Digital Video, MPEG and Associated Artifacts

Mr Shanawaz A. Basith
Mr Stephen R. Done

Departments of Computing & Electrical Engineering
Imperial College, 180 Queen's Gate, London, SW7 2BZ
email : sab@doc.ic.ac.uk, srd2@doc.ic.ac.uk

June 14, 1996

Abstract

Compression and digital video is introduced, transform coding, motion compensation and the need for compression are discussed. The technology and objectives behind the MPEG digital video encoding and decoding standard are described. A comparison of a few competing standard for compression of digital video is presented and the associated advantages and disadvantages are given. A few applications of digital video are discussed. Finally problems with digital video compression are described.

Contents

A.1. Appendix One

1. Introduction

Reducing the amount of data needed to reproduce video saves storage space, increases access speed and is the only way to achieve motion video on digital computers. This document looks at digital video and explains some techniques of reducing the storage space needed.

It was in the late eighties that the audio and video industry faced the prospect of saturated markets and over capacity. What was required were new products and services that would captures consumers' imagination. The data capacity of existing digital storage and their transmission links limited its potential for further exploitation. What was needed were standards that the industry could follow.

This document also looks at one such standard, this methodology was not borne out of the desire for commercial overpowering, but by an independent body that recognised the problems at the time.

After looking at the Moving Pictures Experts Group (MPEG) standard, an objective look at some of its many competitors in the same market place is looked at. Some applications of digital video are presented.

Digital video is not without its problems, many of which are shared by all digital medium. These problems are discussed to some length.

2. Compression and Digital Video

A great deal of research has gone into image and video compression and indeed it is quite difficult to invent something new in this field. A diagram showing the many compression techniques are shown in figure 2.1. The assumption is that the input is always a PCM digitised signal in colour components. The output of the compression process is a bitstream. Lets consider each technique briefly :

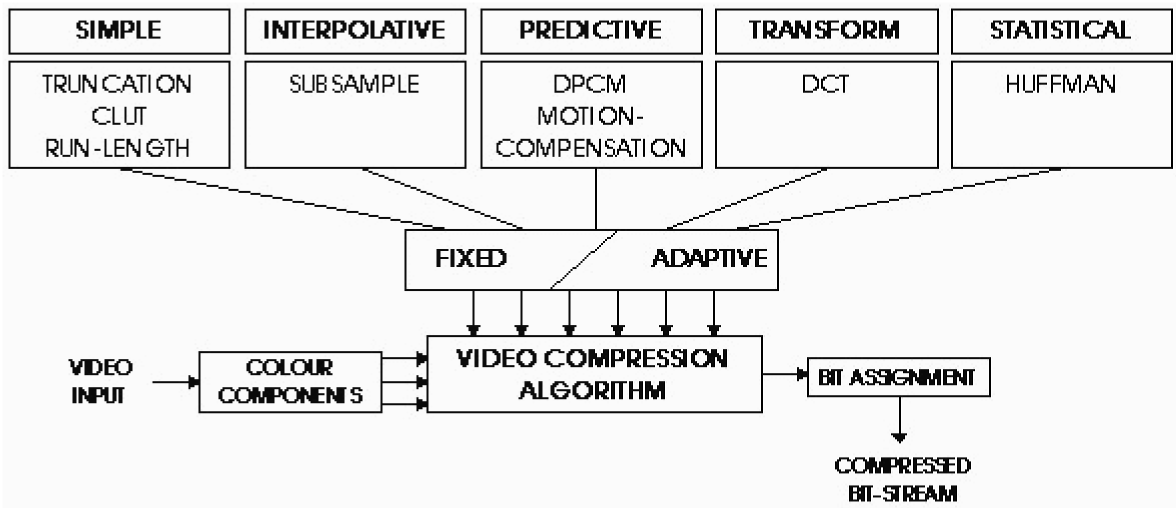

Figure 2.1 - Compression Techniques

Simple Compression Techniques

Various techniques exist, including :

- Truncation

  Reduces data by reducing the number of bits per pixel. This suffers from contouring (resolution loss) but has the advantage that processing is simple.

