1 Projektovanje i razvoj algoritama

U ovom odeljku će biti opisana najjednostavnija varijanta algoritma sa primenom podopsežnog kodovanja. Ulazni signal se deli na dva podopsega, odnosno dve „grane“, primenom idealnog NF i VF filtra, čiji su izlazi definisani izrazima (4.1.9) i (4.1.10). Dakle, nakon propuštanja ulaznog signala kroz filtre dobijamo dva nova signala, obeležena kao  $y_1(n)$  i  $y_2(n)$  (Slika 4.2.1.1), koji se dalje mogu obrađivati nezavisno i nad njima se mogu primeniti različite tehnike kvantizacije.



Slika 4.2.1.1. Blok šema podopsežnog kodovanja sa dve grane

Kvantizeri, obeleženi kao  $Q1$  i  $Q2$ , zaduženi su za kvantovanje odmeraka signala prvog i drugog podopsega, respektivno. U slučaju najjednostavnije kodne šeme zasnovane na podopsežnom kodovanju sa dva podopsega konstruisana su dva kvazilogaritamska kvantizera, odnosno logaritamska kvantizera zasnovana na  $\mu$  zakonu kompresije, sa različitim bitskim brzinama, projektovani za Laplasov izvor. Kvantizer sa višom bitskom brzinom je obeležen sa  $Q1$  i to je kvantizer projektovan za primenu na prvi podopseg, nakon primene NF filtra. Oba kvantizera primenjuju kompresorsku funkciju definisanu izrazom (3.3.3.1), sa identičnim faktorom kompresije za koji su razmatrane dve vrednosti,  $\mu=255$  i  $\mu=511$ . Ovako konstruisane kvantizere odlikuje robusnost, odnosno mogućnost da pruže uravnotežen kvalitet izlaznog signala u širokom opsegu amplituda ulaznih odmeraka. Parametar koji se razlikuje kod dva projektovana kvantizera je amplituda maksimalnog opeterećenja, koja je od velikog značaja za performanse kvantizera. Ovaj

parametar je određen na osnovu varijansi signala u podopsezima, koje se izračunavaju uz pomoć referentne varijanse  $\sigma_x^2$  kao [1]:

$$\sigma_1^2 = \frac{\sigma_x^2}{2}(1 + \rho) \quad (4.2.1.1)$$

$$\sigma_2^2 = \frac{\sigma_x^2}{2}(1 - \rho) \quad (4.2.1.2)$$

Algoritam koristi fiksne vrednosti amplitude maksimalnog opterećenja  $x_{\max 1}$  i  $x_{\max 2}$ , određene na osnovu iskustava iz prethodnih istraživanja, koja su pokazala da je amplituda maksimalnog opterećenja druge grane, koja je izlaz VF filtra, približno tri puta manja od vrednosti istog parametra prve grane, odnosno podopsega. Na osnovu toga određene su fiksne vrednosti parametara  $x_{\max 1} = 10$  i  $x_{\max 2} = 3$ .

Bit-alokacija je određena primenom izraza za teorijsko određivanje bit-alokacije (4.1.2) i (4.1.3) na govorni signal. U ovom slučaju, za algoritam za podopsežno kodovanje sa dva podopsega, dobija se da razlika između bitske brzine korišćene za kodovanje signala frekvencijskih podopsega treba da iznosi 2 bita po odmerku. Kako pretpostavljamo da prva grana sadrži veću količinu korisne informacije ulaznog signala, jasno je da veću bitsku brzinu treba dodeliti kvantizeru koji se nalazi u prvoj grani, nakon NF filtra. To u praksi znači da ako imamo raspoloživu bitsku brzinu od 9 bita po odmerku, najbolje rezultate očekujemo u slučaju kada je bitska brzina prvog kvantizera  $r_1=10$  i  $r_2=8$  bita po odmerku. Ova pretpostavka je eksperimentalno proverena, prilikom čega se pored teorijski određene bit-alokacije ispituju performanse algoritma sa različitom raspodelom bitske brzine.

## 4.2.2 Analiza eksperimentalnih rezultata

Eksperimentalno dobijeni rezultati su prikazani u sledećoj tabeli, dok su pored predloženog algoritma prikazani i rezultati primene klasičnog PCM kodera na iste ulazne signale, sa ciljem poređenja. Ulazni signali su signal muškog govornika, snimljen u akustičkom studiju na Elektronskom fakultetu u Nišu i signal klasične muzike, prilikom čega su oba signala sa frekvencijom odmeravanja od 44 kHz. Prilikom eksperimenta su korišćene različite bit-alokacije, kako bi se praktično utvrdila optimalna raspodela bitske brzine u podopsezima, kao i izvršila provera teorijski određene bit-alokacije.

Iz Tabele 4.2.2.1 možemo videti da se primenom teorijski određene bit-alokacije (koja predlaže da razlika u bitskoj brzini između podopsega bude jednaka 2) zaista ostvaruje najviši SQNR, pa za  $R_1=10$  i  $R_2=8$  i  $\mu = 511$  SQNR iznosi 45.79dB. Vidimo da primenom najjednostavnije varijante podopsežnog kodovanja i teorijske bit-alokacije ostvarujemo dobitak u odnosu na PCM od skoro 5 dB. Interesantno je i to što u slučaju primene jednake bitske brzine u oba podopsega vrednost SQNR-a je niža za 0.1 dB u odnosu na PCM. Ovo

Tabela 4.2.2.1. SQNR performanse algoritma za podopsežno kodovanje sa dva podopsega primenjenog na govorni signal za različite bitske brzine

R	$R_{sr}$	$\mu$	SQNR <sub>SBC</sub> [dB]	SQNR <sub>PCM</sub> [dB]
$R_1=9$	9	255	39.6559	39.7597
$R_2=9$		511	40.7983	40.8280
$R_1=10$	9	255	43.8295	39.7597
$R_2=8$		511	45.7902	40.8280
$R_1=8$	9	255	33.7779	39.7597
$R_2=10$		511	34.8456	40.8280
$R_1=11$	9	255	33.7779	39.7597
$R_2=7$		511	41.3042	40.8280
$R_1=7$	9	255	27.7532	39.7597
$R_2=11$		511	28.8205	40.8280
$R_1=12$	9	255	35.1428	39.7597
$R_2=6$		511	39.4752	40.8280
$R_1=6$	9	255	21.5487	39.7597
$R_2=12$		511	22.7751	40.8280

potvrđuje značaj optimalne bit-alokacije koja omogućava iskorišćenje punog potencijala algoritma za podopsežno kodovanje. Pored toga, udaljavanjem od teorijske bit-alokacije se gubi dobitak u SQNR-u, pa je za slučaj razlike bitske brzine od 4 i 6 bita po odmerku SQNR algoritma zasnovanog na podopsežnom kodovanju značajno niži od slučaja optimalne bit-alokacije kao i PCM-a.

Ovim eksperimentom je proverena i pretpostavka da je veća količina korisne informacije ulaznog signala sadržana u prvom podopsegu, koji je rezultat primene NF filtra. Iz Tabele možemo zaključiti da je i ova pretpostavka tačna, jer zamenom bitskih brzina prvog i drugog podopsega dobijamo SQNR niži za približno 11 dB. Ova velika razlika u dobijenom SQNR-u potvrđuje pretpostavku da je veća količina informacije sadržana u



prvom podopsegu koji je potrebno kodovati sa višom bitskom brzinom u odnosu na podopseg nakon VF filtra. Dodelom niže bitske brzine drugom podopsegu se smanjuje prosečna bitska brzina i time ostvaruje kompresija signala.

Tabela 4.2.2.2. SQNR performanse algoritma za podopsežno kodovanje sa dva podopsega primenjenog na audio signal klasične muzike za različite bitske brzine

R	$R_{sr}$	$\mu$	SQNR <sub>SBC</sub> [dB]	SQNR <sub>PCM</sub> [dB]
$R_1=9$	9	255	41.5437	41.5904
$R_2=9$		511	41.7644	41.7758
$R_1=10$	9	255	46.6482	41.5904
$R_2=8$		511	47.2528	41.7758
$R_1=8$	9	255	35.5722	41.5904
$R_2=10$		511	35.7789	41.7758

Tabela 4.2.2.2 prikazuje eksperimentalne rezultate primene algoritma za podopsežno kodovanje sa dva podopsega na audio signal klasične muzike. Iako se kod ovog signala ostvaruje nešto viši SQNR primenom obeju tehnika kodovanja, rezultati prate iste principe kao i u slučaju govornog signala. Najbolje performanse ostvarene su u slučaju kada je  $R_1=10$ ,  $R_2=8$  i  $\mu = 511$ , kada SQNR iznosi 47.25dB, dok dobitak u odnosu na PCM iznosi približno 5.5 dB.

Na osnovu rezultata predstavljenih u prethodne dve tabele, možemo zaključiti da je predloženi algoritam pogodan za kodovanje audio signala i govornog signala visokog kvaliteta. Ova najjednostavnija verzija algoritma zasnovanog na podopsežnom kodovanju ostvaruje značajni dobitak u SQNR-u, iako primenjuje jednostavne tehnike kvantizacije, identične kao kod PCM-a. Zaključujemo da je prikazani dobitak u SQNR-u direktni rezultat primene podopsežnog kodovanja, uz korišćenje optimalne bit-alokacije.

### 4.3 Adaptivni algoritam za podopsežno kodovanje govornog i audio signala sa dva podopsega

Prethodni odeljak je pokazao primenu jednostavne varijante podopsežnog kodovanja u kombinaciji sa kvazilogaritamskim kvantizerima niske kompleksnosti. Algoritam primenjuje jednostavne tehnike i robusne kvantizere koji pružaju visok kvalitet za govorne i audio signale različitih karakteristika, što je i pokazano eksperimentalno

dobijenim rezultatima. Nadgradnju ovako projektovanog algoritma moguće je vršiti u više domena:

- projektovanjem digitalnih filtara različitih karakteristika
- implementacijom složenijih mehanizama podopsežnog kodovanja
- primenom kompleksnijih kvantizera i tehnika kodovanja

U ovom odeljku će biti predstavljen algoritam zasnovan na podopsežnom kodovanju koji primenjuje adaptivnu kvantizaciju i njegova primena u kodovanju audio signala. Projektovani digitalni filtri i broj podopsega identični su algoritmu opisanom u prethodnom odeljku. Blok šema adaptivnog algoritma takođe odgovara blok šemi algoritma predstavljenog u prethodnom odeljku (Slika 4.2.1.1), pri čemu  $Q_1$  u  $Q_2$  u ovom slučaju predstavljaju adaptivne kvantizere. Recimo najpre šta je adaptivni kvantizer i kakve su potencijalne benefiti primene istih.

#### 4.3.1 Tehnike adaptacije

Dok se neadaptivni kvantizeri projektuju nezavisno od ulaznog signala koji će biti doveden na njihov ulaz, adaptivni kvantizeri su prilagođeni na neki od statističkih parametara ulaznog signala, najčešće na srednju vrednost ili standardnu devijaciju. Na taj način se postiže viši kvalitet kodovanog signala, uz neznatno povećanje kompleksnosti samog kvantizera, pa i algoritma u koji se primenjuje.

Adaptacija se najčešće primenjuje kao adaptacija unapred (*Forward Adatation*, FA) i adaptacija unazad (*Backward Adatation*, BA). Adaptacija unapred se zasniva na korišćenju još nekvantovanih odmeraka ulaznog signala, dok adaptacija unazad koristi već kvantovane odmerke ulaznog signala. Iako primena bilo koje od dveju tehnika adaptacije donosi poboljšanje objektivnog kvaliteta kodovanog signala, obe tehnike imaju svoje prednosti i mane, pa i sam izbor predstavlja kompromis. Prednost adaptacije unapred je ta što pruža veći SQNR u odnosu na adaptaciju unazad i to za oko 1 dB [3]. Ovaj dobitak je prisutan jer kod adaptacije unazad koristimo već kvantovane odmerke koji sadrže određenu grešku kvantizacije, koja se propagira u postupku adaptacije. Sa druge strane, nedostatak adaptacije unapred je taj što je za njenu primenu potrebno prenositi korišćeni statistički parametar kao sporednu informaciju, jer originalni odmerci nisu dostupni prijemniku i kako bi signal mogao uspešno da se dekoduje na prijemnoj

strani. Nasuprot tome, kod adaptacije unazad nema potrebe za prenosom sporedne informacije jer su odmerci na osnovu kojih adaptacija vrši dostupni prijemniku.

Najjednostavnija varijanta primene adaptacije jeste faktorisanje (množenje) reprezentacionih nivoa kvantizera vrednošću parametra koji se koristi u postupku adaptacije, npr. standardna devijacija signala. Kvantizer koji je konstruisan za primenu u podopsežnom kodovanju sa adaptacijom unazad primenjuje adaptaciju same vrednosti amplitude maksimalnog opeterećenja kvantizera, koja se vrši na osnovu standardne devijacije frejma ulaznog signala i optimalne vrednosti amplitude maksimalnog opeterećenja za Laplasov izvor, što će biti opisano u nastavku.

### 4.3.2 Projektovanje algoritma

Algoritam koji će biti predstavljen u ovom odeljku primenjuje podopsežno kodovanje sa dva podopsega i dva nezavisna kvazilogaritamska kvantizera sa adaptacijom unazad. Adaptacija unazad je odabrana kako bi se izbegao prenos dodatne informacije ali i usled očekivanja da će se postići dobitak u kvalitetu kodovanog signala i primenom samih tehnika podopsežnog kodovanja i adaptacije. Kvantizeri su projektovani na način sličan onom opisanom u prethodnom odeljku, sa primenom identične kompresorske funkcije, uz dodatak korišćenja adaptivne amplitude maksimalnog opeterećenja kvantizera, kao i različitih vrednosti faktora kompresije.

Pored adaptacije, algoritam primenjuje podelu ulaznog signala na frejmove, pa se kvantizacija i obrada signala izvršava frejm po frejm. Tako je i parametar korišten za adaptaciju kvantizera standardna devijacija frejma ulaznog signala. Podela signala na frejmove i primena adaptacije unazad unose potencijalni problem u razvoj algoritama, jer za prvi frejm ne postoje parametri na osnovu kojih bi se adaptacija unazad mogla primeniti. Ovaj problem je prevaziđen tako što se na prvi frejm vrši primena ne-adaptivnog kvantizera, odnosno kvantizera čija amplituda maksimalnog opeterećenja nije adaptirana na statistiku ulaznog signala. Ovaj postupak predstavlja inicijalizaciju algoritma, koji će nakon toga dobiti parametre za dalju primenu adaptacije. Algoritam se inicijalizuje vrednostima amplitude maksimalnog opeterećenja koje su korištene kod ne-adaptivnog algoritma opisanog u odeljku 4.2, pa su za prvi frejm korišćene vrednosti  $x_{\max 1} = 10$  i  $x_{\max 2} = 3$ .

Nakon obrade prvog frejma, može se nesmetano primeniti adaptacija unazad, kako su prijemniku dostupni kvantovani odmerci za prethodni (prvi) frejm. Adaptivne amplitude maksimalnog opterećenja se procenjuju iterativno za svaki frejm, na osnovu varijanse podopsega prethodnog frejma i optimalne vrednosti amplitude maksimalnog opterećenja za Laplasov izvor, koja je definisana kao:

$$x_{\max}^{\text{opt}} = \frac{1}{\sqrt{2}} \ln \left( \frac{3\mu N^2}{\ln^2(\mu + 1)} \right), \quad (4.3.2.1)$$

gde  $N$  predstavlja broj kvantizacionih nivoa kvantizera. Nakon određivanja optimalne vrednosti amplitude maksimalnog opterećenja, potrebno je odrediti varijanse trenutnog  $i$ -tog frejma odmeraka podopsega,  $\sigma_1^{2(i)}$  i  $\sigma_2^{2(i)}$ , koji se mogu proceniti primenom sledećih izraza: [6]:

$$\sigma_1^{2(i)} = \frac{\sigma_x^{2(i)}}{2} (1 + \rho^{(i)}), \quad (4.3.2.2)$$

$$\sigma_2^{2(i)} = \frac{\sigma_x^{2(i)}}{2} (1 - \rho^{(i)}), \quad (4.3.2.3)$$

prilikom čega  $\rho^{(i)}$  predstavlja koeficijent korelacije trenutnog frejma ulaznog signala, a  $\sigma_x^{2(i)}$  označava varijansu trenutnog frejma ulaznog signala. Varijansa frejma ulaznog signala se može odrediti kao:

$$\sigma_x^{2(i)} = \frac{1}{L_f} \sum_{r=1}^{L_f} x(r)^2, \quad (4.3.2.4)$$

gde  $L_f$  predstavlja dužinu frejma. Koeficijent korelacije trenutnog frejma odmeraka ulaznog signala se određuje kao:

$$\rho^{(i)} = \frac{\sum_{n=1}^{L-1} x_i(n) \times x_i(n+1)}{\sum_{n=1}^L x_i^2(n)}, \quad (4.3.2.5)$$

gde  $x_i(n)$  i  $x_i(n+1)$  predstavljaju ulazne odmerke trenutnog frejma ulaznog signala. U ovom trenutku možemo odrediti adaptivne vrednosti amplitude maksimalnog

opeterećenja podopsega,  $x_{\max 1}^{\text{BA}(i)}$  i  $x_{\max 2}^{\text{BA}(i)}$ , množenjem vrednosti standardne devijacije podopsega prethodnog frejma sa optimalnom vrednošću amplitude maksimalnog opeterećenja za Laplasov izvor:

$$x_{\max 1}^{\text{BA}(i)} = x_{\max}^{\text{opt}} \sigma_1^{(i-1)}, i = 2, 3, \dots, P, \quad (4.3.2.6)$$

$$x_{\max 2}^{\text{BA}(i)} = x_{\max}^{\text{opt}} \sigma_2^{(i-1)}, i = 2, 3, \dots, P, \quad (4.3.2.7)$$

gde P predstavlja ukupan broj frejmova ulaznog signala. Ovi koraci predstavljaju iterativni postupak, kojim se vrši adaptacija amplitude maksimalnog opeterećenja kvantizera primenjenih u podopsežnom kodovanju, za svaki frejm. Ovaj proces takođe predstavlja primenu adaptacije unazad, kako se za određivanje adaptivne amplitude maksimalnog opeterećenja trenutnog frejma, koriste informacije iz prethodno kodovanog frejma.

U nastavku će biti predstavljena analiza eksperimentalno dobijenih rezultata primene opisanog adaptivnog algoritma za podopsežno kodovanje sa dva podopsega na audio signale.

### 4.3.3 Analiza eksperimentalnih rezultata

Eksperiment je sproveden primenom projektovanog algoritma na dva audio signala različitih karakteristika, prilikom čega jedan predstavlja uzorak klasične muzike, dok je drugi uzorak rok muzike. Oba signala su sa frekvencijom odmeravanja od 44 kHz, dok su ostali parametri od značaja predstavljeni u sledećoj tabeli:

Tabela 4.3.3.1 prikazuje statističke parametre ulaznih signala od značaja za konstrukciju kvantizera i analizu rezultata. Možemo videti da su oba signala okarakterisana visokom amplitudskom dinamikom i relativno niskom varijansom. Takođe, visok koeficijent korelacije nagoveštava uspešnu primenu tehnike adaptacije.

Tabela 4.3.3.1 Statistički parametri ulaznih signala korišćenih u eksperimentu

Ulazni signal	Broj odmeraka	Varijansa signala	Koeficijent korelacije	Amplitudska dinamika [dB]
Klasična muzika	832531	0.0089	0.9938	90.0529
Rok muzika	2741355	0.0361	0.9776	96.32

Kao što je i opisano, adaptivne amplitude maksimalnog opterećenja kvantizera se procenjuju za svaki frejm u oba podopsega kako bismo imali prilagođenje na statistiku ulaznog signala. Nakon primene predloženog algoritma na oba ulazna signala, izvršili smo analizu dobijenih adaptivnih amplituda maksimalnog opterećenja kvantizera, kako bismo videli odstupanje od vrednosti odabrane za inicijalizaciju algoritma i prvi frejm. Naredna tabela predstavlja srednje vrednosti amplituda maksimalnog opterećenja oba kvantizera, nakon primene na realne ulazne signale.

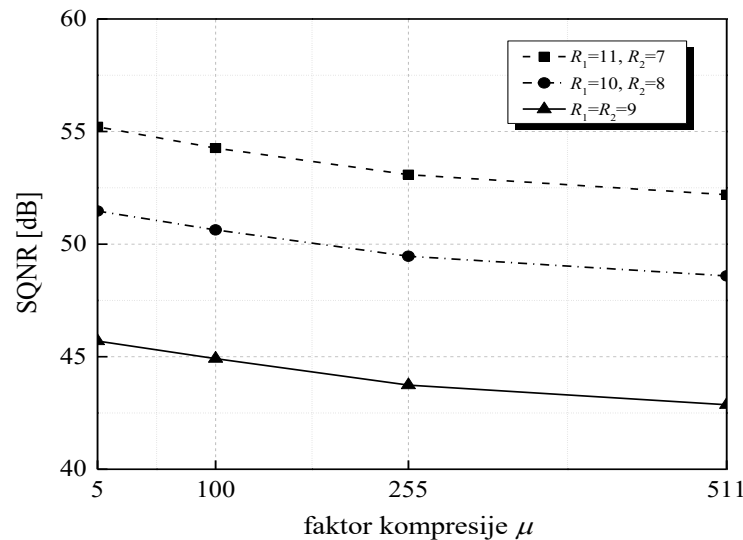
Tabela 4.3.3.2 Prosečne vrednosti amplitude maksimalnog opterećenja kvantizera sa adaptacijom unazad

Ulazni signal	Prosečna sr. vr. 1	Prosečna sr. vr. 2
	$x_{\max sr1}$	$x_{\max sr2}$
Klasična muzika	0.2786	0.0371
Rok muzika	1.1944	0.1275

Tabela pokazuje da su dobijene vrednosti drastično različite od polaznih vrednosti kojima je algoritam inicijalizovan, koje su, da se podsetimo,  $x_{\max 1} = 10$  i  $x_{\max 2} = 3$ . Takođe, iz tabele se može uočiti da se kod signala rok muzike, koji je odlikovan višom amplitudskom dinamikom i nešto višom korelacijom ostvaruju više vrednosti srednje vrednosti adaptivne amplitude maksimalnog opterećenja.

Kao što smo videli kod prethodnog, ne-adaptivnog algoritma, pored parametara kvantizera, u razvoju algoritama zasnovanog na podopsežnom kodovanju veoma važnu ulogu igra bit alokacija. Zamenom vrednosti koeficijenta korelacije iz Tabele 4.3.3.1 u

izraze (4.1.2) i (4.1.3) teorijski određujemo bit-alokaciju koja je u ovom slučaju, za ukupnu bitsku brzinu od 9 bita po odmerku određena kao  $R_1=11$  i  $R_2=7$ . Vidimo da je sada razlika između bitske brzine podopsega 4 bita po odmerku, dok je kod prethodnog algoritma ova razlika iznosila 2 bita po odmerku. Teorijske vrednosti bit-alokacije su takođe proverene eksperimentalno, o čemu će biti reči u nastavku.



Slika 4.3.3.1. SQNR za različite bit-alokacije i ulazni signal klasične muzike

Slika 4.3.3.1 predstavlja SQNR karakteristiku predloženog algoritma za podopsežno kodovanje audio signala sa adaptacijom unazad, za različite bit-alokacije i faktore kompresije, primenjenog na signal klasične muzike. Analiza različitih faktora kompresije se vrši jer se na osnovu literature očekuje da adaptivni kvantizer ostvaruje bolje performanse prilikom korišćenja nižih vrednosti faktora kompresije. Grafik nam to i potvrđuje kako vidimo da kriva SQNR-a ima blagi pad sa povećanjem faktora kompresije i da se najviši SQNR ostvaruje za faktor kompresije  $\mu=5$ . Pored toga, eksperimentalno je potvrđeno da je teorijski određena bit-alokacija ujedno i najbolje rešenje, kako se njenom primenom ostvaruje najviši objektivni kvalitet kodovanog signala. Sa slike lako možemo uočiti značajan rast SQNR-a kada je  $R_1=11$  i  $R_2=7$ , u odnosu na ostale korišćene bit-alokacije, što dodatno ističe značaj pravilne raspodele bitske brzine između podopsega.

Nakon što je utvrđena optimalna bit-alokacija, potrebno je analizirati rezultate za slučajeve korišćenja različitih dužina frejma, kao i napraviti poređenje sa nekim od tehnika kodovanja slične kompleksnosti. Naredna tabela pokazuje uporedne rezultate

predloženog algoritma za podopsežno kodovanje sa adaptacijom unazad, ne-adaptivnog algoritma za podopsežno kodovanje, predloženog u prethodnom odeljku, kao i PCM kodera, primenjenih na iste ulazne signale. Cilj je analiza uticaja primene adaptacije u algoritmu, za koju se očekuje da donese poboljšanje SQNR-a u odnosu na algoritma bez primene adaptacije, pogotovu u slučaju primene nižih vrednosti faktora kompresije. Performanse adaptivnog algoritma su obeležene kao  $SQNR_{AUSBC2}$ , ne-adaptivnog kao  $SQNR_{SBC}$ , a PCM-a kao  $SQNR_{PCM}$ .

Tabela 4.3.3.3 SQNR za različite dužine frejma i faktore kompresije ostvaren za signal klasične muzike

Bitska brzina [bit po odmerku]	Dužina frejma [odmeraka]	Faktor kompresije $\mu$	$SQNR_{AUSBC2}$ [dB]	$SQNR_{SBC}$ [dB]	$SQNR_{PCM}$ [dB]
9	40	5	55.2136	28.5767	26.9639
		100	54.2638	44.1071	40.0344
		255	53.0825	46.6482	41.5904
		511	52.1915	47.2528	41.7758
	80	5	54.5946	28.5767	26.9639
		100	54.4935	44.1071	40.0344
		255	53.3451	46.6549	41.5904
		511	52.4748	47.2552	41.7758

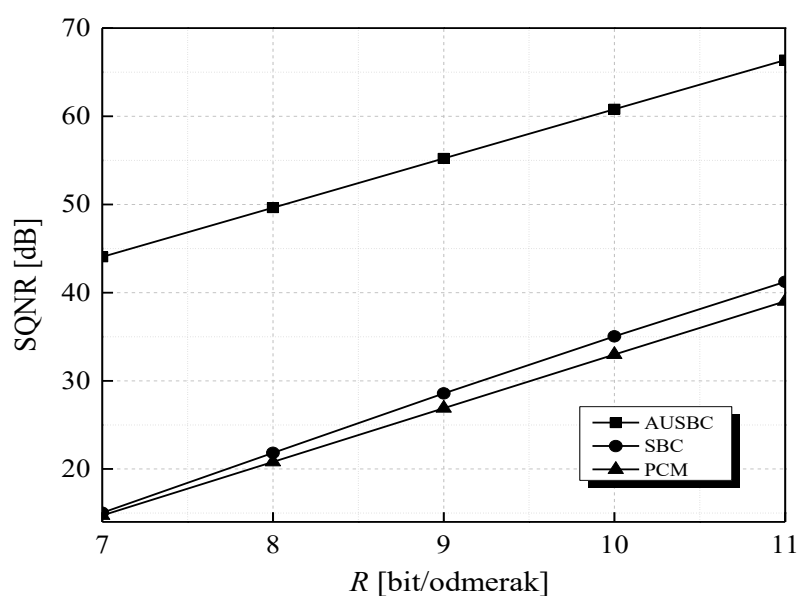
Tabela predstavlja rezultate primene predloženog algoritma na signal klasične muzike, za različite dužine frejma, kao i faktore kompresije. Treba naglasiti da se dužina frejma ne odnosi na PCM i neadaptivni algoritam za podopsežno kodovanje sa dva podopsega, jer oni ne primenjuju podelu signala na frejmove. Takođe, iako je ukupna bitska brzina identična kod svih tehnika koje se porede, dva algoritma za podopsežno kodovanje primenjuju različite bit-alokacije, jer se razmatraju i porede slučajevi kada oba algoritma ostvaruju najbolje performanse. Tako da je  $SQNR_{AUSBC2}$  ostvaren za bit-alokaciju  $R_1=11$  i  $R_2=7$ , dok je  $SQNR_{SBC}$  ostvaren za bit-alokaciju  $R_1=10$  i  $R_2=8$ .

Najpre treba istaći da adaptivni algoritam ostvaruje bolje performanse kada je dužina frejma 40 odmeraka, što se može objasniti boljom adaptacijom kvantizera na manji frejm, kako se na manjem uzorku očekuje veća korelacija između odmeraka i niža amplitudska dinamika u odnosu na celokupan signal. Usled toga, analiza rezultata će biti vršena za

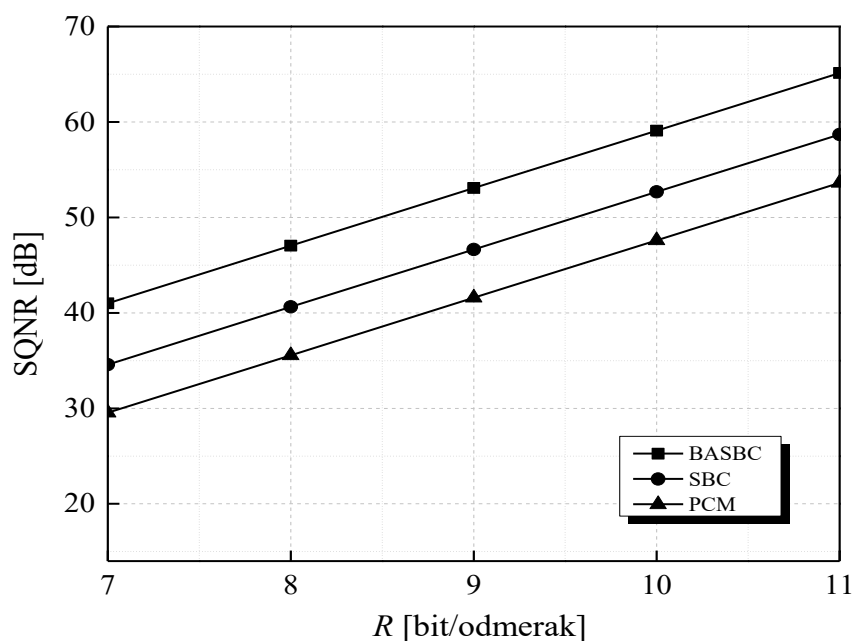


slučaj dužine frejma od 40 odmeraka. Iz Tabele se može uočiti značajan dobitak u SQNR-u adaptivnog algoritma u odnosu na neadaptivni i PCM za slučaj primene faktora kompresije  $\mu=5$ . Ova činjenica predstavlja dobitak koji je očigledni rezultat primene adaptivne kvantizacije, jer su performanse ne-adaptivnog kvantizera na nižim faktorima kompresije dosta slabije. Tako je u ovom slučaju  $SQNR_{AUSBC2}=55.21$  dB i dobitak u SQNR-u iznosi približno 26.6 dB u odnosu na neadaptivni algoritam i 28.25 dB u odnosu na PCM. Sa porastom faktora kompresije dobitak adaptivnog algoritma opada, ali je i dalje prisutan i nije zanemarljiv. Tako je za faktor kompresije  $\mu=511$  dobitak adaptivnog algoritma u odnosu na neadaptivni približno 4.9 dB, dok u odnosu na PCM dobitak iznosi približno 10.4 dB. Možemo zaključiti da se ispravnim odabirom parametara algoritma kao što je dužina frejma, bit-alokacija i faktor kompresije može ostvariti izuzetan kvalitet izlaznog signala i značajan dobitak u SQNR-u u odnosu na poređene tehnike.

