

Drugi projektni zadatak

Linearno prediktivno kodovanje govora

Objavljen: 28.01.2015. godine

Rok za predaju: 13.02.2015. godine

Linearna prediktivna analiza je klasična, ali i jedna od najmoćnijih tehnika analize govora. Najčešće oblasti primjene linearne prediktivne analize su u estimaciji osnovnih parametara govora (fundamentalna frekvencija, formanti, spektar, itd.) i u kodovanju govora u cilju prenosa ili arhiviranja.

Osnovna pretpostavka linearne prediktivne analize je da su odmjerci govora korelisani i da se vrijednost odmjerkera može estimirati korišćenjem linearne kombinacije prethodnih odmjeraka. Ova tehnika je povezana i sa fizičkim modelom generisanja govora prema kojem je govorni signal izlaz iz vremenski promjenljivog linearnog filtra koji se pobuđuje periodičnom sekvencom impulsa (zvučni govor) ili bijelim šumom (bezvučni govor). U literaturi je pokazano da ovaj filter ima oblik:

$$H(z) = \frac{S(z)}{U(z)} = \frac{G}{1 - \sum_{k=1}^p a_k z^{-k}}, \quad (1)$$

gdje su a_k koeficijenti, G pojačanje, a p red filtra. Svi ovi parametri se sporo mijenjaju u funkciji vremena. Jednačina diferencija ovog filtra je oblika

$$s(n) = \sum_{k=1}^p a_k s(n-k) + Gu(n). \quad (2)$$

Linearni prediktor sa koeficijentima α_k je filter čija je jednačina diferencija

$$\hat{s}(n) = \sum_{k=1}^p \alpha_k s(n-k), \quad (3)$$

a funkcija prenosa

$$P(z) = \sum_{k=1}^p \alpha_k z^{-k}. \quad (4)$$

Greška predikcije, $e(n)$, se definiše kao

$$e(n) = s(n) - \hat{s}(n) = s(n) - \sum_{k=1}^p \alpha_k s(n-k). \quad (5)$$

Alternativno, greška predikcije je signal koji se dobija na izlazu sistema (filter greške predikcije) pobuđenog signalom $u(n)$ čija je funkcija prenosa

$$A(z) = 1 - \sum_{k=1}^p \alpha_k z^{-k}. \quad (6)$$

Iz jednačina (1) i (6) slijedi da je, u slučaju kada je $a_k = \alpha_k$, greška predikcije $e(n) = Gu(n)$, pa je filter greške predikcije (6) inverzni filter za filter kojim se generiše govorni signal (1)

$$H(z) = \frac{G}{A(z)}. \quad (7)$$

Zadatak linearne prediktivne analize je da se za govorni signal projektuje prediktor dat jednačinom (3), tj. da se, korišćenjem odmjera govornog signala, odrede njegovi koeficijenti α_k tako da se, uz pretpostavku $a_k = \alpha_k$, pomoću jednačine (7) dobije dobra estimacija spektra govornog signala. Naravno, pošto je govorni signal nestacionaran, ova analiza se mora vršiti na kratkim prozorima.

Koeficijenti prediktora se određuju minimizacijom srednjekvadratne greške, pri čemu se usrednjavanje vrši po vremenu

$$E_n = \sum_{m=0}^{N+p-1} e_n^2(m), \quad (8)$$

gdje indeks n označava vrijednosti dobijene u prozoru analize pomjerenom u trenutak n . Dužina prozora je N , a red prediktora p . Kao što je poznato, minimalna srednjekvadratna greška se postiže rješavanjem sistema Wiener-Hopfovih jednačina. Autokorelacione funkcije potrebne za formiranje sistema Wiener-Hopfovih jednačina se izračunavaju usrednjavanjem po vremenu. Nakon izračunavanja koeficijenata, pojačanje modela, G , se izračunava iz uslova da je energija ulaznog signala jednaka energiji E_n signala greške. Pošto se za svaki prozor analize projektuje novi prediktor, greška predikcije će na početku i na kraju prozora biti vrlo velika zbog prelaznih režima. Kako bi se ovaj problem ublažio neophodno je korišćenje prozorske funkcije koja će prigušiti vrijednosti signala na krajevima prozora analize.