- Colour Lookup Table (CLUT)

  Pixel values represent an index into a table of colours. The processing for this is non-trivial.

- Run length coding

  Blocks of repeated pixels are replaced with a pixel value and a count. This works well on images with blocks of single colours and can achieve a high compression ratio. However it not effective if images contain no repetitive areas.

Interpolative Techniques

This technique aims to send a subset of the pixels and use interpolation to reconstruct the intervening pixels. This technique is particularly useful for motion sequences, as certain frames are compressed by still compression ; the frames between these are compressed by doing an interpolation between the other frames and sending only the data needed to correct the interpolation.

Predictive Techniques

This relies on the fact that there is nearly always some redundancy between frames in a sequence. There are two common methods :

- DPCM (Differential Pulse Code Modulation)

  This operates at the pixel level and sends only the difference between successive pixels. Since there is likely to be very little difference between adjacent pixels we can encode the value into smaller data widths. This technique suffers from slope-overload which causes smearing at high contrast edges in an image.

- ADPCM (Adaptive DPCM)

  This tries to reduces the slope-overload by using smaller steps for the difference values.

Transform Coding Techniques

A transform is a process that converts data into an alternate form which is more convenient for some particular purpose. Transforms are ordinarily designed to be reversible. Useful transforms typically operate on large blocks of data and perform some complex calculations. In general transform coding becomes more useful with larger blocks. The Discrete Cosine Transform (DCT) is especially important for video compression.

The DCT

The DCT is performed on a block of horizontally and vertically adjacent pixels (typically an 8 by 8 block of pixels). The outputs represent amplitudes of two dimensional spatial frequency components. These are called DCT coefficients. The coefficient for zero spatial frequency is called the DC coefficient and it is the average value of all the pixels in the block. The rest of the coefficients represent progressively higher horizontal and vertical spatial frequencies in the block.

Since adjacent pixel values tend to be similar or vary slowly from one to another, the DCT processing provides opportunity for compression by forcing most of the energy into lower spatial frequency components. In most cases, many of the higher frequency coefficients will have zero or new-zero values and therefore can be ignored.

The decoder performs the reverse process, but due to the transcendental nature of the DCT the reverse process can only be approximated and hence some loss takes place. The trick is to use some cunning methods of keeping coefficients so that the loss is minimally visible.

Statistical Coding (or Entropy Coding)

This takes advantage of the statistical distribution of the pixel values. Some data values can occur more frequently then others and therefore we can set up a coding technique that use less bits for these values. One widely used form of this coding is Huffman encoding. This technique has the overhead that a syntax has to be pre-defined or sent for the decoder to work.

Motion Compensation

Consider the case of a video sequence where nothing is moving in the scene. Each frame of the video should be exactly the same as the previous one. In a digital system, it should be clear that, we only need to send one frame and a repetition count. Consider now, a dog walking across the same scene. The scene is the same throughout the sequence, but only the dog moves. If we could find a way of only sending the motion of the dog, then we can save a lot of storage space. This is an oversimplified case of motion video, but it reveals two of the most difficult problems in motion compensation :

- How can we tell if an image is stationery ?
- How do we extract the part of the image that moves ?

We can try to answer these questions by some form of comparison between adjacent frames of the sequence. We can assume that the current and previous frames are available for the comparison. The simple comparison technique is too simple and is like a frame-by-frame DPCM. This has a few problems :

- The pixel compare will rarely produce a zero difference, due to quantisation noise in the system (this can be overcome with a threshold).
- Images are rarely stationery.

Therefore, more sophisticated techniques are needed. This problem is usually addressed by dividing the image into blocks. Each block is examined for motion. If a block is found to contain no motion, a code is sent to the decompresser to leave the block the same as the previous one.

If enough processing power is available, still more powerful techniques may be applied. For examples, blocks may be compared to previous block to see if there is a difference between the two. Only this difference (*motion vector*) is sent.

