

## Sprawozdanie z implementacji algorytmu

Celem programu jest analiza pliku audio w formacie WAV w celu przewidywania płci osoby na podstawie charakterystyki częstotliwości głosu. Program opiera się na analizie transformaty Fouriera i spektrum produktu harmonicznego (HPS) sygnału audio, identyfikując dominujące częstotliwości, które są charakterystyczne dla głosów mężczyzn i kobiet.

Pierwszym krokiem w algorytmie jest wczytanie pliku audio, który jest przekazywany jako argument wiersza poleceń. Program wykorzystuje funkcję `scipy.io.wavfile.read` do załadowania pliku WAV. Jeśli wczytany sygnał jest w formacie stereo (tzn. ma dwa kanały), zostaje przekształcony na sygnał mono. Konwersja polega na uśrednieniu wartości obu kanałów. Aby poprawić analizę widma częstotliwościowego stosowana jest funkcja wygładzająca - okno Kaisera. Wartość beta ustalana jest na podstawie długości nagrania w sekundach.

Kolejnym krokiem jest obliczenie transformaty Fouriera sygnału, która pozwala na reprezentację sygnału w dziedzinie częstotliwości, dodatkowo wyznaczony zostaje jej moduł.

Główną metodą analizy częstotliwości w tym programie jest obliczenie spektrum produktu harmonicznego (HPS). Jest to technika wykorzystywana w celu wyodrębnienia dominujących częstotliwości w sygnale. HPS oblicza się przez decymację (zmniejszanie liczby próbek) sygnału w kilku etapach i obliczanie iloczynów powstałych sygnałów.

Po licznych próbach i analizie wyników, program wypadł dla 4 iteracji.

Po obliczeniu HPS, sprawdzam, czy dominująca częstotliwość z przedziału (65 Hz - 300 Hz) - wartości ustalone na podstawie otrzymywanych wykresów - mieści się w zakresie częstotliwości podstawowej tonu krtaniowego w mowie dla mężczyzn (65 Hz - 171.5 Hz). W przypadku spełniania warunku, program przewiduje, że osoba w nagraniu to mężczyzna, w przeciwnym przypadku, jeśli częstotliwość jest poza tym zakresem, zakładam, że plik dźwiękowy zawiera głos kobiety.

Trafność klasyfikacji dla zbioru testowego: **94.51%**.

Błędnie rozpoznane głosy w plikach:

046\_K.wav  
051\_K.wav  
055\_K.wav  
067\_K.wav  
068\_K.wav

Macierz pomyłek

Rzeczywista klasa	Przewidziano: M	Przewidziano: F
M	100% (45/45)	0% (0/45)
F	10.87% (5/46)	89.13% (41/46)