SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA OSIJEK

**FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA I INFORMACIJSKIH TEHNOLOGIJA OSIJEK**

Kristian Klišković, Damir Bašić i Josip Ilakovac

**Prepoznavanje tipa govornika na osnovu spektra govornog signala**

P R O J E K T N I Z A D A T A K

Osijek, 2022.

Sadržaj

[1. Uvod 1](#_Toc98328220)

[2. Pregled postojećih rješenja u području prepoznavanja tipa govornika na osnovu spektra govornog signala 2](#_Toc98328221)

# Uvod

# Pregled postojećih rješenja u području prepoznavanja tipa govornika na osnovu spektra govornog signala

Spektar govornog signala širok je i odavno poznat pojam te je samim time postojan veliki broj postojećih rješenja. U ovom poglavlju prikazana su 2 postojeća rješenja u području prepoznavanja tipa govornika na osnovu spektra govornog signala na način da su ukratko opisani sami algoritmi prepoznavanja te da su na osnovu toga navedene prednosti i nedostatci postojećih rješenja.

## *Male Female Voice Recognition – Harsh Mittal*

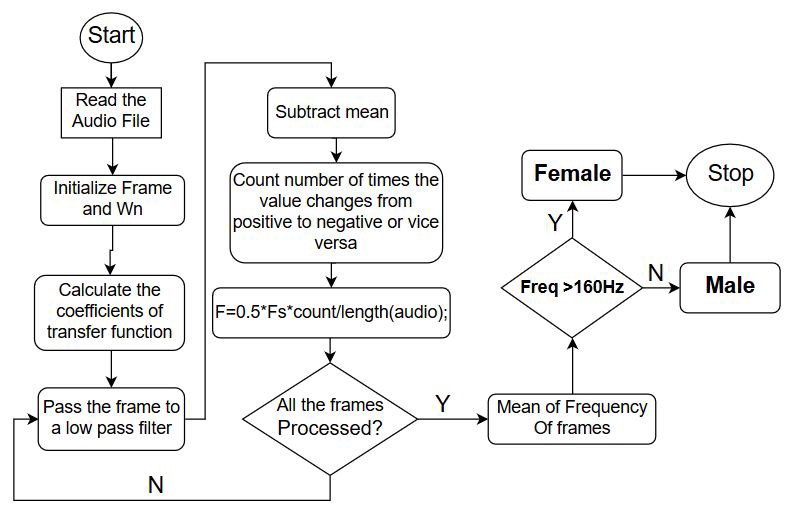
U samom uvodu programskog rješenja autor daje informacije o odabiru frekvencijskih raspona muških i ženskih osoba i to na način kako je prikazano tablicom 1.

|  |  |
| --- | --- |
| Frekvencijski raspon [Hz] | Spol |
| 85 Hz – 180 Hz | Muški |
| 155 Hz – 255 Hz | Ženski |

Tablica 1: Frekvencijski rasponi glasova

S obzirom na navedene frekvencijske raspone definira graničnu frekvenciju od 160 Hz. Drugim riječima, sve frekvencije koje bivaju dobivene kao rješenje algoritma za prepoznavanje tipa govornika veće od 160 Hz predstavljat će ženski glas.

U nastavku rada autor prikazuje algoritam shematskim prikazom kako je prikazano na slici 1.



Slika 1: Algoritam za prepoznavanje tipa govornika na osnovu spektra govornog signala

Autor navodi da na vlastitoj bazi govornih signala algoritam daje ispravan zaključak o tipu govornika u 6 od 10 slučajeva (60 %). Algoritmi i programska rješenja ovakvog tipa najbolje je ispitati sa „slučajnom“ bazom govornih signala. Drugim riječima, u tablici 2 prikazani su zaključci algoritma na nasumičnim uzorcima govornih signala.

|  |  |  |  |
| --- | --- | --- | --- |
| **R. br.** | **Audio zapis (muško/žensko)** | **Frekvencija** | **Odluka algoritma** |
| **1.** | Muška odrasla osoba | 142.9 | Muško |
| **2.** | Muška odrasla osoba | 167.1 | Žensko |
| **3.** | Muška odrasla osoba | 161.3 | Žensko |
| **4.** | Muška odrasla osoba | 180.3 | Žensko |
| **5.** | Muška odrasla osoba | 138.4 | Muško |
| **6.** | Ženska odrasla osoba | 236.2 | Žensko |
| **7.** | Ženska odrasla osoba | 197.2 | Žensko |
| **8.** | Ženska odrasla osoba | 189.4 | Žensko |
| **9.** | Ženska odrasla osoba | 180.5 | Žensko |
| **10.** | Ženska odrasla osoba | 165.3 | Žensko |
|  | | | **Ukupno 7/10 (70 %)** |

Tablica 2: Testiranje rada algoritma

Kao što je vidljivo u tablici 2, algoritam daje ispravne rezultate u 70 % slučajeva što čak predstavlja bolji rezultat nego što je sami autor naveo. Nedostatak algoritma je što podržava samo audio zapise u *.wav* formatu. Nadalje, u samom algoritmu uopće nisu uzeti u obzir dječji glasovi te samim time algoritam nema mogućnost raspoznavanja dječjeg glasa od glasa odrasle osobe.