Performanse algoritma se takođe mogu analizirati posmatranjem SQNR karakteristike. Sledeći grafik predstavlja SQNR karakteristiku adaptivnog algoritma za podopsežno kodovanje, ne-adaptivnog algoritma i PCM-a, primenjenog na signal klasične muzike, za različite ukupne bitske brzine i faktor kompresije  $\mu=5$ .



Slika 4.3.3.2. Uporedna SQNR karakteristika algoritma za ulazni signal klasične muzike i faktor kompresije  $\mu=5$



Slika 4.3.3.3. Uporedna SQNR karakteristika algoritma za ulazni signal klasične muzike i faktor kompresije  $\mu=255$

Na slici se može jasno uočiti značajan dobitak predloženog adaptivnog algoritma za podopsežno kodovanje sa dva podopsega u odnosu na neadaptivni algoritam i PCM, za sve razmatrane bitske brzine i faktor kompresije  $\mu=5$ . Naredna slika prikazuje sličnu uporednu SQNR karakteristiku, za slučaj upotrebe faktora kompresije  $\mu=225$ , korišćenog kod standardnog PCM-a. SQNR karakteristika sa slike 4.3.3.2 pokazuje nešto manji ali značajan dobitak za sve korišćene bitske brzine.

Naredna tabela prikazuje eksperimentalne rezultata primene predloženog algoritma na signal rok muzike, uz uporedni prikaz rezultata ne-adaptivnog algoritma za podopsežno kodovanje sa dva podopsega i PCM-a.  $SQNR_{PCM}$ .

Rezultati prikazani u Tabeli 4.3.3.4 pokazuju slične pravilnosti kao u slučaju primene algoritama na signal klasične muzike. Maksimalni dobitak u SQNR-u adaptivnog algoritma u odnosu na neadaptivni iznosi 16 dB za faktor kompresije  $\mu=5$ . Sa povećanjem faktora kompresije dobitak u SQNR-u očekivano opada, ali je prisutan za sve razmatrane vrednosti faktora kompresije. Nakon analize eksperimentalnih rezultata možemo zaključiti da su sve teorijske pretpostavke eksperimentalno potvrđene i da je predloženi

algoritam za podopsežno kodovanje audio signala sa dva podopsega uz primenu adaptacije unazad pogodan za primenu u kodovanju audio signala, pre svega u slučaju korišćenja niskih faktora kompresije.

Tabela 4.3.3.4 SQNR za različite dužine frejma i faktore kompresije ostvaren za signal rok muzike

Bitska brzina [bit po odmerku]	Dužina frejma [odmeraka]	Faktor kompresije $\mu$	SQNR <sub>AUSBC2</sub> [dB]	SQNR <sub>SBC</sub> [dB]	SQNR <sub>PCM</sub> [dB]
9	40	5	50.6572	34.6374	32.4916
		100	49.9723	46.9182	42.5586
		255	48.8659	47.7685	42.7975
		511	48.0651	47.5891	42.4244
	80	5	50.6080	34.6374	32.4916
		100	50.3481	46.9182	42.5586
		255	49.2487	47.7685	42.7975
		511	48.4233	47.5891	42.4244

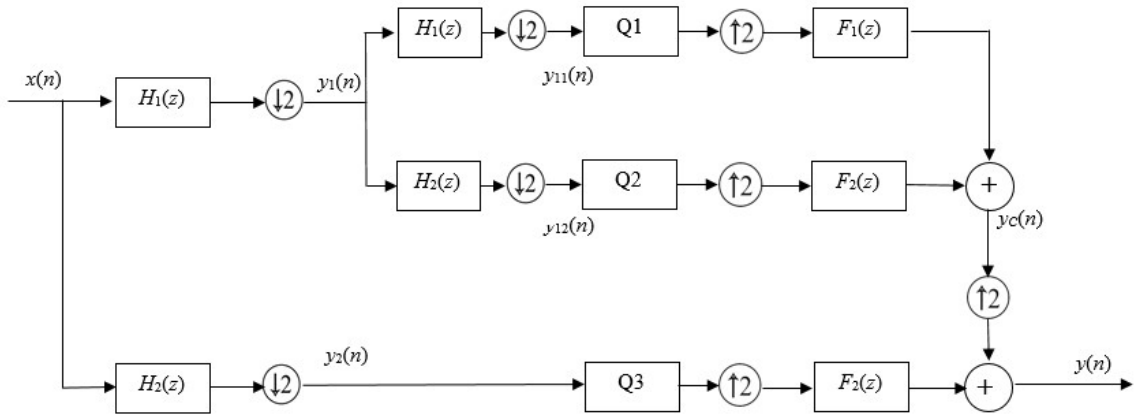
#### 4.4 Algoritam za podopsežno kodovanje audio signala sa tri podopsega i adaptacijom unazad

U ovom odeljku će biti predstavljen algoritam zasnovan na podopsežnom kodovanju koji deli ulazni signal na tri nezavisna podopsega. U pitanju je adaptivni algoritam, koji predstavlja nadgradnju algoritma predstavljenog u odeljku 4.3. Kako je poznato u teoriji i praktično potvrđeno da je veća količina korisne informacije govornog i audio signala sadržana u podopsegu koji sadrži komponente signala na nižim frekvencijama, na ovom podopsegu će biti primenjena dodatna obrada u postupku kodovanja.

##### 4.4.1 Projektovanje algoritma i primena podopsežnog kodovanja

Algoritam za podopsežno kodovanje govornog i audio signala sa tri podopsega, vrši kodovanje ulaznog signala frejm po frejm, uz primenu adaptacije unazad, po istim principima opisanim u prethodnom odeljku. Najpre se primenjuje podela ulaznog signala na dve grane, kao i u prethodno predstavljenim algoritmima. Dodatni, treći podopseg

nastaje daljom podelom prvog podopsega na još dve grane. Na ovaj način je omogućena je dodatna dodela prioriteta u raspodeli bitske brzine između podopsega kako bi se postigla bolja kompresija kodovanog signala. Blok šema algoritma za podopsežno kodovanje sa tri podopsega je prikazana na sledećoj slici:



Slika 4.4.1.1. Blok šema podopsežnog kodovanja sa tri podopsega

Algoritam primenjuje tri nezavisna kvantizera, koji su u prikazanoj blok šemi označeni kao  $Q_1$ ,  $Q_2$  i  $Q_3$ . U pitanju su kvazilogaritamski kvantizeri, odnosno kvantizeri koji primenjuju logaritamski  $\mu$  zakon kompresije. Filtri analize i sinteze su identični onima opisanim u odeljku 4.2. Dodatno, algoritam vrši podelu grane koja je predstavljala prvi podopseg ( $y_1$ ) na još dva podopsega ( $y_{11}$  i  $y_{12}$ ). Konačno, u blok šemi imamo ukupno tri podopsega koji su obeleženi kao  $y_{11}$ ,  $y_{12}$  i  $y_2$ , respektivno. Dva novonastala podopsega dobijena su primenom identične transformacije kao što je slučaj kod algoritma za podopsežno kodovanje sa dva podopsega. Podsećanja i celovitosti radi, biće navedeni izrazi za dobijanje podopsega  $y_1$  i  $y_2$ :

$$y_1(n) = \frac{x(n) + x(n+1)}{2}, n = 1, 2, \dots, M, \quad (4.4.1.1)$$

$$y_2(n) = \frac{x(n) - x(n+1)}{2}, n = 1, 2, \dots, M. \quad (4.4.1.2)$$

Nakon primene ovih transformacija, konačno korišćeni prvi i drugi podopseg se određuju kao:

$$y_{11}(n) = \frac{y_1(n) + y_1(n+1)}{2}, n = 1, 2, \dots, M, \quad (4.4.1.3)$$

$$y_{12}(n) = \frac{y_1(n) - y_1(n+1)}{2}, n = 1, 2, \dots, M, \quad (4.4.1.4)$$

Nakon ovog koraka dobijamo tri nezavisna podopsega koji se dovode na ulaze tri različita kvantizera. Kvantovane signale je potrebno rekonstruisati na prijemu, što predstavlja inverzni proces razdvajanja signala na podopsege. Najpre treba kombinovati signale prva dva opsega, kako bi se rekonstruisao signal koji je rezultat primene NF filtra. Ovo je moguće postignuti naizmeničnim uzimanjem svakog drugog odmerka prva dva podopsega. Na taj način dobijamo rekonstruisani kvantovani signal koji je originalno bio rezultat primene NF filtra na ulazni signal. Ovaj signal je u blok šemi obeležen sa  $y_c$  i definisan je kao:

$$y_c(n) = [y_{11}(n), y_{12}(n+1)], n = 1, 2, \dots, M. \quad (4.4.1.5)$$

Isti princip primenjujemo još jednom, spajanjem signala  $y_c$  sa kvantovanom reprezentacijom trećeg podopsega, čime dobijamo rekonstruisani kvantovani ulazni signal:

$$y(n) = [y_c(n), y_2(n+1)], n = 1, 2, \dots, M. \quad (4.4.1.6)$$

U ovom trenutku se može analizirati objektivni kvalitet kodovanog signala, što će biti predstavljeno u nastavku, nakon definisanja kvantizera parametra kodovanja.

#### 4.4.2 Projektovanje kvantizera i određivanje parametra kodovanja

Algoritam primenjuje kvantizacione tehnike slične onima opisanim u prethodnom poglavlju. Konstruisani su logaritamski kvantizeri sa  $\mu$  zakonom kompresije i primenom adaptacije unazad, projektovani za Laplasov izvor. Adaptacija se primenjuje po istom principu kao u prethodno predstavljenom algoritmu, pri čemu je usled postojanja 3 podopsega postupak nešto složeniji. Primenjuje se adaptacija amplitude maksimalnog opterećenja na osnovu prethodno kvantovanih odmeraka, odnosno adaptacija unazad. Adaptivna amplituda maksimalnog opterećenja se određuje za svaki frejm pojedinačnih podopsega, na osnovu varijanse signala podopsega i optimalne vrednosti amplitude maksimalnog opterećenja za Laplasov izvor ( $x_{max}^{opt}$ ), koja je definisana izrazom (4.3.2.1). Nakon toga je potrebno odrediti varijanse za svaki podopseg, što je moguće učiniti na osnovu sledećih izraza:

$$\sigma_1^{2(i)} = \frac{\sigma_x^{2(i)}}{4} (1 + \rho_1)(1 + \rho_2), \quad (4.4.2.1)$$

$$\sigma_2^{2(i)} = \frac{\sigma_x^{2(i)}}{4} (1 + \rho_1)(1 - \rho_2), \quad (4.4.2.2)$$

$$\sigma_3^{2(i)} = \frac{\sigma_x^{2(i)}}{2} (1 - \rho_1), \quad (4.4.2.3)$$

prilikom čega su korišćene promenljive označene i definisane kao u prethodnom odeljku. Na osnovu ovih parametara moguće je odrediti adaptivne vrednosti amplituda maksimalnog opterećenja za svaki frejm pojedinačnih podopsega:

$$x_{\max 1}^{(i)} = x_{\max}^{\text{opt}} \sigma_1^{(i-1)}, i = 2, 3, \dots, P, \quad (4.4.2.4)$$

$$x_{\max 2}^{(i)} = x_{\max}^{\text{opt}} \sigma_2^{(i-1)}, i = 2, 3, \dots, P, \quad (4.4.2.5)$$

$$x_{\max 3}^{(i)} = x_{\max}^{\text{opt}} \sigma_3^{(i-1)}, i = 2, 3, \dots, P, \quad (4.4.2.6)$$

gde  $P$  označava ukupan broj frejmova. Pored parametara kvantizera, potrebno je odrediti optimalnu bit alokaciju, koja će odgovarati broju podopsega i karakteristikama ulaznih signala za čije kodovanje se algoritam projektuje. U prethodno predstavljenim algoritmima videli smo da je u slučaju primene dva podopsega potrebno dodeliti višu bitsku brzinu prvom podopsegu. Za algoritam koji primenjuje podopsežno kodovanje sa tri podopsega se očekuje da najveći deo korisne informacije ulaznog signala sadrži u prvom podopsegu, koji je praktično rezultat primene dva NF filtra. Vodeći se istom logikom kao kod prethodno predstavljenih algoritama, može se zaključiti da je prvom podopsegu potrebno dodeliti značajno višu bitsku brzinu nego preostalim podopsezima. Ukupna bitska brzina algoritma je definisana kao:

$$R = \frac{\frac{R_1 + R_2}{2} + R_3}{2}, \quad (4.4.2.6)$$

gde  $R_1$ ,  $R_2$  i  $R_3$  označavaju bitske brzine dodeljene prvom, drugom i trećem podopsegu respektivno. Posmatranjem prethodne formule lako možemo uočiti na koji način se postiže kompresija signala primenom predloženog algoritma za podopsežno kodovanje.  $R_1$  predstavlja bitsku brzinu prvog podopsega koji sadrži najveću količinu korisne

informacije. Uočavamo da u ukupnoj bitskoj brzini ona učestvuje podeljena sa 4, jer prvi deo izraza predstavlja bitsku brzinu podopsega koji je rezultat prve dekompozicije signala, nakon NF filtra  $\left(\frac{R_1+R_2}{2}\right)$ , čija aritmetička sredina sa bitskom brzinom trećeg podopsega daje ukupnu bitsku brzinu. Dakle, dodelom više bitske brzine prvom podopsegu ukupna bitska brzina se neće značajno povećati uz primenu kompenzacije i dodele niže bitske brzine ostalim podopsezima.

### 4.4.3 Analiza eksperimentalnih rezultata

Ekperimenti su sprovedeni na realnim ulaznim audio signalima, koji su korišćeni u prethodnim analizi performansi prethodnih algoritama. Takođe, koristi se ista ukupna bitska brzina od 9 bita po odmerku, kako bi eksperimentalni rezultati do sada predloženih algoritama bili uporedivi. Teorijski određena bit-alokacija za oba signala predlaže da bitske brzine podopsega budu:  $R_1=11$ ,  $R_2=9$ ,  $R_3=8$ . Teorijska bit-alokacija će biti praktično testirana, sa ciljem pronalaženja optimalne bit-alokacije, koja se u praksi može razlikovati od teorijske.

Najpre će predloženi algoritam za podopsežno kodovanje sa tri podopsega biti primenjen na audio signal klasične muzike za slučajeve korišćenja različitih bit-alokacija. Nakon pronalaženja optimalne bit-alokacije, ona će se primeniti i na drugi ulazni signal, zajedno sa teorijski predloženom bit-alokacijom. Pored toga eksperimentalno dobijeni rezultati će biti upoređeni sa PCM-om, kao i sa algoritmom za podopsežno kodovanje sa dva podopsega.

Tabela 4.4.3.1. prikazuje eksperimentalno dobijenu objektivnu meru kvaliteta kodovanog signala klasične muzike primenom algoritma za podopsežno kodovanje sa tri podopsega uz primenu adaptacije unazad, kao i PCM-a. Rezultati su prikazani za bitsku brzinu od 9 bita po odmerku i dužinu frejma od 80 odmeraka. Razmatrana je primena dve vrednosti faktora kompresije  $\mu=100$  i  $\mu=255$ . Performanse su razmatrane za različite bit-alokacije, kako bi se utvrdila optimalna raspodela bitske brzine između podopsega.

Tabela 4.4.3.1 SQNR algoritma za podopsežno kodovanje sa tri podopsega za različite faktore kompresije i bit-alokacije i PCM-a za signal klasične muzike

R	$R_{sr}$	$\mu$	SQNR <sub>AUSBC3</sub> [dB]	SQNR <sub>PCM</sub> [dB]
$R_1=11$ $R_2=9$ $R_3=8$	9	100	55.8300	40.0344
		255	54.7967	41.5904
$R_1=9$ $R_2=9$ $R_3=9$	9	100	45.0911	40.0344
		255	43.8863	41.5904
$R_1=12$ $R_2=10$ $R_3=7$	9	100	57.2355	40.0344
		255	56.2075	41.5904
$R_1=13$ $R_2=11$ $R_3=6$	9	100	53.2730	40.0344
		255	52.3010	41.5904
$R_1=13$ $R_2=7$ $R_3=8$	9	100	54.1872	40.0344
		255	54.9302	41.5904
$R_1=12$ $R_2=8$ $R_3=8$	9	100	57.5813	40.0344
		255	57.3567	41.5904
$R_1=13$ $R_2=9$ $R_3=7$	9	100	58.0180	40.0344
		255	57.3135	41.5904

Prvi red tabele predstavlja performanse predloženog algoritma za teorijski određenu bit-alokaciju  $R_1=11$ ,  $R_2=9$ ,  $R_3=8$ . Iako u tom slučaju dobitak u odnosu na PCM iznosi 15.8 dB (za faktor kompresije  $\mu=100$ ) možemo uočiti da se za teorijski određenu bit-alokaciju ne ostvaruje najviši SQNR. Eksperimentalno je ustanovljeno da se za ulazni signal klasične muzike najviša vrednost SQNR-a ostvaruje za bit-alokaciju određenu kombinacijom  $R_1=13$ ,  $R_2=9$ ,  $R_3=7$ , gde se za faktor kompresije  $\mu=100$  ostvaruje SQNR od 58.0180 dB. To predstavlja dobitak od približno 18 dB u odnosu na PCM, što ukazuje na izuzetne performanse algoritma za podopsežno kodovanje sa tri podopsega koji primenjuje adaptaciju unazad. Tabela 4.4.3.1 još jednom potvrđuje prednosti primene tehnika podopsežnog kodovanja u algoritmima za kodovanje govornog i audio signala kojom se postiže značajna kompresija signala, uz zadržavanje relativno niske kompleksnosti algoritma. Tabela pokazuje i veliki doprinos pravilno određene bit-



alokacije, koji možemo brojno izraziti poređenjem performansi algoritama kada je  $R_1=R_2=R_3=9$  sa slučajem  $R_1=13, R_2=9, R_3=7$ . Iz Tabele vidimo da se dodelom više bitske brzine prvom podopsegu a najniže trećem ostvaruje dobitak u SQNR-u od 12.9269 dB u odnosu na slučaj kada su svim podopsezima dodeljene identične bitske brzine. Ovo potvrđuje pretpostavku načinjenu u projektovanju algoritama i pokazuje značaj pravilne bit-alokacije za ostvarivanje više kompresije signala.

Predloženi algoritam je primenjen i na audio signal rok muzike, za nekoliko slučajeva karakterističnih bit alokacija, dok je dužina frejma ostala identična (80 odmeraka), za dva različita faktora kompresije.

Tabela 4.4.3.2 SQNR algoritma za podopsežno kodovanje sa tri podopsega za različite faktore kompresije i bit-alokacije za signal rok muzike

R	$R_{sr}$	$\mu$	SQNR <sub>AUSBC3</sub> [dB]	SQNR <sub>PCM</sub> [dB]
$R_1=11$ $R_2=9$ $R_3=8$	9	100	53.6578	42.5586
		255	52.6046	42.7975
$R_1=9$ $R_2=9$ $R_3=9$	9	100	44.9953	42.5586
		255	43.7754	42.7975
$R_1=12$ $R_2=10$ $R_3=7$	9	100	52.4473	42.5586
		255	51.3561	42.7975
$R_1=12$ $R_2=8$ $R_3=8$	9	100	54.1655	42.5586
		255	53.5793	42.7975
$R_1=13$ $R_2=9$ $R_3=7$	9	100	52.7411	42.5586
		255	51.7719	42.7975

Posmatranjem rezultata možemo uočiti slične pravilnosti kao i kod signala klasične muzike. Najveći dobitak u SQNR-u u odnosu na PCM ostvaruje se za faktor kompresije  $\mu=100$  i bit-alokaciju  $R_1=12, R_2= R_3=8$  i iznosi približno 10.2 dB, dok je u slučaju podjednake raspodele bitske brzine između podopsega dobitak u odnosu na PCM iznosi samo 2.4367 dB. Možemo primetiti da bit alokacija za koju algoritam ostvaruje najviše performanse nije ista kod signala klasične i rok muzike, usled razlika u statističkim parametrima signala. Pred toga, u oba slučaja, najbolji izbor bit-alokacije nije identičan

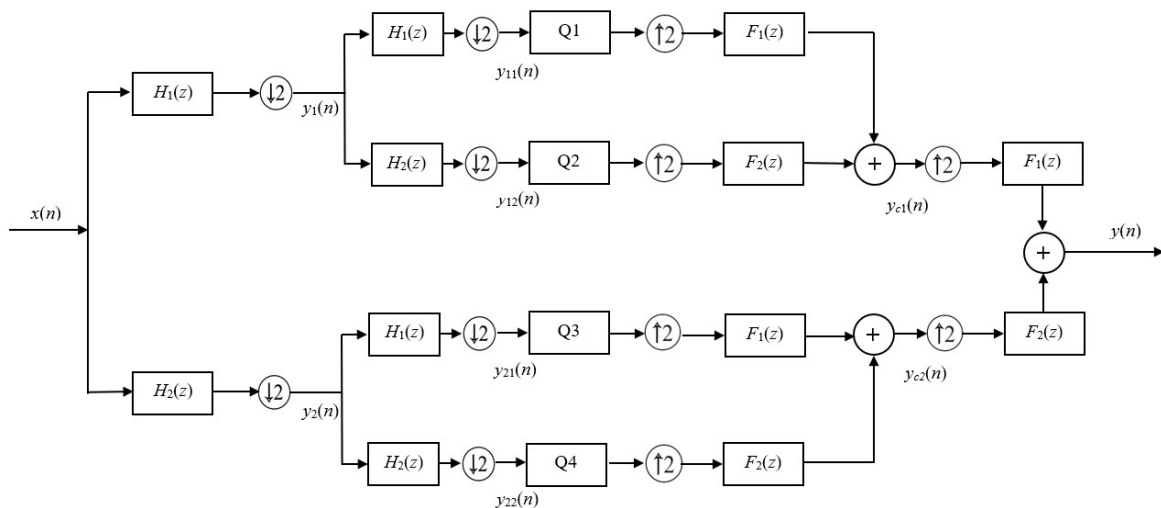
teorijski određenoj bit-alokaciji, za koju se takođe ostvaruje značajan dobitak u SQNR-u u odnosu na PCM.

## 4.5 Podopsežno kodovanje audio signala sa četiri podopsega

U prethodnom odeljku smo videli kako se sukcesivnom primenom idealnih digitalnih filtara na signal prvog frekvencijskog podopsega mogu dobiti tri podopsega i dodatno iskoristiti pogodnosti podopsežnog kodovanja. U ovom odeljku će biti predstavljen algoritam za podopsežno kodovanje audio signala koji vrši podelu ulaznog signala sa četiri frekvencijka podopsega. Algoritam koji će biti predložen u ovom odeljku primenjuje jednostavnu kvazilogaritamsku kvantizaciju sa adaptacijom unapred.

### 4.5.1 Projektovanje algoritma

Frekvencijski podopsezu su dobijeni sukcesivnom primenom idealnih digitalnih filtara, najpre na originalni signal, čime se dobijaju dva podopsega, a nakon toga i na svaki od dobijenih frekvencijskih podopsega. Blok šema algoritma je prikazana na sledećoj slici.



Slika 4.5.1.1. Blok šema algoritma za podopsežno kodovanje sa četiri podopsega

Algoritam primenjuje četiri kvazilogaritamska kvantizera sa adaptacijom unapred (Q1 – Q4) i različitim bitskim brzinama. Kako adaptacija unapred podrazumeva korišćenje odmeraka koji tek treba da se kvantuju, algoritam unosi malu količinu dodatne informacije koja povećava ukupnu bitsku brzinu u određenoj meri.

Kao i kod prethodno predloženih algoritama, ulazni signal se najpre deli na dva frekvencijska podopsega primenom idealnih digitalnih filtara:

$$y_1(n) = \frac{x(n) + x(n+1)}{2}, n = 1, 2, \dots, M, \quad (4.5.1.1)$$

$$y_2(n) = \frac{x(n) - x(n+1)}{2}, n = 1, 2, \dots, M. \quad (4.5.1.2)$$

Nakon primene prethodnih transformacija na ulazni signal, četiri podopsega korišćena u kodovanju signala se dobijaju ponavljanjem istovetnih transformacija na prvobitno dobijena dva podopsega:

$$y_{11}(n) = \frac{y_1(n) + y_1(n+1)}{2}, n = 1, 2, \dots, M, \quad (4.5.1.3)$$

$$y_{12}(n) = \frac{y_1(n) - y_1(n+1)}{2}, n = 1, 2, \dots, M, \quad (4.5.1.4)$$

$$y_{21}(n) = \frac{y_2(n) + y_2(n+1)}{2}, n = 1, 2, \dots, M, \quad (4.5.1.5)$$

$$y_{22}(n) = \frac{y_2(n) - y_2(n+1)}{2}, n = 1, 2, \dots, M, \quad (4.5.1.6)$$

Na signal svakog od dobijenih frekvencijskih podopsega se nezavisno primenjuje kvantizacija, primenom kvazilogaritamskih kvantizera, čija je kompresorska funkcija opisana u odeljku 3.3.3. Nakon primene kvantizera i slanja kvantovanog signala kroz kanal, signal je potrebno rekonstruisati na prijemu. Najpre je potrebno rekonstruisati grupe od po dva podopsega, kako bi se dobila inicijalno korišćena dva frekvencijska podopsega  $y_{c1}$  i  $y_{c2}$ .

$$y_{c1}(n) = [y_{11}(n), y_{12}(n+1)], n = 1, 2, \dots, M. \quad (4.5.1.7)$$

$$y_{c2}(n) = [y_{21}(n), y_{22}(n+1)], n = 1, 2, \dots, M. \quad (4.5.1.8)$$

Konačno rekonstruisani ulazni signal se dobija primenom istog principa kao:

$$y(n) = [y_{c1}(n), y_{c2}(n+1)], n = 1, 2, \dots, M. \quad (4.5.1.9)$$

## 4.5.2 Konstrukcija kvantizera

Algoritam za podopsežno kodovanje sa četiri podopsega primenjuje četiri nezavisna kvazilogaritamska kvantizera sa adaptacijom unapred i različitom bitskom brzinom. U postupku konstrukcije kvantizera, ulazni signal se modeluje Laplasovom raspodelom. Adaptacija se vrši primenom adaptivne amplitude maksimalnog opterećenja, na osnovu varijansi signala pojedinačnih frekvencijskih podopsega, koji tek treba da se kvantuju. Polazna osnova za određivanje adaptivne vrednosti amplitude maksimalnog opterećenja je optimalna vrednost amplitude maksimalnog opterećenja za Laplasov izvor sa jediničnom varijansom:

$$x_{\max}^{\text{opt}} = \frac{1}{\sqrt{2}} \ln \left( \frac{3\mu N^2}{\ln^2(\mu+1)} \right), \quad (4.5.2.1)$$

Nakon toga, potrebno je odrediti varijanse trenutnog,  $i$ -tog frejma odmeraka signala frekvencijskih podopsega [6]:

$$\sigma_1^{2(i)} = \frac{\sigma_x^{2(i)}}{4} (1 + \rho_1)(1 + \rho_2), \quad (4.5.2.2)$$

$$\sigma_2^{2(i)} = \frac{\sigma_x^{2(i)}}{4} (1 + \rho_1)(1 - \rho_2), \quad (4.5.2.3)$$

$$\sigma_3^{2(i)} = \frac{\sigma_x^{2(i)}}{2} (1 - \rho_1)(1 + \rho_2), \quad (4.5.2.4)$$

$$\sigma_4^{2(i)} = \frac{\sigma_x^{2(i)}}{2} (1 - \rho_1)(1 - \rho_2), \quad (4.5.2.5)$$

gde su korišćene promenljive označene i definisane kao u prethodnom odeljku. Na osnovu varijansi signala pojedinačnih frekvencijskih podopsega moguće je odrediti adaptivne vrednosti amplitude maksimalnog opterećenja za svaki frejm:

$$x_{\max 1}^{\text{FA}(i)} = x_{\max}^{\text{opt}} \sigma_1^{(i)}, i = 2, 3, \dots, P, \quad (4.5.2.6)$$

$$x_{\max 2}^{\text{FA}(i)} = x_{\max}^{\text{opt}} \sigma_2^{(i)}, i = 2, 3, \dots, P, \quad (4.5.2.7)$$

$$x_{\max 3}^{\text{FA}(i)} = x_{\max}^{\text{opt}} \sigma_3^{(i)}, i = 2, 3, \dots, P, \quad (4.5.2.8)$$

$$x_{\max 4}^{\text{FA}(i)} = x_{\max 4}^{\text{opt}} \sigma_4^{(i)}, i = 2, 3, \dots, P. \quad (4.5.2.9)$$

Korišćenjem adaptivne amplitude maksimalnog opterećenja kvantizacija se prilagođava ulaznom signalu, sa ciljem poboljšanja kvaliteta kodovanog signala. Adaptacija unapred zahteva prenos dodatne informacije, koja predstavlja varijansu signala frekvencijskih podopsega, koja je kvantovana primenom log-uniformnog kvantizera sa 6 bita po odmerku. Ovaj postupak unosi dodatnu informaciju, pa se ukupna bitska brzina računa kao:

$$R = \frac{R_1 + R_2 + R_3 + R_4}{4} + \frac{4r_v}{L_f}, \quad (4.5.2.10)$$

prilikom čega  $L_f$  označava dužinu frejma, a  $r_v$  predstavlja bitsku brzinu korišćenu za kvantovanje varijanse signala svakog od frekvencijskih podopsega. Kako je veličina frejma 320 odmeraka, a koristi se 6 ( $4 \times 6$ ) bita po odmerku za kvantovanje varijanse, možemo proračunati da dodatna informacije povećava bitsku brzinu za 0.075 bita po odmerku.

