

Sprawozdanie z rozwiązania problemu

Rozpoznawanie płci na podstawie nagrania głosu

Mateusz Urbaniak, I1 127345

Kajetan Zimniak, I1 127229

1) Wstęp

Nasze 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 naszym programie, korzystamy z 3 funkcji (pozostałe używane wyłącznie do testowania skuteczności algorytmu): `voice_gender_recognition` (główna procedura programu), `average_frequency` 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 wykorzystujemy funkcję obliczającą odwrotną transformatę Fouriera, wyciągamy moduł z wyniku tej funkcji i logarytmujemy go, przypisujemy jego wartość jako sygnał. Następnie funkcja `voice_gender_recognition` korzysta z funkcji `filter_frequencies`, która na wejściu przyjmuje sygnał, ilość próbek sygnału i częstotliwość próbkowania. Funkcja ta przeprowadza filtrowanie częstotliwości tak, aby na wyjściu otrzymać tylko częstotliwości z zakresu ludzkiego głosu (od ~ 85 do ~ 255) oraz sygnał dla częstotliwości z tego zakresu. Następnie, przy pomocy funkcji `average_frequency` obliczamy średnią częstotliwość (poprzez odczyt wartości częstotliwości dla sygnału o maksymalnej wartości). Głosy poniżej zadanej wartości progowej interpretowane są przez algorytm jako głosy męskie, głosy powyżej lub równe wartości progowej jako głosy żeńskie. Wartość progową przyjęliśmy w algorytmie jako **170**.

3) Skuteczność

Skuteczność algorytmu obliczaliśmy 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 $\sim 75\%$.