Indian Institute of Technology, Kharagpur

Centre for Educational Technology

**Mid Semester Examination 2018**

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

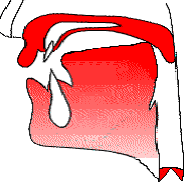
**Time: 2:00 Hours** PART-A:-10\*2=20; PART-B:-5\*6=30 **Full Marks =50**

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

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

**PART-A**

1. Following figure represents the production of a consonant write the manner and place of articulation of the consonant



1.  is the transfer function of a uniform lossless tube (as given in the equation (1)). The length of the tube ***l=17.5 cm*** and speed of sound ***c=350m/s***. Draw the volume velocity vs. Frequency curve for **first** **4 roots**?

(1)



1. Draw the directivity pattern of an unidirectional microphone
2. Table 1 shows an F1 and F2 of the vowels of a language plot the vowels in F1 and F2 plane and indicate the relation of tongue height and tongue position with F1 and F2.

Table-1

|  |  |  |
| --- | --- | --- |
|  | 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 |

1. Determine the F0 of the signal as given in figure-1. The signal is sampled at **16 kHz.**

64 Samples

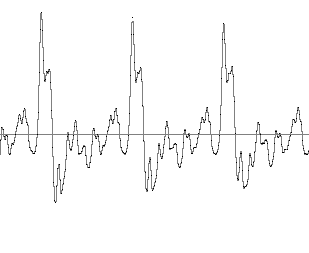


Figure-1

1. Draw the glottal flow waveform and define the following
2. Open phase, b) Closed phase and c) Period
3. An audio signal is recorded using the following format.

***FS = 16 kHz***, encoded with***16 bit*** and recorded in MONO

To store ***100 ms*** signal in PCM WAV format how much memory is required?

1. If figure-2 represent the Frequency response of Uniform tube in no loss condition. Draw the Frequency response of Uniform tube if only radiation loss at the opening is considered

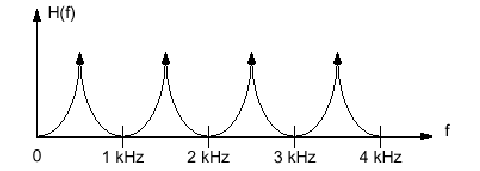


Figure-2

1. If cross section area of a uniform tube is ***A***, and average atmospheric density is ***ρ.*** What will be value of acoustic impedance? Where the speed of sound is ***c***
2. The frequency response of a uniform tube is as given as figure-3. How many complex conjugate poles will be in the tube transfer function?

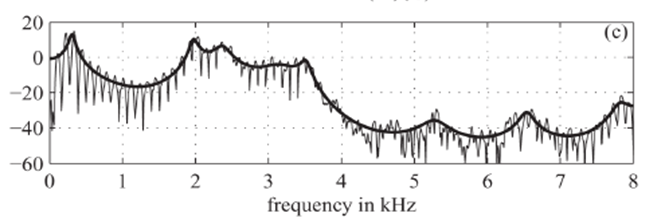


Figure-3

**PART-B**

1. A speech signal is sampled at **16 kHz** and a segment of **1024** samples are selected for frequency analysis using DFT. Length of the DFT analysis is **1024**.
2. Determine the frequency resolution of the frequency analysis
3. Determine the time duration of the selected segment.
4. How many complex multiplications are required to compute the above DFT using FFT algorithm **[2+2+2]**

2. Consider a two tube lossless vocal tract model. Draw the signal flow diagram using reflection coefficient and delay element for the case in which tube cross section and length are as **A1=1 cm2, l1=9 cm and A2 = 7 cm2 and l2= 8 cm**. Include the radiation and glottal boundary condition in the flow diagram. Where velocity of sound **c=340m/s**. **[6]**

3. (a) Describe the state of the Glottis during the pronunciation of the following phonemes?

