

Signals and systems

PROJECT

Systém pro vyhledávání v audio pomocí akustického vzoru

Ladislav Ondris (xondri07)

19. 11. 2019

1 Recordings

File name	Samples read	Length (seconds)
sa1.wav	71658	4.478625
sa2.wav	56298	3.518625
si1446.wav	99818	6.238625
si2076.wav	44778	2.798625
si816.wav	65258	4.078625
sx186.wav	52458	3.278625
sx276.wav	60138	3.758625
sx366.wav	78058	4.878625
sx6.wav	51178	3.198625
sx96.wav	60138	3.758625

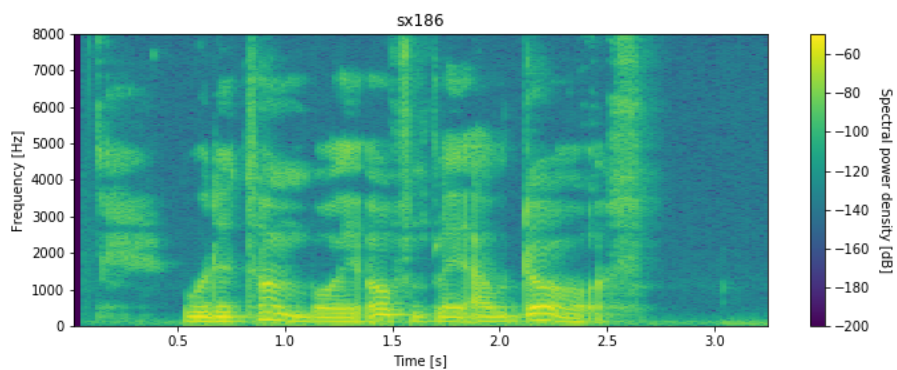
The recordings can be used for:

c) for (a), (b) and for freely available “Czenglish TIMIT” database.

2 Queries

File name	Samples read	Length (seconds)	Query
q1.wav	12196	0.762250	essentially
q2.wav	12604	0.787750	exercise

3 Spectrogram



Obrázek 1: Would a tomboy often play outdoors?

4 Features

I used linear bank of filters to calculate features of a given sentence or query. For that I used the matrix multiplication approach $\mathbf{F} = \mathbf{A}\mathbf{P}$.

The purpose of matrix \mathbf{A} is to sum every B rows (in our case 16 rows). To make this work, we create the matrix \mathbf{A} by filling it with zeros and ones in a specific pattern so that it produces the sum.

The first row of the matrix \mathbf{A} will contain 16 ones and the rest of it will be zeros. The second row will contain 16 zeros, then 16 ones, and the rest will be zeros. And so on. The shape of the matrix \mathbf{A} is $(B, f.size)$ where f is an array of sample frequencies.

5 Correlation score calculation

6 Main output

