CARDIFF UNIVERSITY
EXAMINATION PAPER

SOLUTIONS

Academic Year:	2003-2004
Examination Period:	Autumn 2003
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 does *Nyquist's Sampling Theorem* state?

In order to effectively sample a waveform the sampling frequency must be **at least** twice that of the highest frequency present in the signal

2 Marks --- Bookwork

(b) What are the implications of Nyquist's Sampling Theorem for multimedia data?

Sampling frequency affects the quality of the data ---- higher frequency equals better sampling hence representation of the underlying signal (given fixed frequency range of signal)

Sampling frequency affects size of digitized data --- higher frequency means more samples therefore more data.

4 (2 marks each above) Marks --- Bookwork

(c) *For each of the following media types, graphics, images, audio and video, briefly discuss how Nyquist's Sampling Theorem affects the quality of the data and the form in which sampling effects manifest themselves in the actual data.*

Graphics

- Quality: Not an issue with vector graphics
- Sampling Artifact: Rendering may lead to Aliasing effect in lines etc

Images

- Quality: Image size decreases so less detail or sampling artifacts
- Sampling Artifact: Aliasing effect in blocky images

Audio

- Quality: Lack of clarity in high frequencies, telephonic voices at low sampling frequencies
- Sampling Artifact: Digital noise present in signal, loss of high frequencies or poor representation of high frequencies give audio aliasing (should be filtered out before sampling)

Video

- Quality: Video Frame size decreases so less detail or sampling artifacts, motion blur or loss of motion detail
- Sampling Artifact: Aliasing effect in frame images, jittery motion tracking etc.

12 (3 Marks per media type) Marks --- Unseen: Extended reasoning on a few parts of course

(d) Calculate the uncompressed digital output, i.e. data rate, if a video signal is sampled using the following values:

25 frames per second
160 x 120 pixels
True (Full) colour depth

True color = 24 bits (3 bytes) per pixel

So **number of bytes per second is**

3*160*120*25 = 144000 bytes or 1.37 Mb

3 Marks --- Unseen: Application of basic knowledge

(e) If a suitable CD stereo quality audio signal is included with the video signal in part d what compression ratio would be needed to be able to transmit the signal on a 128 kbps channel?

Stereo audio = 44100*2 (16 bit/s byte)*2 = 176400 bytes per second

So uncompressed bytes stream is 144000 + 176400 = 320400 bytes per second

128 kbps is kilo **bits** per second

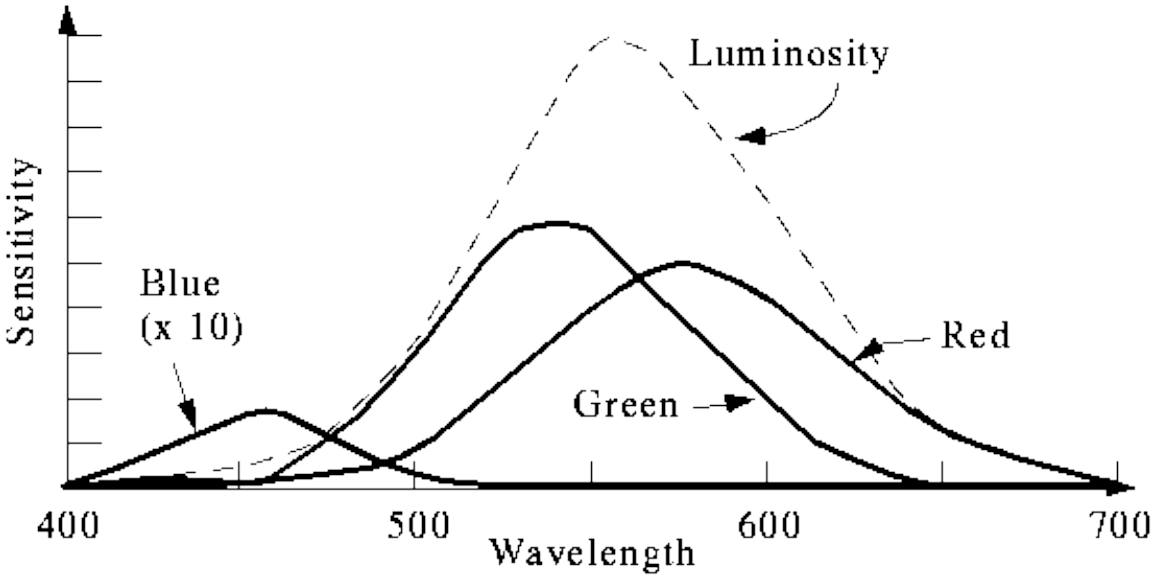
so compression ratio is (128*1024)/(320400 *8) = 0.05.

