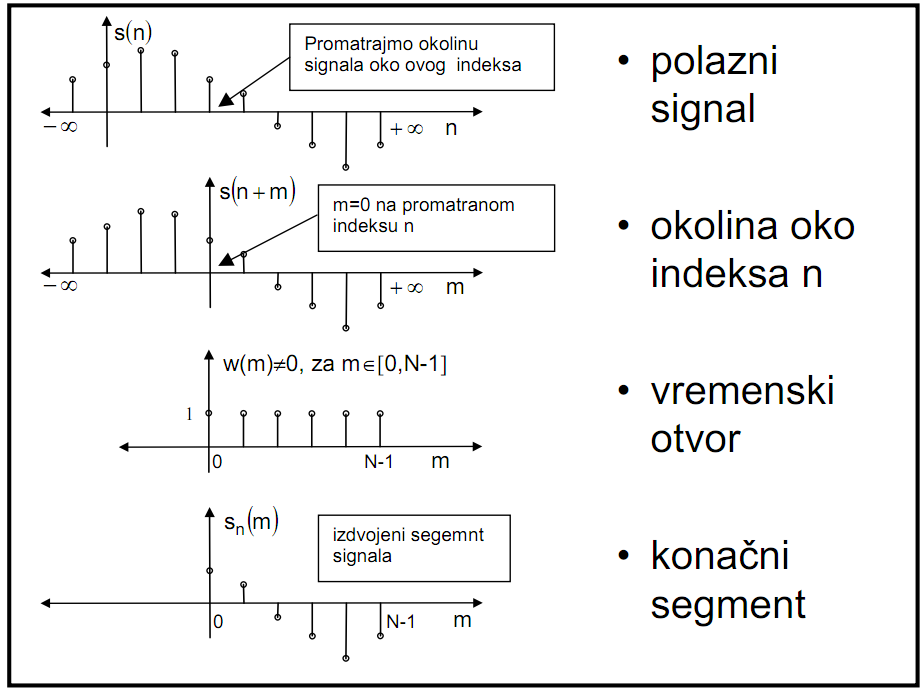
**DOG06**

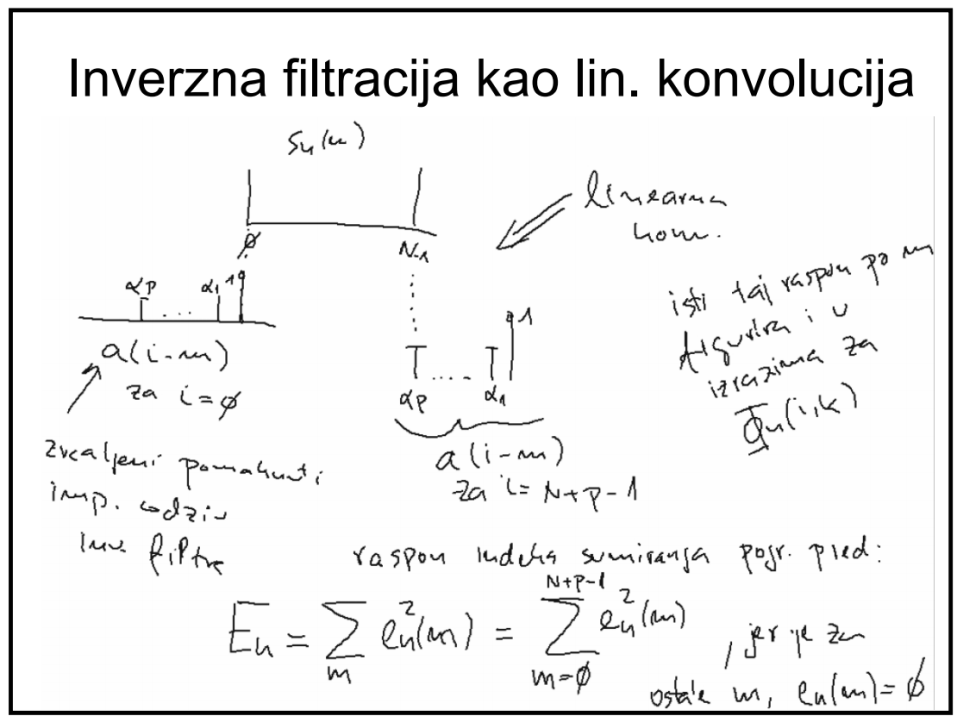
1. Objasni na koji način se kod linearne predikcije postupkom autokorelacije provodi definiranje (ograničavanje) intervala analize u pomoćnom indeksu m? Skiciraj grafički.

Pretpostavka je da je segment signala sn(m) jednak nuli izvan intervala 0 ≤ m ≤ N-1, što odgovara operaciji množenja signala s vremenskim otvorom sn(m) = s(m+n)\*w(m), gdje je w(m) vremenski otvor konačne dužine koji je jednak nuli izvan intervala 0 ≤ m ≤ N-1



2. Za vremenski ograničeni signal konačnog trajanja od N uzoraka i linearni prediktor reda p, skiciraj domenu (raspon vremenskih indeksa) za signal pogreške predikcije en(m). Koja je ukupna dužina signala en(m), tj. koliki je broj uzoraka općenito različit od nule?

Budući da je sn (m) različit od nule samo na intervalu 0 ≤ m ≤ N-1, odgovarajuća pogreška predikcije en(m) za prediktor p-tog reda bit će različita od nule samo na intervalu 0 ≤ m ≤ N-1+p.



3. Zbog čega kod autokorelacijskog postupka linearne predikcije postoji problem na rubovima? Koliko iznosi pogreška predikcije za indeks m=0, a koliko za m=N+p-1? Kakva se rješenja koriste da bi se ovi problemi rubova umanjili?

Pogreška predikcije je velikog iznosa na početku intervala (za 0 ≤ m ≤ p-1), zato što se pokušava provesti predikcija signala na temelju izvornih uzoraka koji su na silu postavljeni na nulu. Pogreška je također velika na kraju intervala (za N ≤ m ≤ N+p-1), jer se pokušava provesti predikcija uzoraka jednakih nuli na temelju uzoraka koji su različiti od nule.  
Jedan od načina kojim je moguće riješiti ovaj problem je primjena vremenskog otvora koji prigušuje signal prema rubovima segmenta sn(m), kao što je npr. Hammingov vremenski otvor.

Za m=0: En = en2 (0)  
Za m = N+p-1: En = en2(N+p-1)

4. Objasni kako i zašto se raspon sume u izrazu za elemente matrice sustava jednadžbi autokorelacijskom postupka može suziti. Koje su konačne granice koje se dobivaju takvim suženjem.

Uvodimo supstituciju m\* = m-i.

Budući da je sn (m\*) jednak nuli izvan intervala 0 ≤ m\* ≤ N-1, jednostavno se može pokazati da se granice sumacije sužavaju. Vrijedi:  
 • za donju granicu sn(m\*) = 0, m\* ≤ 0  
 • analogno za gornju sn(m\*+i-k) = 0, m\* + i -k ≥ 2N  
-> na osnovu toga stvarni raspon indeksa je 0 ≤ m\* ≤ N −1− ( i-k )

5. Izvedi izraz za elemente matrice sustava jednadžbi autokorelacijskog postupka fin(i,k). Koja funkcija se dobiva kao rješenje ovog izvoda i u kojem argumentu?

Dobije se vremenski kratkotrajna autokorelacijska funkcija (eng. short-time autocorrelation function) izračunata za pomak (i-k): Rn(i-k).  
Funkcija Rn(j) je parna funkcija u pomaku j, tj. Rn(j) = Rn(-j), vrijedi:

6. Napiši matrični oblik jednadžbe za optimalne koeficijente linearnog prediktora pomoću autokorelacijskog postupka. Napiši izraz pomoću kojeg se računaju elementi ovog sustava jednadžbi.

