Ryan Chau, Philip Kwan

EE114

Computer Assignment 2

Abstract:

We investigate speech samples to observe the properties of pre-emphasis filters and the speech signal parameters formant frequencies and pitch period.

We are given the pre-emphasis filter which varies by a parameter a ∈ [0,1], and we can look at the transfer function to determine poles and zeros, which dictate which type of filter it is (low-pass or high-pass). We observe the effects of this filter by applying it to a track of a female speaking a sentence, listening to the results as well as seeing them in the spectra.

We look at formants, or resonant frequencies in the vocal tract, by observing the spectrum for the male_a sound. The narrowly spaced lobes are harmonics from the source signal, but the shaping of the spectrum is the vocal tract. Hence, we can approximate the envelope and obtain the formants. We then determine the pitch periods and fundamental frequencies of the male and female voiced speech by looking at time-domain signals.

Pre-emphasis Filters:

1. $y(n) = x(n) - ax(n-1)$

$$Y(z) = X(z) - az^{-1}X(z)$$

$$\frac{Y(z)}{X(z)} = 1 - az^{-1}$$

Pole at $z = 0$, zero at $z = a$

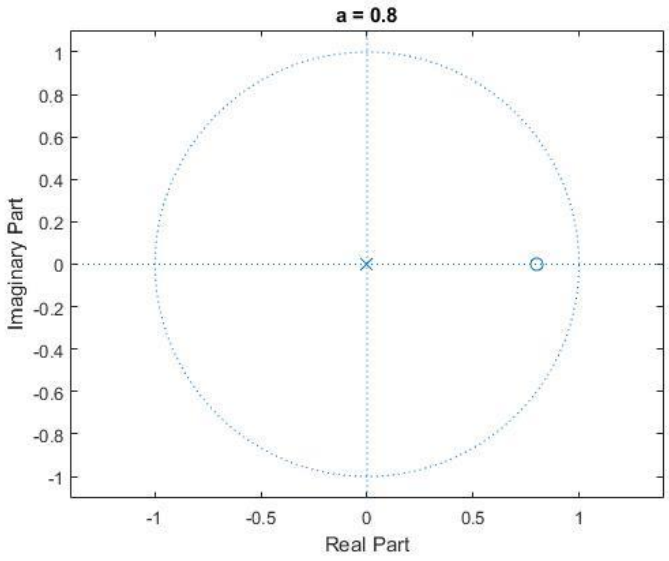

Figure 1: Pole-Zero Plot for a = 0.8

At a = 0.8, the filter is a high-pass filter.

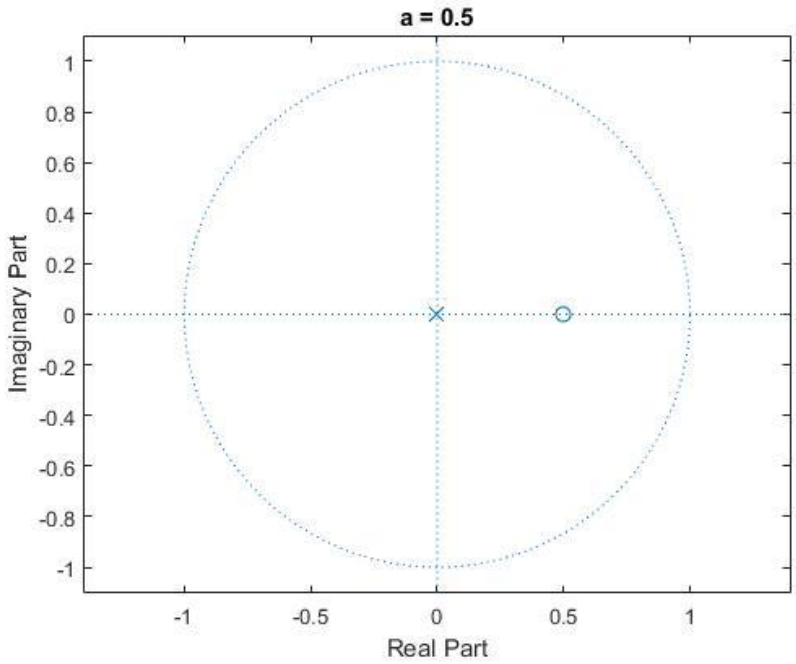

Figure 2: Pole-Zero Plot for a = 0.5

At a = 0.5, the filter is a high-pass filter.

