

Rozpoznawanie płci na podstawie nagrania głosu

1. Opis algorytmu:

W pierwszym kroku wczytany z pliku sygnał jest dzielony na mniejsze, nakładające się na siebie fragmenty (długość części wspólnej kolejnych fragmentów przyjęto równą połowie długości jednego fragmentu). Następnie każdy fragment mnożony jest przez funkcję okna Hanninga oraz obliczana jest dla niego wartość bezwzględna z dyskretnej transformaty Fouriera. Na otrzymanym spektrum stosowany jest algorytm Harmonic Product Spectrum, czyli każdy jego element jest mnożony przez odpowiedni element tego samego spektrum skalowanego przez kolejne liczby naturalne zgodnie z równaniem:

$$Y(\omega) = \prod_{r=1}^R |X(\omega r)|,$$

gdzie X – badane spektrum. Przyjęto parametr $R = 4$. Na podstawie indeksu maksymalnej wartości wektora Y oraz częstotliwości próbkowania pliku wejściowego obliczana jest częstotliwość podstawowa tonu krtaniowego. Przy jej poszukiwaniu nie są brane pod uwagę częstotliwości mniejsze niż 50 Hz z powodu dużych szumów w tym zakresie. Na koniec obliczana jest mediana częstotliwości podstawowych znalezionych dla każdego fragmentu i jeśli jest ona mniejsza niż 170 Hz podejmowana jest decyzja, że osoba na nagraniu to mężczyzna, w przeciwnym wypadku uznana zostaje ona za kobietę.

2. Wyniki działania dla zbioru treningowego:

Rzeczywisty\Wskazany	K	M
K	41	5
M	1	44

3. Bibliografia:

Materiał opracowany na podstawie:

[http://musicweb.ucsd.edu/~trsmyth/analysis/
Harmonic Product Spectrum.html?fbclid=IwAR136GrdjLHgRXBb-7brVysS-
CzLNc8jX0z9OCBHzhEgoD7by8PaFdImD4k](http://musicweb.ucsd.edu/~trsmyth/analysis/Harmonic_Product_Spectrum.html?fbclid=IwAR136GrdjLHgRXBb-7brVysS-CzLNc8jX0z9OCBHzhEgoD7by8PaFdImD4k)