Brunel University

Department of Electronic and Computer Engineering

Final Examination June 2002

EE3052B - MultiMedia Signal Processing

Time allowed 3 Hours

Answer *five* out of eight questions

Ensure that your registration number is written clearly on the front cover.

Q1

(a)

(i) What is MPEG an acronym of?

[1 mark]

(ii) List the different versions of MPEG coders.
What is meant by the term "open standard coder"?

[1 mark]

(iii) State the function and the applications of each version of MPEG coder including the latest version.

[3 marks]

(b)

Draw the detailed block diagram of an MPEG-1 layer 3 (also known as MP3) music encoder. Include typical bit rates for input and output.

[5 marks]

(c)

Describe the following parts of an MP3 coder:

(i) The structure of the filter bank unit: how many filters and what type of filters are used, what are the main advantage of the type of filters employed.

[1 mark]

(ii) The modified discrete cosine transform unit.

[1 mark]

(iii) The Huffman coder unit.

[2 marks]

(iv) The psychoacoustics model.

[2 marks]

(v) The quantisation strategy, the bit allocation pattern and the overall bit rate.

[4 marks]

Q 2

(a)

(i) Write a definition of the term "information" in the context of random variables and information theory (Don't use an equation).

[2 marks]

(ii) Explain why probability is a useful function for measuring or quantifying information. Explain why a function of $\log_2$ of probability is usually employed.

[2 marks]

(iii) Write the equation quantifying the information content of a variable. Draw the curve describing the variation of information content of this variable as a function of its probability.

[4 marks]

(iv) Explain the entropy of a random process, and write down the entropy of an *N*-valued random process.

[4 marks]

(b)

Calculate the entropy of the following sources:

(i) Written English text composed of a combination of 26 letters, 10 numbers and 30 other symbols. Assume all symbols are equi-probable.

[2 marks]

(ii) Spoken English speech composed of 40 phonemic symbols. Assume all symbols are equi-probable.

[2 marks]

(c)

Using the results of part (b) calculate the total number of bits required to encode the following:

(i) A text document composed of 3000 words and 50 other symbols numerical. Assume there are 5 letters per word.

[2 marks]

(ii) Half an hour of speech. Assuming that the average speaking rate is 120 words per minute, and that the average word has 5 phonemes.

[2 marks]

Q3

(a)

Cepstral and dynamical cepstral features are used for speech recognition.

(i) Draw a block diagram for the extraction of FFT-based cepstral coefficients.

[2 marks]

(ii) Write the equation for cepstrum features.

[2 marks]

(iii) Write the equations showing the derivation of the delta and delta-delta cepstrum from the cepstrum. Explain the use of delta and delta-delta cepstrum features.

[2 marks]

(b)

(i) What is the purpose of dynamic time warping (DTW) method.

[1 mark]

(ii) With the aid of an appropriate sketch and the relevant equations explain how DTW works.

[4 marks]

(iii) State two applications of DTW in different areas.

[1 mark]

(c)

Answer the following questions for the design of an automatic voice-dialling system, based on DTW, for a mobile phone handset.

(i) Write the grammar, the vocabulary and the lookup table for name dialling.

[1 mark]

(ii) With the aid of a block diagram, briefly describe the main signal processing sections of a name dialling system.

[3 marks]

(iii) State your choice of the dimensions of the template feature vector and the minimum number of examples of a name required for training a name template.

[1 mark]

(iv) Sketch a flow-chart diagram for the operation of a name-dialling system. Briefly explain the operation of the training unit and the name recognition unit.

[3 marks]

Q4

(a)

The input-output relation of a discrete-time finite impulse response filter is given as

$$y(m) = x(m) \; - 2.5x(m-1) + 5.25x(m-2) - 2.5x(m-3) + x(m-4)$$

(i) Write the z-transfer function of this filter.

[1 mark]

(ii) Show that the filter has a linear phase response,

[3 marks]

(iii) Find and plot the zeros of this filter. Explain the constraints on the position of the zeros of a linear phase filter.

[3 marks]

(b)

The required frequency response of a digital audio digital Bass filter operating at a sampling rate of 40 kHz is given by the expression

$$H(f) = \begin{cases} 1 + 2.1623|f|/400 & |f| \le 400 \\ 1 & |f| > 400 \end{cases}$$

(i) Using the window design technique and the inverse Fourier transform obtain an expression for the discrete-time impulse response of the filter. (You may use the superposition and integration properties to simplify the inverse Fourier transform).

[7 marks]

(ii) Explain how this filter is made finite impulse response and causal.

[2 marks]

(iii) Explain the effects of the choice of filter length on its frequency response

[1 mark]

(iv) Sketch the magnitude frequency response of the filter and calculate its gain in dB at a frequency of 400 Hz.

