Multimedia in the Teaching Space

Dr D G Jameson
University College London

Contents

1 INTRODUCTION

Information displayed and presented in multimedia formats can be a powerful tool in teaching and learning. This paper discusses how information can be introduced into teaching space over communications networks and how it can be displayed. The present-day desk-top or lap-top computer systems have sufficient power that they are capable of digitising media in the analogue form and manipulating this information to be displayed on suitable devices for use in teaching space. The most commonly available media are audio and visual information, and there are many techniques to digitise this information and then process and manipulate it. The term multimedia arises because different media can be used, multiple bit-streams of these media can be used simultaneously and synchronously, and these information streams can be merged or split as required.

Once digitised this information can be stored or transported over networks and thus it is available to be used in teaching and learning. The capacity of the computer systems which can be used in the teaching space is such that this information can be stored within a desk-top or lap-top computer and used in the stand alone mode or it can be stored on network servers, especially large amounts of information, and passed to the teaching space over suitable high speed networks as required.

The introduction of multimedia material and information into teaching space raises both pedagogic and technical questions. The way in which this digital information is converted back to the analogue form for display and presentation, and the nature of the user interface raises important pedagogical issues, especially if an interactive environment is required. Interaction must be considered carefully as the timescale of interaction can vary from the real-time to asynchronous situations, and these will place very different requirements on the computer systems.

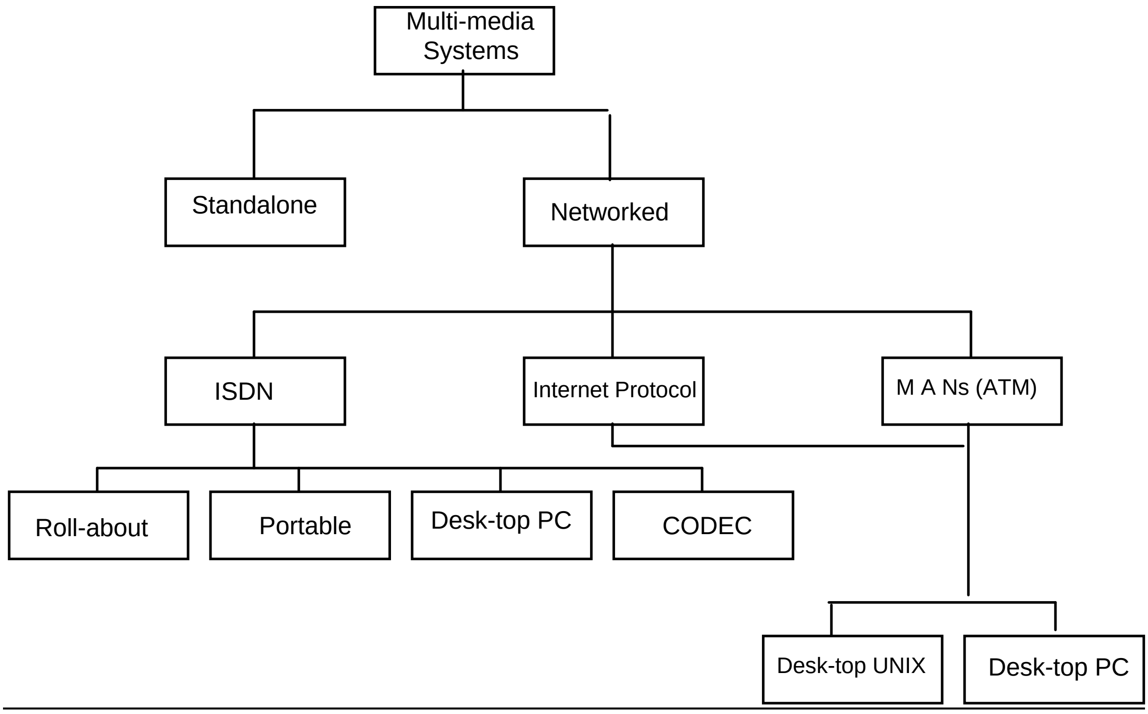
The production and storage of multimedia information is expensive because sophisticated equipment is necessary to digitise, encode and compress the information for communication across networks, and then decompress and decode the information to be able to hear the audio, or display in video. Therefore it is necessary for the benefits for teaching and learning to be justified, especially when it is still commonplace for academic staff in universities to continue to be satisfied with the traditional methods of blackboards, whiteboards and overhead projectors. Teaching applications using multimedia information will be produced and delivered through a collaboration of those people experienced in the technology and pedagogy such that the benefits are immediately apparent and attractive to the teachers. To bring about this change in attitudes the technologists have a delicate task to increase awareness and provide substantial support to the teaching staff at the present stage of the process of these new and innovative ideas and facilities.