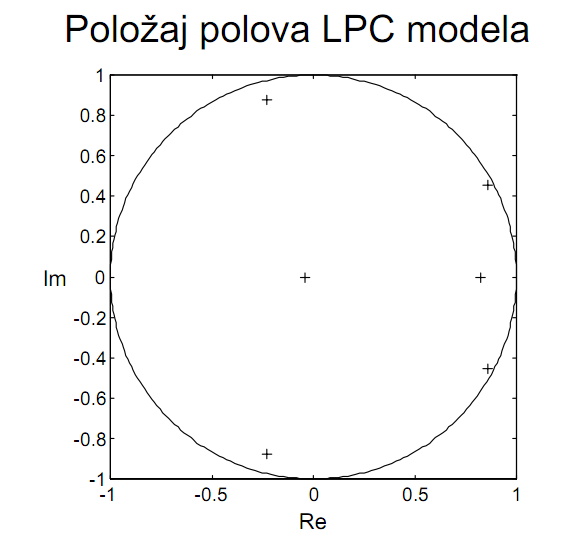
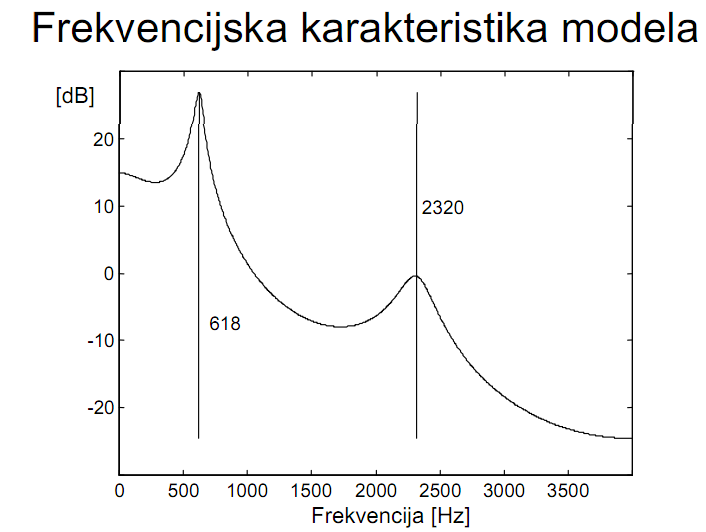
7. Koliki je broj različitih skalarnih veličina koje su potrebne za potpuno formiranje sustava jednadžbi autokorelacijskog postupka? Koje su to veličine? Kako nazivamo tip matrice takve strukture i koje je to njeno temeljno svojstvo strukture?

Postoji p autokorelacijskih koeficijenata koji definiraju cijeli sustav jednadžbi. Te veličine su koeficijenti.  
Matrica autokorelacijskih vrijednosti je Toeplitzove forme (simetrična po obje glavne dijagonale) dimenzija p\*p. Također, vrijednosti elemenata duž svake desno položene dijagonale su jednake.

8. Napiši izraz pomoću kojeg je moguće odrediti ukupnu kvadratnu pogrešku predikcije autokorelacijskog postupka na osnovu optimalnih koeficijenata alfa i koeficijenata autokorelacije za prediktor reda p.

Rn(0) je energija signala, dok su u sumi skalarni produkti vektora i [Rn(k)].

9. Diskutiraj stabilnost modela H(z) koji se dobiva autokorelacijskim postupkom. Kakve polove može imati ovaj model (koje tipove)? Ilustriraj odnos položaja konjugirano kompleksnih polova i formantnih karakteristika modela. Kako na osnovu položaja pola u kompleksnoj ravnini možemo odrediti centralnu frekvenciju i širinu pojasa formanta?

H(z) je stabilan sustav, svi polovi su unutar jedinične kružnice. Ima šest polova: dva realna i dva konjugirano-kompleksna para. Svaki konjugirano-kompleksni par polova određuje jednu rezonantnu karakteristiku koju se poklapa s jednim od formanata govornog signala. Polovi blizu jedinične kružnice imaju veliki Q-faktor i daju izražene maksimume u prijenosnoj funkciji (odgovaraju uskim formantima).

10. Zašto kažemo da LPC model modelira spektralnu ovojnicu vremenski kratkotrajnog spektra govornog signala. Kakva je točnost ovog modeliranja u vrhovima formanata (spektralni maksimumi) u odnosu na točnost modeliranja u spektralnim dolinama između formantnih karakteristika?

LPC model modelira spektralnu ovojnicu vremenski kratkotrajnog spektra govornog signala jer prediktor opisuje sporo promjenjivi dio spektra uzduž frekvencijske osi. Točnost modela najveća je u vrhovima formanata, a manja u spektralnim dolinama.

11. Objasni zašto kažemo da inverzni filtar provodi spektralno izbjeljivanje govornog signala. Koji je odnos log-spektra ulaznog signala i log-spektra signala pogreške predikcije? Kako red prediktora utječe na stupanj izbjeljivanja spektra?

Inverzni filtar provodi spektralno izbjeljivanje jer su vrhovi svih latica približno na istoj visini (unutar 15dB). Što je red prediktora viši, to je bolje izbjeljivanje. Spektar pogreške predikcije računa se kao razlika logaritamskog spektra signala i amplitudno-frekvencijske karakteristike prediktora u logaritamskoj mjeri. Spektar odgovara kvocijentu modela spektra signala i modela frekvencijske karakteristike LPC modela (H(z)).

12. Objasni u kojem su odnosu autokorelacije izdvojenog segmenta govornog signala i autokorelacija impulsnog odziva modela H(z) koji je određen autokorelacijskim postupkom nad tim izdvojenim segmentom.

Autokorelacije izdvojenog segmenta govornog signala i autokorelacija impulsnog odziva modela H(z) su identički jednake za pomake od –p do p, dok se izvan tog područja razlikuju.

**DOG07**

1. Po čemu se razlikuje postupak ograničavanja signala kod autokorelacijskog postupka i kod metode kovarijance? Zašto kod metode kovarijance ne postoje problemi s rubovima?

Kod metode kovarijance ograničavamo samo interval nad kojim računamo pogrešku predikcije, a signal ostavimo kakav jest. Nema problema s rubovima jer svi uzorci signala sudjeluju u određivanju prediktora jednakom “težinom”, te se ne koristi vremenski prozor.

2. Kako se definira raspon indeksa m kod izdvajanja signala u metodi kovarijance? Koji uzorci signala (raspon vremenskih uzoraka) su potrebni za izračunavanje svih elemenata matrice sustava jednadžbi? Napiši izraze iz kojih je to jasno vidljivo.

Odabiremo pomoćnu varijablu m’=m-i, m’=m-k:

Uzorci signala potrebni za izračunavanje sn(m) su vrijednosti van interval

3. Koja svojstva ispunjava matrica sustava jednadžbi za metodu kovarijance? Da li je Toeplitz strukture?

Matrica ispunjava svojstvo simetričnosti.  
Matrica je dimenzije p\*p koja sadrži kroskorelacije za sve pomake i i k je simetrična, ali nije Toeplitzove strukture.

4. Izvedi i opiši odnos elemenata matrice sustava jednadžbi za metodu kovarijance na desno ležećim dijagonalama. Radi čega je ovaj odnos značajan, tj. za što se može iskoristiti?

Na glavnoj dijagonali su elementi koji nose “energiju signala” – energija N susjednih elemenata uz p pomaka.

-> Bilo koja dva elementa matrice sustava koji su u odnosu i su vrlo slični po korelacijskom značenju.  
Svi elementi u desno lež. dijagonalama povezani su izrazom:

5. Opisati efikasni postupak formiranja matrične jednadžbe za metodu kovarijance, na osnovu kros-korelacija fin(i,0) za i=0 do p. Obrazloži numeričku složenost izračuna ovog stupca kros-korelacija, kao i složenost izračuna svih ostalih elemenata matrice sustava jednadžbi.

Korištenjem originalnog izraza izračunamo za i = 0, 1, …, p a sve ostale elemente sustava na glavnoj dijagonali i ispod nje mogu se odrediti rekurzivnim izrazom:  
Elementi iznad glavne dijagonale se pronalaze zrcaljenjem oko glavne dijagonale (zbog simetrije).  
Broj operacija za elemente koji se određuju direktnim izrazom je (p+1)\*(Nmnoženja+(N-1)zbrajanja), a za svaki preostali element, rekurzivni izraz traži p2/2 operacija.