Need for Compression

Compression is needed to simply reduce the amount of space that video would otherwise take to store. There are many factors to consider when choosing a compression technique :

- **Real-Time / Non-Real-Time**

  This refers to capturing, compressing, decompressing and playing back all in real time with no delays. The requirement is to have sufficient frame rate (frames per second)to make sure that there is no jerky motion.

- **Symmetrical / Asymmetrical**

  Symmetrical implies capturing, storing, and playback at the same rate. Asymmetrical uses more time to compress and hence may have an advantage for playback speed.

- **Compression Ratio**

  The compression ratio relates the numerical representation of the original video in comparison to the compressed video. Generally the higher the compression ratio the poorer the video quality.

- **Lossless / Lossy**

  The loss factor determines whether there is a loss of quality between the original image and the image after it has been compressed and played back (decompressed). Again this is affected by the amount of compression.

- **Inter-frame / Intra-frame**

  Inter-frame compresses each frame of the sequence as a discrete picture. Intra-frame is a more powerful method which uses a predictive technique.

Compared to traditional analogue video, digital video provides the following advantages :

- There is no copy from copy loss
- Picture does not get fuzzy
- Signal-to-Noise ratio goes down slowly
- Editing, storage and retrieval is simpler, quicker and cheaper

When dealing with digital video a number of points have to be kept in mind :

- **Frame Rate**

  How many frames are displayed per second, also the method of frame display : *progressive* - each line of video is shown one after the other; *interlaced* - odd lines (fields) are shown then the even fields.

- **Colour Resolution**

This refers to the number of colours displayed at any one time. There are also various colour formats : RGB and YUV are two common formats. Colour depth is the maximum number of colours displayed.

- **Spatial Resolution**

  This deals with the size of the picture.

- **Image quality**

  Does the final sequence match the requirements of the application.

Video 'Standards'

With so many techniques, you would expect many companies to be competing for a position in the market place. This is in-fact the case and there are many competing technologies.

The above discussion of techniques and decisions introduced the building blocks available for creating algorithms. An actual algorithm consists of one or more techniques which operate on the raw digitised images to create a compressed bitstream. The number of algorithms possible is nearly infinite. However, practical applications (see below) require that all users who wish to interchange compressed video must use exactly the same algorithm choice. Further sophisticated algorithms will benefit from the development of special hardware. All this expresses the need for standards to allow the orderly growth of markets which utilise video compression technology.

Driven by these needs, there has been a strong effort to develop international standards for motion video compression algorithms, underway for several years in the International Standards Organisation (ISO) and the International Electrotechnical Commission (IEC). It is the Motion Pictures Expert Group (MPEG) which considers algorithms for motion video compression.

---

3. MPEG : The Standard

The MPEG committee began life in late 1988 by the hand of Leonardo Chairigloione and Hiroshi Yasuda with the immediate goal of standardising video and audio for compact discs. A meeting between the International Standards Organisation (ISO) and the International Electrotechnical Commission (IEC) in 1992 resulted in a standard for audio and video coding, known as MPEG-1. MPEG-2 became a bone fide standard in 1994 after a five day meeting of ISO and ITC in Singapore. The technology behind MPEG-1 and 2 are inherently the same.

The MPEG system consists of two layers :

- System Layer (timing information to synchronise video and audio)
- Compression Layer (includes audio and video streams)

Only the MPEG-1 standard will be described in detail here.

Objectives

The MPEG standard is designed to be generic, meaning that it will support the needs of many applications. The objectives include :

- Delivery of acceptable video quality at compressed data rates in excess of 1.0 Mbits/s
- Support both symmetric and asymmetric compression/decompression applications.
- Random access playback should be possible
- Fast-play and fast reverse in addition to normal playback should be supported
- Audio video synchronisation should be maintained.
- Data errors should not be catastrophic
- The processing requirements should not preclude the development of low-cost semiconductor based implementation which may encode in real time.