### 4.5.3 Analiza eksperimentalnih rezultata

Kao i kod analize performansi prethodnih algoritama, sproveden je eksperiment sa ciljem određivanja objektivne mere kvaliteta kodovanog signala, izraženog SQNR-om u decibelima. Dodatno, rezultati se porede sa performansama PCM kodera, definisanog standardom G.711 [7]. Ulazni signali korišćeni u eksperimentu su već analizirani signal klasične i rok muzike. Sledeća tabela prikazuje rezultate primene predloženog algoritma za podopsežno kodovanje sa četiri podopsega na signal klasične muzike, za različite bit-alokacije, dužinu frejma od 320 odmeraka i dva različita faktora kompresije.

Tabela 4.5.3.1 SQNR algoritma za podopsežno kodovanje sa četiri podopsega za različite faktore kompresije i bit-alokacije za signal klasične muzike

R	$R_{uk}$	$\mu$	SQNR <sub>FASBC4</sub> [dB]	SQNR <sub>PCM</sub> [dB], $\mu=255$
$R_1=9, R_2=9$ $R_3=9, R_4=9$	9.075	40	45.8226	41.5904
		120	44.7263	
$R_1=11, R_2=7$ $R_3=11, R_4=7$	9.075	40	50.5336	
		120	52.1851	
$R_1=12, R_2=8$ $R_3=9, R_4=7$	9.075	40	55.0164	
		120	57.0002	
$R_1=14, R_2=8$ $R_3=7, R_4=7$	9.075	40	53.0315	
		120	55.0074	

Posmatranjem rezultata u prethodnoj tabeli vidimo da se najbolji rezultat postiže za bit-alokaciju  $R_1=12, R_2=8, R_3=9, R_4=7$  i faktor kompresije  $\mu=120$ , za koje SQNR iznosi približno 57 dB. U ovom slučaju se ostvaruje SQNR viši za približno 15.41dB u odnosu na PCM koder. Bit alokacija za koju se ostvaruje maksimalni SQNR potvrđuje pretpostavku da se u prvom frekvencijskom podopesgu nalazi najveća količina informacija, zbog čega je taj frekvencijski podopseg kvantovan sa najvišom bitskom brzinom.

Tabela 4.5.3.2 SQNR algoritma za podopsežno kodovanje sa četiri podopsega za različite bit-alokacije i faktore kompresije za signal rok muzike

R	$R_{uk}$	$\mu$	SQNR <sub>FASBC4</sub> [dB]	SQNR <sub>PCM</sub> [dB], $\mu=255$
$R_1=9, R_2=9$ $R_3=9, R_4=9$	9.075	40	44.0726	39.7597
		120	43.2901	
$R_1=11, R_2=7$ $R_3=11, R_4=7$	9.075	40	46.4577	
		120	47.5904	
$R_1=12, R_2=8$ $R_3=9, R_4=7$	9.075	40	50.8267	
		120	51.8067	
$R_1=14, R_2=8$ $R_3=7, R_4=7$	9.075	40	48.4754	
		120	49.0619	

Rezultati primene predloženog algoritma na signal rok muzike pokazuje da se za istu bit-alokaciju kao i kod signala klasične muzike ostvaruje maksimalni SQNR. U tom

slučaju dobitak u odnosu na PCM koder iznosi približno 12 dB, što potvrđuje da se predloženim algoritmom ostvaruje visok kvalitet kodovanog signala.

Algoritam predložen u ovom odeljku pokazuje kako se uvođenjem dodatne transformacija može ostvariti visok kvalitet kodovanog signala, uz primenu jednostavnog kvazilogaritamskog kvantizera sa adaptacijom unapred.

#### 4.6 Poređenje rezultata predloženih algoritama za podopsežno kodovanje govornog i audio signala

Nakon predstavljanja tri algoritma za kodovanje govornog i audio signala zasnovanih na podopsežnom kodovanju i logaritamskoj kvantizaciji sa  $\mu$  zakonom kompresije možemo uporediti njihove performanse primenom na isti ulazni signal. Predstavljena su četiri algoritma u odeljcima 4.2, 4.3, 4.4 i 4.5, gde je svaki sledeći algoritam za nijansu složeniji od prethodnog, dok posmatrano pojedinačno, svaki ostvaruje visok kvalitet kodovanog signala. Uporedni prikaz rezultata će biti predstavljen za identičnu bitsku brzinu od 9 bita po odmerku, dok će biti uzeti najviša vrednost SQNR-a za svaki algoritam, ostvarena za različite vrednosti faktora kompresije i različite bit-alokacije. Na ovaj način dobijamo fer poređenje, kako različiti algoritmi i tehnike kodovanja ostvaruju najbolje performanse sa različitim faktorom kompresije i bit-alokacijom.

Tabela 4.6.1 SQNR algoritma za podopsežno kodovanje sa različitim brojem podopsega za signal klasične muzike

R	$\mu$	SQNR <sub>SBC</sub> [dB]	$\mu$	SQNR <sub>AUSBC2</sub> [dB]	$\mu$	SQNR <sub>AUSBC3</sub> [dB]	$\mu$	SQNR <sub>FASBC4</sub> [dB]
9	511	47.2528	5	55.2136	100	58.0180	120	57.0002

Posmatranjem Tabele uočavamo blagi rast SQNR-a sa povećanjem složenosti algoritma. Ipak, najbolje performanse postignute su primenom algoritma za podopsežno kodovanje sa tri podopsega i kvazilogaritamskim kvantizerima sa adaptacijom unazad, dok očekivano, najniži SQNR ostvarujemo primenom algoritma za podopsežno kodovanje sa dva podopsega i kvazilogaritamskim kvantizerima bez adaptacije. Treba istaći da se primenom PCM-a na isti ulazni signal ostvaruje SQNR od 41.5904 dB, za faktor kompresije  $\mu=255$ . Na taj način primenom predloženih algoritama za podopsežno

kodovanje možemo ostvariti dobitak u SQNR-u od 5.6624 do 16.4276 dB u odnosu na PCM. Možemo videti da broj podopsega ne mora nužno doneti povećanje SQNR-a, pa se algoritmom za podopsežno kodovanje sa četiri podopsega ostvaruje maksimalni SQNR za 1 dB niži u odnosu na algoritam sa tri podopsega. Ovo se može pripisati efikasnijoj bit-alokaciji ostvarenoj prilikom podele ulaznog signala na tri grane, koja u većoj meri odgovara prirodi ulaznog signala.

Na osnovu opisa algoritama i eksperimentalnih rezultata možemo zaključiti da se u disertaciji predlažu novi algoritmi za kodovanje govornog i audio signala zasnovani na tehnikama podopsežnog kodovanja. Algoritmi su okarakterisani niskom kompleksnošću i primenom relativno jednostavnih tehnika kvantizacije, dok se istovremeno njihovom primenom dobija kodovani signal visokog kvaliteta. Analizom eksperimentalno dobijenih rezultata primene predloženih algoritama na realne signale možemo zaključiti da se predloženi algoritmi mogu uspešno primeniti u kodovanju govornog i audio signala.

## **5. KODOVANJE GOVORNOG I AUDIO SIGNALA PRIMENOM MODIFIKOVANOG BTC ALGORITMA I DIFERENCIJALNOG KODOVANJA**

U ovom poglavlju će biti predstavljen algoritam za kodovanje govornog i audio signala zasnovan na Block Truncation Coding (BTC) algoritmu, koji je originalno namenjen kodovanju crno-bele slike. Svedoci smo čestih primena sličnih ili istih tehnika kodovanja na različite tipove signala, koji se u nekim slučajevima mogu i zajedno prenositi za potrebe određene multimedijalne aplikacije. Ovo je motivisalo primenu algoritma za kodovanje slike na audio signal.

Kako crno-bela slika i audio signal imaju različite karakteristike, potrebno je primeniti određene modifikacije algoritma i korišćenih tehnika kodovanja, kako bi se algoritam mogao uspešno primeniti u kodovanju govornog i audio signala. Uz primenu značajnih modifikacija, zadržana je osnovna ideja originalnog algoritma, kao i osnovni principi na kojima je originalni algoritam zasnovan. Nakon kratkog opisa originalnog BTC algoritma, biće predstavljen modifikovani algoritam, za koji će se analizirati eksperimentalni rezultati primene u kodovanju govornog i audio signala.



## 5.1 Originalni BTC algoritam

BTC algoritam je predstavljen 1977 od strane Delpa i Mičela [15], kao algoritam za kompresiju crno-bele slike. Algoritam je doživeo veliki uspeh i ostvario je mnogobrojne primene, među kojima su i poznati i izuzetno rasprostranjeni standardi za kodovanje slike JPG i JPG2000.

BTC algoritam se zasniva na dekompoziciji signala u nepreklapajuće blokove odmeraka (frejmove) koji su dvodimenzionalni, odnosno dimenzija  $n \times n$ , prilikom čega je tipična vrednost  $n=4$ . Osnovna verzija algoritma primenjuje kvantizaciju sa dva nivoa, što je moguće uraditi kako vrednost odmerka crno-bele slike može da uzima vrednost u opsegu od 0 do 255. Ideja BTC algoritma je da se kod bloka kompresovanog signala očuvaju identične vrednosti standardne devijacije i srednje vrednosti sa originalnim blokom ulaznog signala. Osnovni BTC algoritam se sastoji iz sledećih koraka [15]:

1. Slika se najpre deli na nepreklapajuće blokove dimenzija  $m \times n$ .
2. Određuju se statistički parametri blokova, srednja vrednost ( $\bar{x}$ ) i varijansa ( $\sigma$ ), koji predstavljaju sporednu informaciju o bloku.
3. Prag odluke kvantizera predstavlja srednja vrednost, sa kojom se upoređuje vrednost svakog piksela, odnosno odmerka. Vrednosti veće ili jednake  $\bar{x}$  se predstavljaju jedinicom, dok se manje vrednosti predstavljaju nulom. Time se blokovi predstavljaju koristeći dva bita po pikselu.
4. Vrš se prebrojavanje piksela čija je vrednost viša od praga i rezultat se sumira u promenljivu  $q$ .
5. U ovom trenutku, blok je opisan srednjom vrednošću, varijansom i dodeljenim jedinicama i nulama umesto originalne vrednosti odmerka.
6. Blok se na strani dekodera rekonstruiše tako što se „1“ zamenjuje sa  $a$ , a „0“ sa  $b$ , koji su određeni kao:

$$a = \bar{x} - \bar{\sigma} \sqrt{\left[ \frac{q}{m-q} \right]}, \quad (5.1.1)$$

$$b = \bar{x} - \bar{\sigma} \sqrt{\left[ \frac{m-q}{q} \right]}, \quad (5.1.2)$$

gde  $m$  predstavlja broj odmeraka u bloku ( $m=n^2$ )

Dakle, možemo uočiti da BTC algoritam vrši kompresiju signala time što ne prenosi originalne vrednosti odmeraka, već samo dva različita reprezentata („0“ i „1“), kao i statističke parametre koje dodatno opisuju blok. Kako se osnovna verzija algoritma pokazala efikasnom a pri tome je i okarakterisana izuzetno niskom kompleksnošću primene, doživela je široku praktičnu primenu. Pored toga ovakav algoritam poseduje veliki potencijal za nadgradnju, što je motivisalo istraživače na razvoj novih modifikovanih verzija osnovnog algoritma. Predstavljen je veliki broj algoritama za kodovanje slike zasnovanih na BTC algoritmu [55]-[61], koji primenjuju različite tehnike kvantizacije i tehnike kodovanja. Takođe, principi BTC algoritma primenjeni su i u kodovanju drugih tipova ulaznih signala, kao što je slučaj sa algoritmom za kodovanje satelitske slike [62] i slike u boji [63], [64].

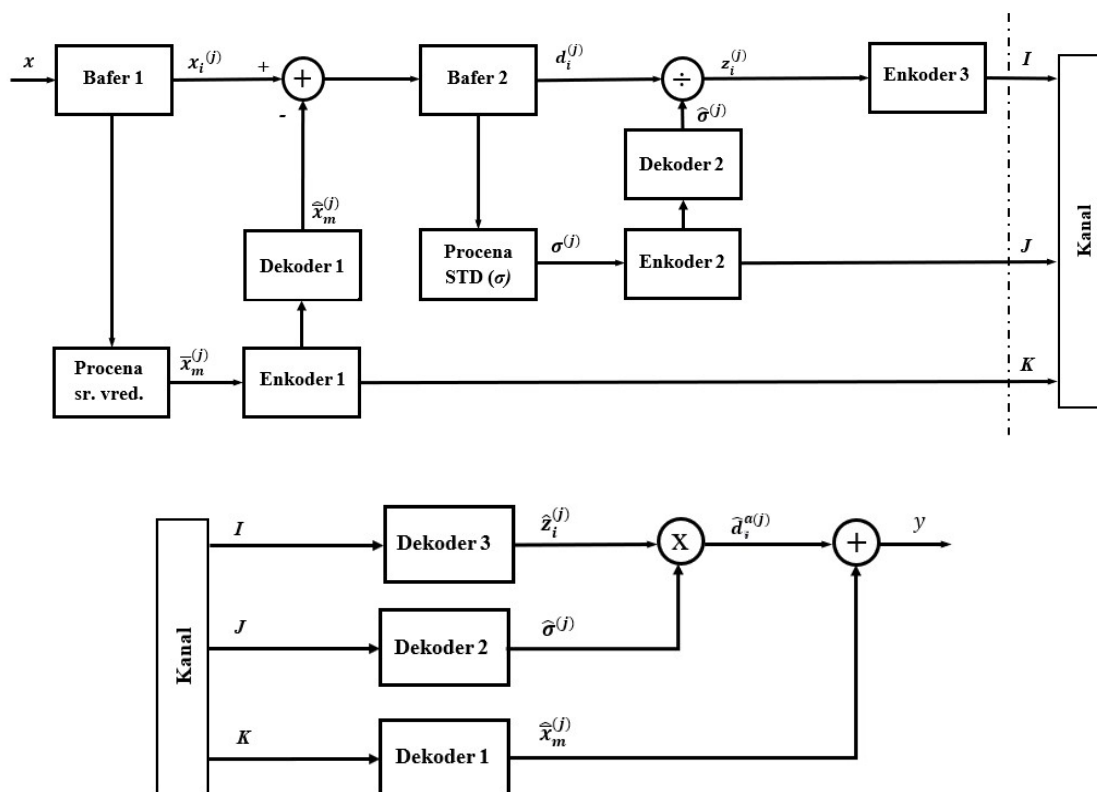
Široka primena i različiti slučajevi upotrebe BTC algoritma u kodovanju slike, uz dosta varijacija algoritma, ali i činjenica da u modernim algoritmima primena sličnih tehnika kodovanja na različite tipove signala nije strana, motivisali su primenu principa BTC algoritma u kodovanju govornog i audio signala. Kako se vrši primena algoritma na ulazni signal različitih karakteristika, potrebno je izvršiti određene modifikacije, koje će biti opisane u nastavku.

## **5.2 Modifikovani BCT algoritam za kodovanje audio signala**

Kao što je već pomenuto, originalni BTC algoritam je razvijen za kompresiju crno-bele slike i neophodno je primeniti određene modifikacije kako bi se mogao primeniti u kodovanju audio signala. Algoritam je zasnovan na podeli ulaznog signala na blokove (frejmove), koji se obrađuju pojedinačno. Za svaki blok se izračunava srednja vrednost i standardna devijacija, koje predstavljaju sporednu informaciju, koja se u originalnom algoritmu koristi u postupku dekodovanja kvantovanih odmeraka ulaznog signala. To nam daje ukupno tri promenljive sa kojima algoritam radi, ako odmerke ulaznog signala posmatramo kao jednu promenljivu, uz dva prethodno navedena statistička parametra koji dodatno opisuju blok podataka. Ove tri promenljive se koduju nezavisno, što pruža prostor za potencijalnu nadgradnju algoritma. Originalni BTC algoritam primenjuje tri uniformna kvantizera, sa bitskom brzinom od 2 bita po odmerku. Ovakva realizacija algoritma sa primenom dvo-nivovske kvantizacije je moguća zbog samog dizajna

algoritma i karakteristika crno-bele slike. Očigledno je da bitska brzina od dva bita po odmerku neće biti dovoljna za kodovanje audio signala, pa je prva modifikacija korišćenje više bitske brzine. Pored toga, kako se odmerci ulaznog signala i njegovi statistički parametri koduju nezavisno, moguće je primeniti različite tehnike kodovanja i bitsku brzinu na svaki od promenljivih.

### 5.2.1 Projektovanje modifikovanog BTC algoritma za kodovanje audio signala



Slika 5.2.1.1. Blok šema enkodera (iznad) i dekodera (ispod) modifikovanog BTC algoritma za kodovanje audio signala

Prva promena koju modifikovani BTC algoritam primenjuje korišćenje dužih frejmova odmeraka ulaznog signala od 40 do 320 odmeraka, za razliku od originalnog BTC algoritma koji koristi blokove od 16 (4 x 4) odmerka. Možemo primetiti da se kod slike koriste blokovi podataka, dok se kod audio signala koriste jednodimenzijalni nizovi ulaznih odmeraka. Korišćenje dužih frejmova nosi niz pogodnosti, od kojih je najočiglednija smanjenje ukupnog broja frejmova, a samim tim i broja izračunavanja

srednje vrednosti i standardne devijacije koje je potrebno proračunati za svaki frejm. Pored toga, korišćenjem dužih frejmova se smanjuje količina dodatne informacije koja se generiše po odmerku, jer korišćenjem dužih frejmova jedna srednja vrednost opisuje veći broj odmeraka ulaznog signala, a isto važi i za standardnu devijaciju. Nedostatak korišćenja statističkih parametara frejma u postupku dekodovanja je samo uvođenje dodatne informacija koja povećava bitsku brzinu usled prenosa statističkih parametara do prijemnika. Korišćenje dužih frejmova umanjuje uticaj ove pojave, pogotovo za slučaj frejmova dužih od 160 odmeraka.

Pored razlike u dimenzijama frejma, modifikovani algoritam ne vrši kvantizaciju ulaznih odmeraka, već signala koji predstavlja razliku amplitude ulaznog odmerka i srednje vrednosti frejma. Na ovaj način se smanjuje amplitudska dinamika ulaznog signala i ovakav signal je pogodniji za primenu tehnika kvantizacije.

Značajne modifikacije su primenjene u strukturi algoritma i oblasti kvantizacije, obzirom da bi primena uniformnih kvantizera sa dva bita po odmerku na audio signal rezultirala niskim kvalitetom kodovanog signala. Blok šema modifikovanog BTC algoritma je prikazana na Slici 5.2.1.1. Konstruisana su tri različita kvantizera zadužena za kvantovanje različitih promenljivih. Srednja vrednost frejma ulaznih odmeraka se kvantuje pomoću logaritamskog komandora definisanog  $\mu$  zakonom kompresije (na slici označen kao Enkoder 1 i Dekoder 1), dok se kvantizacija standardne devijacije frejma ulaznog signala vrši primenom log-uniformnog kvantizera [49] (Enkoder 2 i Dekoder 2). Razlika između amplitude ulaznog signala i kvantovane srednje vrednosti frejma ulaznih odmeraka se kvantuje korišćenjem optimalnog komandora sa adaptacijom unapred [65], [66] (Enkoder 3 i Dekoder 3). Detaljan opis izvršenja algoritma i postupka kvantizacije biće predstavljen u nastavku.

Modifikovani BTC algoritam prikazan na Slici 5.2.1.1 funkcioniše na sledeći način. Ulazni signal  $x = x_i^{(j)}$ ,  $i = 1, 2, \dots, M$ ,  $j = 1, 2, \dots, L$  je podeljen na  $L$  frejma od kojih svaki sadrži po  $M$  odmeraka. Odmerci se dovode na Bafer 1, koji učitava jedan po jedan frejm. Zatim se za svaki frejm procenjuje srednja vrednost odmerka ulaznog signala  $\bar{x}_m^{(j)}$ , i kvantuje primenom komandora sa  $\mu$  zakonom kompresije (Enkoder 1 i Dekoder 1). Kvantovana srednja vrednost se koristi za formiranje signala razlike, oduzimanjem od

trenutne amplitude ulaznih odmeraka trenutnog frejma  $x_i^{(j)}$ . Signal razlike se formira za svaki frejm kao:

$$d_i^{(j)} = x_i^{(j)} - \hat{x}_m^{(j)}, i = 1, 2, \dots, M, j = 1, 2, \dots, L, \quad (5.2.1.1)$$

gde  $M$  označava veličinu frejma dok  $L$  označava ukupan broj frejmova. Nakon ovog koraka, signal razlike se dalje prenosi i praktično zamenjuje ulazni signal u daljem toku algoritma. Činjenica da signal razlike signala ima manju amplitudsku dinamiku od prvobitnog ulaznog signala čini ga pogodnijim za kodovanje. Frejm signala razlike se najpre dovodi na ulaz Bafera 2, nakon čega se procenjuje standardna devijacija frejma koja se koristi za adaptaciju kvantizera koji se koristi za kvantovanje signala razlike. Standardna devijacija frejma signala razlike se kvantuje pomoću log-uniformnog kvantizera, tj. kombinacije Kodera 2 i Dekodera 2 prikazanog na slici.

Kao što je pomenuto, frejm signala razlike  $d_i^{(j)}$  se kvantuje primenom optimalnog komandora sa adaptacijom unapred. Adaptacija se može izvesti u dve faze [1], [3], [4], [67]. Najpre se određuju fiksni reprezentanti, koji su rezultat fiksnog (ne-adaptivnog) optimalnog komandora. Adaptivni reprezentanti, se zatim dobijaju množenjem fiksnih reprezentata kvantovanom vrednošću standardne devijacije frejma signala razlike:

$$\hat{d}_i^a{}^{(j)} = \hat{d}_i^{(j)} \times \hat{\sigma}^{(j)}, i = 1, 2, \dots, M, j = 1, 2, \dots, L, \quad (5.2.1.2)$$

gde  $\hat{\sigma}^{(j)}$  predstavlja kvantovanu vrednost standardne devijacije trenutnog  $j$ -tog frejma signala razlike. Adaptacija unapred se takođe može realizovati normalizacijom signala razlike vrednošću  $\hat{\sigma}^{(j)}$ , nakon čega se normalizovani odmerci  $z_i^{(j)}$ , mogu kvantovati primenom fiksnog optimalnog komandora (Enkoder 1 i Dekoder 1), tipično projektovanog za jediničnu varijansu.

Izlazi ovako definisanih enkodera se šalju kroz kanal kao binarne informacije, označene sa  $I, J$  i  $K$ . Proces dekodovanja započinje sporednom informacijom, nakon čijeg se dekodovanja vrši proces denormalizacije, kako bi se dobila kvantovana vrednost odmerka signala razlike  $\hat{d}_i^a{}^{(j)}$ . Kvantovani odmerci ulaznog signala  $y = y_i^{(j)}, i = 1, 2, \dots, M, j = 1, 2, \dots, L$  se rekonstruišu sabiranjem dekodovane vrednosti adaptivnih reprezentata i dekodovane srednje vrednosti trenutnog frejma:

$$y_i^{(j)} = \hat{d}_i^{a(j)} + \hat{x}_m^{(j)}, i = 1, 2, \dots, M, j = 1, 2, \dots, L, \quad (5.2.1.3)$$

Ovim se završava postupak kvantizacije odmeraka ulaznog signala, nakon čega je moguće analizirati performanse algoritma. Performanse će biti analizirane kao i do sada, posmatranjem SQNR-a, kao i poređenjem sa tehnikama slične složenosti.

Kako algoritam koristi adaptaciju unapred i vrši prenos sporedne informacije u vidu srednje vrednosti i standardne devijacije, povećava se bitska brzina u odnosu na bitsku brzinu koja se koristi za kvantovanje samih odmeraka ulaznog signala. Ukupna bitska brzina koji modifikovani BTC algoritam koristi se može odrediti kao:

$$R = r_3 + \frac{r_1 + r_2}{M}, \quad (5.2.1.4)$$

gde  $r_3$  označava bitsku brzinu koja se koristi za fiksno kodovanje signala razlike,  $r_1$  predstavlja bitsku brzinu korišćenu za kodovanje srednje vrednosti odmerka ulaznog signala,  $r_2$  označava bitsku brzinu koja se koristi za kodovanje standardne devijacije signala razlike, dok  $M$  predstavlja veličinu frejma.

## 5.2.2 Projektovanje kvantizera

Dakle, modifikovani BTC algoritam koristi tri različita kvantizera za kvantovanje različitih korišćenih parametara. Priroda audio signala nalaže da se u algoritmu primenjuju robusne tehnike kodovanja, koje imaju cilj da odgovore što uravnoteženijim performansama za ulazne signale različitih karakteristika. Prilikom projektovanja kvantizera ulazni audio signal je modelovan Laplasovom raspodelom jedinične varijanse.

Srednja vrednost frejma je kvantovana primenom kvazilogaritamskog kvantizera, koji primenjuje kompresorsku funkciju definisanu  $\mu$  zakonom kompresije, kao što je definisano u odeljku 3.3.4. Projektovan je kvantizer sa 64 reprezentacionih, što znači da se za kodovanje srednje vrednosti koristi 6 bita po odmerku. Kako imamo samo jednu srednju vrednost po frejmu, zaključujemo da je dodatna informacija u ovom slučaju 6 bita po frejmu. Kvantizer je projektovan za faktor kompresije  $\mu=255$ .

Varijansa frejma signala razlike se kvantuje primenom log-uniformnog kvantizera, projektovanog za  $N_g$  kvantizacionih nivoa [49]:

$$20 \log(\hat{\sigma}^{(j)} = \hat{\sigma}_k^{(j)}) = 20 \log_{10}(\sigma_{\min}) + \frac{2k-1}{2} \Delta^{\text{lu}}, \quad (5.2.2.1)$$

gde je  $k=1, \dots, N_g$ , a širina kvanta je definisana kao:

$$\Delta^{\text{lu}} = \frac{20 \log_{10} \left( \frac{\sigma_{\max}}{\sigma_{\min}} \right)}{N_g}, \quad (5.2.2.2)$$

Log-uniformni kvantizer vrši uniformnu raspodelu opsega standardne devijacije ( $B$ ) u dB  $[20 \log(\sigma_{\min}), 20 \log(\sigma_{\max})]$  u  $N_g$  amplitudskih kvantata širine  $\Delta^{\text{lu}}$ :

$$\hat{\sigma}^{(j)} = \hat{\sigma}_k^{(j)} \mid 20 \log_{10}(\sigma^{(j)}) \in [V_k^{\text{l}} [\text{dB}], V_k^{\text{h}} [\text{dB}]], \quad (5.2.2.3)$$

$$V_k [\text{dB}] = 20 \log_{10}(\hat{\sigma}_k^{(j)}) = (V_k^{\text{l}} [\text{dB}] + V_k^{\text{h}} [\text{dB}]) / 2, \quad (5.2.2.4)$$

prilikom čega je  $k \in \{1, 2, \dots, N_g\}$ , dok se kvant kome trenutna standardna devijacija pripada određuje kao:

$$V_k^{\text{l}} [\text{dB}] = 20 \log_{10}(\sigma_k^{\text{l}}) = V_1^{\text{l}} [\text{dB}] + (k-1) |B| / N_g, \quad (5.2.2.5)$$

$$V_k^{\text{h}} [\text{dB}] = 20 \log_{10}(\sigma_k^{\text{h}}) = V_1^{\text{l}} [\text{dB}] + k |B| / N_g, \quad (5.2.2.6)$$

$$B = \bigcup_{k=1}^{N_g} [V_k^{\text{l}} [\text{dB}], V_k^{\text{h}} [\text{dB}]], \quad (5.2.2.7)$$

Za potrebe kvantizacije standardne devijacije frejma signala razlike projektovan je log-uniformni kvantizer sa  $N_g = 16$  kvantizacionih nivoa, pa se za prenos standardne devijacije dodaje 4 bita po frejmu ulaznog signala.

