

END Semester Examination 2014

Subject: **INTRODUCTION TO DIGITAL SPEECH PROCESSING**

Code: ET60007

Time: 3:00 Hours

PART-A:-10*2=20; PART-B:-5*16=80

Full Marks =100

Answer all the questions of PART-A and PART-B

(Please enclose the Annexure-1 along with the answer script)

PART-A

1. An audio signal is recorded using the $F_s = 16$ kHz, encoded with 16 bit and recoded in MONO format. To store 200 ms signal in PCM WAV format how much memory is required? If the above signal fundamental frequency is 250 Hz. How many samples will be in one pitch period?
2. Draw the polar pattern which shows the variation of sensitivity in 360 degrees around a unidirectional microphone.
3. Derive the expression of $y[n]$ for the system given in Figure-1

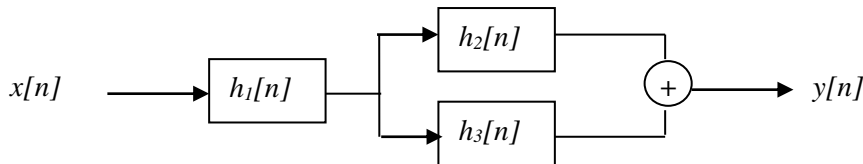


Figure-1

4. A low-pass digital FIR filter is designed using the following specification.
Transition bandwidth = 100 KHz, Cut off frequency = 2 kHz, Sampling frequency of the input signal is 8 kHz and the filter is implemented using HAMMING window. Where the Transition bandwidth for HAMMING window is $8\pi/M$, M is the order of the filter.
Determine the order of the FIR filter?
5. If the centroid of two data class A (1.1, 2.5) with variance (0.4, 0.5) and class B (2, 2.5) with variance (0.5, 0.1) the value of a test data x is (2.5 1.5). Using Mahalanobis distance function find out x is closed to which class.
6. Write the name of two time domain methods for F_0 extraction?
7. Which of the following is true for narrow band spectra?
 - a. Time resolution low, frequency resolution high
 - b. Time resolution high, frequency resolution low
 - c. None of the above.
8. Write the phonetic transcription of the last word of your surname and write the place and manner of articulation of any two phonemes of your name.
9. Why the child speech has high F_0 and formant compare to a adult
10. Write the name of three Supra-segmental Speech parameters

PART-B

1. (a) A voiced operated wheelchair operation is designed using the following words

[BACKWARD; FORWARD; LEFT; RIGHT; STOP; TURN; SLOW; FIRST]

Annexure-1 shows wideband spectrograms of one version of each of these words. Using your knowledge of acoustic phonetics, determine which wideband spectrogram corresponds to which word. [8]

- b) Draw Schematic representation of human speech production system and level the following place of articulation [4+4]

I. Velar, II. Alveolar, III. Palatal, IV. Dental

2. Consider a Uniform tube of length l is closed at one end with a volume velocity source $u(0,t)$ as in figure-2

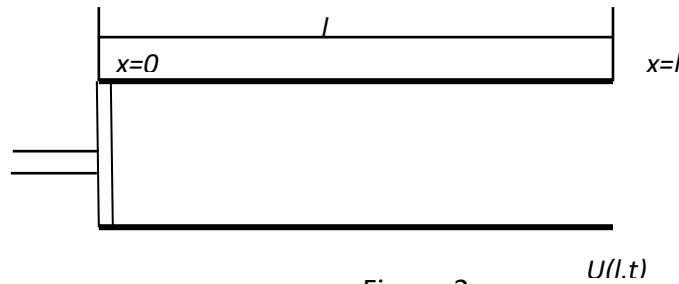


Figure-2

Consider $u(0,t) = U(\Omega)e^{j\Omega t}$ and the velocity and pressure wave solution is as below

$$u(x,t) = u^+(t - x/c) - u^-(t + x/c)$$

$$p(x,t) = \frac{\rho c}{A} [p^+(t - x/c) + p^-(t + x/c)]$$

- (a) Find the volume velocity and pressure along the tube length in terms of source volume velocity [12]
- (b) Plot the volume velocity and pressure envelop along the tube? [4]
3. (a) Derive expression of LPC model gain for a voiced signal? Draw a functional block diagram a LPC decoder. [4+4]
- (b) Deriving the lattice formulation i^{th} order prediction error filter was define as given in equation-1. If the predictor coefficient satisfies the equation-2 show that the backward prediction error can be written as in equation-3 [8]

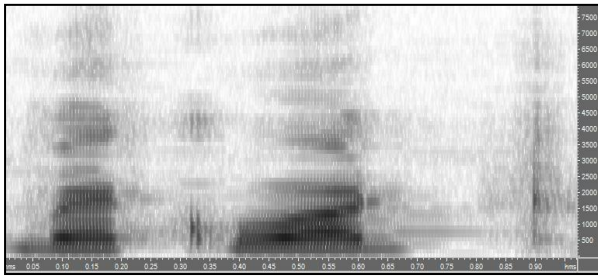
$$A^i(z) = 1 - \sum_{k=1}^i \alpha_k z^{-k} \quad (1)$$

$$\begin{aligned} \alpha_i^i &= k_i \\ \alpha_j^i &= \alpha_j^{i-1} - k_i \alpha_{i-j}^{i-1} \end{aligned} \quad (2)$$

