

Izvještaj za prvi ciklus iz predmeta DOG

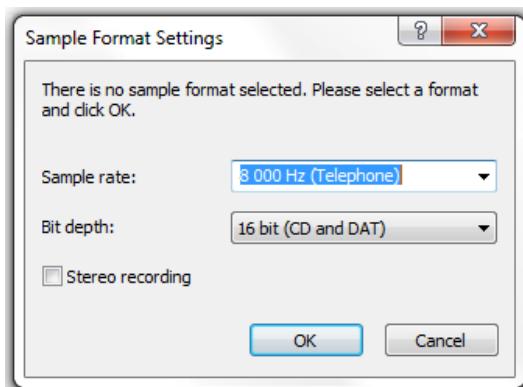
Domaća zadaća za prvi ciklus:

Domaća zadaća prvog ciklusa je izvještaj o obavljenom samostalnom radu na prvih sedam vježbi iz udžbenika „Uvod u digitalnu obradbu govora“. To su vježbe:

1.	<i>Snimanje govornih signala korištenjem zvučnih kartica.....</i>	2
2.	<i>Analiza svojstava govornih signala u vremenskoj domeni.....</i>	3
3.	<i>Spektralna analiza govora i spektrogrami.....</i>	13
4.	<i>Analiza formantne strukture govora.....</i>	19
5.	<i>Automatska klasifikacija samoglasnika na osnovu formantne strukture.....</i>	22
6.	<i>Automatska klasifikacija govornika na osnovu formantne strukture.....</i>	27
7.	<i>Postupci linearne predikcije</i>	28

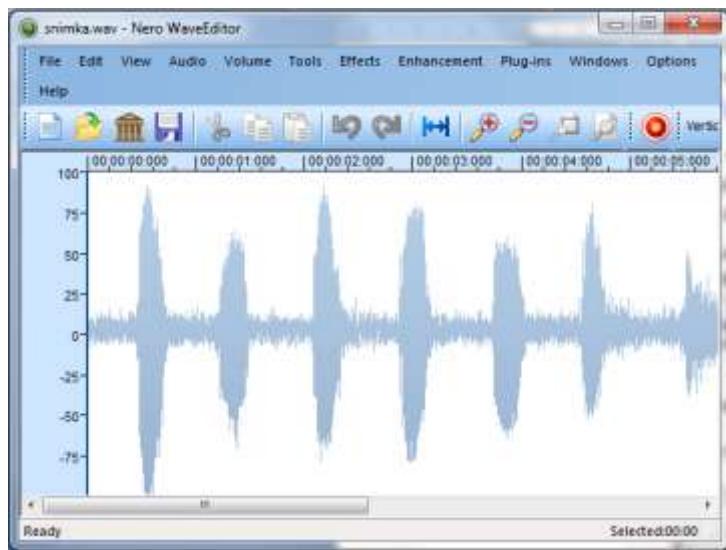
1. Snimanje govornih signala korištenjem zvučnih kartica

U sklopu laboratorijskih vježbi i domaće zadaće obrađuju se prethodno snimljeni uzorci govornog signala. Sama obrada odvija se pomoću *Matlab* programske potpore, dok se govorni signal snima na proizvoljan način. Budući da operacijski sustav Windows 7 ne sadrži zadovoljavajući alat za snimanje zvuka po određenim postavkama, moguće je slobodno preuzeti besplatnu verziju programa „[Nero WaveEditor](#)“ te uz njegovu pomoć snimiti odgovarajuće snimke govora. Za snimanje govora potrebno je unutar programa *WaveEditor* postaviti odgovarajući mikrofon za snimanje te je potrebno postaviti frekvenciju otiskivanja na 8 kHz-a, amplitudnu rezoluciju otiskivanja na 16 bita te mono način snimanja (slika 1.1).



1.1 Postavke snimanja govora

Nakon snimanja snimku je potrebno spremiti u .wav datoteku. Primjer jedne snimke nalazi se na slici 1.2.



1.2 Primjer snimke u programu WaveEditor

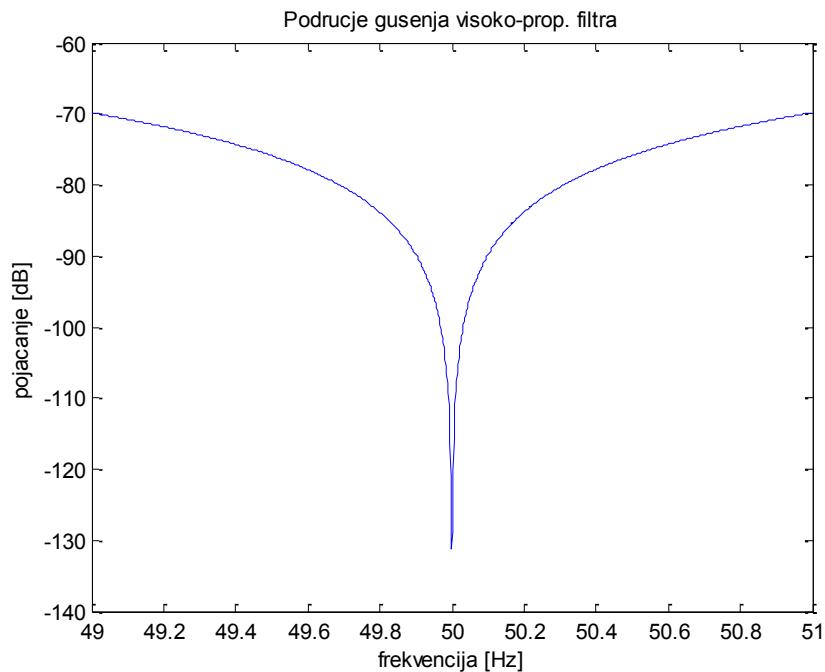
2. Analiza svojstava govornih signala u vremenskoj domeni

Zadatak 1.: Odrediti frekvencijsku karakteristiku filtra u pojasu 49 do 51 Hz i prikazati je grafički.

Određivanje frekvencijske karakteristike visokopropusnog filtra ostvaruje se naredbom:

```
[h, om]=freqz(bh, ah, linspace(49,51,1024)/fs*2*pi);
```

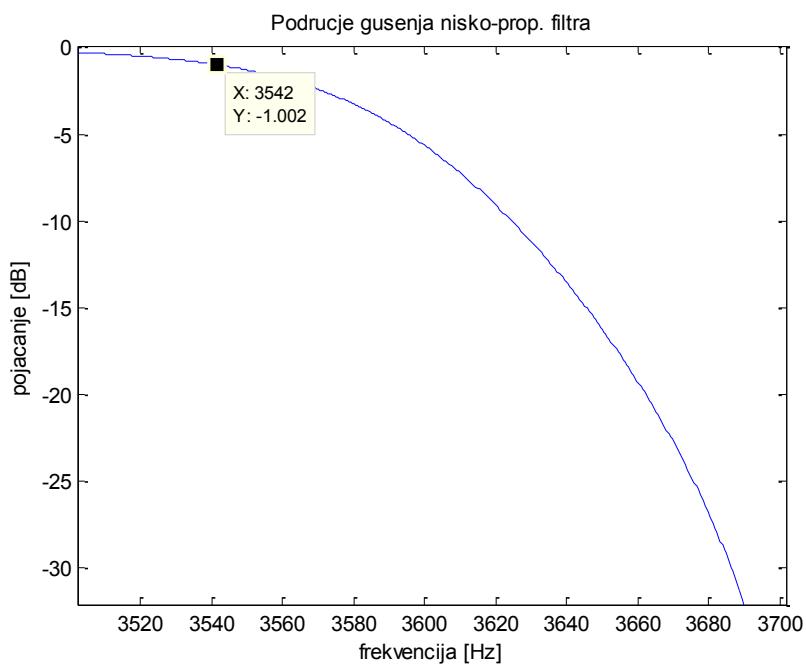
Rezultat:



Slika 2.1 Frekvencijska karakteristaika u pojasu 49 do 51 HZ

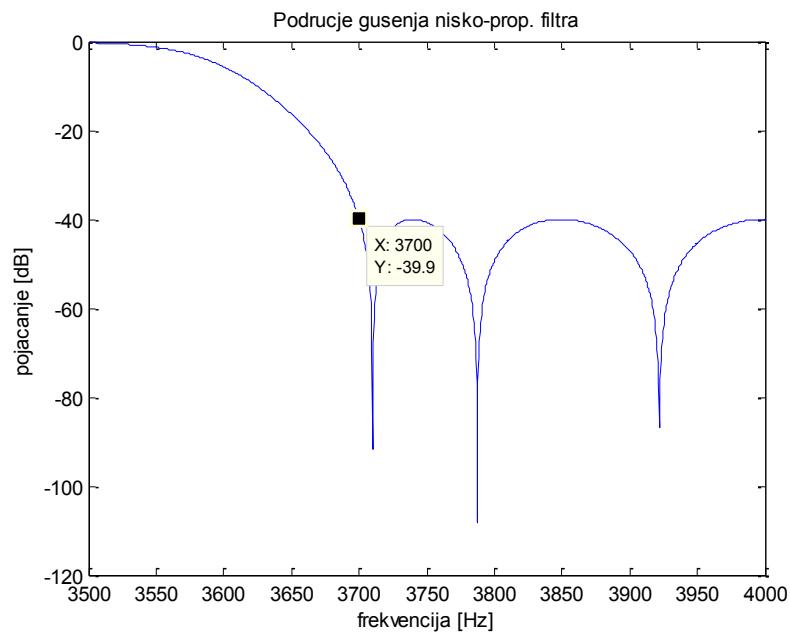
Zadatak 2. : Variranjem f_{min} i f_{max} , te izračunavanjem frekvencijske karakteristike filtra u tom pojasu odrediti gornju graničnu frekvenciju područja propuštanja, tj. Krajnju frekvenciju na kojoj gušenje filtra još uvijek nije veće od 1dB. Na osnovu određene granične frekvencije definirati područje propuštanja, prijelazno područje i područje gušenja filtra, kao što je to bilo učinjeno za visokopropusni filter.

Na slici 2.3 vidljivo je kako je krajnja frekvencija gušenja manja od -1 dB oko 3542 Hz.



Slika 2.2 Područje gušenja ne veće od -1 dB

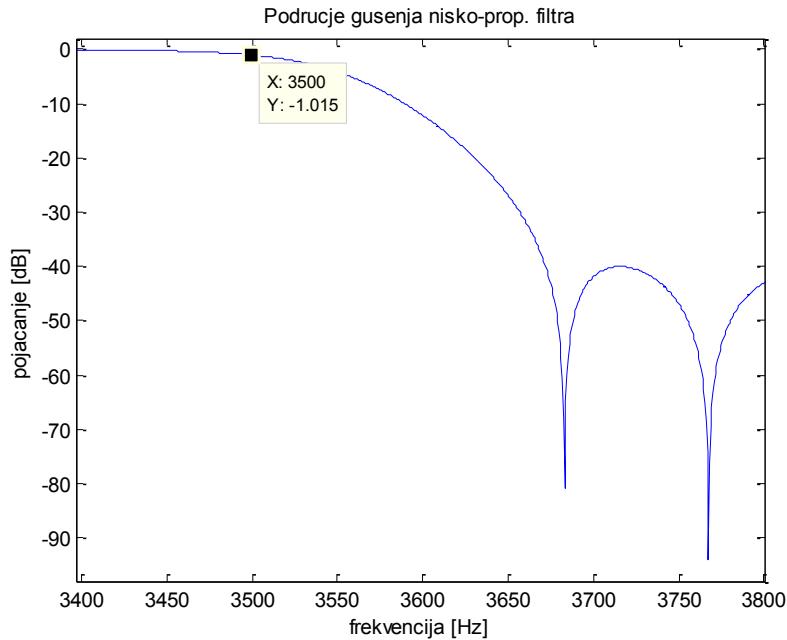
Područje propuštanja iznosi sve frekvencije manje od 3540 Hz, prijelazno područje sadrži frekvencije između 3540 i 3710 Hz dok je područje gušenja na 3700 Hz. Ove frekvencije vidljive su iz slike 2.4.