The technological issues at one level are concerned with the delivery of multimedia material and information at a reasonable cost, combined with simplicity of operation. At another level the opportunities exist to be able to do things which have not been done before, and therefore to develop new and innovative ideas for teaching and learning. A good example is the opportunities and openings in distance learning, to attract to universities a spectrum of students which was not possible previously and to provide tuition and tutorials to students at times and places convenient to them, i.e. the teaching is being taken to the student as opposed to the student attending a teaching institution. The methods being developed for distance learning should also be able to bring benefits to conventional teaching.

In practice how multimedia information is introduced into teaching space is largely a question of what communications technology network is used. There may well be local issues involved in installing networks in teaching space, i.e. lecture theatres, seminar rooms and other specialised teaching rooms. The networks used to bring multimedia materials into teaching space can be expected to use a variety of different network technologies; e.g. ISDN and IP. Already in UK we have a number of experiments supported by JISC using different network technologies, e.g. the MANs which are high speed broadband networks, the INTERNET and the public networks such as ISDN.

One precaution to be recognised for teachers is that introducing a wide variety of multimedia materials into teaching can provide a multitude of options to the student and teacher which do nothing to enhance the learning process; there must be a balance of the effort put into producing multimedia materials and their use, otherwise the technology intrudes into the teaching and learning processes.

The following logical approach should help the user to understand the options available.



Various claims have been made about multimedia information and communications, and it is worth quoting the statement of Nicholas Negroponte, of the Media Lab, at MIT (Massacheusect Institute of technology) when he said "multimedia is the slayer of boredom, the seducer of the senses and the arch nemesis of the 'been there, done that' attitude". Multimedia information is more interactive than print and today's student must beside being equipped with the basic skills of learning, must also learn new skills to be capable of dealing with the information rich society. The teacher and student can interact using more flexible tools, but the essential information in teaching and learning must be decided by the teacher, hence the performance of the technology must develop in such a manner that there is a sympathy and understanding between the teacher and the technologist.

2 WHAT IS MULTIMEDIA

2.1 DEFINITIONS

The title of this document contains two words which need to be defined; the term ‘multimedia’ has a variety of meanings to different people and an attempt is made in the following section to define the term within the ideas of this document.

The term ‘teaching space’ is basically self explanatory, but is not frequently used, and so a definition has been given.

2.1.1. Multi-Media

There is common confusion due to the different meanings of the term “Media”, which can refer to communications industry such as the press, newspapers and television, or it can refer to the plurality of medium by which information is conveyed to and between humans. In this latter case the information is conveyed by text, sounds and vision, touch and smell. The computer has made it possible to handle several of these medium simultaneously, in particular those which can be digitised easily.

Documents can be scanned or typed into electronic formats which can be handled in sophisticated ways, altering the layout, font styles and sizes etc. with ease. These documents can include pictorial and graphic information. Pictures and graphics can be digitised and handled either as bitmaps or as vector drawings. The JPEG standards exist for bitmap images. Animations can also be produced to support text information.

Visual information is very flexible and if the time factor is taken into account moving images can be handled as well as still pictures. Much recent research has been carried out into the handling of moving images, and when this information is synchronised with digitised audio information video is available. This technology is covered by a number of MPEG standards, some hybrids such as Motion-JPEG and the H.26X standards used in video-conferencing.

Multimedia has become a generic term for “multimedia computing” or “interactive multimedia”. The computer and its software are used to control and navigate through the communications medium, not only one at the time, but several simultaneously which simulates the real-world and presents unique and innovative opportunities to captivate human senses. Computer systems are most developed in using vision and hearing to interface between the digital and analogue worlds, e.g. still and moving images, text and graphics use the visual senses, audio uses hearing. In this paper “Multimedia” is defined as visual, audio and textual information which can be presented separately or simultaneously to convey and present information interactively to users. Multimedia is now possible because it is technically easy to digitise the analogue forms of these common media and handle this by computers which are easily available and which are small enough to be used on the desk-top.

The applications of multimedia are various and include virtual reality and 3-D presentations.