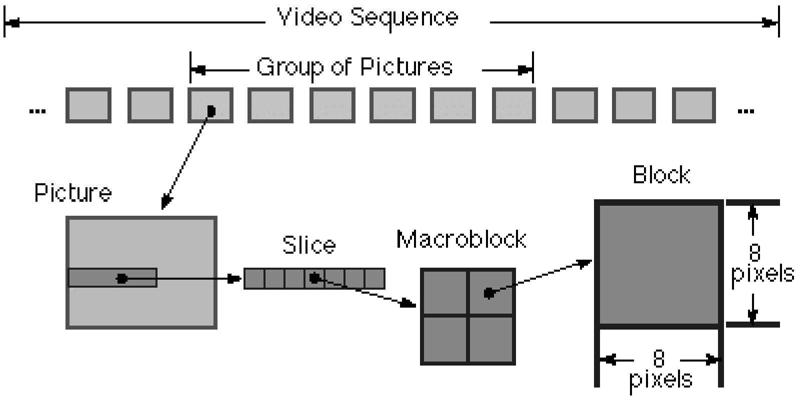

Figure 3.1 - The MPEG data hierarchy

Architecture

The MPEG standard is primarily a bitstream specification, although it also specifies a typical decoding process to assist in interpreting the bitstream specification. This approach supports data interchange, but does not restrict innovation in the means for creating or decoding that bitstream. The bitstream specification is based on a data hierarchy, shown in figure 2. The data hierarchy is pretty self-explaining and is useful for the following reasons :

- Groups of pictures allow random access into a sequence
- Slices aid error recovery, in that if one slice contains an error then it can be skipped.

The bitstream architecture is based on a sequence of pictures, each of which contains the data needed to create a single display-able image. There are four different kinds of picture, depending on how each picture is to be decoded :

- **I-Picture**

  These are intra-coded, which means they are coded independent of any other pictures.

- **P-Picture**

  These are predicted pictures and are coded using motion compensation from previous I or P-pictures.

- **B-Picture**

  These are interpolated pictures, which are coded by interpolating between a previous and a future I or P-picture. This process is called bi-directional prediction.

- **D-Picture**

These pictures are a special format and are used to implement sequence searches.

The above description of MPEG has been very terse and a lot of detail has had to be left out. However MPEG is not the only compression technique on the market, there are many algorithms available. The next section presents a brief overview of the competitors to MPEG and looks at the advantages and disadvantages for each method.

---

4. Contendors in the Compression Market

There are several popular digital video formats (or CODECS) in use today.

- **Apple QuickTime**

  This was a milestone for digital video. This is a software only playback scheme that requires no special hardware. This format supports scalability in that some quality loss can be sacrificed for the sake of synchronisation.

- **Intel Indeo**

  Shortly after Apples' QuickTime was released, Intel introduced Indeo. This also allows scalability. This format is well supported by all the major software vendors in the world and hardware version exist.

- **Radius Cinepak**

  This is another software only CODEC and again is well supported.

When choosing a particular format, it is worthwhile bearing in mind the following points :

- Is is easy to convert the source format (eg. Video tape)
- Will the format handle the source content
- What is the quality of the encoding system

With this in mind, the following table outlines the advantages and disadvantages of each digital format.

Format	Advantages	Disadvantages
Intel Indeo	• Uses an uncompressed audio format that does not need decoding • Inexpensive to compress analogue video • Video can be compressed at low frame rates and picture sizes • Supports low data rates	• When using low frame rate, the quality cannot be improved with faster hardware • Frame rates are slow
Cinepak	• Uses an uncompressed audio format that does not need decoding • Better colour resolution then Indeo	• Suffers from the same problems as Indeo • Generally lower frame rates then Indeo

QuickTime	• Uses an uncompressed audio format that does not need decoding • Has the same advantages as Indeo	• Requires apple platform for encoding and editing
MPEG-1	• Higher resolution and frame rates • Better scaling • CD Quality audio • Low data rates	• May require extra hardware • High cost encoding
MPEG-2	• Currently the highest resolution and frame rate available • Supports MPEG-1 format	• Higher data rates • Large storage requirements • Requires extra hardware for decoding • Very high cost encoding

It is very much obvious that all the formats have their problems. However, MPEG is currently the highest quality digital video CODEC around and hence will will be used applications requiring high quality video (see below).