Treći kvantizer koji se primenjuje u modifikovanom BTC algoritmu je optimalni komparator sa adaptacijom unapred, koji se koristi za kvantovanje signala razlike. Ovaj kvantizer je zasnovan na kompadding tehnici, opisanoj u odeljku 3.3. Kvantizer primenjuje modifikovanu kompresorsku funkciju, koja vrši kopiranje opsega  $(-\infty, \infty)$  u opseg  $[-1, 1]$  i definisana je kao:

$$c_{\text{mod}}(x) = \begin{cases} -1, & x < -t_{\text{max}} \\ -1 + 2 \frac{\int_{-t_{\text{max}}}^x p^{1/3}(t) dt}{\int_{-t_{\text{max}}}^{t_{\text{max}}} p^{1/3}(t) dt}, & x \in [-t_{\text{max}}, t_{\text{max}}], \\ 1, & x > t_{\text{max}} \end{cases} \quad (5.2.2.8)$$

Kako se u postupku kvantizacije ulazni signal modeluje Laplasovom raspodelom, zamenom izraza definisanog u (3.1) u prethodni izraz, dobija se izraz za modifikovanu kompresorsku funkciju za Laplasovu raspodelu:

$$c_{\text{mod}}(x) = \begin{cases} -1, & x < -t_{\text{max}} \\ -1 + 2 \frac{\exp\left\{\frac{\sqrt{2}}{3}t_{\text{max}}\right\} - \exp\left\{-\frac{\sqrt{2}}{3}x\right\}}{\exp\left\{\frac{\sqrt{2}}{3}t_{\text{max}}\right\} - \exp\left\{-\frac{\sqrt{2}}{3}t_{\text{max}}\right\}}, & x \in [-t_{\text{max}}, t_{\text{max}}], \\ 1, & x > t_{\text{max}} \end{cases} \quad (5.2.2.9)$$

Kao što je prethodno opisano, nakon primene kompresorske funkcije primenjuje se uniformna kvantizacija. Granice odlučivanja i reprezentacioni nivoi konstruisanog uniformnog kvantizera se mogu odrediti na sledeći način:

$$t_{u,m} = -1 + \frac{2m}{N_{oc}}, m = 0, 1, \dots, N, \quad (5.2.2.10)$$

$$y_{u,m} = -1 + \frac{2m-1}{N_{oc}}, m = 1, 2, \dots, N, \quad (5.2.2.11)$$

gde  $N_{oc}$  označava broj reprezentacionih nivoa optimalnog komandora. Primenom prethodna dva izraza, granice odlučivanja i reprezentacioni nivoi optimalnog komandora se mogu odrediti kao rešenja sledećih jednačina:

$$t_m = c^{-1}(t_{u,m}), m = 0, 1, \dots, N, \quad (5.2.2.12)$$

$$y_m = c^{-1}(y_{u,m}), m = 1, 2, \dots, N, \quad (5.2.2.13)$$

gde usled simetrije kvantizera u odnosu na nulu važi da je  $t_m = t_{N-m}$ ,  $m = 1, 2, \dots, N/2 - 1$ ,  $y_m = y_{N-m+1}$ ,  $m = 1, 2, \dots, N/2$ . Simetrija se očekuje kako je funkcija gustine verovatnoće



Laplasove raspodele simetrična u odnosu na nulu. Zamenom izraza (5.2.2.10) i (5.2.2.11) u izraze (5.2.2.12) i (5.2.2.13) dobijaju se izrazi za granice odlučivanja i reprezentacioni nivoi optimalnog komandora kao:

$$t_m = \frac{3}{\sqrt{2}} \ln \left( \frac{N}{2N - 2m - (2m - N) \exp \left\{ -\frac{\sqrt{2}}{3} t_{\max} \right\}} \right), \quad (5.2.2.14)$$

$$y_m = \frac{3}{\sqrt{2}} \ln \left( \frac{N}{2N - 2m + 1 + (2m - 1 - N) \exp \left\{ -\frac{\sqrt{2}}{3} t_{\max} \right\}} \right), \quad (5.2.2.15)$$

gde je  $m$  definisano kao  $\frac{N}{2} \leq m \leq N$  dok je amplituda maksimalnog opterećenja

kvantizera,  $t_{\max}$ , definisana kao:

$$t_{\max} = \frac{3}{\sqrt{2}} \ln(N + 1). \quad (5.2.2.16)$$

U ovom trenutku su određeni fiksni, odnosno neadaptivni reprezentacioni nivoi i granice odlučivanja, koji se nakon toga adaptiraju na standardnu devijaciju frejma. Adaptacija se praktično vrši množenjem kvantovane vrednosti standardne devijacije frejma sa neadaptivnim vrednostima reprezentacionih nivoa i granica odlučivanja:

$$y_m^a = \hat{\sigma}^{(j)} y_m. \quad (5.2.2.17)$$

$$t_m^a = \hat{\sigma}^{(j)} t_m. \quad (5.2.2.18)$$

Nakon primene opisane tehnike adaptivne kvantizacije, izlazni odmerak se dobija sabiranjem kvantovane vrednosti odmerka frejma signala razlike i kvantovane srednje vrednosti tog frejma. Postupak se ponavlja za svaki frejm ulaznog signala.

### 5.2.3 Analiza eksperimentalnih rezultata

Eksperimentalni rezultati su dobijeni primenom modifikovanog BTC algoritma na realne audio signale, uzorke klasične i rok muzike, čiji su statistički parametri prikazani

u Tabeli 4.3.3.1. Pored toga, kao i u prethodnim odeljcima, rezultati se porede sa algoritmima slične kompleksnosti. Kao što je opisano, predloženi modifikovani algoritam implementira robusne tehnike kvantizacije, kako bi odgovorio prirodi audio signala. Prvi eksperiment ima zadatak da ispita tu činjenicu poređenjem performansi dve varijante modifikovanog BTC algoritma, BTC1U i BTC3U.

BTC3U predstavlja modifikovani BTC algoritam koji koristi 3 uniformna kvantizera, što je sličnije originalnom algoritmu. Kako nema smisla probati kodovanje audio signala sa dva bita po odmerku, korišćene su više bitske brzine i to: 8 bita po odmerku za kodovanje frejma signala razlike, 6 bita po frejmu za kvantovanje srednje vrednosti i 4 bita po frejmu za kvantovanje standardne devijacije. Dodatno, kvantovanje signala razlike se vrši adaptivnim uniformnim kvantizerom, kako bi se ostvarile bolje performanse.

BTC1U predstavlja algoritam sličan predloženom modifikovanom BTC algoritmu, koji se od njega razlikuje primenom adaptivne uniformne kvantizacije za kvantovanje frejma signala razlike, dok se ostali parametri kvantuju identično kao u predloženom modifikovanom BTC algoritmu. Srednja vrednost frejma ulaznog signala se kvantuje primenom kvazilogaritamskom kvantizera, dok se standardna devijacija signala razlike kvantuje log-uniformnim kvantizerom, uz bitsku brzinu od 6 i 4 bita po odmerku respektivno. Tabela prikazuje ukupnu bitsku brzinu koja zavisi od dužine frejma, kao što je opisano u prethodnim odeljcima.

Amplituda maksimalnog opterećenja projektovanih uniformnih kvantizera je određena kao:

$$x_{\max u} = \sqrt{2} \ln(N_u) - \sqrt{2} \ln\left(\frac{2}{3} \ln(N_u)\right). \quad (5.2.3.1)$$

Tabela 5.2.3.1. SQNR za različite dužine frejma i bitske brzine dobijen za signal klasične muzike primenom varijacija BTC algoritma

Dužina frejma [odmeraka]	R [bit/odmerak]	SQNR [dB] klasična muzika		SQNR [dB] rok muzika	
		BTC3U	BTC1U	BTC3U	BTC1U
40	8.25	28.8709	34.307	34.886	38.5205
80	8.25	28.8573	34.145	34.875	37.8494
120	8.125	28.8355	33.9998	34.865	37.5655
160	8.083	28.8303	33.9024	34.855	37.2777
200	8.062	28.8173	33.8242	34.834	37.045
240	8.05	28.8221	33.7667	34.839	36.8659
280	8.042	28.8162	33.783	34.819	36.6875
320	8.036	28.8274	33.7814	34.789	36.5425

Posmatranjem Tabele 5.2.3.1 se može primetiti da, za istu ukupnu bitsku brzinu, BTC1U ostvaruje dobitak u SQNR-u u odnosu na BTC3U od oko 5 dB do 5.4 dB za signal klasične muzike i od oko 1.7 dB do 3.6 dB za signal rok muzike, u zavisnosti od veličine frejma. Ovaj dobitak je direktno rezultat primene robusnijih kvantizera u kvantizaciji sporedne informacije. Kasnijom analizom eksperimentalnih rezultata primene predloženog modifikovanog BTC algoritma sa verzijama koje primenjuju uniformnu kvantizaciju se može još efikasnije utvrditi uticaj primene robusnih tehnika kvantizacije. Takođe, naredni eksperiment ima za cilj da pokaže i potencijalni dobitak u SQNR-u u odnosu na PCM tehniku, kao što je bio slučaj u prethodnim eksperimentima.

Naredna tabela prikazuje eksperimentalne rezultate primene predloženog BTC algoritma na signal klasične muzike. Pored toga, prikazane su i performanse optimalnog komandora sa adaptacijom unapred, kao i PCM tehnike za slučaj primene na isti ulazni signal. Rezultati su upareni sa tačnom vrednošću ukupne bitske brzine kako bi se ostvarilo fer poređenje. Bitska brzina optimalnog komandora sa adaptacijom unapred je određena primenom izraza (5.2.1.4), prilikom čega je  $r_1=0$ , dok je bitska brzina PCM-a konstanta i iznosi 8 bita po odmerku. Možemo primetiti u tabeli da je bitska brzina modifikovanog BTC algoritma i optimalnog komandora viša od 8 bita po odmerku i upravo ta vrednost za koju je bitska brzina viša od 8 predstavlja dodatak usled korišćenja dodatne informacije i adaptacije unapred.

Tabela 5.2.3.2. SQNR za različite dužine frejma i bitske brzine dobijen za signal klasične muzike primenom varijacija BTC algoritma

Dužina frejma [odmeraka]	SQNR <sub>BTC</sub> [dB], R <sub>BTC</sub> [bit/odmerak]	SQNR <sub>FAOC</sub> [dB], R <sub>FAOC</sub> [bit/ odmerak]	SQNR <sub>PCM</sub> [dB], R <sub>PCM</sub> [bit/ odmerak]
40	43.7468, 8.25	41.8704, 8.1	35.5672, 8
80	42.8037, 8.125	41.8118, 8.05	35.5672, 8
120	42.223, 8.0833	41.721, 8.0333	35.5672, 8
160	41.926, 8.0625	41.6449, 8.025	35.5672, 8
200	41.6867, 8.05	41.5733, 8.02	35.5672, 8
240	41.553, 8.0417	41.5297, 8.0167	35.5672, 8
280	41.524, 8.036	41.5258, 8.0143	35.5672, 8
320	41.504, 8.031	41.5329, 8.0125	35.5672, 8

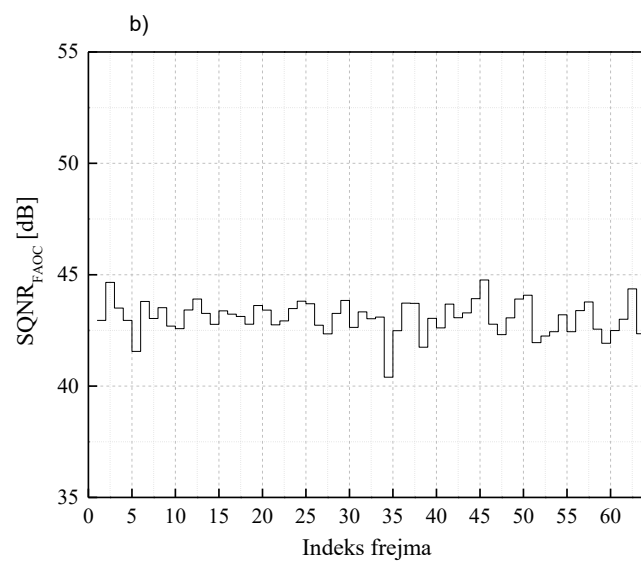
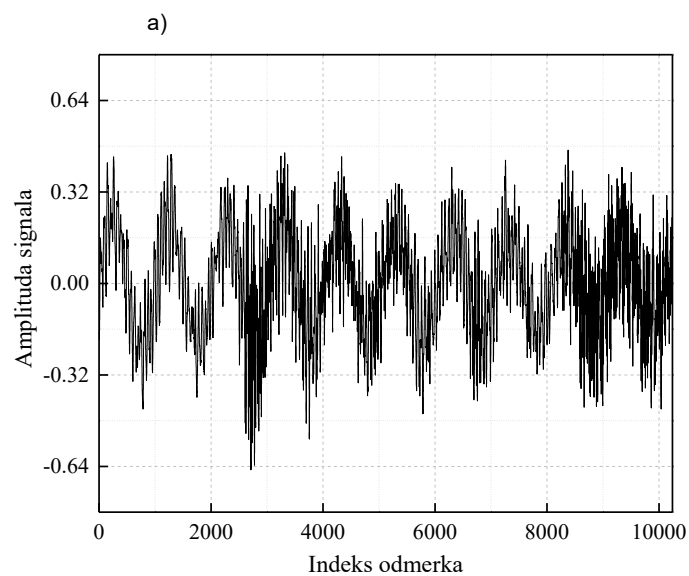
Iz tabele možemo uočiti da se primenom predloženog modifikovanog BTC algoritma ostvaruje značajan dobitak u SQNR-u, u poređenju sa PCM-om za sve razmatrane veličine frejma, kao i u poređenju sa FAOC-om za slučaj korišćenja manjih veličina frejma. Predloženi modifikovani BTC algoritam ostvaruje dobitak od oko 8.1 dB do 5.9 dB u odnosu na PCM. U poređenju sa FAOC-om, predloženi modifikovani BTC algoritam ostvaruje dobitak od oko 1.9 dB do 0.5 dB za manje veličine frejma ( $M \leq 120$ ), dok u slučaju korišćenja dužih frejmova dobitak opada. Ova pojava se može objasniti karakteristikama ulaznog signala, čijom je analizom utvrđeno da ima ukupnu srednju vrednost vrlo blisku nuli. U slučaju primene dužih frejmova, srednja vrednost frejma zapravo bliža nuli nego svojoj kvantovanoj vrednosti uz primenu log-uniformog kvantizera sa 64 nivoa. Ako se fokusiramo na manje veličine frejma do 120 odmeraka, možemo primetiti da dobitak varira od oko 8.1 dB do 6.6 dB u odnosu na PCM, dok iznosi od oko 1.9 dB do 0.5 dB u odnosu na FAOC. Stoga možemo zaključiti da je modifikovani BTC algoritam pogodan za kodovanje signala klasične muzike, naročito u slučaju korišćenja manjih veličina frejma.

Tabela 5.2.3.3 prikazuje performanse predloženog modifikovanog BTC algoritma na signal rok muzike, koji je okarakterisan višom amplitudskom dinamikom i varijansom. Pored toga, u tabeli 4 prikazane su vrednosti SQNR-a, dobijene za isti ulazni signal, primenom optimalnog komandora sa adaptacijom unapred i PCM-a, kao i kod prethodnog ulaznog signala korišćenog u eksperimentu. Možemo primetiti da predloženi

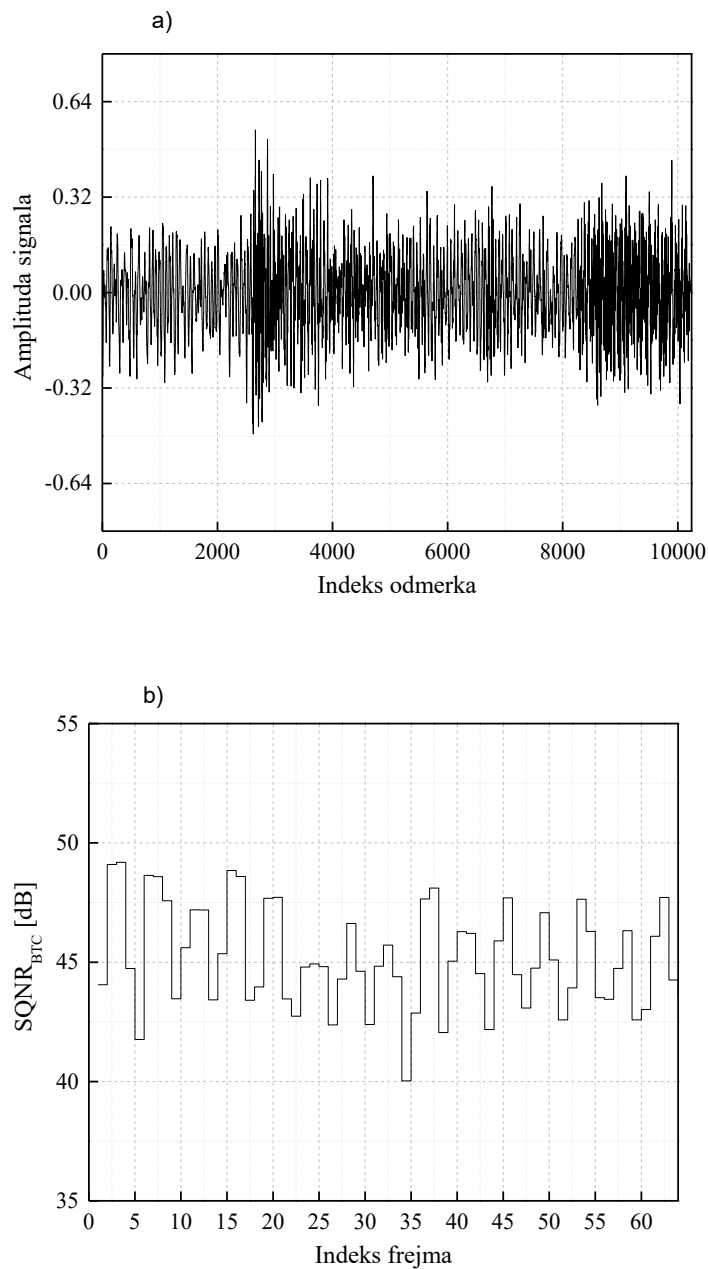
modifikovani BTC algoritam ostvaruje značajan dobitak u SQNR u odnosu na PCM, od oko 8.6 dB kada koristimo veličinu frejma od 40 odmeraka, do oko 6.1 dB u slučaju 320 odmeraka po frejmu. U poređenju sa FAOC, modifikovani BTC algoritam obezbeđuje maksimalni dobitak od 3.9 dB, za veličinu frejma od 40 odmeraka, dok dobitak iznosi oko 0.5 dB, kada je veličina frejma jednaka 320 odmeraka.

Glavna razlika između predloženog modifikovanog BTC algoritma i FAOC-a je u korišćenju informacije o srednjoj vrednosti frejma. Kod FAOC ne izračunavamo i kvantujemo srednju vrednost pa imamo nešto manje informacija o ulaznom signalu. Izračunavanjem srednje vrednosti i uvođenjem signala razlike smanjuje se amplitudska dinamika ulaznog signala, što ga čini pogodnijim za obradu. Ovo ima veliki uticaj, naročito za slučaj korišćenja manjih dužina frejma, kada srednja vrednost frejma tačnije opisuje pojedine odmerke nego što je slučaj sa dužim frejmovima.

Uticaj i benefiti korišćenja srednje vrednosti ulaznog signala u predloženom modifikovanom BTC algoritmu se mogu posmatrati uporednom analizom talasnog oblika ulaznog signala i signala razlike. Za ovu analizu je korišćen isečak signala rok muzike, već korišćenog u analizi performansi predloženog algoritma. Ovaj isečak se sastoji od 10240 odmeraka, koji čine 64 frejma od po 160 odmeraka. Slika 5.2.3.1.a) prikazuje talasni oblik amplitude posmatranog isečka signala rok muzike, dok je na slici 5.2.3.1.b) prikazan SQNR dobijen primenom optimalnog komandora sa adaptacijom unapred. Sa druge strane, Slika 5.2.3.2.a) prikazuje talasni oblik amplitude signala razlike, dok je na slici 5.2.3.2.b) prikazan SQNR dobijen primenom modifikovanog BTC algoritma, koji u postupku kodovanja koristi srednju vrednost signala razlike. U oba slučaja, SQNR se izračunava i prikazuje za pojedinačne frejmove posmatrane sekvence ulaznog signala.



Slika 5.2.3.1. a) Talasni oblik amplituda isečka signala rok muzike, b) SQNR individualnih frejmova dobijen primenom optimalnog komandora sa adaptacijom unapred



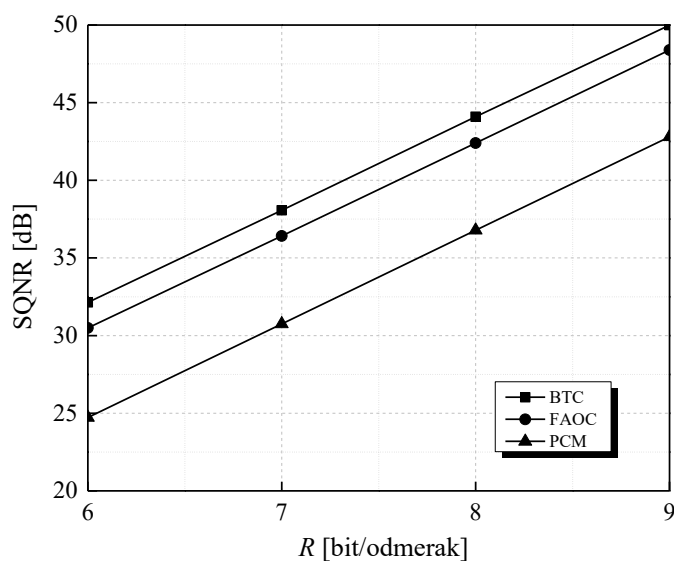
Slika 5.2.3.2. a) Talasni oblik amplituda isečka signala razlike rok muzike, b) SQNR<sub>BTC</sub> po frejmu dobijen primenom modifikovanog BTC algoritma za 64 frejma od po 160 odmeraka

Može se primetiti da signal razlike ima manju amplitudsku dinamiku, tj. uži raspon vrednosti amplituda, što ga čini pogodnijim za kvantizaciju. U ovom trenutku performanse predloženog modifikovanog BTC algoritma se oslanjaju proces kvantizacije

srednje vrednosti, koja se vrši primenom kvazilogaritamskog kvantizera sa 6 bita po odmerku. Upoređivanjem Slika 5.2.3.1.b) i 5.2.3.2.b) jasno možemo videti povećanje u SQNR-u, što je direktna posledica korišćenja srednje vrednosti signala razlike u modifikovanom BTC algoritmu, koji kvantizaciju odmeraka signala razlike takođe vrši primenom optimalnog komandora sa adaptacijom unapred. Dodatno, može se primetiti razlika u SQNR-u od frejma do frejma, što je rezultat prirode ulaznog signala, koji karakteriše visoka amplitudska dinamika. Ova analiza potvrđuje da kvantizacija i upotreba srednje vrednosti frejma za uvođenje signala razlike poboljšava performanse algoritma. Pored toga, pokazano je da korišćenje robusnijih kvantizera obezbeđuje značajan dobitak u SQNR-u.

Kao što je prikazano u analizi prethodno predloženih algoritama, performanse algoritma za kompresiju se mogu predstaviti i pomoću SQNR karakteristike. Slika 5.2.3.3 prikazuje SQNR karakteristiku predloženog modifikovanog BTC algoritma, zajedno sa optimalnim komandorom sa adaptacijom unapred i PCM koderom za poređenje. SQNR karakteristika je dobijena primenom algoritma na signal rok muzike, sa veličinom frejma od 160 odmeraka, za bitsku brzinu od 6, 7, 8 i 9 bita po odmerku. Dobitak u SQNR-u u korist predloženog modifikovanog BTC algoritma je lako uočljiv, posmatrano za sve razmatrane bitske brzine. Sa slike se može primetiti i sledeće. Poređenjem SQNR-a PCM-a u slučaju korišćenja 8 bita po odmerku, sa SQNR-om predloženog modifikovanog BTC algoritma kada se koristi 7 bita po odmerku, vidimo da predloženi modifikovani BTC algoritam ostvaruje viši SQNR. To znači da se predloženim algoritmom ostvaruje kompresija od više od 1 bita po odmerku, u poređenju sa PCM, pri čemu se zadržava visok kvalitet signala. Ova činjenica dodatno potvrđuje da je predloženi modifikovani BTC algoritam pogodan za primenu u kodovanju i kompresiji audio signala.





Slika 5.2.3.3. SQNR karakteristika modifikovanog BTC algoritma, optimalnog komparatora sa adaptacijom unapred i PCM-a za različite bitske brzine i signal klasične muzike

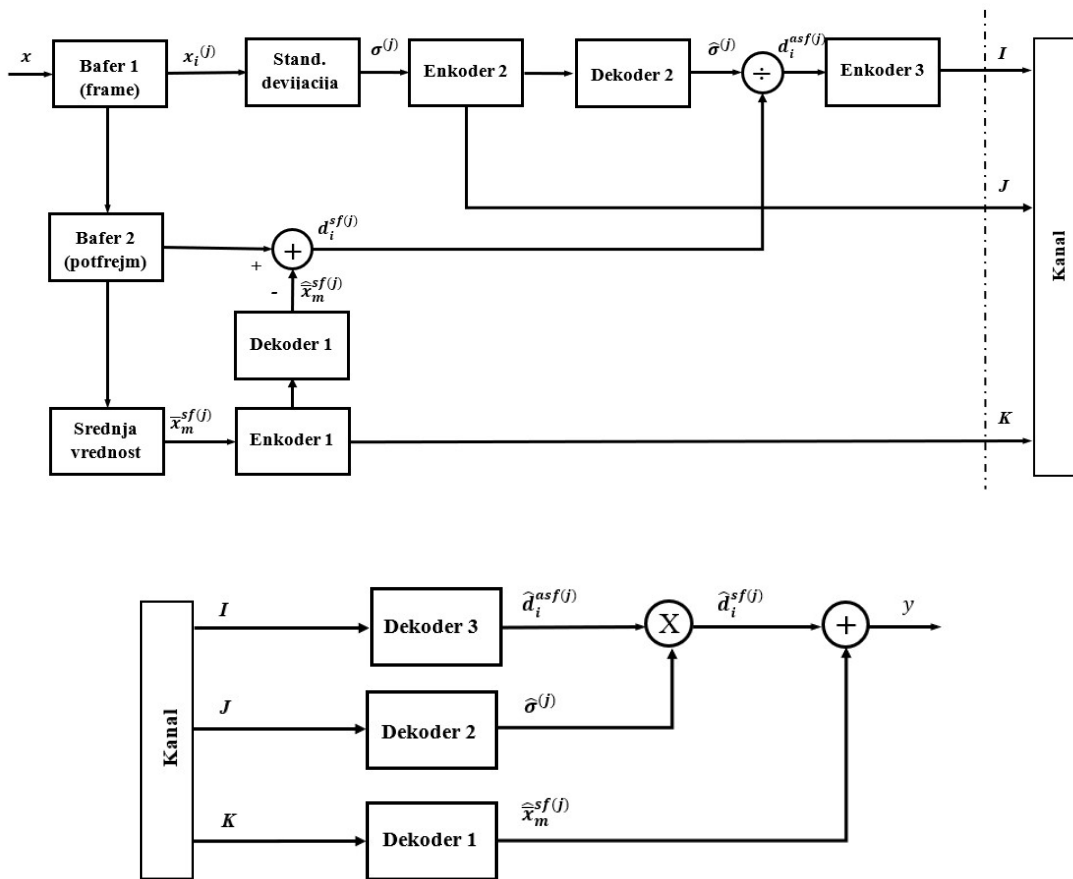
### 5.3 Modifikovani BTC algoritam za kodovanje širokopojasnog govornog signala visokog kvaliteta

Modifikovani BTC algoritam koristi postojanje korelacije između odmeraka, kako bi se postigla viša kompresija signala. Pored audio signala, korelacija između susednih odmeraka postoji i kod širokopojasnog govornog signala, koji se uglavnom uzorkuje frekvencijom odmeravanja od 16 kHz. To je motivisalo razvoj modifikovanog BTC algoritma sa ciljem primene na govorni signal.

#### 5.3.1 Projektovanje modifikovanog BTC algoritma za kodovanje širokopojasnog govornog signala

Za razliku od prethodno opisanog modifikovanog BTC algoritma za kodovanje audio signala, širokopojasni BTC algoritam primenjuje poddelu signala na frejmove i podfrejmove. Ovaj pristup je rezultat pretpostavke da se može ostvariti dodatna kompresija govornog signala i bolja adaptacije algoritma na ulazni signal drugačijim korišćenjem sporedne informacije. Uvođenjem i kvantovanjem podfrejmova algoritam

vrši bolje prilagođenje kvantizera na podfrejm, koji sadrži manji broj odmeraka, dok se ušteda u korišćenju bitskoj brzini ostvaruje korišćenjem standardne devijacije čitavog frejma. Blok šema modifikovanog BTC algoritma za kodovanje širokopojasnog govornog signala je predstavljena na slici 5.3.1.1.



Slika 5.3.1.1. Blok šema modifikovanog BTC algoritma za kodovanje širokopojasnog govornog signala, enkoder (iznad) i dekoder (ispod)

Kako amplituda odmerka širokopojasnog govornog signala može imati nepredvidljivu vrednost, BTC algoritam za kodovanje širokopojasnog govornog signala primenjuje robusnije metode kvantizacije, uz korišćenje viših bitskih brzina u odnosu na originalni BTC algoritam. Projektuju se tri različita kvantizera koji se primenjuju na različite parametre ulaznog signala. Pored toga, slično kao i u prethodno predstavljenom algoritmu, vrši se smanjenje amplitudske dinamike ulaznog signala stvaranjem signala razlike nad kojim se primenjuju tehnike kodovanja. Za razliku od prethodnog algoritma, ovde se signal razlike dobija oduzimanjem kvantovane srednje vrednosti podfrejma od

svakog odmerka u podfrejmu, za frejm koji je sačinjen od ukupno  $M_f$  odmeraka i  $L$  podfrejmova, gde svaki podfrejm sadrži  $M_{sf}$  odmerka.

$$d_i^{sf(j)} = x_i^{(j)} - \hat{x}_m^{sf(j)}, i = 1, 2, \dots, M_{sf}, j = 1, 2, \dots, L, \quad (5.3.1.1)$$

prilikom čega  $x_i^{(j)}$  označava  $i$ -ti odmerak ulaznog signala,  $j$ -tog podfrejma, dok  $\hat{x}_m^{sf(j)}$  predstavlja kvantovanu srednju vrednost  $j$ -tog podfrejma. Srednja vrednost podfrejma se kvantuje primenom logaritamskog kvantizera sa  $\mu$  zakonom kompresije, opisanog u odeljku 3.3.4 (Enkoder 1 i Dekoder 1). Za kvantovanje srednje vrednosti podfrejma se koristi 5 bita po odmerku uz faktor kompresije  $\mu=255$ . Kvantovana srednja vrednost podfrejma se dalje vodi na ulaz Enkodera 3, zajedno sa standardnom devijacijom frejma. Standardna devijacija se računa za svaki frejm i kvantuje primenom log-uniformnog kvantizera opisanog u odeljku 5.2.2, na blok šemi označenog kao Enkoder 2 i Dekoder 2, sa bitskom brzinom od 4 bita po odmerku. Standardna devijacija frejma se koristi u adaptivnoj kvantizaciji podfrejma signala razlike koji trenutni frejm sadrži. Podfrejm signala razlike se kvantuje primenom kvazilogaritamskog kvantizera sa adaptacijom unapred, koji primenjuje faktor kompresije  $\mu=80$  i osnovnu bitsku brzinu od 7 bita po odmerku. Rezultat adaptivne kvantizacije se dobija adaptivni signal razlike podfrejma,  $\hat{d}_i^{asf(j)}$ , prilagođen statistici ulaznog signala.