2.1.2. Teaching Space

Teaching space is space within a teaching environment where teaching and learning can take place. It would include lecture theatres, laboratories, seminar rooms and tutors rooms where tutorials might take place. The size of these room can vary considerably and the size can pose problems when presenting multimedia materials, especially in large lecture theatres. This therefore also includes learning centres and libraries where the student may be working alone, because these locations would be within institutional space and under some kind of supervised control. A liberal definition of teaching space would include the students own living quarters, whether it is in a home or student residence. This discussion will deliberately not included that situation.

If multimedia material are to be available in the teaching space then in most cases a computer, lap-top or desk-top should be available to be operated in the space. If legacy audio-visual networks are available, then the computer is not necessary, as the information is delivered in analogue format and will be displayed without having to be digitised. The analogue signals are capable of producing high resolution images, and do not suffer from the delays introduced by the compression and encoding procedures.

However the signals cannot be manipulated easily in the analogue form, usually requiring dedicated hardware. If the signals are digitised then a large variety of digital tools can be applied, and a more flexible system is made available to the teacher. One of the problems which arises from introducing greater flexibility of presentation into the teaching space is that the teacher's attention is diverted away from teaching into managing the various devices and facilities. Experience has shown in the Teaching and Learning Technology Project, INSURRECT (http:/www.mmscc.ucl.ac.uk/INSURRECT/) for teaching undergraduate surgery on the SuperJANET ATM video network that the students are quickly aware that the teacher's attention is divided and would prefer their attention to the technology.

As will be indicated elsewhere multimedia material can be presented in the standalone mode or it can be networked. The standalone mode gives limited opportunities, generally due to the limitations in the storage capacity of the computer, and the speed with which information can be retrieved from floppy disk or CD-ROM. Greater flexibility comes with the use of networks, and the higher the bandwidth of these networks the more sophisticated the multimedia presentation can be, and the images have higher resolution.

2.2 THE ANALOGUE LEGACY NETWORKS

Most people assume that a computer is essential to introduce multimedia information into teaching material, but moving and still images and audio have been transferred over audio-visual analogue networks direct into lecture theatres for some years, since the mid 1970s. These applications are sometimes called legacy networks and work is being carried out to link these analogue networks with new high speed broadband networks. As these legacy network were analogue they supplied high resolution images which were near broadcast quality.

Examples of these analogue video networks are LIVENET (London University Interactive Video Network for Education, established in the early 1980s) and the Charing Cross Hospital Video Teaching Network. These network used optical fibre to carry video and audio signals over limited distances up to about 25 miles, and have been used for teaching for 10-15 years. Their use of analogue technology in fact gave better resolution than the early ISDN networks, and for that reason were favoured for medical applications.

In the late 1980s Sony and other video disk manufacturers produced systems that did not require pre-mastering. The SONY LVR laser video recording system was available in 1989 and was used to store and replay still and moving images (with audio) into lecture theatres under the control of the teacher. The LVR was located outside the lecture theatre and remotely controlled. By using the INTERNET signals were sent from other parts of Europe to control video being played over satellite. This system is still in use on the LIVENET network which is used for teaching in University College London, and handles up to 1000 hours of teaching per annum in a variety of subjects; including Medicine (Surgery), Classic, and Physics. Likewise in the SuperJANET video demonstrators and teaching programme video was transferred over SuperJANET to a number of remote sites, with the teacher at one of the sites controlling the system.

These systems used special video disks cartridges which allowed images, slides etc. to be laid down on the disc in real-time. The output from a video camera could be fed directly to the LVR recording unit and recording made onto pre-determined frames on the disk. The playback unit could then be set to the requisite frame numbers and the video played back to a monitor. This was a WORM (Write Once Read Many times) system and so the recording was available for some time. This was not a digital system and the quality of recording did deteriorate with time, due to deterioration in the surface of the disks. It was possible to record video directly from camera, or from previous recording on cassette or other video disks. The system was also able to input still images to a specified frame number and replay as selected. Thus a mixture of still and moving images could be stored on the disk cartridge to a maximum of 54,000 frames which corresponded to 36 minutes of video material. The LVR video disk system could be controlled remotely using RS232A protocol, which could be delivered over the INTERNET. The command set permitted the following commands:-

START

STOP

PAUSE

PLAY FORWARD

PLAY REVERSE

PLAY frame number XXX to YYY (Video sequence)

