CARDIFF CARDIFF UNIVERSITY
EXAMINATION PAPER

SOLUTIONS

CHECKER: OFR

Academic Year:	2001-2002
Examination Period:	Lent 2002
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 multimedia and a multimedia system.*

   Multimedia is the field concerned with the computer-controlled integration of text, graphics, drawings, still and moving images (Video), animation, audio, and any other media where every type of information can be represented, stored, transmitted and processed digitally.

   A **Multimedia System** is a system capable of processing multimedia data and applications.

2 Marks - BOOKWORK

(b) What are the key distinctions between multimedia data and more conventional types of media?

Multimedia systems deal with the generation, manipulation, storage, presentation, and communication of information in digital form.

The data may be in a variety of formats: text, graphics, images, audio, video.

A majority of this data is large and the different media may need synchronisation -- the data may have **temporal** relationships as an integral property.

Some media is **time independent** or **static** or **discrete** media: normal data, text, single images, graphics are examples.

Video, animation and audio are examples of **continuous** media

4 Marks Bookwork

(c) What key issues or problems does a multimedia system have to deal with when handling multimedia data?

A Multimedia system has four basic characteristics:

- Multimedia systems must be computer controlled.
- Multimedia systems are integrated.
- The information they handle must be represented digitally.
- The interface to the final presentation of media is usually interactive.

Multimedia systems may have to render a variety of media at the same instant -- a distinction from normal applications. There is a temporal relationship between many forms of media (e.g. Video and Audio. There 2 are forms of problems here

- Sequencing within the media -- playing frames in correct order/time frame in video
- Synchronisation -- inter-media scheduling (e.g. Video and Audio). Lip synchronisation is clearly important for humans to watch playback of video and audio and even animation and audio. Ever tried watching an out of (lip) sync film for a long time?

The key issues multimedia systems need to deal with here are:

- How to represent and store temporal information.
- How to strictly maintain the temporal relationships on play back/retrieval
- What process are involved in the above.

Data has to represented digitally so many initial sources of data needs to be digitised -- translated from analog source to digital representation. The will involve scanning (graphics, still images), sampling (audio/video) although digital cameras now exist for direct scene to digital capture of images and video.

The data is large, several Mb easily for audio and video -- therefore storage, transfer (bandwidth) and processing overheads are high. Data compression techniques very common.

7 Marks BOOK WORK

(d) An analog signal has bandwidth that ranges from 15Hz to 10 KHz. What is the rate of sampler and the bandwidth of bandlimiting filter required if:

(i) the signal is to be stored within computer memory.

Nyquist Sample Theorem rate says that sampling must be **at least twice** the highest frequency component of signal or transmission channel

Highest frequency is 10 KHz so
Sampling rate = 20 KHz or 20,000 sample per second. **1 Mark**

Bandwidth of bandlimiting filter = 0 – 10 KHZ **2 Marks**

(ii) the signal is to be transmitted over a network which has a bandwidth from 200Hz to 3.4 KHz.

Channel has lower rate than max in signal so must choose this a limiting high frequency so

Sampling rate = 6.8 KHz or 6,800 sample per second. **2 Marks**

Bandwidth of bandlimiting filter = 0 – 3.4 KHZ **2 Marks**

7 Marks TOTAL: ALL UNSEEN

(e) Assuming that each signal is sampled at 8bits per sample what is the difference in the quantisation noise and signal to noise ratio expected for the transmission of the signals in (i) and (ii).

Quantisation noise = $V_{max}/2^{n-1}$

SNR = 20 log ($V_{maxa}l/V_{min}$)

So for (i) Quantisation noise = 78.125
SNR = 20 log (10,000/15) = 56.48 Db
3 Marks

And (ii) Quantisation noise = 26.56
SNR = 20 log (3,400/15) = 47.11 Db
4 Marks

2. (a) What is MIDI?

