

Prepoznavanje govora

Prvi domaći zadatak

Opis zadatka

Implementirati alat koji prikazuje frekventni spektar unapred snimljenih reči. Dozvoljeno je koristiti gotovu implementaciju DFT algoritma.

Ulaz

Ulazni zvučni signal treba da bude sirov .wav fajl. Na snimku treba da se nalazi jedna izgovorena reč, s time da je moguće da pre i posle reči postoji tišina proizvoljne dužine. Ovu tišinu treba iseći pre početka obrade i zatim signal obraditi algoritmom (varijanta E - poravnati rešenje). Treba obezbediti istovremeno učitavanje 2 zvučna signala.

Obezbediti 5 .wav datoteka za testiranje, svaka dužine po par sekundi:

- Govorni signal - muški glas
- Govorni signal - ženski glas
- Jedan ton na muzičkom instrumentu
- Nasumični šum
- Generisan signal koji sadrži bar 10 harmonika (napisati program za generisanje ovog signala)

Napomena

Sistem treba da prijavi poruku o grešci ako smatra da na snimku nema izgovorene reči (npr prilikom obrade .wav fajla na kome se nalazi šum).

Izlaz

Alat treba da omogući prikaz frekventnog spektra na 4 načina:

1. Prikaz čitavog signala iz odabranog .wav fajla sa označenom lokacijom početka i kraja reči.
 - Prikaz lokacije isečene reči treba na horizontalnoj osi da ima vreme, a na vertikalnoj osi očitane vrednosti iz .wav fajla. Potrebno je prikazati dve crvene vertikalne linije koje predstavljaju početak i kraj pronađene reči. U slučaju da nije nadjena reč u ulaznom signalu prikazati samo ceo signal, bez linija koje bi označavale početak i kraj reči.
2. Prikaz čitavih signala iz oba .wav fajla sa označenim lokacijama početka i kraja reči na isti način kao u prvoj tački. Svaki signal i granične linije predstaviti drugačijom bojom.
3. Prikaz frekventnog spektra odabranog zvučnog signala za jedan prozor, gde se spektar prikazuje kao histogram.

- Dijagram na horizontalnoj osi ima frekvenciju, a na vertikalnoj osi ima magnitudu. Frekvenciju prikazati u Hz, a magnitudu skaliranu na opseg 0-100.
- 4. Prikaz frekventnog spektra čitavog odabranog signala, gde se spektar prikazuje kao spektrogram.
 - Dijagram na horizontalnoj osi ima vreme, na vertikalnoj frekvenciju, a intenzitet tačke na dijagramu predstavlja magnitudu. I horizontalna i vertikalna osa moraju da budu numerisane, uz fiksni broj ravnomerno raspoređenih pomoćnih horizontalnih linija. Prikazati dve crvene vertikalne linije koje bi označavale početak i kraj pronadjene reči kao na prvom dijagramu. U slučaju da reč nije nadjena, izostaviti linije. Napraviti legendu boja koje predstavljaju magnitude.

GUI

Alat treba da sadrži jednostavan GUI koji ima potrebne funkcionalnosti. Sam dizajn nije preterano bitan, sve dok su podržane sve funkcionalnosti.

Parametri

Alat treba da obezbedi:

- Odabir 2 .wav fajla za obradu
- Izbor prozorske funkcije (Hamming, Hanning ili ni jedna prozorska funkcija)
- Izbor širine DFT prozora
- Odabir vrste prikaza frekventnog spektra i konkretnog zvučnog signala za prikaz (ili oba za slučaj 2)

Rok za izradu i bodovanje

Dozvoljeno je raditi u programskom jeziku po želji.

Domaći zadatak mora da se odbrani da bi bio bodovan. Odbrana podrazumeva odgovaranje na pitanja u vezi izrade domaćeg zadatka, ili izvršavanja neke jednostavne izmene nad zadatkom na licu mesta. Domaći se radi isključivo individualno.

Termini za odbranu prvog domaćeg zadatka će biti objavljeni nakon predaje. Svaki student će imati svoj termin u kojem vrši odbranu. Ovi termini će biti van časova vežbi, i moći će da se obavljaju isključivo uživo.

Bodovanje:

- Generisanje ulaza, odsecanje, konfiguracija DFT: 6 poena
- Poređenje signala: 3 poena
- Histogram: 3 poena
- Sonogram: 3 poena

Zadatak se predaje putem email-a na ajeremic@raf.rs. Rok za izradu domaćeg je ponoć, nedelja-ponedeljak, 6-7 novembar.