Marking Scheme **COMP28512**

<u>Two hours</u>

UNIVERSITY OF MANCHESTER
SCHOOL OF COMPUTER SCIENCE

Mobile Systems (June 2014)

Date: June 2014

Time: **.45 – **.45

Please answer Question ONE and TWO other Questions from the remaining THREE questions provided

This is a CLOSED book examination

The use of electronic calculators is permitted provided they are not programmable and do not store text

[PTO]

1.

a) When looking at the magnitude spectrum obtained from the fast Fourier Transform of N samples of a real valued signal, why do we only need to plot about half the magnitudes?

The FFT gives us N samples of the spectrum from 0 Hz up to the sampling frequency (Fs). Because of the Sampling theorem, the frequency content is limited to the range 0 to Fs/2, and these are represented by frequency samples 0, 1, …, N/2. Therefore we only need to plot these, i.e. N/2 + 1 magnitudes.

b) What is meant by non-uniform quantisation and how is it implemented?

Instead of having a constant step size between quantisation levels, the steps are small for small samples and become progressively larger for larger samples. It is implemented by a compressor followed by a uniform quantiser.

c) What is meant by periodicity, and pseudo-periodicity?

A periodic signal repeats itself exactly at time intervals which are called the period T. Pseudo-periodic waveforms are not exactly periodic but are approximately so in the short term: say over a short segment.

d) What is meant by 'frequency-domain processing'? Give a simple example of what it can be used for.

Frequency-domain processing is done by converting to the frequency-domain using a transform such as the FFT or DCT, applying some modification to the spectral components, and then transforming back to the time domain. It can be used for filtering out an unwanted sine-wave component of a signal.

e) Why is run-length coding used in both MP3 music coders? How could the following sequence of integers be run-length coded?

0 0 8 0 0 0 0 0 4 5 0 0 0 0 1

The MP3 masking process sets to zero frequency domain samples that will be masked and will not be heard. Therefore there will be long strings of zero valued samples.

2, 8, 5, 4, 0, 5, 4, 1

f) Explain the mechanism of '1-persistent carrier sensing multiple access' (CSMA).

Each transmitter continually monitors the channel before deciding whether to transmit.
Each transmitter waits until no carrier signal is detected, then transmits immediately.
Waits for random time after any collision, then starts again.

g) What are the main goals of the '4G IMT-Advanced' standard as proposed by the International Telecommunications Union (ITU)?

Up to 100 Mbit/s for high mobility access
Up to 1 Gbit/s for low mobility/nomadic access

All-IP packet switched network.
Smooth handover across different networks
High spectral efficiency with dynamic sharing of network resources

h) A 2 minute video-clip encoded using MPEG-1 at 1.2 Mbit/s is being downloaded to your mobile phone at 1Mbit/s. What length of buffer would allow you to watch it in real time without interruptions (frame freezing)? How much delay would this cause before the video-clip started to play?

Need to download 120 x 1.2 = 144 Mbit which takes 144 seconds at 1 Mbit/s
Therefore we need 24 seconds of delay with a buffer size of 24x1M = 24 Mbit

i) What is the 'cellular' concept of spatial multiplexing?

City is divided into small areas called cells from 0.1 to 35 km in diameter.
Each cell is given a frequency band, e.g. *f1*, *f2*, *f3* Frequency bands must be different in adjacent cells. Frequency bands are re-used when cells are far away. Users must not transmit 'too loud. 'Seamless hand-over' is required as user moves from cell to cell.

j) A mobile communication system uses a radio channel of bandwidth 6000 Hz. The reception is affected by 'additive white Gaussian noise' (AWGN) whose constant level is such that the signal-to-noise ratio is 40 dB. According to the Shannon-Hartley Law, what is the maximum bit-rate that can be conveyed with arbitrary low bit-error probability over this radio channel? What is the maximum bit-rate that could be achieved over this channel with binary frequency shift keying as used by 2G-GSM telephony?

By Shannon-Hartley Law, channel capacity

$$C = B \log_2(1 + S/N)$$
$$= B \times \log_{10}(1+S/N) / \log_{10}(2)$$
$$\approx 3.32\ B \log_{10}(S/N) \quad \text{assuming } S/N >>1$$
$$\approx 0.332\ B\ SNR \text{ with SNR in dB, assuming } S/N >>1$$

(Many students may just remember this)

$$= 0.332 \times 6000 \times 40$$
$$\approx 80000 \text{ bit/s}$$

With binary fsk signalling, only get 2 bit/s per Hz, i.e. 12000 bit/s.