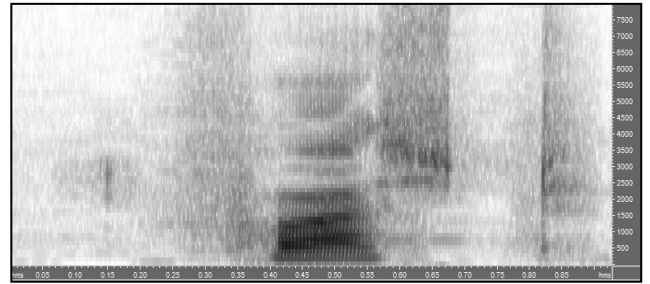
$$b^i[m] = b^{i-1}[m-1] - k_i e^{i-1}[m] \quad (3)$$

4. (a) Draw the block diagram of MFCC parameter extraction method.
- (b) Describe the need of delta and double delta of MFCC parameter in speech application.
- (c) What are the differences between a spoken language and written language and how does it affect automatic speech recognition? [6+4+6]
5. (a) Find the minimum edit- distance between the words 'SPEAKER' and first word of your surname.
- (b) Draw the functional block diagram of speech synthesis system and explain each block.
- (c) Name the signal segment that required for synthesized your first name using Diphone based Concatenative synthesizer. [6+6+4]

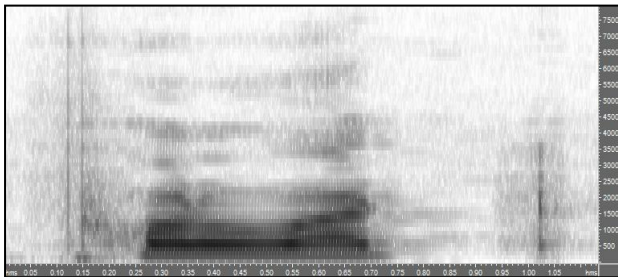
Annexure-1



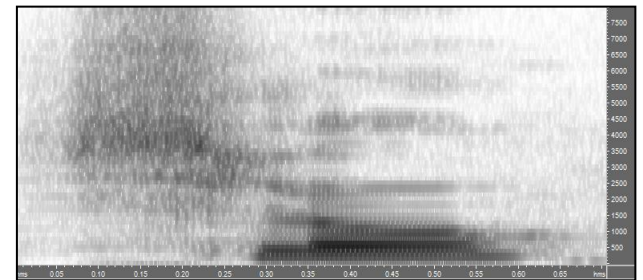
a).....



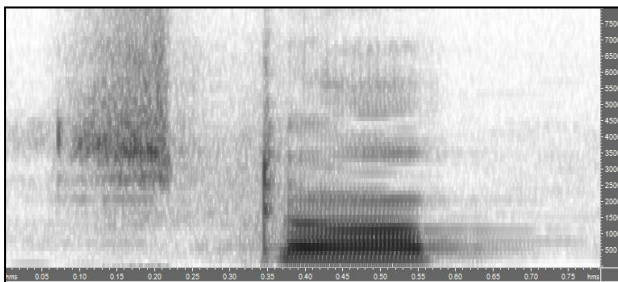
b)



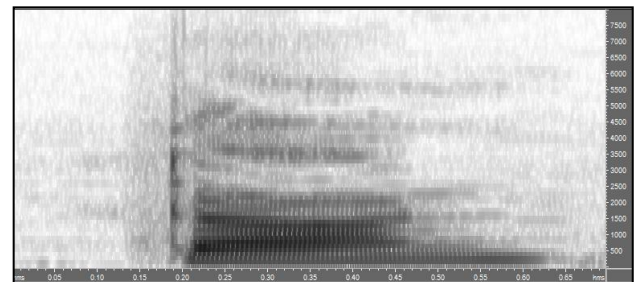
c).....



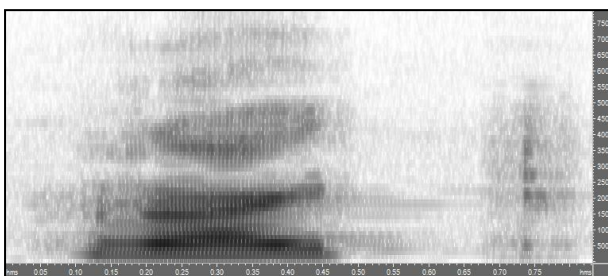
d).....



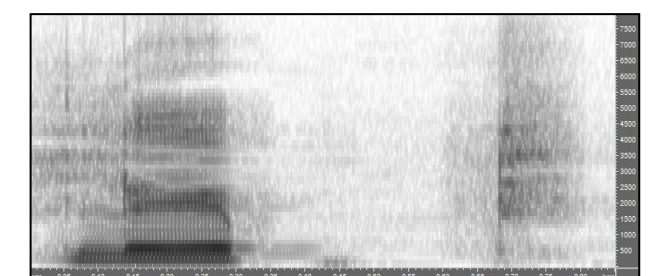
e).....



f).....



g).....



h).....