Voiced and Unvoiced Classification using ZFF

By: Shreya Gupta, Naila Fatima, Gourishetty Shirisha, Pavan Kumar

Zero Frequency Filter

An ideal zero-frequency digital resonator is an infinite impulse response filter with a pair of poles located on the unit circle.

A cascade of two such resonators is used for making the zero frequency filter.

Advantage

Zero-frequency resonator is used because the characteristics of the time-varying vocal-tract system will not affect the characteristics of the discontinuities in the output of the resonator.

How to get ZFF signal?

- Calculate differenced signal to remove any slowly varying components.

$$x[n] = s[n] - s[n-1]$$

- Pass the differenced signal through a cascade of two ideal zero-frequency resonators.

$$y_0[n] = -\sum_{k=1}^{4} a_k y_0[n-k] + x[n]$$

Continue...

- The average pitch period is computed using the autocorrelation function of 30 ms segments of x[n].
- The trend $y_0[n]$ is removed by subtracting the local mean computed over the average pitch period, at each sample. The resulting signal

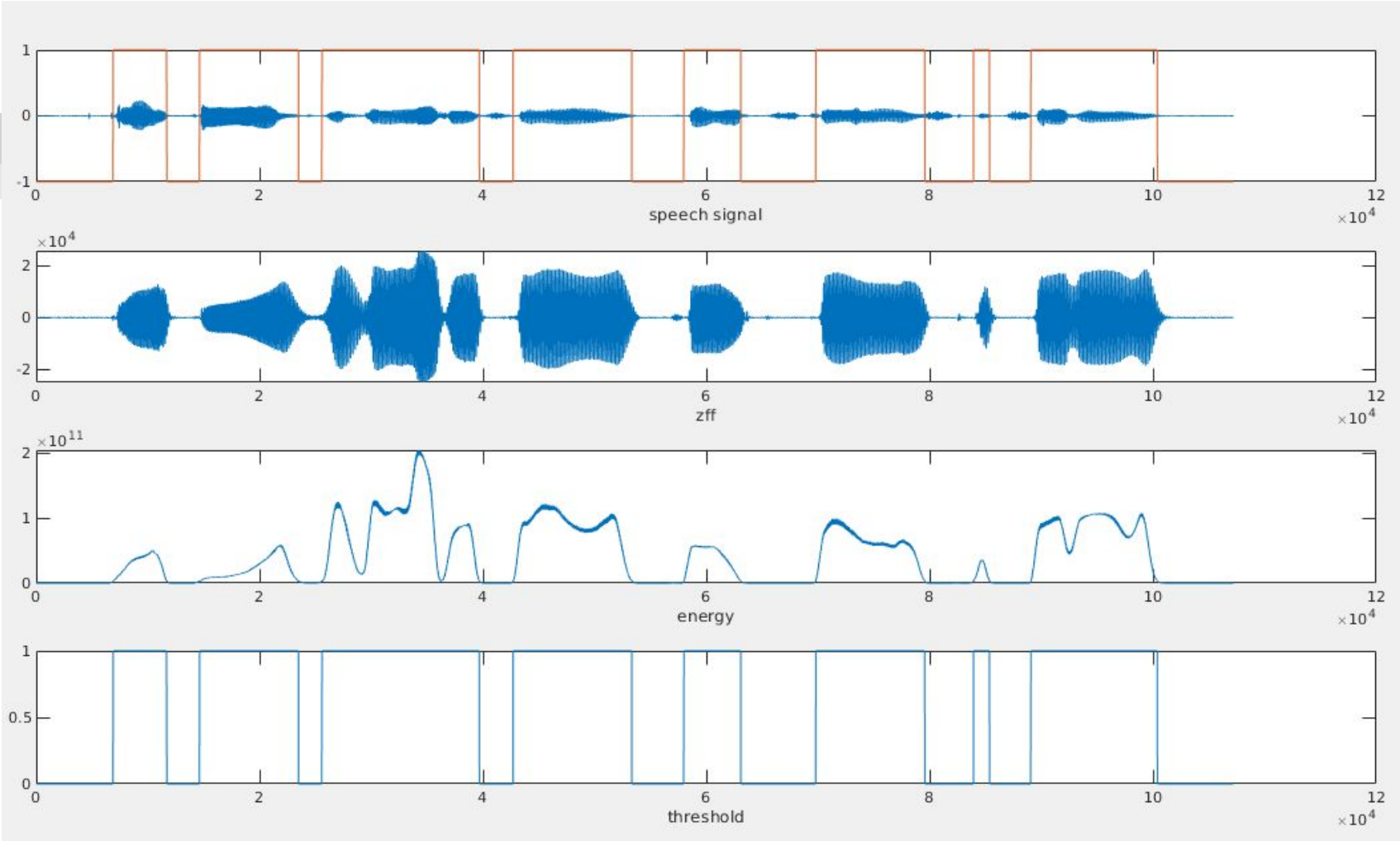
$$y[n] = y_0[n] - \frac{1}{2N+1} \sum_{m=-N}^{N} y_0[n+m]$$

is the ZFF signal.

- 2N+1 corresponds to the number of samples in the window used for trend removal.

Voiced and unvoiced detection using ZFF

- We know that the ZFF filtered signal gives us information about the excitation source of the signal.
- In order to distinguish between the voiced (sounds with glottal vibrations) and unvoiced (sounds without glottal vibrations), we find the energy of the filtered signal for 20 ms.
- Using a threshold, we can distinguish between the two as voiced regions have greater energy whereas unvoiced regions have lesser energy.


speech signal
zff
energy
threshold

Thank You