EC516 Lab 4 (Fall 2016)

Due Thursday Dec 1st, 2016

Lab 4

This lab is the fourth in a series of MATLAB assignments designed to explore Digital Signal Processing from an **alternative patternization** perspective and its importance in computational tasks such as **Learning, Enhancement, Authentication, Detection, Encoding, Recognition or Separation (LEADERS).** In this particular lab you will get a sense of how alternative patternization of a speech signal enables us to efficiently compress and encode the speech signal.

In this lab, the alternative patternization used for the purpose of encoding the speech signal is known as Linear Predictive Coding (LPC). LPC is a form of parametric signal modelling and is widely used in the field of speech coding. In order to perform LPC, the speech signal is first broken into 20 msec overlapping frames and windowed. Each of these 20 msec frames (160 samples at 8KHz sampling rate) are then alternatively patternized as 19 numbers. These 19 numbers represent the parameters(18 Poles and the Gain) of an all-pole model, just as in the case of Indirect Least Squares Modelling. There are a few additional parameters extracted from each frame which can be encoded with very few bits[1].

[1] The additional parameters include whether the frame is voiced/unvoiced and also the pitch period. Voiced sounds are the sounds produced due to the vibration of the vocal chords. Example of voiced sounds include the sounds produced when we pronounce vowels (ä,æ,e). Conversely, unvoiced sounds do not involve the vibration of vocal chords. Examples of unvoiced sounds include ‘s,sh’.