3 Marks --- Unseen: Application of basic knowledge

2. (a) *What characteristics of the human visual system can be exploited for the compression of colour images and video?*

The eye is basically sensitive to colour intensity

- Each neuron is either a *rod* or a *cone* . Rods are not sensitive to colour.
- Cones come in 3 types: red, green and blue.
- Each responds differently --- Non linearly and not equally for RGB differently to various frequencies of light.



5 Marks --- Bookwork

(b) What is the YIQ color model and why is this an appropriate color model used in conjunction with compression methods such as JPEG and MPEG?

- YIQ is origins in colour TV broadcasting
- Y (luminance) is the CIE Y primary.
- $Y = 0.299R + 0.587G + 0.114B$
- the other two vectors:
- $I = 0.596R - 0.275G - 0.321B \qquad Q = 0.212R - 0.528G + 0.311B$
- The YIQ transform:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.528 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

How to exploit to compression:

- Eye is most sensitive to Y, next to I, next to Q.
- Quantise with more bits for Y than I or Q.

4 (2 for Transform (Matrix or Eqn) and 2 for Compression scheme) Marks --- Bookwork

(c) Given the following YIQ image values:

128	*126*	*127*	*129*
124	*123*	*124*	*124*
130	*136*	*132*	*132*
154	*143*	*132*	*132*

Y

55	*66*	*54*	*54*
56	*57*	*56*	*56*
45	*56*	*58*	*49*
34	*36*	*39*	*37*

I

44	*44*	*55*	*55*
44	*44*	*55*	*55*
34	*34*	*36*	*35*
35	*35*	*34*	*34*

Q

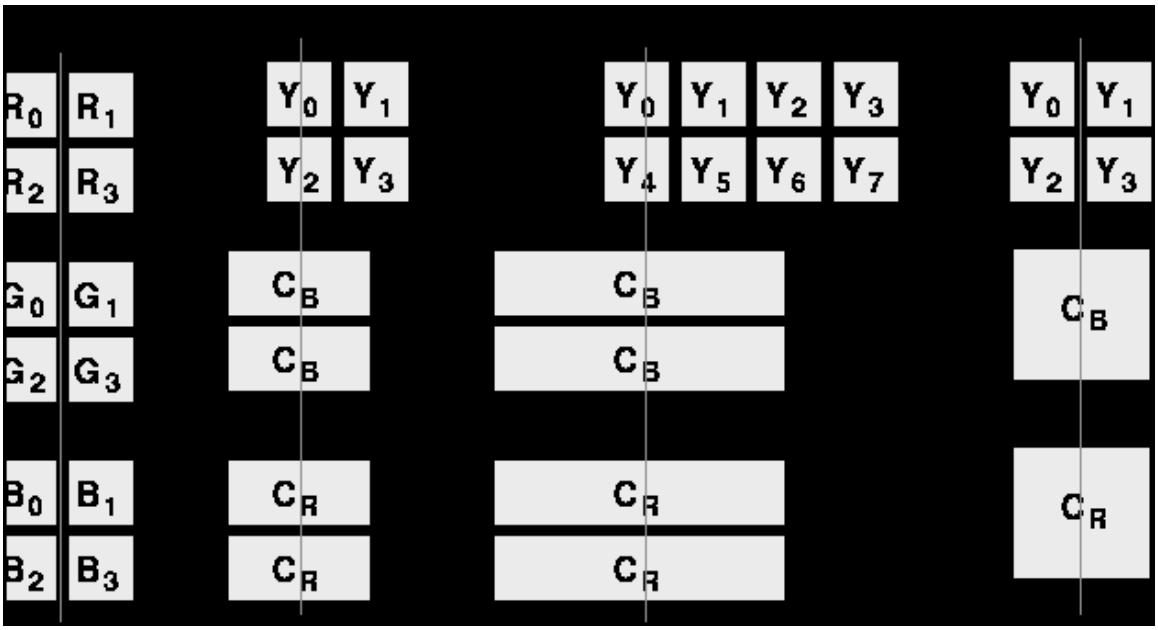
What are the corresponding chroma subsampled values for a