---

5. Applications of Digital Video

Digital video has many and varied applications, here we briefly look at some applications. The number of applications is growing rapidly as the need for compression and digital transmission grows.

- **High Definition Television (HDTV)**

  HDTV is defined as having twice the horizontal and vertical resolution of conventional television, a 16:9 picture ratio and at least 24 frames per second. Using this definition, HDTV has approximately double the number of lines of current broadcast television. This combined with the resolution increase means that 6 times more bandwidth is needed for transmission.

  This is an ideal place for compression, as this will reduce the data rate and hence the bandwidth.

- **Presentations**

  This is the number one application for digital video. This application includes video kiosks, training, corporate presentations and video libraries. The advantages of using digital video (and particularly MPEG) are :

  - Footage can be updated or changed with ease
  - MPEG has network capabilities which means the presentation can be distributed
  - Digital video adds a whole new dimension to presentation. Moving pictures can be incorporated into computer presentations with ease.

  Multimedia used in student training has also been shown to improve achievement by an average of 38 percent.

- **Digital Video Databases**B

Since digital video clips are stored in files, they can be easily integrated into many databases just like text or numeric fields. For example, a travel agency can keep video clips of their holiday locations as well as more mundane information and really show what it is like to go for a holiday in a particular resort.

6. Problems with Digital Video

Distortions that get added to a video signal during digital encoding are known as *artifacts*. There are several types of artifact that explain the degradation in a video signal quality during digitisation. This section of the report will look at the various artifacts. These will be demonstrated by applying them the picture shown in figure 6.1. Please note that the comparison of these pictures is best done on a machine running at least 16-bit colour.

Figure 6.1 - 24-bit Colour Reference Image

General Problems

- **Aliasing**

  Aliasing occurs when a signal being sampled contains frequencies that are too high to be successfully digitised at a given sampling frequency. When sampled these high frequencies **fold back** on top of the lower frequencies producing distortion. In most methods of video digitising, this will produced pronounced vertical lines in the picture. This problem can be reduced by applying a low pass filter to the video signal before it is digitised to remove the unwanted high frequency components. This is tricky to do without removing some of the wanted high frequency components, and results in softer edges in the picture due to the slower permitted transitions in the signal level. See figures 6.2(a) and 6.2(b).

(a) **(b)**

Figure 6.2 - (a) Aliasing and (b) Effect of Low-Pass Filtering before Digitising

- **Quantisation Noise**

  This form of distortion occurs because, when digitised, the continuously variable analogue waveform must be quantised into a fixed finite number of levels. It is the coarseness of these levels that causes quantisation noise. A 24-bit colour picture (composed of an 8-bit value for each of the red, green and blue components of each pixel) suffers from virtually no quantisation noise, since the number of available colours is so high - 16.7 million. Reasonable results can be obtained from an 8-bits per pixel picture, especially if the picture is greyscale rather than colour. Figure 6.3 (a), (b), (c) and (d) show some examples of the same picture represented with varying colour resolutions.

(a) - Colour (b) - Colour

(c) - Greyscale (d) - Greyscale

Figure 6.3 - (a) 8-bits per Pixel, (b) 4-bits, (c) 8-bits and (d) 1-bit

- **Overload**

  Like quantisation noise, overload is related to the finite number of levels that the signal can take. If a signal is digitised that is too high in amplitude, then the picture will appear *bleached*. For example, if the signal level of a greyscale image is too high for the conversion process to cope with, then all levels above the maximum will be converted to white, causing the washed out appearance. Figure 6.4 shows one possible outcome of overloading the analogue to digital conversion process.

