Reg. No.										

MANIPAL INSTITUTE OF TECHNOLOGY

MANIPAL

(A constituent unit of MAHE, Manipal)

SEVENTH SEMESTER BTECH. (E & C) DEGREE END SEMESTER EXAMINATION
MARCH 2021
SUBJECT: DIGITAL SPEECH PROCESSING (ECE - 4007)

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

Instructions to candidates

- Answer **ALL** questions.
- Missing data may be suitably assumed.

1A. A commonly used approximation to the glottal pulse is, $g[n] = \begin{cases} na^n, & n \geq 0 \\ 0, & n < 0 \end{cases}$

i. Find the z-transform of g[n].

ii. Sketch the Fourier transform $G(e^{j\omega})$, as a function of ω.

iii. Write your comments on part (i) and (ii) above. (iv) The value of 'a' is normally chosen using the criteria, $20\log_{10}\left(|G(\omega)|\right)_{\omega=0} - 20\log_{10}\left(|G(\omega)|\right)_{\omega=\pi} = 20dB$. Using this criterion, find the value of 'a'.

1B. Describe an algorithm for speech versus silence discrimination using short-time parameters of speech.

(5+5)

2A. Consider a vocal fold oscillation in a vocal fry or diplophonic state, where a secondary glottal flow pulse occurs within a glottal cycle. We model this condition over one pitch period as $\tilde{g}[n] = g[n] - \alpha g[n - n_0]$, where n_0 is the delay between primary and secondary pulses. The resulting periodic glottal flow waveform is given by, $u[n] = \sum_{k=-\infty}^{\infty} \tilde{g}[n - kN]$ where N is the pitch period.

i. Determine the spectrum of in terms of spectrum of $\tilde{g}[n]$. Write the Fourier transform of the periodic glottal flow waveform u[n].

ii. Suppose that in a diplophonic state, $n_0 = N/2$.

Describe how the presence of $g[n-n_0]$ affects at the harmonic frequencies, the squared magnitude of U(w). Describe the effect when α changes from 0 to 1.

2B. A certain vocal tract is represented by two resonance frequencies at 1.5kHz and at 2.5kHz. Give an all-pole model representation for vocal tract and show cascade (second order) implementation of the all-pole transfer function. Assume sampling frequency of 8kHz.

(5+5)

3A. Explain the basic principles of linear predictive analysis. How all-pole model for vocal tract can be obtained through LPC? Also show that the total minimum predictor error

consists of a fixed component and a component which depends on the predictor coefficients.

3B. Describe the Fourier transform and linear interpretations of short-time Fourier transform (STFT). Give the necessary block diagrams and mathematical steps.

(5+5)

4A. With the help of block-diagram, describe the simple inverse filtering tracking (SIFT) algorithm for pitch estimation. What are its advantages and disadvantages?

4B. Highlight the main problem with LPC coder. Describe voice excited LPC vocoder for low bit rate coding of speech.

(5+5)

5A. Explain the concept of sub-coding of speech signal. Speech signal is band limited to 3.4kHz and sampled at 8kHz. It is coded with four sub-bands which are 0-425Hz, 425-850Hz, 850-1700Hz and 1700-3400Hz. Four bits are used to encode first two band and two bits are used for the last two bands. Draw the frequency response of the filter bank and estimate the bit rate of the coded speech.

5B. What are the considerations of an unrestricted text-to-speech system? Explain with block diagram, how synthetic speech output can be produced by a general TTS system.

(5+5)