(i) 4:2:2 subsampling scheme
(ii) 4:1:1 subsampling scheme
(iii) 4:2:0 subsampling scheme

Basic Idea required (from notes):

Chroma Subsampling

- 4:2:2 -> Horizontally subsampled colour signals by a factor of 2. Each pixel is two bytes, e.g., (Cb0, Y0)(Cr0, Y1)(Cb2, Y2)(Cr2, Y3)(Cb4, Y4) ...
- 4:1:1 -> Horizontally subsampled by a factor of 4
- 4:2:0 -> Subsampled in both the horizontal and vertical axes by a factor of 2 between pixels



(i) 4:2:2 subsampling scheme

Take every two horizontal pixels in I Q Space

128	126	127	129
124	123	124	124
130	136	132	132
154	143	132	132

Full YIQ

55	66	54	54
56	57	56	56
45	56	58	49
34	36	39	37

44	44	55	55
44	44	55	55
34	34	36	35
35	35	34	34

128	126	127	129
124	123	124	124
130	136	132	132
154	143	132	132

4:2:2
YIQ

55	54
56	56
45	58
34	39

44	55
44	55
34	36
35	34

(ii) 4:1:1 subsampling scheme Take every 4 pixels in the horizonatl

128	126	127	129
124	123	124	124
130	136	132	132
154	143	132	132

55	66	54	54
56	57	56	56
45	56	58	49
34	36	39	37

44	44	55	55
44	44	55	55
34	34	36	35
35	35	34	34

Full YIQ

128	126	127	129
124	123	124	124
130	136	132	132
154	143	132	132

55
56
45
34

44
44
34
35

4:1:1
YIQ

(iii) 4:2:0 subsampling scheme   AVERAGE in every 2x2 block

128	1126	1127	129
124	1123	1124	124
130	1136	132	132
154	1143	132	132

55	66	54	54
56	57	56	46
45	56	58	49
34	36	39	37

44	44	55	55
44	44	55	55
34	34	36	35
35	35	34	34

Full
YIQ

128	126	127	129
124	123	124	124
130	136	132	132
154	143	132	132

59	55
43	46

44	55
35	35

4:2:0
YIQ

15 Marks --- Unseen: Practical Application of Bookwork Knowledge

3. (a) What is the distinction between lossy and lossless data compression?

Lossless Compression

Where data is compressed and can be reconstituted (uncompressed) without loss of detail or information. These are referred to as bit-preserving or reversible compression systems also.

Lossy Compression

where the aim is to obtain the best possible *fidelity* for a given bit-rate or minimizing the bit-rate to achieve a given fidelity measure. Video and audio compression techniques are most suited to this form of compression.

2 Marks Bookwork

(b) Briefly describe two repetitive suppression algorithms and give one practical use of each algorithm.

1. **Simple Repetition Suppresion**

If in a sequence a series on *n* successive tokens appears we can replace these with a token and a count number of occurences. We usually need to have a special ***flag*** to denote when the repated token appears

For Example

894000000000000000000000000000000

we can replace with

894f32

where fis the flag for zero.

Compression savings depend on the content of the data.

