# OGK/D4 – Prepoznavanje govora

## Postavka vezbe

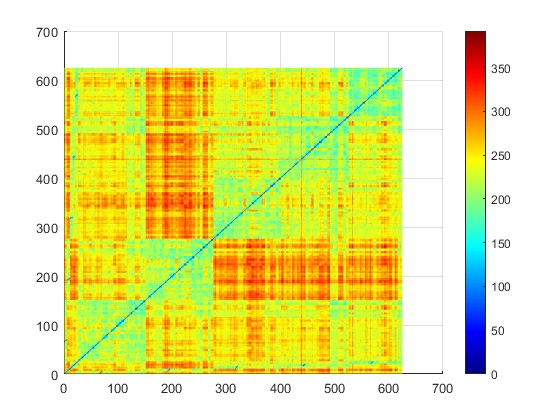
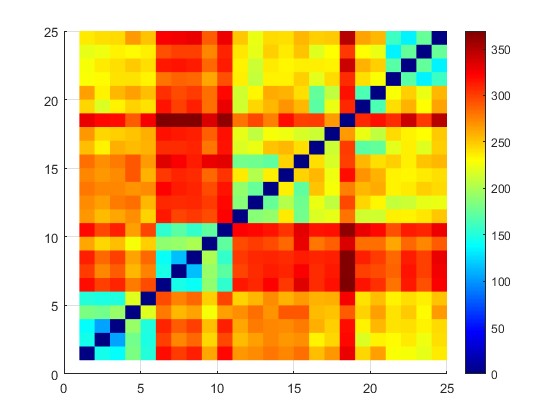
Napraviti korigovanu verziju MFCC funkcije koja rad isa promenljivim prozorima. Odrediti MFFC koeficijente signala snimljenih na drugoj laboratorijskoj vezbi. Porediti dobijene koeficijente sa koeficijentima ostatka baze i odrediti slicnosti medju signalima. Na osnovu slicnosti odrediti kriterijum kojim odredjujemo da li su signali dovoljno slicni da ih mozemo smatrati istom recju. Iscrtati grafik zavisnosti uspeha od kriterijuma procene.

Baza se nalazi na sledecem linku:

https://www.dropbox.com/scl/fo/erzil48k7ix6ovaoznw8b/h?dl=0&rlkey=09rfhl4exb0n8v94t0c2j204h

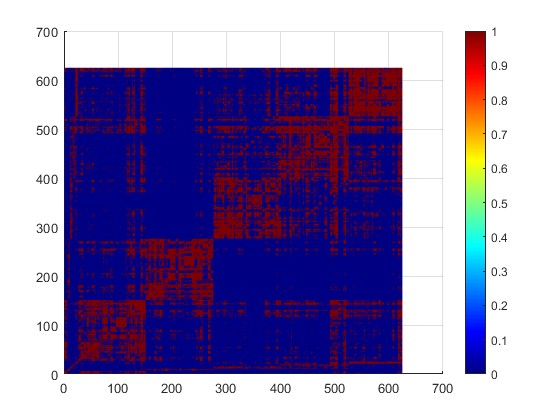
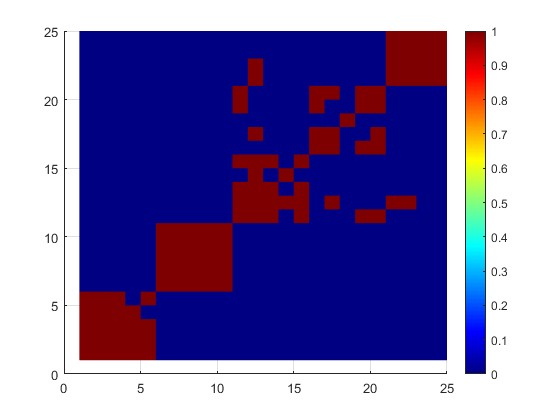
## Zapazanja