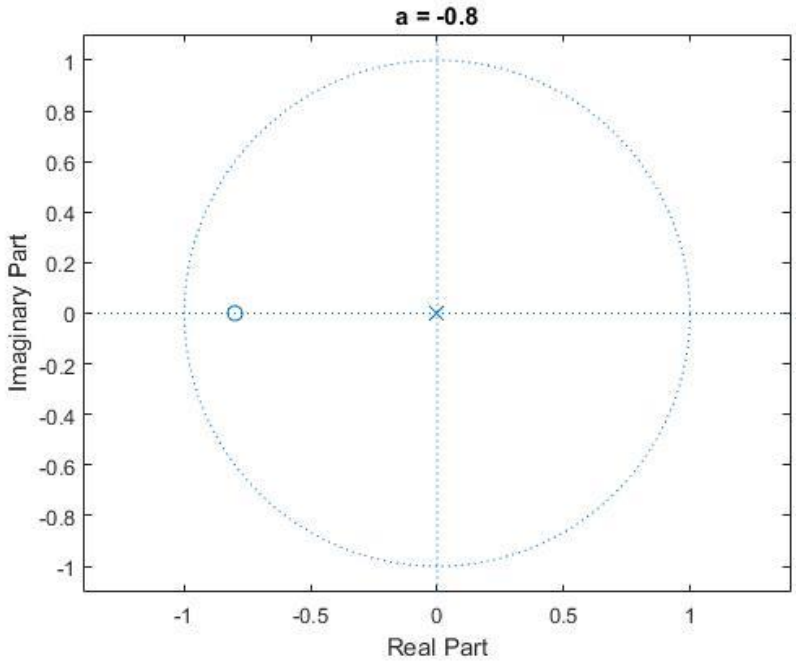

Figure 3: Pole-Zero Plot for a = -0.8

At a = -0.8, the filter is low-pass.

2. The track sounds a lot less deep-pitched; sounds like a lack of bass (or low frequencies). There's a similar effect in removing the subwoofer in speakers while playing music.

3.

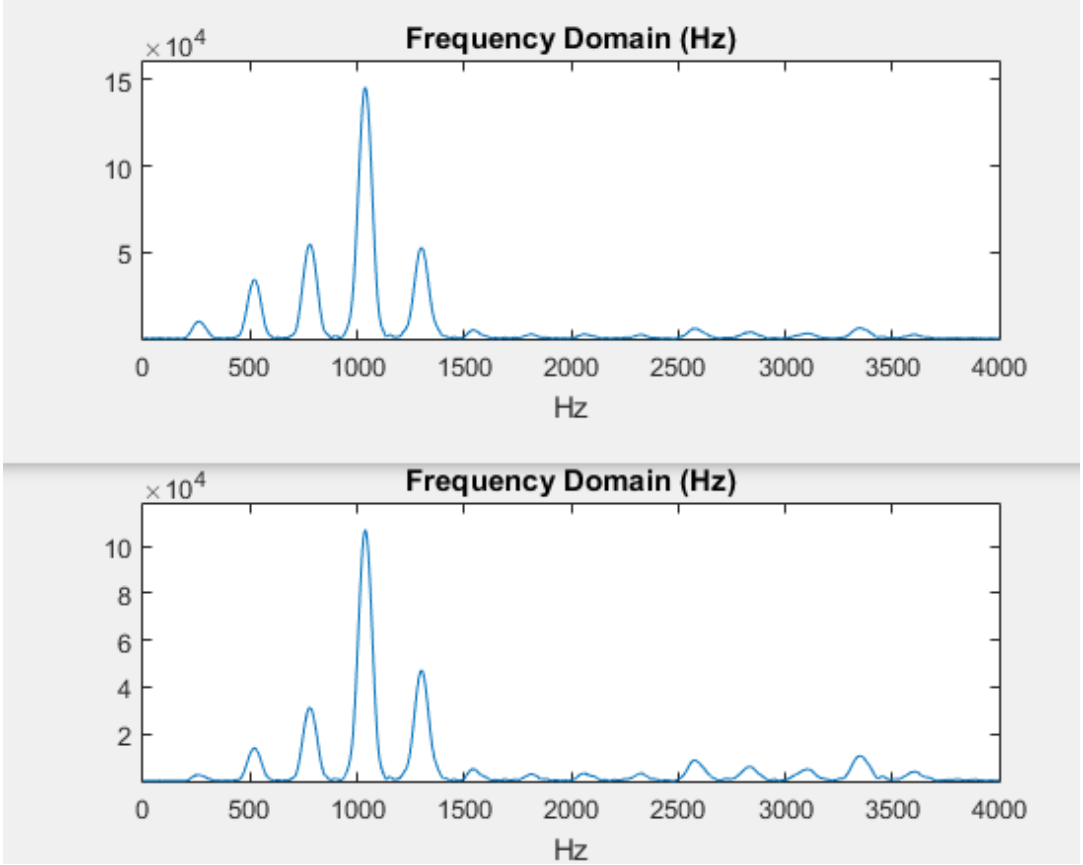

Figure 4: Frequency spectra for the steady-state /a/ sound, original (top) and after the pre-emphasis filter (bottom)

Time domain: amplitudes generally dropped throughout the signal. Frequency domain: peaks/amplitudes in the spectrum are lower than the original. There's a general decrease of gain to about 66%~80% of the original signal's.

