SSP Assignment - 2

- Srihari Bandarupalli
- 2021112006

Question 1

a) Autocorrelation

- It measures the similarity of a signal with time delayed version of itself. It is used to identify the periodicity in a signal. We can get pitch of a signal by analyzing the peaks of autocorrelation with different time lags.

b) ZCR

- Zero crossing rate is the rate at which signal crosses the zero axis. We use this measure to identify voiced and unvoiced parts of speech as unvoiced speech being more random and noise like will have high ZCR.

c) Mel Spectrogram

- Mel Spectrogram is a the representation of signal in frequency domain where the frequency axis is linear upto 1KHz and then follows a logarithmic scale. This approximately represents how human auditory system perceives sound

d) LP spectrum

- Linear prediction spectrum gives the envelope or shape of the given signal. It estimates the coefficients of the linear filter that can be used to estimate the signal from its past samples. These coefficients give LP spectrum that captures the spectral envelope of the signal that corresponds to the resonances in the signal.

Question 2

Voiced / Unvoiced Speech

- Voiced speech refers to the speech sounds produced by the vocal fold vibrations. Hence, they are more energetic, periodic and harmonic. Generally, vowels and consonants come under this category.
- Unvoiced speech refers to the speech sounds produced without the vocal fold vibrations. Hence, they are less energetic , non-periodic and more noise like. Generally , some consonants come under this category.
- Methods for Identifying Voiced and Unvoiced Speech
    - Energy Based Method:- Voiced segments tend to have high energy due to vocal fold vibrations and unvoiced segments generally have low energy. This method might not work in presence of noise
    - Auto Correlation:- When we multiply a delayed version of signal with itself a periodic signal will give higher peaks. Thus Voiced signal will have high auto correlation and unvoiced signal will have low autocorrelation
    - Normalized Linear Prediction error:- In LP analysis we try to estimate current signal using previous samples of signal. We used normalized error as Voiced signal tends to have high energy. Voiced signal being periodic will have low error and Unvoiced signal being random will have high error.