

**CARDIFF UNIVERSITY
EXAMINATION PAPER**

SOLUTIONS

Academic Year:	2003-2004
Examination Period:	Postgraduate
Examination Paper Number:	CMP632
Examination Paper Title:	Multimedia Systems
Duration:	2 hours

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

Structure of Examination Paper:

There are **THREE** 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) *Give a definition of a Multimedia Authoring System.*

An Multimedia Authoring System is a program which has pre-programmed elements for the development of interactive multimedia software titles. Authoring systems vary widely in orientation, capabilities, and learning curve.

2 Marks --- Bookwork

- (b) *Briefly describe **five** ways in which content can be formatted and delivered in a Multimedia Authoring System.*

1. Scripting (writing)
Standard Text --- say what you want with word
2. Graphics (illustrating)
“A picture is worth a thousand words” say what you want with a graphic illustrations
3. Animation (wiggling)
Now we approach multimedia --- say what you want with a graphic animation or video
4. Audio (hearing)
Sounds can convey alerts, ambience and contents say what you want with a narration
5. Interactivity (interacting)
True mulitedia immerse yourself in am interactive presentation, possibly more instructive. Interactive actions can start animations, audio, move to new parts of presentation, control simulations etc.

10 Marks --- Bookwork (2 Marks per point)

(c) *What extra information is multimedia good at conveying with respect to conventional media.*

(i) *What can spoken text convey that written text cannot?*

Spoken Text can convey:

Emotion more readily

Or Feelings more readily

Unwritten Sounds to express feelings or emotions example “tut tutting”, sharp intakes of breath

Accents/dialects readily apparent

Important when more than speaker is present

Perhaps in getting certain messages across

E.g. assimilating a radio play is easier when we can distinguish speakers more easily

If in stereo position in 3D sound space is discernable

Possibly useful in creating a feeling of space or locating people in immersive 3D environments

Easier to author/synchronise with Video

Since Audio is already time dependent media, aligning media type on a timeline in a multimedia authoring package is usually simple

Written text will need to be “animated” e.g. rollover credits, subtitles --- not difficult but some additional multimedia editing of raw text is required

(ii) *When might written text be better than spoken text?*

Hard to assimilate/remember a lot of spoken word

Written text easier to reread (as opposed to replay)

Additional information, such as

what a person looks like, what he is wearing, general appearance can be possibly more easily conveyed

Directions as to what else is happening, e.g a person moving around, waving arms more easily conveyed

e.g. Comparison a (radio) screenplay has plenty of stage directions

Written text has negligible bandwidth, high quality audio has significant bandwidth requirements

12 Marks Unseen

2. (a) *Briefly explain how the human visual system senses colour. How is colour exploited in the compression of multimedia graphics, images and video?*

- The eye is basically just a “biological camera”
- Eye through lens etc focused light onto the Retina (back of eye)
- Retina consists of neurons that fire nerve signals proportional to incident light
- Each neuron is either a rod or a cone. Rods are not sensitive to colour.
- Cones organized in banks that sense red green and blue

Multimedia Context:

Since Rods do not sense colour only sense luminosity intensity Eye is more sensitive to luminance than colour. Also Eye is more sensitive to red and green and blue (this is due to evolution need to see fellow humans where blue is not prevalent in skin hues)

So any Multimedia compression techniques should use colour representation that presents colour in a way that models Human Visual system. We can then encode luminance is high bandwidth (more bits) than colour as this is much more perceptually relevant.

5 Marks -- Bookwork

(b) *List **three** distinct models of colour used in Multimedia. Explain why there are a number of different colour models exploited in multimedia data formats.*

Possible models:

- *RGB*
- *CIE Chromaticity*
- *YIQ Colour Space*
- *YUV (YCrCb)*
- *CMY/CMYK*

Different models reflect need to represent colour in a perceptually relevant model for effective compression.

Different models also due to evolution of colour from Video (YIQ,YUV), Display (RGB) and Print (CMYK) media requirements.

9 Marks Bookwork ---- 3 marks per model

(c) Suppose we have 24 bits per pixel available for a colour image. We also note that humans are more sensitive the red and green than to blue, by a factor of approximately 1.5 times. How may we design a simple colour representation to make use of the bits available?

