Chapter 17. Multimedia

Table of Contents

Introduction to Multimedia

Context

This chapter unit looks back to chapter 5, which dealt with coding. While coding can be used to digitise discrete media such as text, it cannot cope with continuous media such as audio and video. Analogue-to-digital conversion is introduced here as a way of digitising continuous media, and after this we can digitise each of the various media, and so are in a position to digitise multimedia. This leads on to the following units by giving a key aspect of the technical background necessary to understand the World Wide Web and its uses.

Introduction

Discrete media, such as text, may be converted to data by a coding process in which their successive symbols are, in turn, converted to a representative data sequence. Continuous media, such as audio and video, may also be converted to data. The process by which this is done is known as analogue-to-digital conversion. In this unit, the technique known as pulse-code modulation is described. With the ability to convert the various media to data, the data exchange capability that has already been established becomes the capability to act as a multimedia network.

The pulse-code modulation process is illustrated by examples relating to CDs and CD-ROMs. The multimedia capability is illustrated with reference to the World Wide Web.

Objectives

At the end of this module, you should be able to:

- show how discrete media can be converted to data;
- show how continuous media can be converted to data;
- show that, since the various media can be converted to data, the Internet, as a data transmission network, is also a multi-media network.

From a data network to a multimedia network

In parallel with this topic, you should read relevant sections of your textbooks.

It has already been established that the Internet is a data delivery network. A major aim of this unit is to show that, by the same token, it is also a multimedia network. To do this, we demonstrate that the various media that are generally taken to comprise multimedia, that is, text, images, graphics, speech and video, can be converted to data.

The key ideas in the conversion of the different media to data are coding and analogue-to-digital conversion. Coding is used to convert media, such as text, that consist of discrete symbols. Analogue-to-digital conversion is needed for media that are analogue in nature.

Discrete media

With discrete media, the 'messages' are composed of symbols that are drawn from some finite repertoire. With text, for example, the textual messages are composed of alphanumeric symbols and each symbol is taken from a finite alphabet. With bit-mapped computer images, each 'image message' is composed of coloured dots (pixels) and the colours are taken from the finite range that is supported by the computer's graphics card. We consider in turn the digitisation of text, images and graphics.

Text

The standard way of converting text to data is to use the ASCII code which, as we know, provides an 8-bit pattern for each alphanumeric symbol, punctuation symbols and some other symbols, including ACK and NAK. A part of the code is:

Symbol	**Code**
A	01000001
B	01000010
C	01000011
D	01000100
E	01000101
F	01000110
<space>	00100000
...	...

So, to transmit text over a data link (perhaps as an e-mail), or to store it on a data storage device (as a text file), the text is coded symbol by symbol using this code. Subsequently, the text can be recovered by decoding the data, using the coding table to convert successive 8-bit patterns to symbols.

(Note that although an 8-bit code allows us to code 256 symbols, in the older ASCII code the first bit was fixed at 0 so that it provided codes for only 128 symbols on any host machine.)

To Do

Do Review Questions 1, 2 and 3.

Images

An image intended for display by a computer is a rectangular grid of dots. The resolution of the image is given by the number of rows and columns in the array. The graininess of the image will decrease as the resolution increases. Each dot has its own colour drawn from a finite palette. The accuracy of the colour of the image will increase as the size of the palette increases. An image of this kind can be described by giving the colour of each of its dots in some well defined order. The usual way of ordering the list is to start at the top-left corner and proceed row by row to the bottom-right corner. (Just as this is the usual way of reading a page of text.)

An image, then, can be converted to data by listing the codes for the colours of the dots in the prescribed order. Assuming that there are 128 colours in the palette, we can code them using the same table as for the ASCII code, except that now we take A to mean Colour A (which can be red) rather than letter A, B to mean colour B (which can be orange) rather than letter B, and so on. The coding table becomes:

Color	**Code**
A (red)	01000001
B (orange)	01000010

Color	Code
C (yellow)	01000011
D (green)	01000100
E (blue)	01000101
F (indigo)	01000110
...	...