6. Koje svojstvo mora ispunjavati matica sustava jednadžbi da bi postupak određivanja rješenja mogli provesti metodom Cholesky dekompozicije? U koji oblik se rastavlja matrica FIn? Opiši strukturu pojedinih dijelova ovog rastava.

Mora ispunjavati svojstvo da je pozitivno definitna, tj. da su njezine vlastite vrijednosti pozitivne, te da je simetrična.

* matricu rastavljamo kao produkt tri matrice V, D i VT
* svojstva tih matrica:
  + V je donja trokutasta matrica s jedinicama na glavnoj dijagonali
  + D je dijagonalna matrica
  + VT je transponirana matrica V, odnosno gornja trokutasta matrica s jedinicama na glavnoj dijagonali

7. Izvedi i opiši dva linearna sustava jednadžbi koji se dobivaju postupkom Cholesky dekompozicije. Koji je redoslijed rješavanja ova dva sustava?

Uvodimo pomoćni stupac dimenzije p, napravimo dva sustava jednadžbi, prvi u pomoćnom stupcu Y i drugi u stupcu konačnog rješenja:  
1. sustav:   
2. sustav:

Prvo odredimo Y iz prvog sustava, a zatim iz drugog.

8. Ilustriraj postupak rješavanja prvog (pomoćnog) sustava jednadžbi za metodu Cholesky dekompozicije u pomoćnoj vektorskoj varijabli Y. Kako nazivamo takav postupak rješavanja linearnog sustava? Napiši izraz za rješenje yi.

– za općenit red p vrijedi rekurznivni izraz:

Unaprijedna supstitucija – rješenje iz prošlog koraka uvrštavamo u izraz za sljedeće rješenje.

9. Ilustriraj postupak rješavanja drugog sustava jednadžbi za metodu Cholesky dekompozicije u konačnoj vektorskoj varijabli alfa. Kako nazivamo takav postupak rješavanja linearnog sustava? Napiši izraz za rješenje alfa\_i.

Nakon određivanja rješenja Y, uvrštavamo ga u drugu jednadžbu: i dobivamo konačni stupac s koeficijentima predikcije .  
Za općenit red p vrijedi rekurzivni izraz:

Koristi se unazadna supstitucija jer se kreće od zadnjeg elementa, te iterativno ide prema prvom.

10. Za primjer linearnog prediktora trećeg reda prikaži matricu sustava jednadžbi FIn čiji su elementi izraženi na osnovu elemenata pomoćne matrice V i dijagonalne matrice D (za metodu Cholesky dekompozicije). Kako nalazimo prvi dijagonalni element matrice D?

Određivanje pomoćnih matrica:  
 =

11. Napiši pomoćne izraze A i B koji se koriste za izračunavanje elemenata matrica V i D za metodu Cholesky dekompozicije. Objasni način primjene ovih izraza, tj. redoslijed kojim se dobivaju nepoznati elementi matrica D i V.

Primjena izraza: naizmjenična primjena B i A:  
 - iz izraza B odredimo d1 - zatim iz izraza A odredimo Vi,1 za i=2, …, p: V1,1=, ,   
 - iz prvog stupca i određenog elementa d1 određujemo prvi stupac matrice V:

1. d1 2. V21, V31, … Vp1 3. d2 4. V32, V42, … Vp2 itd. dok ne nađemo sve elemente.

12. Izvedi izraz za sumarnu kvadratnu pogrešku predikcije optimalnog linearnog prediktora reda p za metodu kovarijance koji je izračunat pomoću Cholesky dekompozicije. Kako se može odrediti pogreška predikcije za prediktor čiji je red p-1?

općenito

Za metodu kovarijance:

Gornjim izrazom se također može računati pogreška predikcije za prediktor reda p-1.

13. Objasni kojim postupkom se složenost postupka izračuna korelacijskog stupca fin(i,0) za i=0 do p za metodu kovarijance može dodatno smanjiti. Koja je ušteda u broju potrebnih operacija ako je širina okvira analize jednaka 250 uzoraka, a korak analize jednak 100 uzoraka.

Složenost se može smanjiti postupkom pomaka okvira analize za N/3 koraka u desno. Tada prve dvije parcijalne sume (u već imamo od prošlog okvira analize i ne računamo ih ponovno.

Time smanjujemo složenost s faktorom 3 (za preklom 2/3). Ušteda u broju operacija je tada:  
N\*(p+1) = 250\*(10+1) = 2750 MAC operacija  
Ušteda je 2750/3 = 916 operacija

14. Diskutiraj stabilnost prediktora određenog postupkom kovarijance, odnosno utjecaj valnog oblika ulaznog signala koji se modelira LPC prediktorom na stabilnost modela.

Ne postoji garancija na stabilnost prediktora. Do pojave nestabilnog LPC modela dolazi u slučajevima kada je okvirom analize zahvaćen segment signala koji se raspiruju prema kraju.

15. Kako širina vremenskog otvora analize N utječe na sličnost postupka kovarijance i autokorelacije? Koja od ove dvije metode daje točnija rješenja ako se koriste vrlo uski vremenski otvori analize koji su jednaki ili neznatno širi od reda prediktora? Koja od tih metoda može savršeno identificirati koeficijente nepoznatog all-pole modela i pod kojim uvjetima?

Ako je širina okvira analize N jako velika u odnosu na red p, tada je rezultat koji se dobiva metodom kovarijance vrlo blizak rezultatu autokorelacijskog postupka s pravokutnim otvorom. Za uske vremenske otvore, točnija rješenja dobit ćemo ako koristimo kovarijantni postupak. Kad je signal nad kojim se provodi analiza uistinu nastao kao impulsni odziv all-pole sustava, takva kovarijantna analiza na okviru od N=p uzoraka će savršeno točno identificirati nepoznate koeficijente modela.

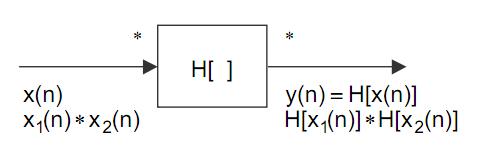
**DOG08**

1. Objasni ključne sličnosti i razlike postupka identifikacije modela vokalnog trakta pomoću linearne predikcije u usporedbi s postupkom homomorfne analize.

Model LPC određuje model vokalnog trakta na osnovu njegovod odziva (pobudni signal nije poznat). Izlaz (govor) je jednak konvoluciji pobude i impulsnog odziva modela. Dekonvolucija – razdvojiti pobudu i impulsni odziv nepoznatog sustava.  
Homomorfna analiza je postupak razdvajanja dvije konvoluirane komponente, tako da se jedna ukloni (zbog svojstva signala da ima relativno ravan spektar: bijeli šum, niz harmonika).

2. Definiraj pojam sustava koji su homomorfni za konvoluciju. skiciraj grafički i označi ulaze i izlaze takvih sustava. Što je generalizirani princip superpozicije?

To su sustavi kji podliježu poopćenom principu superpozicije, tj. sustavi H() koji zadovoljavaju svojstvo da možemo + zamijeniti s \* (generalizirani princip superpozicije).

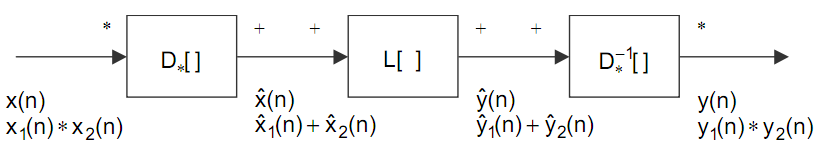


Prolaskom dvije konvoluirane komponente kroz H() na izlazu se dobiva konvolucija individualnih transformacija svake pojedine komp.

3. Što je homomorfni filtar i koja je njegova uloga?

