

Четврти домаћи задатак из предмета Основи Говорне Комуникације
Марко Мијаиловић
Извештај 28.05.2023

Поставка четвртог домаћег задатка:

Направити допуњену верзију MFCC функције која прихвата промењливе вредности прозора. Одредити MFCC коефицијенте сигнала снимљених од стране студената на трећој лабораторијској вежби, потом одредити сличност између добијених коефицијената за целу базу. На основу добијене сличности одређујемо критеријум по коме гледамо да ли су сигнали из базе довољно слични да буду иста реч. На крају исцртамо график зависности успеха од критеријума процене.

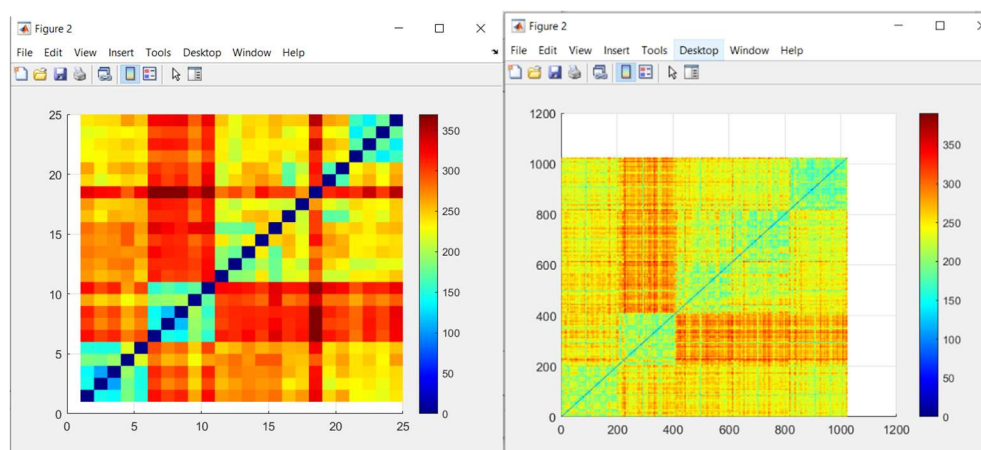
Резултати и запажања:

При коришћењу мање базе сигнала је приметан бољи резултат у смислу препознавања истих речи из више разлога. Пре свега у малој бази се налазе сигнали који су снимљени у доста уредном окружењу, без значајних сметњи и шума. Поред тога говорник је исти, па је и сам изговор датих речи приближнији што доста побољшава наше резултате.

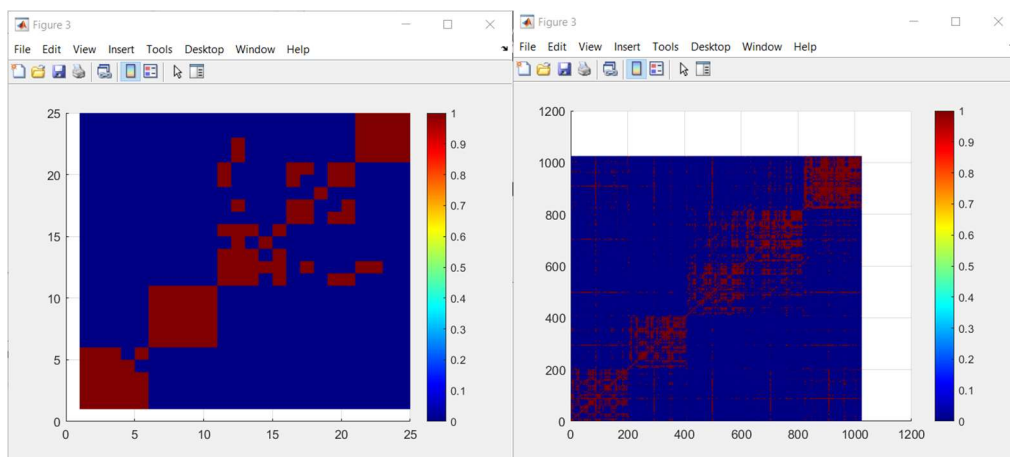
У ситуацији где користимо велику базу, у којој се налазе снимци који су направљени од стране студената примећујемо мању успешност препознавања сигнала из више разлога. Пре свега је то неуједначеност изговора код самих студената, број различитих говорника као и разлику у полу говорника. Поред ових главних утицаја, примећује се и мањи квалитет при обради сигнала, те су они у одређеним случајевима превише или премало скраћени. На крају је врло приметан утицај различитости нивоа сигнала, јер су говорници сви били различито удаљени од микрофона, због чега је било потребно урадити нормализацију сигнала.

Примећујемо велику разлику између речи „један“ и „два“ од осталих речи, док речи „три“ и „четири“ имају доста сличности због изговора речи „четири“ код разних говорника.

На графицику 1. видимо разлику у сличностима сигнала при коришћењу мале и велике базе, при чему се види да ниједна база неће дати идеалан резултат. Такође се примећује знатно мања успешност у великој бази и велики број лажних препознавања сличности. Предност велике базе је покривање већи број случајева, док је код мале базе поузданије препознавање сличних сигнала.

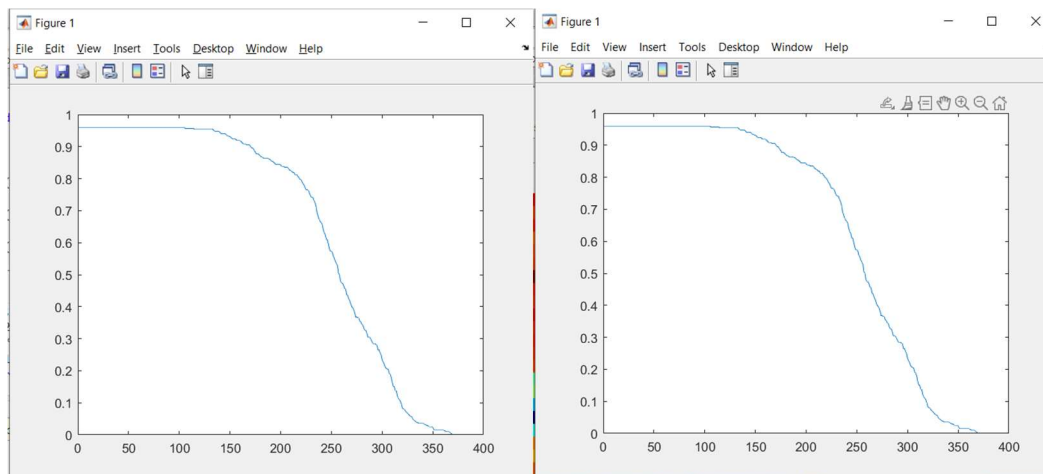


1 Матрица сличности за малу базу и велику базу



2 Матрице истих речи након критеријума за малу и велику базу

На графицима који приказују проценат успешности препознавања речи (график 5. и 6.) нема већих разлика између велике и мале базе, где велика база има нешто глађу криву и почиње од већих процената успешности због већег броја снимака.



3 Проценат успешности за малу и велику базу

Закључак домаћег:

Као што сам претходно изнео, велика база снимака омогућава боље покривање разних случајева, али такође доприноси смањењу препознавања сличних речи због различитости говорника и начина изговора речи. Код мале базе која је снимана од стране једног говорника се види доста успешније препознавање, али такође и мање покривених случајева. Са порастом критеријума за исту реч нам опада проценат различитих сигнала, што нам даје боље резултате.