

## Prepoznavanje govora

### Drugi domaći zadatak - Reprezentacija signala

Zadatak se predaje putem email-a na [ajeremic@raf.rs](mailto:ajeremic@raf.rs). Rok za izradu domaćeg je ponoć, četvrtak - petak, 24 - 25 novembar.

#### Opis zadatka

Implementirati alat koji:

- Snima niz reči od govornika, i za svaku od tako unetih reči računa LPC i MFCC vektore.
- Omogućava podešavanje parametara za enkodiranje.
- Proizvodi izveštaj o sličnosti unetih reči pomoću DTW.

#### Ulaz

Korisnik može da unosi proizvoljan broj reči kroz .wav datoteke ili pomoću mikrofona. Za sve unete reči se računaju koeficijenti i pamte se u memoriji. Pri unosu svake nove reči, bilo kroz .wav datoteku ili sa mikrofona, alat treba da pita korisnika da imenuje reč, radi kasnijeg prikaza.

Alat mora na ulazu da obezbedi sledeće stavke:

- Korišćenje sirovog .wav fajla kao ulazni zvučni signal.
- Korišćenje mikrofona kao ulazni zvučni signal.
- Na snimku može da ima više od jedne reči, s time da će pre i posle svake reči postojati tišina minimalne dužine od 3 sekunde. Ovu tišinu treba iseći pre početka obrade, primenom algoritma opisanog na vežbi 2. Sistem treba da prijavi poruku o grešci ako smatra da na snimku nema izgovorenih reči.
- Izbor Hamming, Hanning ili ni jedne prozorske funkcije.
- Izbor širine LPC prozora i pomeraja između dva susedna prozora. Ponuditi korisniku izbor između milisekundi ili smplova kao merne jedinice.
- Podešavanje parametra  $p$  za LPC enkodiranje, tj. broj LPC koeficijenata.
- Izbor širine DFT prozora za MFCC. Ponuditi korisniku izbor između milisekundi ili smplova kao merne jedinice.
- Broj filtera u MFCC enkodiranju.
- Izbor da li se koriste  $\Delta$  i  $\Delta\Delta$  koeficijenti ili ne. (opisano ispod)

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

- Govorni signal - muški glas - 5 datoteka.
- Govorni signal - ženski glas - 5 datoteka.
- Svaki zapis treba da ima tri različite reči unutar sebe, s time da:
  - Tri muška i tri ženska zapisa treba da imaju isti skup reči (9 različitih reči muškim glasom, i istih 9 reči ženskim glasom). Za ove test primere treba iskoristiti dvoje muških i dvoje ženskih govornika.
  - Dva muška i dva ženska zapisa treba da imaju sasvim različite reči (6 različitih reči muškim glasom, i drugačijih 6 različitih reči ženskim glasom).

## Dinamički parametri

Kod MFCC koeficijenata se često koriste tzv. dinamički ili “delta” parametri, koji služe da enkodiraju informaciju o promeni vrednosti MFCC koeficijenata u vremenu. Korisnik bira parametar  $t$  (tipično u opsegu 2-4), koji informiše širinu dinamičkih parametara. Ovi parametri se računaju kao prvi i drugi izvod MFCC koeficijenata u vremenu. Za prvi izvod se računa razlika koeficijenata koji su  $t$  vremenskih indeksa u prošlosti i  $t$  vremenskih indeksa u budućnosti u odnosu na posmatrani:

$$\Delta c_m(q) = c_{m+\tau}(q) - c_{m-\tau}(q); q = 0, 1, \dots, Q - 1$$

Drugi izvod se računa po istom principu, samo što se računa razlika prvih izvoda, a ne originalnih koeficijenata:

$$\Delta \Delta c_m(q) = \Delta c_{m+\tau}(q) - \Delta c_{m-\tau}(q); q = 0, 1, \dots, Q - 1$$

Ovi parametri se dodaju na već postojeće parametre, i ima ih isto koliko je bilo originalnih. Primera radi, ako smo imali 13 filtera pri računanju MFCC koeficijenata i dobili vektor od 13 vrednosti, uvođenjem dinamičkih parametara ćemo imati 39 vrednosti u vektoru.

## Izlaz

Nakon što je uneo reči, korisnik može da računa LPC ili MFCC vektore za reči. Ovi vektori se čuvaju pod novim imenom, pošto će korisnik možda želeći za isti snimak da računa više različitih reprezentacija.

Nakon što je korisnik dobio vektorske reprezentacije reči, on može da izabere:

- Kolekciju reprezentacija reči koje će da služe kao template baza.
- Reprezentaciju jedne reči koja će da se testira nad tom bazom.

Kao rezultat pretrage, korisnik treba da dobije koliko je izabrana reč udaljena od svake iz template baze, po DTW algoritmu, kao i da vidi putanju unutar DTW trelisa za bilo koju od ovih pretraga.

## Napomene

Domaći zadatak ne mora da ima GUI komponente. Domaći zadatak mora da se odbrani da bi bio bodovan.

## Bodovanje:

- Unos reči: 2 poena
- LPC parametrizacija: 3 poena
- MFCC parametrizacija: 5 poena
- DTW: 5 poena