PLAY FAST

PLAY SLOW

SCAN FORWARD

SCAN REVERSE

DISPLAY frame number (Still image)

This command structure permits interaction between the user and the images being displayed by the video disk. The system recorded video in PAL and NTSC formats

2.3 DIGITAL SYSTEMS

2.3.1. The Multimedia Computer

It is now possible to digitise video and audio with desk-top computers such as the PC or Macintosh, which makes possible the handling of multimedia material which can be used in support of teaching. The resolution of the video and still images is now available at SVGA and even XVGA quality which results in images of very high quality, which are better than that seen on domestic television. If this multimedia material is to be displayed in the teaching space, often to several students at once, then large displays are necessary. This is achieved by the use of video projectors with computer input, by using LCD projectors displays or large television monitors connected to the computer by scanning devices. These scanning devices can handle high resolution, higher than that produced by PCs and their cost is affordable. Until recently this hardware was expensive and the result were barely acceptable at VGA level, e.g. the Mediator.

As multimedia computer systems handle digitised audio and video they need to be powerful systems and will commonly be using a fast Pentium processor. A multimedia computer system will be comprised of the following components:-

Powerful high speed central processor with clock speed of over 200 Mhz.

Random Access memory 32 Mbytes

Hard disk drive 2-5 Gbytes

Input through Floppy Disk Drive or CD-ROM

Audio input and output devices, e.g. microphones and loudspeakers

Video input and output devices, e.g. VCRs, video-cameras, video disks etc.

Still image input and output, e.g. slide scanner, monitors etc.

Graphics capability to handle both bitmap and vector graphics

Display of visual output on SVGA/XSVGA quality monitors

User input through Mouse, tracker ball etc.

CODECs - hardware but may become software in the near future.

The essential capability of any multimedia computer system is the ability to convert the analogue signal to a digital format and using standard algorithms compress this information. The power of the CPU will determine whether this process can be carried out in real-time or whether it has to be done off-line. Originally this power only existed in UNIX systems, but more recently it has become available in PCs and Macintosh. The compression process is necessary otherwise the quantity of data to be stored and transmitted would be excessive. Most PCs have until recently depended upon hardware encoding/decoding systems (CODECs), but the speed of current processors is such that software compression systems have been developed. The advantage of the software systems is that the compatibility and interoperability issues can be handle more easily, and the cost of the equipment is not raised by the need to purchase expensive hardware devices.

2.3.2. Audio Compressed File Formats

The computer system must handle compressed file formats. For audio common types of format include:

- WAV- a digital sound file for Windows.
- MIDI (Musical Instrument Digital Interface) - audio files created by connecting the computer to musical instruments and control devices.
- Real-Time Audio - supercompressed digital audio information.

2.3.3. Video Compressed File Formats

Two types of file are used to store still images:-

JPEG - designed for compressing either full-colour or grey-scale images.

GIF - designed for compressing images with a few distinct colours and for line drawing and simple cartoons. GIF files can only handle 8-bit colour.

In the video domain it is desirable to have full-screen, full motion video straight form the video source, but this would require a data-stream with a bandwidth of 27Mbps, which is significantly higher than most PCs can handle.

Popular video files are Quicktime (.MOV) for the Macintosh and Video for Windows (.AVI) for PC systems. Both these compression files are proprietary system and it is more advisable to use the common standards such as H.261, H.263, MPEG-1, MPEG-2, MPEG-4, Moving-JPEG etc. (See below for more detailed discussion of these standards).

2.4 THE COMPRESSION STANDARDS.

This section briefly discusses the principle compression standards. A more detailed discussion takes place in section 2.7.

2.4.1. Video Standards

The acronym MPEG stands for the Moving Picture Experts Group, which meets to generate standards for digital video (sequences of images in time), and audio, and for the development and defining of the conditions for compression algorithms. The standard defines a compressed bit-stream, which implicitly defines the function of the compressor. The compression algorithm, design by individual manufacturers must meet certain conditions laid down by the standard and this has led to the development of proprietary algorithms optimising the characteristics of a particular manufacturer’s equipment. These proprietary algorithms have largely merged into the background and everyone produces systems that conform to the MPEG standard, and as a result compatibility between different systems has improved.

