Reg. No.										

MANIPAL INSTITUTE OF TECHNOLOGY
Manipal University

SIXTH SEMESTER B.TECH (ECE) DEGREE END SEMESTER EXAMINATION - April/May 2017

SUBJECT: DIGITAL SPEECH PROCESSING (ECE -326)

TIME: 3 HOURS **MAX. MARKS: 50**

Instructions to candidates

- Answer **ANY FIVE** full questions.
- Missing data may be suitably assumed.

1A. Starting from the wave equations derive the expression for the transfer function of uniform lossless tube model.

1B. Briefly explain the role of vocal tract in speech production.

1C. Briefly explain the role of vocal cords in speech production.

(5+3+2)

2A. i) With a block diagram explain the pitch period estimation using simple inverse filtering and tracking method.

ii) Explain how formant frequencies can be estimated using LPC parameters.

2B. Assume that STAMF has to be calculated on a speech signal sampled at 12.8KHz using a Hamming window with length of 31.25ms at a rate of 160 times/second. Find the amount of overlapping between two successive frames.

2C. Explain how pitch period can be estimated using short time average magnitude difference function.

(5+3+2)

3A. Consider a sequence $x(n)$ with DTFT $X(e^{jw})$. Let $\tilde{X}(k)$ be the DFT of $x(n)$ obtained by sampling $X(e^{jw})$ at frequencies $w_k = \frac{2\pi k}{N}, k = 0,1,2,..N-1$. If $\tilde{x}(n)$ represent the inverse DFT of $\tilde{X}(k)$, show that $\tilde{x}(n) = \sum_{r=-\infty}^{\infty} x(n - rN)$.

3B. Consider a speech signal sampled at 12KHz. Find the minimum total sampling rate required to obtain a STFT representation in the following cases.

i) If a Rectangular window of length 25ms is used.

ii) If a Hamming window of length 25ms is used.

3C. Let $y[n] = x[n] - x[n-1]$. Show that STFT of the given signal can be represented as

$$Y_n(e^{jw}) = X_n(e^{jw}) - e^{jw}\, X_{n-1}(e^{jw})$$

(5+3+2)

4A. Explain a method of speech vs. silence discrimination using Energy and Zero-crossings.

4B. Briefly explain the autocorrelation method of LPC analysis.

4C. Consider a first order linear predictor whose output is given by $y(n) = \alpha s(n-1)$.
Let $e(n) = s(n) - y(n)$ be the error signal. Derive the expression for the optimum value of α by minimizing the mean squared error.

(5+3+2)

5A. What is a Mel filter bank? Explain how MFCC feature vectors are obtained from speech signal.

5B. Explain briefly the different methods available for evaluating speech quality.

5C. Explain working of feedback adaptive quantiser with time varying gain.

(5+3+2)

6A. Explain how speech signal can be synthesized using LPC synthesizer. Also explain the working of LPC vocoder.

6B. What are the different elements of an HMM?

6C. Consider a signal $x(n) = [-2, -1, -1, 2, 3, 4]$. If $Rn(k)$ represents autocorrelation of $x(n)$, find $Rn(0)$.

(5+3+2)