Slika 2.3 Frekvencijska karakteristika između 3500 i 4000 Hz

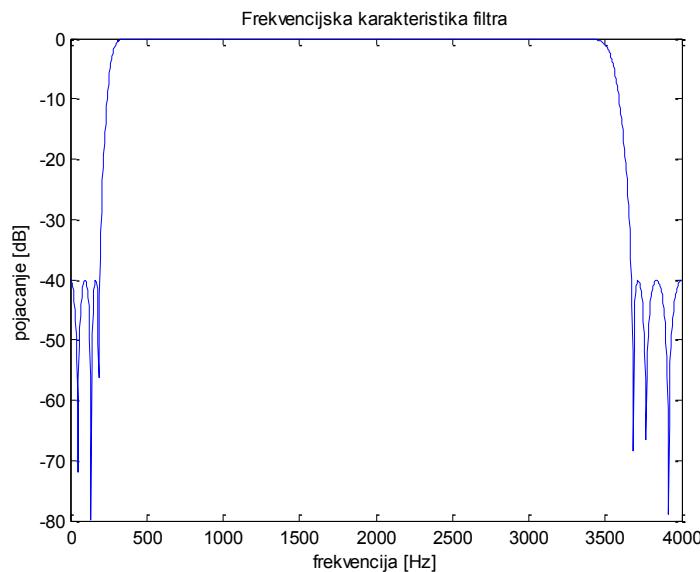
Zadatak 3.: Variranjem donje granične frekvencije f_g , pokušati odrediti iznos f_g uz koji će gornja granična frekvencija područja propuštanja biti točno jednaka 3500 Hz.

Ukoliko se frekvencija f_g namjesti na vrijednost od 3672 Hz, tada gornja granična frekvencija iznosi točno 3500 Hz.



Slika 2.4 Slika uz zadatok 3

Zadatak 4.: Odrediti frekvencijske karakteristike kaskade u prijelaznim područjima i usporediti ih s pripadnim karakteristikama svakog pojedinog filtra (NP i VP). Zbog čega su prijelazne karakteristike ostale nepromijenjene?



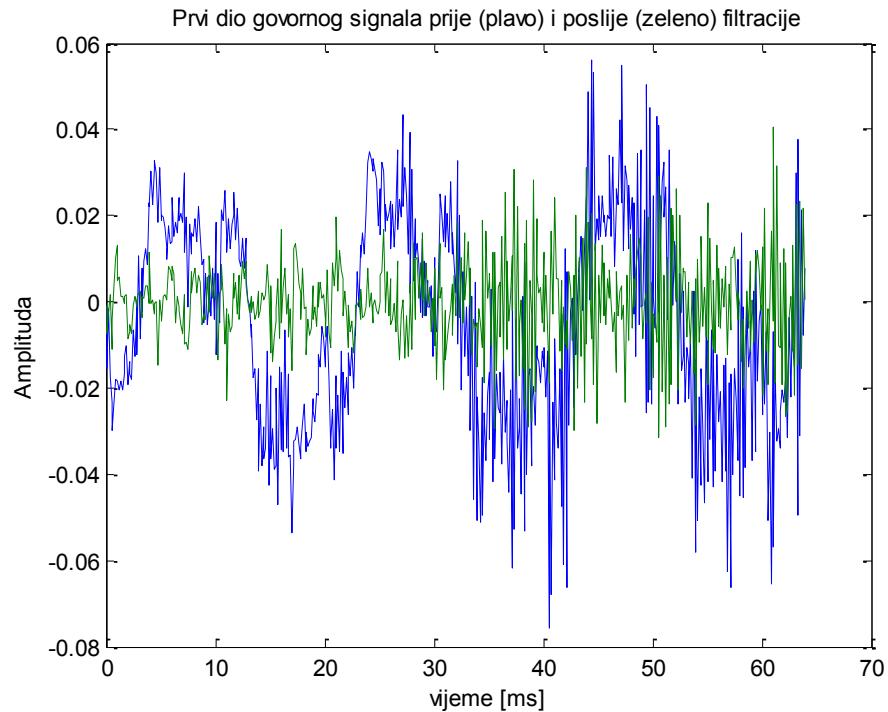
Slika 2.5 Kaskada NP i VP filtra

Frekvencijske karakteristike kaskade u prijelaznim područjima iznose jednake iznose kao i pripadni NP odnosno VP filtri. Razlog tomu je što primjerice NP filter nema gušenje u prijelaznom području VP filtra i obratno.

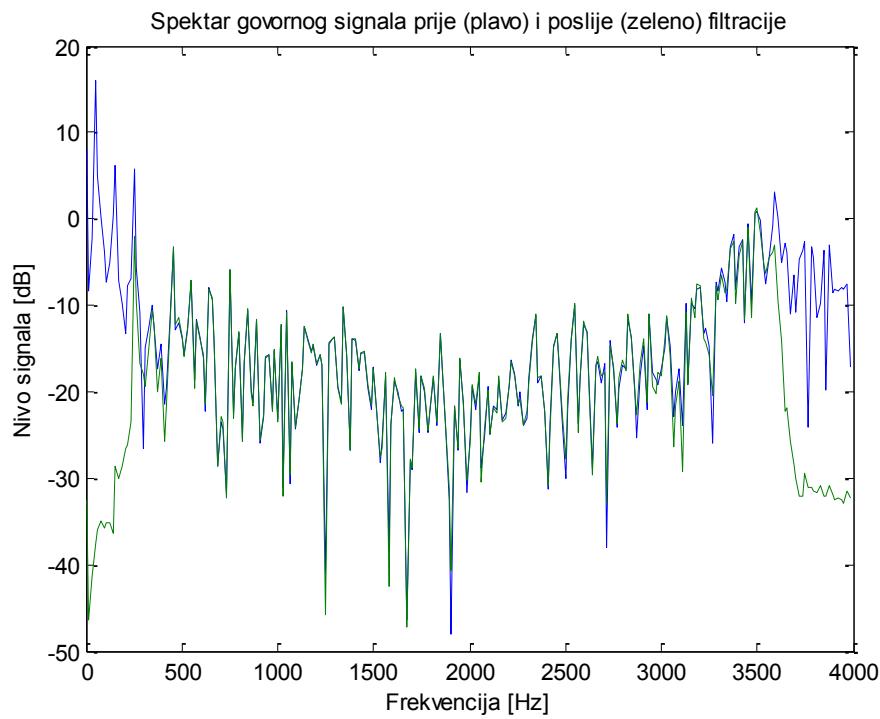
Zadatak 5.: Naizmjence primjenom funkcije `soundsc` reproducirati originalni signal y i filtrirani signal yf . Da li čujete kakvu razliku između ova dva signala?

Razlika između originalne snimke i filtrirane snimke jest ta da filtrirana snimka ne sadrži duboke tonove te također zvuči tiše od originalne snimke.

Zadatak 6.: Nakon što program usporedno prikaže originalni i filtrirani signal u vremenskoj domeni, na slijedećem prikazu su ta dva signala prikazana u frekvencijskoj domeni. Pohraniti ovu sliku i diskutirati područja gušenja filtra, te kako se ona manifestiraju u prikazanim spektrima.



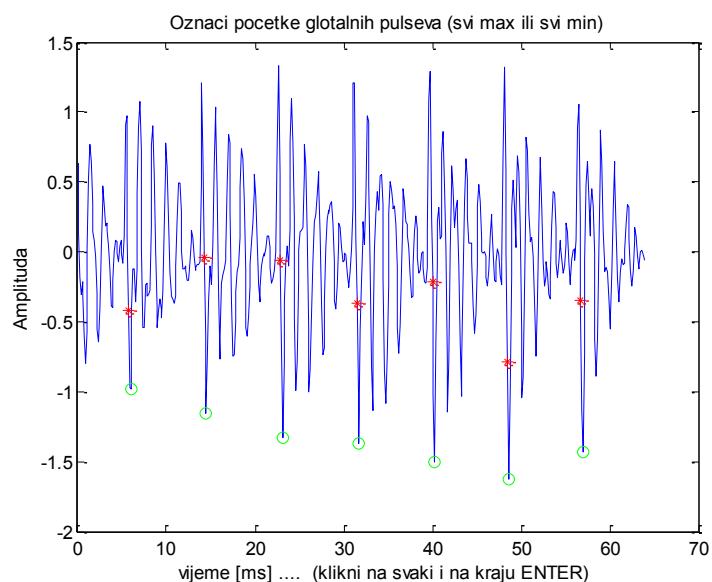
Slika 2.6 Filtrirani i originalan signal



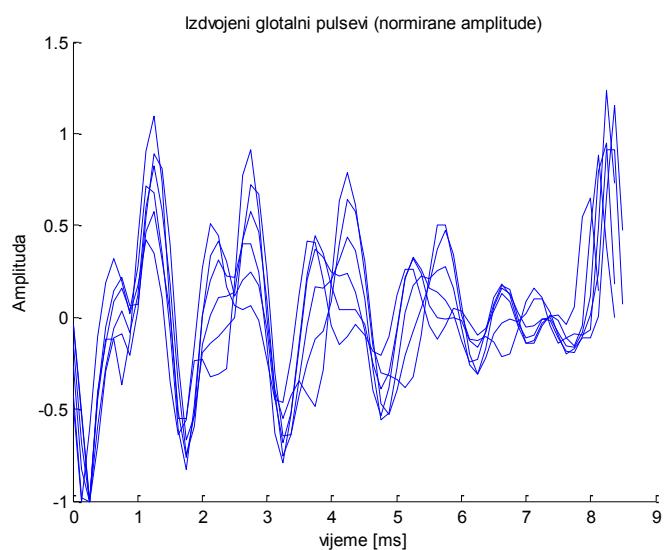
Slika 2.7 Spektar filtriranog i originalnog signala

Na slici signala vidljivo je kako je kako je uklonjena smetnja izazvana gradskom mrežom na 5 Hz odnosno 20 ms. Na slici spektra vidljivo je područje gušenja koje smo projektirali NP i VP filtrom odnosno u području ispod 300 i iznad 3500 Hz.

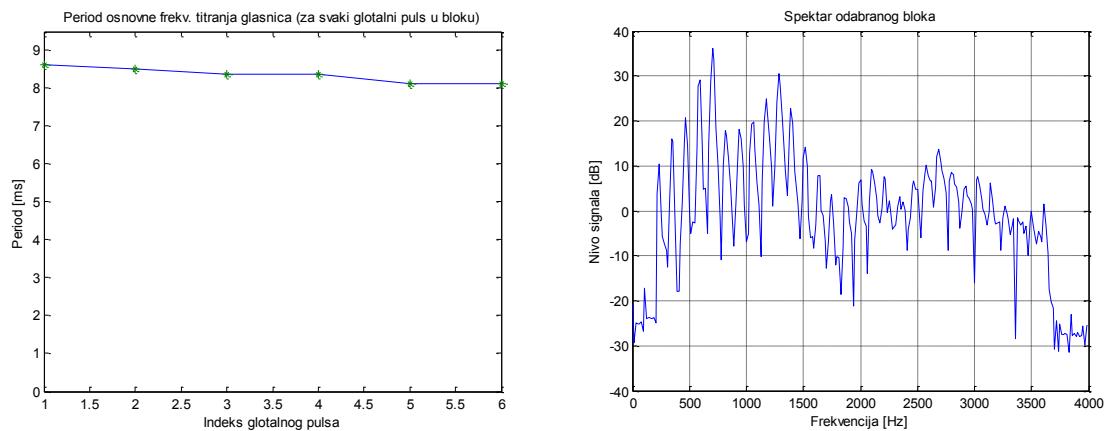
Zadatak 7.: Gore opisanim postupkom odrediti segment govornog signala koji odgovara glasu "a" na kraju izgovora riječi "subota". Prikazani valni oblik odabranog segmenta pohraniti u Word datoteku izvještaja s rezultatima samostalnog rada.



