

# Prepoznavanje govora

## Drugi domaći zadatak

### LPC parametrizacija

#### Opis zadatka

Implementirati alat koji snima niz reči od govornika, i za svaku od tako unetih reči radi LPC enkodiranje.

Alat treba da omogući podešavanje parametara enkodiranja.

Alat treba da proizvodi jednostavan izveštaj o sličnosti između LPC koeficijenata za različite reči.

# Ulaz

Alat mora na ulazu da obezbedi sledeće stavke:

- Korišćenje sirovog .WAV fajla kao ulazni zvučni signal.
- Korišćenje mikrofona kao ulazni zvučni signal.
- Na snimku (WAV i mikrofoni) može da ima više od jedne reči. Pretpostaviti da je nivo šuma na nivou jednog snimka konstantan, kao i da postoji fiksni period "tišine", tj. ne-govora na početku snimka.
- Izbor Hamming, Hanning ili ni jedne prozorske funkcije.
- Izbor širine LPC prozora i pomeraja između dva susedna prozora (prozori mogu da se preklapaju). Ponuditi korisniku izbor između milisekundi i smplova kao merne jedinice.
- Izbor dužine vektora koeficijenata (parametar **p** sa vežbi)

Obezbediti makar 10 wav datoteka, sa po tri reči za testiranje.

- Muški glas - 5 WAV zapisa
- Ženski glas - 5 WAV zapisa
- Svaki zapis treba da ima različite tri reči unutar sebe, s time da:
  - Tri muška i tri ženska zapisa treba da imaju isti skup od tri reči (9 reči muški, istih 9 reči ženski).
  - Dva muška i dva ženska zapisa treba da imaju sasvim jedinstvene reči (6 reči muški, drugačijih 6 reči ženski).
  - Ukupno ima  $9 + 6 + 6 = 21$  različitih reči u test skupu.

# Funkcionalnost

Endpointing treba da bude dovoljno dobar da pravilno odseca reči u učionici na Računarskom Fakultetu, u kojoj su vrata i prozori zatvoreni, i niko ne govori u pozadini.

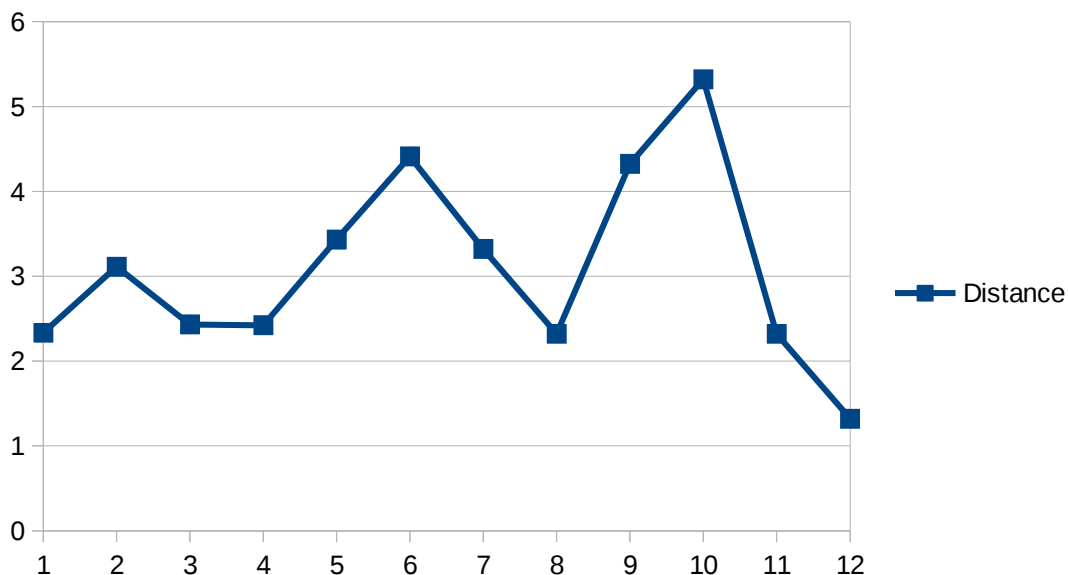
Nakon što je korisnik zadao ulazne parametre i snimio reči / uneo .WAV datoteke, alat treba da sračuna LPC koeficijente za zadati ulaz.

Za računanje LPC koeficijenata, koristiti autokorelacionu jednačinu, i računati rezultat matričnim množenjem, ne primenom Levinson-Durbin algoritma.

## Izlaz

Omogućiti kreiranje izveštaja gde korisnik bira:

- Dve reči koje će se porediti (reči mogu da budu snimljene mikrofonom, date kao WAV, ili oba).
- Prihvatljivo je da reči prethodno izanalizirane i skladištene u "bazi", kao i da se analiziraju u trenutku traženja izveštaja.
- Pretpostaviti da je parametar **p** za ove dve reči isti.
- Skup vektora za poređenje u jednoj reči i u drugoj reči. Skupovi moraju biti iste dužine, ali ne moraju biti na istoj poziciji.
- Kreira se linijski grafik koji po X osi ima onoliko tačaka koliko ima vektora u skupovima. Visina tačke na grafiku predstavlja euklidsku udaljenost između odgovarajućih vektora.
- U primeru ispod je korisnik izabrao opseg od 12 vektora.



## Napomene

Grafik je jedini deo sistema koji mora da bude grafički predstavljen. Dozvoljeno je da se grafik predstavi kao prozor u aplikaciji, ili kao eksportovana rasterska slika, pdf, itd. Podešavanje sistema, snimanje ulaza i podešavanje izveštaja mogu da budu konzolni ili GUI, po diskreciji studenta.

Dozvoljeni jezici za implementaciju su: C, C++, C#, Java. Dozvoljeno je predložiti jezik za implementaciju.

Implementacija i odbrana se rade isključivo individualno. Prikupljanje i generisanje ulaznih datoteka se takođe radi isključivo individualno.

## Bodovanje

- Endpointing: 7 poena
- LPC implementacija: 5 poena
- Izveštaj: 3 poena

## Rok za predaju zadatka

Zadatak se predaje putem email-a na [bmilojkovic@raf.rs](mailto:bmilojkovic@raf.rs). Rok za izradu domaćeg je ponoć, ponedeljak-utorak, 27-28 novembar. Odbrana domaćeg je obavezna. Termin vežbi nakon isteka roka je rezervisan za odbranu, uz mogućnost naknadne odbrane u posebnim slučajevima.