To Do

Do Review Questions 4, 5 and 6.

Graphics

Now consider an image created with an object-oriented drawing programme. Each item in the image, be it a line, a circle or a rectangle, is an object. Each image created with such a programme is a collection of objects. The problem, then is to represent a collection of objects as data.

First, treating the objects that can be drawn by the programme as symbols, we can devise a code along the lines of the ASCII code for them. As long as there are not more than 128 distinct types of object, by calling a line Object A, a circle Object B, and so on we can use the following code.

Color	Code
A (line)	01000001
B (circle)	01000010
C (rectangle)	01000011
D (arc)	01000100
E (polygon)	01000101
F (text)	01000110
...	...

Then various properties of each object need to be represented, including, for example, its position, size and colour. An object's position can be given by its co-ordinates, that is, by a pair of numbers, and a number can be expressed as a binary number. The size is numeric, and can be dealt with similarly. The colour can be limited to one from a finite palette, and then each colour can then be assigned a code.

In this way, an image can be represented by a list of objects. Each object can be represented by a formatted structure giving its type, size, position and colour. Each item in the structure has a binary representation either through a code or a direct conversion. And, by these stages, an object-oriented graphic can be converted to data.

To Do

Do Review Questions 7, 8 and 9.

Audio and Video

This section should be read in conjuction with your textbooks and other resources that are available to you.

Analogue to digital conversion

Audio (for example, speech and music) and video (film and TV) are, as far as human perception is concerned, both analogue in nature. We accept them as waveforms, either audible or visible, that are essentially continuous and can vary continuously. To illustrate this, an audible medium such as music, can vary continuously through time in volume and pitch, while visible media can vary continuously through time in brightness and colour. It is more difficult to convert analogue media to data than to convert media such as text which are composed from a finite repertoire of discrete symbols. The range of possibilities for any instant is infinite with analogue media whereas with text and similar media it is finite.

One of the most widely used methods for analogue-to-digital conversion is known as pulse-code modulation. It works by reducing the infinite possibility of an analogue medium to a finite number of possibilities. This is done by first reducing time from a continuum to a set of discrete instants, and then by reducing the values that can be assumed at each instant (which can correspond to volume or to brightness) from a continuum of possibilities to one of a set of discrete possibilities. Each of the possible resulting values can then coded (by treating it as a symbol) in the way described above for discrete media.

The stages of analogue-to-digital conversion when it is achieved by pulse-code modulation are as follows:

1. Sample the analogue signal.

   This converts it from a continuous waveform to a sequence of samples.

   To avoid any loss of fidelity, the samples must be taken at a rate that exceeds twice the bandwidth. For example, the bandwidth of telephone-quality speech is 3.4 kHz, so that it must be sampled at a rate exceeding 6.8 kHz, which means that at least 6.8 thousand samples per second are required. In practice, a standard sampling rate of 8 kHz is used.

2. Classify each sample.

   The sequence of samples becomes a sequence of class names (which are actually interval identifiers). Each name can be treated as a symbol.

   The range of values assumed by the signal is divided into a finite number of intervals, and each sample from the signal is classified according to the interval in which it falls. (We can call the intervals Interval A, Interval B, and so on.)

   In telephony, 256 intervals are used.

3. Code the interval identifiers.

   The analogue medium has been converted to a sequence of symbols which can be coded in the same way as text and images. (Interval A can be coded as A, Interval B as B, and so on.)

   With telephony, 8 bit codes are used (256 = 28).

   This gives the data rate for telephony as 8k [samples per second] * 8 [bits per sample] = 64 kbits per second.

The following table gives the bandwidth, standard sampling rate and number of quantisation intervals (the technical term for the classification process quantisation) used when various types of analogue signals are converted to data. The signal types are, respectively, telephone-quality speech, high-quality music such as that recorded on a compact disc, and video, that is, colour television.

