EEC 201 Interim Report

Team members:

Chen Wang	916712699	wacwang@ucdavis.edu
Ken Yang	916682693	lsyang@ucdavis.edu

Project Objectives:

For this project, we need to build a graphical user interface letting people upload a segment of voice recorded and get a new voice segment which is generated by MATLAB according to the original voice.

After receiving the void segment, LPC vocoder will analyze it, compress it and generate a new file with lower bit rate, at the same time the vocoder will output the pinch of the voice. According to the pinch, the program will create a new sound.

What we need to do is to find at least two ways to estimate the pinch of the voice and compare their spectrums, then synthesis a new void using this pinch. Furthermore, we can convert a female voice to a male voice by simply changing the pinch we got.

Plan and Tasks

1. Record & play, read & write	Done
2. Generate a spectrogram	Done
3. Build a basic LPC vocoder	Done
4. Optimize the LPC vocoder to generate a low rate output	Mar. 2
5. Build a synthesizer	Mar. 4
6. Estimate the pitch & compare	Mar. 4
7. GUI	Mar. 6
8. Test program on reporting and singing voices	Mar. 8
9. Test on two people talking	Mar. 8
10. Convert male & female voice	Mar. 8
11. Finish the final report	Mar. 11

Methodology

$$\text{vocal cords} \quad \longrightarrow \quad \boxed{\text{vocal tract}} \quad \longrightarrow \quad \text{speech}$$

Vocal cords represents excitation signal e(n).

Vocal tract represents H(z).

So, we can get:

$$e(n) \quad \longrightarrow \quad \boxed{H(z)} \quad \longrightarrow \quad y(n)$$

(Typical order of H(z) is 8~12)

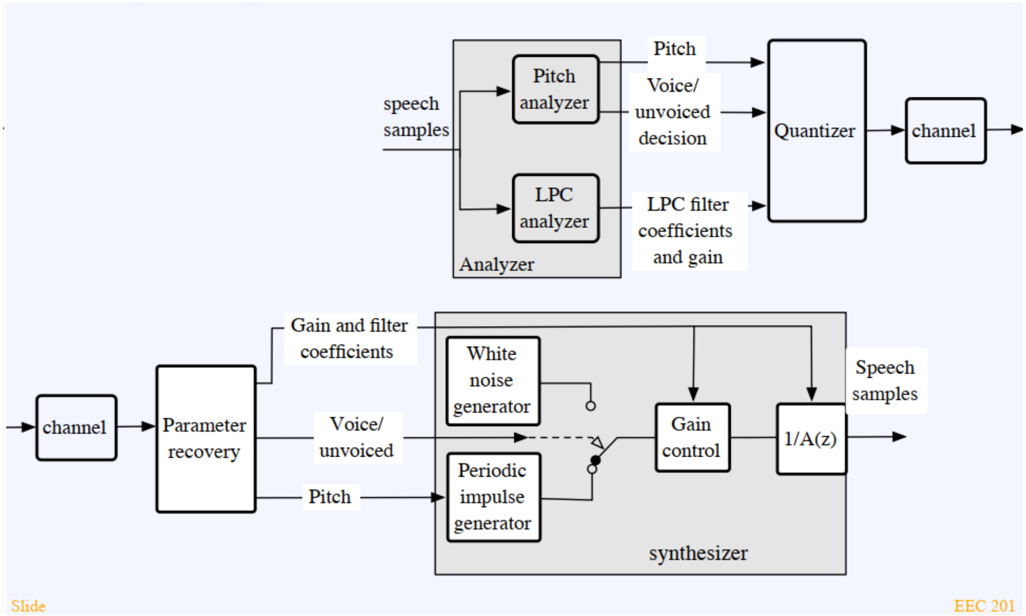
To get e(n):

$$y(n) \quad \longrightarrow \quad \boxed{1/H(z)} \quad \longrightarrow \quad e(n)$$

And since y(n) is the estimation of original voice signal s(n), we can replace y(n) by s(n) to compute e(n).[1]

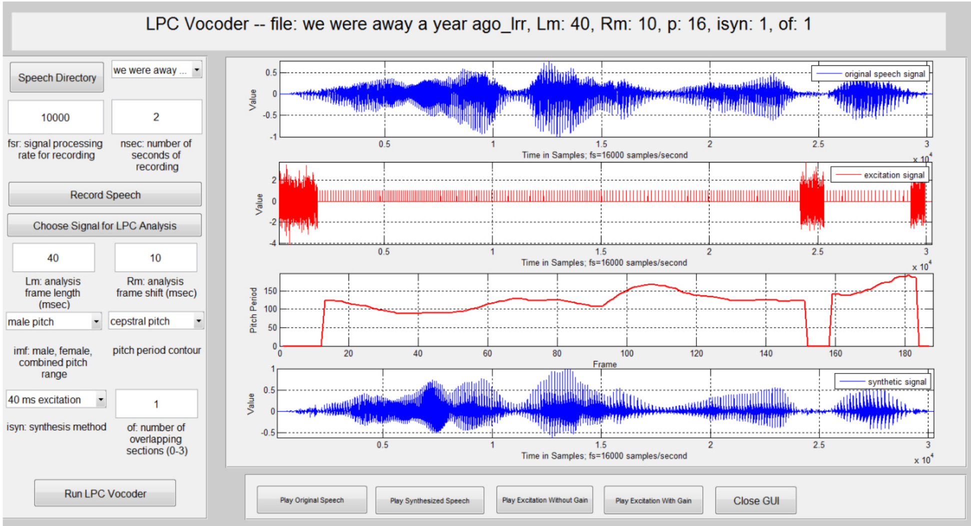
LPC vocoder: LPC vocoder can generate excitation signal and coefficients of H(z) in the LTI system. By this process, bit per second is supposed to decrease dramatically. We will only store data generated from LPC vocoder in files instead of the original voice signal.