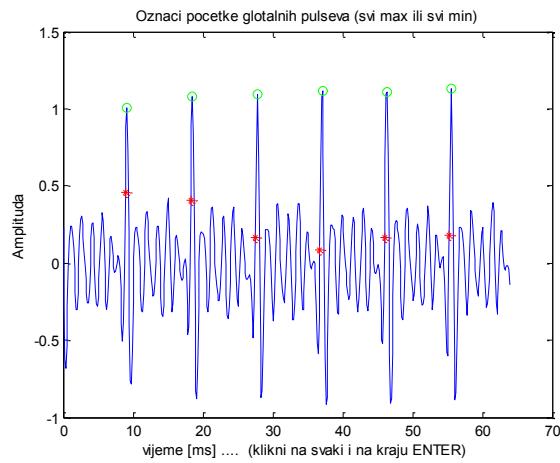
Slika 2.8 Vremenski odsječak glasa a



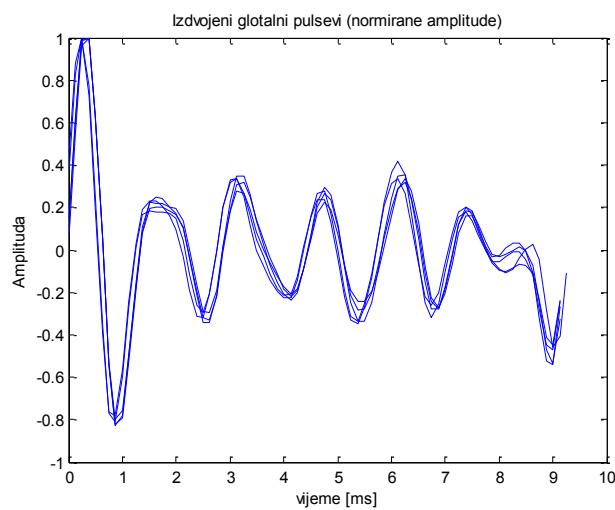
Slika 2.9 Preklopjeni periodi



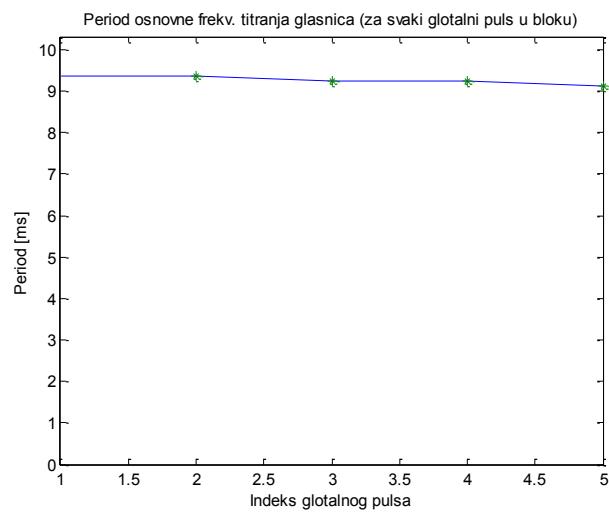
Zadatak 8.: Opisanu analizu primjenom programa "pri.m" ponoviti na Wav-datoteci sa samoglasnicima koju ste snimili u okviru samostalnih zadataka u poglavlju 5. Analizu provesti za samoglasnik "a".



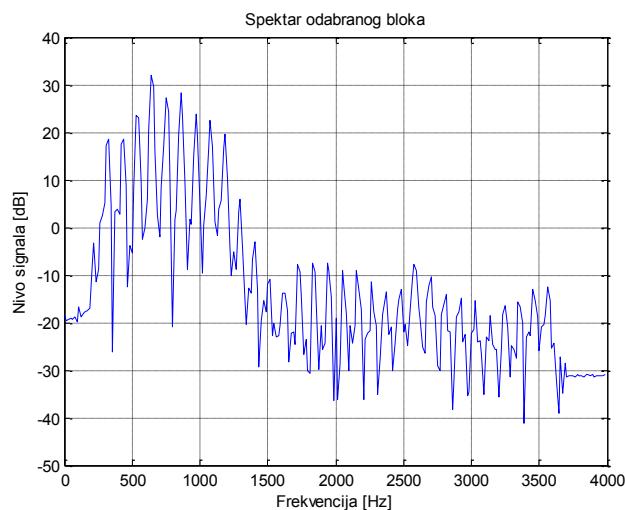
S obzirom da smo zahvatili nazovimo to čišći glas „a“ od prethodno izdvojenog glasa iz riječi subota, vidljivo je kako su amplitude ujednačenje od prethodnog primjera.



Također je vidljivo da se izdvojeni glotalni pulsevi relativno bolje preklapaju od glasa „a“ izdvojenog iz riječi subota.

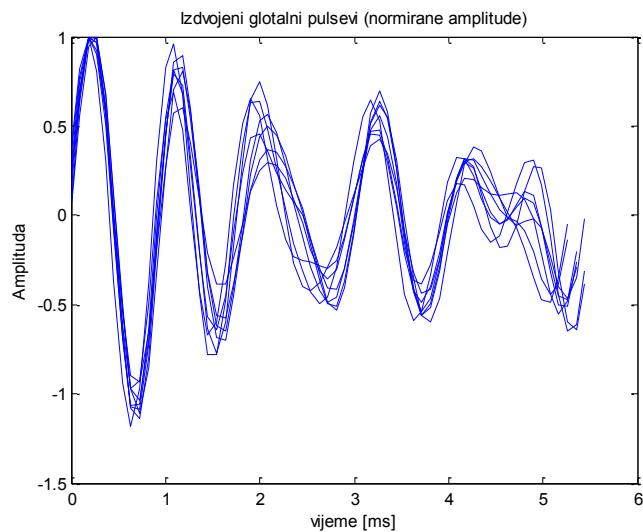
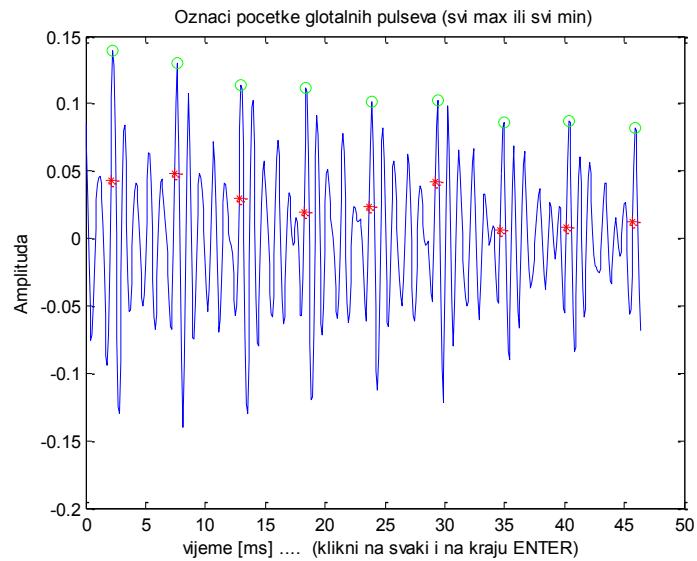


Period frekvencija titranja glasnice s nije mjenjao što mso i očekivali s obzirom da se radi o istom glasu „a“.

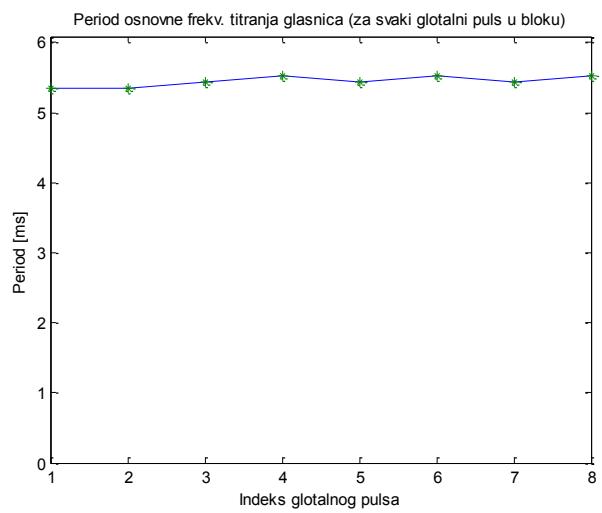


Kod spektra odabranog bloka je vidljiva razlika jedino u ujednačenijem iznosu novoga signala za razliku od primjera s izdvojenim glasom „a“.

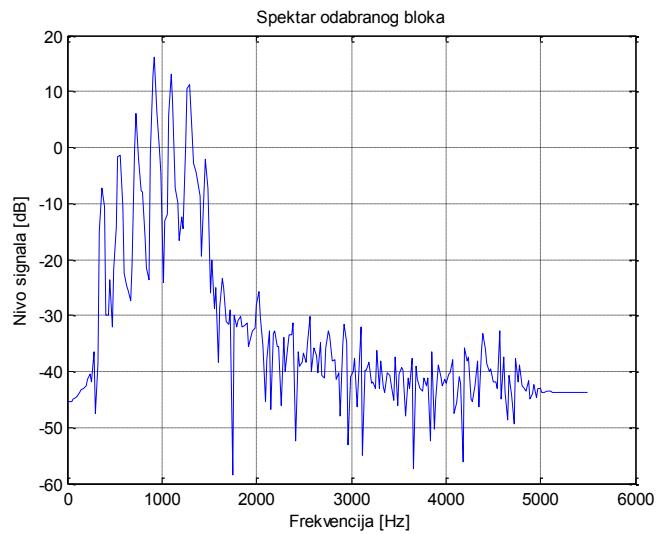
Zadatak 9.: Za primjer treba snimiti glas "a", ali ovaj put sa silaznom intonacijom (visina glasa se spušta prema kraju izgovora). Prilikom snimanja, neka brzina promjene osnove frekvencije bude velika (brza), tako da ta promjena bude očita unutar okvira analize. Kakva će biti periodičnost osnovne frekvencije titranja glasnica ako se mijenja visina glasa? Koliki je period titranja?



Na ovoj slici je vidljivo kako se pred kraj periodi produžuju.

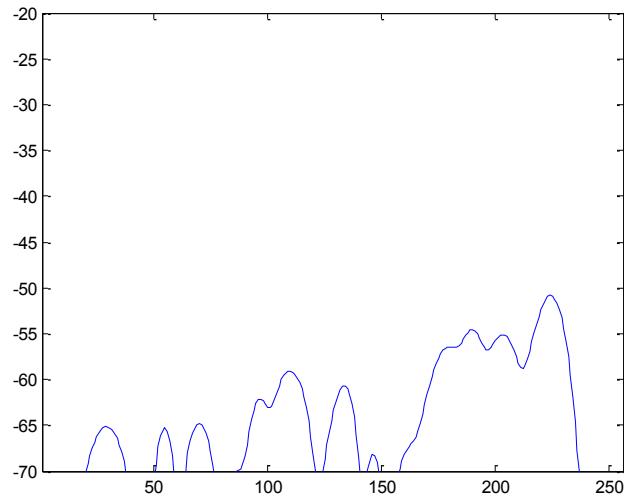


Na ovoj slici je vidljivo kako osnovni period titranja raste. NA početku iznosi oko 5.2 ms dok na kraju iznosi oko 5.5 ms. Na snimci nije moj glas zbog privremene nemogućnosti normalnog govora.