[3 marks]

Q5

(a)

With the aid of a block diagram describe a source-filter model of speech production and explain how each block of the model relates to the physics of speech production.

[3 marks]

(b)

State the following:

(i) The current standard sampling rate and the compressed speech bit rate, in bits per second, for GSM digital mobile telephony.

(ii) The speech compression ratio of linear coded uncompressed speech relative to compressed speech.

(iii) Three benefits of speech compression for mobile phone applications.

[3 marks]

(c)

Draw the block diagram of a code excited linear prediction (CELP) system

[4 marks]

Describe the following components of a CELP coder, and providing parameter values where appropriate:

(i) The signal segmentation windowing operation,

[2 marks]

(ii) The pitch estimation method,

[2 marks]

(iii) The choice of model orders for the predictor filter and pitch filter.

[2 marks]

(iv) The excitation estimation method.

[4 marks]

Q6

(a) Figure Q6 is a block diagram illustration of a linear prediction model of signal $x(m)$.

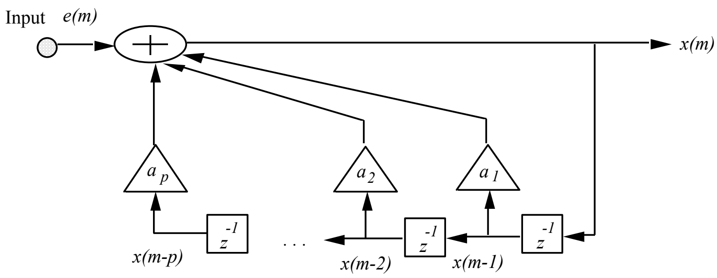

Figure Q6

(i) Describe three applications for a linear prediction (LP) model.

[1 Mark]

(ii) Write the time-domain equation describing the input-output relation of a linear prediction model.

[1 Mark]

(iii) Taking the z-transform of the input-output equation obtain the transfer function of the linear prediction model. Express the transfer function in polar form.

[2 Mark]

(iv) Write the equation for the inverse linear predictor filter.

[1 Mark]

(v) Derive the least squared error solution for the linear predictor coefficients.

[5 Mark]

(b) The first three auto-correlation coefficients of a signal process are as follows

$$r(0) = 1.0,\ r(1) = 0.865,\ r(2) = 0.521.$$

(i) Obtain the coefficients of a second order linear prediction model for the process, and express the coefficients in polar form.

[6 Marks]

(ii) Use the model to write an expression for the frequency response of the process and sketch the spectrum of this predictor.

[2 Marks]

(iii) Sketch a pole-zero diagram showing the position of the poles and the zeros of the predictor and its inverse.

[2 Marks]

Q7

(a)

Using the Fourier transform show the effect of decimation by a factor of L on the spectrum of a discrete-time signal.

[4 Marks]

(b)

With the aid of a block diagram, describe the outline of a system for re-sampling a HiFi digital audio signal $x(m)$ originally sampled at a rate of 44 kHz to a new sampling rate of 10 kHz.

What is the quantitive effect of this resampling process on the spectrum of the signal.

[6 Marks]

(c)

(i) Using the window design technique and the inverse Fourier transform, design lowpass and a highpass digital finite impulse response (FIR) filters to split a total bandwidth of 20 kHz into 2 equal bandwidth sub-bands.

Write the impulse response of each sub-band filter. State how this filter can be made causal.

[6 marks]

(ii) Explain how you can use down sampling and further applications of the high pass and low pass filters to split the signal into four bands.

[4 marks]

Q8

(a)

The discrete Fourier transform (DFT) is given by

$$X(k) = \sum_{m=0}^{N-1} x(m) e^{-j\frac{2\pi}{N}mk} \qquad k = 0, \ldots, N-1$$

The DFT is used in a spectrogram to plot the time-variation of the spectrum of a signal.

(i) Show that the DFT spectrum is periodic along the frequency axis with a period of N.

[5 marks]

(ii) With the aid of a block diagram, briefly state the main signal processing steps in the design of a spectrogram.

[3 Marks]

(iii) A music signal is sampled at 44100 Hz. Assuming that music is relatively stationary for 50 milliseconds what are the best choice of window length and the resulting frequency resolution of the spectrogram?

[3 Marks]

(iv) For a speech signal sampled at 8 kHz, what is the best choice of time window length, and what is the resulting frequency resolution of the spectrogram?
Briefly justify your choice of the window length?

[3 Marks]

(b)

(i) A segment of N samples of a signal is padded with N zeros. Derive the DFT equation for the zero-padded signal. Explain how the frequency resolution of the DFT of a signal changes by zero-padding.

[4 Marks]

(ii) Assuming that a segment of 2000 samples of this signal is zero padded with 2000 extra zeros, calculate the *actual* frequency resolution and the interpolated *apparent* frequency resolution. Assume the signal is sampled at 44100 Hz.

[2 Marks]