Applications of this simple compression technique include:

- Silence in audio data, Pauses in conversation ***etc.***
- Bitmaps
- Blanks in text or program source files
- Backgrounds in images
- other regular image or data tokens

2. Run-length Encoding

In this method, sequences of (image) elements X1, X2 … Xn are mapped to pairs (C1,l1), (C2,L2) …… (Cn,Ln) where *ci* represent image intensity or colour and *li* the length of the *i*th run of pixels.

For example:

Original Sequence:

111122233333311112222

can be encoded as:

(1,4),(2,3),(3,6),(1,4),(2,4)

The savings are dependent on the data. In the worst case (Random Noise) encoding is more heavy than original file: 2*integer rather 1* integer if data is represented as integers.

Applications:

- This encoding method is frequently applied to images (or pixels in a scan line).
- It is a small compression component used in JPEG compression

10 Marks Bookwork

(c) *Briefly state the LZW compression algorithm and show how you would use it to encode the following stream of characters*:

MYMEMYMO

You may assume that single character tokens are coded by their ASCII codes, as per the original LZW algorithm. However, for the purpose of the solution you may simply output the character rather than the ASCI value.

The LZW Compression Algorithm can summarised as follows:

```
w = NIL;
   while ( read a character k )
       {
         if wk exists in the dictionary
          w = wk;
         else
           add wk to the dictionary;
           output the code for w;
           w = k;
       }
```

Original LZW used dictionary with 4K entries, first 256 (0-255) are ASCII codes.

Encoding of *MYMEMYMO:*

W	K	Output	Index	Symbol
nil	M			
M	Y	'M' (ASCII)	256	MY
Y	M	'Y'	257	YM
M	E	'M'	258	ME
E	M	'E'	259	EM
M	Y			
MY	M	256	260	MYM
M	O	'M'	261	MO

So Token Stream is

MYME<256>M

12 Marks Unseen: Application of Algorithm

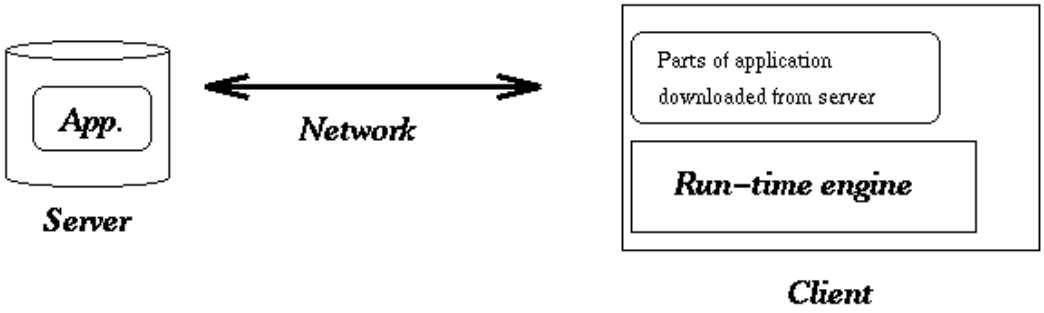
4. (a) What is the basic format of an MHEG application?

An MHEG-5 application is made up of Scenes and objects that are common to all Scenes.

- A Scene contains a group of objects, called **Ingredients**, that represent information (graphics, sound, video, etc.) along with localized behavior based on events firing (e.g., the 'Left' button being pushed activating a sound).
- At most one Scene is active at any one time.
- Navigation in an application is done by transitioning between Scenes.

2 Marks Bookwork

(b) Briefly describe the MHEG Client-Server interaction and the role that the MHEG Engine plays in this process.



Client Server Interaction **(4 Marks)**

- Server streams out content requested by MHEG application.
- Client/Run Time Engine (RTE) embedded in firmware process MHEG, Deals with streaming of sourced data and formatting for presentation

Run Time Engine (RTE) main functions: **(6 Marks)**

- RTE is the kernel of the client's architecture.
- Issues I/O and data access and requests to client
- Prepares the presentation and handles accessing, decoding,
- and managing MHEG-5 objects in their internal format.
  - Interpretation of MHEG objects
- Actual presentation, which is based on an event loop where events trigger actions.
- These actions then become requests to the Presentation layer along with other actions that internally affect the engine.