2. This question is about speech digitisation, coding and multiple access for mobile telephony

(a) What is the bandwidth of 'narrowband' telephone quality speech, and at what frequency is it normally sampled? What bit-rate did 2G-GSM mobile telephony originally use for speech, and why is it not possible to digitise narrowband speech with reasonable quality at this bit-rate using waveform coding. (2 marks)

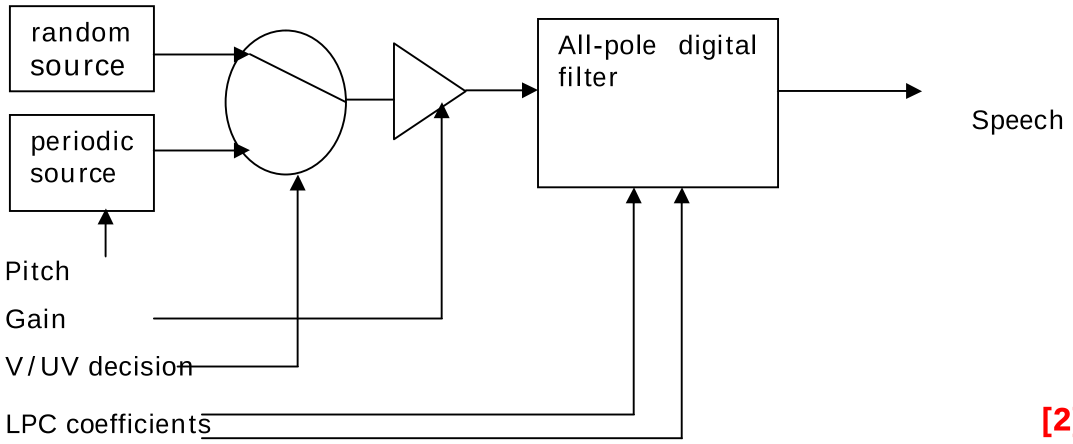
2 (a) The bandwidth of narrowband telephone speech is normally quoted as being 300 Hz to 3400 kHz. It is normally sampled at 8 kHz. **[1]**
To digitise at 13 kbit/s as required for the original 2G-GSM technology would give only about 1.5 bits per sample. We cannot accurately represent a waveform in this way. **[1]**

(b) What is 'voiced speech' and 'unvoiced speech'? (2 marks)

2(b) Voiced speech (vowels) is produced by periodically vibrating vocal cords producing pseudo-periodic sound that is 'coloured' by the vocal tract. **[1]**
Unvoiced speech (consonants) is produced by non-periodic (random) turbulent air-flow created at some constriction within the vocal tract. **[1]**

(c) Explain how the early version of linear predictive coding known as LPC-10 models the human speech production mechanism to achieve bit-rate reduction on a mobile telephone, (8 marks)

2(c) LPC-10 is a parametric coding technique which exploits the characteristics of speech according to a 'source-filter' model of the human speech production mechanism as illustrated below: **[1]**



[2]

At suitable intervals of time (typically about 20ms) the encoder measures and parameterizes the vocal tract resonances as a set of 10, LPC coefficients, **[1]**

It determines whether the speech is voiced or unvoiced (to generate a 1 bit V/UV decision), **[1]**
It measures the speech loudness (to produce a gain control) **[1]**
And, for voiced speech, it determines a fundamental frequency. **[1]**

This data may be transmitted as a low bit-rate representation of a frame of speech which may be re-synthesised by the model above. **[1]**

2(d) What is meant by 'vector-quantisation' as used by 'code-excited LPC' (CELP) coding? (3 marks)

2(d) With CELP, the excitation is a signal segment read from a code-book. **[1]**
An 'analysis by synthesis' approach is adopted to derive the excitation for each frame: Each of the codebook entries is tried until the best one is found. **[1]**
The codebook index is transmitted and an identical codebook is available at the receiver to allow the same excitation signal to be read from it. **[1]**

2(e) By what mechanism is the available spectrum within each cell shared by multiple users in 2G-GSM technology and how has this changed with the introduction of third generation (3G) mobile phones. What is the main advantage and disadvantage of the '3G' mechanism? (5 marks)

2(e) 2G-GSM uses a combination of time-division multiple access (TDMA) and frequency division multiple access (FDMA). **[1]**
It uses FDMA to divide up the available spectrum into 128 up-link bands and 128 down-link bands, and shares each band among 8 users by giving each user a specific time-slot in which to transmit and receive bit-sequences.(TDMA). **[1]**
'3G' uses code-division multiple access which allows all users within a cell to transmit simultaneously, each using a unique 'spread spectrum' code. **[1]**
Main advantage of CDMA: soft limit to the number of users (with greater spectral efficiency). **[1]**
Main disadvantage of CDMA: Power control is difficult due to the 'near-far- problem. **[1]**

3. This question is concerned with bit-error control.

(a) Explain the mechanism of a 'cyclic redundancy check' (CRC) as used for bit-error detection in a mobile system. If a CRC has generator polynomial $G(x) = x^4 + x + 1$, calculate the CRC of the short bit-stream which has already been augmented with '0000':

1 0 1 0 0 1 0 0 0 0.