Sustav H() tako projektiran, da transformacija jedne komponente bude bliska jediničnom impulsu , tada je na izlazu konvolucija što je jednako ; tako da se odvaja samo druga komponenta, a prva eliminira.

4. Prikaži i objasni kanonsku izvedbu sustava homomorfnih za konvoluciju. Koje su prednosti takve izvedbe?



D\*() – karakteristični sustav za homomorfnu dekonvoluciju

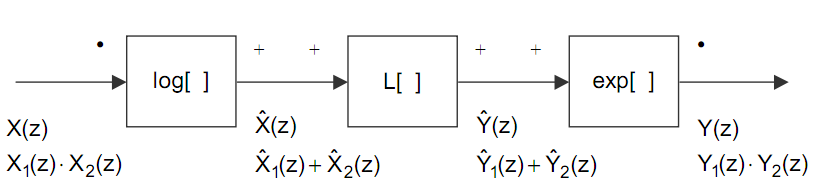
Kaskada tri sustava:  
D\*() konvoluciju pretvara u sumu, L() običan linearni sustav, D\*-1() inverzni sustav za homomorfnu dekonvoluciju, sumu pretvara u konv.

Prednosti: karakteristika ukupnog sustava određena je središnjim linearnih sustavom L() koji se projektira za željenu primjenu: sustav D\*() se ne mijenja, a podliježe generaliziranom principu superpozicije, gdje je ulaz konvolucija, a izlaz zbrajanje. Inverzni karakteristični sustav D\*-1() također je fiksiran sustav, koji transformira aditivnu kombinaciju signala natrag u konvolucijsku kombinaciju izlaza.

5. Prikaži i objasni što je karakteristični sustav za homomorfnu dekonvoluciju. Kako se definira njegov inverz?

To je sustav D\*() koji konvoluciju pretvara u sumu.

6. Skiciraj izvedbu karakterističnog sustava za homomorfnu dekonvoluciju temeljenu na z-transformaciji. Kojom operacijom se umnožak z-transformacija pretvara u aditivnu kombinaciju?



Umnožak z-transformacije pretvara se u aditivnu kombinaciju primjenom logaritma.

7. Kako se definira logaritam kompleksne veličine koji ispunjava tražene uvjete generaliziranog principa superpozicije. Dokaži da je uz takvu definiciju traženi uvjet uistinu ispunjen.

Dokaz:

8. Zbog čega problem fazne višeznačnosti predstavlja problem u definiciji kompleksnog logaritma i kako se taj problem rješava u sustavim homomorfnim za konvoluciju?

Problem fazne višeznačnosti predstavlja problem jer je imaginarni dio jednak kutu z-transformacije na jediničnoj kružnici, dok je kod numeričkog izračuna taj kut uvijek u rasponu od 0 do 2pi, pa dva ulazna argumenta X1 i X2 mogu dati jednaki kompleksni logaritam umnoška.  
To se rješava tako da fazna frekvencijska funkcija oba spektra X1 i X2 mora biti kontinuirana neparna funkcija od ; i na mjestima faznih skokova potrebno je izračunatu osnovnu faznu vrijednost nadopuniti višekratnicima od 2pi -> PHASE-UNWRAPPING.

9. Skiciraj i objasni izvedbu karakterističnog sustava za homomorfnu dekonvoluciju temeljenu na Fourierovoj transformaciji vremenski diskretnog signala (DTFT).

Fourierova transformacija također pretvara konvoluciju u umnožak, ali modelira ponašanje sustava samo u stacionarnom stanju. Izlaz karakterističnog sustava D\*() je:

10. Što je Kepstar i kako se definira? Kako se definira kompleksni, a kako realni kepstar? Tko je zaslužan za uvođenje pojma Kepstar?

Kepstar je rezultat Fourierove transformacije logaritma spektra signala. Pojam kepstar uveo je A. Michael Noll 1967. godine.

Kompleksni kepstar je inverzna Fourierova transformacija kompleksnog logaritma Fourierove transformacije , tj. izlaz karakterističnog sustava za homomorfnu dekonvoluciju:  
Realni kepstar je inverzna Fourierova transformacija logaritma spektra snage signala:

11. U kojem su odnosu realni i kompleksni kepstar? Napiši izraz pomoću kojeg je moguće odrediti realni kepstar iz kompleksnog kepstra.

12. Koja svojstva ispunjavaju realni i kompleksni kepstar (simetričnost, realnost, ...)?

Oba su realne sekvence, jer i običan i kompleksni logaritam spektra daju konjugirane simetrične funkcije, pa je njihova IFT realna. Realni kepstar je uvijek simetrična sekvenca oko ishodišta zbog realnosti logaritma modula c[n] = c[-n].

13. Za koju svrhu koristimo kepstralnu analizu na govornom signalu? Koji su rezultati ove analize, te kako treba odabrati linearni operator u kanonskom obliku karakterističnog sustava za homomorfnu dekonvoluciju u svrhu nalaženja traženih rezultata?

Kepstralnu analizu koristimo zbog toga da saznamo koji dio pripada pobudi, a koji modelu vokalnog trakta. To je moguće jer je taj signal jednak konvoluciji pobudnog signala i impulsnog odziva vokalnog trakta.  
Rezultat analize je odvajanje dijela pobude i dijela koji pripada vokalnom traktu, a linearni susta nužno ovisi o svojstivma njegovih ulaznih signala, pa se pri izboru linearnog operatora moraju uzeti u obzir i svojstva ulaznih signala.

14. Za općenit rekurzivni sustav sa polovima i nulama unutar i izvan jedinične kružnice, napiši opći oblik prijenosne funkcije X(z) kao umnožak korjenih faktora. Izračunaj logaritam ove prijenosne funkcije.

Logaritam:

15. Pokaži kako se u vremensku domenu (tj. u domenu kompleksnog kepstra) preslikava logaritam jednog korjenog faktora brojnika općenitog modela X(z), za k-tu nulu unutar jedinične kružnice (koristiti Taylorov razvoj).

-> niz uzoraka na pozitivnim vremenskim indeksima n čije su amplitude

Doprinos k-te nule: i

16. Napiši konačni izraz za kompleksni kepstar implusnog odziva sustava X(z) definiranog na osnovu pozicije polova i nula unutar i izvan jedinične kružnice. Koji koeficijenti modela određuju iznos kepstra na pozitivnim vremenskim indeksima, a koji na negativnim?

ak i ck na pozitivnim vremenskim indeksima, bk i dk na negativnim vremenskim indeksima.

17. Izračunaj kompleksni kepstar impulsnog odziva FIR sustava s dvije konjugirano kompleksne nule unutar jedinične kružnice, na radijusu r, po kutom +/-fi.

18. Izračunaj kompleksni kepstar impulsnog odziva all-pole IIR sustava drugog reda s dva konjugirano kompleksna pola unutar jedinične kružnice, na radijusu r, po kutom +/-fi. (?)

19. O čemu ovisi brzina opadanja modula kepstra s udaljavanjem od ishodišta? Napiši izraz kojim se definira gornja granica (garantirano opadanje). Kako izgleda kepstar sustava čiji su polovi ili nule na jediničnoj kružnici? (?)

Brzina opadanja modula kepstra s udaljavanjem od ishodišta ovisi o maksimalnim modulima svih koeficijanata za (eksponencijalno prigušenje kepstralnih uzoraka).

– maksimalni modul koeficijenata

- faktor proporcionalnosti

20. Što su sustavi minimalne faze? Koja svojstva ispunjava kompleksni kepstar impulsnog odziva sustava minimalne faze?

Sustavi minimalne faze su sustavi koji imaju samo polove i nule unutar jedinične kružnice (potpuno su opisani koeficijentima ak i ck, te faktorom pojačanja A).  
Kompleksni kepstar takvog sustava ispunjuje svojstvo da je fazno-frekvencijska karakteristika potpuno određena s amplitudno-frekvencijskom karakteristikom.

Ovaj kepstar može se odrediti direktno iz cmin[n]

21. Napiši i obrazloži kako i za koje sustave se kompleksni kepstar može jednoznačno odrediti iz realnog kepstra c[n].

Kompleksni kepstar može se jednoznačno odrediti iz realnog za sustave minimalne faze, tj. realni kepstar predstavlja parni dio kompleksnog kepstra.

