

| Sprawozdanie z Projektu                       |   | Rok akademicki:<br>2018/2019 |
|---|---|------------------------------|
| Przedmiot:<br>Komunikacja Człowiek - Komputer | Prowadzący:<br>Paweł Liskowski                  |                              |
| Tytuł projektu:                               | Rozpoznawanie płci na podstawie nagrania głosu. |                              |
| Autorzy                                       | Kinga Karasiewicz, nr indeksu 132247            |                              |
|   | Robert Dudek, nr indeksu 132211                 |                              |
| Termin:<br>11.01.2019 r.                      | Data oddania sprawozdania:<br>10.01.2019 r.     | Ocena:                       |

## 1 Wstęp

Celem naszego projektu było stworzenie algorytmu do rozpoznawania płci na podstawie podanego jako argument nagrania głosu. W naszym algorytmie zastosowaliśmy metodę HPS (Harmonic Product Spectrum).

## 2 Metoda HPS

W naszym algorytmie płeć rozpoznajemy za pomocą wysokości głosu w danym nagraniu. Jednakże nagranie jest w pewnym sensie zanieczyszczone przez nasze struny głosowe, poprzez nakładające się na siebie dźwięki harmoniczne i zniekształcenia. Aby móc o tym zdecydować należy najpierw znaleźć ton podstawowy z otrzymanego przez nas nagrania.

Najpierw musimy przygotować sygnał do uzyskania tonu głosu. Wpierw dzielimy sygnał przez  $2^{16}$  w celu znormalizowania sygnału. Następnie zdecydowaliśmy się wyciąć początkową i końcową  $\frac{1}{4}$  sygnału, by uzyskać mniej zanieczyszczony sygnał. W kolejnym kroku zastosowaliśmy okienkowanie kaisera ( $\beta = 5$ ), które pozwala na wygładzenie sygnału, po czym skorzystaliśmy z dyskretnej transformaty Fouriera i obliczenia wartości bezwzględnej.

Przy tak przygotowanym sygnale możemy zacząć decymację, czyli zachowanie  $n$ -tej próbki, a odrzuceniu pozostałych. W naszym algorytmie skorzystaliśmy z drugiej, trzeciej i czwartej decymacji. Uzyskane przez nas wyniki dla każdej decymacji, a także sygnał początkowy przemnożyliśmy przez siebie, co sprawia, że wzmocniliśmy ton podstawowy, a wytłumiliśmy szumy. Teraz, by dowiedzieć się jaki ton głosu miało nagranie, ignorujemy pierwsze częstotliwości i szukamy wartości o najwyższej amplitudzie. Następnie dzielimy przez rozdzielczość częstotliwości, by uzyskać częstotliwość ostateczną.

Jeżeli uzyskana przez nas częstotliwość jest wyższa od 165 Hz, to rozpoznajemy tę osobę jako kobietę, jeżeli niższą to jako mężczyznę.

## 3 Wyniki

Skuteczność naszego algorytmu na zbiorze treningowy przy decymacji do czterech wynosiła: 83/91, czyli skuteczność na poziomie 91.2%

|           | Liczba wszystkich | Prawidłowo rozpoznanych |       |
|-----------|-------------------|-------------------------|-------|
| Kobiety   | 46                | 42                      | 91.3% |
| Mężczyźni | 45                | 41                      | 91.1% |

Tabela 1: Wyniki skuteczności algorytmu

## 4 Literatura

- (1) Notatki z zajęć
- (2) [http://www.cs.put.poznan.pl/mkomosinski/lectures/kck/KCK\\_mowa\\_plec.pdf](http://www.cs.put.poznan.pl/mkomosinski/lectures/kck/KCK_mowa_plec.pdf)
- (3) Film pana doktora Komosińskiego <https://youtu.be/4R5uvJW9fXs>