(6 marks)

3(a): Represent bit-stream by a polynomial, p(x) say, after appending N zeros where N is the order of the 'generator polynomial' g(x) known at transmitter & receiver. Then, using polynomial 'division' in modulo 2 (xor) arithmetic, divide p(x) by g(x) and note the remainder. The remainder, represented by
N-1 bits, is the CRC check. Append this and send. Do same polynomial 'division' at receiver. If remainder is not identical, the CRC fails. **[2]**
Since the generator polynomial is of order 4, append four zeros to 1 0 1 0 0 1 to produce
1 0 1 0 0 1 0 0 0 0 with polynomial $x^9 + x^7 + x^4$

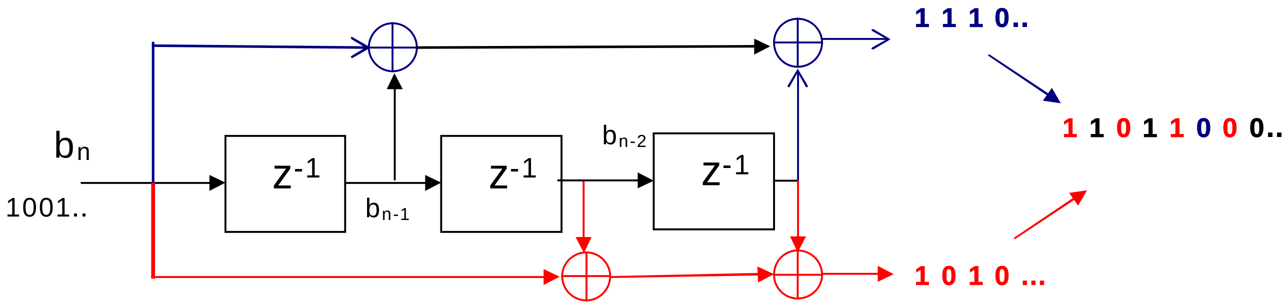
$$\begin{array}{r l l}
 & \underline{x^5 + x^3 + x^2 + x + 1} & \\
x^4 + x + 1 \;) & x^9 + x^7 + x^4 & \\
 & \underline{x^9 + x^6 + x^5} & \oplus \\
 & x^7 + x^6 + x^5 + x^4 & \\
 & \underline{x^7 + x^4 + x^3} & \oplus \\
 & x^6 + x^5 + x^3 & \\
 & \underline{x^6 + x^3 + x^2} & \oplus \\
 & x^5 + x^2 & \\
 & \text{------------------------} & \\
 & x^5 + x^2 + x & \\
 & \text{------------------------} & \\
 & x &
\end{array}$$

[3]

So remainder is x which means that the '4-bit' CRC check is 0010. **[1]**

3(b) What are the essential differences between block codes and convolutional codes for forward error correction (FEC)? If a convolutional coder has two generator functions expressed in octal as (13) and (15), what is the 'rate' of the coder and what is its 'constraint length'? Draw a diagram for the coder and calculate the first 8 bits of its output when the first 4 bits of the input are '1 0 0 1', and the coder starts in zero memory state. What is a 'systematic' coder, and could this particular convolutional coder be described as systematic? (6 marks)

3(b) Block codes require the whole block of data to be available before it can be coded at the transmitter, and similarly the complete block of coded data must have been received before decoding can begin at the receiver. Convolutional coding can start as soon as a few bits have been received and can go on uninterrupted in principle for ever. Similarly, a convolutional decoder can start producing its error-corrected bit-stream once about 50 bits (or so) bits have been received from the channel, and can go on decoding for as long as the transmitter continues to send it data. **[3]**
This is a 'half-rate coder with constraint length 4 : there are three delays as shown:



[2]

Tabulate as follows to calculate output:

b_n	b_{n-1}	b_{n-2}	b_{n-3}	$b_n \oplus b_{n-2} \oplus b_{n-3}$	$b_n \oplus b_{n-1} \oplus b_{n-3}$
1	0	0	0	1	1
0	1	0	0	0	1
0	0	1	0	1	0
1	0	0	1	0	0