22. Napiši rekurzivni izraz pomoću kojeg se uzorci kompleksnog kepstra impulsnog odziva sustava minimalne faze mogu odrediti direktno iz uzoraka tog impulsnog odziva.

23. Objasni kako analitički možemo odrediti kompleksni kepstar idealiziranog pobudnog signala vokalnog trakta koji se sastoji od niza kroneckerovih delta impulsa na razmaku Np. U kojim vremenski trenutcima se javljaju uzorci kepstra ovog signala?

Izrazimo P(z) kao produkt korjenih faktora oblika   
Kompleksni kepstar ima uzorke u vremenskim trenucima koji su cjelobrojni višekratnici Np.

24. Izvedi kompleksni kepstar i njegovu z-transformaciju za signal koji se sastoji od dva impulsa: delta[n] i alfa\*delta[n-Np]. Diskutiraj oblik i brzinu opadanja ovog kepstra.

Dobiveni beskonačni periodični niz impulsa s periodom Np i alternirajućim eksponencijalno padajućim amplitudama , amplituda impulsa teži prema nuli kako r raste (za ). Ne postoji impuls u n=0, nego tek u n=Np.

25. Objasni ukratko kako se kepstralna analiza može koristiti za utvrđivanje zvučnosti govornog signala i određivanje fundamentalne frekvencije titranja glasnica.

Analiza nad segmentima konačnog trajanja

Za zvučne glasove: V(z) ima samo polove (ak i bk su nule za svaki k).  
Kod nazalnih i bezvučnih glasova: V(z) ima i polove i nule  
 - za zvučne glasove (vokalni trakta -> FIR sustav) -> Hv(z) imat će polove unutar jedinične kružnice, a nule unutar i izvan jedinične kružnice (nije sustav minimalne faze). U kepstru šiljci na indeksima Np i njegovim cjelobrojnim višekratnicima. c(n) je simetričan oko nule.

Na osnovu pozicije izraženih maksimuma u kepstru zvučnih glasova moguće je odrediti frekvencije titranja glasnica (pitch detection), a u kepstru bezvučnih glasova šiljci ne postoje. Ako je veličina šiljka iznad nekog praga, ulazni govorni segment je zvučni pri čemu pozicija šiljka predstavlja dobru estimaciju osnovnog perioda, a ako je veličina šiljka ispod praga -> bezvučni.

26. Objasni kako izgleda i kako se odabire linearni sustav za izdvajanje pobudnog signala vokalnog trakta pomoću homomorfnog filtra. Koje svojstvo kepstra govora koristimo za ovo izdvajanje?

Koristimo svojstvo da su komponente kepstra razdvojene u vremenu, tj. dio blizak nuli (low time part) je određen s hv(n), a preostali dio koji je udaljen od nule (high time part), je određen s p(n) -> vremenska dislokacija.  
Govorni signal s(n) množi se s vremenskim otvorom w(n), pa se na tom segmentu primjenjuje D\*() čiji je izlaz kepstar cp(n), zatim na cp(n) primjenjujemo dva linearna operatora L1() i L2(), pa se ta dva dijela propuštaju kroz inverzni karakteristični sustav D\*-1(). Tako dobijemo dva izlazna signala h(n) i e(n).  
Ako pravilno odaberemo L1() i L2(), h(n) je impulsni odziv linearnog sustava, a e(n) je njegova pobuda.

**DOG09**

1. Objasni kako se kod kepstralne analize govora rješava problem nestacionarnosti govornog signala. Koju transformaciju koristimo umjesto sve-vremenske Fourierove transformacije vremenski diskretnog signala (DTFT)?

Problem nestacionarnosti rješava se u postupcima vremenski kratkotrajne analize – provodi se na kratkim odsječcima govora. Umjesto svevremenske FT, koristimo DFT i IDFT transformaciju koja je opisana sumama u oba smjera transformacije.

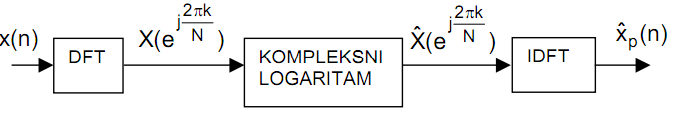
2. Koji su temeljni parametri kojima se definira vremenski kratkotrajna homomorfna analiza? Koje su tipične vrijednosti ovih parametara za analizu govornog signala?

Temeljni parametri: segment signala se izdvaja pomoću vremenskog otvora odgovarajuće širine i oblika (pomak ishodišta analize za odabrani broj uzoraka – korak analize), širina vremenskog otvora < koraka analize, dva susjedna izdvojena segmenta za dva susjedna okvira analize se nužno preklapaju, učestalost analize (broj okvira analize u sekundi), učestalost mora biti odabrana u skladu s vremenskom dinamikom izmjene spektralnih svojstava govora (i pobude vokalnog trakta).

3. Opiši i objasni numeričke postupke izračuna kepstra govora korištenjem diskretne Fourierove transformacije (DFT). Kako se bira broj uzoraka DFT-a i kako on utječe na svojstva tako izračunatog kepstra. (?)

Aproksimiramo računanje spektra pomoću DFT-a tako da zamijenomo Fourierove transformacije DFT-om:

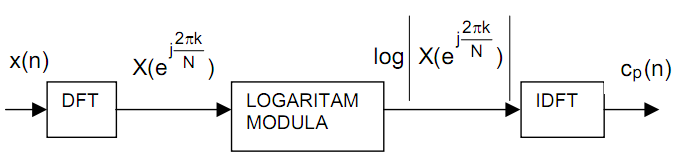
Otipkavanje u spektru -> periodičko proširenje u vremenu.



Kompleksni kepstar nalazimo (za segment signala x(n) dužine N uzoraka) kao:

Aproksimacija stvarnog kepstra:

Aproksimacija cp(n) i stvarni kepstar:



N mora biti velik (zbog aliasinga), da se osigura kontinuitet fazne karakteristike (N=512 ili više).

4. Objasni i formulom izrazi koji je odnos kepstra izračunatog pomoću svevremenske Fourierove transformacije (DTFT) izdvojenog odsječka govornog signala i kepstra izračunatog DFT-om u N točaka za isti taj segment trajanja manjeg od N uzoraka. (?)

Kepstar izračunat pomoću DFT-a možemo dobiti periodičkim ponavljanjem kepstra izračunatog korištenjem prave Fourierove transformacije svakih N uzoraka, tj. prema slijedećem izrazu:

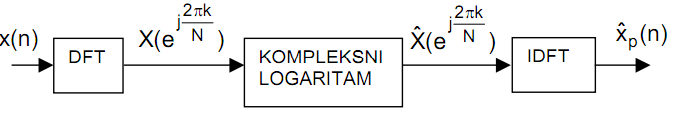
DTFT:

DFT:

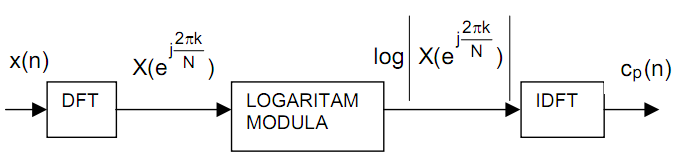
5. Koje svojstvo kepstra govornog signala opravdava korištenje DFT-a umjesto DTFTa, tj. zašto možemo ignorirati pojavu aliasinga u vremenskoj domeni? Za koji tip govornog signala je problem aliasinga izraženiji?

Pojavu aliasinga možemo ignorirati zbog toga što je model formiranja govornog signala temeljen na vrlo sporo promjenjivom linearnom sustavu pobuđenim s periodičnim nizom impulsa (ili slučajnim šumom).  
Najizraženiji je a zvučne govorne signale (sve osim nazalnih).