Figure 6.4 - Overloading During Conversion

Another possible result is known as wrap-around. This is where all out of range values are converted to the lowest value i.e. black. This can result in poor quality video and should be avoided if at all possible. See figure 6.5.

Figure 6.5 - Wraparound due to Overload

The actual outcome of overloading the converter depends mainly on the design of the hardware.

- **Digital Signal Degradation**

Video in digital form degrades far less gracefully than its analogue counterpart. While digital information may in theory be duplicated an infinite number of times without any degradation, once that degradation does occur, it is very noticeable. Due to the compression techniques used, a single bit error in the data stream could for example cause a large block of pixels to be displayed in a completely different colour to that intended.

Figure 6.6 - An MPEG video frame with multiple bit errors

Artifacts Caused by Compression

- **The Gibbs Effect**

One of the most common artifacts that afflicts both MPEG and JPEG compression is the Gibbs effect. This is most noticeable around artificial objects such as plain coloured, large text and geometric shapes such as squares. It shows up as a blurring or haze around the object, where the sudden transition is made from the artificial object to the background. It is caused by the discrete cosine transform used to digitise chrominance and luminance information. This phenomena is also apparent around more natural shapes like a human figure. The area of the background around the subject appears to shimmer as the subject moves slightly. This shimmering has been nicknamed mosquitos. See figures 6.7 (a) and (b).

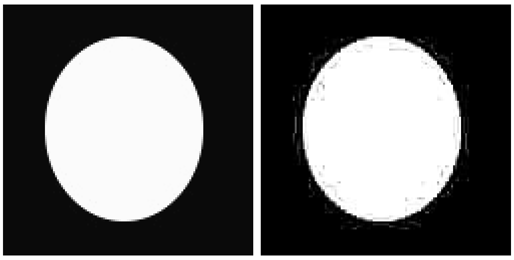

Figure 6.7 - (a) A Geometric Shape and (b) The Gibbs Effect

- **Blockiness**

  Another artifact that affects JPEG and MPEG is blockiness. When video footage involving high speed motion is digitised, the individual 8x8 blocks that make up the picture become more pronounced.

Figure 6.8 - Blockiness caused by Compression

- **Lossy Compression**

  A lossy compression method allows a system to produce much higher compression ratios. This removes some of the information contained in the signal, hopefully information that will go unnoticed. For example, an encoder may be designed with the criteria of providing output with say a 98% similarity to the input signal. Under most circumstances this may produce an acceptable picture, but if the video footage is a tennis match, then it may quite justifiably ignore the tennis ball (according to the encoding criteria) since it is so small! This kind of behaviour is obviously unacceptable, but lossy compression is very difficult to get right.

Implementation Problems

Both encoding and decoding of video information requires a significant amount of processing power. In general though, the encoding is far more demanding.

- **Compression and Encoding**

  For an author to tap the potential digital video market, he must transfer his video into a compressed digital form. Currently, there are three main methods of doing this:

  - B Use a service bureau. Several large companies including IBM offer the service of transferring analogue video into formats such as MPEG, Indeo and CinePak using 'high end' supercomputers and software which can achieve this in real time.
  - Purchase dedicated hardware, usually in the form of an expansion card for a computer. These can also encode in real time, but cost many thousands of pounds.
  - Use a video capture card to digitise the video into uncompressed form stored on hard drive, then use software based encoding for off-line compression. This is the cheapest option, though it is very slow. It can take 100 times longer to compress the information, even on the most powerful Pentium PCs.

- **Decompression and Decoding**

  Although dedicated hardware replay cards are available for certain digital video standards, all common formats can be decoded at a reasonable frame rate by a 100 MHz Pentium PC. Nowadays, this kind of machine is virtually entry level in the PC world, meaning that there are millions of users capable of replaying digital video material.

