

Ovaj projekat digitalnom obradom zvuka prepoznaje izgovaranje reči "hej kompjuteru".

Prvo se obrađuje glas mog brata koji izgovara reči za paljenje. Iz zvučnog fajla se uzimaju vremenski isečci od 20ms. Nad njima se radi brza Furijeova transformacija da bi se dobile frekvencije koje sačinjavaju zvučni signal. Nakon toga se koriste Melovi filtri. Oni nam omogućavaju da podelimo raspon frekvencija u 40 regiona i odredimo jačinu zvuka u svakoj od njih. Na kraju se nad dobijenim niz propušta kroz DCT filtar da bi se dobila karakteristika zvučnog signala u vremenskom isečku.

Na isti način se sa mikrofona uživo obrađuju vremenski isečci. Metodom dynamic time warping se nalazi niz isečaka sličan prvobitnom fajlu. U svakom trenutku se izračunava cena spajanja tj. koliko je sličan fajlu niz zvukova koji je algoritam našao, koji se završava u tom trenutku. Kada cena spajanja padne ispod određene granice, program zaključuje da se izgovorilo "hej kompjuteru" i pokazuje zeleno svetlo.

Koristi se programski jezik python sa bibliotekom numpy.

Detekcija reci

Novak Stijepić