6. Skiciraj blok shemu i napiši izraze za postupak izračunavanja kompleksnog kepstra korištenjem DFT-a



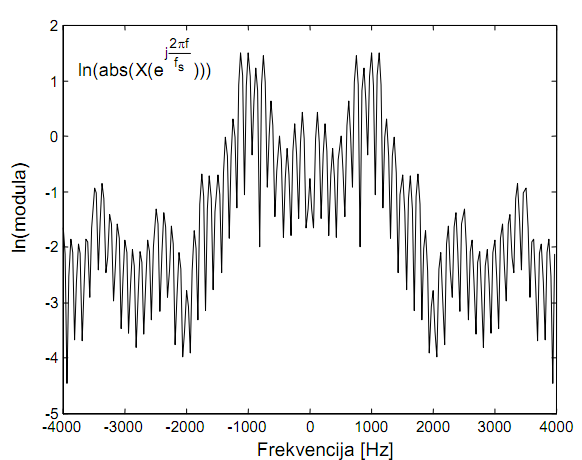
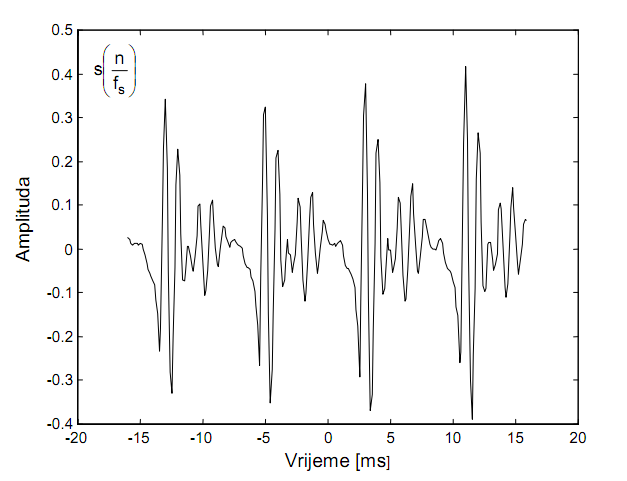
7. Skiciraj blok shemu i napiši izraze za postupak izračunavanja realnog kepstra korištenjem DFT-a. Koji je odnos ovog kepstra i kepstra koji se dobiva DTFTom?

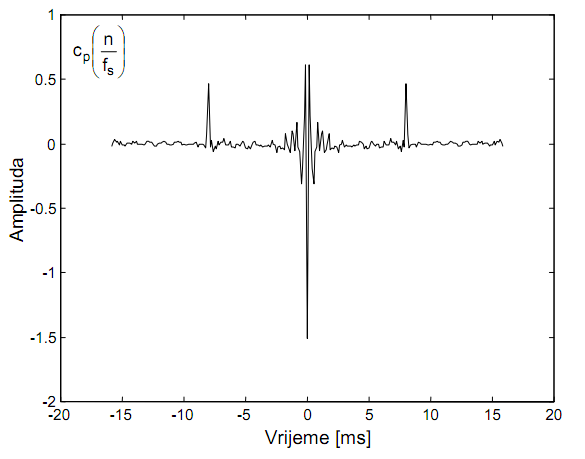


DTFT:

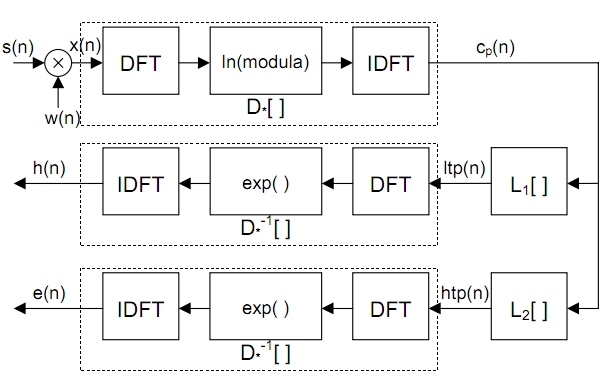
DFT:

8. Za primjer samoglasnika kvalitativno skiciraj: valni oblik izdvojenog segmenta govora, modul njegovog DFTa u log mjerilu i oblik realnog kepstra ovog segmenta. Označi koji dio realnog kepstra odgovara pobudi, a koji dio vokalnom traktu.



- veliki šiljak: pobuda, mali šiljci: vokalni trakt

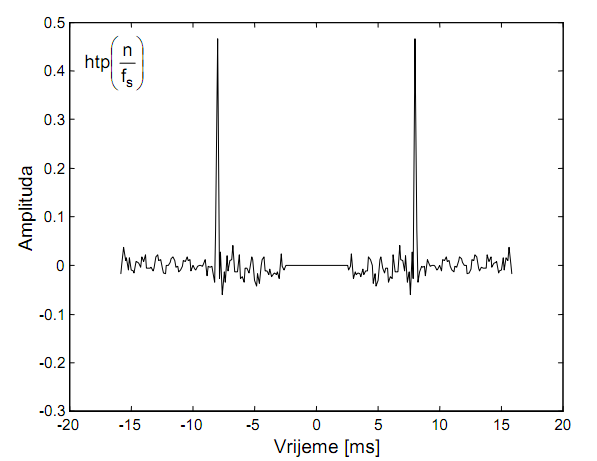
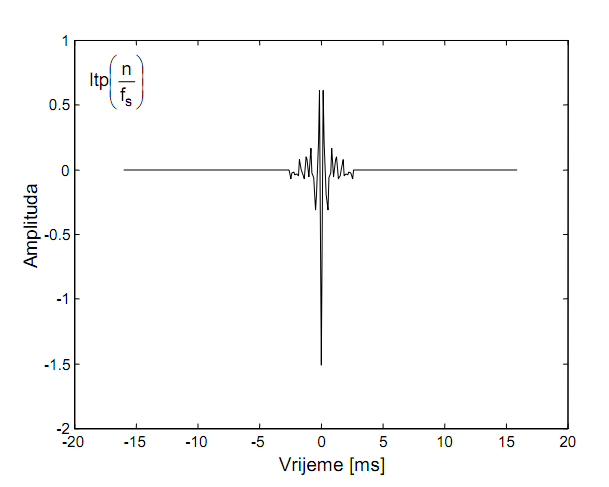
9. Skiciraj cjelokupnu blok shemu sustava za homomorfnu analizu govora kojim se govorni signal rastavlja u signal pobude i impulsni odziv vokalnog trakta. Kako su definirani linearni sustavi koji obavljaju ovo razdvajanje (L1, L2)?



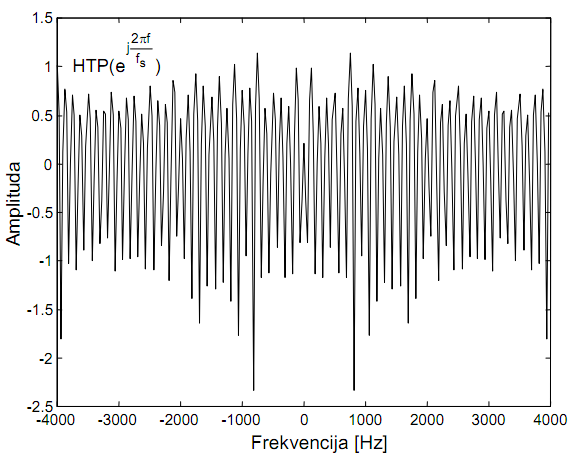
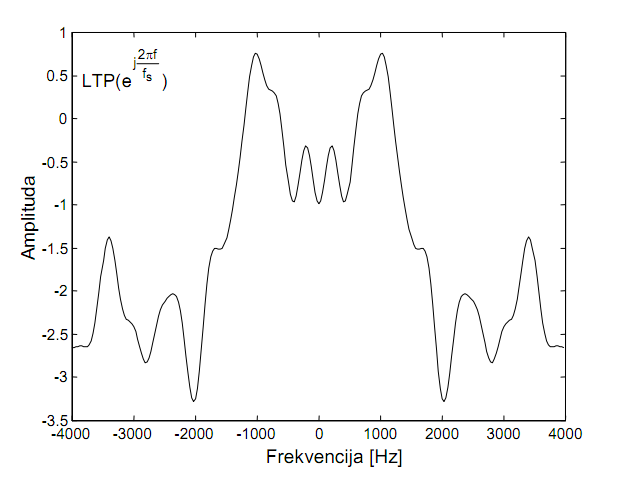
Niskovremenski dio, ltp, izdvaja dio oko ishodišta, a drugi dio je njegov komplement http:

|  |  |
| --- | --- |
|  |  |
|  |  |

10. Skiciraj oblik spektra niskovremenskog i visokovremenskog dijela kepstra samoglasnika. Što se dobiva zbrajanjem ova dva log-spektra? Skiciraj konačni vremenski oblik ova dva razdvojena dijela nakon propuštanja kroz inverzni sustav za homomorfnu dekonvoluciju. Čemu odgovara svaki pojedini od ta dva dobivena signala?