Type of signal	**Bandwidth**	**Standard sampling rate**	**Number of quantisation intervals**
speech	0.3 - 3.4 kHz	8 kHz	256
music	15 kHz	32 kHz	2048
Colour TV	5.5 MHz	13 MHz	512

Note that for each type of signal the standard sampling rate exceeds twice the bandwidth. Standard rates are needed to ensure that different systems can work with each other. The way in which the number of quantisation intervals is determined will be dealt with later.

In practical systems of interest for audio and image the analogue signal is almost always pre-processed so as to take advantage of the properties of the human senses that allow significant reductions in the required data rates.

To Do

Do Review Questions 10, 11 and 12.

Carry out Activity 1.

Digital-to-analogue conversion

Audio and video can be recovered from their data representations by, as far as it is possible, reversing the analogue-to-digital conversion process. This is done by:

1. Decoding, to recover a sequence of interval identifiers.

2. Reconstructing the prototypical sample for each interval identifier. Once a sample has been classified, its exact value cannot be recovered by knowing the class it belongs to. The best approximation to its value is provided by the value in the centre of the range of values for its class. For this reason, each sample is reconstructed with the value that corresponds to the centre of its interval. In this way, the error that inevitably follows from classifying a sample (that is, from replacing its precise value by a classification that records that this value falls in a certain range of values) is restricted so that it cannot exceed half the width of the interval. This form of error is known as quantisation error.

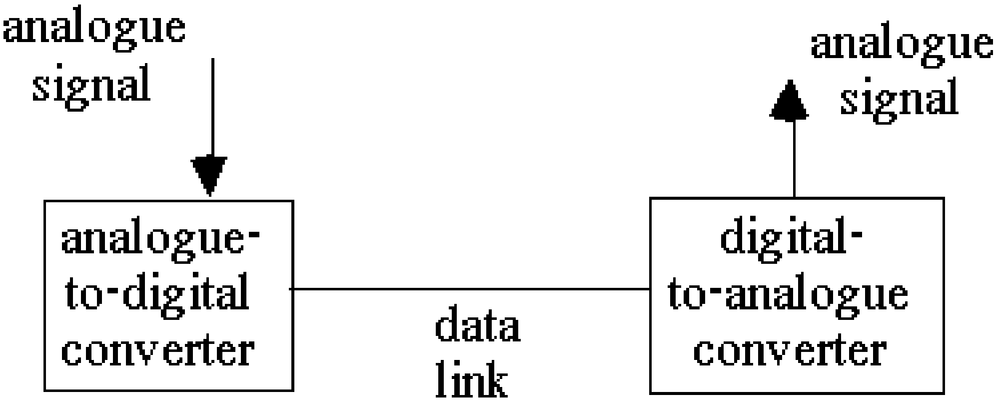
3. Reconstructing the continuous waveform passing through these reconstructed samples. In practice, this waveform recovery is achieved by filtering the samples, which acts to smooth them out into the waveform.

Since pulse code modulation is used to convert music to the data that is recorded on a CD, a CD player is an example of a device that performs this reverse (digital-to-analogue conversion) process.

To Do

Do Review Questions 13 and 14.

The situation when an analogue signal has to be carried from its source to its destination by a data link, can be represented in the following way:



This diagram could represent, for example, the use of a data link to convey a television signal from a television transmitter to a television set. Referring back to the table given above, one way to determine the number of quantisation levels needed for colour television would be simply to try different numbers of levels, and to increase the number until the viewer could not tell that a data link had been used to transport the signal. So, if 32 quantisation levels are used, the viewer can see that there is something wrong with the picture on the television set. The picture is still not satisfactory when the number of intervals is increased to 64, 128 and 256. But with 512 levels, the viewer cannot see any difference between the pictures displayed on the television set and those shown when television signals are transmitted in the normal way.

To Do

Carry out Activity 2.

The Internet as a multimedia network

Given that the Internet can be regarded as a data delivery network. This unit has shown that the various media can all, by one means or another, be converted to data. Consequently, the Internet may also be regarded as a multimedia network. Its support of the World Wide Web is, of course, a demonstration of the truth of this assertion. Internet telephony, RealAudio and RealVideo, and WebTV are further illustrations of the multimedia capabilities of the Internet.