Definition of MIDI: a protocol that enables computer, synthesizers, keyboards, and other musical device to communicate with each other**.**

2 MARKS - BOOKWORK

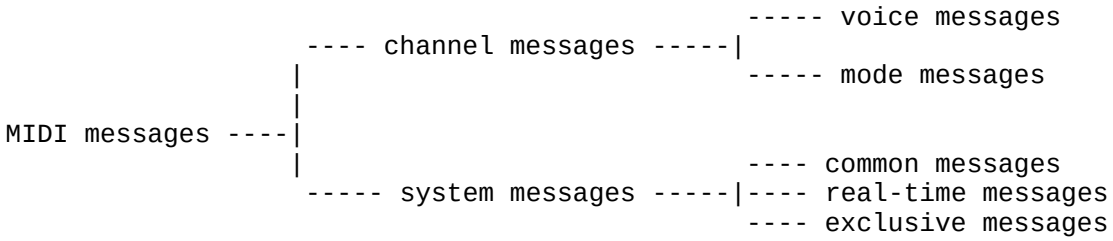
(b) How is a basic MIDI message structured? How is a MIDI message used to control how musical notes are played and also what instrument sounds the note?

MIDI messages are used by MIDI devices to communicate with each other.

Structure of MIDI messages:

- MIDI message includes a status byte and up to two data bytes.
- Status byte
- The most significant bit of status byte is set to 1.
- The 4 low-order bits identify which channel it belongs to (four bits produce 16 possible channels).
- The 3 remaining bits identify the message.
- The most significant bit of data byte is set to 0.

Classification of MIDI messages:



Channel messages:

- messages that are transmitted on individual channels rather that globally to all devices in the MIDI network.

TO SET SOUND USE Channel voice messages:

- Instruct the receiving instrument to assign particular sounds to its voice
- Turn notes on and off
- Alter the sound of the currently active note or notes

ESSENTIALLY TO PLAY SOUND USE

- NOTE ON
- NOTE OFF

Voice Message	Status Byte	Data Byte1	Data Byte2
Note off	8x	Key number	Note Off velocity
Note on	9x	Key number	Note on velocity

TO effect the sound of the note velocity, pitch bend, controller, pressure may be set:

Voice Message	Status Byte	Data Byte1	Data Byte2
Polyphonic Key Pressure	Ax	Key number	Amount of pressure
Control Change	Bx	Controller number	Controller value
Channel Pressure	Dx	Pressure value	None
Pitch Bend	Ex	MSB	LSB

TO CHOOSE A SOUND SET THE PRGRAM CHANGE :

Voice Message	Status Byte	Data Byte1	Data Byte2
Program Change	Cx	Program number	None

Note IN ALL ABOVE TABLES: `x' in status byte hex value stands for a channel number.

Example: a Note On message is followed by two bytes, one to identify the note, and one to specify the velocity.

To play note number 80 with maximum velocity on channel 13, the MIDI device would send these three hexadecimal byte values: 9C 50 7F

Program change usually sets a patch name up. Program Change values usually adhere to general MIDI sound set or the synthesizers internal patch settings.

General MIDI Instrument Patch Map (FULL DETAILS NOT REQUIRED FOR SOLUTIONS BUT SOME EGS MAY BE GIVEN)

