CSE534 Spring 2017 Homework Set1

Zhe Ye 50207918

This exercise is intended to practice both recording of speech signal as well as encoding of digital speech signal. Matlab is used to implement the encoding process. Display of the speech signals at various stage of speech creation, processing, and reconstruction is presented.

The 9-second audio speech file is ‘male.wav’ with sample rate at 8000Hz. Before partition the recorded speech, the noise is reduced first using ‘removeNoise.m’. Then partition the speech into segments of 20ms for display and processing. Voice segmentation, including filtering and windowing is finished in ‘segment.m’ function. The graphic rendering of the original recorded samples are as figure 1 below.

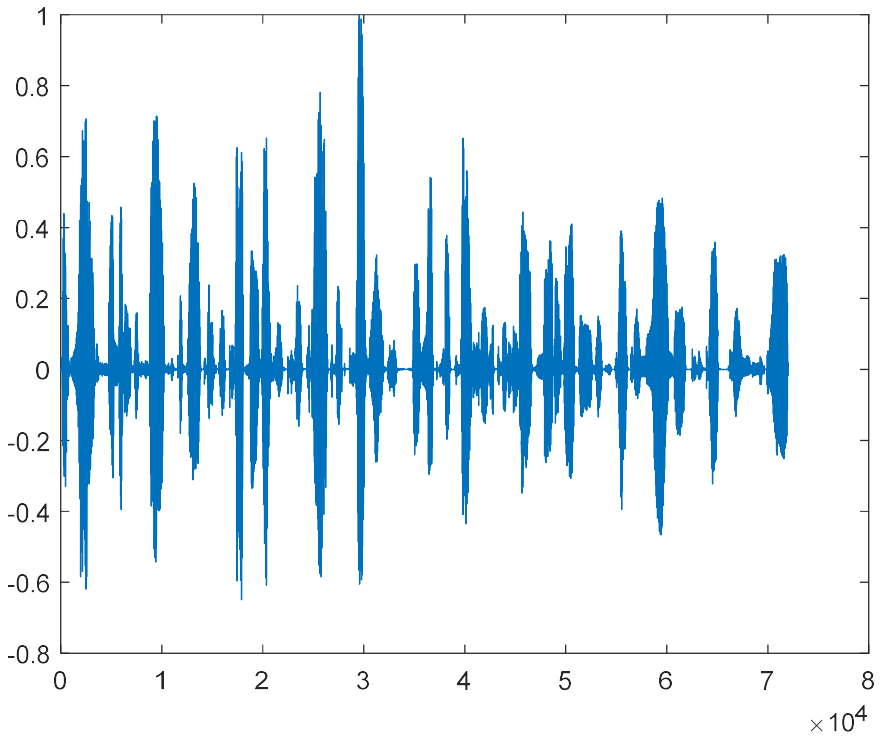


Figure 1 Original speech spectrum

For each speech segment, Voiced/Unvoiced decision and pitch detection is done for the audio for each segment. Pitch data is saved in Matlab workspace (after running through the proj.m program successfully). Voiced/Unvoiced decision is made based on the following figure 2 through a simple Autocorrelation calculation method. LPC coefficient and Gain estimation is calculated by ‘lpcFeatures.m’ function. The table 1 shows the LPC coefficient.

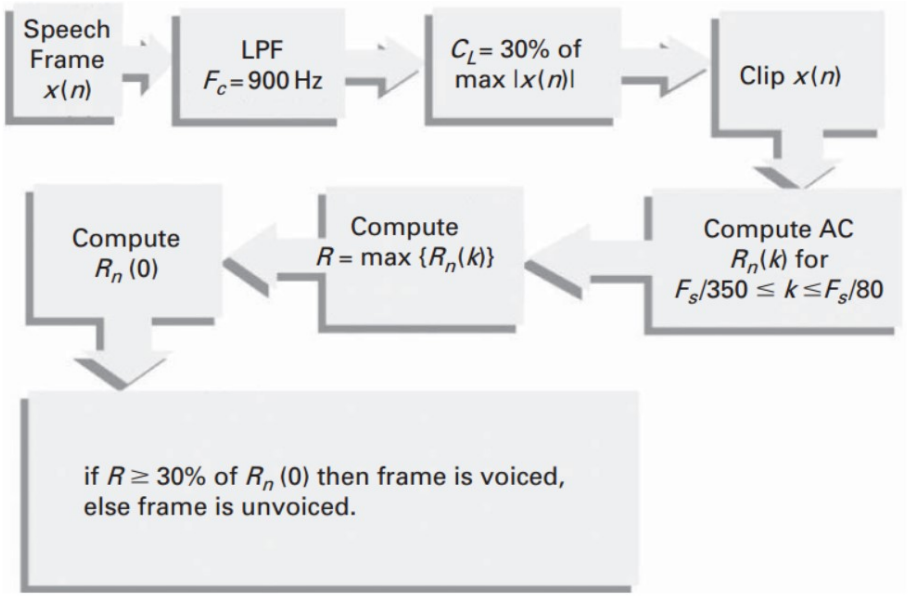
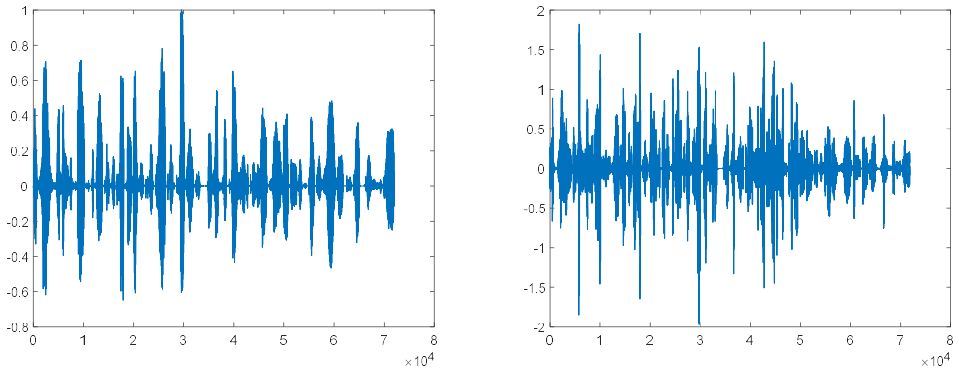


Figure 2 Voiced/Unvoiced decision method

alpha	0.3460	-1.0870	-0.3378	-0.7035	0.0436	-0.9414	-0.1812	-0.2784	-0.3690	-0.1686

Table 1 LPC coefficient

The last step is to reconstruct the audio with the coefficient we got in the steps before. This is finished with function ‘reconstruct.m’ . The output data is saved in workspace with variable name ‘rec’.



Orignal After LPC

Figure 3

The Linear predictive coding (LPC) is a tool used mostly in audio signal processing and speech processing for representing the spectral envelope of a digital signal of speech in compressed form, using the information of a linear predictive model. In this assignment, the audio after LPC is compared with original audio in figure 3. We can see that LPC performs much better for encoding good quality at a low bite rate and provides extremely accurate estimates.

Reference: The voiced or unvoiced decision with pitch period estimation is achieved through a simplified autocorrelation calculation method (http://www.ee.ucla.edu/ee213a/speech/speech.htm)