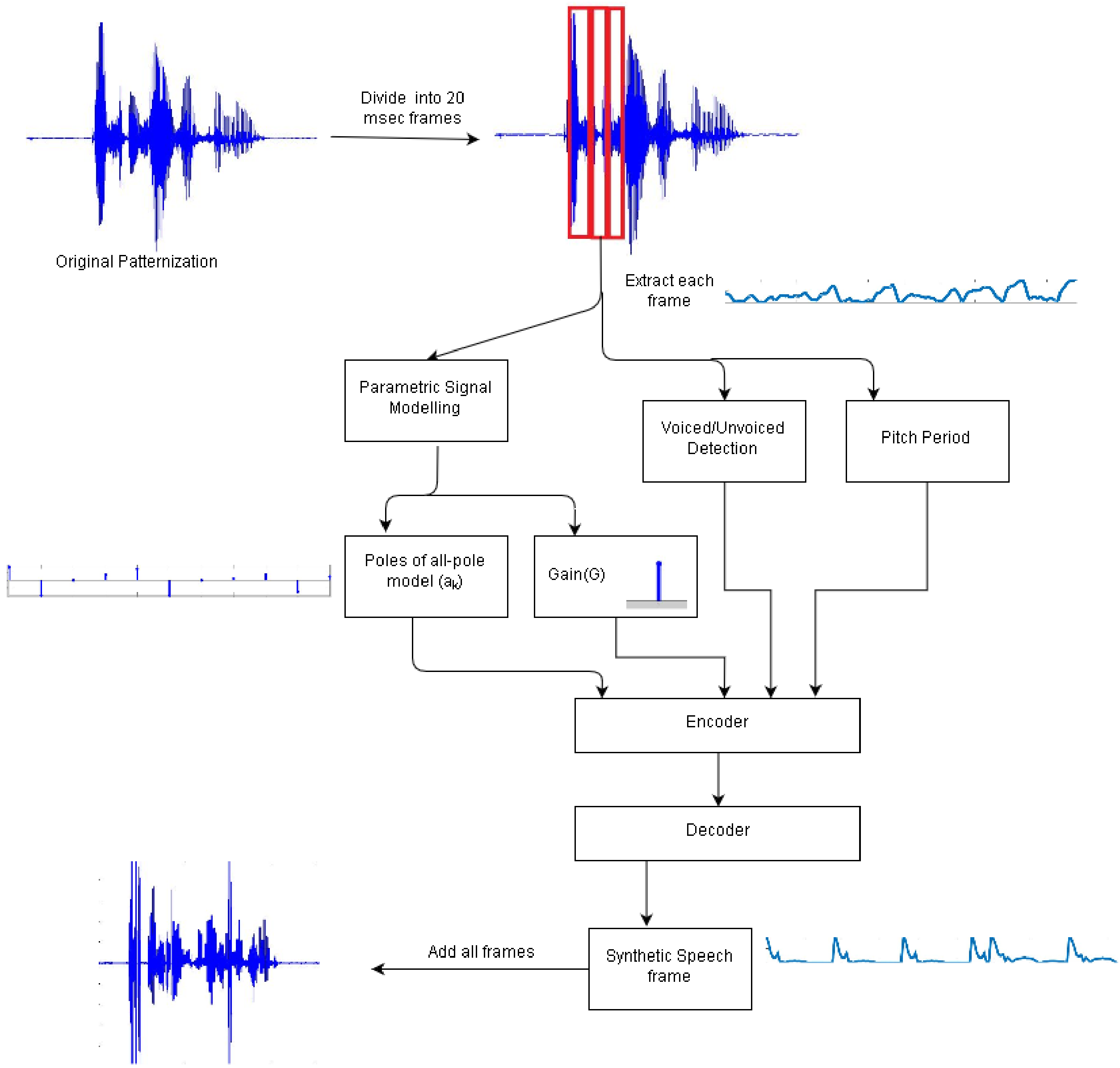


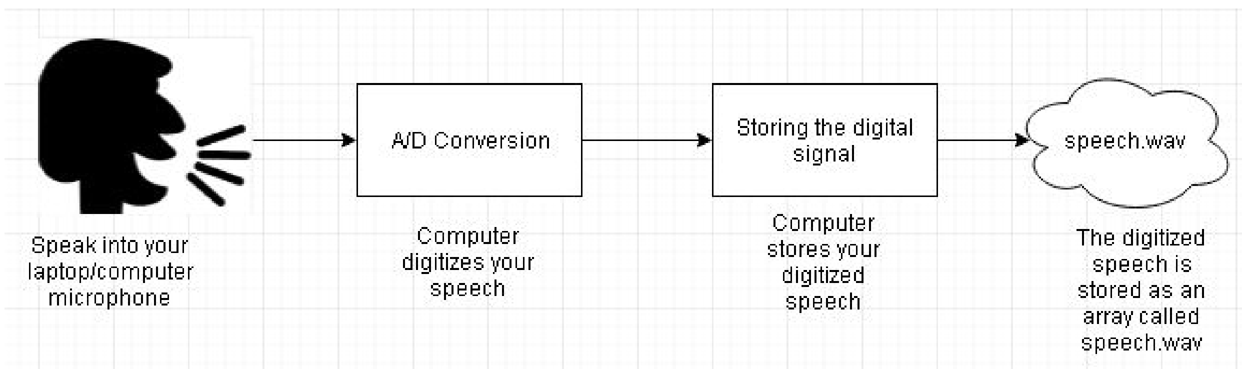
Figure 1: Computation of synthetic speech using Alternative patternization (LPC)

The original patternization is recorded at a sampling rate of 8 KHz. Assuming each sample is represented by 16 bits, the total number of bits per second of the original patternization =

16*8000 = 128 kbits/second. In this lab, the goal is to represent the original patternization at much fewer bits per second(32, 16, 8, 4 kbits/second) using our alternative patternization (LPC).

