Ahmet Erdem Çağatay – 49826 10.03.2019

ELEC404 – Assignment #4

Engin Erzin

1. Autocorrelation functions should have some properties:

* Θ[k] is even.
* Θ[k] has maximum at k=0.
* If x[n]=x[n+P] then Θ[k]= Θ[k+P].

Thus, for the options:

1. is a *valid autocorrelation function*. It is even and has maximum point at = 0.
2. is *not a valid autocorrelation function*. It is even but it doesn’t have a maximum at =0.
3. is a *valid autocorrelation function*. It is the sinc function and has maximum at = 0. Also, it is even.
4. is a valid autocorrelation function. Since cos is an even function and it has maximum at = 0.

2. I chose Hamming window with L = 1001 and R = 20.

I wanted to choose L as high as possible because this decreases window bandwidth and thus, it avoids the edgy segments on the magnitude and energy measures. On the other hand, L should be smaller than a sound duration; otherwise, it won’t adequately reflect the changes in sound. To find an optimal L, I did fine-tune and found L=401.

The shift amount, R is proportional to the precision of the windowed signal, but it comes with more complex calculations. I reached R=20 by fine-tune.

I chose Hamming window over rectangular window because it functions better with its sharper stopband frequencies.

Here are the short-time parameters;

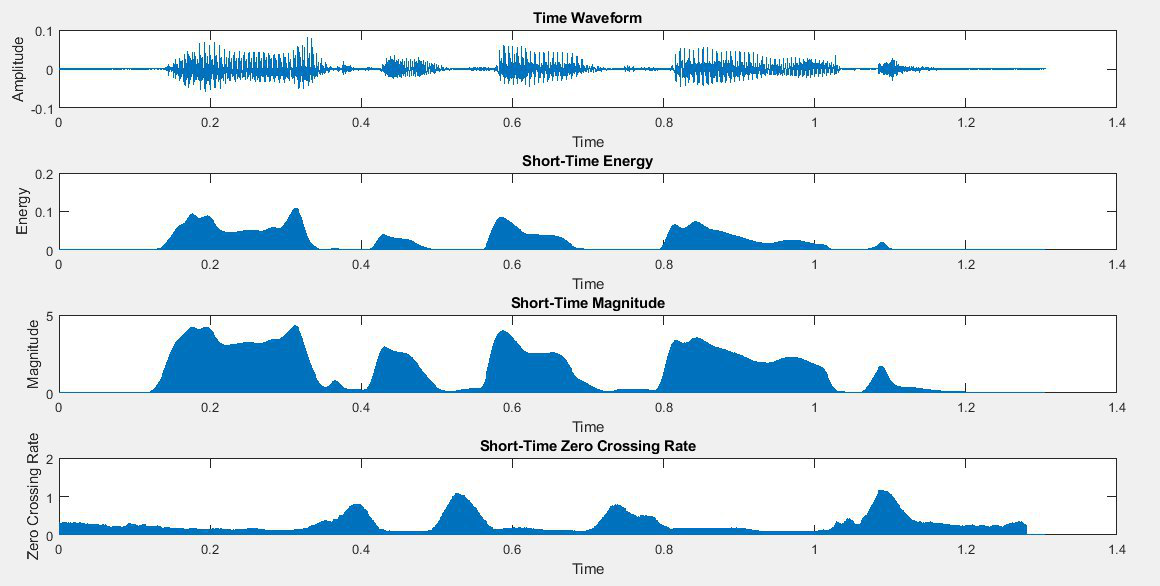


Figure 1: short-time parameters

As it is seen, energy and magnitude parameters are high in the voiced regions as expected. On the other hand, zero crossing rate is higher in the unvoiced regions.

3.

a) To be able to estimate pitch period precisely, the window length L must be longer than twice of the longest period in the waveform. Also, it increases the calculation complexity as it increases. Therefore, I picked L=1500 as an optimal number.

Similarly, window shift R should be as high as possible to bring higher precision to measurements. Therefore, I again picked R to be 1000.