Prog No.	Instrument	Prog No.	Instrument
(1-8	PIANO)	(9-16	CHROM PERCUSSION)
1	Acoustic Grand	9	Celesta
2	Bright Acoustic	10	Glockenspiel
3	Electric Grand	11	Music Box
4	Honky-Tonk	12	Vibraphone
5	Electric Piano 1	13	Marimba
6	Electric Piano 2	14	Xylophone
7	Harpsichord	15	Tubular Bells
8	Clav	16	Dulcimer
(17-24	ORGAN)	(25-32	GUITAR)
17	Drawbar Organ	25	Acoustic Guitar(nylon)
18	Percussive Organ	26	Acoustic Guitar(steel)
19	Rock Organ	27	Electric Guitar(jazz)
20	Church Organ	28	Electric Guitar(clean)
21	Reed Organ	29	Electric Guitar(muted)
22	Accoridan	30	Overdriven Guitar
23	Harmonica	31	Distortion Guitar
24	Tango Accordian	32	Guitar Harmonics
(33-40	BASS)	(41-48	STRINGS)
33	Acoustic Bass	41	Violin
34	Electric Bass(finger)	42	Viola
35	Electric Bass(pick)	43	Cello
36	Fretless Bass	44	Contrabass
37	Slap Bass 1	45	Tremolo Strings
38	Slap Bass 2	46	Pizzicato Strings
39	Synth Bass 1	47	Orchestral Strings
40	Synth Bass 2	48	Timpani
(49-56	ENSEMBLE)	(57-64	BRASS)
49	String Ensemble 1	57	Trumpet
50	String Ensemble 2	58	Trombone
51	SynthStrings 1	59	Tuba
52	SynthStrings 2	60	Muted Trumpet
53	Choir Aahs	61	French Horn
54	Voice Oohs	62	Brass Section
55	Synth Voice	63	SynthBrass 1
56	Orchestra Hit	64	SynthBrass 2
(65-72	REED)	(73-80	PIPE)
65	Soprano Sax	73	Piccolo
66	Alto Sax	74	Flute
67	Tenor Sax	75	Recorder
68	Baritone Sax	76	Pan Flute
69	Oboe	77	Blown Bottle
70	English Horn	78	Skakuhachi
71	Bassoon	79	Whistle
72	Clarinet	80	Ocarina
(81-88	SYNTH LEAD)	(89-96	SYNTH PAD)
81	Lead 1 (square)	89	Pad 1 (new age)
82	Lead 2 (sawtooth)	90	Pad 2 (warm)
83	Lead 3 (calliope)	91	Pad 3 (polysynth)
84	Lead 4 (chiff)	92	Pad 4 (choir)
85	Lead 5 (charang)	93	Pad 5 (bowed)
86	Lead 6 (voice)	94	Pad 6 (metallic)
87	Lead 7 (fifths)	95	Pad 7 (halo)
88	Lead 8 (bass+lead)	96	Pad 8 (sweep)

```
   (97-104  SYNTH EFFECTS)
 97        FX 1 (rain)
 98        FX 2 (soundtrack)
 99        FX 3 (crystal)
100        FX 4 (atmosphere)
101        FX 5 (brightness)
102        FX 6 (goblins)
103        FX 7 (echoes)
104        FX 8 (sci-fi)

   (113-120    PERCUSSIVE)
113        Tinkle Bell
114        Agogo
115        Steel Drums
116        Woodblock
117        Taiko Drum
118        Melodic Tom
119        Synth Drum
120        Reverse Cymbal

   (105-112  ETHNIC)
105        Sitar
106        Banjo
107        Shamisen
108        Koto
109        Kalimba
110        Bagpipe
111        Fiddle
112        Shanai

   (121-128  SOUND EFFECTS)
121        Guitar Fret Noise
122        Breath Noise
123        Seashore
124        Bird Tweet
125        Telephone Ring
126        Helicopter
127        Applause
128        Gunshot
```

8 MARKS – BOOKWORK
3 marks for basic midi message
5 marks Midi control .

(c) You have been asked to advise on a Multimedia production that is to send an audio production of a network. The transmission is to occupy as low a bandwidth as possible but yet maintain as high as possible audio fidelity. The audio production will involve voices, for narration and dialog, music both electronic and acoustic instruments and sound effects.

Comment on the most appropriate methods in which to deliver such a presentation and also address to relative merits and drawbacks of the methods you suggest in your solution.

Sketch of solution with breakdown of marks for each.

Choice between using

- sampled sounds and
- MIDI sound synthesizers

4 Marks for basic idea.

Sample Sounds :

used for audio that can't be synthesized

- Narration
- Dialog
- Some Acoustic instruments if fidelity or availability is not on synthesizer
- Some Sound effects instruments if fidelity or availability is not on synthesizer

MIDI Synthesiser:

Can be Sofware (Web Browser, Quicktime) or soundcard or external hardware sound source
Most musical instruments but NOT Vocals and some specialised instruments
Some sound effects (GM sound set (above) and other samples.

