

# Sprawozdanie z projektu – rozpoznawanie głosu

Wiktoria Keller 145189

## 1 Wstęp

Celem projektu było napisanie programu rozpoznającego płeć osoby na podstawie nagrania głosu.

## 2 Opis metody rozpoznawania głosu

Płeć osoby mówiącej w nagraniu jest rozpoznawana poprzez znalezienie jej podstawowego tonu głosu, do tego celu została wykorzystana między innymi dyskretna transformata Fouriera, a także implementacja algorytmu HPS (Harmonic Product Spectrum). Pełny algorytm przedstawia się następująco:

1. obliczenie średniej z dwóch kanałów (tylko w przypadku gdy wczytany plik audio zawiera dwa kanały),
2. wycięcie części nagrania z jego początku i końca w celu pozbycia się zanieczyszczeń,
3. pomnożenie sygnału przez okno Kaisera ( $\beta = 4$ ) w celu jego wygładzenia,
4. użycie dyskretnego transformaty Fouriera na tak przygotowanym sygnale, a następnie obliczenie wartości bezwzględnej,
5. wykonanie decymacji (downsampling), w tym przypadku została zachowana 2, 3 oraz 4 próbka,
6. pomnożenie dyskretnego widma uzyskanego przy pomocy DFFT przez sygnały uzyskane po 2, 3 oraz 4 decymacji,
7. znalezienie częstotliwości o najwyższej amplitudzie (przy czym początkowe częstotliwości są ignorowane),
8. jeżeli tak uzyskana częstotliwość jest większa lub równa 165Hz to rozpoznana została kobieta, w przeciwnym wypadku rozpoznany został mężczyzna.

## 3 Wyniki

Skuteczność algorytmu na zbiorze treningowym wyniosła 93.4% – poprawnie zostało rozpoznanych 85 osób na 91.

Rzeczywista klasa	Przewidziana klasa	
	Kobieta	Mężczyzna
Kobieta	43	3
Mężczyzna	3	42

Tabela 1: Macierz pomyłek