Kada koristimo bazu u kojoj se nalaze snimci samo jedne osobe ili vise osoba istog pola razlike su znatno manje nego u slucaju vecih baza. U nasoj bazi snimci iz druge laboratorijske vezbe su neidealni: Puni sumova, lose iseceni, snimani sa razlicitih udaljenosti od mikrofona… Samim time i razultati znatno variraju u razlicitim delovima grafika i unose veci broj gresaka pri pravljenju grupacija kojoj reci pripada dati signal(Slika1). Pri radu sa Malom bazom rezultati su mnogo bolji, ali I dalje neidealni usled razlicitih akcentovanja pri snimanju(Slika2). Takodje broj “dva” se najvise razlikuje od ostalih u izgovoru, a “jedan” je blisko drugo mesto i u slucaju male baze vidimo najgore snimljen snimak po crvenom plusu u cijem se centru on nalazi. Slicni su brojevi “tri” i “cetiri” jer sui m prelazne zone maltene stopljene na nekim delovima.

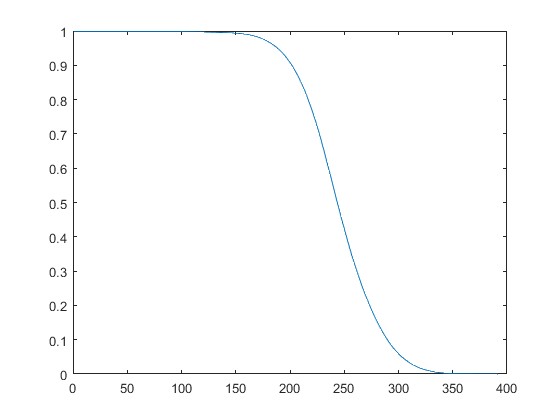
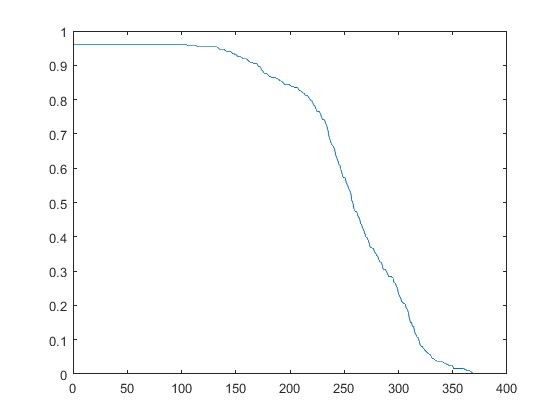
Slika 1. Slicnosti Velika baza Slika 2. Slicnosti Mala baza

Odstupanja u slicnosti snimaka su znatno manje pri koriscenju samo jedne osobe nego cele grupe ljudi(razlicitih kalibara da pritisnu dugme). Korisceni kriterijumi jesu priblizni, ali daju znatno drugacije rezultate. Birani su po principu da najbolje predstave “clustere” istih reci, da prikazu slicne delove, ali da greske i dalje ne budu prevelike(ogranicen borj “false posititve”)(Slike 3 i 4).

Slika 3. Priblizne reci Velike baze Slika 4. Priblizne reci Male baze

Na graficima koji prikazuju procenat koji je prosao kao isti drugoj reci vidimo vrlo slicne slike za veliku i malu bazu. Slika vece baze je samo malo vise glatka i krece od veceg procenta usled vecih tolerancija, tj. Veceg broja snimaka(Slike 5 i 6).



Slika 5. Zavisnost slicnosti od kriterijuma MB Slika 6. Zavisnost slicnosti od kriterijuma VB

## Zakljucak

Pri koriscenju vecih baza imamo vece greske u proceni ali pokrivamo i veci broj slucajeva i mozemo bolje da optimizujemo program za veci broj scenarija. Najvise odstupaju brojevi “jedan” i “dva” od ostalih, dok su “tri” i “cetiri” najslicniji jedan drugom. Sa porastom kriterijuma nam opada procenat “nejednakih” signala.