MULTIMEDIA DATA PROCESSING
SCSV 3223-02

LAB 02- SPEECH PATTERN REPRESENTATION [WAVEFORM]

Prepared by: TAN SEE JOU
A17CS0218

INTRODUCTION

The aims for this lab is to find out and differentiate the segment of the speech waveform for voiced speech, unvoiced speech and silence.
The measurements used:

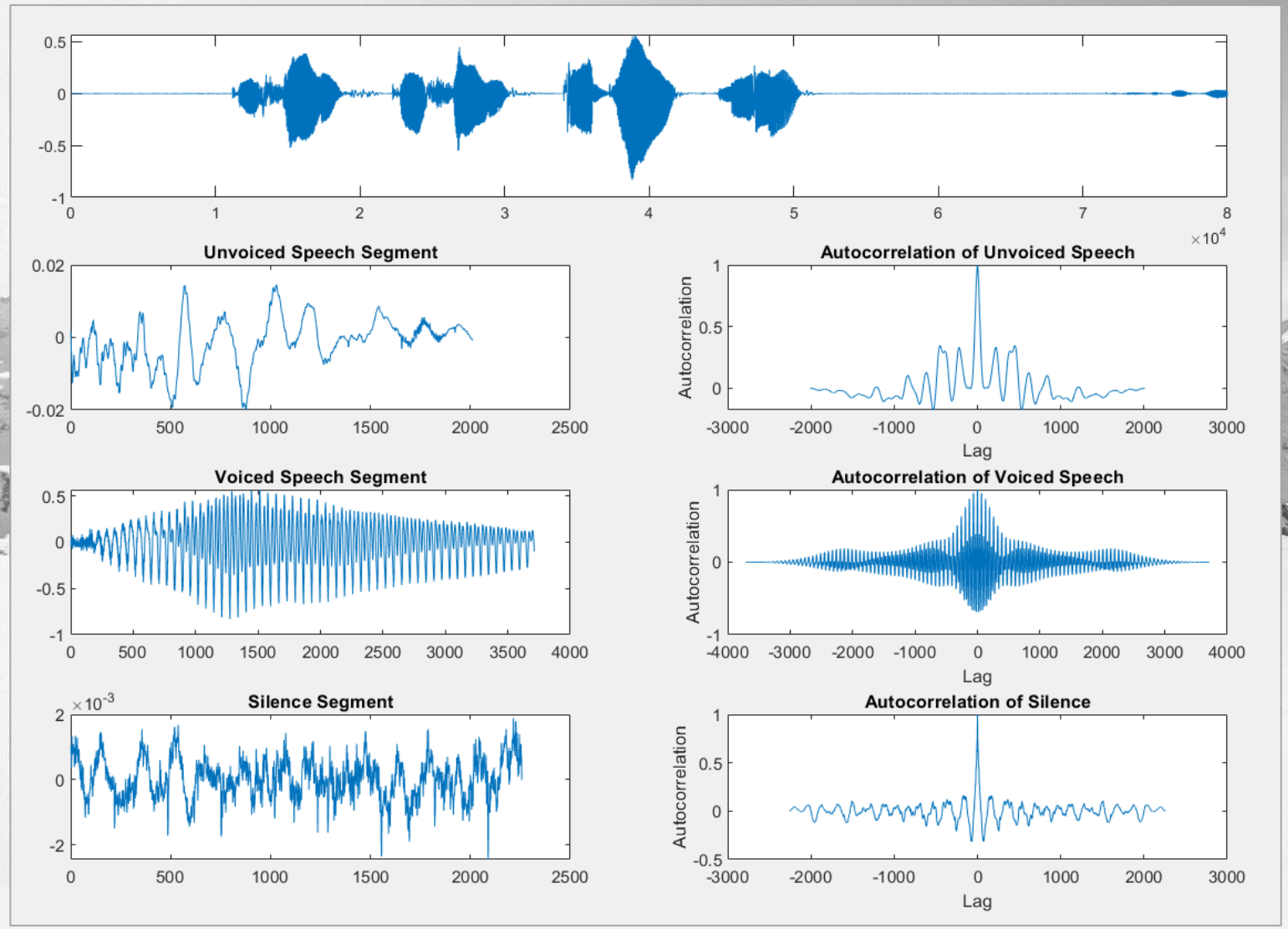
- Energy of the signal
- Autocorrelation coefficient

The data used is import from set C-0075.wav.