Hence encoder output bit-stream is: 1 1 0 1 1 0 0 0 … **[1]**
Although this ordering of the interleaving is the norm, alternative ordering will be accepted

3(c) In principle, how is a convolutionally coded transmission decoded at the receiver, assuming that it may have been affected by bit-errors. (2 marks)

The decoder would normally be a Viterbi decoder which, in principle, generates all possible input sequences and compares the output sequences obtained with the received coded sequence. It selects the one with the minimum Hamming distance to the received sequence. In practice input sequences that, after a few bits are seen to be unlikely to be selected, are discarded to greatly decrease computational complexity. **[2]**

3(d) Explain why bit-error detection and forward error correction (FEC) are used simultaneously at the data-link layer on IEEE802.11 WLAN networks, whereas only error detection is generally used on wired networks**.** (4 marks)

3(d) Error detection with ARQ is effective and economic on wired links which do not introduce many errors. **[1]**
Occasional retransmissions over wire do not require much channel capacity and the extra bit-rate required for introducing a forward error correction capability into every frame may prove expensive in channel capacity and unnecessary most of the time. **[1]**
On radio links, bit-errors occur much more frequently and the many re-transmissions that may be required with error detection and ARQ could be too expensive. **[1]**
The radio channel resources are more precious and limited than wired capacity, and since the physical layer synchronising preamble has to be much longer for radio (about 180 μs for IEEE802.11b WLANs) than for wired (6.4 μs for 10MHz Ethernet), re-transmission packets are more expensive to send by radio. **[1]**

3(e) How does the use of forward error correction (FEC) in cellular mobile telephone systems increase their energy efficiency and also the effectiveness of spatial multiplexing by frequency re-use? (5 marks)

3(e) Transmitting at higher power is an obvious way of making sure a signal is received with fewer errors. But high power signals carry further and cause interference over a wider range, thus making the re-use of frequency bands some distance away more difficult. They also quickly deplete a battery powered transmitter. **[1]**
Reducing the transmission power and dealing the resulting increase in bit-error rate using FEC provides a solution to the cellular frequency re-use problem and also the need to reduce power consumption. **[1]**
---------- ---------

4. This question is about image compression for mobile transmission, the need for Huffman coding and the derivation of a Huffman code.

a) *How is bit-rate compression achieved for images according to the JPEG standard?* (7 marks)

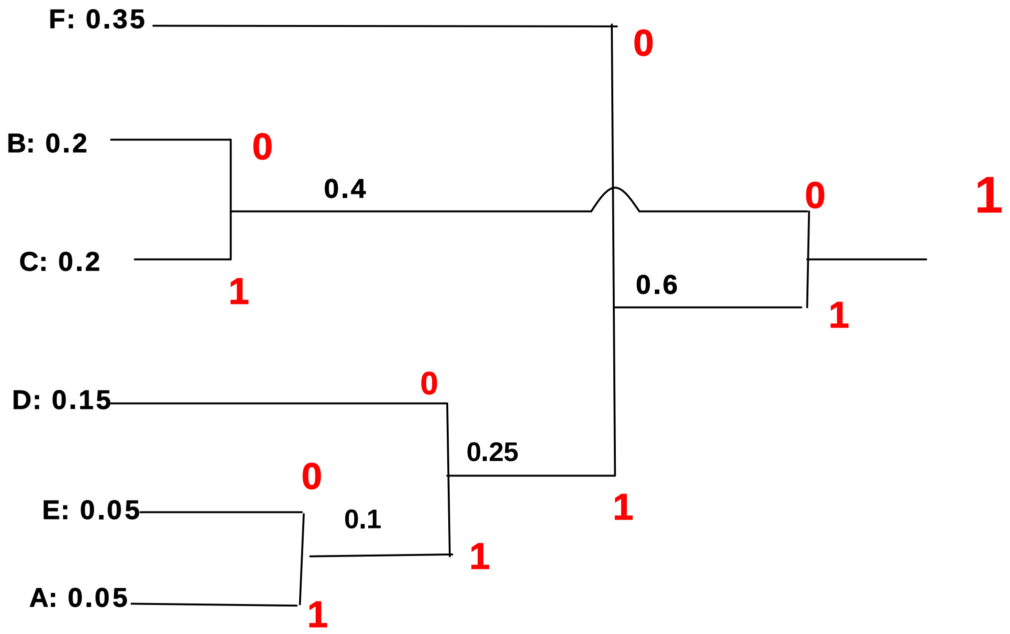
Bit-map, with Red, Green & Blue values for each pixel, split into 8 by 8 tiles. **[1]**
For each tile obtain a value of luminance Y and 2 values of chrominance I and Q for each of the 64 pixels.
Approximately, for each pixel, Y = (R+B+G)/3, I = R – Y ('Red difference' chrominance)
Q = Y – G (Green difference chroma) **[1]**
The 64 values for I are reduced to 16 by dividing the tile into 16 sets of 4 pixels and taking all 4 pixels within each set to have the same value (an average of all 4). Similarly for Q **[1]**
This is done because the eye is less spatially aware of chrominance when luminance is accurately represented. This achieves a degree of bit-rate compression. **[1]**
A 2-D DCT is applied to the Y, I & Q for each tile.
Only the lower 2D frequency terms need be conveyed accurately since the eye is not discerning of the rapidly changing detail represented by the higher 2D frequencies. **[1]**
The lower frequency terms are quantized with more bits than the higher ones.
In fact most of the higher frequency terms become zero when quantized. **[1]**
Therefore we produce a representation of each tile requiring variable length code-words and long runs of zero valued terms. By efficiently encoding this representation, significant bit-rate savings are achievable. **[1]**

