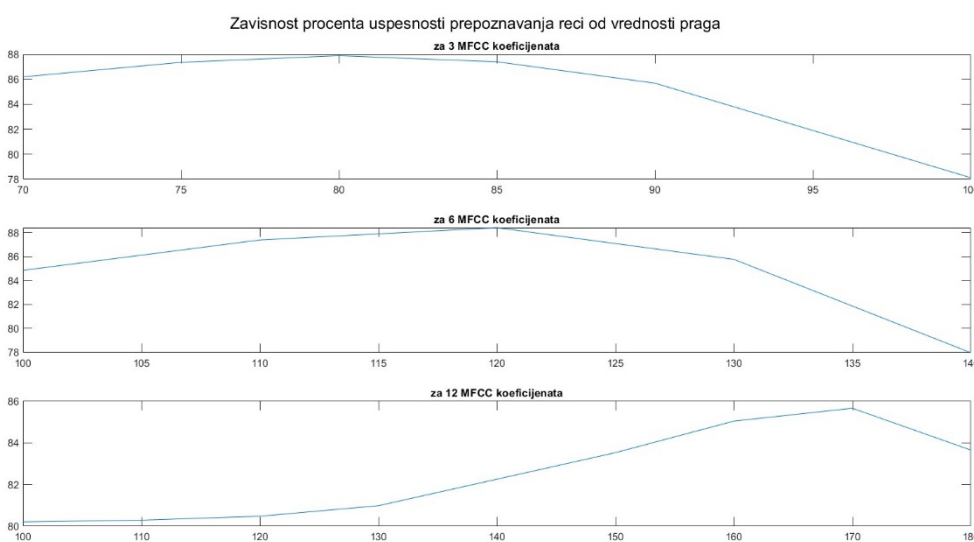
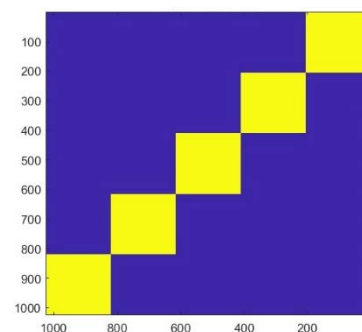


Diskusija rešenja

Reči koje treba da budu prepoznate kao tačne (žuti kvadratići) čine petinu ukupnog broja odluka, a kao pogrešne 4/5. Ako je prepoznavanje ugašeno i svim odlukama dodelimo vrednost 0 tj. da su automatski prepoznate kao pogrešne, bez ikakvog poređenja, bili bismo u pravu za sva ljubičasta polja tj. u 80% slučajeva. Zato program ima smisla samo ako je prepoznavanje iznad 80%.



Maksimalan procenat tačnih odluka je bio 88,4% za prag 120 i 6 MFCC koeficijenata, zatim, 87,9% za prag 80 i 3 MFCC koeficijenta i 85,7% za prag 170 i 12 MFCC koeficijenata.

Najbolji su rezultati za 6 pa 3 i na kraju 12

MFCC koeficijenata. Treći slučaj zahteva mnogo više podataka i vremena za računanje, a ne daje bolje rezultate. Pošto je mala razlika u procentu tačnih odluka za 3 i 6 koeficijenata, mislim da bi optimalan metod bio sa 3 jer barata sa manje podataka i potrebno mu je manje vremena za izvršavanje.

Matrica distanci za 3 MFCC koeficijenta:

Upoređivala sam rezultate prve osobe (žensko, ja) sa ostalima. Euklidska distanca je bila veća kada se poredi sa muškim osobama. Najveća razlika je bila kada je neki signal odsečen, zatim ako je trajanje signala previše različito, a malo manja razlika ako je dikcija drugačija ili ako je osoba suprotnog pola. Srednja vrednost poređenja prvog izgovora broja jedan prve sa ostalim osobama je 82.3061, što je u skladu sa odabranim pragom.

Srednja vrednost poređenja prvog izgovora broja jedan prve osobe sa prvim izgovorima broja 2 ostalih osoba je 90.4547, taman iznad praga, za broj 3 je 122.0132, za 4 je 111.6863 za 5 je 111.9094.

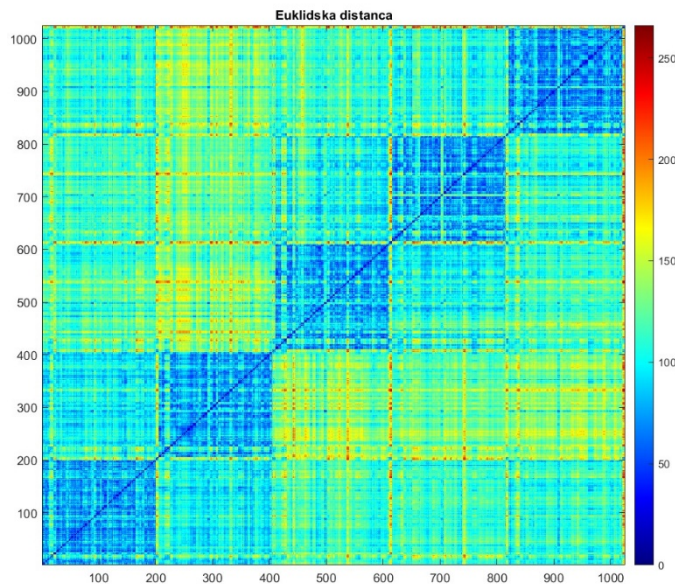
Jedan ima najviše sličnosti sa sobom, očekivano, zatim sa brojem 2, najmanje ima sa brojem 3.

Poređiću prve izgovore broja 2 prve osobe sa ostalim osobama. Srednja vrednost kada se poredi sa 2 je 86.8344, sa 3 je 160.7507, sa 4 je 147.0516, sa 5 je 150.1898. Najveća je razlika u izgovoru brojeva 2 i 3 što je i bilo očekivano.

Poređenje broja 3 sa samim sobom je 72.2120, sa 4 je 98.8429, sa 5 je 105.0160. Poređenje 4 sa samim sobom je 89.1764, a sa 5 je 110.8151. Pet sa samim sobom je 72.9761.

Uvek sam poredila distance prve osobe sa ostalima jer mi nije trebala precizna estimacija, a bilo je znatno brže.

Kada se broj poredi sam sa sobom najveća je razlika kod 4 pošto može da se izgovori na više načina “četiri”, “četri” i 2 jer dužina signala najviše varira.



Kada gledamo matricu distanci, primećujemo da su najveće razlike kada se poredi 2 sa nekim brojem, pogotovo sa 3. Najtamnija je dijagonala sa euklidskim distancama 0 jer se reč poredila sama sa sobom.,

Tražila sam gde je prepoznavanje najgore, npr za broj jedan od 15 do 20 se vidi plava štrafta, gotovo da se ni sa jednim izgovorom neka poklapanja, tu je npr amplituda signala manja od 0.01. Zatim od 85 do 90 osoba nije uklonila šum tako da ima smisla da preklapanje bude manje.

Ako je neka reč pogrešno prepoznata razlog je uglavnom klipovanje, preveliko odsecanje, nedovoljno odsecanje, jačina.

Jedan i dva imaju dosta grešaka u prepoznavanju baš zato što su im manje euklidske distance. Najgore je za tri i četiri baš zato što u nekim izgovorima “četiri” sadrži tri. I u matrici distanci ta polja su najtamnija. (2 slike su invertovane, gledajte xy ose)