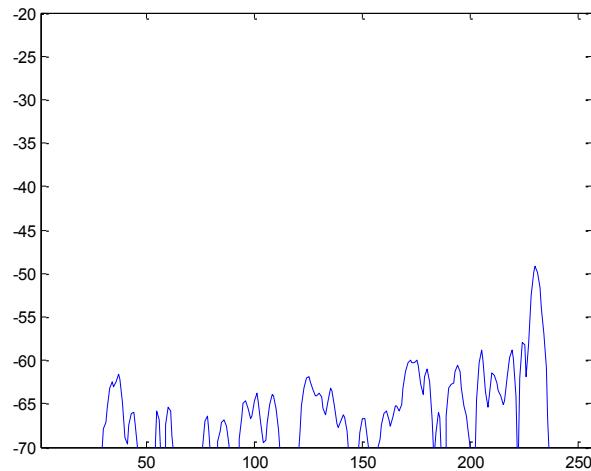


3. Spektralna analiza govora i spektrogrami

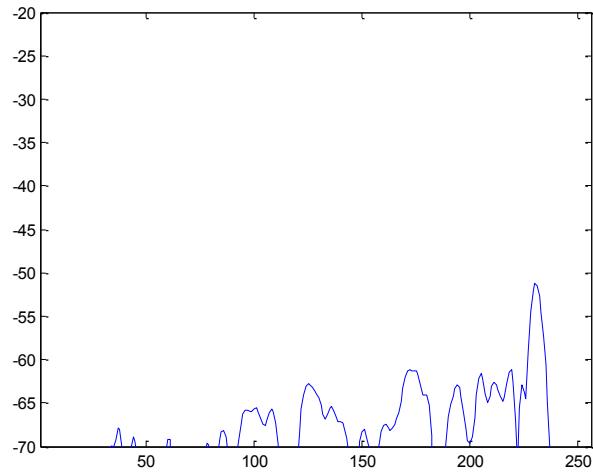
Zadatak 1: Uz tako fiksiran indeks okvira analize, izvoditi naredbe unutar petlje jednu po jednu i prikazivati signal (odnosno spektar) nakon svake pojedine linije, te tako dobivene slike pohranjivati u vašu Doc-datoteku.



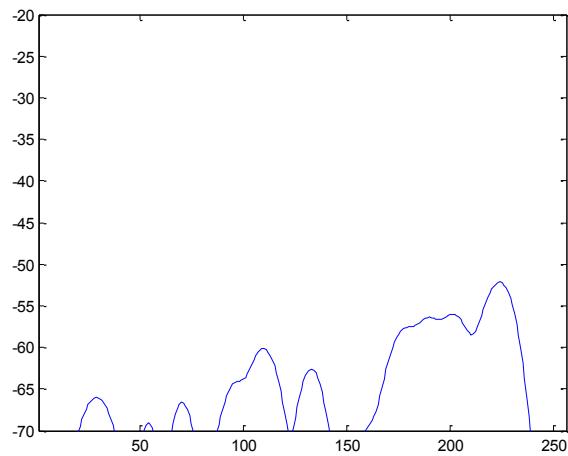
Slika 3.1 Okvir broj 60 s Hammingovim prozorom veličine 80 uzoraka



Slika 3.2 Okvir broj 60 s Hammingovim prozorom veličine 200 uzoraka



Slika 3.3 Okvir broj 60 s Blackmanovim prozorom veličine 200 uzorka



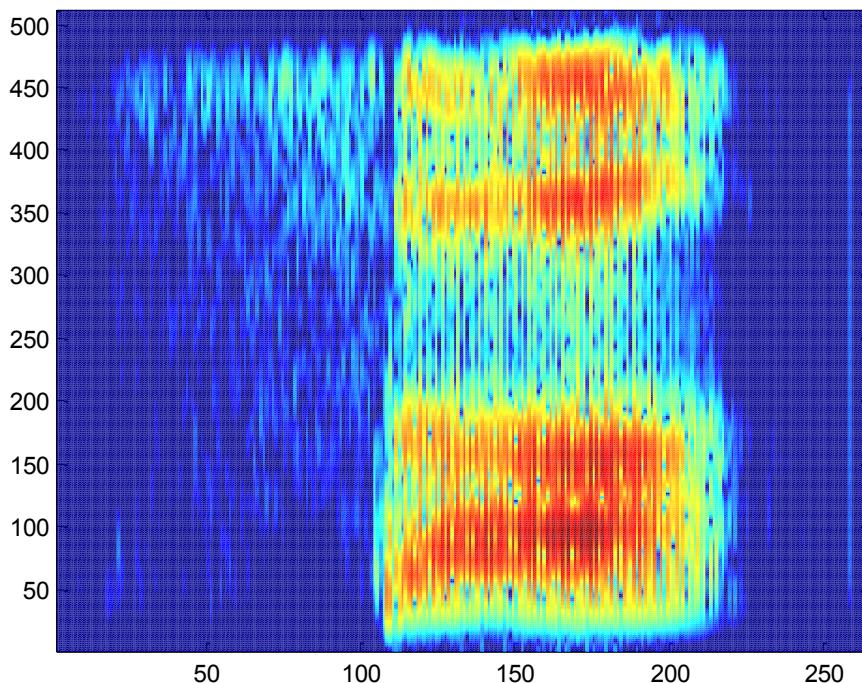
Slika 3.4 Okvir broj 60 s Blackmanovim prozorom veličine 80 uzorka

Na slikama je vidljivo kako se povećanjem veličine prozora spektar snage preciznije određuje.

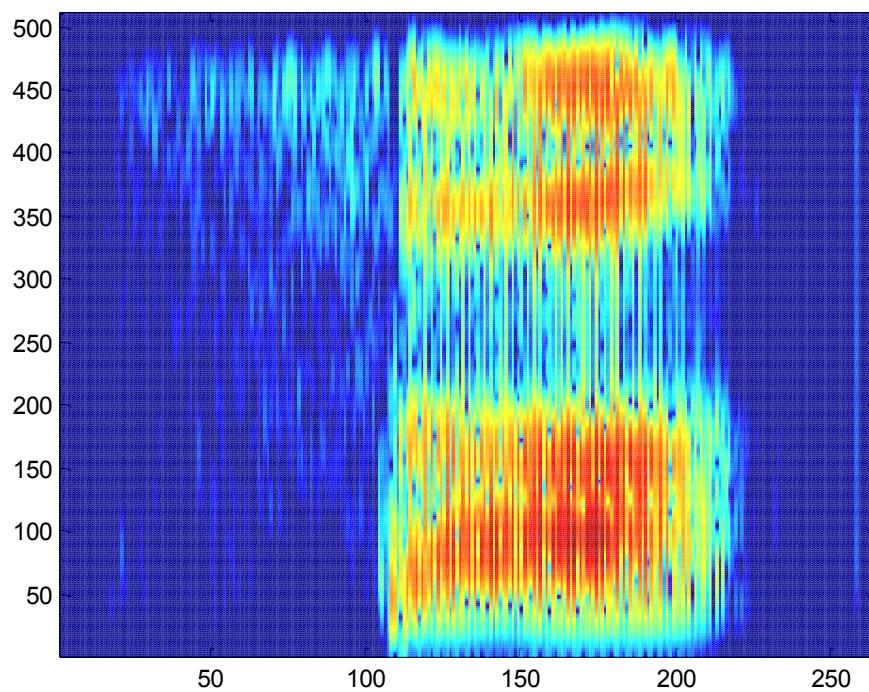
Zadatak 2.: Prije pokretanja programa specgrm.m potrebno je izmjeniti vrijednosti N, kor i otv, kao što je navedeno u tablici da bi se dobio određeni spektrogram.

	N	Kor	Otv
Širokopojasni	1024	30	50
Uskopojasni	1024	150	200

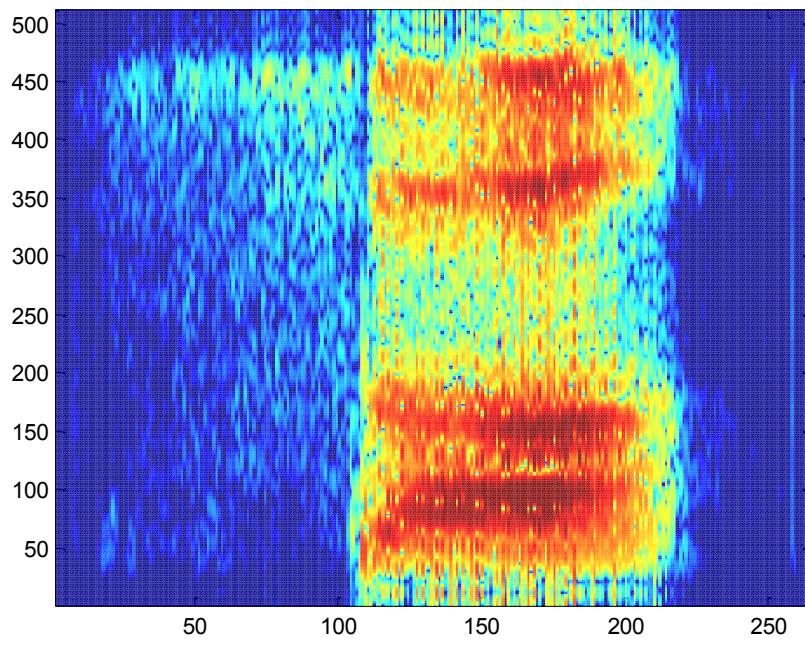
- a) Potrebno je pokušati odrediti, na uskopojasnom spektrogramu, razmak među harmonicima (uzimajući u obzir da je frekvencija prikazana u rasponu 0-4 kHz, uz pomoć prebrojavanja pruga i širine područja u kojem su prebrojene).
 - b) Što se događa kada variramo korak između dva susjedna bloka analize kor? Kada je prikaz bolji?
 - c) Prikažite spektrograme za različite tipove vremenskih otvora w (Hamming, Blackman, pravokutni...), ...koji je najbolji?
 - d) Pogledajte što se događa kada uzmemmo mali otv i Blackmanov otvor w, je li rezultat spajanja pruga dobra stvar?
 - e) Izračunajte i prikažite spektrograme za svoje glasove snimljene u okviru praktičnih zadataka prethodnih poglavlja.
-
- a) Razmak među harmonicima iznosi oko 140 Hz. ($3200/23$ pruge = 140 Hz)
 - b) Manjim korakom dobivamo precizniji spektrogram (slika je oštrega).
 - c) Na slici 3.5 nalazi se širokopojasni spektrogram dobiven Hammingovim otvorom širine 50 uzoraka i brojem koraka 30, dok se na slici 3.6 nalazi širokopojasni spektrogram dobiven Blackmanovim otvorom te na slici 3.7 širokopojasni spektrogram dobiven pravokutnim otvorom uz iste parametre.