---

7. The Future of MPEG

The digital video market in which MPEG is a contender is not without competition. It has many competitors including Cinepak and Intel's Indeo. No single standard has yet attained supremacy in the marketplace. So far, two MPEG standards have been implemented (1 and 2) with support for multiple resolutions and channels of audio. By 1998, MPEG 4 will become a ratified standard for very low bitrate compression, increasing the range of applications to which the MPEG standards may be applied. Also the multimedia standard MHEG is currently being designed and will integrate a lot of media into one format.

With real time software decoding now feasible on most machines, and 'add-on' hardware cards available for the estimated 80 million legacy machines not powerful enough, MPEG has the potential to reach an extremely large market. Given a powerful PC, the quality of reproduction using MPEG is superior to any of its competitors. But Indeo and Cinepak do perform better on low-end machine. This causes an obvious split in the market. Most businesses involved in digital video appear to be 'sitting on the fence', waiting to see which way the market will go. The uptake of MPEG has not been as fast as some might have wished. But this is a problem for the whole digital video industry. It is a 'Catch-22' situation. Consumers will not buy digital video playing equipment without something to use it for, and suppliers will not provide their titles in digital form without a large, stable market in which to sell their products.B

A.1. Appendix One

This appendix lists any related articles produced by I.S.E for SURPRISE '96, references used in the project and further reading.

Related Articles

Our Articles

- Compression of Audio and Video Information (Stephen Done)
- MHEG - A Multimedia Presentation Standard (Stephen Done)
- Digital Video: An Introduction (Shanawaz Basith)
- MPEG : Standards, Technology and Applications (Shanawaz Basith)

Other Articles

- MPEG Image Compression and ATM Networks
- Interactive Television
- Technology and Clinical Applications

References Used

Internet

These references were obtained by using the Alta Vista search engine. All references in the sub-sections are listed by order of readability, usefulness, presentation and articulation.

Powerwebs MPEG site - Contains various links including the MPEG FAQ
MPEG Pointers and Resources (Tristan Savatier) - Contains a wealth of links to all the MPEG related sites
MHEG Information - A lot of information on MHEG
MHEG Related Links

Papers and Magazine Articles

D. Ruiu 'MPEG-2 Digital Video Technology and Testing' BSTS Solution Note 5963-7511E (Hewlett Packard)
R. Rubenstein 'Unleashing a Broadcasting Revolution' New Electronics on Campus, Autumn 1995
D. Thon 'Multimedia Design Reaches a Higher Level' New Electronics on Campus, Autumn 1995
D. Boothroyd 'Never Mind the Quality, Look at the Quantity' New Electronics on Campus, Autumn 1995

Books

J. Koegel Buford Multimedia Systems Addison-Wesley, 1994 - a very useful book, covers all aspects of multimedia

H-M. Hang and J. Woods Handbook of Visual Communications Academic Press, 1992 - a huge collection of papers, covering all aspects of the subject
J. Showrank Multimedia Exploration CPM Books, 1994 - covers a lot of issues concerning multimedia
R. Gonzalez and R. Woods Digital Image Processing Addison-Wesley, 1993 - a wealth of information about image processing and compression
A. Luther Using Digital Video AP Professional, 1995

Further Reading

For economy of space, the further reading section of Shanawaz Basith's second article has not been put here, please consult that document for those references. The format here has been somewhat tidied up.

J. Buford and C. Gopal 'Standardizing a Multimedia Interchange Format: A Comparison of OMFI and MHEG' 1994 IEEE Intl Conf on Multimedia Computing and Systems, May 1994
C. Gopal and R. Price 'Multimedia Information Delivery and the MHEG Standard' Massachusetts Telecommunication R&D Conference
R. Clarke Digital Compression of Still Images and Video Book - 1995
J. Watkinson Compression in Video and Audio Book - 1995

Acknowledgements

A special thanks to Dr J Barria for the kind support that he has given throughout this project and Dipan Patel for the advice and help. Finally thanks to Dr N Dulay for his efforts in co-ordinating SURPRISE '96, we hope it is as enlightening for future I.S.E students as it was for us.