b) *Why is run-length coding and Huffman coding required by JPEG?* (4 marks)

Run-length coding is required for efficiently representing the long strings of zero valued terms. **[1]**
Huffman coding is needed to efficiently represent the quantized DCT coeffs
with SELF TERMINATING variable bit-length codewords **[1]**
chosen according to how often each quantization level occurs. **[1]**
Commonly occurring quantization levels are given shorter code-words than less commonly occurring ones. **[1]**

c) Symbols A,B,C,D E,F have probabilities:
0.05, 0.2, 0.2, 0.15, 0.05, 0.35
Devise a Huffman code and consider how it would be decoded. (7 marks)

Firstly the symbols need to be arranged in decreasing order of probability as below. **[1]**
Combine the least probable pair (A and E)**, t**hen rearrange order if necessary (can do this visually without redrawing) and continue as in diagram.



Self terminating codes are:
A: 1 1 1 1
B: 0 0
C: 0 1
D: 1 1 0
E: 1 1 1 0
F: 1 0

NB: Alternative solutions exist

For continuing correct method **[3]**
For correct answer **[2]**
Note that reordering has been necessary and it has been done visually.
Decoding is done by looking for the shortest codes first: i.e. 00, 01, 10. If these are not found, look for next shortest ones, i.e. 110 for D. If this is not found go on to next longest code-words, and so on. **[1]**

4 (d) Why would you expect a JPEG compressed image to be more sensitive to the effect of bit-errors than an uncompressed image such as a bit-map? (2 marks)

If a bit-error occurs, not only will the variable length code-word containing the bit be affected, but, quite possibly, all following variable length codewords may be misinterpreted. The self-terminating properties may be lost. In this compressed form, each bit is carrying more information than in an uncompressed representation. **[2]**

End of answers

Not asked:
Estimate the expected bit-saving obtained by using Huffman coding in comparison to using equal length codes.
Assume we have a large number (N) of samples
For equal length, we need 3 bits per symbol, therefore 3xN bits
For Huffman coding, we get 0.75xN samples with 2 bits (1.5 x N bits)
0.15x N samples with 3 bits (0.45 x N bits)
0.1 x N samples samples with 4 bits. (0.4 x N bits)
Total = 2.35 x N bits. A saving of about 22 %

Short questions not used:

What is meant by 'bandwidth efficiency' with respect to the transmission of digital information over radio channels?
Ans: The number of bits per second that can be transmitted per Hz of bandwidth.

Explain the terms 'frequency masking' and 'temporal masking' as used in MP3 compression.
Definition of frequency masking: 'A strong tonal audio signal at a given frequency will mask, i.e. render inaudible, quieter tones at nearby frequencies, above and below that of the strong tone, the closer the frequency the more effective the matching'.
Definition of temporal masking: 'A loud sound will mask i.e. render inaudible a quieter sound occurring shortly before or shortly after it. The time difference depends on the amplitude difference'.

k) State the 'Sampling Theorem' or 'Nyquist Sampling Criterion' for the regular sampling of an analogue signal at some frequency Fs Hz. Explain what happens when the constraints of the theorem or criterion are not satisfied.
If the sampling frequency is more than twice the maximum frequency component of the signal being sampled, no information is lost about the signal. If we know the sampled values exactly, we can always recreate the signal exactly from these samples. If the sampling Theorem is not obeyed, aliasing occurs, changing the frequencies of components above half the sampling frequency.

Would be a nice exercise to program the Huffman decoding efficiently (next year maybe)