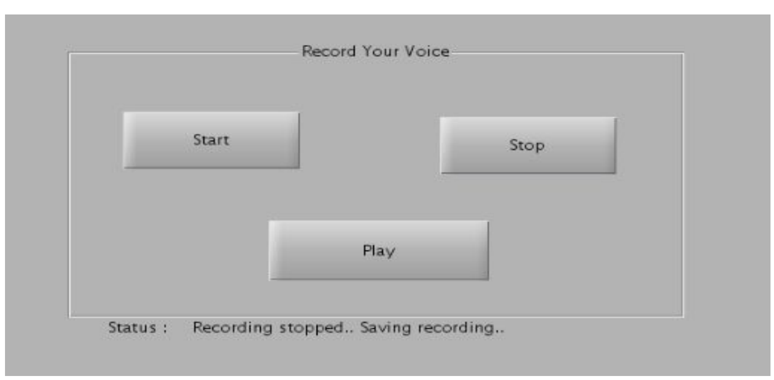
This lab is divided into 5 sections as described below:

a) **Recording your voice(ROV):**



The first step is to record your voice into MATLAB using your laptop/computer microphone. Since our computer can only work on digital data, it digitizes your voice into a digital speech signal and stores it into a file named speech.wav, which is an array of real values, that form a pattern as a function of time.

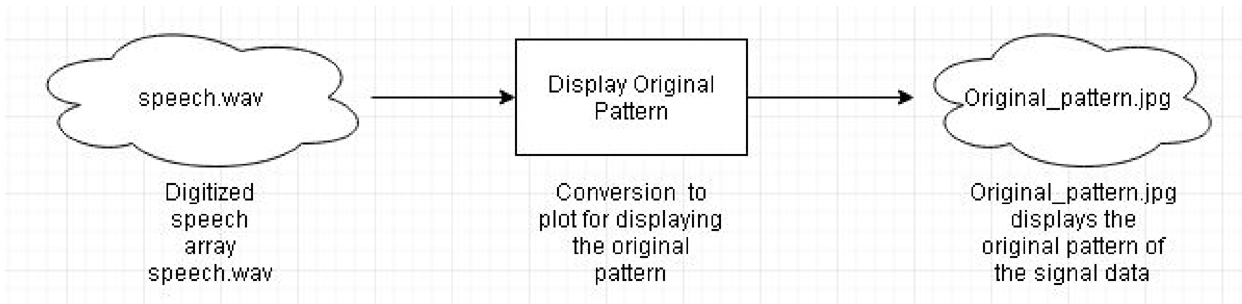
GUI Usage for ROV:



- Download and unzip EC516_Lab4.zip provided to you into your current directory in MATLAB.

- Open and run EC516Lab4.m provided to you. This will open a GUI named 'EC516Lab4gui'.
- Click on "Start Recording" when you are ready. Speak a sentence such as "Two plus seven is less than ten" into your laptop's microphone and click on "End Recording". Please make sure you do this in a quiet room. Also ensure that your laptop has a microphone and an audio outlet.
- You can now play the recorded audio signal by clicking on the "Play" button.
- You will see that an audio file 'speech.wav' has been created and saved in your current directory.

b) <u>Display Original Pattern (DOP)</u>



Now that we have stored the original pattern, we would like to display the original pattern(DOP) of our digitized speech signal (speech.wav). An example of 'original_pattern.jpg' is shown below:

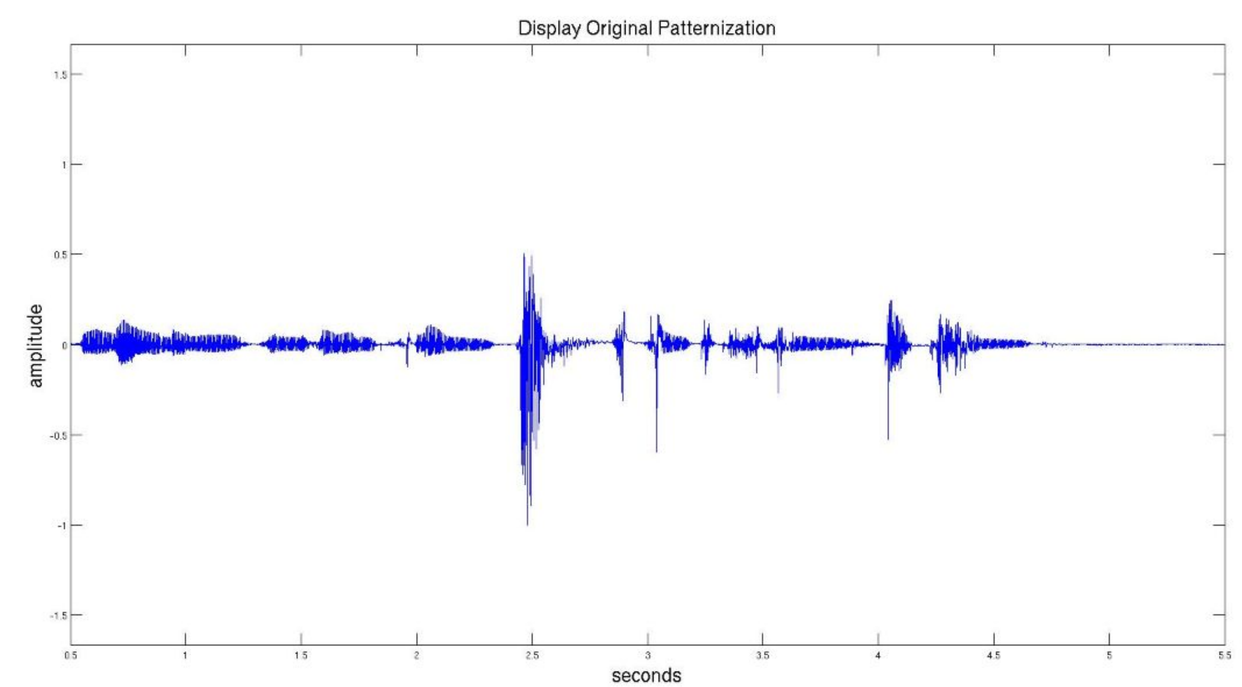
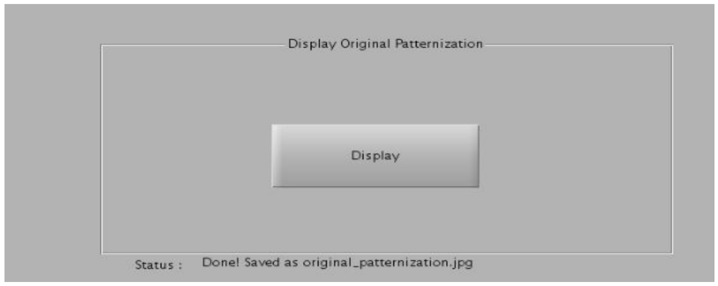


Figure 2 : Display Original Patternization

GUI Usage for DOP:



- Click on the "Display Original Pattern" button on the GUI. A plot of the original pattern pops up and this plot will be saved as "original_pattern.jpg" in your current directory.