U ovom domaćem zadatku potrebno je:

1. Formulirati Wiener-Hopfove jednačine za problem predikcije govora;
2. Napisati program u MATLAB-u koji će usrednjavanjem po vremenu, na osnovu odmjera govornog signala, izračunati autokorelacione funkcije za prozor analize i formirati sistem Wiener-Hopfovih jednačina;
3. Napisati program u MATLAB-u za rješavanjem sistema Wiener-Hopfovih jednačina i određivanje koeficijenata prediktora α_k , odnosno, modela a_k , kao i pojačanja modela G ;
4. Dobijeni model govornog signala testirati najprije na jednom segmentu govornog signala na kojem se može smatrati da je govorni signal stacionaran. Zanimljivo je uporediti rezultate dobijene npr. za segmente glasova A i S. Korišćenjem programa napisanih u prethodnim tačkama izračunati koeficijente modela;
5. Na istom grafiku nacrtati amplitudni spektar posmatranog segmenta govornog signala dobijen Furijeovom transformacijom i amplitudnu karakteristiku dobijenog modela (1). Frekvencijska osa treba da bude u Hz, a vrijednosti spektra u dB.
6. Izračunati i nacrtati signal greške (5). Uporediti dobijene rezultate za različite glasove.

7. Kako bi se ublažila neravnomjerna raspodjela energije u spektru govornog signala često se prije analize govorni signal filtrira visokopropusnim filtrom, kao što je na primjer

$$H_{vp}(z) = 1 - 0,98z^{-1} \quad (9)$$

Ponoviti analizu u tačkama 4. i 5. za signale filtrirane filtrom (9). Komentarisati dobijene rezultate, npr. sa stanovišta određivanja formantnih frekvencija.

8. U fizičkoj akustici se pokazuje da su polovi filtra (1), odnosno, nule polinoma $A(z)$ datog jednačinom (6) povezani sa formantnim frekvencijama. Zapravo se formantne frekvencije (izražene u rad/s) mogu estimirati kao argumenti konjugovano-kompleksnih polova filtra (1). Napisati program u MATLAB-u koji će za jedan primjer govornog signala određivati formantne frekvencije za pojedine segmente tog signala. Ulazni argumenti treba da budu: vektor odmjeraka signala, korak analize, dužina prozora i red prediktora. Izlaz treba da bude matrica u kojoj svaka kolona odgovara jednom prozoru analize i u njoj se nalaze formantne frekvencije izražene u Hz. Predložite mogući način vizuelizacije ovako estimiranih formantnih frekvencija kako bi se dobijeni rezultati mogli uporediti sa spektrogramom signala. U ovoj tački potrebno je da izaberete dužinu i korak prozora analize, prozorsku funkciju, red modela, kao i da li ćete koristiti visokopropusni predfilter (9).

Napomene: Ne zaboravite korišćenje prozorske funkcije koja će prigušiti vrijednosti odmjeraka na krajevima prozora analize.

Vrijednosti autokorelacione funkcije odredite usrednjavanjem po vremenu, bez normalizacije.

Iako postoje efikasni algoritmi za rješavanje sistema Wiener-Hopfovih jednačina, jednostavni pristup korišćenjem MATLAB-ovih matričnih funkcija je sasvim zadovoljavajući za ovaj zadatak.

Heurističko pravilo za određivanje reda modela je da postoji po jedna formantna frekvencija u svakom kHz frekvencijskog opsega signala, a da svakoj formantnoj frekvenciji odgovara po jedan par konjugovano-kompleksnih polova. Pošto pobudni signal modela odstupa od idealnog onda se za modeliranje spektra pobudnog signala obično koriste još dva pola.

Preklapanje prozora analize od oko 50% je uobičajeno.

Detalji predaje radova: Nakon završetka kompletnog projekta, najkasnije do navedenog roka, predaju se sledeći rezultati rada:

- program u MATLAB-u,
- odštampan rad maksimalne dužine 10 stranica koji sadrži sve važne detalje o projektovanju i implementaciji vašeg linearnog prediktivnog modela:
 - obrazloženja izbora pojedinih parametara algoritma,
 - talasne oblike i spektre/spektrograme ulaznih i izlaznih testnih signala,
 - analizu dobrih strana i nedostataka,
 - prijedloge poboljšanja,
 - uputstvo za upotrebu,
 - ovaj dokument treba da bude obrađen na računaru.

- nakon predaje radova biće organizovano testiranje modela na novim signalima.

Način rada i ocjenjivanje: Zadatak se radi u grupama po dva studenta. Konsultacije između grupa su dozvoljene, ali svaka grupa treba da preda originalno rješenje. U slučaju prepisivanja odgovarajući radovi će biti ocijenjeni sa nula poena.

Oprema: Za rad na ovom zadatku studentima je na raspolaganju računarska oprema u Laboratoriji za digitalnu obradu signala.

Predmetni nastavnik i asistent