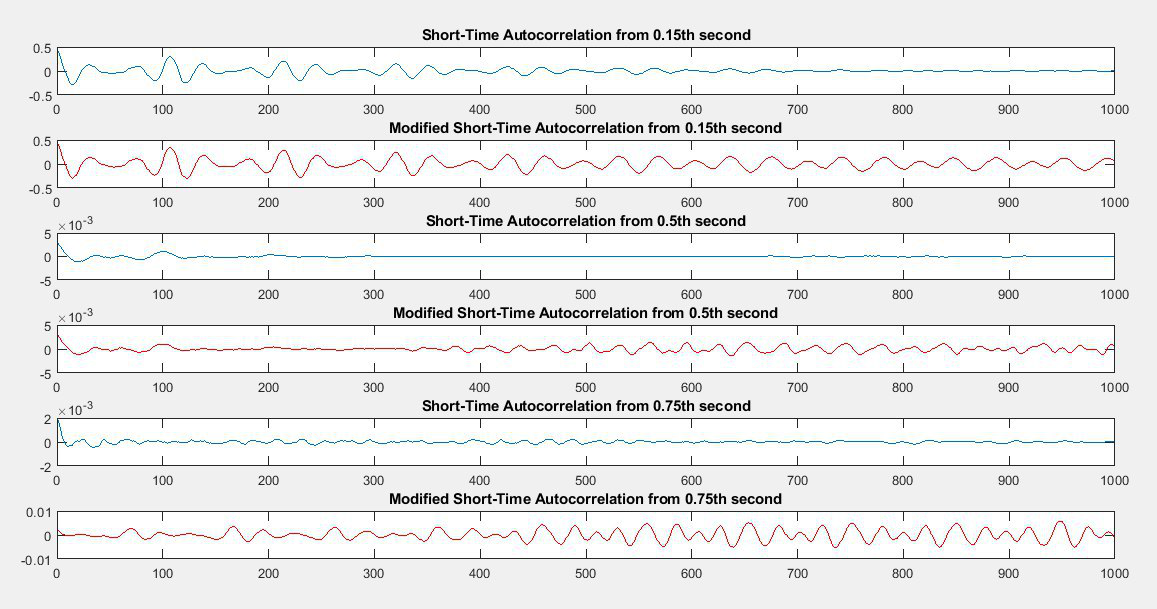


Figure 2: short-time autocorrelations from various regions

As you can see from the figure above, it makes more sense to pick modified short-time autocorrelation to estimate pitch period, and so on. Short-time autocorrelation extinguishes around the middle.

b) To estimate a suitable threshold for selecting the peak autocorrelation in the voiced regions, I used modified autocorrelation’s formula. I tried various threshold values and the best one was 0.3 because it gave the peak values at the voiced regions.

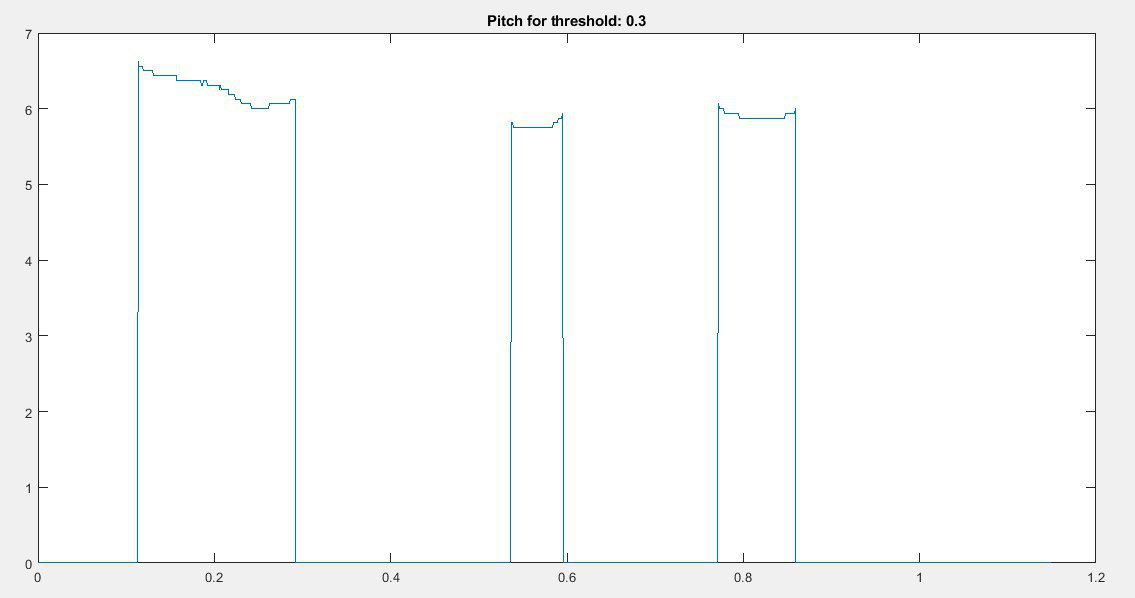


Figure 3: Pitch with threshold 0.3