6 marks for knowing basic uses of each in application

Sampled Sound Versus Synthesiser:

Sampled Sound:

- Consists of analog waves that are digitally sampled. Sample rate and sample bit quantisation have huge effect in fidelity and also the bandwidth of sample.
- Size of file can be large
- Can represent any audio sound in sample
- Sample file format supported may vary over different client machines
- Synchronisation of multiple audio sample a potential problem

Synthesised sound:

- MIDI data is very compact, less than 100 Kb for half hour of audio
- Client is responsible for making sound via software or hardware synthesisier
- Synthesised is not actual audio sample of sound source.
- Midi Synthesiser sound will vary between client. Cannot guarantee fidelity of playback.
- Many MIDI tracks can be synchronized with ease

7 Marks for analysis Could include other ideas

TOTAL 17 MARKS – TOTALLY UNSEEN

3 (a) Why is integration of multimedia data a potential problem for multimedia systems? Briefly how are these problems addressed in such systems?

Many formats and types of media type or format individually form text, graphics to audio, animation and video. Many may be compressed as part of their format and require uncompression before they can be rendered.

Certain media (individually) are based on spatial and/or temporal representations (e.g audio, video, animations), other may be static (Text, graphics).

Multimedia applications need integrate media spacial and temporal implications become even more critical. E.g. static text may need to index or label a portion of video at a given instant or segment of time and there the integration becomes temporal and spatial if the label is placed at a given location (or locations moving over time).

Clearly, it is important to know the tolerance and limits for each medium as integration will require knowledge of these for **synchronisation** and indeed create further limits (e.g. bandwidth of two media types increase, if audio is encoded at a 48Khz sampling rate and it needs to accompany video being streamed out at 60 frames per second then inter-stream synchronisation is not necessarily straightforward.

Spatial, temporal structural and procedural constraints will exist between the media.

This especially true now that interaction is a common feature of multimedia**.**

So many problems are

- Dealing with lots of media and formats
- Potential computational overhead to decompress data
- Synchronisation of data due to
  - Different formats
  - Different sample rates
  - Bandwidth problems may lead to dropping of data

7 Marks Bookwork

(b) What is MHEG and what is its target application domain?

MHEG is a family of standards defined by Multimedia and Hypermedia Information Encoding Expert Group (MHEG) . MHEG specifies an encoding format for multimedia applications independently of service paradigms and network protocols.

The development of MHEG arose directly out of the increasing convergence of broadcast and interactive technologies.

3 Marks Bookwork

(c) Describe giving suitable code fragments how you would effectively combine a video clip and an audio clip in an MHEG application and start a subtitle text display 20,000 milliseconds into the video clip. You may assume that both clips are of the same duration and must start at the same instant. The Video should be displayed in the whole of the presentation and the subtitles should run along the bottom of the presentation.

TheMHEG code listing below is illustrates the solution:

```
{:Application ("SYNC_APP.mh5" 0)
:OnStartUp ( // sequence of initialization
actions
}
{:Scene ("main_scene.mh5" 0)
:OnStartUp ( // sequence of initialization
actions
preload (2) // the connection to the
source of the video clip is set up
preload (3) // the connection to the
source of the audio clip is set up

...
setCounterTrigger (2 3 20000) // book a
time code event at 20000 msec for example
...
)
:Items ( // both presentable ingredients and
links
{:Bitmap 1 // background bitmap  NOT IMPORTANT FOR SOLN
:InitiallyActive true
:CHook 3 // JPEG
:OrigContent
:ContentRef ("background.jpg")
:OrigBoxSize 800 600
:OrigPosition 0 0
}
{:Stream 2 // video clip
:InitiallyActive false
:CHook 101 // MPEG-1
:OrigContent
:ContentRef ("video.mpg")
:Multiplex (
```

```
{:Audio 3 // audio component of the
video clip
:ComponentTag 1 // refers to audio
elementary stream
:InitiallyActive true
}
... // possibly  more presentable ingredients

}
{:Link 49 // the video clip crosses a pre
defined time code position
:EventSource (2) // video clip
:EventType CounterTrigger
:EventData 3 // booked at startup by
setCounterTrigger (2 3 20000)
:LinkEffect (
:SetData (5 // text subtitle is set
to a new string, that is
:NewRefContent ("subtitle.txt")) //
"Subtitle text"
:SetHighlightStatus (20 true) //
hotspot 20 is highlighted
:setPosition (100,550)
)
}
... //  more links
)
:SceneCS 800 600 // size of the scene's
presentation space
}
```