Slika 3.5 širokopojasni spektrogram dobiven Hammingovim otvorom



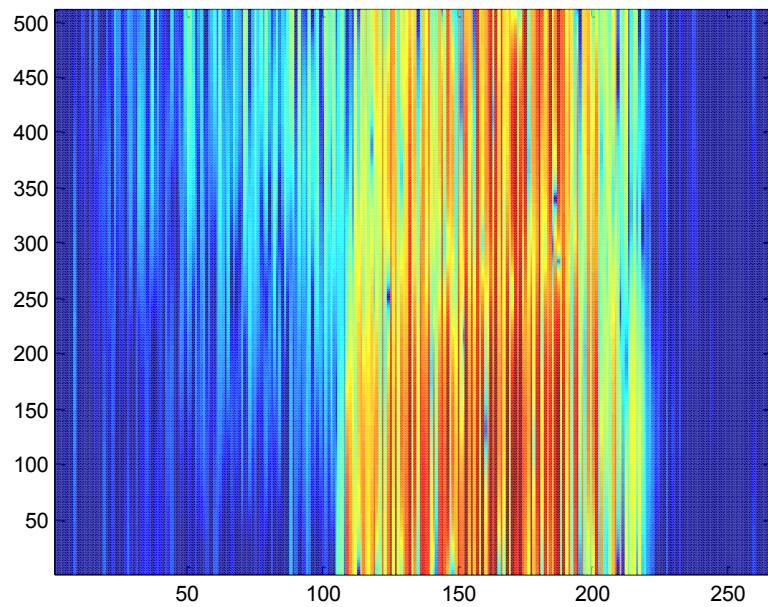
Slika 3.6 Širokopojasni spektrogram dobiven s Blackmanovim otvorom



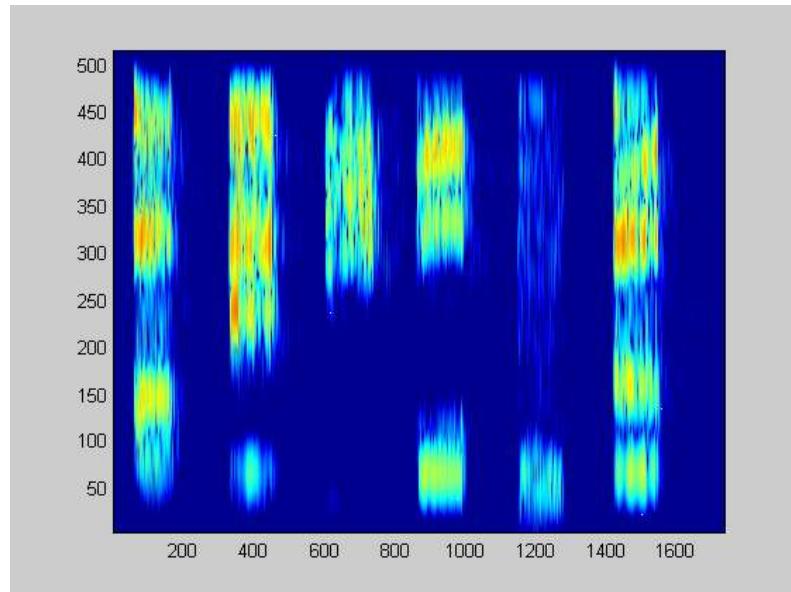
Slika 3.7 Širokopojasni spektrogram dobiven pravokutnim otvorom

Na slikama se razlika gotovo ni ne primjećuje osim za pravokutni otvor kod kojeg je slika nešto lošija.

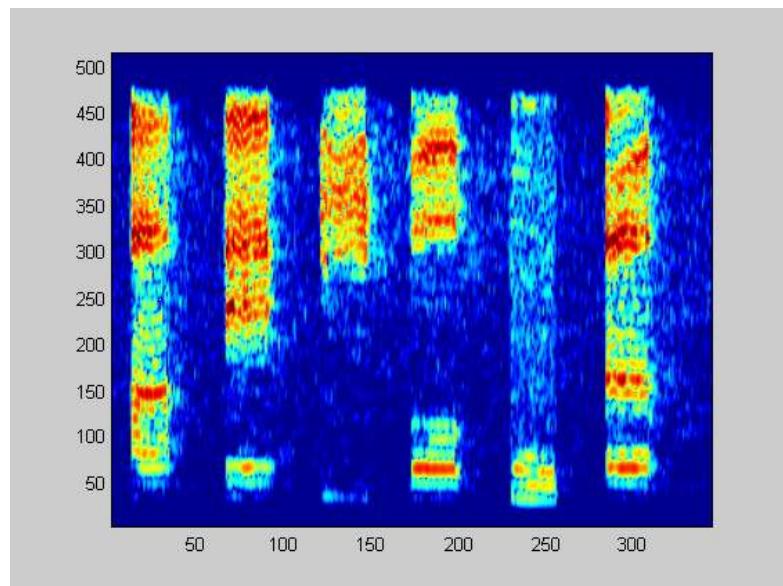
- d) Ukoliko smanjimo veličinu prozora na vrlo malu vrijednost dolazi do spajanja pruga što naravno nije dobra stvar jer tada ne možemo ništa korisno isčitati iz samog grafa. Primjer se nalazi na slici 3.8 s otvorom veličine 10.



- e) Na sljedećim slikama nalaze se uskopojasni i širokopojasni spektrogrami samoglasnika i neutralnog glasa 3:



[3.8 Uskopojasni spektrogram](#)

