Indian Institute of Technology, Kharagpur

Centre for Educational Technology

**Mid Semester Examination 2014**

Subject**: INTRODUCTION TO DIGITAL SPEECH PROCESSING** Code: ET60007

**Time: 2:00 Hours** PART-A:-10\*2=20; PART-B:-5\*6=30 **Full Marks =50**

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

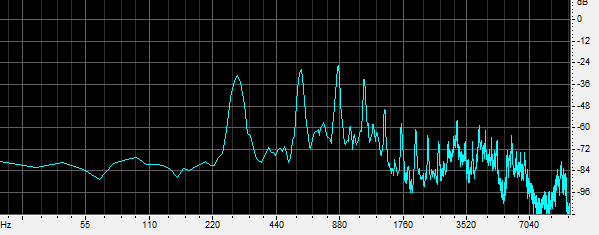
*(Please enclose the Annexure-1 along with the answer script)*

**PART-A**

1. What are the perceptual parameters of speech?
2. Length of a vocal tract is *17.5cm* and it model using single lossless acoustic tube. Derive the first three formant frequency and bandwidth for a voice sound. Where the speed of sound *c=350m/s*.
3. Two source of speech signal producing a vowel /o/. The formant frequencies of the sources are differing by 3-4%. If a human being perceives the two sounds as same sound explains why this happen?
4. *5* sec. speech segment is encoded using LPC coefficient and the LPC coefficient are extracted for each frame (frame length(L)= *5* pitch period) with a frame rate *100 frame/s*. Determine the required order of the LPC analysis and how many frame’s LPC coefficient can be extract from the above speech signal. where the *F0* of the speech segment is *250 Hz* and sampling frequency *Fs=16 kHz*
5. Which of the following pair of tones is perceived as louder tone?
   1. *25dB* level at *500Hz* and *25* db at *800 Hz* (b) *5dB* level at *7 KHz* and *5dB* level at *1 KHz*
6. Suppose an electric fan produces an intensity of *40 dB*. How many times more intense is the sound of a conversation if it produces an intensity of *60 dB*?
7. Draw the glottal flow wave from and define the following
8. Open phase, b) Closed phase and c) Period
9. Describe the state of the Glottis during the pronunciation of the following phoneme?

/k/, /b/, /a/, /l/, /ʃ/

1. Figure-1 represents the frequency analysis of a voice signal segment. Find out the value of *F0* of the voice signal



*Figure-1*

*880 HZ*

1. Write the place and manner of articulation of the following phonemes?

/gh/, /r/, /m/, /ɖ/

**PART-B**

1. (a) Table 1 shows an F1 and F2 of the vowels of a language plot the vowels in F1 and F2 plane in Bark scale mark the axis intern of tongue height tongue position.

Table-1

|  |  |  |
| --- | --- | --- |
| Vowels | F1 [Hz] | F2[Hz] |
| /u/ | 325 | 1035 |
| /o/ | 378 | 1025 |
| /ɔ/ | 543 | 1019 |
| /a/ | 866 | 1530 |
| /æ/ | 591 | 1846 |
| /e/ | 383 | 1978 |
| /i/ | 309 | 2131 |

(b) What is equal loudness curve or phone curve? Draw an equal loudness curve for *10 dB*

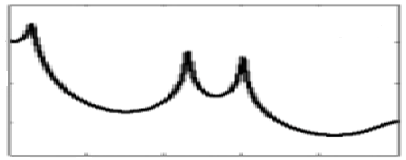
1. (a) An analog signal *xa(t)=6sin(320πt)+4cos(820πt)* is sampled at *600* Hz. Dertermine the F0 of the signal. How many sampled will be present in one fundamenal period?

(b) A discrete signal *x[n]=6.35 cos(πn/5)* is quantized with aresolution *Δ=0.01*. How many bits are required in A/D conversion?

1. Linear prediction analysis is used to obtain a 6-order all-pole model for a segment of voiced speech that was sampled at a rate of FS = 10000 Hz. Figure -1 shows the LPC spectrum of the above extracted LPC coefficient. Determine the root magnitude and root angle of the pole corresponding to first three formant. F1=288Hz , F2=719Hz, F3,=2294Hz BW1=92 Hz, BW2=65Hz, BW3=50Hz

F2

F1



Frequency

BW3

BW2

BW1

F3

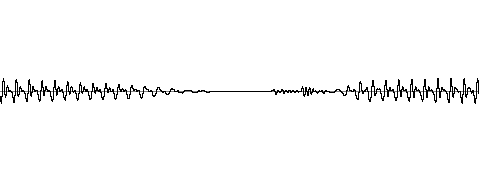
Figure-1

1. (a) Figure-2 (a) and (b) in annexure-1 represents waveform and spectrogram of a VCV speech segment where C represent consonant and V represent Vowel. Mark the occlusion period, burst and VOT part? Write the manner of articulation of the consonant represented by the figure-2.

(b) Figure-3 in annexure-1 represent wideband spectrograms of a speech segment mark the following region your knowledge of acoustic phonetics i) voiced segment ii) unvoiced segment iii) any one consonant to vowel transition segment

1. To produce a voiced speech signal of bandwidth *5 KHz* how many section of lossless tubes are required. Where length of the tube is *17.5cm* and c=35000cm/s. prove that one resonant frequency will be appear in every *1000Hz*. Draw the digital equivalent circuit diagram of the above voiced speech signal production system where *rg* is load at glottisand *rl* is the radiation load.

Annexure-1

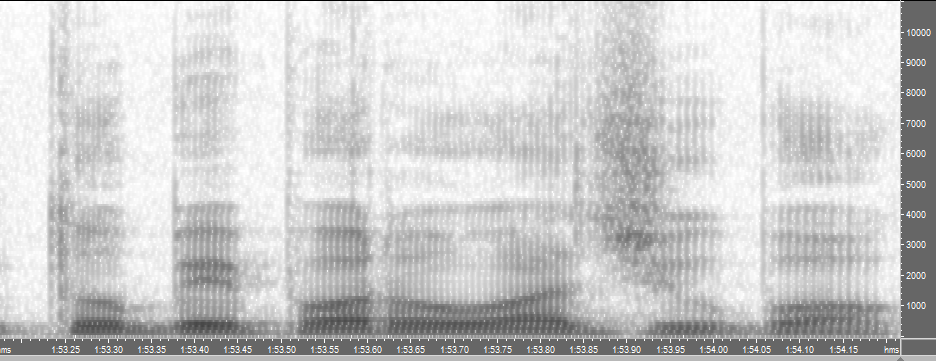


(a)



(b)

*Figure-2*



*Figure-3*