



Odczytywanie płci na podstawie próbek głosu - projekt

Marek Rydlewski 117214
Marcin Jabłoński 117270
Politechnika Poznańska, Wydział Informatyki
Poznań

17 grudnia 2015

I. Wprowadzenie

Naszym zadaniem było stworzenie aplikacji która dla zadanego pliku .wav decyduje czy nagrany głos należy do kobiety czy mężczyzny. Komunikacja z użytkownikiem odbywa się poprzez konsolę. Aby to osiągnąć korzystaliśmy z dobrodziejstw języka Python w wersji 3.4 a także biblioteki scipy.

II. Opis metody && zastosowane funkcje

Analizę dźwięku podzieliśmy na dwa główne etapy:

- Znalezienie głównej częstotliwości
- Na podstawie empirycznie wyznaczonych progów, zdecydowanie czy głos jest głosem kobiecym czy męskim

Jeśli chodzi o pierwszy etap, stworzyliśmy funkcję `get_lead_frequency`:

```
def get_lead_frequency(signal, samples_count, audio_duration, no=0, mono=0):
    signal = signal if mono else [s[no] for s in signal]
    signal = signal * kaiser(samples_count, 100)
    spectrum = np.log(abs(np.fft.rfft(signal)))
    spectrum_enriched = copy.copy(spectrum)
    for beta in range(2, 6):
        decimated_spectrum = decimate(spectrum, beta)
        spectrum_enriched[:len(decimated_spectrum)] += decimated_spectrum

    peak_start = 50 * audio_duration
    peak = np.argmax(spectrum_enriched[peak_start:])
    lead_frequency = (peak_start + peak) / audio_duration
    return lead_frequency
```

Opis metody:

- wybieramy sygnał z jednego kanału (w razie gdyby nagranie posiadało dwie)
- filtrujemy sygnał
- obliczamy spektrum sygnału
- dodajemy do sygnału jego zdziesiątkowane wersje
- wyznaczamy dominującą częstotliwość

W drugim etapie bazujemy korzystamy z progów:

```
if lead_frequency < 165:
    result = 'M'
else:
    result = 'K'
```

- przez proste porównanie głównej częstotliwości decydujemy do kogo należy próbek głosu

Progi zostały przez nas wyznaczone empirycznie na podstawie dostarczonych danych.

III. **Sprawność i podsumowanie**

W zastosowany przez nas algorytmie uzyskaliśmy sprawność około 90% (z 90 poprawnie rozpoznanych zostało 80 próbek - a więc uzyskaliśmy całkiem zadowalający wynik