CARDIFF CARDIFF UNIVERSITY
EXAMINATION PAPER

Academic Year:	2002-2003
Examination Period:	Autumn 2002
Examination Paper Number:	CM0340
Examination Paper Title:	Multimedia
Duration:	2 hours

Do not turn this page over until instructed to do so by the Senior Invigilator.

Structure of Examination Paper:
There are four pages.
There are four questions in total.

There are no appendices.

The maximum mark for the examination paper is 100% and the mark obtainable for a question or part of a question is shown in brackets alongside the question.

Students to be provided with:
The following items of stationery are to be provided:
One answer book.

Instructions to Students:
Answer THREE questions.

The use of translation dictionaries between English or Welsh and a foreign language bearing an appropriate departmental stamp is permitted in this examination.

1. (a) What is MIDI? [2]

   (b) How is a basic MIDI message structured? [4]

   (c) A piece of music that lasts 3 minutes is to be transmitted over a network. The piece of music has 4 constituent instruments: Drums, Bass, Piano and Trumpet. The music has been recorded at CD quality (44.1 Khz, 16 bit, Stereo) and also as MIDI information, where on average the drums play 180 notes per minute, the Bass 140 notes per minute, the Piano 600 notes per minute and the trumpet 80 notes per minute.

      (i) Estimate the number of bytes required for the storage of a full performance at CD quality audio and the number of bytes for the Midi performance. You should assume that the general midi set of instruments is available for any performance of the recorded MIDI data. [8]
      (ii) Estimate the time it would take to transmit each performance over a network with 64 kbps. [2]
      (iii) Briefly comment on the merits and drawbacks of each method of transmission of the performance. [4]

   (d) Suppose vocals (where actual lyrics were to be sung) were required to be added to the each performance in (c) above. How might each performance be broadcast over a network? [7]

2. (a) What is meant by the terms *frequency* and *temporal masking* of two or more audio signals? Briefly, what is the cause of this masking? [8]

   (b) How does MPEG audio compression exploit such phenomena? Give a schematic diagram of the MPEG audio perceptual encoder. [8]

   (c) The *critical bandwidth* for average human hearing is a constant 100Hz for frequencies less than 500Hz and increases (approximately) linearly by 100 Hz for each additional 500Hz.

      (i) Given a frequency of 300 Hz, what is the next highest (integer) frequency signal that is distinguishable by the human ear assuming the latter signal is of a substantially lower amplitude? [4]
      (ii) Given a frequency of 5000 Hz, what is the next highest (integer) frequency signal that is distinguishable by the human ear assuming the latter signal is of a substantially lower amplitude? [7]

3. (a) What is the main difference between the *H.261* and *MPEG* video compression algorithms? [2]

(b) MPEG has a variety of different standards, *i.e.* MPEG-1, MPEG-2, MPEG-4, MPEG-7 and MPEG-21. Why have such standards evolved? Give an example target application for each variant of the MPEG standard. [8]

(c) Given the following two frames of an input video show how MPEG would estimate the motion of the macroblock, highlighted in the first image, to the next frame.

1	1	1	1	1	1	1	1
1	1	2	3	3	2	1	1
1	1	2	2	2	2	1	1
1	1	2	**4**	**5**	2	1	1
1	1	2	**5**	**3**	2	1	1
1	1	2	3	3	2	1	1
1	1	1	3	3	2	1	1
1	1	1	3	3	1	1	1

Frame *n*

1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	2	1	2	2	2	2
1	1	2	1	4	3	3	2
1	1	2	1	4	3	4	3
1	1	2	1	4	4	5	4
1	1	2	1	4	5	4	5
1	1	2	1	2	4	4	4

Frame *n+1*

For ease of computation in your solution: you may assume that all macroblock calculations may be performed over 4x4 windows. You may also restrict your search to ± 2 pixels in horizontal and vertical direction around the original macroblock. [12]

(d) Based upon the motion estimation a decision is made on whether INTRA or INTER coding is made. What is the decision based for the coding of the macroblock motion in (c)? [5]

4. (a) What is the distinction between *lossy* and *lossless* data compression? [2]

(b) Briefly describe the four basic types of data redundancy that data compression algorithms can apply to audio, image and video signals. [8]

(c) Encode the following steam of characters using **decimal** *arithmetic coding* compression:

MEDIA

You may assume that characters occur with probabilities of M = 0.1, E = 0.3, D = 0.3, I = 0.2 and A = 0.1. [12]

(d) Show how your solution to (c) would be decoded. [5]