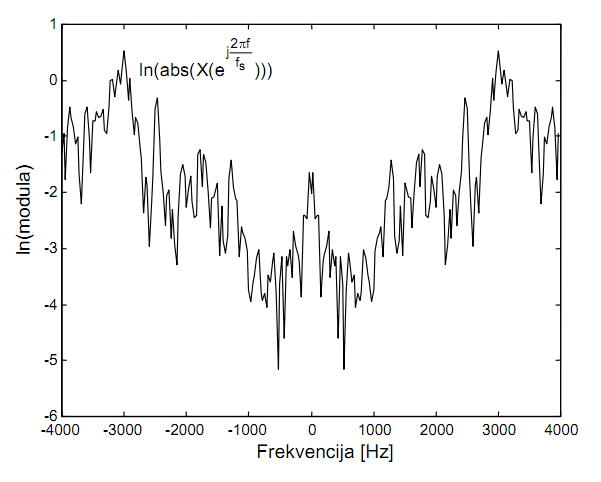
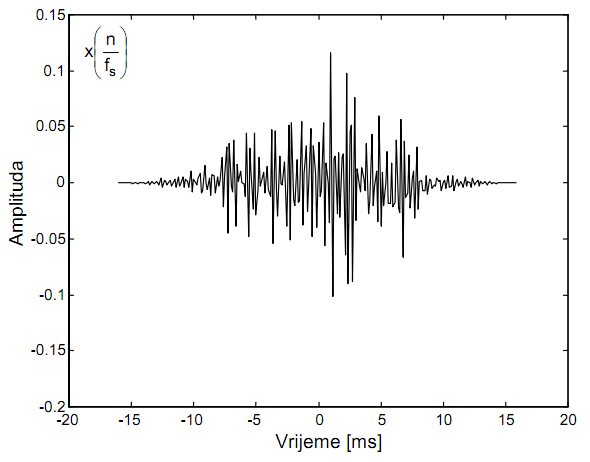
Niskovremenski (lijevo) i visokovremenski (desno) dio spektra. Njihovim zbrajanjem dobiva se jedinični operator, suma DFT-a od lpt i htp dijela jednaka je logaritmu modula DFT-a.

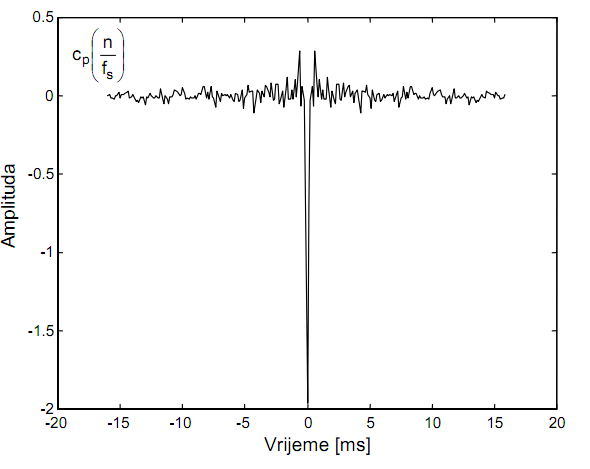
  
LTP dio kepstra opisuju sporo promjenjivu spektralnu ovojnicu govornog signala ( log – magnituda frekvencijske karakteristike modula vokalnog trakta), a HTP dio kepstra trebao bi biti jednag log spektru pobudnog signala.

11. Zbog čega se prilikom primjene sustava za homomorfnu analizu govora korištenjem realnog kepstra konvolucijom između estimirane pobude vokalnog trakta i njegovog impulsnog odziva dobiva signal koji nema jednak valni oblik kao i ulazni segment govora (tj. zašto nemamo potpunu rekonstrukciju)? U kojem smislu su ova dva signala ipak jednaki?

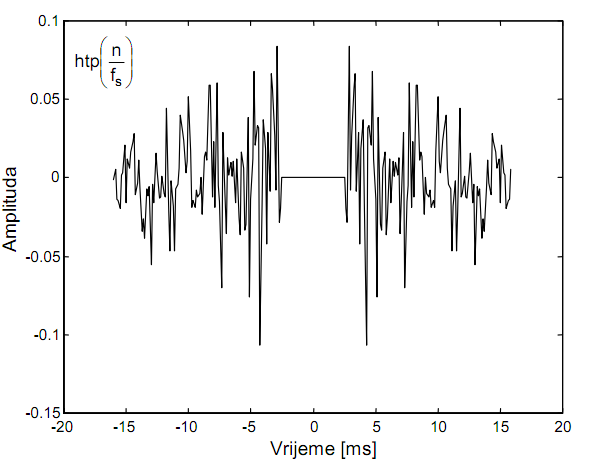
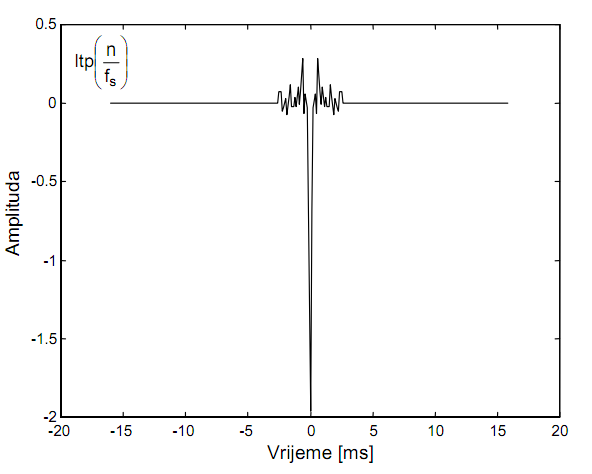
Dobiva se signal koji nema jednak valni oblik kao i ulazni segment govora jer u procesu odbacujemo faznu karakteristiku.  
Međutim, ova dva signala su ipak jednaka u smislu modula spektra: modul spektra konvolucije e(n)\*h(n) je identički jednak modulu spektra izdvojenog segmenta x(n).

12. Za primjer bezvučnog suglasnika (npr. glas 's') kvalitativno skiciraj: valni oblik izdvojenog segmenta govora, modul njegovog DFTa u log mjerilu i oblik realnog kepstra ovog segmenta. Označi koji dio realnog kepstra odgovara pobudi, a koji dio vokalnom traktu.

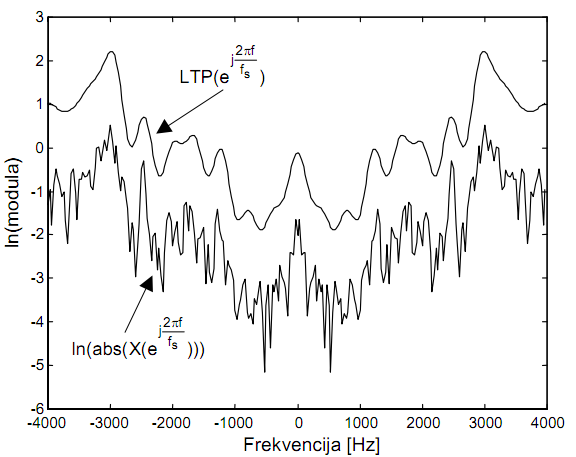


- centar: pobuda, nema šiljaka u htp dijelu

13. Skiciraj oblik spektra niskovremenskog i visokovremenskog dijela kepstra bezvučnog suglasnika. Što se dobiva zbrajanjem ova dva log-spektra? Skiciraj konačni vremenski oblik ova dva razdvojena dijela nakon propuštanja kroz inverzni sustav za homomorfnu dekonvoluciju. Čemu odgovara svaki pojedini od ta dva signala?