(i) /g/, (ii) /d/, (iii) /e/, (iv) /ʃ/ **[2]**

(b) Table-2 represent the consonants of a language are grouped based on the manner of articulation. Write the manner of articulation of each of the group. Group all the consonant of table-1 based on place of articulation. **[2+2]**

Table-2

|  |  |
| --- | --- |
| Group | Phoneme in IPA symbols |
| Group-I | /k/, /t/, /ʈ/, /p/ |
| Group-II | /kh/, /th/, /ʈh/, /ph/ |
| Group-III | /g/, /d/, /ɖ/, /b/, |
| Group-IV | /gh/, /dh/, /ɖh/, /bh/ |

4. (a) Figure-4 in annexure-1 is a time domain representation of a speech segment for a language. The phonemes of the language are grouped based on the manner of articulations (as given in table-3) and each of the grouped assigned a pseudo name. Using your knowledge of acoustic phonetics, level the speech segment based on the pseudo name **[4]**

Table-3

|  |  |
| --- | --- |
| Pseudo name | Phonemes group in IPA symbols |
| S | /ʃ/, /s/ |
| P | /k/, /t/, /ʈ/, /p/ |
| F | /kh/, /th/, /ʈh/, /ph/, /ʧ/, /ʧʰ/ |
| A | /g/, /d/, /ɖ/, /b/, /gh/, /dh/, /ɖh/, /bh/, /ʤ/, /ʤh/ |
| L | /l/, /m/, /n/ |
| V | /ɔ/, /ɐ/, /i/, /u/, /e/, /o/, /æ/ |

1. Figure-5 (a) and (b) in annexure-1 represents waveform and spectrogram of a VCV speech segment where C represent consonant and V represent Vowel. Level the occlusion period and VOT part. Write the manner of articulation of the consonant represented by the figure-5. **[1+1]**

5. Figure-6 represents the magnitude spectra of a steady-state vowel segment. The envelope of the spectral magnitude is sketched with a dashed line. The figure showed the four formants considering only the main lobe of the window spectra. The sampling rate of the signal is **16 kHz** meet the Nyquist rate.

(a) Determine the first formant frequency.  **[2]**

(b) If the vowel is produce by a single uniform lossless tube determine the length of the tube. Where velocity of sound **c=340m/s**. **[2]**

(c) Determine the pitch period (F0) of the vowel in milliseconds  **[2]**

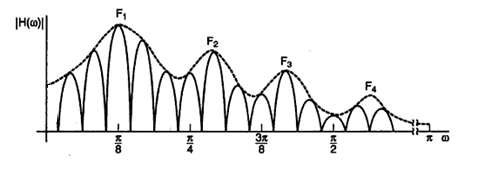
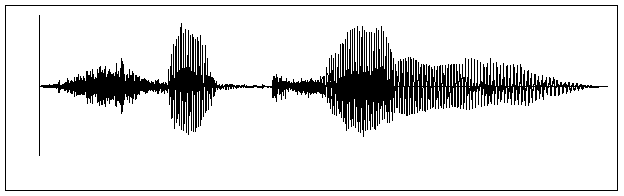


Figure-6

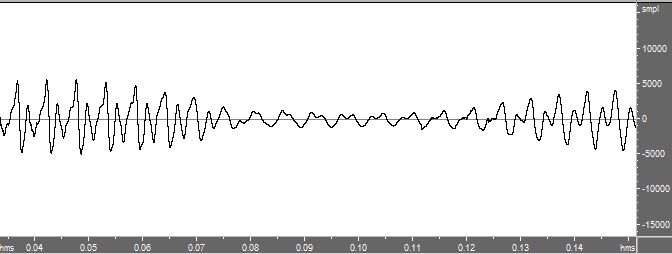
**Annexure-1**



Amplitude

Time

Figure-4



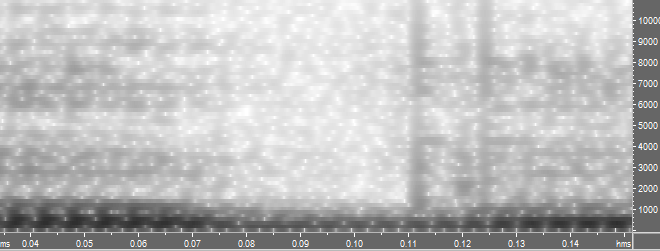


Figure-5