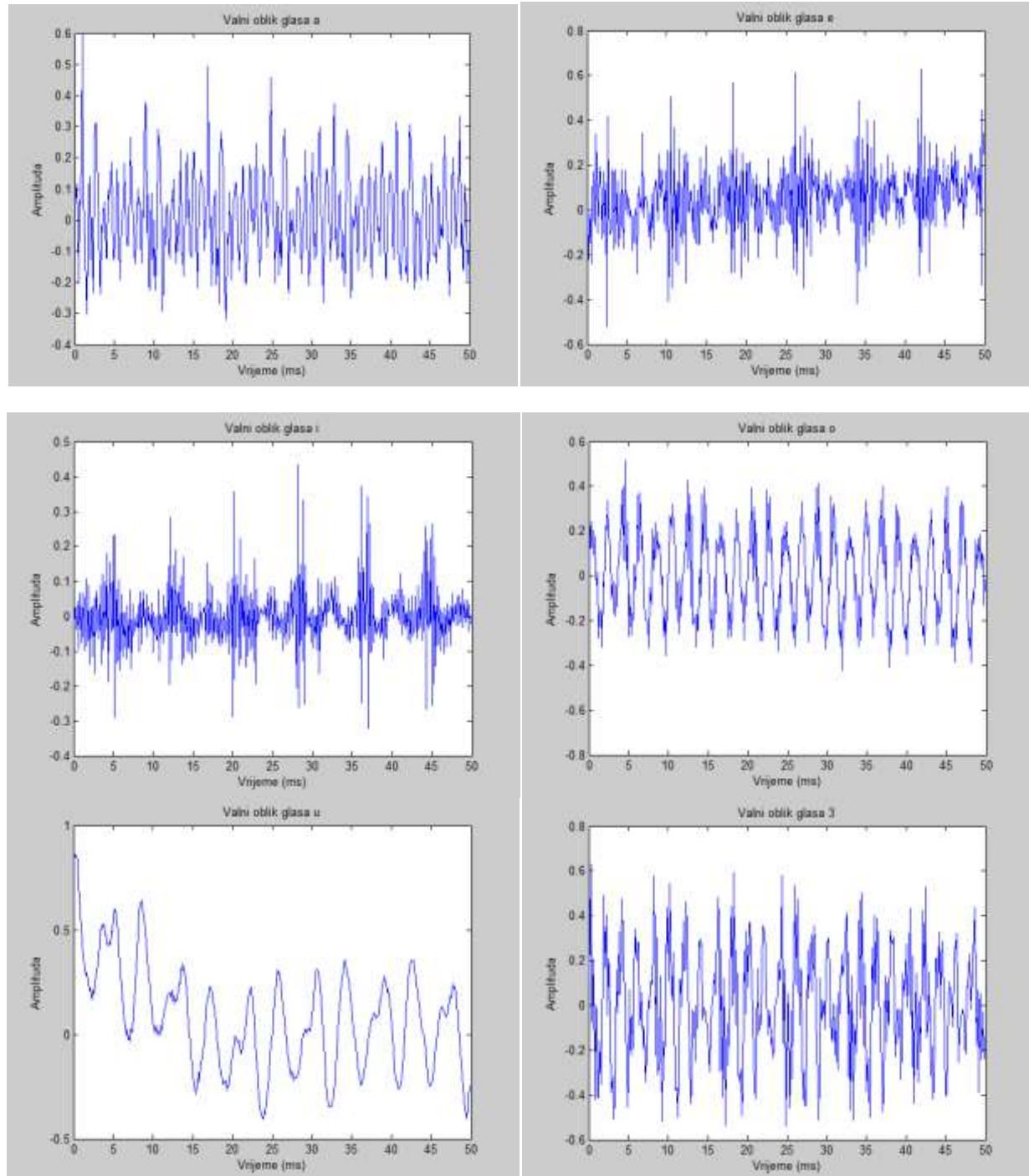


3.9 Širokopojasni spektrogram

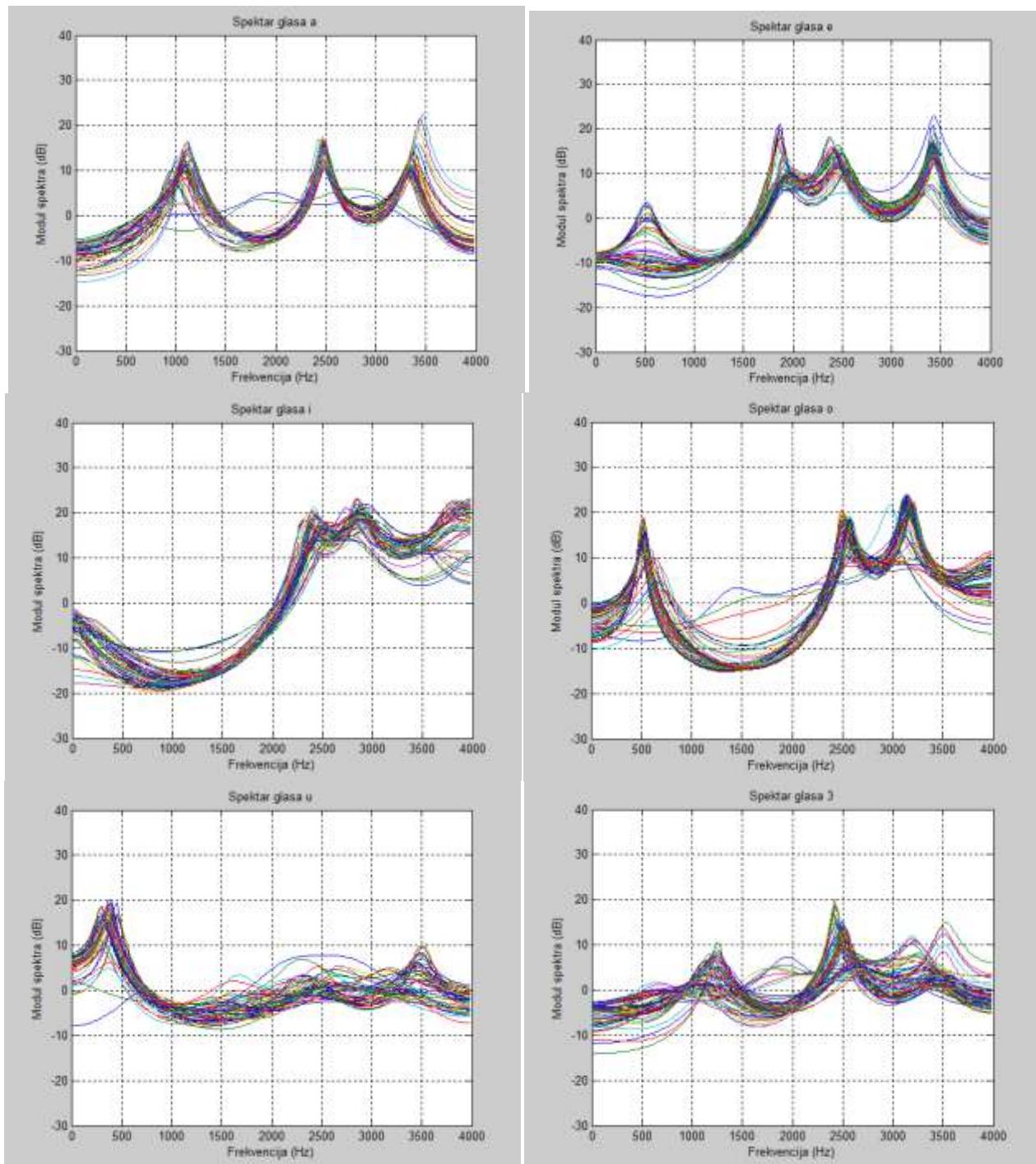
4. Analiza formantne strukture govora

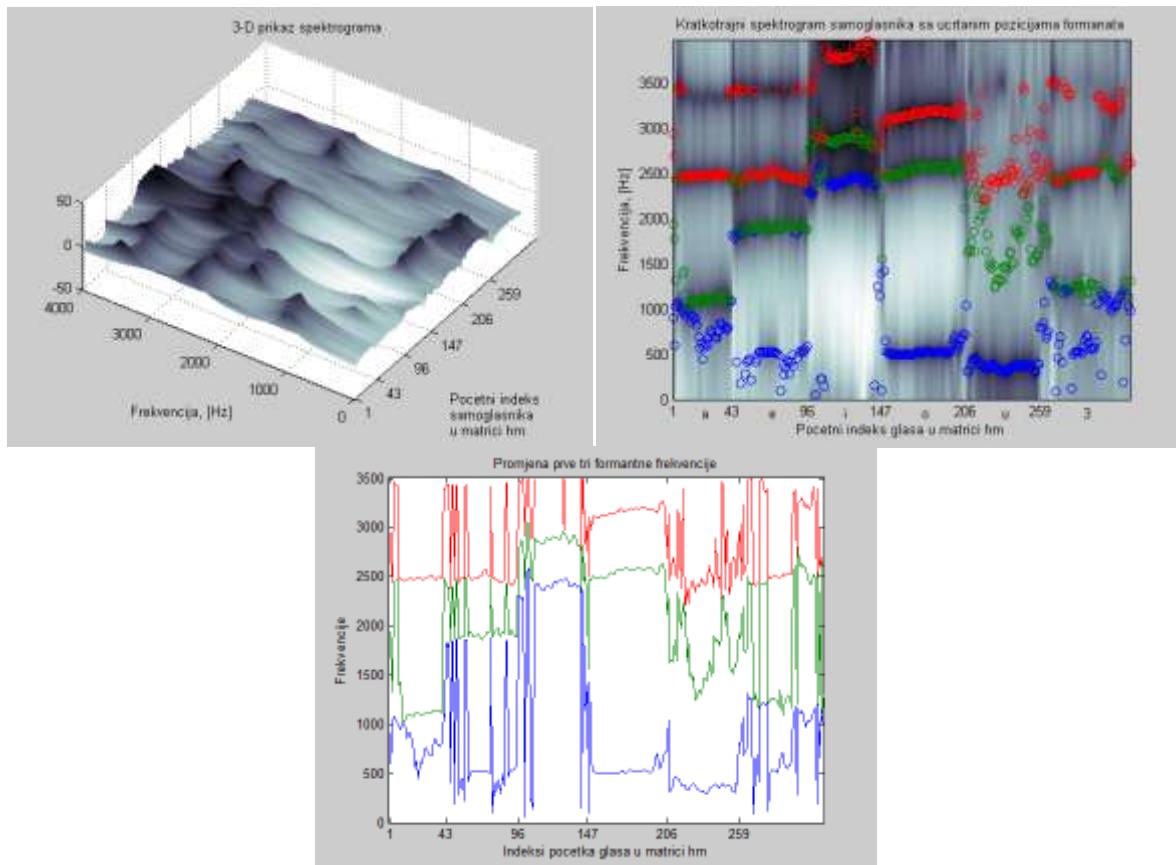
Zadatak 1.: Snimiti wav-datoteku s nizom samoglasnika, obraditi ih pomoću poziv:cov2 i dog5. Isprobajte nekoliko redova prediktora (npr. 9., 10. i 11i) dok ne dobijete minimalno rasipanje ovojnica spektrograma i centralnih frekvencija formanata. LPC ovojnice kratkotrajnih spektara signala u 1D, 2D i 3D prikazu prenesite u MS word pomoću copy+paste postupka i snimite u izvještaj. U tablicu upišite početne i završne okvire analize u kojima su samoglasnici.

Valni oblici samoglasnika (slike su po redu a,e,i,o,u,3):



U nastavku slijede slike spektara za pojedine glasove. Najbolja moguća analiza, odnosno rasipanje provedeno je s osmim (8) redom prediktora:



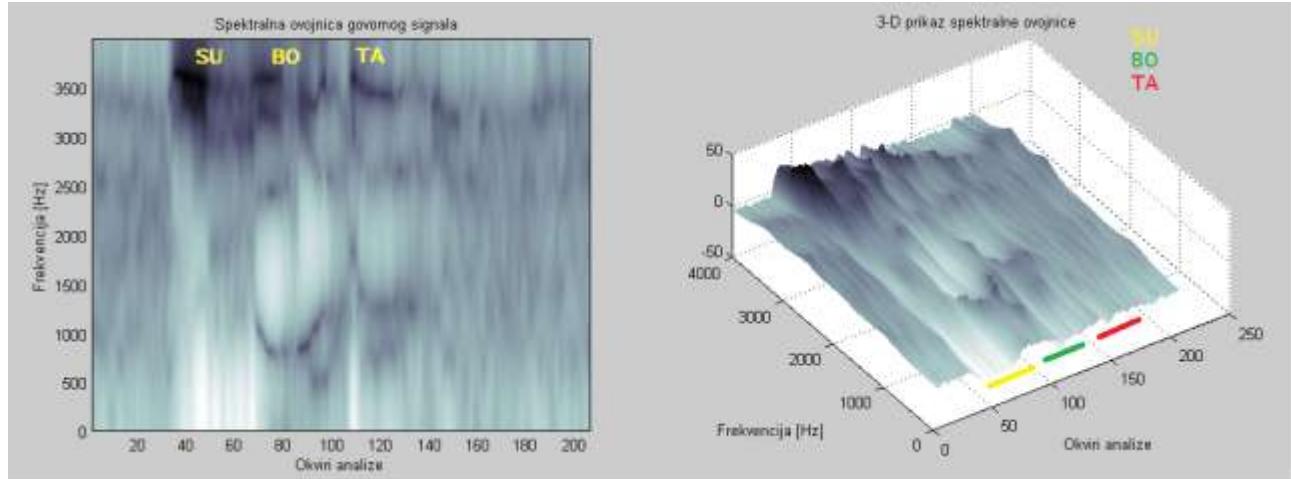


Vokal	Početak	Završetak
a	1	43
e	44	96
i	97	147
o	148	206
u	207	259
3	260	309

Razlog ovako velikog rasipanja jest i u tome što je snimka samoglasnika vrlo loša zbog loše kvalitete mikrofona.

Zadatak 2.: Provedite analizu programom rijec2.m koji je već prilagođen da učita wav-datoteku sa snimkom vašeg imena da se odredi spektralna ovojnica za vaš glas. Snimite kraću riječ s dovoljno samoglasnika, npr. subota, utorak ili slično. Obradite ju datotekom riječ, rezultate pohranite u Word datoteku. Pomoći 2D i 3D slike pokušajte pronaći i označiti samoglasnike.

Korištena je snimka riječi „SUBOTA“



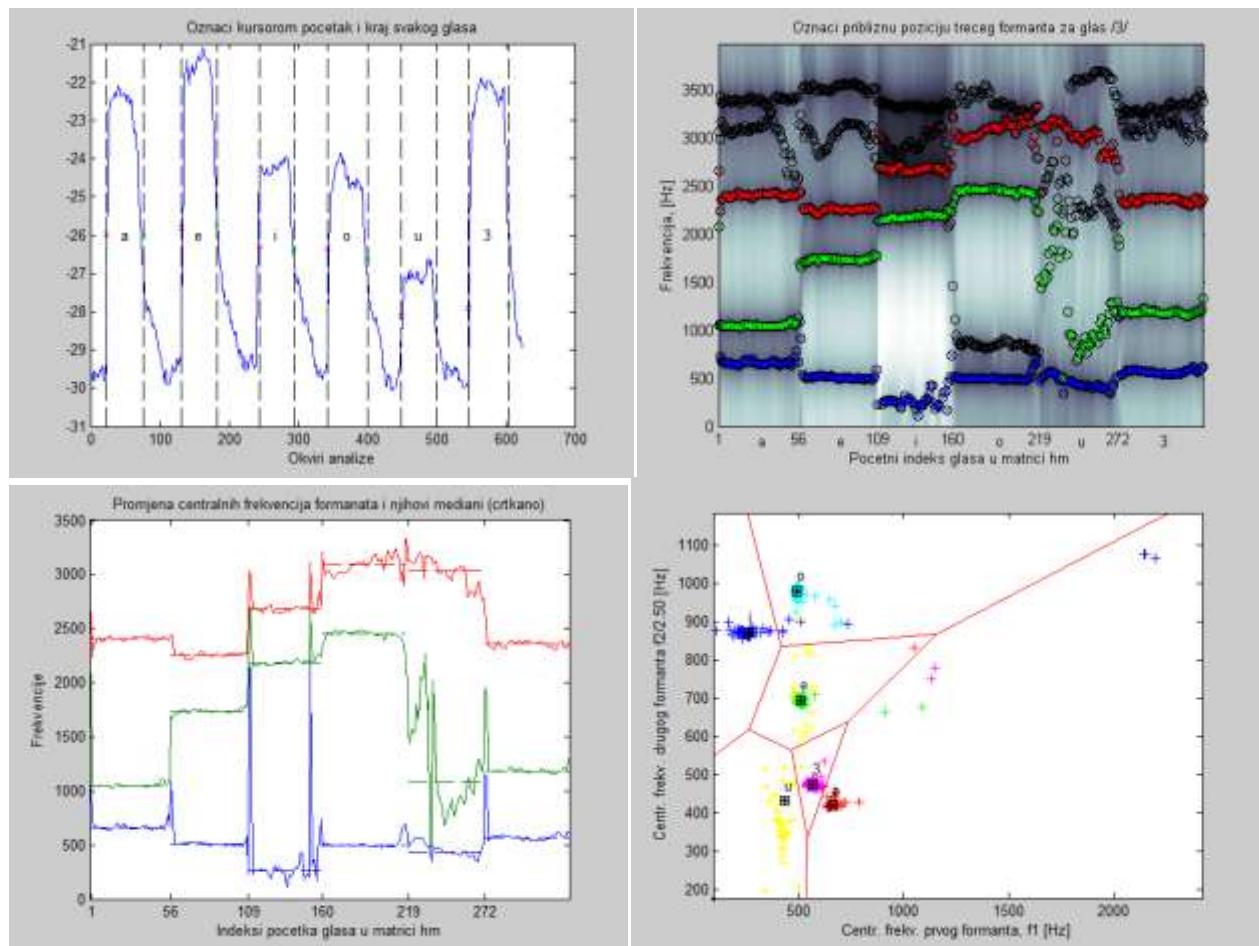
5. Automatska klasifikacija samoglasnika na osnovu formantne strukture

Zadatak 1.:

- a) Provesti LPC analizu vaših izgovora samoglasnika kako je opisano u pripremnom dijelu. Vaše datoteke numerirati rednim brojem 0. (*sp0_1, sp0_2 i sp0_3*).
- b) Provesti formantnu analizu sa sva tri vaša izgovora primjenom programa 'dog6_formanti.m', te od tih tri izgovora odabrati barem dva koja imaju najčišću formantnu strukturu. Dva odabrana izgovora i sve pripadne datoteke preimenovati tako da dva najbolja imaju indekse izgovora 1 i 2.

U word datoteku pohraniti sve slike.

Iako je u zadatku zadano da se odaberu samo dva izgovora, ovdje su sva tri.