Multimedia storage

Just as the capability to convert multimedia to data allows it to be transmitted over a data network, so it allows it to be stored on a data storage device such as a CD-ROM. A CD-ROM has a storage capacity of up to 650 Mbytes.

To Do

Do Review Questions 15, 16 and 17.

Carry out Activities 3, 4 and 5.

Activities>

Activity 1 - pulse code modulation

Explain how the first two stages of the pulse code modulation process progressively reduce a continuous source to a discrete source so that the latter can then be treated in the way described in chapter 5.

You can find a discussion of this activity at the end of the chapter.

Activity 2 - transmission rate

Signals with a 10 kHz bandwidth are to be transmitted by pulse code modulation. An accuracy of 0.25% of the dynamic range is required for the received signals. What is the necessary transmission rate?

You can find a discussion of this activity at the end of the chapter.

Activity 3 - cd-rom

How do you explain the answer to Review Question 16, given that a music CD holds at most something just over 70 minutes of music?

You can find a discussion of this activity at the end of the chapter.

Activity 4 - cd-rom 2

What conclusion do you draw from the answer to Review Question 17?

You can find a discussion of this activity at the end of the chapter.

Activity 5 - cd-rom 3

A multimedia CD-ROM allots equal amounts of storage to speech, music and video. How much of each medium does it contain?

If a CD-ROM held amounts of speech, music and video of equal durations, how much storage would it have to allocate to each?

You can find a discussion of this activity at the end of the chapter.

Review Questions

Review Question 1

Code the message: A BAD FACE.

You can find an answer/comment for this review question at the end of the chapter.

Review Question 2

How long will it take to transmit the message from question 1 over a data link operating at 20 bits per second?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 3

How much storage is needed to store the message from question 1 in computer memory?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 4

Describe the result of digitising an image with a resolution of 100 by 100 that is completely blue.

You can find an answer/comment for this review question at the end of the chapter.

Review Question 5

How long will it take to transmit the image from question 4 over a data link operating at 10 thousand bits per second?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 6

How much storage is needed to store the message from question 4 in computer memory?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 7

In an object oriented graphics system, each object is represented by its type, size, position and colour. The type (Object A, Object B and so on) is coded as above. Objects can have one of 64 sizes ranging from size 0 for the smallest to size 63 for the largest. The screen resolution is 512 by 340. The colours are coded using the table from the previous section. How many bits are needed to represent each object?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 8

A graphic image contains ten lines, four circles and two rectangles. How may bits are needed to represent this graphic?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 9

How long will it take to transmit the graphic from question 9 over a link operating at 64 kbits per second?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 10

Speech signals ‘keep coming’ as long as someone is speaking, so that when they are converted to data the data ‘keeps coming’ and, what, is more comes at a characteristic rate.

What are the data rates for high-quality music and for video?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 11

How much storage is needed to hold a minute of speech, a minute of high-quality music, and a minute of video?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 12

To compare the amount of storage needed by discrete media items with the amounts needed by continuous media items, answer the following.

How much storage is needed for a page of text? (Assume, say, 300 words to a page and that the average word contains 5 letters.)

How much storage is needed for a bit-mapped image with a resolution of 512 by 340 when each pixel can assume a colour drawn from a palette of 256 colours?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 13

What causes lack of precision when analogue-to-digital conversion is carried out using pulse code modulation?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 14

What is the maximum extent of the waveform error caused by the pulse code modulation process?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 15

How much telephone-quality speech can be stored on a CD-ROM?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 16

How much high-quality music can be stored on a CD-ROM?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 17

How much video can be stored on a CD-ROM?

You can find an answer/comment for this review question at the end of the chapter.

Discussion Topics

1. Are there ways to reduce the amounts of data produced by using pulse code modulation for analogue to digital conversion? Would it be better to reduce the amount at source by having a better means of conversion, or would it be better to retain pulse code modulation and use data compression on its results?
2. We have seen that the ability to convert the various different media to data leads to a ‘convergence’ as far as transmission is concerned. Could there be a corresponding convergence in the end-user devices needed to deal with the various media?
3. Would it be sensible to have digital end-user devices so as to avoid some of the conversion stages that are otherwise needed when exchanging media?
4. What is meant by ‘streaming’ media?