Formants:

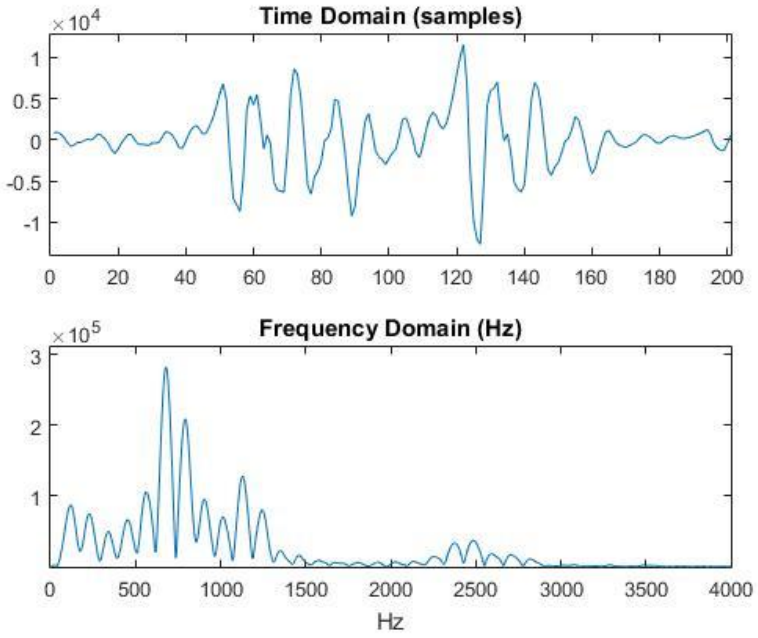

Figure 5: Spectrum of "male_a"

1. By looking at where three noticeable peaks may be formed (mainly looking at where lobes reach local maxima), formant frequencies may be approximated at 200Hz, 600Hz, and 1100Hz. Harmonic spacings are around 125Hz each.

Pitch Period Estimation:

1.

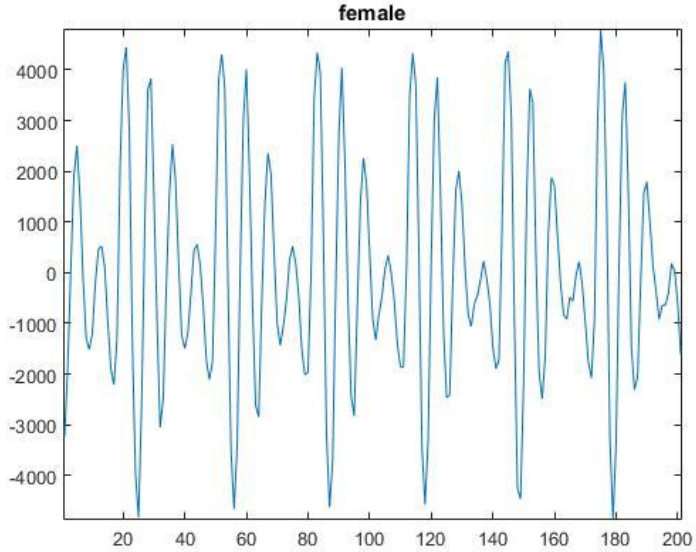

Figure 6: Steady-State Female /a/ Sound

Pitch period (samples) ≈ 30 samples, $\frac{30\ samples}{8000\frac{samples}{s}} = 3.75\text{ms}$. Fundamental frequency $= \frac{1}{3.75ms} =$ 266.67Hz

2.

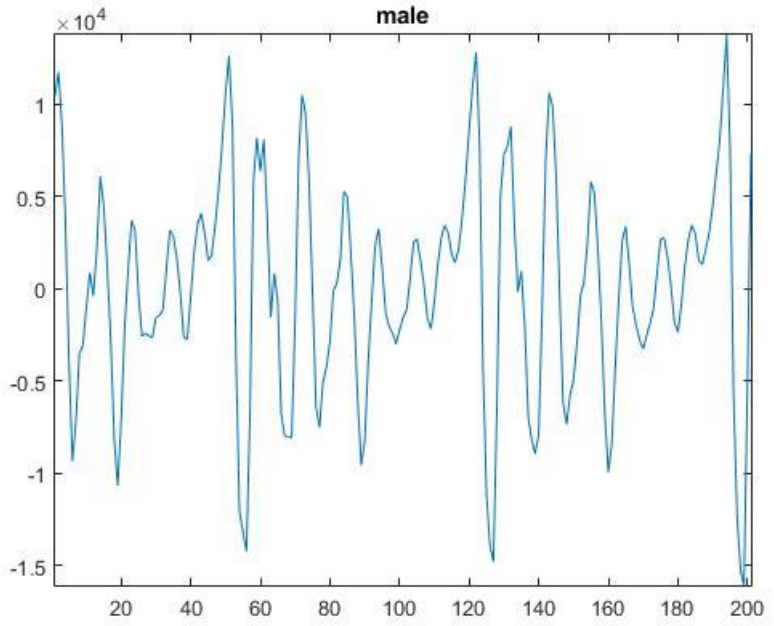

Figure 7: Steady-State Male /a/ Sound

Pitch period (samples) ≈ 70 samples, $\frac{70\ samples}{8000\frac{samples}{s}} = 8.75\text{ms}$. Fundamental frequency $= \frac{1}{8.75ms} =$ 114.3Hz