Seat No.: ________ Enrolment No.____________

GUJARAT TECHNOLOGICAL UNIVERSITY

M. E. - SEMESTER – I • EXAMINATION – WINTER 2012

Subject code: 710406N **Date: 16-01-2013**

Subject Name: Speech Processing

Time: 02.30 pm – 05.00 pm **Total Marks: 70**

Instructions:

1. **Attempt all questions.**
2. **Make suitable assumptions wherever necessary.**
3. **Figures to the right indicate full marks.**

Q.1 **(a)** Explain following terms of speech signal. **07**
(1)Voiced-Unvoiced speech (2) Fricatives (3) Diphthongs
(4)Affricates (5) Stops (6)Nasals (7) Whisper

(b) Explain vowel Triangle. **07**

Q.2 **(a)** Explain speech perception model. **07**

(b) Explain lossless tube model for speech production. **07**

OR

(b) A commonly used approximation to the glottal pulse is **07**

$$g[n] = \begin{cases} na^n & n \geq 0 \\ 0 & n < 0 \end{cases}$$

(a) Find the z-transform of g[n].
(b) Sketch the Fourier transform, $G(e^{j\omega})$, as a function of ω.
(c) Show how a should be chosen so that
$20\log10|G(e^{j0})| - 20\log10|G(e^{j\pi})| = 60\text{db}$

Q.3 **(a)** Explain **07**
(1) Pitch (2) Formant (3) cepstrum (4) LPC

(b) Explain speech coding. **07**

OR

Q.3 **(a)** Enlist frequency domain techniques for speech analysis. Explain one of them. **07**

(b) The complex cepstrum, $\hat{x}[n]$, of a sequence x[n] is the inverse Fourier transform of the complex log spectrum **07**
$\hat{X}(e^{j\omega}) = \log|X(e^{j\omega})| + j\arg[X(e^{j\omega})]$
Show that the cepstrum c[n], defined as the inverse Fourier transform of the log magnitude, is the even part of $\hat{x}[n]$; i.e., show that
$$c[n] = \frac{\hat{x}[n] + \hat{x}[-n]}{2}$$

Q.4 **(a)** Explain noise cancellation and echo suppression in speech processing. **07**

(b) Explain basic parameters of speech used for recognition. **07**

OR

Q.4 **(a)** Compare S/N ratios for different types of speech coding. **07**

(b) Consider the design of a high-quality digital audio system. **07**
The specifications are: 60 dB signal-to-noise ratio must be maintained over a range of peak signal levels of 100 to 1. The useful signal bandwidth must be at least 8 kHz.
(a) Draw a block diagram of the basic components needed for A/D and D/A conversion.

(b) How many bits are required in the A/D and D/A converter?
(c) What are the main considerations in choosing the sampling rate? What types of analog filters should be used prior to the A/D converter and following the D/A converter? Estimate the lowest sampling rate that would be possible in a practical system.
(d) How would the specifications and answers change if the objective was only to maintain telephone quality representation of speech?

Q.5	**(a)**	Enlist application of LPC parameters.	**07**
	(b)	Differentiate High level and low level speech synthesis.	**07**

OR

Q.5	**(a)**	Explain Text to speech system.	**07**
	(b)	Enlist application of speech processing and explain one of them.	**07**

****