Quite a simple scheme:

- Since Blue is less perceptually important use less bits to represent blue colour.
- Use proportionately more bits for red and green rather than blue
- Therefore
 - Red Green use 9 bits each and Blue 6 bits to represent values
 - Need to quantise at different levels for blue and Red/green

4 Marks --- Unseen

(d) Briefly explain why we need to have less than 24-bit colour representations (typically down to 8-bit) and why this is a sometimes a problem. Give one example where 8-bit colour representation have an advantage in terms of image/video processing?

Reasons:

Need more video memory

24 bit colour is overkill for even most photorealistic images

1 million colour at max present 24-bit allows for 16-17 Million colours!!

There a waste of video memory and bandwidth

Some video cards do not support 24-bit color (more a legacy issue, most modern PCs now support high end graphics)

Better to use 8-bit:

(Simple Answer) bandwidth friendly

(More complex answer) Some video effects more easily coded by colour table manipulation rather than pixel manipulation e.g. Making objects fade to background, color overlay

6 marks --- unseen

- 3 (a) Give **one** example each of a lossless and a lossy compression technique.

Possible Lossless Methods:

- Zero Length Suppression
- Pattern Substitution
- Run Length Encoding
- Shano-Fannon Encoding
- Huffman Coding
- LZW/GIF Coding
- Arithmetic Coding

Possible Lossy Methods

- Difference Encoding/Quantisation
- Discrete Cosine Transform Coding
- Vector Quantisation
- JPEG Coding (? Mix of above but acceptable?)

2 Marks --- Bookwork

- (b) Briefly explain the basic approach of entropy coding algorithms. Give **two** examples of entropy coding algorithms

Entropy measures come for the basics of Information Theory

The entropy of an information source S is defined as:

$$H(S) = \eta = \sum_i p_i \log_2 \frac{1}{p_i}$$

where p_i is the probability that symbol S_i in S will occur.

- $\log (1/p_i)$ indicates the amount of information contained in S_i , i.e., the number of bits needed to code S_i .

Basic Coding Idea

1. Measure entropy
2. Sort Data into some binary tree structure
3. Traverse Tree in top-down fashion assigning a 0/1 for each branch
4. At root node code is the path of 0/1 branches to the root
5. Assign codes for all root nodes

Two examples:

- Shano-Fannon Encoding (top-down binary tree sort)
- Huffman Coding (bottom-up based sort of data entropy)

7 Marks --- Bookwork

- (c) *Breifly state the Huffman coding algorithm. Show how you would use Huffman coding to encode the following set of tokens:*

AAABDCEFBBAADCDF

Huffman coding is based on the frequency of occurrence of a data item (Token). The principle is to use a lower number of bits to encode the data that occurs more frequently. Codes are stored in a Code Book which may be constructed for each image or a set of images. In all cases the code book plus encoded data must be transmitted to enable decoding.

The Huffman algorithm is now briefly summarised:

- A bottom-up approach
 1. Initialization: Put all nodes in an OPEN list, keep it sorted at all times (e.g., ABCDE).
 2. Repeat until the OPEN list has only one node left:
 - (a) From OPEN pick two nodes having the lowest frequencies/probabilities, create a parent node of them.
 - (b) Assign the sum of the children's frequencies/probabilities to the parent node and insert it into OPEN.
 - (c) Assign code 0, 1 to the two branches of the tree, and delete the children from OPEN.

For Problem of AAABDCEFBBAADCDF

Token count:

A 5
B 3
C 2
D 3
E 1
F 2

First Iteration Merge E and F

A 5
B 3
C 2
D 3
EF 3

2nd Iteration Merge B and C

A 5
BC 5
D 3
EF 3

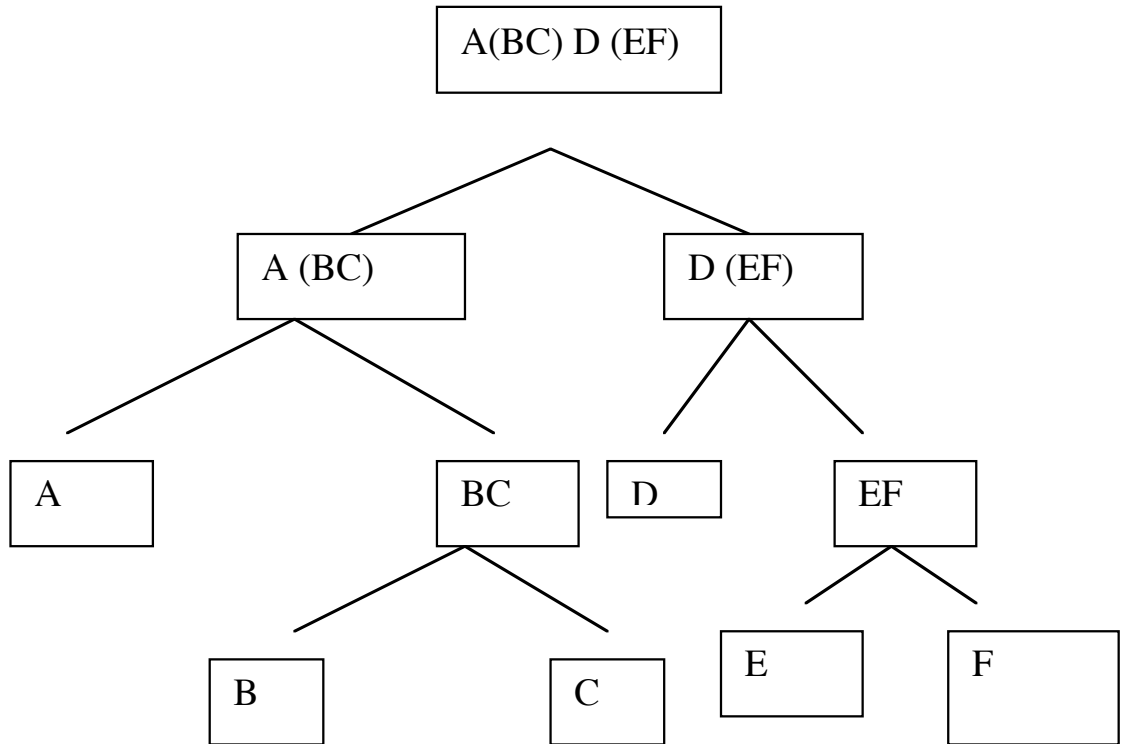
3rd Iteration Merge D with EF

A 5
BC 5
D(EF) 6

4th Iteration Merge A with BC

A (BC) 10
D(EF) 6

SO Tree is



Assigning 0 for left branch 1 for right we get codes

A 00
 B 010
 C 011
 D 10
 E 110
 F 111

9 Marks --- Unseen problem

- (d) *Briefly explain how the Huffman coding algorithm can be adapted for the streaming of live token streams.*

The basic Huffman algorithm has been extended, for the following reasons:

- (a) The previous algorithms require the statistical knowledge which is often not available (e.g., live audio, video).
- (b) Even when it is available, it could be a heavy overhead especially when many tables had to be sent when a non-order0 model is used, i.e. taking into account the impact of the previous symbol to the probability of the current symbol (e.g., "qu" often come together,

...).

The solution is to use adaptive algorithms. As an example, the Adaptive Huffman Coding is examined below. The idea is however applicable to other adaptive compression algorithms.

ENCODER

DECODER

```

Initialize_model();          Initialize_model();
while ((c = getc (input)) != eof)  while ((c = decode (input)) != eof)
{
    encode (c, output);          {
    update_model (c);              putc (c, output);
    }                             update_model (c);
}
}

```

- The key is to have both encoder and decoder to use exactly the same initialization and update_model routines

6 Marks --- application of bookwork

4. (a) What **two** broad classes of data compression techniques are applied to video compression? How does each class type typically get applied in video compression methods?