c) <u>Compute Alternative Patternization(CAP) :</u>

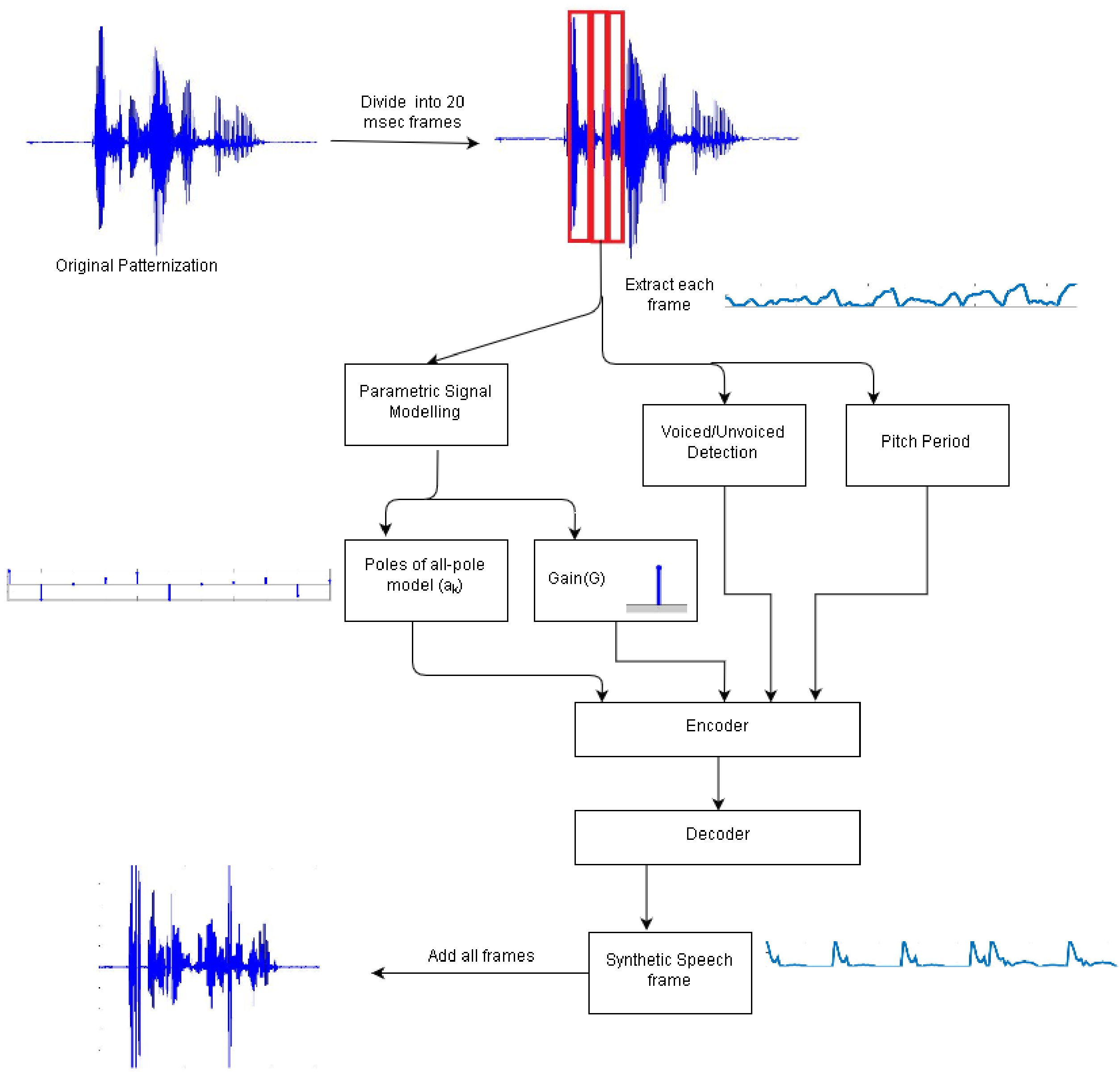


Figure 3: Computation of Alternative Patternization

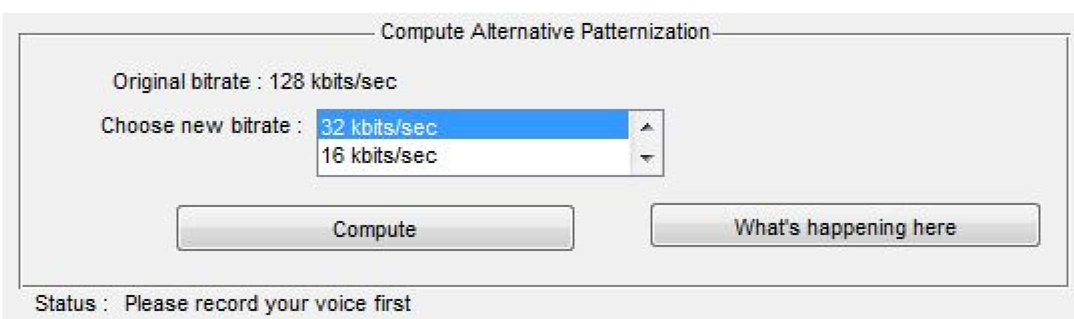
In this step we convert the Original pattern of our speech signal data into an alternative pattern. This process is known as the Alternative Patternization of Signal Data (APSD). The alternative patternization is achieved using a form of Parametric Signal Modelling known as Linear Predictive Coding. The alternative patternization then enables us to represent the Original Patternization using much fewer bits.