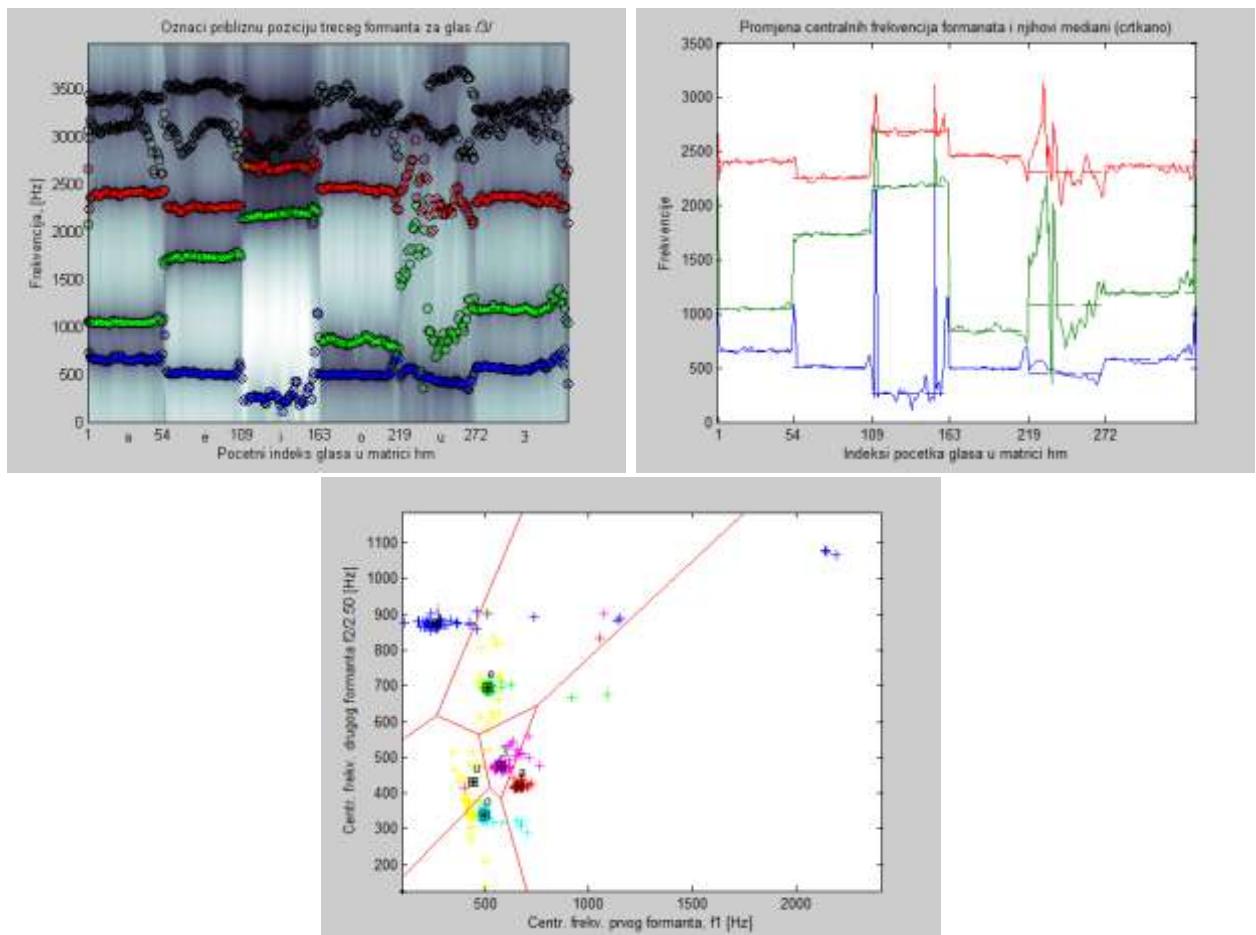
Median za svaki formant i svaki glas

glas	f1	f2	f3
------	----	----	----

-----+-----

a	659	1047	2410
e	509	1737	2261
i	264	2181	2687
o	496	2453	3095
u	439	1076	3038
3	572	1184	2364

rez_files\sp0_1



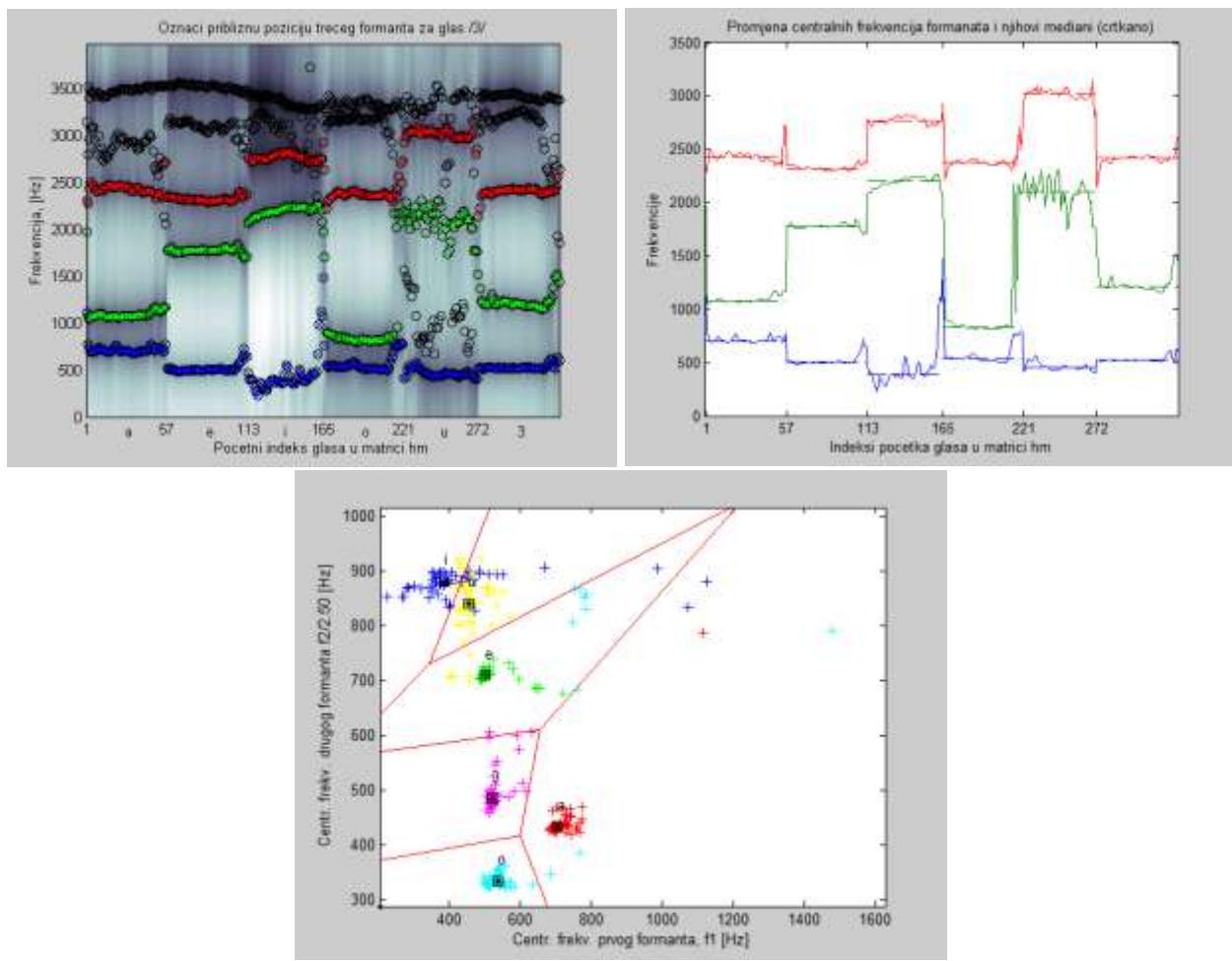
Median za svaki formant i svaki glas

glas	f1	f2	f3
------	----	----	----

-----+-----

a	658	1047	2410
e	510	1737	2261
i	265	2182	2689
o	496	845	2457
u	441	1076	2309
ɜ	574	1188	2358

rez_files\sp0_2



Median za svaki formant i svaki glas

glas	f1	f2	f3
------	----	----	----

-----+-----

a	706	1083	2436
---	-----	------	------

e	502	1776	2320
---	-----	------	------

i	389	2205	2767
---	-----	------	------

o	539	835	2375
---	-----	-----	------

u	457	2104	3029
---	-----	------	------

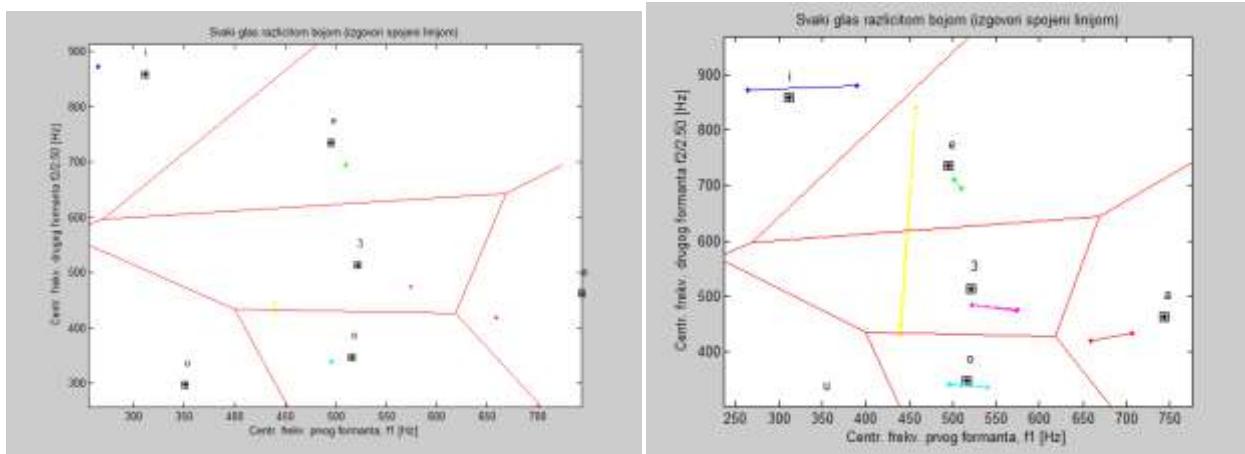
ɔ	522	1211	2417
---	-----	------	------

rez_files\sp0_3

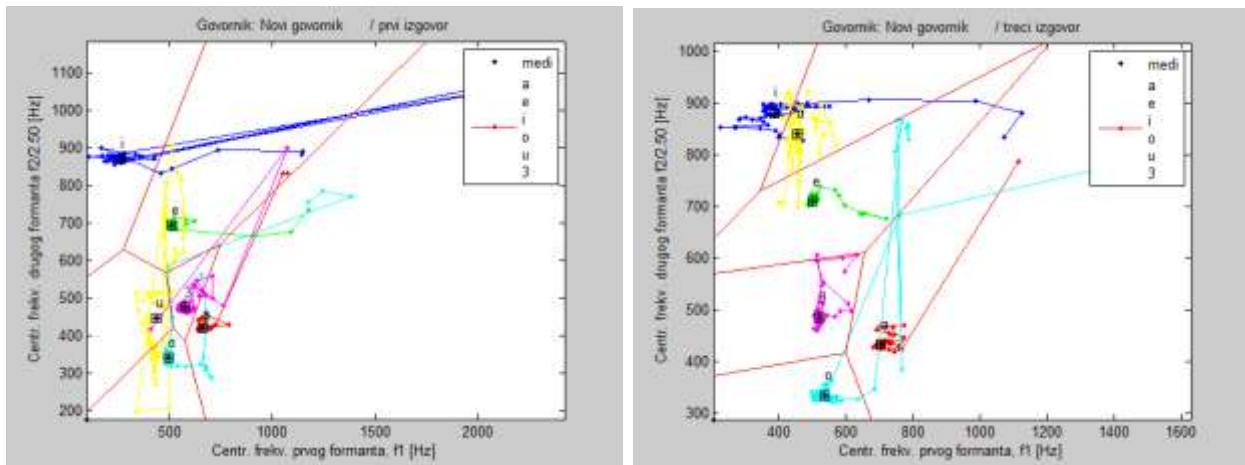
Zadatak 2.:

- a) Nakon uspješno provedene formantne analize svih vaših izgovora samoglasnika, primjenom programa 'dog6_prikaz.m' prikazati zbirne rezultate za sve vaše pravilne izgovore. Provesti prikate za oba slučaja klasifikatora (fiksног, kao i klasifikatora određenog na osnovu vaših podataka).