Zbrajanjem se dobiva spektralna ovojnica signala zajedno s prirodnim logaritmom modula DFT-a signala x(n) pomnoženog s vremenskim otvorom. LTP odgovara spektralnoj ovojnici signala (formatnoj strukturi govornog signala), a HTP dio ima oblik spektra bijelog šuma.



14. Objasni zašto se kao pobudi signal vokalnog trakta za slučaj homomorfne analize bezvučnog suglasnika korištenjem realnog kepstra dobiva signal koji izgleda kao jedinični impuls sa pribrojenim šumom male amplitude? Kako ovaj pobudni signal učiniti sličnijim stvarnoj bezvučnoj pobudi vokalnog trakta?

Kad se pobudi signal vokalnog trakta za slučaj homomorfne analize bezvučnog suglasnika korištenjem realnog kepstra dobiva se signal koji izgleda kao jedinični impuls s pribrojenim šumom male amplitude jer sve komponente DFT-a HTP dijela imaju nultu fazu, pa se idealno pribrajaju na vremenskom indeksu n=0 (koherentne su).  
Promjenom fazne karakteristike pobudnog signala (odabrali je slučajno između 0 i 2pi – fazu pojedinog spektralnog uzorka)

**DOG10 i DOG11**

1. Koje zadatke i elemente obuhvaća sustav za komunikaciju čovjeka i računala korištenjem govornog jezika? Koje su glave prednosti govora kao modaliteta komunikacije s računalom?

Glavni elementi su prepoznavanje govora, transformacija govora u tekst, razumijevanje teksta i shvaćanje konteksta, zatim generiranje teksta iz kojeg se radi sinteza govora.

Glavne prednosti govora kao modaliteta komunikacije: **Govor je prirodan** (Nije potrebna edukacija o korištenju jezika), **fleksibilan** (ruke i oči su slobodne), **efikasan** (ima visok prijenos podataka) i **ekonomičan** (transmisija i prijam govora su jeftini.

1. Nabroji i na primjerima objasni ograničenja koja proizlaze iz korištenja prirodnog govora u svrhu upravljanja računalom.

**Ograničenja su**: **akustična** (ljudski vokalni trakt), **fonetska** („let us pray“ zvuči slično kao „lettuce spray“), **fonološka** („gas shortage“ i „fish sandwich“), **fonotaktika** („blit vnuk“), **sintaksna** („I am flying to Chicago tomorrow“ i „tomorroy I am flying Chicago am to“), **semantična** („Is the baby crying“ i „Is the bay bee crying“) i **kontekstualna** („It is easy to recognize speech“ i „It is easy to wreck a nice beach“)

1. Nabroji moguće konačne primjene koje koriste prirodni govor kao modalitet komunikacije (u oba moguća smjera). Koji su glavni izazovi i problemi koji ograničavaju točnost automatskog prepoznavanja prirodnog govora?

**Primjene** su: *samo za upravljanje računalom* (npr. Glasovno upravljanje umjesto miša i tipkovnice, unos podataka i diktiranje umjesto pisanja), *interaktivna komunikacija čovjeka i računala* (npr. Automat za informacije, inteligentni posrednici).

Glavni izazovi i problemi su: ko-artikulacija, neovisnost govirnika (dialekti i ljudi čiji je govoreni jezik strani), spontani govor (sleng i sl.), modeliranje jezika i otpornost na buku.

1. Nabroji glavne parametre koji karakteriziraju sustav za automatsko prepoznavanje govora. Što je modeliranje jezika i zbog čega se koristi u sustavima za automatsko prepoznavanje?

Glavni parametri su akustični, leksički i jezični modeli i pretraživanje.

Modeliranje jezika je pridjeljivanje ocjene vjerojatnosti nekom nizu riječi. Ono pomaže ograničiti pretraživanje među alternativni objašnjenima riječi prilikom prepoznavanja.

1. Kako točnost prepoznavanja ovisi o broju riječi u rječniku, o tipu govora (prirodni vezani izgovor ili diktirani govor), o broju različitih korisnika sustava (vezano uz treniranje modela)?

Točnost prepoznavanja raste s porastom broja riječi u rječniku, porastom broja različitih korisnika sustava i kada se koristi diktirani govor.

1. Diskutiraj postupke treniranja (automatiziranog učenja) sustava za automatsko prepoznavanje govora. Zbog čega je nužno u svrhu gradnje akustičkih i leksičkih modela jezika koristiti govorne audio zapise i pripadne transkripcije iz stvarne očekivane primjene sustava?

Nužno je koristiti govorne audio zapise jer se iz tih snimki određuje skup riječi korišten u konverzacijama te se sustav trenira da konkretno prepoznaje samo riječi iz tog govornog skupa. Iz svih prikupljenih transkripcija razgovora potrebno je odrediti i tipove upita upućene od različitih korisnika.

Objedinjavanjem audio zapisa i transkripcija formiraju se statistički modeli koji se koriste za prepoznavanje svih elemenata dijaloga: osnovnih fonemskih grupa, riječi i rečenica.

1. Diskutiraj točnost prepoznavanja sustava za automatsko prepoznavanje govora u usporedbi s točnosti prepoznavanja ljudskih slušača. Kako točnost ovisi o tipu govora (čitanje znamenaka, radio vijesti, nevezani telefonski razgovor o proizvoljnoj temi, itd.)

Ljudski slušači imaju veću točnost.

1. Opiši utjecaj kvalitete snimljenog govornog signala na točnost prepoznavanja. Koje su tipične primjene gdje kvaliteta signala predstavlja ograničavajući faktor? Koje su mjere kojima se ova kvaliteta može poboljšati?

Mjere kojima se kvaliteta može poboljšati su poništenje akustičke povratne veze, poništenje jeke i povećanje kvalitete govora.

1. Opiši utjecaj tipa vremenskog otvora i njegove vremenske širine (broja uzoraka) na karakteristike vremenski kratkotrajnog spektra govornog signala, odnosno na pripadne spektrograme (uskopojasni i širokopojasni).

Ovisno o tipu vremenskog otvora, može doći do pojave spektralnog rasipanja,tj. spektralne komponente velikih amplituda prekriju svojim bočnim laticama spektralne komponente malih amplituda.

1. Objasni i slikom ilustriraj razliku u načinu modeliranja spektralne ovojnice govornog signala LPC postupkom u usporedbi sa vremenski kratkotrajni spektrogramom koji se dobiva korištenjem uskih vremenskih otvora (tj. za širokopojasni slučaj).
2. Objasni i ilustriraj način izdvajanja nisko-vremenskog dijela kepstra pravokutnim otvorom, u usporedbi s otvorom koji ima postepen prijelaz iz jedinice u nulu (tzv. prigušeni otvor ili Tapering Window). Kojim se otvorom dobiva bolje poklapanje sa spektralnom ovojnicom?

Kepstra se množi sa prvaokutnim otvorom (c(n)\*w(n)) za koji vrijedi da je abs(n)<n0

Nisko-vremenski dio kepstra jednoznačno opisuje sporo promjenjivu spektralnu ovojnicu.

1. Koji parametri se najčešće koriste kao vektor značajki u sustavima za automatsko prepoznavanje govora? Koji je tipičan broj tih parametara i njihova vremenska učestalost? Zbog kojeg razloga se koriste prve i druge vremenske derivacije ovih parametara?
2. Zbog čega linearna frekvencijska skala nije dobar odabir u sustavima za automatsko prepoznavanje govora? Kako se zovu često korišteni parametri za modeliranje vokalnog trakta u svrhu automatskog prepoznavanja koji su izračunati iznad nelinearne frekvencijske osi, te kako se provodi sam postupak analize?