Izlazi sva tri enkodera se kroz kanal šalju kao binarna informacija, koja se dekoduje na strani prijema. Dekodovana vrednost izlaznog signala se dobija sabiranjem kvantovane vrednosti podfrejma i adaptivno kvantovanih odmeraka podfrejma ulaznog signala.

$$y_i^{(j)} = \hat{x}_m^{sf(j)} + \hat{d}_i^{asf(j)}, i = 1, 2, \dots, M_{sf}, j = 1, 2, \dots, L. \quad (5.3.1.2)$$

Ovaj postupak se ponavlja za svaki podfrejm i frejm ulaznog signala čime se dobija rekonstruisani signal čiji se kvalitet može dalje analizirati.

Kako algoritam primenjuje podelu ulaznog signala na frejmove i podfrejmove, potrebno je definisati ukupnu bitsku brzinu. Bitska brzina je određena veličinama frejma i podfrejma, kao i bitskih brzina korišćenih od strane pojedinačnih kvantizera. Na osnovu toga, bitska brzina je definisana sledećim izrazom:

$$R = r_3 + \frac{r_1}{M_{sf}} + \frac{r_2}{M_f}, \quad (5.3.1.3)$$

prilikom čega  $r_3$  označava bitsku brzinu koja se koristi za kvantovanje podfrejma signala razlike,  $r_1$  predstavlja bitsku brzinu korišćenu za kodovanje srednje vrednosti podfrejma ulaznog signala,  $r_2$  označava bitsku brzinu koja se koristi za kodovanje standardne devijacije frejma, dok  $M_{sf}$  i  $M_f$  označavaju broj odmeraka koji podfrejm i frejm sadrže, respektivno.

### 5.3.2 Analiza eksperimentalnih rezultata

Performanse predloženog modifikovanog BTC algoritma za kodovanje širokopoljarnog govornog signala su određene njegovom primenom na realne ulazne signale. U eksperimentu su korišćeni signali muškog i ženskog govornika, snimljeni u akustičkom studiju na Elektronskom fakultetu u Nišu, sa frekvencijom odmeravanja od 16 kHz. Kao i prilikom analize performansi prethodno predstavljenih algoritama, korišćena objektivna mera kvaliteta je SQNR. Dobijene performanse se porede sa performansama PCM algoritma.

Veliki uticaj na performanse i ukupnu bitsku brzinu imaju veličine frejma i podfrejma. Odabrana veličina frejma iznosi 320 odmeraka, dok veličina podfrejma varira u opsegu od 5 do 40 odmeraka. Kako se pre obrade vrši učitavanje frejma od 320 odmeraka u bafer, algoritam unosi kašnjenje signala od 0.02 sekunde. Kada se odabere da veličina frejma i podfrejma budu jednake, dobija se specijalni slučaj u kome se praktično ne primenjuje podela ulaznog signala na podfrejmove, jer u ovom slučaju podfrejm sadrži čitav frejm. Posmatranje ovog specijalnog slučaja pogodno je za analizu doprinosa podele signala na podfrejmove, poređenjem performansi algoritma. Naredna tabela prikazuje eksperimentalne rezultate primene predloženog algoritma na signal muškog govornika.

Pored SQNR-a, veoma bitna stavka za analizu performansi širokopoljarnog BTC algoritma u tabeli je i ukupna bitska brzina. Lako je uočiti uticaj dužine podfrejma na ukupnu bitsku brzinu i da se korišćenjem podfrejma dužine 5 odmeraka unosi 1 bit po odmerku dodatne informacije. Takođe, možemo uočiti da modifikovani BTC algoritam ostvaruje dobitak u SQNR-u u odnosu na PCM za svaki od razmatranih dužina podfrejma. U slučaju kada se ne implementira podela frejma na podfrejmove (za jednake dužine frejma i podfrejma), dobitak u SQNR-u iznosi oko 4.5 dB.

Tabela 5.3.2.1. SQNR modifikovanog BTC algoritma ostvaren za različite dužine podfrejma i PCM-a za signal muškog govornika sa frekvencijom odmeravanja od 16 kHz

Dužina frejma [odmeraka]	Dužina podfrejma [odmeraka]	Bitska brzina [bit/odmerak]	SQNR <sub>BTC</sub> [dB]	SQNR <sub>PCM</sub> [dB]
320	5	8.0125	39.7409	28.13
320	10	7.5125	37.0506	28.13
320	20	7.2625	34.5204	28.13
320	40	7.1375	33.2855	28.13
320	320	7.0281	32.666	28.13

Kako bi se analizirao uticaj podfrejmova, možemo posmatrati slučaj kada je veličina podfrejma jednaka 10 odmeraka. U ovom slučaju bitska brzina se povećava za 0.5 bita po odmerku usled dodatne informacije. Kao što je poznato, korišćenje jednog bita po odmerku više prilikom kodovanja PCM tehnikom povećava SQNR za približno 6 dB [3], pa bi stoga dodatnih 0.5 bita po odmerku, trebalo da poveća SQNR za oko 3 dB. Upoređivanjem vrednosti iz Tabele I, vidimo da je SQNR modifikovanog BTC algoritma povećan za oko 4,4 dB. To znači da dobijamo dobitak u SQNR-u od oko 1,4 dB, samom primenom podfrejma u algoritam. Primenom iste logike na PCM, možemo proceniti SQNR za bitsku brzinu od 7.51 bita po odmerku, koji bi iznosio približno 31.13 dB. Možemo zaključiti da modifikovani BTC algoritam ostvaruje dobitak u SQNR-u od oko 5.9 dB, u slučaju podfrejma veličine 10 odmeraka i performansi PCM-a procenjenih za 7.51 bita po odmerku: 37.05 dB - (28,13 dB + 3 dB).

Tabela 5.3.2.2 prikazuje performanse modifikovanog BTC algoritma za kodovanje širokopojasnog govornog signala i PCM-a, za slučaj primene na signal ženskog govornika, sa frekvencijom odmeravanja od 16 KHz. Primenom istih principa kao kod analize prethodne tabele uočavamo da modifikovani BTC algoritam bez implementacije podfrejma ostvaruje dobitak SQNR-u od oko 3.6 dB, u poređenju sa PCM-om. U slučaju podfrejma koji se sastoji od 10 odmeraka, stvarni dobitak u SQNR-u iznosi oko 5.7 dB, u poređenju sa PCM-om sa performansama procenjenim za istu bitsku brzinu koja je jednaka 7.51 bita po odmerku. Uticaj primene podfrejma analiziramo poređenjem performansi modifikovanog BTC algoritma bez implementacije podfrejma sa slučajem primene podfrejma dužine 10 odmeraka. Kako se povećavanjem bitske brzine za 0.5 bita

po odmerku očekuje dobitak od oko 3 dB, a primenom modifikovanog BTC algoritma sa 7.5125 bita po odmerku se ostvaruje dobitak od oko 5 dB, zaključujemo da se samo uvođenjem podfrejma u algoritam ostvaruje dobitak od 2 dB u odnosu na slučaj kada se ne vrši podela frejma na podfrejmove.

Tabela 5.3.2.2. SQNR modifikovanog BTC algoritma ostvaren za različite dužine podfrejma i PCM-a za signal ženskog govornika sa frekvencijom odmeravanja od 16 kHz

Dužina frejma [odmeraka]	Dužina podfrejma [odmeraka]	Bitska brzina [bit/odmerak]	SQNR <sub>BTC</sub> [dB]	SQNR <sub>PCM</sub> [dB]
320	5	8.0125	41.2025	29.60
320	10	7.5125	38.2895	29.60
320	20	7.2625	35.4304	29.60
320	40	7.1375	33.8531	29.60
320	320	7.0281	33.2202	29.60

Dakle, u prethodna dva odeljka su predložena dva algoritma za kodovanje audio signala i širokopolasnog govornog signala, zasnovanih na BTC algoritmu koji je originalno namenjen kompresiji crno-bele slike. Pokazano je kako se primenom jednostavnih modifikacija i robusnijih tehnika kodovanja dobijaju algoritmi pogodni za kodovanje govornog i audio signala i kojima se ostvaruje visok kvalitet kodovanog signala.

#### **5.4 Algoritam za jednostavno diferencijalno kodovanje govornog signala zasnovano na adaptaciji unazad**

U ovom odeljku će biti predložen algoritam za kodovanje govornog signala zasnovan na diferencijalnom kodovanju i adaptaciji unapred uz primenu kvazilogaritamske kvantizacije. Diferencijalno kodovanje se zasniva na predikciji vrednosti sledećeg odmerka na osnovu prethodno obrađenih odmeraka. Razlika između predviđene i stvarne vrednosti odmerka se naziva greška predikcije (signal razlike) koja se prenosi umesto stvarnih vrednosti odmerka ulaznog signala. Neke od prednosti diferencijalnog kodovanja su već opisane u prethodnim odeljcima. Pokazano je kako se uvođenjem signala razlike

(diferencije) dobija signal koji je pogodniji za kodovanje od originalnog signala, koji je okarakterisan manjom varijansom i nižom amplitudskom dinamikom.

Pored diferencijalnog kodovanja algoritam koji će biti predstavljen u ovom odeljku primenjuje i adaptaciju unazad. O adaptaciji unazad je već bilo reči u disertaciji. Kako se adaptacija vrši na osnovu kvantovanih odmeraka ulaznog signala, koji su prilikom rekonstrukcije dostupni prijemniku nije potrebno prenositi dodatnu informaciju. Kako se diferencijalno kodovanje i adaptacija vrše unazad, bitska brzina se ne povećava primenom ovih tehnika, dok se očekuje značajno povećanje kvaliteta kodovanog signala. Takođe, kako se ne prenosi sporedna informacija, algoritam može biti primenjen u okolnostima koje zahtevaju malo kašnjenje.

Razmatrani algoritam modeluje ulazni signal Gausovom funkcijom gustine verovatnoće, koja je za slučaj srednje vrednosti jednakoj nuli i varijansom  $\sigma^2$  definisana izrazom (3.2). Algoritam koristi optimalni granularni region kvantizacije, čije se granice određuju iterativnom metodom. Optimalne amplituda maksimalnog i minimalnog opterećenja kvantizera se dobijaju minimizacijom izraza za ukupnu distorziju. U slučaju kvazilogaritamskog kvantizera i ulaznog signala modelovanog Gausovim izvorom sa srednjom vrednošću jednakom nuli i varijansi  $\sigma^2=1$ , granularna distorzija je definisana kao [3]:

$$D_g = \frac{\ln^2(\mu+1)}{3N^2} \left[ \frac{t_{N-1}^2}{\mu^2} + t_{N-1} \frac{2}{\mu} \sqrt{\frac{2}{\pi}} + 1 \right]. \quad (5.4.1)$$

U slučaju primene veće vrednosti faktora kompresije, kao što je slučaj sa kvantizerin definisanim G.711 standardom gde je  $\mu=255$ , ( $\mu \gg 1$ ) deo izraza  $\frac{1}{\mu^2}$  teži nuli, pa se prethodna jednakost može aproksimirati kao:

$$D_g = \frac{\ln^2(\mu+1)}{3N^2} \left[ \frac{2}{\mu} \sqrt{\frac{2}{\pi}} t_{N-1} + 1 \right]. \quad (5.4.2)$$

Za kvantizer sa amplitudom maksimalnog opterećenja  $t_{N-1}$  i ulaznim signalom modelovanim Gausovim izvorom sa nultom srednjom vrednošću i jediničnom varijansom, distorzija prekoračenja se može izračunati kao [68]:

$$D_o = \sqrt{\frac{2}{\pi}} \frac{2}{t_{N-1}^3} \exp\left\{-\frac{t_{N-1}^2}{2}\right\}. \quad (5.4.3)$$

Ukupna distorzija je određena kao zbir granularne distozije i distorzije prekoračenja  $D=D_g+D_o$ . Minimizacija distorzije po  $t_{N-1}$  se vrši na sledeći način:

$$\frac{\partial D}{\partial t_{N-1}} = 0. \quad (5.4.4)$$

Zamenom izraza za distorziju u prethodnu jednakost dobija se izraz za optimalnu amplitudu maksimalnog opterećenja, koja se iterativno proračunava kao:

$$t_{N-1}^{(i)} = \sqrt{\left| 2 \ln \left( \frac{1}{\mu} \frac{\ln^2(\mu+1)}{3N^2} \frac{t_{N-1}^{4(i-1)}}{t_{N-1}^{2(i-1)} + 3} \right) \right|}. \quad (5.4.5)$$

Tabela 5.4.1. SQNR modifikovanog BTC algoritma ostvaren za različite dužine podfrejma i PCM-a za signal ženskog govornika sa frekvencijom odmeravanja od 16 kHz

$R$ bitska brzina	Faktor kompresije $\mu$	Početna vrednost $t_{N-1}$	1. iteracija	2. iteracija	3. iteracija
6	255	3	4.3869	4.1760	4.2028
		3.5	4.3001	4.1869	4.2014
		4	4.2264	4.1963	4.2002
		4.5	4.1622	4.2046	4.1991
		5	4.1052	4.2122	4.1981

Kako bi se ubrzalo izvršenje algoritma koriste se samo tri iteracije, kojima se dobija rezultat veoma blizak optimalnom. Početne vrednosti iterativnog postupka su uzete intuitivno, kako je poznato da će one konvergirati blizu optimalnog rešenje za različite polazne vrednosti. Sledeća tabela pokazuje konvergenciju amplitude maksimalnog opterećenja za različite polazne vrednosti. Možemo videti da se nakon treće iteracije dobijaju veoma bliske vrednosti, za sve razmatrane početne vrednosti.



Kvantizacija ulaznih odmeraka se vrši primenom kvazilogaritamskog kvantizera sa adaptacijom unazad, čija je kompresorska funkcija data izrazom (3.3.3.1). Razmatrano je više faktora kompresije, uz očekivanje da adaptivni kvantizer postiže bolje performanse sa nižim vrednostima faktora kompresije. Time se uz primenu optimalne amplitude maksimalnog opeterećenja i faktora kompresije, uz primenu diferencijalnog kodovanja i adaptacije unazad dobija jednostavan algoritam sposoban da ostvari visok kvalitet kodovanog signala, što će biti provereno primenom na realne ulazne signale.

#### 5.4.1 Tok algoritma

Kako algoritam primenjuje adaptaciju unazad, potrebno je izvršiti inicijalizaciju, usled toga što za prvi frejm ne postoji informacija o prethodno kodovanom frejmu. Ovaj problem je prevaziđen primenom PCM enkodera [7] na prvi frejm, uz korišćenje identične bitske brzine koja je na raspolaganju adaptivnom kvazilogaritamskom kvantizeru. Time se dobija prvi kodovani frejm, čiji se statistički parametri mogu koristiti u postupku diferencijalnog i adaptivnog kodovanja preostalih frejmova. Najpre se, počevši od drugog frejma, računa koeficijent korelacije koji se koristi u kodovanju trenutnog  $j$ -tog frejma:

$$\rho^{(j-1)} = \frac{\sum_{n=1}^{M-1} \hat{x}^{(j-1)}[n] \times \hat{x}^{(j-1)}[n+1]}{\sum_{n=1}^M (\hat{x}^{(j-1)}[n])^2}, j=2, \dots, L, \quad (5.4.1.1)$$

gde  $L$  predstavlja ukupan broj frejmova,  $M$  broj odmeraka koji svaki frejm sadrži, dok  $\hat{x}[n]$  and  $\hat{x}[n+1]$  označavaju kvantovane vrednosti trenutnog i sledećeg odmerka ulaznog signala, respektivno. Sledeći korišćeni parametar je variansa prethodno kvantovanog frejma, koja je definisana kao:

$$\left[ \sigma_{\hat{x}}^{(j-1)} \right]^2 = \frac{1}{M} \sum_{n=1}^M (\hat{x}^{(j-1)}[n] - \eta^{(j-1)})^2, \quad (5.4.1.2)$$

gde  $\eta^{(j-1)} = \frac{1}{M} \sum_{n=1}^M \hat{x}^{(j-1)}[n]$ ,  $j = 2, \dots, L$  predstavlja srednju vrednost  $(j-1)$ og kvantovanog frejma.

Frejm signala razlike se formira odmerak po odmerak:

$$d^{(j)}[n] = x^{(j)}[n] - \rho^{(j-1)} \hat{x}^{(j)}[n-1]. \quad (5.4.1.3)$$

Sledeći korak je računanje varijanse frejma razlike, koja je definisana na sledeći način:

$$\left[\sigma_d^{(j)}\right]^2 = \left[\sigma_{\hat{x}}^{(j-1)}\right]^2 (1 - \left[\rho^{(j-1)}\right]^2). \quad (5.4.1.4)$$

Poslednji korak enkodovanja je primena projektovanog kvazilogaritamskog kvantizera na odmerke signala razlike:

$$\hat{d}^{(j)}[n] = Q_{ad} \left[ d^{(j)}[n] \right], \quad (5.4.1.5)$$

gde se adaptivne amplitude maksimalnog opterećenja određuju kao:

$$t_{N-1}^{(j)} = t_{N-1} \times \sigma_d^{(j)}, \quad (5.4.1.6)$$

gde se  $t_{N-1}$  dobija primenom izraza (5.4.5). Na ovaj način se vrši primena adaptacije unazad, gde se postupak kvantizacije prilagođava statistici ulaznog signala, u ovom slučaju, statistici prethodno kvantovanog frejma signala razlike. Izlaz opisanog enkodera se prenosi kroz kanal kao binarna informacija, koja se rekonstruiše na prijemu.

Dekodovanje predstavlja inverzan proces enkodovanju, sa ciljem rekonstrukcije originalnog ulaznog signala. Proces dekodovanja započinje rekonstrukcijom prvog frejma, koja se vrši primenom PCM dekodera. Nakon rekonstrukcije prvog frejma, može se pristupiti statističkim podacima koji su potrebni za rekonstrukciju sledećeg frejma. Kao i u postupku enkodovanja, najpre se proračunava koeficijent korelacije prethodnog frejma, primenom izraza (5.4.1.1). Sledeći korišćeni statistički parametar je varijansa prethodnog frejma, koja se dobija primenom izraza (5.4.1.2). Nakon dobijanja informacije o varijansi prethodnog frejma, može se primeniti inverzna kvantizacija, kako bi se dobio odmerak signala razlike:

$$\hat{d}^{(j)}[n] = Q_{ad}^{-1} [I]. \quad (5.4.1.7)$$

Nakon toga, vrednost ulaznog odmerka se može rekonstruisati primenom sledećeg izraza:

$$\hat{x}^{(j)}[n] = \hat{d}^{(j)}[n] + \rho^{(j-1)} \hat{x}^{(j)}[n-1]. \quad (5.4.1.8)$$

Ovaj postupak se ponavlja za svaki odmerak i frejm ulaznog signala, nakon čega se dobija kompletno rekonstruisani ulazni signal.

## 5.4.2 Analiza numeričkih rezultata

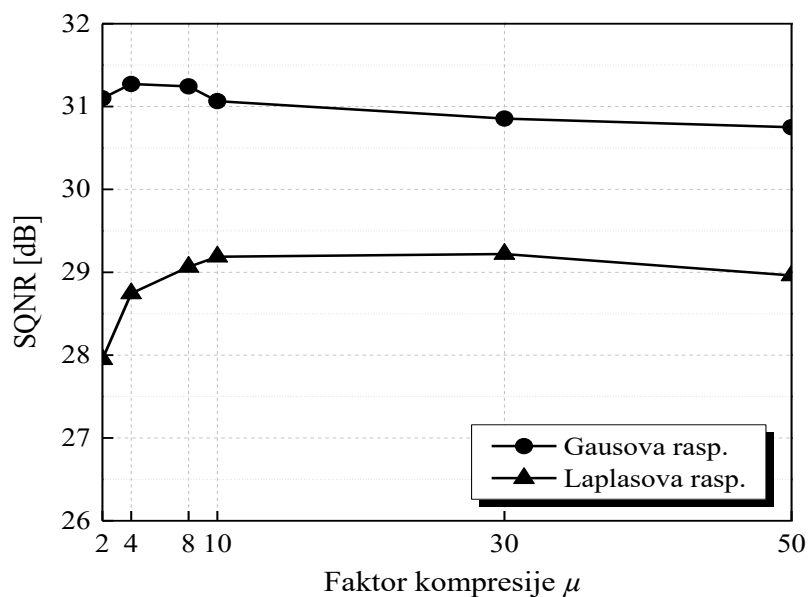
Nakon kompletiranja postupka dekodovanja može se vrši analiza kvaliteta kodovanog signala. Kako i kod do sada razmatranih algoritama korišće se SQNR kao objektivna mera kvaliteta kodovanog signala. Kako se primenjuje kvazilogaritamski kvantizer, najpre će biti analizirane teorijske performanse kvantizera. Performanse su ispitane za kvazilogaritamski kvantizer uz korišćenu bitsku brzinu od 6 bita po odmerku i različite faktore kompresije. Govorni signal je modelovan Gausovim izvorom sa nultom srednjom vrednošću i jediničnom varijansom, kao što je pretpostavljeno u projektovanju algoritama. Dodatno, prikazane su teorijske performanse kvazilogaritamskog kvantizera za slučaj modelovanja ulaznog signala Laplasovim izvorom nulte srednje vrednosti i jedinične varijanse.

U slučaju modelovanja ulaznog signala Laplasovom funkcijom gustine verovatnoće, optimalna amplituda maksimalnog opterećenja kvantizera može biti definisana kao [1]:

$$t_{N-1}^{Lap} = \begin{cases} \frac{1}{\sqrt{2}} \ln \left( \frac{3N^2 \mu}{\ln^2(1+\mu)} \right), & \mu \geq 255 \\ \frac{1}{\sqrt{2}} \ln \left( \frac{\frac{3N^2 \mu}{\ln^2(1+\mu)}}{1 + \frac{1}{\mu} \ln \left( \frac{3N^2 \mu}{\ln^2(1+\mu)} \right)} \right), & \mu < 255 \end{cases} \quad (5.4.2.1)$$

Ukupna distorzija kvazilogaritamskog kvantizera projektovanog za Laplasov izvor se može proračunati na sledeći način [2], [69]:

$$D^{Lap} = \frac{\ln^2(\mu+1)}{3N^2} \left[ \frac{(t_{N-1}^{Lap})^2}{\mu^2} + \frac{\sqrt{2}t_{N-1}^{Lap}}{\mu} + 1 \right] + \exp\{-\sqrt{2}t_{N-1}^{Lap}\}. \quad (5.4.2.2)$$



Slika 5.4.2.1. Teorijska zavisnost SQNR-a od faktora kompresije za kvazilogaritamski kvantizer projektovan za Laplasov i Gausov izvor

Dakle, na sledećoj slici su prikazani rezultati teorijskih performansi kvazilogaritamskog kvantizera sa adaptacijom unazad, za slučaj modelovanja ulaznog signala Laplasovom i Gausovom funkcijom gustine verovatnoće. Korišćeni su faktori kompresije u opsegu od 2 do 50, kako se primenjuje adaptivna kvantizacija i najbolje performanse se očekuju za niže vrednosti faktora kompresije.

Sa slike se može uočiti da se modelovanjem ulaznog signala Gausovim izvorom ostvaruje viši SQNR u odnosu na slučaj modelovanja Laplasovim izvorom, za svaki od razmatranih faktora kompresije. Dobitak je najveći za slučaj faktora kompresije jednakom 4, kada dobitak iznosi 2.1 dB. Ova analiza potvrđuje da je u postupku projektovanja kvantizera govorni signal pogodnije modelovati Gausovim izvorom, kako se teorijski ostvaruje viši SQNR u odnosu na slučaj korišćenja Laplasovog izvora.

Kako bi se ispitale performanse predloženog algoritma sproveden je eksperiment. Algoritam je primenjen na realni signal muškog govornika, u trajanju od 15 sekundi, sa frekvencijom odmeravanja od 16 kHz. Objektivna mera kvaliteta je takođe SQNR, izražen u decibelima. Pored predloženog adaptivnog algoritma, rezultati su prikazani i za identičan algoritam, koji ne primenjuje adaptivnu kvantizaciju, kao i za PCM za potrebe poređenja.

Tabela 5.4.2.1. SQNR predloženog adaptivnog algoritma ostvaren za različite faktore kompresije i PCM-a za signal ženskog govornika sa frekvencijom odmeravanja od 16 kHz

$R$ bitska brzina	Dužina frejma	Faktor kompresije $\mu$	SQNR <sub>PCM</sub> [dB]	SQNR <sub>NON-AD</sub> [dB]	SQNR <sub>AD</sub> [dB]
<b>6</b>	<b>40</b>	<b>10</b>	24.0995	3.7499	35.7590
		<b>20</b>	24.0995	7.4873	36.3638
		<b>30</b>	24.0995	9.7586	36.4615
		<b>40</b>	24.0995	11.3872	36.4439
		<b>50</b>	24.0995	12.6633	36.4267
		<b>100</b>	24.0995	16.5379	36.0445
		<b>255</b>	24.0995	21.2701	35.2108

Iz tabele se može uočiti da predloženi algoritam (SQNR<sub>AD</sub>) zadovoljava G.712 standard uz korišćenje 6 bita po odmerku, kako ostvaruje SQNR veći od 34 dB [9], za sve razmatrane vrednosti faktora kompresije. U poređenju sa PCM-om, predloženi algoritam ostvaruje dobitak u SQNR-u u opsegu od 11.1 do približno 12.36 dB. Treba napomenuti da PCM koristi fiksni faktor kompresije  $\mu=255$  i ne primenjuje podelu ulaznog signala na frejmove. Takođe, jasno se vidi značaj primene adaptivne kvantizacije, kako se ostvaruje veliki dobitak u SQNR-u u odnosu na neadaptivnu verziju algoritma. Taj dobitak je očekivano najviši za najnižu vrednost faktora kompresije, dok za slučaj faktora kompresije jednakom 255, dobitak iznosi približno 14 dB u korist predloženog adaptivnog algoritma.

Dodatno, izvršeno je poređenje performansi predloženog algoritma i drugih rešenja iz literature i prethodnih istraživanja koja su slične kompleksnosti izvršavanja. Pored performansi predloženog algoritma, sledeća tabela prikazuje i rezultat primene DPCM tehnike sa uniformnom kvantizacijom i prediktorom drugog reda [70]. Prikazan je i rezultat primene adaptivnog komandring kvantizera sa različitom dužinom kodnih reči [71], kao i transformacionog kodovanja sa primenom adaptacije unapred [72].

Tabela 5.4.2.2. Komparativne SQNR performanse kodnih šema koje ispunjavaju nivo kvaliteta propisan G.712 standardom

Bitska brzina [70]	SQNR <sup>[70]</sup> [dB]	Bitska brzina [71]	SQNR <sup>[71]</sup> [dB]	Bitska brzina [72]	SQNR <sup>[72]</sup> [dB]	Bit rate (predloženo)	SQNR <sub>AD</sub> [dB] (predloženo)
6	30	6.5	35.143	6.52	34.603	6	36.46

Može se videti da predloženi algoritam ostvaruje dobitak u SQNR-u od 6.46 dB u odnosu na DPCM koder predložen u [70]. Iz tabele takođe možemo uočiti da predloženi algoritam jedini zadovoljava G.712 standard za bitsku brzinu od 6 bita po odmerku, dok kodne šeme predstavljene u [71] i [72] zadovoljavaju G.712 standard za bitsku brzinu od 6.5 bita po odmerku. Ako bismo aproksimirali performanse predloženog algoritma za 6.5 bita po odmerku, kao srednju vrednost performansi dobijenih za 6 i 7 bita po odmerku, dobili smo vrednost SQNR-a od približno 39.51 dB. Vidimo da u tom slučaju predloženi algoritam ostvaruje dobitak u SQNR-u iznosu od 4.3 dB i 4.9 dB, u poređenju sa kodnim šemama predloženim u [71] i [72] respektivno.

Na osnovu analize performansi predloženog algoritma i poređenjem sa drugim kodnim tehnikama iz literature, možemo zaključiti da je predloženi algoritam pogodan za primenu u kodovanju govornog signala. Ostvaruje se visok kvalitet kodovanog signala, koji zadovoljava G.712 standard pri bitskoj brzini od samo 6 bita po odmerku.

## **6. ANALIZA KORELISANOSTI REZULTATA SUBJEKTIVNIH I OBJEKTIVNIH MERA KVALITETA KODOVANOG SIGNALA**

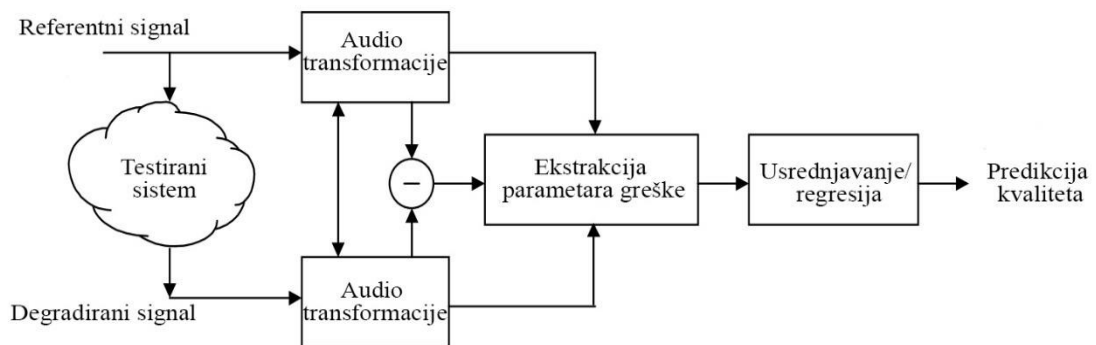
Kvalitet kodovanog signala se može utvrditi subjektivnim i objektivnim merama kvaliteta. U disertaciji je kao mera kvaliteta korišten SQNR, koji predstavlja odnos signal-šum kvantizacije i široko rasprostranjenu meru kvaliteta. Puno je pogodnosti primene objektivnih mera kvaliteta, kao što su ponovljivost, pouzdanost, nepristrasnost, jednostavna primena, brzo vreme izvršenja uz relativno male resurse koji su potrebni. Nasuprot tome, primena subjektivnih ocena kvaliteta govornog signala se i dalje smatra najpouzdanijom, jer, konačno, sve tehnologije i algoritmi se razvijaju upravo za ljudsku upotrebu. Primena subjektivnim mera kvaliteta očigledno nosi sa sobom određene zahteve i poteškoće primene, kao što su obezbeđivanje odgovarajućih tehničkih uslova u kojima se eksperiment može izvesti, pronalaženje i koordinacija sa većim brojem subjekata i vreme potrebno da se ispitivanje sprovede. Sve ovo čini objektivne mere poželjnim za primenu, a njihova pouzdanost i tačnost garantuju da dobijeni rezultati zaista oslikavaju kvalitet kodovanog signala.

