

Sprawozdanie z rozwiązania problemu

Rozpoznawanie płci na podstawie nagrania głosu

Mateusz Urbaniak, I1 127345

1) Wstęp

Zadanie polegało na rozpoznawaniu płci na podstawie pliku WAV zawierającego nagranie głosu powiązanej osoby. W rozwiązaniu tego problemu skorzystaliśmy z bibliotek:

- scipy
- numpy
- glob (wyłącznie do testów – uproszczone wczytanie całego folderu z plikami WAV)

Rozwiązanie otrzymuje się po podaniu jako argumentu wejściowego programu nazwy pliku o rozszerzeniu WAV. Znak 'M' oznacza rozpoznany głos męski, znak 'K' rozpoznanie kobiecego głosu.

2) Opis algorytmu

W programie, korzystam z 3 funkcji (pozostałe używane wyłącznie do testowania skuteczności algorytmu): `voice_gender_recognition` (główna procedura programu), `HPS` i `filter_frequencies`. Funkcja `voice_gender_recognition` przyjmuje jako argument nazwę pliku WAV, a następnie otwiera zadany plik korzystając modułu biblioteki `scipy`. Wyciągamy średnią z obu kanałów dźwięku, a następnie następuje funkcja `voice_gender_recognition` korzysta z funkcji `hps`, która na wejściu przyjmuje sygnał i częstotliwość próbkowania. Funkcja ta jest implementacją algorytmu HPS (Harmonic Product Spectrum). Jako parametry przyjąłem liczbę downsamplingów równą 5, dla FFT liczbę próbek równą 524288, wartości dobrane testowo. Funkcja zwraca 1/5 próbek. Następnie, jeśli najczęściej występująca częstotliwość jest w zakresie ludzkiego głosu (między 50 a 255), wartości większe od 160 to głosy żeńskie, wartości mniejsze lub równe 160 to męskie. W przeciwnym wypadku, sumujemy wartości głosów z przedziału żeńskiego (od 160 do 255) i wartości głosów z przedziału męskiego (od 50 do 160). Jeśli suma wartości męskich jest większa od sumy wartości żeńskich, to oznaczamy głos jako męski. W przeciwnym wypadku, oznaczamy głos jako żeński.

3) Skuteczność

Skuteczność algorytmu obliczyłem na podstawie zadanego zbioru plików spod adresu:

[https://www.google.com/url?](https://www.google.com/url?q=http://www.cs.put.poznan.pl/wjaskowski/pub/teaching/kck/gender/train.zip&sa=D&ust=1516057941097000&usg=AFQjCNEU2Orl1WDDKtPhyYohTLmWpG7kJg)

[q=http://www.cs.put.poznan.pl/wjaskowski/pub/teaching/kck/gender/train.zip&sa=D&ust=1516057941097000&usg=AFQjCNEU2Orl1WDDKtPhyYohTLmWpG7kJg](http://www.cs.put.poznan.pl/wjaskowski/pub/teaching/kck/gender/train.zip&sa=D&ust=1516057941097000&usg=AFQjCNEU2Orl1WDDKtPhyYohTLmWpG7kJg)

Osiągnięta skuteczność algorytmu wynosi ~83%.