The distribution of bits using the alternative patternization is shown below :

	32 kbits/sec	16 kbits/sec	8 kbits/sec	4 kbits/sec
Predictor Coefficients	18*16=288 bits/frame	18*8=144 bits/frame	18*4=72 bits/frame	18*2=36 bits/frame
Gain	16 bits/frame	8 bits/frame	4 bits/frame	2 bits/frame
Pitch Period	16 bits/frame	8 bits/frame	4 bits/frame	2 bits/frame
Voiced / Unvoiced	1 bit/frame	1 bit/frame	1 bit/frame	1 bit/frame
Frames	100 frames/sec	100 frames/sec	100 frames/sec	100 frames/sec
Total	321 bits/frame *100 frames/sec= 32 kbits/sec	161 bits/frame*100 frames/sec= 16 kbits/sec	81 bits/frame*100 frames/sec= 8.1 kbits/sec	41 bits/frame*100 frames/sec= 4.1 kbits/sec

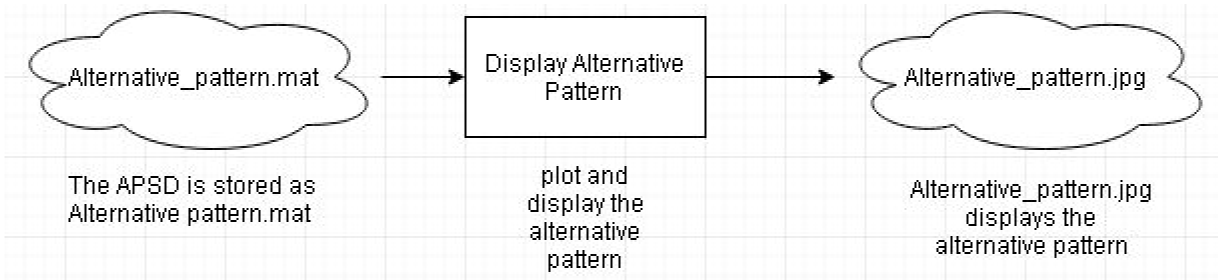
Table 1: Distribution of bits used for encoding using alternative patternization.

GUI Usage for CAP:



- Select a bitrate out of 32, 16, 8 or 4 kbits/sec and click on compute. This will encode the original patternization at 128 kbits/sec to the bitrate selected by you.
- You will find a new file ‘Alternative_pattern.mat’ created and stored in your current directory.

d) Display Alternative Patternization(DAP):



Now that we have computed an alternative pattern, we would like to display the alternative patternization of our signal data. An example of ‘alternative_patternization.jpg’ is shown below:

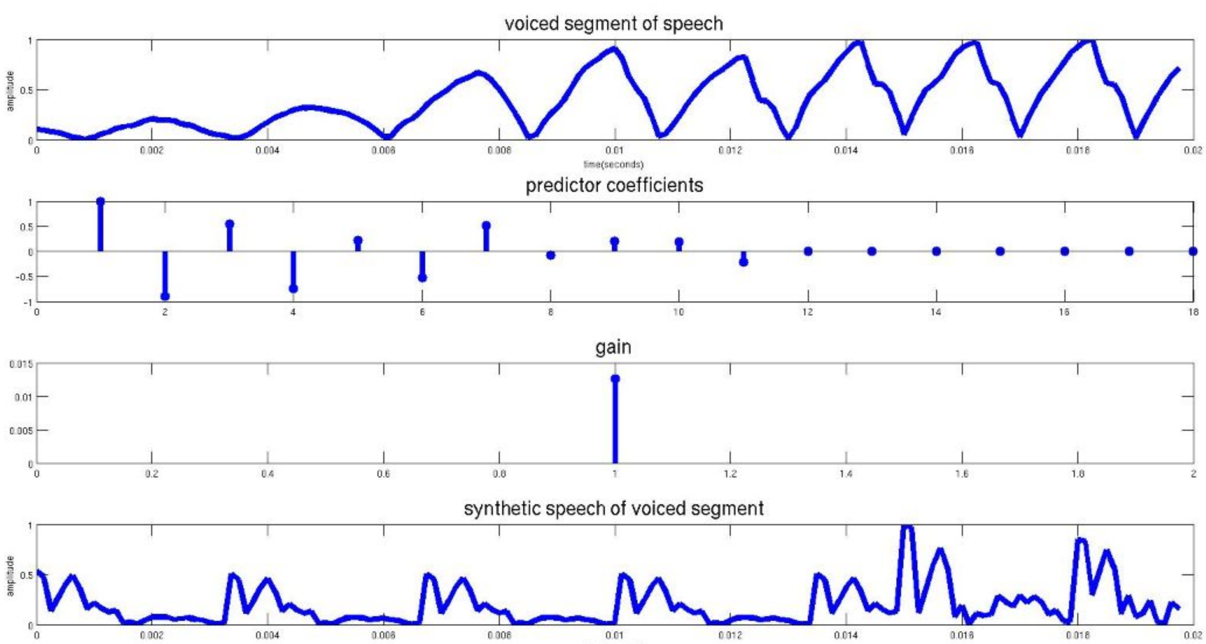
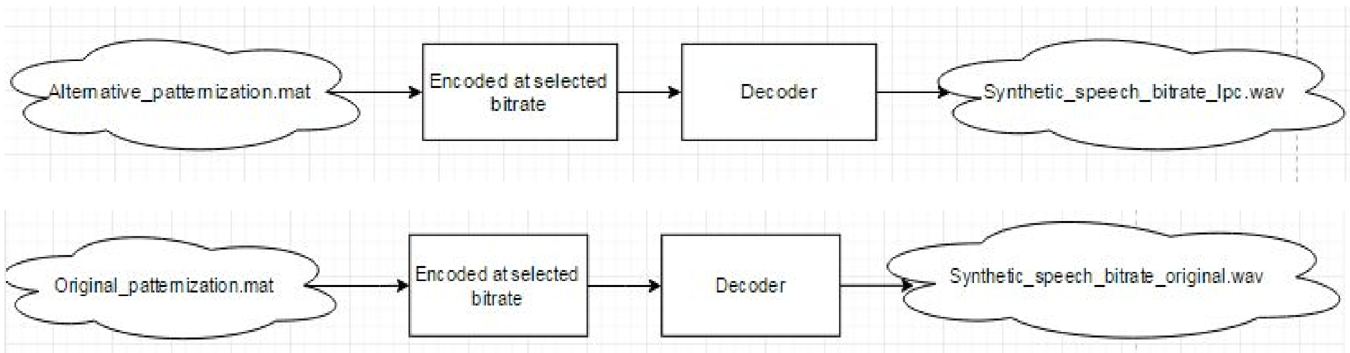


Figure 4 : Display Alternative Patternization. This figure shows a voiced segment of the speech signal. We perform LPC, which is a form of Parametric Signal Modelling on this frame and obtain the predictor coefficients (a_k) and Gain(G). These values represent the alternative patternization. The alternative patternization is then encoded and decoded at the selected bit rate to obtain the synthetic form of the voiced frame.

e) Play Encoded and Decoded Voice (PEDV):



In this step, we will play the voice encoded and decoded at the selected bitrates using the alternative patternization (LPC) and compare it with the voice encoded and decoded by quantizing the original patternization to the same bitrate.