Postavlja se pitanje korelisanosti subjektivnih i objektivnih mera kvaliteta kodovanog signala, odnosno da li će kodna tehnika kojom se ostvaruje viši kvalitet meren objektivnom merom kvaliteta istu razliku u kvalitetu pokazati i analizom subjektivne mere kvaliteta. Odgovor na ovo pitanje će biti uporedna analiza rezultata dobijenih primenom različitih tehnika za određivanje kvaliteta kodovanog signala. Pored toga, u ovom poglavlju će kvalitet kodovanog signala ostvaren primenom nekim od algoritama predloženim u disertaciji biti ispitan primenom dodatnih mera procene kvaliteta.

### **6.1 Objektivne metode procene kvaliteta kodovanog signala**

Uporedna analiza tehnika procene kvaliteta kodovanog signala obuhvata ukupno tri različite metode, od kojih su dve objektivne metode za procenu kvaliteta kodovanog signala. Prva je, do sada korišćena metoda za procenu objektivne mere kvaliteta kodovanog signala, odnos signal-šum kvantizacije, odnosno SQNR. SQNR je definisan u poglavlju 3.1.

Druga objektivna metoda procene kvaliteta koja će biti analizirana je PESQ mera koja predstavlja perceptualnu procenu kvaliteta kodovanog signala. PESQ ima zadatak da pruži objektivnu meru kvaliteta kodovanog signala koja će odgovarati primeni subjektivne metode ocene kvaliteta [73]- [74]. Standarizovan je od strane ITU-T kao P.862 i namenjen je proceni performansi kodera za primenu u govornim komunikacijama [75]. ITU-T preporuka P.862.2 proširuje PESQ standard za primenu u analizi kvaliteta širokopojasnog govornog signala [76]. Iako ne može da zameni subjektivne metode procene kvaliteta PESQ se koristi u proceni performansi kodera u inicijalnoj fazi razvoja. Dodatno, usvojen je i standard za perceptualnu procenu kvaliteta kodovanog audio signala, PEAQ (*Perceptual Evaluation of Audio Quality*) [77].



Slika 6.1.1. Uprošćeni model procene PESQ mere kvaliteta

PESQ se zasniva na poređenju originalnog (referentnog) ulaznog signala sa signalom na koji je dodat šum ili neki drugi negativni elementi nastali usled prenosa određenim sistemom (Slika 6.1.1). Poređenje se vrši frejm po frejm, nakon vremenskog poravnanja signala kako bi se neutralisalo kašnjenje. Razvijen je kognitivni model, koji na osnovu razlike između originalnog i degradiranog signala procenjuje kvalitet koji bi bio opažen od strane slušaoca [74]. Kvalitet je izražen u opsegu od -0.5 do 4.5 (često od 1 do 4.5), gde 4.5 označava najviši kvalitet signala koji se ispituje. Maksimalna vrednost od 4.5 je određena analizom rezultata subjektivnih testova u kojima je upravo 4.5 uglavnom najviša ostvarena vrednost subjektivne mere kvaliteta. PESQ skala je razvijena na osnovu ACR (*Absolute Category Rating*) skale kvaliteta korišćene u testovima slušanja, koja će biti prikazana u nastavku [74], [78], [79].



Tabela 6.1.1. ACR skala ocene kvaliteta prilikom slušanja korišćena u razvoju PESQ mere

Kvalitet govora	Rezultat
Odličan	5
Dobar	4
Prosečan	3
Slab	2
Loš	1

Česta primena PESQ metodologije je i prilikom ocenjivanja uspešnosti algoritama koji imaju zadatak da eliminišu šum iz već kodovanog i prenetog signala (*denoising*), kao što je slučaj sa u radovima [80]-[82]. Pored toga, PESQ je moguće primeniti i u proceni razumljivosti govora [83].

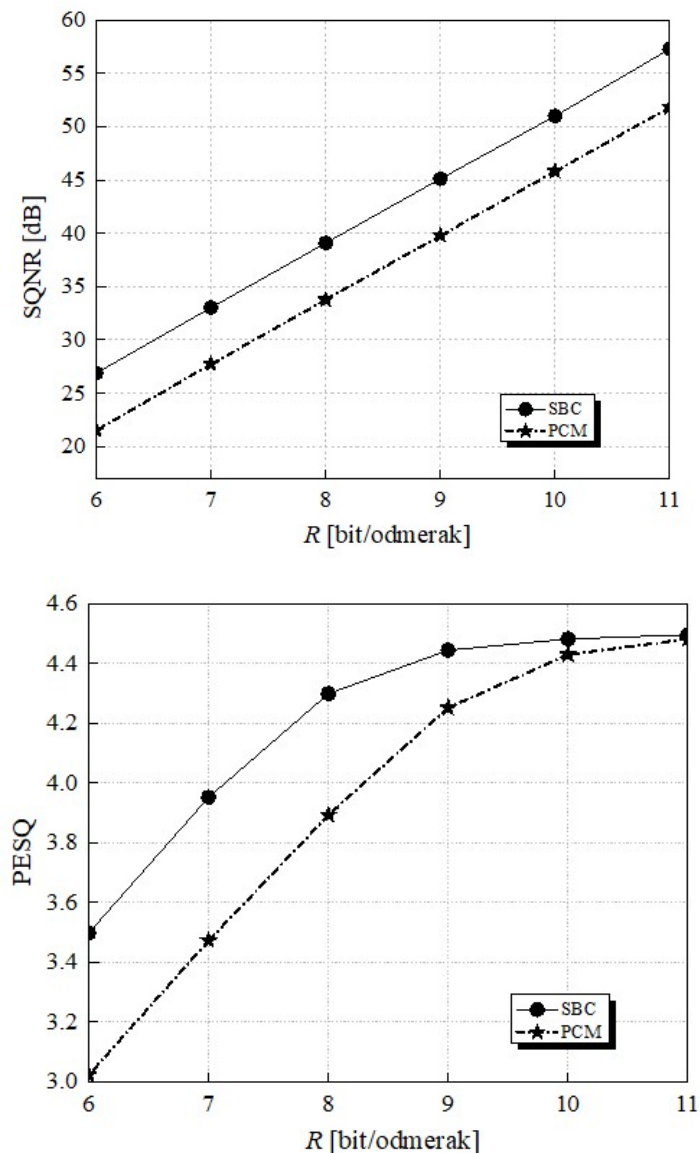
U nastavku će biti prikazana uporedna analiza SQNR-a i PESQ mere za signal muškog govornika, dobijeni primenom podopsežnog kodovanja sa dve grane, predstavljenog u odeljku 4.2.1 (označen kao SBC) i PCM-a. Pored same analize korelisanosti između ove dve objektivne mere kvaliteta, biće dodatno ispitan kvalitet kodovanog signala ostvaren primenom predloženog algoritma za podopsežno kodovanje u odnosu na PCM.

Tabela 6.1.2. SQNR i PESQ mera za podopsežno kodovanje sa dva podopsega i PCM-a primenjenog na signal muškog govornika za različite bitske brzine

R [bit/odmerak]	SQNR <sub>SBC</sub> [dB]	PESQ <sub>SBC</sub>	SQNR <sub>PCM</sub> [dB]	PESQ <sub>PCM</sub>
6	26.9038	3.4980	21.5288	3.0235
7	33.0344	3.9517	27.7341	3.4723
8	39.0932	4.2988	33.7695	3.8926
9	45.1124	4.4439	39.7847	4.2511
10	50.9915	4.4817	45.8096	4.4292
11	57.2893	4.4953	51.7989	4.4812

Tabela 6.1.1 prikazuje uporedne rezultate SQNR-a u decibelima i PESQ mere, ostvarene primenom dva navedena algoritma na signal muškog govornika, sa frekvencijom odmeravanja od 16 kHz, za bitsku brzinu u opsegu od 6 do 11 bita po odmerku. PESQ mera je dobijena primenom softverskog koda koji su autori standarda učinili dostupnim za korišćenje [73], [74]. Tabela pokazuje da algoritam za podopsežno kodovanje sa dva

podopsega ostvaruje bolje rezultate izražene SQNR-om i PESQ merom. Možemo uočiti da je rezultat primene podopsežnog kodovanja sa dva podopsega uz korišćenje 10 bita po odmerku izuzetno blizak maksimalnoj vrednosti PESQ skale, koja iznosi 4.5, dok je kod PCM-a ovo slučaj tek uz primenu bitske brzine od 11 bita po odmerku. Ovo potvrđuje činjenicu da se primenom predloženog algoritma za podopsežno kodovanje sa dva podopsega ostvaruje kompresija od 1 bita po odmerku u odnosu na PCM i to uz nisku kompleksnost i bez primene adaptivne kvantizacije.



Slika 6.1.2. Uporedna SQNR i PESQ karakteristika za slučaj primene na signal muškog govornika za različite bitske brzine

Slika 6.1.2. prikazuje SQNR i PESQ karakteristike za različite bitske brzine i slučaj primene na signal muškog govornika. Možemo primeti da vrednosti SQNR-a imaju stalni rast za oba razmatrana algoritma, dok algoritam za podopsežno kodovanje sa dva podopsega ostvaruje dobitak u SQNR-u između 5 i 6 dB za sve razmatrane bitske brzine. Sa druge strane, kako PESQ mera ima maksimalnu vrednost, primećujemo da sa povećanjem bitske brzine rezultati konvergiraju ka maksimalnoj vrednosti od 4.5. Takođe, PESQ karakteristika pokazuje da je poboljšanje kvaliteta signala kodovanog primenom tehnike podopsežnog kodovanja više izraženo prilikom korišćenja nižih bitskih brzina, pa za slučaj bitske brzine od 6 bita po odmerku razlika u PESQ meri iznosi 0.4745, dok u slučaju bitske brzine od 10 bita po odmerku ova razlika iznosi 0.0525 u korist algoritma za podopsežno kodovanje sa dva podopsega. Konačno, primenom PESQ mere je potvrđen kvalitet signala kodovanog predloženim algoritmom za podopsežno kodovanje sa dva podopsega.

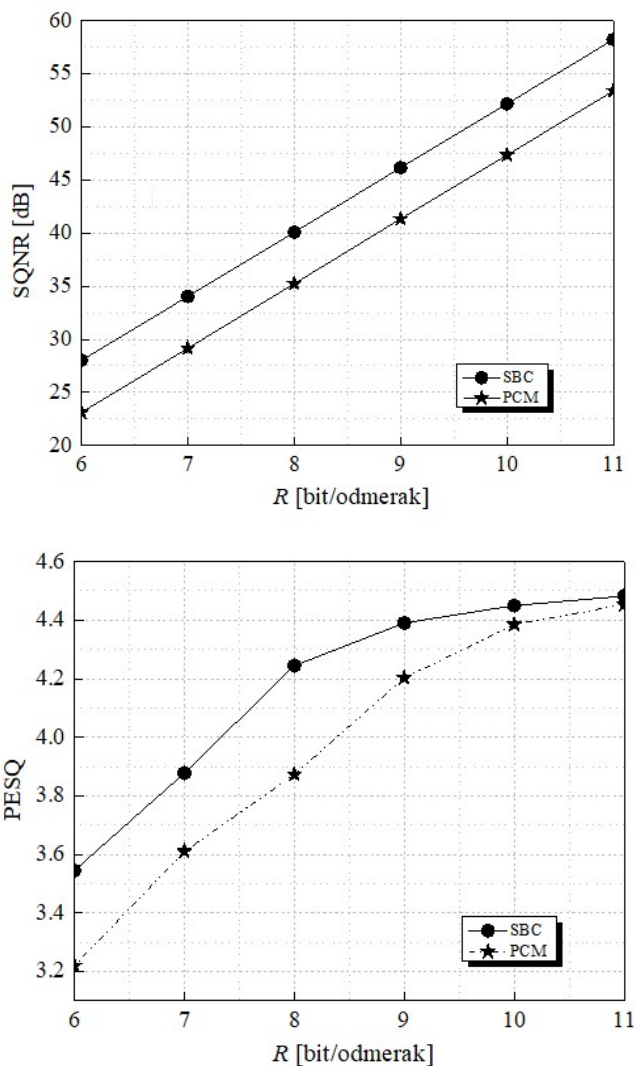
U nastavku će biti prikazana analiza SQNR-a i PESQ mere predloženog algoritma za podopsežno kodovanje sa dva podopsega i PCM-a u slučaju primene na signal ženskog govornika. U pitanju je takođe širokopojasni govorni signal, sa frekvencijom odmeravanja od 16 kHz, snimljen u akustičkom studiju na Elektronskom fakultetu u Nišu. Sledeća tabela prikazuje uporedne rezultate objektivnog kvaliteta signala kodovanim primenom podopsežnog kodovanja i PCM-a.

Tabela 6.1.3. SQNR i PESQ mera za podopsežno kodovanje sa dva podopsega i PCM-a primenjenog na signal ženskog govornika za različite bitske brzine

R [bit/odmerak]	SQNR <sub>SBC</sub> [dB]	PESQ <sub>SBC</sub>	SQNR <sub>PCM</sub> [dB]	PESQ <sub>PCM</sub>
6	28.0180	3.5452	23.1107	3.2192
7	34.0355	3.8774	29.1635	3.6096
8	40.0752	4.2442	35.2398	3.8729
9	46.1525	4.3904	41.3417	4.2030
10	52.1768	4.4498	47.3486	4.3847
11	58.2216	4.4825	53.3959	4.4536

Možemo primetiti veći skok PESQ mere u prelasku od 7 na 8 bita po odmerku u slučaju algoritma za podopsežno kodovanje sa dva podopsega, od 3.8774 na 4.2442. Sa druge strane PESQ mera PCM-a u slučaju korišćenja 9 bita po odmerku iznosi 4.2030, što je

manje u odnosu na podopsežno kodovanje za jedan bit po odmerku manje. Time je potvrđena kompresija signala u odnosu na PCM od 1 bit po odmerku i slučaju ženskog govornog signala, analizom SQNR-a i PESQ mere. Nakon tabele, biće prikazana uporedna SQNR i PESQ karakteristika primenom algoritma za podopsežno kodovanje sa dva podopsega i PCM-a na signal ženskog govornika.



Slika 6.1.3. Uporedna SQNR i PESQ karakteristika za slučaj primene podopsežnog kodovanja i PCM-a na signal ženskog govornika za različite bitske brzine

SQNR karakteristika pokazuje gotovo konstantnu razliku u SQNR-u od približno 5 dB u korist algoritma za podopsežno kodovanje sa dva podopsega. Sa druge strane PESQ karakteristika ne raste ravnomerno sa povećanjem bitske brzine. Možemo primetiti da algoritam za podopsežno kodovanje ostvaruje odličan kvilitet kodovanog signala pri

bitskoj brzini od 8 bita po odmerku, dok PCM ostvaruje ovaj nivo kvaliteta za 1 bit po odmerku više.

Možemo zaključiti da postoji visoka korelacija između rezultata dveju posmatranih objektivnih mera kvaliteta kodovanog signala, kao i da predloženi algoritam za podopsežno kodovanje sa dva podopsega ostvaruje značajano bolji SQNR i PESQ meru u slučaju primene na oba govorna signala u eksperimentu. U nastavku će biti analiziran subjektivni kvalitet signala kodovanih uz primenu dveju tehnika razmatranih u ovom odeljku.

## **6.2 Subjektivna mera kvaliteta kodovanog signala**

### **6.2.1 MOS- *Mean Opinion Score***

Treća razmatrana mera procene kvaliteta kodovanog signala je ujedno i najrasprostranjenija subjektivna mera kvaliteta – MOS (*Mean Opinion Score*) [3], [6], [84]. Internacionalna telekomunikaciona unija definiše MOS kao „vrednost na predefinisanoj skali na kojoj subjekt dodeljuje svoje mišljenje o učinku sistema“ [85]. MOS praktično predstavlja usrednjeni rezultat testova slušanja sprovedenog nad više subjekata (ispitanika), koji procenjuju kvalitet signala. MOS je jedna od najpopularnijih mera procene kvaliteta kodovanog signala. Koristi se sa velikim uspehom u proceni kvaliteta govornog signala, ali i kod drugih tipova signala, kao što su audio i video signali, kao i kod raznih laboratorijskih ispitivanja. Najčešće se koristi MOS skala sa 5 nivoa, gde ispitanik zaokružuje redni broj koji odgovara opisu koji sledi nakon numeričke vrednosti.

Za potrebe ove analize je sproveden eksperiment u akustičkom studiju na Elektronskom fakultetu u Nišu. Iako su nakon sprovođenja testova nad dovoljnim brojem ispitanika rezultati veoma pouzdani, oni ipak dosta zavise i od uslova u kojima su sprovedeni. Bilo kakva buka iz okruženja, šum koji proizvodi oprema, pa i kvalitet same opreme može uticati na rezultate testova slušanja. Merenjima je utvrđeno da akustički studio zadovoljava kriterijume po pitanju spoljašnje buke u skladu sa preporukom međunarodne telekomunikacione unije ITU-T P800 [86]. Kao dodatna mera opreza testovi su sprovedeni u poslepodnevnim časovima, kako bi se minimizirala mogućnost uticaja buke iz okruženja.

Različite ITU preporuke predlažu različiti broj ispitanika, pa tako ITU-T preporuka P.911 predlaže da moguć broj ispitanika treba biti u opsegu između 6 i 40 ljudi [87], dok ITU-T preporuka BT.500 predlaže da je za eksperiment potrebno minimalno 15-oro ispitanika [88].

### 6.2.2 Tok eksperimenta

Za potrebe analize subjektivne procene kvaliteta kodovanog signala, testovi su sprovedeni nad 17-oro ispitanika, bez prethodne obuke, koji su različitog pola, uzrasta i obrazovanja. Ispitanicima je pre testa pušten referentni signal, za koji im nije rečeno kog je kvaliteta. Nakon toga su im puštani signali, koji su bili različitog kvaliteta, odnosno, signali koji predstavljaju izlazni signal algoritma za kodovanje za slučajeve korišćenja različitih bitskih brzina. Pored signala, u eksperimentu je korišćen šum, koji je proizvod algoritma za kodovanje. Šum je izdvojen iz kodovanog signala, takođe za slučaj korišćenja različitih bitskih brzina. Signal i šum, kao i korišćene bitske brzine su bile u nasumičnom redosledu, kako bi rezultati bili što objektivniji. Treba napomenuti da je sam kodovani signal sadržao i šum, ali se uticaj šuma nešto teže procenjuje kada je zajedno sa signalom, jer je signal značajno više amplitude, koja lako može da zamaskira šum.

U zavisnosti od toga da li je ispitaniku pušten signal ili šum pred njim je bilo sledeće pitanje i ponuđeni odgovori. U slučaju procene kvaliteta signala pitanje koje je postavljeno ispitaniku glasi:

- Odnoseći se samo na govorni signal, odaberite kategoriju koja najbolje opisuje isečak koji ste upravo čuli za potrebe svakodnevnog komunikacije.
  - Kvalitet govornog signala u ovom isečku je bio:
    5. Odličan
    4. Dobar
    3. Solidan
    2. Slab
    1. Loš
- Odnoseći se samo na pozadinski šum, odaberite kategoriju koja najbolje opisuje isečak koji ste upravo čuli.

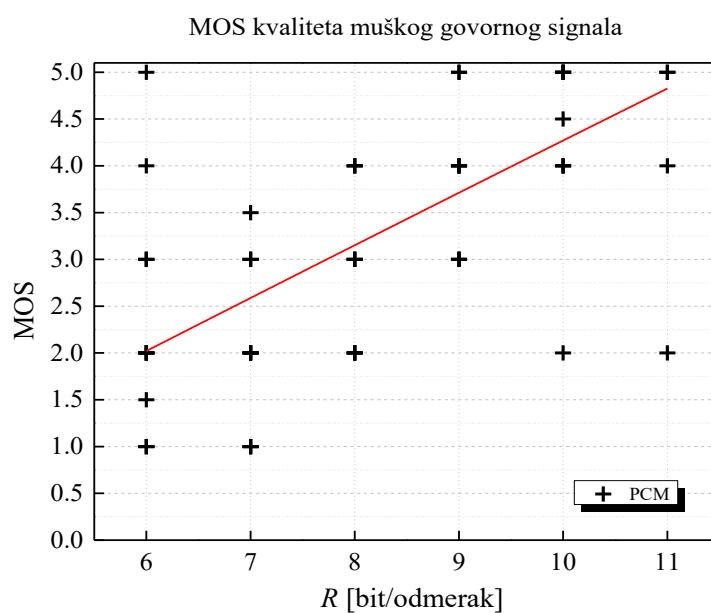
- Pozadinski šum u ovom isečku je bio:

5. Neprimetan
4. Blago primetan
3. Primetan ali ne smeta
2. Pomalo smeta (nametljiv)
1. Veoma smeta

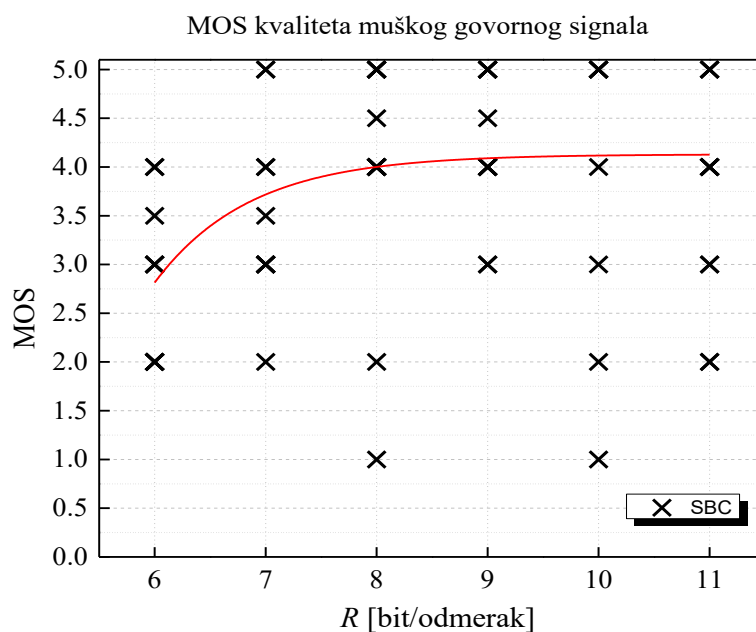
Ispitanicima je puštan jedan po jedan test signal i nakon svakog je ostavljeno vreme za dodeljivanje ocene. Nakon vizuelne potvrde o završetku odgovora nastavlja se sa sledećim test signalom. Usrednjavanjem odgovora svih ispitanika za svaki test signal, dobija se po jedna vrednost na MOS skali za svaki signal. U nastavku su grafički prikazani rezultati testova slušanja, odnosno MOS skor za ispitane signale.

### **6.2.3 Rezultati eksperimenta subjektivne procene kvaliteta kodovanog signala**

Kako određivanje MOS skora podrazumeva dobijanje većeg broja odgovora (brojnih vrednosti) na isto pitanje, rezultati se mogu prikazati na više načina. Rezultati eksperimenta će najpre biti predstavljeni u vidu grafika rasejanja (*scatter plot*). Kako pojedinačni rezultati mogu uzeti samo celobrojne vrednosti, u velikom broju slučajeva dolazi do njihovog preklapanja. Dodatno, grafik prikazuje i ekstrapolacionu krivu, koja na osnovu promena u pojedinačnim rezultatima procenjuje ponašanje krive i prikazuje je na osnovu raspoloživih podataka. Ova kriva praktično prati vrednosti odgovora koji imaju najveće preklapanje, što čini rezultate nešto jasnijim. Grafik rasejanja je prikazan za MOS meru kvaliteta muškog govornog signala, za slučajeve primene algoritma za podopsežno kodovanje sa dva podopsega i PCM-a, za različite bitske brzine. Kako rezultat testova slušanja predstavlja usrednjeni MOS skor, srednje vrednosti rezultata će biti grafički prikazane uz nastavku.



Slika 6.2.3.1. Grafik rasejanja pojedinačnih rezultata testova slušanja za PCM koder



Slika 6.2.3.2. Grafik rasejanja pojedinačnih rezultata testova slušanja algoritma za podopsežno kodovanje sa dva podopsega

Posmatranjem grafika rasejanja možemo uočiti da se kod subjektivne procene kvaliteta rezultati, odnosno odgovori pojedinih ispitanika drastično razlikuju i da isti signal dobija

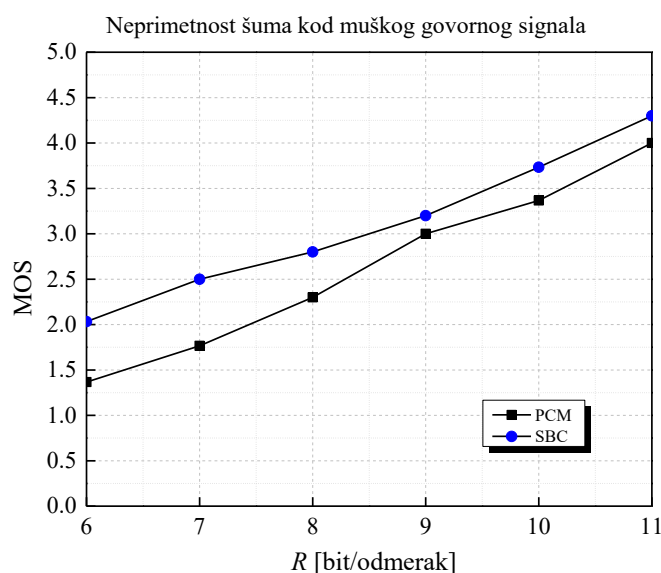
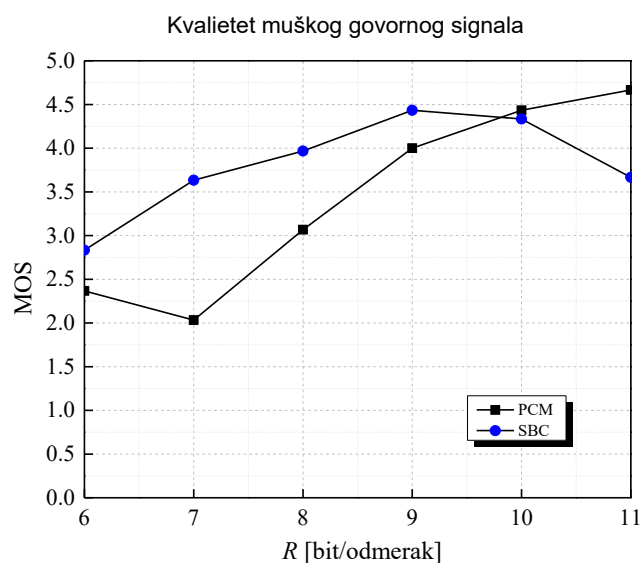


značajno različite ocene. To čini grafik rasejanje tesko preglednim, ša preglednost u određenoj meri olakšava kriva koja prati odgovore sa najvećom učestalošću. Zbog toga će u nastavku biti analizirane srednje vrednosti odgovora svih subjekata koji su učestvovali u eksperimentu.

Ekperimentalni rezultati su prikazani i u vidu grafika MOS skora (srednje vrednosti subjektivne procene kvaliteta) signala kodovanog algoritmom za podopsežno kodovanje sa dva podopsega i PCM-om, za iste bitske brzine, u opsegu od 6 do 11 bita po odmerku. Najpre je prikazan uporedni MOS skor kvaliteta i uticaja šuma kod muškog govornog signala, analiziranog u prethodnim graficima rasejanja.

Posmatranjem MOS skora kvaliteta govornog signala, možemo uočiti da za većinu posmatranih bitskih brzina algoritam za podopsežno kodovanje (SBC) ostvaruje viši subjektivni kvalitet signala. U slučaju korišćenja 11 bita po odmerku signal kodovan PCM-om je ostvario značajno viši MOS skor, što je u suprotnosti sa rezultatima dobijenim objektivnim merama kvaliteta (Slika 6.1.2). MOS skor uticaja šuma je usklađeniji sa objektivnom merom kvaliteta i pokazuje da signal kodovan primenom algoritma za podopsežno kodovanje ima manje primetan šum u odnosi na PCM.