Marking Scheme:

Basic Outline of MHEG presentation **2 Marks**
Basic Objects set up **2 Marks**
Key Points:

- Preloading both clips is essential to start streaming: **3 Marks**
- Need to “book” 20000 msec event for subtitles **3 Marks**
- Content loaded and Video and audio is multiplexed **2 Marks**
- Set a link transition for subtiltle **3 Marks**
- Set up the position for the subtile `:setPosition (100,550) at bottom of page. Note Coordinates are only example any values like this will suffice` **2 Marks**

TOTAL 17 MARKS TOTALLY UNSEEN

4. (a) Why is data compression necessary for Multimedia activities?

Audio and Video and Images take up too large memory,disk space or bandwidth uncompressed.

3 Marks BookWork

(b) What is the distinction between lossless and lossy compression? What broad types of multimedia data are each most suited to?

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.

Types suitability

Lossless
Computer data fles (compression)
Graphcs and graphical images lossless (GIF/LZW)

Lossy
Audio MP3
Photographic images (JPEG)
Video (Mpeg)

5 Marks Bookwork

(c) Briefly explain the compression techniques of zero length suppression and run length encoding. Give one example of a real world application of each compression technique.

Simple Repetition Suppresion

Simplest Suppression of zero's in a --- ***Zero Length Supression***

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 *code* to

denote when the repated token appears

For Example

894000000000000000000000000000000

we can replace with

894f32

where f is the code for zero.

Example:

- Silence in audio data, Pauses in conversation
- Bitmaps
- Blanks in text or program source files
- Backgrounds in images

Run-length Encoding

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

In this instance, sequences of image elements (X_1,X_2, …X_n)are mapped to pairs (c_1,l_1, c_2,l_2…,(c_n,l_n) where c_i represent image intensity or colour and l_i the length of the ith run of pixels (Not dissimilar to zero length supression above).

For example:

Original Sequence:

11112223333311112222

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.

Examples

- Simple audo
- graphics
- Images

7 Marks Bookwork

(d) Show how you would encode the following token stream using zero length suppression and run length encoding:

ABC000AAB00000000DEFAB00000

Total length of token stream = 27

Zero Length Suppresson Code

ABCf3AABf8DEFABf5

Number of tokens 17 where f is code for 0

Run Length Encoding

A1B1C103A2B108D1E1F1A1B105

Number of tokens 26

(i) What is the compression ratio for each method when applied to the above token stream?

Total length of token stream = 27

Zero Length Suppresson Code

ABCf3AABf7DEFABf5

Number of tokens 17 where f is code for 0

Run Length Encoding

A1B1C103A2B107D1E1F1A1B105

Number of tokens 26

Compresson ratios:

Zero length Supresson = 17/27

Run Length Encoding = 26/27

3 Marks each for correct encoding
2 Mark for each ratio
10 Marks Total

(ii) Explain why one has a better compression ratio than the other. What properties of the data lead to this result?

Data has only one repeated token the 0. So coding is wasted on rapidly changing remainder of data in run length encoding where every token frequency count needs recording.

2 Marks

12 Marks for all of PART (d) ALL WORK UNSEEN