Overall method:



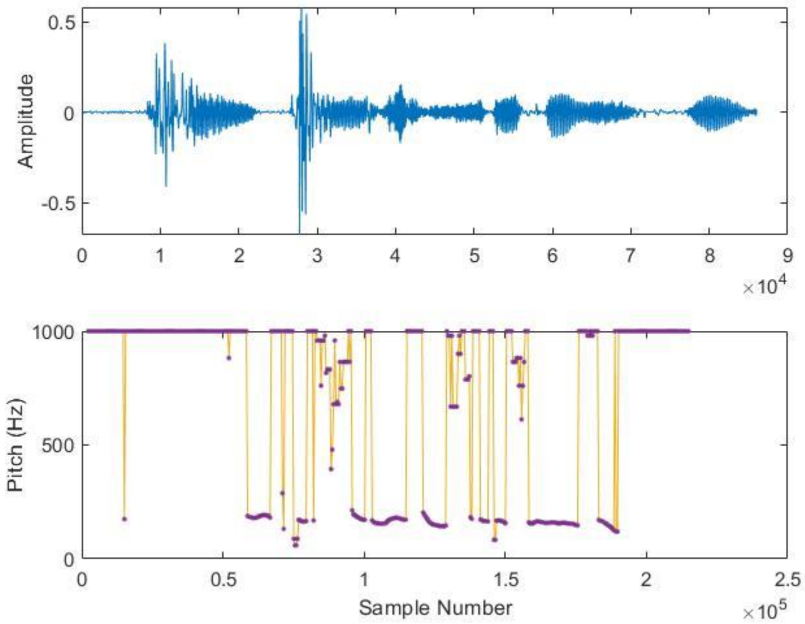
Since voices are actually frequency variance system, so we need to divide the original voice signal into many parts which are less than 3 milliseconds.

Expect GUI Design:[2]

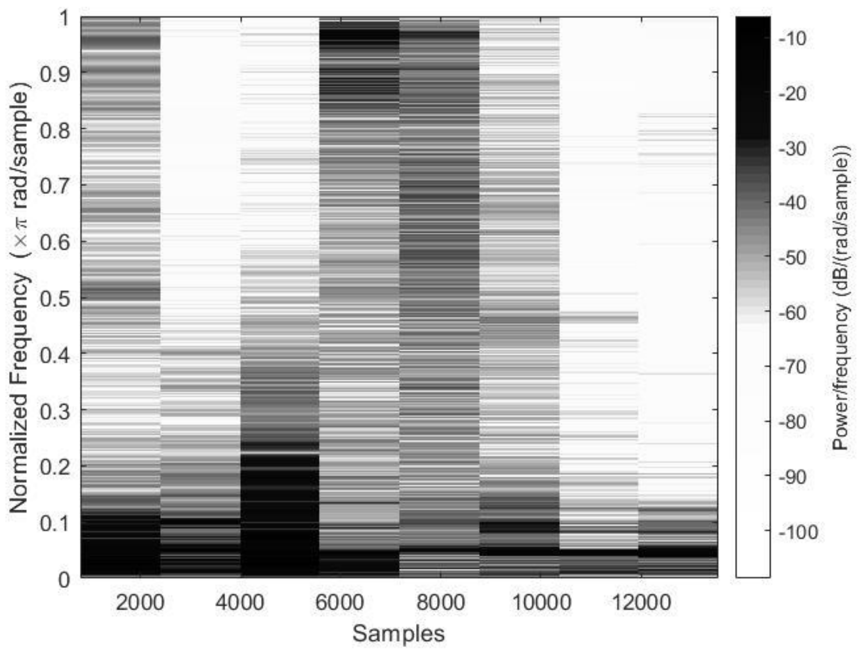


Preliminary Findings and Results:

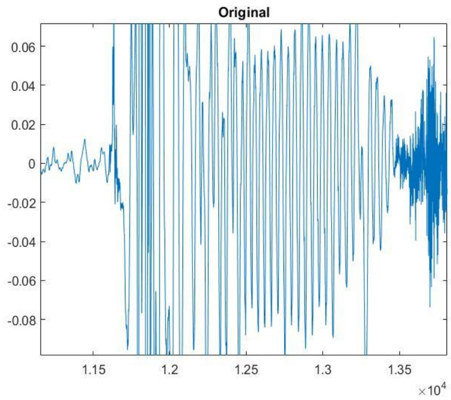
1. A(t) and Pitch

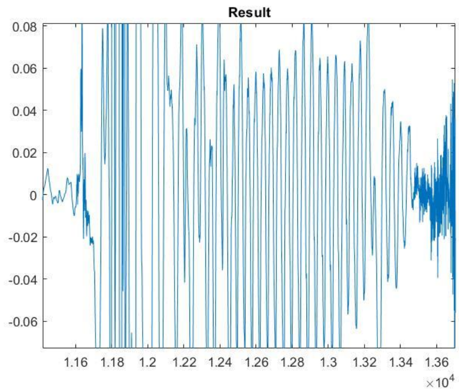


2. Spectrogram



3. Basic LPC vocoder





Duty Division

Chen Wang	Task 4,6,7,8,9,10,11
Ken Yang	Task 1,2,3,4,5,9,11

References

[1] http://eeweb.poly.edu/iselesni/EL713/Speech/speech.pdf

[2] https://www.mathworks.com/matlabcentral/fileexchange/45321-lpc-vocoder

[3] http://www.seas.ucla.edu/spapl/projects/ee214aW2002/1/report.html