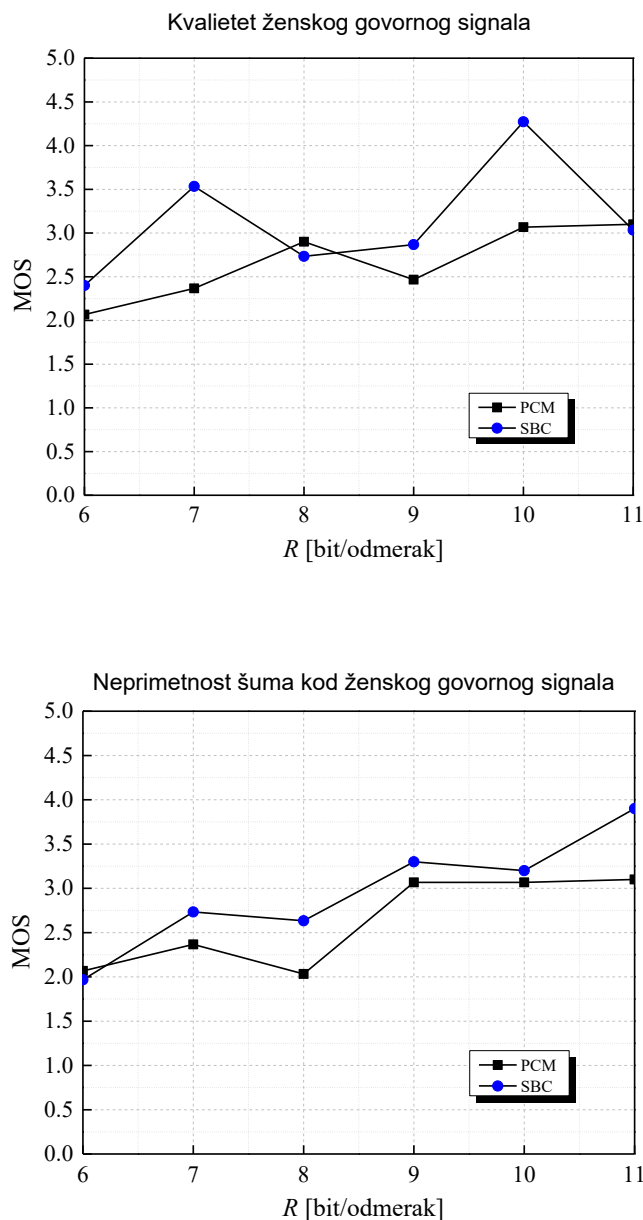
U ovom trenutku možemo izvesti sveobuhvatnu procenu kvaliteta signala kodovanog algoritmom za podopsežno kodovanje sa dva podopsega, na osnovu tri razmatrane mere kvaliteta. Primenom predloženog algoritma za podopsežno kodovanje na signal muškog govornika, za bitsku brzinu od 8 bita po odmerku ostvaruje se SQNR od približno 40 dB, PESQ mera koja iznosi približno 4.25 i MOS skor od približno 4. Možemo zaključiti da kvalitet signala kodovanog predloženim algoritmom za podopsežno kodovanje sa 8 bita po odmerku zadovoljava standard G.712 i postiže dobar kvalitet signala po PESQ meri i MOS skali. Treba napomenuti da je u pitanju najjednostavniji od predloženih algoritama za podopsežno kodovanje sa dva podopsega bez adaptacije, dok se sa preostala dva predložena algoritma za podopsežno kodovanje ostvaruju još bolje performanse i kvalitet kodovanog signala.



Slika 6.2.3.3. MOS ocena kvaliteta signala (iznad) i uticaja šuma (ispod) za signal muškog govornika

Posmatranjem MOS skora u slučaju signal ženskog govornika (Slika 6.2.3.3) uočavamo da skor ne raste ravnomerno sa povećanjem bitske brzine, već poseduje određene oscilacija, za razliku od objektivne procene kvaliteta kodovanog signala. Takođe, uočljivo je da je predloženi algoritam za podopsežno kodovanje ostvario višu MOS ocenu u odnosu na PCM, za gotovo sve razmatrane bitske brzine. Grafik pokazuje da je signal kodovan algoritmom za podopsežno kodovanje dobio višu ocenu za slučaj

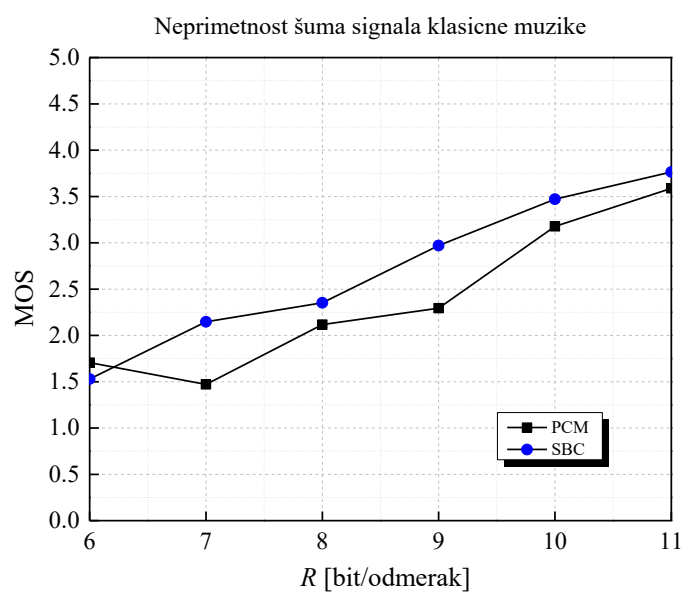
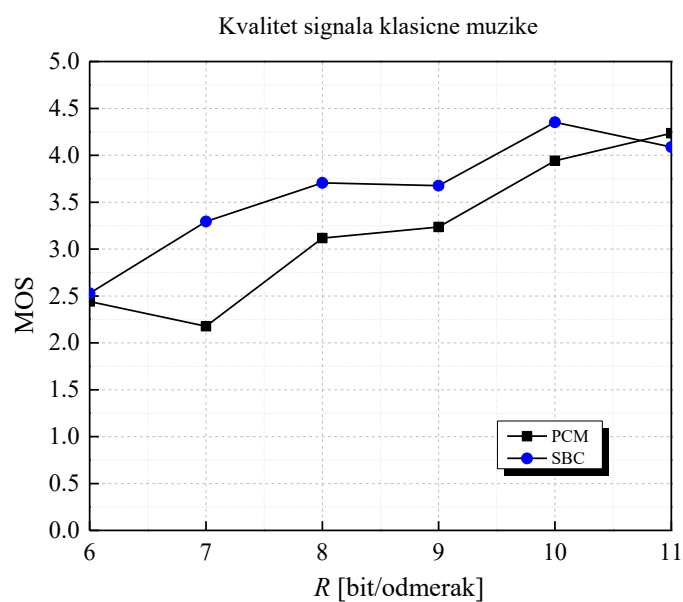
korišćenja 10 bita po odmerku u odnosu na slučaj korišćenja 11 bita po odmerku, što ne odgovara objektivnoj proceni kvaliteta signala. Ova „anomalija“ se može pripisati postojanjem tzv. „*bias-a*“, odnosno uticaja redosleda reprodukcije test signala ispitanicima. Takođe u pitanju su subjektivni testovi slušanja i dva signala kodovana visokom bitskom brzinom, oba izuzetno visokog kvaliteta sa malom međusobnom razlikom.



Slika 6.2.3.4. MOS ocena kvaliteta signala (iznad) i uticaja šuma (ispod) za signal ženskog govornika

Posmatranjem uporedne ocene uticaja šuma PCM-a i algoritma za podopsežno kodovanje vidimo da je ona više međusobno korelisana u odnosu na kvalitet samog kodovanog signala. U ovom delu eksperimenta ispitanicima reprodukovano samo šum izdvojen od korisnog signala i viša ocena označava manju smetnju (nelagodnost) koju po mišljenju ispitanika šum proizvodi. MOS skor pokazuje da algoritam za podopsežno kodovanje unosi manje šuma u odnosu na PCM za sve razmatrane bitske brzine.

Treći signal čiji je subjektivni kvalitet određen testovima slušanja je isečak signala klasične muzike koji je korišćen u prethodnim analizama. Isečak je istog trajanja od 8 sekundi, kao i govor, dok se pitanja koja su postavljena ispitanicima odnose na kvalitet audio signala. Sa grafika se može primetiti nešto viša korelisanost rezultata signala klasične muzike sa objektivnim merama u odnosu na prethodno analizirane govorne signale. MOS skor kvaliteta kodovanog signala raste sa povećanjem bitske brzine, dok se predloženim algoritmom za podopsežno kodovanje ostvaruje viši MOS skor za gotovo sve razmatrane bitske brzine. Neprimetnost šuma raste sa povećanjem bitske brzine i možemo primetiti da je kod signala klasične muzike rast krive više blizak linearnom. Takođe kao i kod do sada analiziranih signala, subjektivnim testovima je utvrđeno da predloženi algoritam za podopsežno kodovanje sa dva podopsega unosi manje šuma u odnosu na PCM, za gotovo sve razmatrane bitske brzine.



Slika 6.2.3.5. MOS ocena kvaliteta signala (iznad) i uticaja šuma (ispod) za signal klasične muzike

Nakon prikaza i analize rezultata subjektivnih testova slušanja za sva tri signala, možemo napraviti grubu analizu usklađenosti rezultata subjektivnih testova sa objektivnim merama kvaliteta. Kako prema objektivnim merama kvaliteta algoritam za podopsežno kodovanje ostvaruje dobitak u odnosu na PCM za sve razmatrane bitske brzine, možemo reći da su rezultati korelisani (usklađeni) ukoliko subjektivni testovi

pokažu isti ili približan odnos kvaliteta, odnosno viši kvalitet signala kodovanog predloženim algoritmom za podopsežno kodovanje. Sumiranjem rezultata možemo zaključiti da je MOS skor algoritma za podopsežno kodovanje viši u 72% slučajeva u odnosu na PCM, dok su ispitanici procenili da se algoritmom za podopsežno kodovanje dodaje manje primetnog šuma u 83% slučajeva. Iako analiza nije dovoljno obimna kako bi se izvukli precizniji zaključci, ona je pokazala doprinose predloženog algoritma za podopsežno kodovanje u primeni na govorni i audio signal, time što su i subjektivni testovi pokazali viši kvalitet predloženog algoritma u odnosu na PCM. Takođe izvodi se zaključak da je subjektivna procena uticaja šuma u većoj meri odgovara objektivim merama kvaliteta, nego što je slučaj sa subjektivnom procenom kvaliteta kodovanog signala.

## 7. ZAKLJUČAK

U ovom poglavlju su sumirani i navedeni doprinosi doktorske disertacije:

1. U disertaciji je predložen algoritam za podopsežno kodovanje govornog i audio signala sa dva podopsega. Projektovani su idealni digitalni filtri koji vrše dekompoziciju signala na dva frekventijska podopsega nad kojima se nezavisno primenjuje kvantizacija. Projektovana su dva robusna kvazilogaritamska kvantizera i određena optimalna bit-alokacija koja pruža maksimalno iskorišćenje korišćene ukupne bitske brzine. Eksperimentalno je utvrđen objektivni kvalitet kodovanog signala, koji je pokazao da predloženi algoritam ostvaruje dobitak u SQNR-u od približno 5 dB u odnosu na PCM za istu bitsku brzinu. Pokazano je da se najbolje performanse ostvaruju kada se prvom od dva podopsega dodela bitska brzina veća za dva bita po odmerku u odnosu na drugi podopseg, odnosno da se za ukupnu bitsku brzinu od  $R = 9$  bita po odmerku najbolje performanse ostvaruju kada je  $R_1=10$  i  $R_2=8$ .
2. Predložen je adaptivni algoritam za podopsežno kodovanje govornog i audio signala sa dva podopsega koji vrši podelu signala na frejmove. Predstavljena je primena adaptacije unazad u podopsežnom kodovanju, gde se adaptira amplituda maksimalnog opeterećenja kvantizera na osnovu standardne devijacije prethodno kvantovanog frejma. Dati su izrazi za određivanje standardne devijacije signala podopsega. Prikazan je i metod inicijalizacije algoritma, kako za prvi frejm ne postoje statistički parametri prethodnog frejma. Pokazano je odstupanje adaptivnih amplituda maksimalnog opeterećenja u odnosu na početne vrednosti, što predstavlja prilagođenje algoritma na ulazni signal. Prikazana je objektivna mera kvaliteta kodovanog signala (SQNR) primenom algoritma na realne audio signale. Određena je i optimalna bit-alokacija adaptivnog algoritma za podopsežno kodovanje sa dva podopsega koja za ukupnu bitsku brzinu od  $R = 9$  iznosi je  $R_1=11$  i  $R_2=7$  bita po odmerku. Analizom eksperimentalnih rezultata pokazan je dobitak u SQNR-u, pogotovo izražen za niske faktore kompresije i u slučaju signala klasične muzike iznosi približno 26.6 dB u odnosu na neadaptivni algoritam i 28.25 dB u odnosu na PCM.

3. Predstavljen je algoritam za podopsežno kodovanje audio signala sa tri podopsega. Prikazan je metod primene podopsežnog kodovanja gde se ulazni signal deli najpre na dva podopsega, da bi se nakon toga prvi podopseg idealnih digitalnih NF i VF filtara podelio još na dva podopsega. Dati su izrazi za digitalne filtre za sva tri podopsega, kao i izrazi za određivanje standardne devijacije signala unutar podopsega. Pored toga, prikazana je metoda adaptivnog kodovanja, uz primenu adaptacije unazad na sva tri projektovana kvantizera. Dat je izraz za ukupnu bitSKU brzinu, kao i analiza određivanja optimalne bit-alokacije koja je određena kao  $R_1=13$ ,  $R_2=9$ ,  $R_3=7$ . Uz primenu objektivne bit-alokacije, eksperimentalni rezultati pokazuju dobitak predloženog algoritma u odnosu na PCM u iznosu od maksimalno 10, odnosno 18 dB, za dva signala obuhvaćena eksperimentom.
4. Prikazani su uporedni rezultati objektivne mere kvaliteta kodovanog signala klasične muzike, dobijeni primenom 3 predložena algoritma za podopsežno kodovanje. Analiza je vršena za slučajeve faktora kompresije i bit-alokacije kada svaki od algoritama ostvaruje najbolje performanse. Rezultati su pokazali da složeniji algoritam ostvaruje bolje performanse i razlika u SQNR-u između ne-adaptivnog algoritma za podopsežno kodovanje sa dva podopsega i adaptivnog algoritma za podopsežno kodovanje sa tri podopsega iznosi približno 10.77 dB. Dodatno, poređenjem performansi sa PCM-om za faktor kompresije  $\mu=255$  pokazuje se da algoritmi za podopsežno kodovanje ostvaruju dobitak u SQNR-u u opsegu od 5.6624 do 16.4276 dB.
5. Predstavljen je algoritam za kodovanje govornog signala zasnovan na BTC algoritmu, namenjenom kodovanju crno-bele slike. Predstavljene su primenjene modifikacije i dat blok dijagram modifikovanog algoritma. Algoritam koristi tri različita kvantizera, za koje je predstavljen postupak projektovanja. Prikazana je primena adaptacije unapred i signala razlike koji omogućavaju efikasniju kvantizaciju. Pored toga opisan je i celokupan postupak enkodovanja i dekodovanja signala. Sproveden je eksperiment sa ciljem određivanja performansi predloženog algoritma i predstavljena detaljna analiza rezultata i doprinosa algoritma.



6. Predložen je modifikovani BTC algoritam za kodovanje širokopojasnog govornog signala visokog kvaliteta koji primenjuje podelu ulaznog signala na frejmove i podfrejmove. Dvostruka podela ulaznog signala obezbeđuje efikasniju adaptaciju projektovanih kvantizera na statistiku ulaznog signala. Opisan je postupak adaptacije u kome se koristi srednja vrednost frejma i standardna devijacija podfrejma, kako bi se kvantizer bolje prilagodio statistici signala i efikasno iskoristila raspoloživa bitska brzina. Dobijeni eksperimentalni rezultati pokazuju dobitak u SQNR-u u iznosu od 1.4 do 2 dB u odnosu na slučaj kada se ne vrši podela ulaznog signala na podfrejmove.
7. Doprinosa disertacije leži i u uporednoj analizi subjektivne i objektivne mere kvaliteta kodovanog signala. Pored SQNR-a, kvalitet kodovanog signala je određen procenom objektivne PESQ mere. Uporedna analiza dveju objektivnih mera kvaliteta kodovanog signala pokazala je visoku korelisanost i potvrdila bolji kvalitet kodovanog signala ostvaren primenom predloženog algoritma za podopsežno kodovanje sa dva podopsega u odnosu na PCM.
8. Sprovedeni su testovi slušanja sa ciljem određivanja subjektivne mere kvaliteta kodovanog signala. Dobijeni su i prikazani rezultati MOS (*Mean Opinion Score*) skora za signale čija je objektivna mera kvaliteta prethodno određena. Rezultati su prikazani kroz grafik rasejanja i MOS karakteristiku i izveden je zaključak o vezi između subjektivnih i objektivnih mera kvaliteta kodovanog signala. Iako su dobijeni rezultati subjektivne i objektivne mere korelisani u određenoj meri, uočeno je da subjektivna mera kvaliteta ne ostvaruje linearan rast sa povećanjem bitske brzine, kao što je slučaj sa objektivnim merama kvaliteta.
9. Rezultati prikazani u ovoj disertaciji pokazuju da je predložene algoritme moguće praktično primeniti u kodovanju govornog i audio signala. Pored toga, algoritme je moguće nadgraditi u budućnosti i time povećati ostvarene performanse. Dati uvid u kvalitet kodovanog signala kroz tri različite metode procene kvaliteta predstavlja doprinos koji nije puno viđen u literaturi.

## 8. LITERATURA

- [1] K. Sayood, *Introduction to data compression*, Newnes, Chapter 1, pp. 3-38, 2012.
- [2] A. Kondo, *Digital Speech: Coding for Low Bit Rate Communication Systems*, John Wiley & Sons, 2nd edition, October 20, 2004.
- [3] NS. Jayant, P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*, Prentice Hall, 1st edition, March 1984.
- [4] D. Salomon, *Data Compression – The Complete Reference*, 4th edition, Springer, 2007.
- [5] K. Sayood, *Lossless Compression Handbook*, Academic Press, 2003.
- [6] L. Hanzo, C. Somerville, J. Woodard, *Voice and Audio Compression for Wireless Communications*, John Wiley & Sons - IEEE Press, 2nd edition, October 8, 2007.
- [7] ITU-T, Recommendation G.711, *Pulse Code Modulation (PCM) of Voice Frequencies*, International Telecommunication Union, 1972.
- [8] ITU-T, Recommendation G.711.1, *Wideband embedded extension for ITU-T G.711 pulse code modulation*, International Telecommunication Union, 2008.
- [9] ITU-T, Recommendation G.712. (2001). *Transmission Performance Characteristics of Pulse Code Modulation Channels*.
- [10] A.W. Rix, J.G. Beerends, M.P. Hollier, A. P. Hekstra. "Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs." In *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP'01). 2001 IEEE International Conference on*, vol. 2, pp. 749-752. IEEE, 2001.
- [11] J. Mitchell, Introduction to digital audio coding and standards. *Journal of Electronic Imaging*, 13(2), 399.
- [12] Z. Peric, M. Tancic, S. Tomic, D. Ciric, Subband coding of audio signal with logarithmic companders. 12th International Conference on Telecommunication in Modern Satellite, Cable and Broadcasting Services (TELSIKS).
- [13] S. Tomic, Z. Peric, M. Tancic, J. Nikolic, "Backward Adaptive and Quasi-Logarithmic Quantizer for Sub-Band Coding of Audio", *Journal of Information Technology and Control*, vol. 47, no. 1, 2018, pp. 131-139,

- [14] S. Tomić , M. Tančić, D. Ćirić, Z. Perić, “Effects of subband coding with quasilogarithmic quantizers of characteristics of audio signal” Proceedings of 2nd International Conference on Electrical, Electronic and Computing Engineering IcETRAN 2015, Silver Lake, Serbia, June 8 – 11, 2015, ISBN 978-86-80509-71-6, pp. AK11.4.1-5.
- [15] E. J. Delp, O. R. Mithcell, “Image compression using block truncation coding”, *IEEE Trans. Commun.*, vol. 27 no. 9, pp. 1335-1342, 1979. doi: 10.1109/tcom.1979.1094560.
- [16] S. Tomic, Z. Peric, J. Nikolic, Wideband Speech Signal Coding with the Implementation of Modified BTC Algorithm, 52nd International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST), 2017, Nis, Serbia.
- [17] S. Tomic, Zoran Peric, Jelena Nikolic, Wideband Speech Signal Coding with the Implementation of Modified BTC Algorithm, 52nd International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST), 2017, Nis, Serbia.
- [18] NS. Jayant, S. Nuggehally S, “Digital coding of speech waveforms: PCM, DPCM, and DM quantizers.” *Proceedings of the IEEE*62, no. 5 (1974): 611-632.
- [19] V. Despotovic, Z. Peric, L. Velimirovic, V. Delic. "DPCM with forward gain-adaptive quantizer and simple switched predictor for high quality speech signals.", *Advances in Electrical and Computer Engineering*, vol.10, no. 4 pp/ 95-98, 2010.
- [20] Z. Peric, D. Denic, J. Nikolic, A. Jovic, A. Jovanovic, "DPCM quantizer adaptation method for efficient ECG signal compression.", *Journal of Communications Technology and Electronics* vol. 58, no. 12, pp. 1241-1250, 2013.
- [21] W. C. Chu, Speech coding algorithms, John Wiley & Sons, New Jersey, 2003, Chapters 5-6, pp. 143-183.
- [22] D. Salomon, *Data compression: the complete reference*. Springer Science & Business Media, 2004.
- [23] H. S. Black, *Modulation Theory*. Princeton, N. J.: Van Nostrand, 1953.
- [24] M. Carasso, J. Peek, J. Sinjou, „The compact disc digital audio system”, *Philips Tech. Rev*, 40(6), 151, 1987.

- [25] M. Hasegawa-Johnson, A. Alwan, *Speech coding: fundamentals and applications. Encyclopedia of Telecommunications*, 2003.
- [26] S. Uddin, I.R. Ansari, S. Naaz, “Low Bit Rate Speech Coding Using Differential Pulse Code Modulation”, *Advances in Research, SCIENCEDOMAIN International*, 8(3), pp.2348-0394, 2016.
- [27] M.L. George, G.S. Tomar, T. Sharma, T. B. Kumar, I. Berry, „Hardware Implementation of Adaptive-Differential Pulse Code Modulation Speech Compression Algorithm”, *International Journal of Smart Device and Appliance*, 2(2), pp.1-10, 2014.
- [28] ITU-T, Recommendation G.721, *32 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)*, International Telecommunication Union, 1984.
- [29] ITU-T, Recommendation G.726, *40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)*, International Telecommunication Union, 1990.
- [30] ITU-T, Recommendation G.727, *5-, 4-, 3- and 2-bit/sample Embedded Adaptive Differential Pulse Code Modulation (ADPCM)*, International Telecommunication Union, 1990.
- [31] ITU-T, Recommendation G.722, *7 kHz Audio-coding within 64 kbit/s*, International Telecommunication Union, 1988.
- [32] ITU-T, Recommendation G.722.1, *Low-complexity Coding at 24 and 32 kbit/s for Hands-free Operation in Systems with Low Frame Loss*, International Telecommunication Union, 1999.
- [33] ITU-T, Recommendation G.722.2, *Wideband Coding of Speech at Around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)*, International Telecommunication Union, 2003.
- [34] A. Spanias, T. Painter, V. Atti, *Audio Signal Processing and Coding*, Wiley-Interscience, February 9, 2007.
- [35] GSM 06.60, “GSM Digital Cellular Communication Standards: Enhanced Full-Rate Transcoding,” ETSI/GSM, 1996.
- [36] GSM 06.20, “GSM Digital Cellular Communication Standards: Half Rate Speech; Half Rate Speech Transcoding,” ETSI/GSM, 1996.

- [37] Sinder, I. Varga, V. Krishnan, V. Rajendran, S. Villette, "Recent speech coding technologies and standards." *Speech and Audio Processing for Coding, Enhancement and Recognition*, pp. 75-109. Springer, New York, NY, 2015.
- [38] B. Bessette, et al. The adaptive multirate wideband speech codec (AMR-WB). *IEEE Trans. on Speech and Audio Processing*, 2002, 10, pp. 620-636.
- [39] M. Jelinek, R. Salami, R. Wideband Speech Coding Advances in VMR-WB Standard. *IEEE Trans. on Audio, Speech and Language Processing*, 2007, 15, pp. 1167-1179.
- [40] M. Dietz, et al. Overview of the EVS Codec Architecture. *Proc. of the Int. Conf. on Acoust. Speech and Signal 782 Processing*, USA, 2015, pp. 5698-5702.
- [41] L. Chiariglione. The development an integrated audio-visual coing standard, *Proceedings of the IEEE*, Vol. 83 n. 2, pp. 1 5 1 - 1 57, February 1995.
- [42] T. Liebchen, Y. A. Reznik "MPEG-4 ALS: An emerging standard for lossless audio coding", In *Data Compression Conference, 2004. Proceedings. DCC 2004* (pp. 439-448). IEEE.
- [43] T. Liebchen, "MPEG-4 ALS-the standard for lossless audio coding." *the Journal of the Acoustical Society of Korea* 28, no. 7 (2009): 618-629.
- [44] P. Thang. Four-channel sub-band coding with combination of [8 8 4 2] for digital audio signal processing. *Science Journal of Circuits, Systems and Signal Processing*, 2013, Vol. 2, No. 5, 106-111.
- [45] A. Aziz. Subband coding of speech signals using decimation and interpolation. In: *13th international conference on aerospace sciences and aviation technology, ASAT-13*, Cairo (Egypt). 2009, pp. 26-28.
- [46] L. Naik, D. Rl. Sub-band coding of noisy speech signals using digital signal processing. *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)*, 2015, Vol. 4, No. 4, 901-904.
- [47] H. Gharavi, A. Tabatabai. Sub-band coding of monochrome and colour images. *IEEE Journal of Circuits and Systems*, 1988, Vol. 35, No: 2, 207-214.
- [48] D. Taubman, A. Zakhor. Multirate 3-D subband coding of video. *IEEE Journal of Image Processing*, 1994, Vol. 3, No: 5, 572-588.

- [49] J. Nikolic, Z. Perić, “Lloyd-Max’s algorithm implementation in speech coding algorithm based on forward adaptive technique”, *Informatica*, 2008, vol. 19, no. 2, pp. 255–270.
- [50] Bennett, “Spectra of Quantized Signals”, *The Bell Systems Technical Journal*, Vol. 27, pp. 446 - 472, July, 1948.
- [51] Allen Gersho, “Quantization”, *IEEE Communications Society Magazine*, pp. 20 - 29, September, 1977.
- [52] Z. Perić, M. Savić, M. Dinčić, D. Denić, M. Prašcević, “Forward Adaptation of Novel Semilogarithmic Quantizer and Lossless Coder for Speech Signals Compression”, *Informatica*, Vol. 21, No. 3, pp. 375-391, 2010.
- [53] S. Agrawal, O. Sahu, “Polyphase structure based eigen design of two-channel quadrature mirror filter bank”, *International Journal of Image, Graphics and Signal Processing (IJIGSP)*, 2014, Vol. 6, No. 10, pp 29-35.
- [54] S. Tan, T. Fischer, Linear prediction of subband signals. *IEEE Journal on Selected Areas in Communications*, 1994, 12(9), 1576–1583. doi: 10.1109/49.339926
- [55] L. Hui, "An adaptive block truncation coding algorithm for image compression." *In Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990 International Conference on*, pp. 2233-2236. IEEE, 1990.
- [56] D. Halverson, N. Griswold, G. Wise, A generalized block truncation coding algorithm for image compression. *IEEE transactions on acoustics, speech, and signal processing*, 32(3), pp.664-668, 1984.
- [57] E. Oshri, N. Shelly, H. B. Mitchell, "Interpolative three-level block truncation coding algorithm." *Electronics Letters*, vol.29, no.14, pp. 1267-1268, 1993.
- [58] C. Wang, C. Chin-Hsing, C. I-Hong Chen, "Modified VQ-BTC algorithm for image compression." *Electronics Letters*, vol. 34, no. 14, pp. 1390-1391, 1998.
- [59] A. Kumar, S. Pardeep, "Enhanced Block Truncation Coding for Gray Scale Image." *Int. J. Comput. Techn. Appl* vol. 2, no. 3, pp. 525-530, 2011.
- [60] Y. Yang, C. Qiqiang, W. Yi, "A fast near-optimum block truncation coding method using a truncated K-means algorithm and inter-block correlation." *AEU-International Journal of Electronics and Communications* vol. 65, no. 6, pp. 576-581, 2011.

- [61] S. Vimala, P. Uma, B. Abidha, "Improved adaptive block truncation coding for image compression." *International Journal of Computer Applications* vol. 19, no. 7, pp. 975-988, 2011.
- [62] G. Rozinaj, S. Herrera, and J. Mikula. "Modified BTC algorithm for satellite image coding." In *Video/Image Processing and Multimedia Communications 4th EURASIP-IEEE Region 8 International Symposium on VIPromCom*, pp. 212-215. IEEE, 2002.
- [63] B. C. Dhara, C. Bhabatosh, "Color image compression based on block truncation coding using pattern fitting principle.", *Pattern Recognition*, vol. 40, no. 9, pp 2408-2417, 2007.
- [64] M. Lema, O. Mitchell, "Absolute moment block truncation coding and its application to color images." *IEEE Transactions on communications* vol. 32, no. 10 pp. 1148-1157, 1984.
- [65] Z. Perić, M. Petković, M. Dinčić, "Simple compression algorithm for memoryless Laplacian source based on the optimal companding technique", *Informatica*, vol. 20, no. 1, pp. 99-114, 2009.
- [66] N. Judell, L. Scharf, "A simple derivation of Lloyd's classical result for the optimum scalar quantizer", *IEEE Trans. Inform. Theory*, vol. 32, no. 2, pp. 326-328, 1986. doi: 10.1109/tit.1986.1057150.
- [67] R. Farias, J. Brossier, "Adaptive quantizers for estimation", *Signal Processing*, vol. 93, no. 11, pp. 3076-3087, 2013. doi: 10.1016/j.sigpro.2013.04.012
- [68] S. Na, D. Neuhoff, "On the support of MSE-optimal, fixed-rate, scalar quantizers", *IEEE Transactions on information theory*, vol. 47 no. 7, pp. 2972-2982, 2001.
- [69] Z. Peric, D. Aleksic, M. Stefanovic, J. Nikolic, "New Approach to Support Region Determination of the  $\mu$ -law Quantizer"; *Elektronika ir Elektrotechnika*, vol. 19 no. 8, pp. 111-114, 2013.
- [70] M. N. Suma, "Performance Analysis of DPCM and ADPCM", In *International Conference on Electronic Design and Signal Processing (ICEDSP)*, pp. 19-23, 2012.
- [71] Z. Perić, J. Nikolić, A. Mosić, M. Petković, "Design of fixed and adaptive companding quantizer with variable-length codeword for memoryless Gaussian source", *Informatica*, vol. 24, no. 1, pp. 71-86, 2013.