DATA USED AND SCRPT

```
A = audioread('0075.wav');
subplot(4,2,1:2); plot(A);

B=A(3.084e+04:3.285e+04);  % UnVoiced
subplot(4,2,3); plot(B); title('Unvoiced Speech Segment');

C = A(3.757e+04:4.128e+04); % Voiced
subplot(4,2,5); plot(C); title('Voiced Speech Segment');

D = A(5579:7836); % Silence
subplot(4,2,7); plot(D); title('Silence Segment');

subplot(4,2,4); plotCorr(B); title('Autocorrelation of Unvoiced Speech');
subplot(4,2,6); plotCorr(C); title('Autocorrelation of Voiced Speech');
subplot(4,2,8); plotCorr(D);title('Autocorrelation of Silence Speech');
```

RESULT FOR VOICED, UNVOICED AND SILENCE SEGMENTS



RESULTED ENERGY:

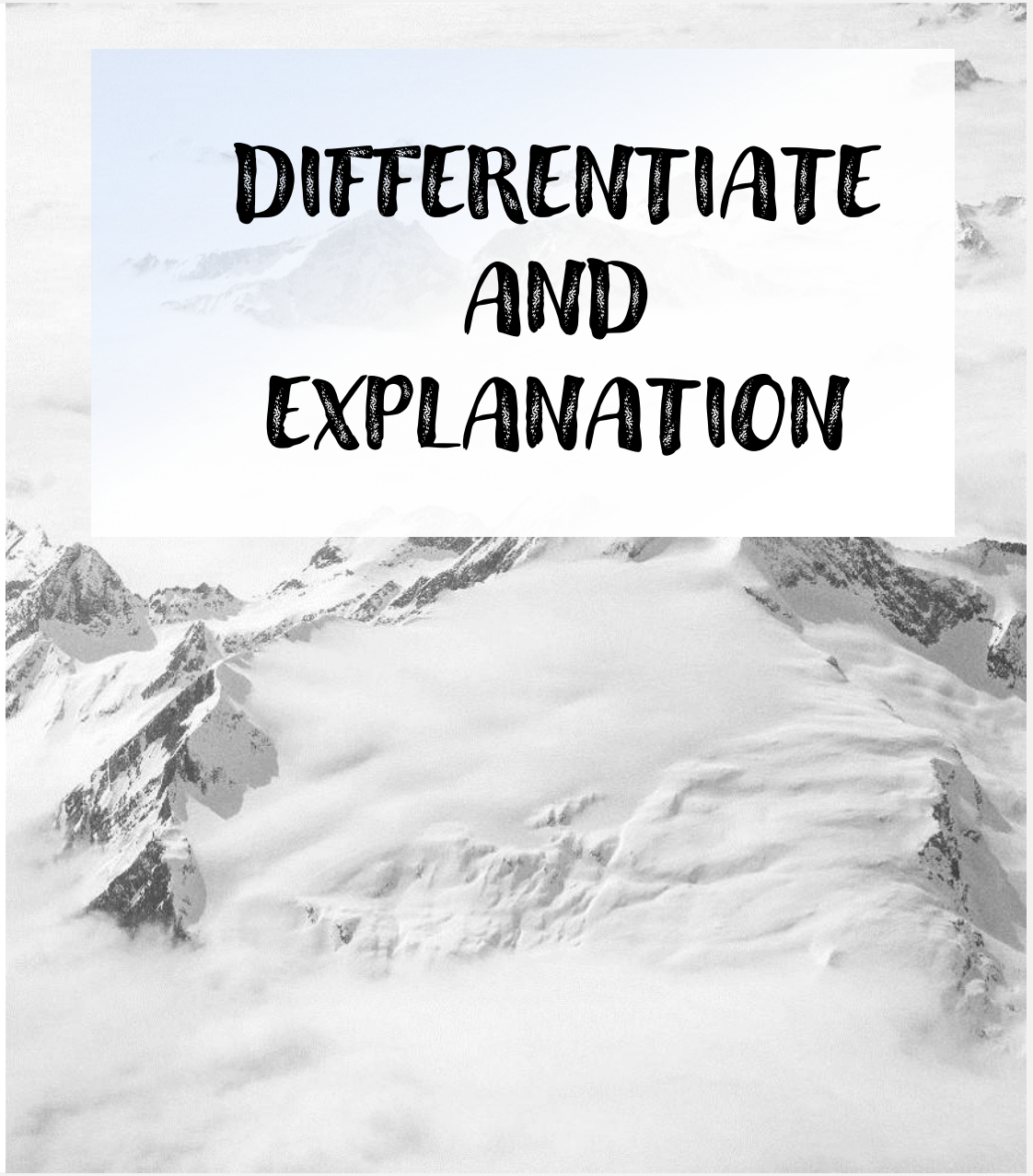
ENERGY		
	SUM	AVG
UNVOICED	362.8003	0.0902
VOICED	715.8952	0.0965
SILENCE	267.5691	0.0503

VOICED > UNVOICED > SILENCE

From visual perception, these 3 segments can be differentiated from the waveform, for

- voiced speech segment, the wave is nearly periodic, as the input sound is also periodically.
- unvoiced speech segment is non-periodic .
- silence is also non periodic and the main point that differentiate it from unvoiced speech is the associated energy or known as the amplitude. Silence has a very low or nearly negligible amplitude.

From the energy table, it is clear to notice that the voiced data had much higher energy compare to the rest because of its periodicity. The energy of unvoiced data is lower than the voiced but higher than the silence. Therefore, energy is also one of the way to differentiate between them.



Do you think the segmentation task in assignment 2 can be enhanced using this method? If can, How ?

Yes. This is because segmentation help us to determine the voiced, unvoiced speech and silence. Therefore, we can focus on the voiced parts.

THANK YOU