H.261 and MPEG-1:

H.261 is the standard used in most ISDN Videoconferencing systems and handles the bandwidth range form 64Kbps to 2Mbps. This range overlaps with the MPEG-1 standard which was developed for encoding video to be stored on CD-ROM. The common data rate for CD-ROM is 1.2Mbps which is higher than that used in Videoconferencing, i.e. 128 and 384Kbps. Visually the quality of video from the MPEG-1 and the H.261 standards are comparable when using the same bandwidth.

MPEG-1, MPEG-2 Motion-JPEG and MPEG-4:

Video digitisation uses the standards MPEG-1, MPEG-2, Motion-JPEG, MPEG-4 and near standards such as AVI. It is now possible to have both hardware and software encoder and decoders. This results in simpler hardware for terminal as it is not necessary to have special cards inserted in the computers to be able to handle these various standards. MPEG-1 equipment is cheaply available, but the other standards are more costly to implement. The user needs to appreciate that it is costly to encode material in the MPEG-2 standard, but the decoders are cheap because this standard has been developed with video-on-demand in mind. The Motion-JPEG requires considerable computing power because it is digitising each video frame in real-time, i.e. pictures are being encoded at 25 frames per second.

Fullscreen in real-time, is now possible, but it must be remembered that the encoding and display of information is a 2-stage process. In standards like MPEG-2 where the encoding is slower and more complicated it may be necessary to carry this out off-line

2.4.2 Audio Standards

These refer to the audio standards for the compression/decompression and transmission of audio signals. G.xxx classifications gives these standards.

There are 3 principal standards which are most commonly used, namely G.771, G.722, and G.728

G.711 A standard for compressing and decompressing audio (50 -3000Hz) into a 48, 56, or 64Kbps stream.

G.722 A standard for compressing and decompressing audio (50 - 7000Hz) into a 48, 56, or 64Kbps stream.

G.728 A standard for compressing and decompressing audio (50 - 3000 Hz) into a 16Kbps stream.

G.728 is particularly interesting as it uses on 16Kbps of the available bandwidth compared with 48Kbps in G.711 and G.722. This means that more of the bandwidth can be devoted to video and the picture quality is therefore more acceptable. The proportionate effect is greater at 128Kbps than at higher bandwidths.

This is all brought together in the Videoconferencing standard H.320 AND H.323 which will be dealt with in a separate section.

2.4.3. Still Picture Standards

JPEG

This is a standardised image compression standard from the Joint Photographic Experts Group. This standard is designed to compress either a full-colour or grey-scale images of the real-world. It works well on photographs and artwork, but not so well on lettering, simple cartoons and line drawings. It only handles still images. It is a “lossy” standard in that the compressed image is not exactly the same as the original. This standard deliberately exploits characteristics of the human eye, in that it is more sensitive to changes in the brightness of colours than in changes of colour - i.e. it is design to produce compressed images to be looked at by humans.

An important aspect of JPEG is that the decoders can trade off decoding speed against image quality.

GIF

GIF can only store images with 8 bits per pixel, so its is suitable for inexpensive computer displays. As full colour hardware is becoming cheaper then GIF should be less in demand, but for certain types of image GIF is superior in image quality, and file size.

2.5 CD-ROM FORMATS

CD-ROM is a means of delivering multimedia to the desk-top, and into teaching space and is suitable for stand-alone systems. This medium is able to handle audio and video, and the disks can contain large amounts of data as well as. Technological advances mean that it is possible to write on CD-ROM using equipment little more expensive than the computer itself. Lap-top computers now have CD-ROM drives and small loudspeakers which make these system highly portable. This portability can be very valuable in teaching as the technology can be carried into the lecture theatre when it is required, which does away with the problems of keeping expensive computer hardware permanently in lecture theatres and seminar rooms. This is particularly convenient as many lecture theatres are not wired to local networks.

Browsing the World Wide Web (WWW) with Netscape and the other available web browsers, has given students the opportunity to work alone accessing very extensive sources of information. The problem now is not whether the student has access to information, but how the student can assess the quality of the information and cope with the volume of information they can so easily access. The WWW is here to stay and we are still learning new methods of using this technology, which is accessible from many desk-top computers. The CD-ROM holds selected information, very often by an expert in the field and with a mind to an academic course. The WWW is “indiscriminate” information which has been made available because someone considers it will be useful to someone else, and the use envisaged in many cases is not related to academic pursuits at all; e.g. suppliers wishing to give detailed information about their products and how they function. The relative roles of CD-ROM and the WWW will be interesting to watch in the near future.