- [72] M. Tancic, Z. Peric, S. Tomic, N. Simic, "Speech Signal Coding Using Forward Adaptive Quantization and Simple Transform Coding, *Elektronika ir Elektrotechnika*, vol. 22, no. 3, pp. 74-77, 2016.
- [73] A. W. Rix, M. P. Hollier, A. P. Hekstra, J. G. Beerends, "Perceptual Evaluation of Speech Quality (PESQ) The New ITU Standard for End-to-End Speech Quality Assessment Part I--Time-Delay Compensation.", *Journal of the Audio Engineering Society*, vol. 50 no. 10, 755-764, 2002.
- [74] J. G. Beerends, A. P. Hekstra, A. W. Rix, M. P. Hollier, "Perceptual evaluation of speech quality (PESQ) the new ITU standard for end-to-end speech quality assessment part ii: psychoacoustic model.", *Journal of the Audio Engineering Society*, vol. 50 no.10, pp. 765-778, 2002.
- [75] ITU-T, Recommendation P.862, Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs, International Telecommunication Union, 2001.
- [76] ITU-T Recommendation P.862.2. Wideband extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs, International Telecommunication Union, 2007.
- [77] T. Thilo, W. C. Treurniet, R. Bitto, C. Schmidmer, T. Sporer, J. G. Beerends, C. Colomes, "PEAQ-The ITU standard for objective measurement of perceived audio quality." *Journal of the Audio Engineering Society*, vol. 48, no. ½, pp. 3-29, 2000.
- [78] ITU-T Rec. P.800, "Methods for Subjective Determination of Transmission Quality," International Telecommunication Union, Geneva, Switzerland, 1996.
- [79] ITU-T Rec. P.830, "Subjective performance assessment of telephone-band and wideband digital codecs," International Telecommunication Union, Geneva, Switzerland, 1996.
- [80] K. Wilson, B. Raj, P. Smaragdis, A. Divakaran. "Speech denoising using nonnegative matrix factorization with priors." In *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pp. 4029-4032. IEEE, 2008.
- [81] X. Lu, Y. Tsao, S. Matsuda, C. Hori, "Speech enhancement based on deep denoising autoencoder." In *Interspeech 2013*, pp. 436-440, 2013.



- [82] M. Russo, M. Stella, M. Kurajica, "Cochlear Model based Enhancement of Noisy Speech Signals." *International Journal of Circuits, Systems and Signal Processing* vol. 9, pp. 446-451, 2015.
- [83] J. G. Beerends, E. Larsen, I. Nandini, J. M. van Vugt. "Measurement of speech intelligibility based on the PESQ approach." *Measurement of Speech and Audio Quality in Networks (MESAQIN)*, 2004.
- [84] J. Gibson, "Speech compression.", *Information*, vol. 7, no. 2. pp. 32-54, 2016.
- [85] ITU-T Rec. P.10, "Vocabulary for performance and quality of service", International Telecommunication Union, 2006
- [86] ITU-T Recommendation P.800, "Methods for subjective determination of transmission quality", International Telecommunication Union, 1996.
- [87] ITU-T Rec. P.911, "Subjective audiovisual quality assessment methods for multimedia applications", International Telecommunication Union, 1998.
- [88] A. P. Frøyland, "Statistical approach to the rational selection of experimental subjects", *Accreditation and Quality Assurance*, vol. 10, no. 5, pp. 185-189, 2004.

## 9. SPISAK REFERENCI AUTORA

1. **Stefan Tomić**, Zoran Perić, Jelena Nikolić: „Modified BTC Algorithm for Audio Signal Coding.” *Advances in Electrical and Computer Engineering*, (2016), vol. 16, no. 4, pp. 31-38, M23.
2. **Stefan Tomic**, Zoran Peric, Milan Tancic, Jelena Nikolic, Backward Adaptive and Quasi-Logarithmic Quantizer for Sub-Band Coding of Audio, *Information Technology and Control*, vol. 47, no. 1, 2018, pp. 131-139, doi:10.5755/j01.itc.47.1.16190, M23.
3. Milan Tančić, Zoran Perić, **Stefan Tomić**, Nikola Simić: „Speech Signal Coding Using Forward Adaptive Quantization and Simple Transform Coding”, *Elektronika Ir Elektrotehnika*, (2016), vol. 22 br. 3, str. 74-77, M23.
4. Milan Tančić, Zoran Perić, **Stefan Tomić**, “Performance of Quasi-logarithmic Quantizer for Discrete Input Signal”, *Information Technology and Control*, 2017, 46(3), 395-402, M23.
5. Milan Tančić, Zoran Perić, Aleksandra Jovanović, **Stefan Tomić**: „Determining compression factor of quasilogarithmic quantizer for Laplacian source in narrow dynamic variance range”, *Facta Universitatis, Series: Automatic Control and Robotics*, (2016), vol. 15, br. 3, str. 217-226, M24.
6. **Stefan Tomić**, Zoran Perić, Milan Tančić: Forward Mean-Adaptive Quasilogarithmic Quantizer for Coding of Correlated Sources”, *Facta Universitatis, Series: Automatic Control and Robotics*, (2016), vol. 15, br.3, str. 205-215, M24.
7. **Stefan Tomic**, Zoran Peric, Jelena Nikolic, Wideband Speech Signal Coding with the Implementation of Modified BTC Algorithm, 52nd International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST), 2017, Nis, Serbia, M34.
8. Zoran Perić, Milan Tančić, **Stefan Tomić**, D. Ćirić, “Subband coding of audio signal with logarithmic compandors”, *International Conference on Advanced Technologies, Systems and Services in Telecommunications - TELSISKS*, Niš, Serbia, October 14 - 17, 2015, pp. 19-22, M33.
9. Zoran Perić, **Stefan Tomić**, Milan Tančić, “High quality speech signal coding with the application of BTC algorithm”, *International Conference on Advanced Technologies, Systems and Services in Telecommunications - TELSISKS*, Nis, Serbia, October 2015, pp. 23-26, M33.
10. **Stefan Tomić** , Milan Tančić, Dejan Ćirić, Zoran Perić, „Effects of subband coding with quasilogarithmic quantizers of characteristics of audio signal” *Proceedings of 2nd International Conference on Electrical, Electronic and*

Computing Engineering IcETTRAN 2015, Silver Lake, Serbia, June 8 – 11, 2015, ISBN 978-86-80509-71-6, pp. AK11.4.1-5, M33.

11. Vera Marković, **Stefan Tomić**, Zlatica Marinković, “Trendovi u primeni indoor sistema u mobilnim komunikacijama”, XXXII Simpozijum o novim tehnologijama u poštanskom i telekomunikacionom saobraćaju- POSTEL, Belgrade, Serbia, December 2-3. 2014, pp. 199-208, M63.
12. Јелена Здравковић, Дејан Ћирић, Марко Јеленковић, **Стефан Томић**, “Добијање различитих карактеристика усмерености зрачења звучног извора са независним елементима у облику додекаедра”, *Зборник радова LVI Конференције ЕТРАН*, Златибор, Србија, 11-14 јун 2012, издавач: Друштво за ЕТРАН, стр. АК2.2-1-АК2.2-4 (на CD-у, рад АК2.2), 2012, ISBN 978-86-80509-67-9, M63.
13. Marko Jelenković, Dejan Ćirić, Jelena Zdravković, **Stefan Tomić**, “Control of radiation directivity applying independent element dodecahedral loudspeaker”, *Proceedings of XLVII International Scientific Conference on Information, Communication, Energy Systems and Technologies - ICEST 2012*, Veliko Tarnovo, Bulgaria, 28-30 June 2012, Publisher: Faculty of Telecommunications, Technical University of Sofia, pp. 131-134, 2012, ISBN 978-619-167-002-4, M34.

## LITERATURA

- [1] K. Sayood, *Introduction to data compression*, Newnes, Chapter 1, pp. 3-38, 2012.
- [2] A. Kondo, *Digital Speech: Coding for Low Bit Rate Communication Systems*, John Wiley & Sons, 2nd edition, October 20, 2004.
- [3] NS. Jayant, P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*, Prentice Hall, 1st edition, March 1984.
- [4] D. Salomon, *Data Compression – The Complete Reference*, 4th edition, Springer, 2007.
- [5] K. Sayood, *Lossless Compression Handbook*, Academic Press, 2003.
- [6] L. Hanzo, C. Somerville, J. Woodard, *Voice and Audio Compression for Wireless Communications*, John Wiley & Sons - IEEE Press, 2nd edition, October 8, 2007.
- [7] ITU-T, Recommendation G.711, *Pulse Code Modulation (PCM) of Voice Frequencies*, International Telecommunication Union, 1972.
- [8] ITU-T, Recommendation G.711.1, *Wideband embedded extension for ITU-T G.711 pulse code modulation*, International Telecommunication Union, 2008.
- [9] ITU-T, Recommendation G.712. (2001). *Transmission Performance Characteristics of Pulse Code Modulation Channels*.
- [10] A.W. Rix, J.G. Beerends, M.P. Hollier, A. P. Hekstra. "Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs." In *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP'01). 2001 IEEE International Conference on*, vol. 2, pp. 749-752. IEEE, 2001.
- [11] J. Mitchell, Introduction to digital audio coding and standards. *Journal of Electronic Imaging*, 13(2), 399.
- [12] Z. Peric, M. Tancic, S. Tomic, D. Ciric, Subband coding of audio signal with logarithmic companders. 12th International Conference on Telecommunication in Modern Satellite, Cable and Broadcasting Services (TELSIKS).
- [13] S. Tomic, Z. Peric, M. Tancic, J. Nikolic, "Backward Adaptive and Quasi-Logarithmic Quantizer for Sub-Band Coding of Audio", *Journal of Information Technology and Control*, vol. 47, no. 1, 2018, pp. 131-139,
- [14] S. Tomić , M. Tančić, D. Ćirić, Z. Perić, "Effects of subband coding with quasilogarithmic quantizers of characteristics of audio signal" Proceedings of 2nd

International Conference on Electrical, Electronic and Computing Engineering IcETRAN 2015, Silver Lake, Serbia, June 8 – 11, 2015, ISBN 978-86-80509-71-6, pp. AKI1.4.1-5.

[15] E. J. Delp, O. R. Mithcell, "Image compression using block truncation coding", *IEEE Trans. Commun.*, vol. 27 no. 9, pp. 1335-1342, 1979. doi: 10.1109/tcom.1979.1094560.

[16] S. Tomic, Z. Peric, J. Nikolic, Wideband Speech Signal Coding with the Implementation of Modified BTC Algorithm, 52nd International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST), 2017, Nis, Serbia.

[17] S. Tomic, Zoran Peric, Jelena Nikolic, Wideband Speech Signal Coding with the Implementation of Modified BTC Algorithm, 52nd International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST), 2017, Nis, Serbia.

[18] NS. Jayant, S. Nuggehally S, "Digital coding of speech waveforms: PCM, DPCM, and DM quantizers." *Proceedings of the IEEE* 62, no. 5 (1974): 611-632.

[19] V. Despotovic, Z. Peric, L. Velimirovic, V. Delic. "DPCM with forward gain-adaptive quantizer and simple switched predictor for high quality speech signals.", *Advances in Electrical and Computer Engineering*, vol.10, no. 4 pp/ 95-98, 2010.

[20] Z. Peric, D. Denic, J. Nikolic, A. Jovic, A. Jovanovic, "DPCM quantizer adaptation method for efficient ECG signal compression.", *Journal of Communications Technology and Electronics* vol. 58, no. 12, pp. 1241-1250, 2013.

[21] W. C. Chu, *Speech coding algorithms*, John Wiley & Sons, New Jersey, 2003, Chapters 5-6, pp. 143-183.

[22] D. Salomon, *Data compression: the complete reference*. Springer Science & Business Media, 2004.

[23] H. S. Black, *Modulation Theory*. Princeton, N. J.: Van Nostrand, 1953.

[24] M. Carasso, J. Peek, J. Sinjou, „The compact disc digital audio system”, *Philips Tech. Rev.*, 40(6), 151, 1987.

[25] M. Hasegawa-Johnson, A. Alwan, *Speech coding: fundamentals and applications. Encyclopedia of Telecommunications*, 2003.

[26] S. Uddin, I.R. Ansari, S. Naaz, "Low Bit Rate Speech Coding Using Differential Pulse Code Modulation", *Advances in Research, SCIENCEDOMAIN International*, 8(3), pp.2348-0394, 2016.

- [27] M.L. George, G.S. Tomar, T. Sharma, T. B. Kumar, I. Berry, „Hardware Implementation of Adaptive-Differential Pulse Code Modulation Speech Compression Algorithm”, *International Journal of Smart Device and Appliance*, 2(2), pp.1-10, 2014.
- [28] ITU-T, Recommendation G.721, *32 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)*, International Telecommunication Union, 1984.
- [29] ITU-T, Recommendation G.726, *40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)*, International Telecommunication Union, 1990.
- [30] ITU-T, Recommendation G.727, *5-, 4-, 3- and 2-bit/sample Embedded Adaptive Differential Pulse Code Modulation (ADPCM)*, International Telecommunication Union, 1990.
- [31] ITU-T, Recommendation G.722, *7 kHz Audio-coding within 64 kbit/s*, International Telecommunication Union, 1988.
- [32] ITU-T, Recommendation G.722.1, *Low-complexity Coding at 24 and 32 kbit/s for Hands-free Operation in Systems with Low Frame Loss*, International Telecommunication Union, 1999.
- [33] ITU-T, Recommendation G.722.2, *Wideband Coding of Speech at Around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)*, International Telecommunication Union, 2003.
- [34] A. Spanias, T. Painter, V. Atti, *Audio Signal Processing and Coding*, Wiley-Interscience, February 9, 2007.
- [35] GSM 06.60, “GSM Digital Cellular Communication Standards: Enhanced Full-Rate Transcoding,” ETSI/GSM, 1996.
- [36] GSM 06.20, “GSM Digital Cellular Communication Standards: Half Rate Speech; Half Rate Speech Transcoding,” ETSI/GSM, 1996.
- [37] Sinder, I. Varga, V. Krishnan, V. Rajendran, S. Villette, “Recent speech coding technologies and standards.” *Speech and Audio Processing for Coding, Enhancement and Recognition*, pp. 75-109. Springer, New York, NY, 2015.
- [38] B. Bessette, et al. The adaptive multirate wideband speech codec (AMR-WB). *IEEE Trans. on Speech and Audio Processing*, 2002, 10, pp. 620-636.
- [39] M. Jelinek, R. Salami, R. Wideband Speech Coding Advances in VMR-WB Standard. *IEEE Trans. on Audio, Speech and Language Processing*, 2007, 15, pp. 1167-1179.
- [40] M. Dietz, et al. Overview of the EVS Codec Architecture. *Proc. of the Int. Conf. on Acoust. Speech and Signal 782 Processing*, USA, 2015, pp. 5698-5702.

- [41] L. Chiariglione. The development an integrated audio-visual coding standard, *Proceedings of the IEEE*, Vol. 83 n. 2, pp. 151 - 157, February 1995.
- [42] T. Liebchen, Y. A. Reznik “MPEG-4 ALS: An emerging standard for lossless audio coding”, In *Data Compression Conference, 2004. Proceedings. DCC 2004* (pp. 439-448). IEEE.
- [43] T. Liebchen, “MPEG-4 ALS-the standard for lossless audio coding.” *the Journal of the Acoustical Society of Korea* 28, no. 7 (2009): 618-629.
- [44] P. Thang. Four-channel sub-band coding with combination of [8 8 4 2] for digital audio signal processing. *Science Journal of Circuits, Systems and Signal Processing*, 2013, Vol. 2, No. 5, 106-111.
- [45] A. Aziz. Subband coding of speech signals using decimation and interpolation. In: *13th international conference on aerospace sciences and aviation technology, ASAT-13*, Cairo (Egypt). 2009, pp. 26-28.
- [46] L. Naik, D. Rl. Sub-band coding of noisy speech signals using digital signal processing. *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)*, 2015, Vol. 4, No. 4, 901-904.
- [47] H. Gharavi, A. Tabatabai. Sub-band coding of monochrome and colour images. *IEEE Journal of Circuits and Systems*, 1988, Vol. 35, No: 2, 207–214.
- [48] D. Taubman, A. Zakhor. Multirate 3-D subband coding of video. *IEEE Journal of Image Processing*, 1994, Vol. 3, No: 5, 572-588.
- [49] J. Nikolic, Z. Perić, “Lloyd-Max’s algorithm implementation in speech coding algorithm based on forward adaptive technique”, *Informatica*, 2008, vol. 19, no. 2, pp. 255–270.
- [50] Bennett, “Spectra of Quantized Signals”, *The Bell Systems Technical Journal*, Vol. 27, pp. 446 - 472, July, 1948.
- [51] Allen Gersho, “Quantization”, *IEEE Communications Society Magazine*, pp. 20 - 29, September, 1977.
- [52] Z. Perić, M. Savić, M. Dinčić, D. Denić, M. Prašćević, “Forward Adaptation of Novel Semilogarithmic Quantizer and Lossless Coder for Speech Signals Compression”, *Informatica*, Vol. 21, No. 3, pp. 375-391, 2010.

- [53] S. Agrawal, O. Sahu, "Polyphase structure based eigen design of two-channel quadrature mirror filter bank", *International Journal of Image, Graphics and Signal Processing (IJIGSP)*, 2014, Vol. 6, No. 10, pp 29-35.
- [54] S. Tan, T. Fischer, Linear prediction of subband signals. *IEEE Journal on Selected Areas in Communications*, 1994, 12(9), 1576–1583. doi: 10.1109/49.339926
- [55] L. Hui, "An adaptive block truncation coding algorithm for image compression." In *Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990 International Conference on*, pp. 2233-2236. IEEE, 1990.
- [56] D. Halverson, N. Griswold, G. Wise, A generalized block truncation coding algorithm for image compression. *IEEE transactions on acoustics, speech, and signal processing*, 32(3), pp.664-668, 1984.
- [57] E. Oshri, N. Shelly, H. B. Mitchell, "Interpolative three-level block truncation coding algorithm." *Electronics Letters*, vol.29, no.14, pp. 1267-1268, 1993.
- [58] C. Wang, C. Chin-Hsing, C. I-Hong Chen, "Modified VQ-BTC algorithm for image compression." *Electronics Letters*, vol. 34, no. 14, pp. 1390-1391, 1998.
- [59] A. Kumar, S. Pardeep, "Enhanced Block Truncation Coding for Gray Scale Image." *Int. J. Comput. Techn. Appl* vol. 2, no. 3, pp. 525-530, 2011.
- [60] Y. Yang, C. Qiqiang, W. Yi, "A fast near-optimum block truncation coding method using a truncated K-means algorithm and inter-block correlation." *AEU-International Journal of Electronics and Communications* vol. 65, no. 6, pp. 576-581, 2011.
- [61] S. Vimala, P. Uma, B. Abidha, "Improved adaptive block truncation coding for image compression." *International Journal of Computer Applications* vol. 19, no. 7, pp. 975-988, 2011.
- [62] G. Rozinaj, S. Herrera, and J. Mikula. "Modified BTC algorithm for satellite image coding." In *Video/Image Processing and Multimedia Communications 4th EURASIP-IEEE Region 8 International Symposium on VIPromCom*, pp. 212-215. IEEE, 2002.
- [63] B. C. Dhara, C. Bhabatosh, "Color image compression based on block truncation coding using pattern fitting principle.", *Pattern Recognition*, vol. 40, no. 9, pp 2408-2417, 2007.
- [64] M. Lema, O. Mitchell, "Absolute moment block truncation coding and its application to color images." *IEEE Transactions on communications* vol. 32, no. 10 pp. 1148-1157, 1984.



- [65] Z. Perić, M. Petković, M. Dinčić, “Simple compression algorithm for memoryless Laplacian source based on the optimal companding technique”, *Informatica*, vol. 20, no. 1, pp. 99-114, 2009.
- [66] N. Judell, L. Scharf, “A simple derivation of Lloyd's classical result for the optimum scalar quantizer”, *IEEE Trans. Inform. Theory*, vol. 32, no. 2, pp. 326-328, 1986. doi: 10.1109/tit.1986.1057150.
- [67] R. Farias, J. Brossier, “Adaptive quantizers for estimation”, *Signal Processing*, vol. 93, no. 11, pp. 3076-3087, 2013. doi: 10.1016/j.sigpro.2013.04.012
- [68] S. Na, D. Neuhoff, “On the support of MSE-optimal, fixed-rate, scalar quantizers”, *IEEE Transactions on information theory*, vol. 47 no. 7, pp. 2972-2982, 2001.
- [69] Z. Peric, D. Aleksic, M. Stefanovic, J. Nikolic, “New Approach to Support Region Determination of the  $\mu$ -law Quantizer”; *Elektronika ir Elektrotechnika*, vol. 19 no. 8, pp. 111-114, 2013.
- [70] M. N. Suma, “Performance Analysis of DPCM and ADPCM”, In *International Conference on Electronic Design and Signal Processing (ICEDSP)*, pp. 19-23, 2012.
- [71] Z. Perić, J. Nikolić, A. Mosić, M. Petković, “Design of fixed and adaptive companding quantizer with variable-length codeword for memoryless Gaussian source”, *Informatica*, vol. 24, no. 1, pp. 71-86, 2013.
- [72] M. Tancic, Z. Peric, S. Tomic, N. Simic, “Speech Signal Coding Using Forward Adaptive Quantization and Simple Transform Coding, *Elektronika ir Elektrotechnika*, vol. 22, no. 3, pp. 74-77, 2016.
- [73] A. W. Rix, M. P. Hollier, A. P. Hekstra, J. G. Beerends, “Perceptual Evaluation of Speech Quality (PESQ) The New ITU Standard for End-to-End Speech Quality Assessment Part I--Time-Delay Compensation.”, *Journal of the Audio Engineering Society*, vol. 50 no. 10, 755-764, 2002.
- [74] J. G. Beerends, A. P. Hekstra, A. W. Rix, M. P. Hollier, “Perceptual evaluation of speech quality (PESQ) the new ITU standard for end-to-end speech quality assessment part ii: psychoacoustic model.”, *Journal of the Audio Engineering Society*, vol. 50 no.10, pp. 765-778, 2002.
- [75] ITU-T, Recommendation P.862, Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs, International Telecommunication Union, 2001.

- [76] ITU-T Recommendation P.862.2. Wideband extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs, International Telecommunication Union, 2007.
- [77] T. Thilo, W. C. Treurniet, R. Bitto, C. Schmidmer, T. Sporer, J. G. Beerends, C. Colomes, "PEAQ-The ITU standard for objective measurement of perceived audio quality." *Journal of the Audio Engineering Society*, vol. 48, no. ½, pp. 3-29, 2000.
- [78] ITU-T Rec. P.800, "Methods for Subjective Determination of Transmission Quality," International Telecommunication Union, Geneva, Switzerland, 1996.
- [79] ITU-T Rec. P.830, "Subjective performance assessment of telephone-band and wideband digital codecs," International Telecommunication Union, Geneva, Switzerland, 1996.
- [80] K. Wilson, B. Raj, P. Smaragdis, A. Divakaran. "Speech denoising using nonnegative matrix factorization with priors." In *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pp. 4029-4032. IEEE, 2008.
- [81] X. Lu, Y. Tsao, S. Matsuda, C. Hori, "Speech enhancement based on deep denoising autoencoder." In *Interspeech 2013*, pp. 436-440, 2013.
- [82] M. Russo, M. Stella, M. Kurajica, "Cochlear Model based Enhancement of Noisy Speech Signals." *International Journal of Circuits, Systems and Signal Processing* vol. 9, pp. 446-451, 2015.
- [83] J. G. Beerends, E. Larsen, I. Nandini, J. M. van Vugt. "Measurement of speech intelligibility based on the PESQ approach." *Measurement of Speech and Audio Quality in Networks (MESAQIN)*, 2004.
- [84] J. Gibson, "Speech compression.", *Information*, vol. 7, no. 2. pp. 32-54, 2016.
- [85] ITU-T Rec. P.10, "Vocabulary for performance and quality of service", International Telecommunication Union, 2006
- [86] ITU-T Recommendation P.800, "Methods for subjective determination of transmission quality", International Telecommunication Union, 1996.
- [87] ITU-T Rec. P.911, "Subjective audiovisual quality assessment methods for multimedia applications", International Telecommunication Union, 1998.
- [88] A. P. Frøyland, "Statistical approach to the rational selection of experimental subjects", *Accreditation and Quality Assurance*, vol. 10, no. 5, pp. 185-189, 2004.

## SPISAK REFERENCI AUTORA

1. **Stefan Tomić**, Zoran Perić, Jelena Nikolić: „Modified BTC Algorithm for Audio Signal Coding.” *Advances in Electrical and Computer Engineering*, (2016), vol. 16, no. 4, pp. 31-38, M23.
2. **Stefan Tomic**, Zoran Peric, Milan Tancic, Jelena Nikolic, Backward Adaptive and Quasi-Logarithmic Quantizer for Sub-Band Coding of Audio, *Information Technology and Control*, vol. 47, no. 1, 2018, pp. 131-139, doi:10.5755/j01.itc.47.1.16190, M23.
3. Milan Tančić, Zoran Perić, **Stefan Tomić**, Nikola Simić: „Speech Signal Coding Using Forward Adaptive Quantization and Simple Transform Coding”, *Elektronika Ii Elektrotehnika*, (2016), vol. 22 br. 3, str. 74-77, M23.
4. Milan Tančić, Zoran Perić, **Stefan Tomić**, “Performance of Quasi-logarithmic Quantizer for Discrete Input Signal”, *Information Technology and Control*, 2017, 46(3), 395-402, M23.
5. Milan Tančić, Zoran Perić, Aleksandra Jovanović, **Stefan Tomić**: „Determining compression factor of quasilogarithmic quantizer for Laplacian source in narrow dynamic variance range”, *Facta Universitatis, Series: Automatic Control and Robotics*, (2016), vol. 15, br. 3, str. 217-226, M24.
6. **Stefan Tomić**, Zoran Perić, Milan Tančić: Forward Mean-Adaptive Quasilogarithmic Quantizer for Coding of Correlated Sources”, *Facta Universitatis, Series: Automatic Control and Robotics*, (2016), vol. 15, br.3, str. 205-215, M24.
7. **Stefan Tomic**, Zoran Peric, Jelena Nikolic, Wideband Speech Signal Coding with the Implementation of Modified BTC Algorithm, 52nd International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST), 2017, Nis, Serbia, M34.
8. Zoran Perić, Milan Tančić, **Stefan Tomić**, D. Ćirić, “Subband coding of audio signal with logarithmic compandors”, International Conference on Advanced Technologies, Systems and Services in Telecommunications - TELSIS, Niš, Serbia, October 14 - 17, 2015, pp. 19-22, M33.
9. Zoran Perić, **Stefan Tomić**, Milan Tančić, “High quality speech signal coding with the application of BTC algorithm”, International Conference on Advanced Technologies, Systems and Services in Telecommunications - TELSIS, Nis, Serbia, October 2015, pp. 23-26, M33.
10. **Stefan Tomić** , Milan Tančić, Dejan Ćirić, Zoran Perić, „Effects of subband coding with quasilogarithmic quantizers of characteristics of audio signal” Proceedings of 2nd International Conference on Electrical, Electronic and Computing Engineering IcETRAN 2015, Silver Lake, Serbia, June 8 – 11, 2015, ISBN 978-86-80509-71-6, pp. AKI1.4.1-5, M33.

11. Vera Marković, **Stefan Tomić**, Zlatica Marinković, “Trendovi u primeni indoor sistema u mobilnim komunikacijama”, XXXII Simpozijum o novim tehnologijama u poštanskom i telekomunikacionom saobraćaju- POSTEL, Belgrade, Serbia, December 2-3. 2014, pp. 199-208, M63.
12. Јелена Здравковић, Дејан Ћирић, Марко Јеленковић, **Стефан Томић**, “Добијање различитих карактеристика усмерености зрачења звучног извора са независним елементима у облику додекаедра”, *Зборник радова LVI Конференције ЕТРАН*, Златибор, Србија, 11-14 јун 2012, издавач: Друштво за ЕТРАН, стр. АК2.2-1-АК2.2-4 (на CD-у, рад АК2.2), 2012, ISBN 978-86-80509-67-9, M63.
13. Marko Jelenković, Dejan Ćirić, Jelena Zdravković, **Stefan Tomić**, “Control of radiation directivity applying independent element dodecahedral loudspeaker”, *Proceedings of XLVII International Scientific Conference on Information, Communication, Energy Systems and Technologies - ICEST 2012*, Veliko Tarnovo, Bulgaria, 28-30 June 2012, Publisher: Faculty of Telecommunications, Technical University of Sofia, pp. 131-134, 2012, ISBN 978-619-167-002-4, M34.

## **IZJAVE AUTORA**



## IZJAVA O AUTORSTVU

Izjavljujem da je doktorska disertacija, pod naslovom  
*Razvoj algoritama za kodovanje govornog i audio signala*  
rezultat sopstvenog istraživanja

- da predložena disertacija, niti u celini, niti u delovima, nije bila predložena za dobijanje bilo koje diplome, prema studijskim programima drugih visokoškolskih ustanova,
- da su rezultati korektno navedeni i
- da nisam kršio autorska prava, niti zloupotrebio intelektualnu svojinu drugih lica.

U Nišu, \_\_\_\_\_

Autor disertacije

\_\_\_\_\_ *Stefan Tomić* \_\_\_\_\_

Potpis doktoranda:

\_\_\_\_\_



## IZJAVA O ISTOVETNOSTI ŠTAMPANE I ELEKTRONSKE VERZIJE DOKTORSKE DISERTACIJE

Ime i prezime autora: Stefan Tomić

Srudijski program: Elektrotehnika i računarstvo

Naslov rada: Razvoj algoritama za kodovanje govornog i audio signala

Mentor: dr Zoran Perić, redovni profesor

Izjavljujem da je štampana verzija moje doktorske disertacije istovetna elektronskoj verziji, koju sam predao/la za unošenje u **Digitalni repozitorijum Univerziteta u Nišu**.

Dozvoljavam da se objave moji lični podaci, koji su u vezi sa 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, \_\_\_\_\_

Autor disertacije

Stefan Tomić

Potpis doktoranda:

\_\_\_\_\_



## IZJAVA O KORIŠĆENJU

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

*Razvoj algoritama za kodovanje govornog i audio signala*

koja je moje autorsko delo.

Disertaciju sa svim priložima predao/la sam u elektronskom formatu, 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/la.

1. Autorstvo
2. Autorstvo – nekomercijalno
3. Autorstvo – nekomercijalno – bez prerade
4. Autorstvo – nekomercijalno – deliti pod istim uslovima
5. Autorstvo – bez prerade
6. Autorstvo – deliti pod istim uslovima

U Nišu, \_\_\_\_\_

Autor disertacije

*Stefan Tomić*

Potpis doktoranda:

\_\_\_\_\_