Klasifikacija fiksним klasifikatorom:



Klasifikacija s izračunatim klasifikatorom:



- b) U tablicu upišite koji su samoglasnici prepoznati (u koju su regiju pali) za oba slučaja klasifikatora.

		Prepoznati samoglasnik	
Stvarno izgovoreni glas		Prvi izgovor	Drugi izgovor
A	A	A	A
E	E	E	E
I	I	I	I
O	O	O	O
u	3	E	E
3	3	3	3

6. Automatska klasifikacija govornika na osnovu formantne strukture

Zadatak 1.: Za isti primjer s početka ovog poglavlja, sa 5 odabranih govornika i tri izgovora po govorniku, potrebno je nakon što program izračuna matricu ms prikazati težine $w_{i,j}$ normirane tako da najmanja težina iznosi 1. Diskutirati koji parametri imaju najveću ponovljivost, a koji najmanju (težina je recipročna vrijednost elementa matrice ms).

S obzirom da u zadnjem stupcu tablice minimalnih udaljenosti moraju biti strogo veće od maksimalnih udaljenosti različitih izgovora istih govornika, odabirom 5 govornika koji to zadovoljavaju možemo pronaći odgovarajuće vrijednosti. Tako odabirom najmanje vrijednosti matrice ms vidimo da je najmanja težina vezana uz glas „*u*“, dok je najveća vrijednost vezana uz glas „*e*“

Matrica =

0.1653	0.0602	0.2150	0.1212	0.0212	0.0736
0.2444	0.4530	0.5029	0.1206	0.0638	0.1526
0.0667	1.0000	0.1164	0.1847	0.1238	0.5601

Zadatak 2.

Pored navedenih govornika, u skup govornika uključiti i rezultate formantne analize za vaše izgovore samoglasnika (govornik s identifikacijskim brojem 0). Napomena, ... koristiti rezultate iz praktičnih zadataka prethodnog poglavlja. Provjeriti da li klasifikator radi pravilno i nakon uključenja vaših izgovora, te odrediti tko vam je najbliži i najdalji par.

Izgleda kako nije zadovoljen uvjet naveden u prethodnom zadatku, odnosno da u zadnjem stupcu tablice minimalnih udaljenosti moraju biti strogo veće od maksimalnih udaljenosti različitih izgovora istih govornika, klasifikator ne radi pravilno nakon uključenja vlastitog glasa.

Zadatak 3.

Odabrati sljedeći skup govornika:

odabrani=[1 2 3 4 5 6 7 8 12 14 15 16 17 18 20];

i fiksirati broj izgovora na 2, te ponoviti izvršenje programa za tako odabrane govornike.

Analizirati u koliko slučaja klasifikator ne radi pravilno i do kojih zamjena dolazi.

Klasifikator ne radi ispravno za snimku „Tomislava Ribarić“ rednog broja 12 ya koji je pogrešno klasificiran glas 5 (12 -> 5).

Zadatak 4.

Za isti gore navedeni skup govornika, prikazati matricu težina i usporediti je s rezultatima prvog zadatka koji je bio određen na osnovu samo pet govornika. Koji su glasovi najviše a koji najmanje ponovljivi ?

Vidljivo je kako je opet za glas „e“ najveća ponovljivost dok je za glas „u“ najmanja.

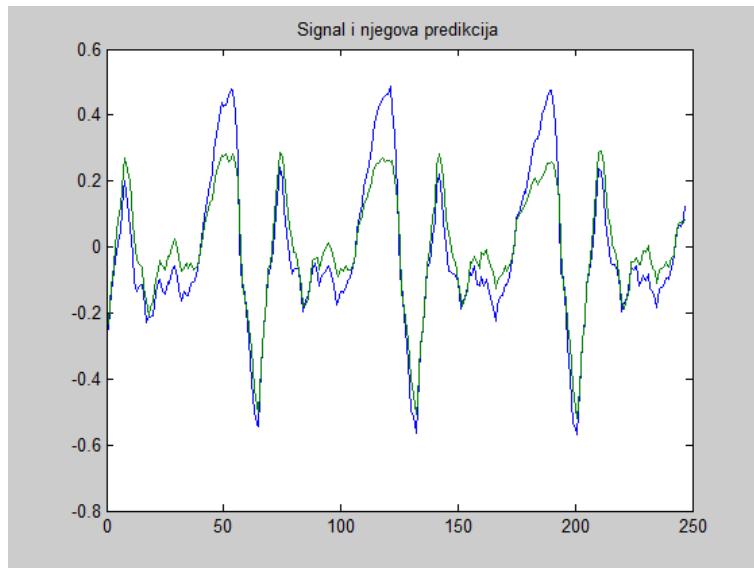
Matrica =

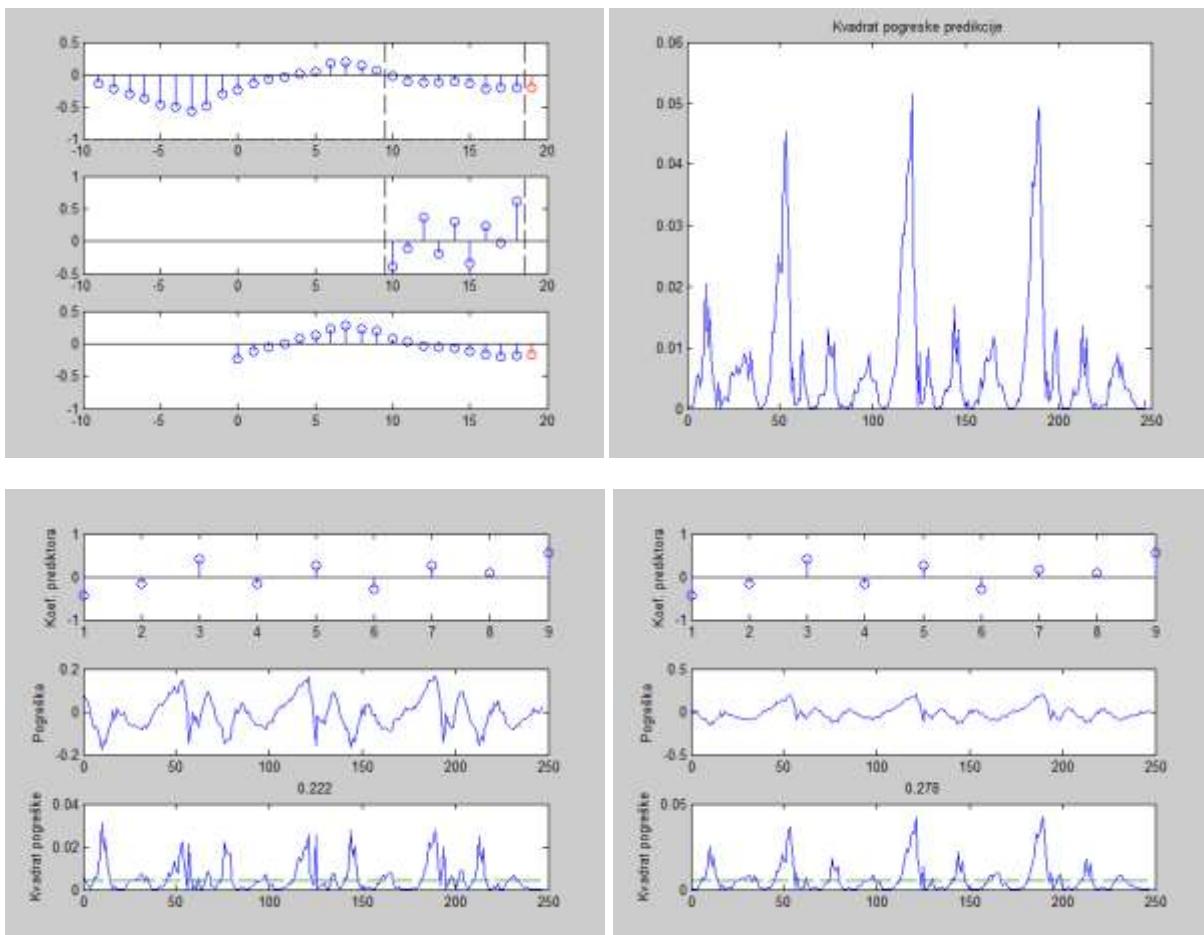
0.2951	0.1836	0.1320	0.2670	0.0653	0.2083
0.4065	0.6294	0.8227	0.2162	0.0886	0.2270
0.5239	1.0000	0.2668	0.4169	0.1900	0.4811

7. Postupci linearne predikcije

Zadaci za samostalno izvođenje

1. Prvi dio simulacija provedite za zajedničkoj govornoj sekvenci „amir1a.wav“ i „amir1a_9.mat“ i to na okviru analize 283, koji odgovara neutralnom samoglasniku „3“ (Schwa) uz pomoć programa „dog2.m“ i „dodatak.m“.



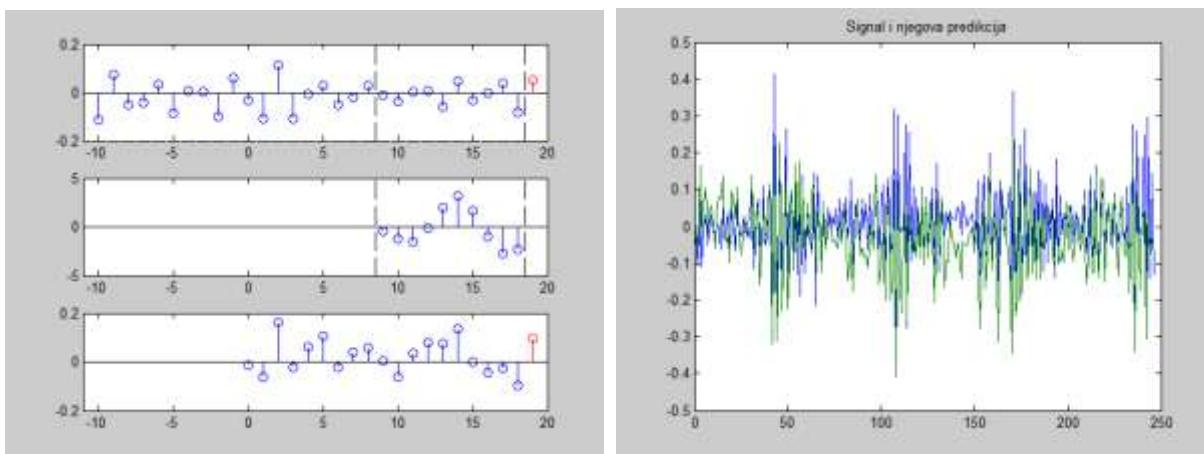


2. Ponovite eksperiment i za ostale glasove, te diskutirajte razlike.

Za glas „*a*“ sumarna predikcijska pogreška iznosi oko 13, glas „*e*“ oko 18, glas „*i*“ oko 8, glas „*o*“ oko 7 te za glas „*u*“ oko samo 0.5. Ovo smo i očekivali s obzirom da je glas „*u*“ u prethodnom zadatku pokazao najmanju težinu.

3. Ponovite simulaciju i na svojim glasovima

Provedena je simulacija nad glasom „*i*“.



Sumarna greska predikcije

1.9201

4. Kakva rješenja dobivate modificiranom mjerom pogreške predikcije?

Modifikacijom mjere pogreške možemo primjetiti kako se pogreška povećava na mjestima gdje se nalaze impulsi.