10 Marks Bookwork

(c) Using suitable fragments of MHEG code, illustrate how you would code the following presentation in MHEG:

Scene 1

Scene 2

*The above presentation consists of two scenes. **Scene 1** plays some Video and is overlayed by some text information and a next button is provided so that the user may elect to jump to **Scene 2**. **Scene 2** plays some video and is overlayed by a visual prompt which when selected displays some further text information.*

***Note** that the precise MHEG syntax and object attributes and attribute values is not absolutely required in your solutions. Rather you should concentrate on giving indicative object attributes values. In essence the structure of the code is more important than precise syntax.*

Basic Idea

Need startup of application --- here do it in **startup.mheg**
This fires up **scene1.mheg** --- only essential MHEG objects lists. **Button1** event triggers **scence2.mheg.** *Important Point is button and link transition*
scene2.mheg --- fires up on **button1** event fires up **moreinfo.mheg** on visual prompt
event trigger. Only essential MHEG objects listed. *Important point is some graphics/bitmap overlay icon plus hot spot for link*
moreinfo.mheg --- simply full of text. Not that important transition to is what Q. requires

startup.mheg (3 Marks):

```
{:Application  ("startup" 0)
  //:OnStartUp    (:TransitionTo(("scene1.mheg" 0)))
  :Items(
          {:Link 1
           :EventSource 1
           :EventType IsRunning
           :LinkEffect (
            :TransitionTo(("scene1.mheg" 0)))
          }
        )
  }
```

scene1.mheg (Object attribute and attribute values just indicative) 5Marks:

```
{:Scene ( "scene1.mheg" 0 )

 :Items
 (
   {:video 0

   :InitiallyActive true
   :OrigContent
   :ContentRef ( "waterfall.mov" )
   :OrigBoxSize 120 120
   :OrigPosition 225 175
   :ComponentTag 100
   :Termination loop
    }

 {:Text 1
    :OrigContent  'Some Text ……'
    :OrigBoxSize 95 95
    :OrigPosition 0 175
    :FontAttributes Bold.14
    :TextColour black
    :HJustification centre
    :VJustification centre
    :TextWrapping true
     }
  {:PushButton 3
    :OrigBoxSize   100  60
    :OrigPosition  540 280
    :ButtonRefColour gray
    :OrigLabel "back to main"
   }

   {:Link 4
    :EventSource 2 :EventType IsSelected
    :LinkEffect (  :TransitionTo( ( "scene2.mhg" 0) )
)
     }
  }
  )
```

scene2.mheg (Object attribute and attribute values just indicative) 5Marks:

```
{:Scene ( "scene2.mheg" 0 )

 :Items
 (
   {:video 0

   :InitiallyActive true
   :OrigContent
   :ContentRef ( "painter.mov" )
   :OrigBoxSize 120 120
   :OrigPosition 225 175
   :ComponentTag 100
   :Termination loop
    }

 {:Bitmap 1
    :OrigContent :ContentRef ( "overlay.gif" )
              :OrigBoxSize 51 39     // 0 0
              :OrigPosition 10 15
              :Tiling false
              }

  {:Hot Spot 2
    :OrigBoxSize   100  60
    :OrigPosition  540 280
    :ButtonRefColour gray
    :OrigLabel "back to main"
   }

   {:Link 3
    :EventSource 2 :EventType IsSelected
    :LinkEffect (  :TransitionTo( ( "moreinfo.mhg" 0)
) )
     }
  }
  )
```

moreinfo.mheg (2 Marks)

```
{:Scene ( "moreinfo.mheg" 0 )

 :Items
 (


 {:Text 1
    :OrigContent  'Some Text ……'
    :OrigBoxSize 95 95
    :OrigPosition 0 175
    :FontAttributes Bold.14
    :TextColour black
    :HJustification centre
    :VJustification centre
    :TextWrapping true
     }
 )
 }
```

15 Marks Unseen