Answers and Comments

Activity 1

Waveforms are continuous in time. The first stage of pulse code modulation, the sampling, serves to convert this continuous dimension to a discrete one. Waveforms can assume any value within the range extending from their maximum to their minimum values. (This is known as the dynamic range of the signal.) The second stage, classification or quantisation, acts to convert this continuous dimension to a discrete one

Activity 2

An accuracy of 1/400 of the signal range means that each quantisation interval must cover at least 1/200 of the range, so that at least 200 quantisation intervals are needed. This requires an 8-bit code. The sampler must take at least 20 thousand samples per second, leaving the transmission rate at a minimum of 20k*8 = 160 kbits per second.

Activity 3

The calculation for Review Question 16 is made on the assumption that there is nothing on the CD but music. In fact, there is an index track, and quite a lot of storage contains handling bits. This reduces the effective storage capacity. In addition to this, two music tracks are stored to allow stereo reproduction, which halves the capacity of the remainder.

Activity 4

That nobody would buy a CD-ROM that contained only 44 seconds of video, and that ways of either compressing the data or expanding the capacity of the disc are required.

Activity 5

Allocating one third of the capacity to each medium would mean that the CD-ROM held just over 7 hours of speech, about one hour and 20 minutes of music and 15 seconds of video.

Equal times for the respective media require storage amounts with the ratios 64:352:117,000. So the fraction of the storage needed by the video is 117,000/117,461, which amounts to 647.7 Mbytes. The fraction needed by the music is 352/117,461, which amounts to 1.95 Mbytes. The remaining 0.35 Mbytes is for speech.

Review Question 1

010000010010000000100001001000001 . . .

Review Question 2

It takes 4 seconds to send 80 bits at 20 bits per second.

Review Question 3

80 bits or 10 bytes of storage are needed.

Review Question 4

The sequence 01000101 is repeated 10,000 times.

Review Question 5

It takes 8 seconds to send 80,000 bits (80 kbits) at 10 kbits per second.

Review Question 6

80 kbits or 10 kbytes.

Review Question 7

8 bits for the object type, plus 6 bits for the size, plus 18 bits for the position (9 each for the row and column numbers), plus 8 bits for the colour gives a total of 40 bits per object.

Review Question 8

16 objects need 16*40 = 640 bits

Review Question 9

It takes 1/100 of a second to send 640 bits at 64 kbits per second.

Review Question 10

Using the table, high-quality music is sampled 32 thousand times per second, and each sample is classified into one of 2048 intervals. This means that an 11-bit code is needed. So, 32 thousand times a second a sample is replaced by 11 bits, which gives a data rate of 32k*11 = 352 kbits per second.

The data rate for video is 13 M*9 = 117 Mbits per second.

Review Question 11

A minute of speech needs 60*64 kbits = 3840 kbits = 480 kbytes.

A minute of music needs 60*352 kbits = 21120 kbits = 2640 kbytes.

A minute of video needs 60*117 Mbits = 7020 Mbits = 877.5 Mbytes.

Review Question 12

A page of text needs 300*(5+1) = 1800 bytes. (The one is for the space that separates the words.)

The image would need 640*640*8 bits = 409.6 kbytes.

Review Question 13

The quantisation, or classification stage.

Review Question 14

Half the width of an interval.

Review Question 15

One second of speech needs 8 kbytes of storage, so the amount of speech is 650M/8k = 81.25 k seconds = 22.5 hours approximately.

Review Question 16

One minute of music needs 2640 kbytes, so the amount of music is 650 M/2.64 M = 246 minutes or approximately 4 hours.

Review Question 17

One minute of video needs 877.5 Mbytes, so the amount of video is 650 M/877.5 M = 0.74 minutes or 44 seconds.