- Spatial Redundancy Removal - Intraframe coding (JPEG)
- Spatial and Temporal Redundancy Removal - Intraframe and Interframe coding (H.261, MPEG)

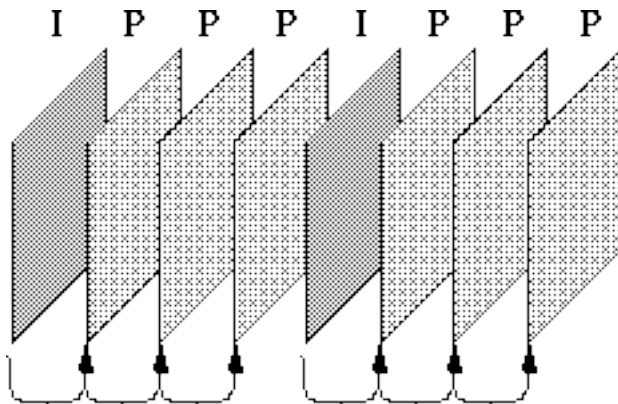
4 Marks --- Bookwork

(b) What is meant by a group of pictures in H.261 and MPEG video encoding? Briefly explain how each type of frame, in the group of pictures, achieves some form of video compression.

A group of pictures is a collection of coded video frames:

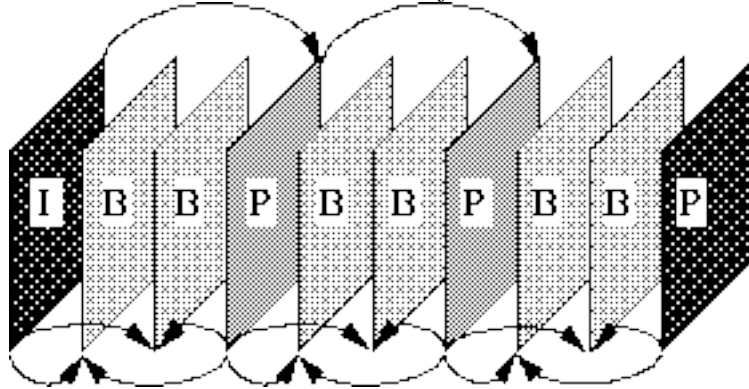
H.261 Two frame types: Intraframes (I-frames) and Interframes (P-frames)

1. I-frames use basically JPEG
2. P-frames use pseudo-differences from previous frame (predicted), so frames depend on each other.



I-frame provides us with an accessing point or key frame for each group of pictures

MPEG also has a *bidirectional B-frame*



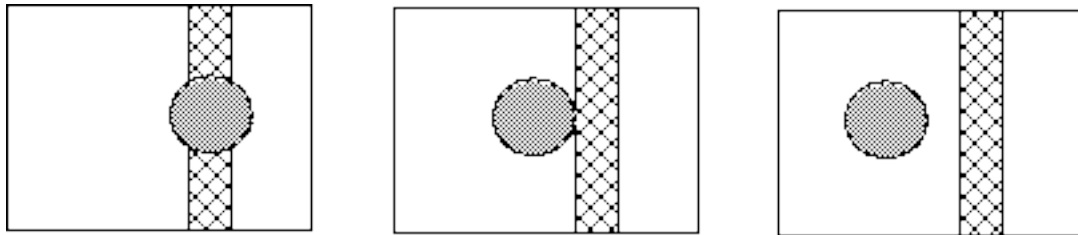
7 Marks --- Bookwork

- (c) *Briefly explain why a bidirectional B-frame improves video compression rates. What drawbacks are there with using B-frames?*

B-frame

Motion compensation works by tracking macroblocks from one frame to the next (in H/261 always the next forward frame)

However if we have occlusion (and possibly unexpected movements) then tracking prediction will fail



Solution:

Predict backwards as well as forwards. If occlusion in one direction track the other way.

Advantage is in coding efficiency

Disadvantage Overheads in maintaining B-Frames: The frame reconstruction memory buffers within the encoder and decoder must be doubled in size to accommodate the 2 anchor frames. Frames transmitted out of order.

8 Marks --- extended reasoning on bookwork

- (d) *If the display order of a group of pictures is IBBPBBPPBBI, what is the order of transmission and coding of this group?*

Coding and transmission order is

IPBBPBBIBB

Since the knowledge of each P frame and second I frame is needed to display using the B frames

5 Unseen Problem --- application of base knowledge but example unseen in notes