GUI Usage for PEDV:

- Click on Play Voice Encoded using Original Patternization. This plays the voice encoded and decoded by quantizing the original patternization to the bitrate selected in CAP.
- Click on Play Voice Encoded using Alternative Patternization. This plays the voice encoded and decoded using the alternative patternization to the bitrate selected in CAP.
- A plot of the synthetic speech using the alternative patternization at the bit rate selected in CAP pops up and is saved as synthetic_speech.jpg.

An example of synthetic_speech.jpg is shown below :

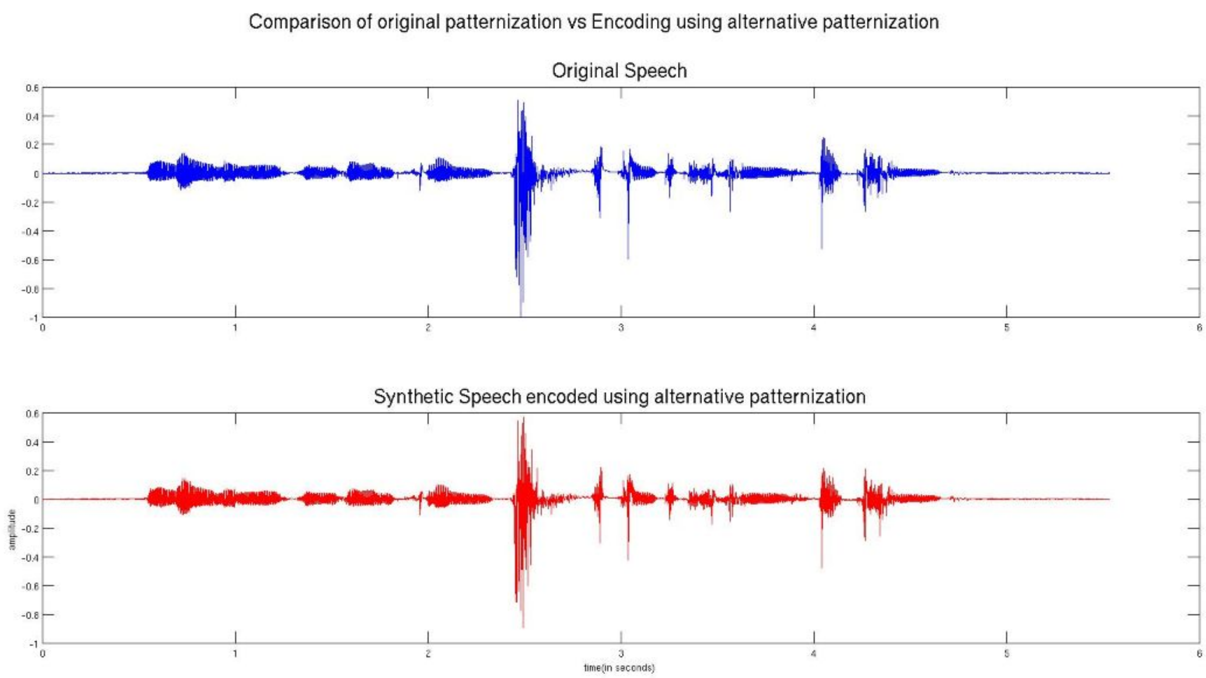


Figure 5: This figure shows the Synthetic speech signal obtained after encoding and decoding the recorded speech signal using the alternative patternization in red and the original patternization in blue.

<u>Submission:</u>

1) Record the sentence 'Two plus Seven is less than 10' and perform all the steps from a) to e) above. Print and submit the 'Alternative_patternization.jpg' and the 'synthetic_speech.jpg' at a bitrate of 4 kbits/second.
2) Repeat the same task as above with 2 sentences of your choice. Print and submit the 'Alternative_patternization.jpg' and the 'synthetic_speech.jpg' at a bitrate of 4 kbits/second in each case.