The original single speed CD-ROM drive had transfer rates of 150 bytes per second, which corresponds with 1.2Mbps which is the speed of the MPEG-1 standard. Today’s CD-ROM drives are much faster with transfer rates as high as 1.5Mbytes per second, i.e. 12Mbps. The throughput is not the only measure of playback speed in a random access system. The size of the cache will have some influence on the effective transfer rate, and the access time is measured by the seek time and the latency; the seek time measures how long it takes to find the appropriate track and the latency how long it takes for the start point to rotate under the read-head. A CD-ROM drive is a constant linear velocity device, which means the read head must speed up to read the out sector of the disk, and slow down to read the inner tracks.

CD-Recording systems (CD-R) have become cheap and common; they are invaluable because they provide the ability to prototype to see how a product runs at CD speeds, which is important with interactive systems, Also they give the ability to produce a one-off disk quickly.

The CD-formats that a user should expect CD-ROM software to support should include DOS or Windows compatible disks, HFS (Macintosh files structures), CD-audio and mixed modes, e.g. audio plus data. It is possible to create disk for Mac and PC systems, where the same dataset is to be read on Mac or PC as the directory structures are set up for each platform.

There are other formats such as CD-I (CD Interactive), and CD-ROM XA, an extended architecture which combines CD-data with interleaved compressed audio or video. Other include the Kodak

Photo-CD and Video-CD. CD-R software should support multi-session writing which means the recording of the disk can be carried out putting down different tracks at different times.

2.6 SCANNERS or SCAN CONVERTERS

Visual information is presented on a computer as pictures made up of picture elements (pixels) sufficiently small that each individual element is not easily visible. The resolution, quality or detail of a picture depends upon how many pixels are used. A television picture however is made up of lines of information and is not divided into pixels. When the computer derived image is presented upon a television monitor there is a compromise between the two technologies. It is recognised that an image derived from a certain density of pixels will appear very similar to the television line picture. The pixels are usually expressed in terms of the number in a horizontal-vertical array; for example SVGA is defined as 800 x 600 pixels, which converts into the 650 line television image with lines to spare. For this reason the SVGA image usually is cut off at the bottom when displayed on a television monitor. The aspect ratio (the ratio of height to width of the picture) of the SVGA image is different to the television picture which also explains why the images are not simply displayed.

To display this image a scan converter can be used, and now scan converters are becoming available which will transfer the images from high quality computer displays such as the Sun and Silicon Graphics to the television monitor. In these latter cases there will be some loss of resolution as information has to be thrown away to display on a television monitor.

Older equipment would have worked with VGA resolution, and this is the case with much video conferencing equipment. The higher resolution computer displays require more information and therefore more bandwidth, which is not available on ISDN and current IP systems.

In the audio-visual and communications industry, we have seen the transformation of a PC computer with the capability of doing number crunching using a spread sheet, record keeping with the data base and glorified typing using a word processor to the advent of the powerful workstation that appears limitless. We have also seen the introduction of the multimedia PC allowing the gap between computer and video to be closed further.

Equipment is available which allows computer output to appear on a single scan monitor or television. With multimedia information, which can be processed on desktop-computers, systems are required which will permit the multimedia output to be presented on large screens for use in teaching space. Only one or two people can easily observe a computer display at the same time, but in teaching situations, a significant number will need to observe the display screen, and this can be done cheaply by using television monitors or projectors.

