INDIAN INSTITUTE OF INFORMATION TECHNOLOGY TIRUCHIRAPPALLI

Department of Electronics and Communication Engineering

ECOE18- Digital Speech Processing

Assignment 01

Posted: Feb.24,2020 **Due: Mar.02,2020**

Note: You can use either MATLAB or Python to solve the first two questions.

1. Write a program to play out a sequence of audio files (including file separation beep tones) as specified below.
    a. s5_synthetic.wav
    b. beep_fs_10000.wav
    c. s5.wav
    d. beep_fs_16000.wav
    e. s5_synthetic.wav
2. Write a program to analyse and plot the following measurements:
    a. Entire Speech Waveform
    b. Short-Time Energy, En
    c. Short-Time Magnitude, Mn
    d. Short-Time Zero-Crossing, Zn

    Use the speech waveforms s5.wav and should.wav to test your program. Choose appropriate window sizes, window shifts, and window for the analysis. Explain your choice of these parameters. (Don't forget to normalize your short-time zero crossing analysis depending on the sampling rate of the speech signal in each file).
3. Which of the following functions cannot be a valid autocorrelation function? For all invalid autocorrelations, explain the reason why.

$$\text{(a) } R_x(\tau) = 2e^{-\tau^2},\ -\infty < \tau < \infty.$$
$$\text{(b) } R_x(\tau) = |\tau| e^{-|\tau|},\ -\infty < \tau < \infty$$
$$\text{(c) } R_x(\tau) = \left(\frac{\sin(\pi\tau)}{\pi\tau}\right)^2,\ -\infty < \tau < \infty$$

4. Consider the calculation of short-time energy via the relation

$$E_n = \sum_{m=-\infty}^{\infty} x^2[m]h[n-m]$$

Consider the use of a system with impulse response, h[n], of the form

$$h[n] = \left[-\frac{a}{b-a}a^n + \frac{b}{b-a}b^n\right]u[n]$$

where a and b are distinct with |a| < 1 and |b| < 1

a) Determine a difference equation implementation for the computation of short-time energy using the impulse response given above.

b) What general property must h[n] have in order that it be possible to find a recursive implementation.

5. A digital signal, x[n], is defined as:

Use DFT to determine

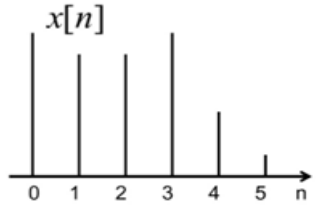
a) Xˆ[k] and inverse transform of Xˆ[k]

b) Obtain the sequence xˆ[n].

c) Plot xˆ[n] for N = 40, N = 10, N = 5, and N = 3

d) Compare the plots with the plot of x[n]

$$x[n] = \begin{cases} 7 & n=0 \\ 6 & n=1 \\ 6 & n=2 \\ 7 & n=3 \\ 3 & n=4 \\ 1 & n=5 \\ 0 & n \geq 6, n \leq 0 \end{cases}$$



6. The figure below shows plots of 6 speech short-time log magnitude spectra as obtained using a Hamming window of an appropriate length. The set of 6 spectra include vowel and consonant regions by a male, a female and a child talker.

a) Which of the 6 spectra are most likely to have been uttered by a child? What leads you to this conclusion?

b) Which of the 6 spectra correspond to voiced sounds?

c) Which of the voiced speech spectra most likely come from an adult male; which from an adult female?