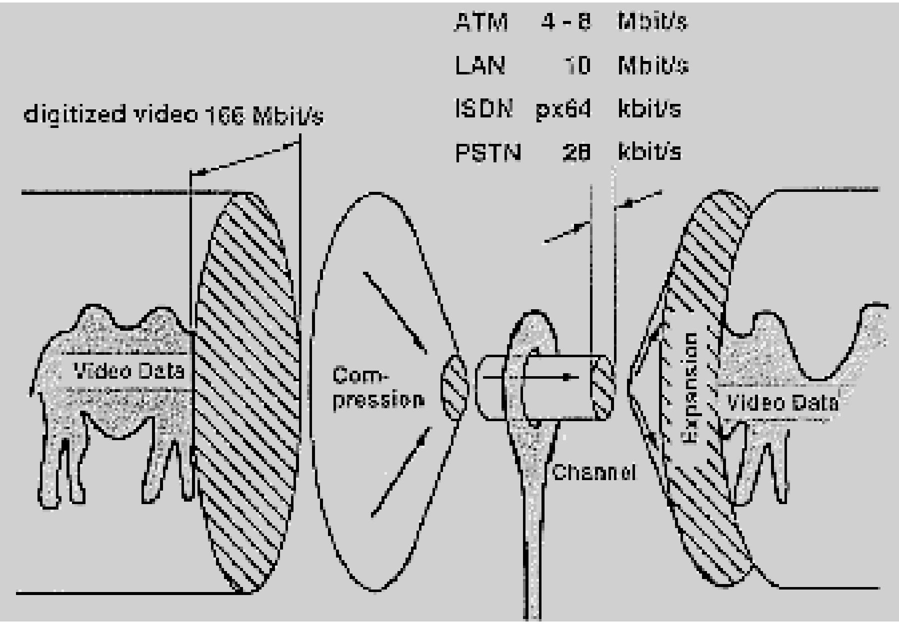
Scan converters can be used to convert the signal from a computer for display on an NTSC or PAL consumer television/composite video monitor. If a resolution of 640x480 pixels is adequate, the MAC or VGA screen from a PC can be displayed on the monitor. There are also scan converters designed to deal with higher resolutions up to 1024x768 from a SUN and Silicon Graphics terminal at 48Khz. The NTSC or PAL Monitor must have a video input jack (line video like the output of a VCR). These devices can be situated at several feet distance from the monitor allowing their use in teaching space. In some cases the signal can be extended for a long distance with very little visible degradation. In other cases the distance can be very short.

It is possible also to link the MAC or PC to an Liquid Crystal Display (LCD) panel as well as to television monitors. As LCD panels can be portable and provide a bright image it is now increasingly common for them to be used in lecture theatres and seminar rooms.

2.7 DIGITAL VIDEO COMPRESSION

2.7.1 Overview

Digital video compression is one of the key issues in video coding, which enables efficient interchange and distribution of visual information. New applications in the field of communication, multimedia and digital television broadcasting require highly efficient and robust digital; video compression and encoding techniques. The integration of motion video as an integral part of the multimedia environment is technologically one of the most demanding tasks due to the high data-rates and real-time constraints.



The rate required to encode a full-motion video signal at VHS quality had come down from 20Mbps to well below 1Mbps . For a typical head and shoulders video conferncing application the data rates are substantially lower e.g. 0.128Mbps.

The common video compression standards currently available are MPEG-1, MPEG-2, H.261, and H.263.

MPEG-1 refers to the delivery of video for a CD-ROM quality presentation.

MPEG-2 refers to broadcast quality compressed video, and would work with HDTV.

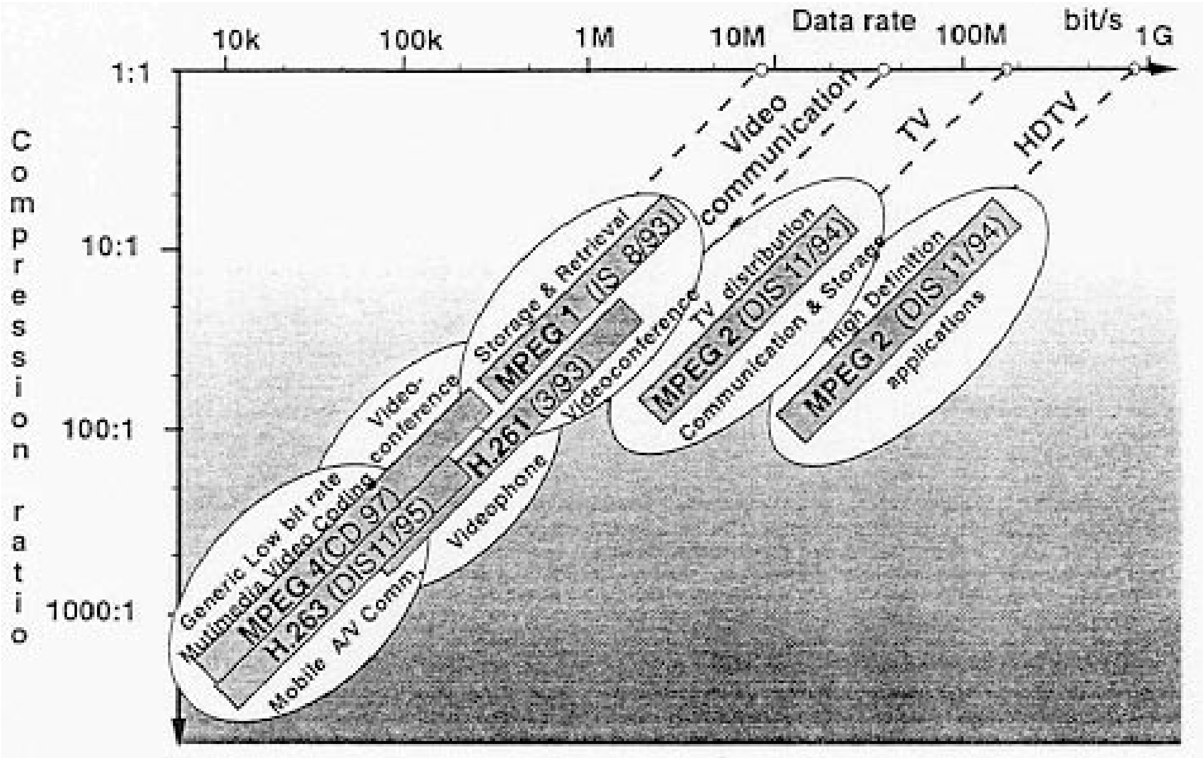
MPEG-3 was targeted for HDTV, however it was discovered that MPEG-2 could handle HDTV.

H.261 is the most widely used international video compression standard, but as the bandwidth of communications links increases it is likely to be superseded. The standard cover the bandwidth from 64Kbps to 2Mbps.

H.263 coding standard is oriented to videoconferencing, and is a descendant of the motion compensated DCT methodology used to support the existing standards H.261, MPEG-1 and MPEG-2.

H.263 has emerged to a high compression standard for moving images, and does not focus exclusively on very low bit-rates. The improvements in H.263 compared with H.261 are mainly obtained by improvements to the motion compensations scheme.

Technology developments show an increase in interactivity in all kinds of application, especially audio-visual ones. Currently much interactivity is restricted to synthetic content, and natural audio-visual content is not supported by existing standards, nor is the combination of natural and synthetic content. Future multimedia applications will require the implementation standards to cover these types of application.



2.7.2 Compression Algorithms:

There are a number of standard in common use in the multimedia and videoconferencing domain. These will be discuss in turn as follows:

2.7.2.1. H.261

This is a video coding standard published by the ITU (International Telecom Unit) in 1990. It was designed for data-rates which were multiples of 64Kbps. These data-rates are suitable for ISDN(Integrated Systems Digital Network) lines. H.261 is the most widely used international video compression standard, but as the bandwidth of communications links increases it is likely to be superseded. The standard cover the bandwidth from 64Kbps to 2Mbps. H.261 is used in conjunction with other control standards such as H.221, H.230, H.242.

Picture Formats Supported

Picture format	Lumin pixels	Lumin lines	H.261 support	H.263 support	Uncompressed Bit-rate Mbps			
					10 frames/s		30 frames/s	
					Grey	Colour	Grey	Colour
QCIF	176	144	Yes	Yes	2.0	3.0	6.1	9.1
CIF	352	288	Optional	Optional	8.1	12.2	24.3	36.5

The encoder on works on non-interlaced pictures. The pictures are coded in term of their luminance and two colour components. H.261 support two image resolution QCIF which is 144 x 176 pixels and optionally CIF which is 288 x 352 pixels.

2.7.2.2. H.263

This is a video coding standard and was published around 1995-6 It was designed for low-bit-rate communications and early drafts specified data-rates less than 64Kbps. However this limitation was removed, and it is expected the standard will be used over a wide range of bit-rates, and it will eventually replace H.261.

H.263 differs from H.261 in the following ways:-

It uses half-pixel precision for motion compensation where H.261 used full pixel precision. Some parts of the hierarchical structure of the data-stream are now optional, so that the CODEC can be configured for a lower bit-rate., or better error recovery.

There are four negotiable options to improve performance;

1. Unrestricted Motion Vectors
2. Syntax-based arithmetic coding
3. Advanced prediction.
4. Forward and backward frame prediction similar to MPEG.

H.263 is supposed to be capable of providing the same quality at half the bit rate to H.261. Also H.263 support 5 resolutions which enables it to compete with the MPEG standards:-

SQCIF

QCIF

CIF

4CIF

16CIF

Picture Formats Supported

Picture format	Lumin pixels	Lumin lines	H.261 support	H.263 support	Uncompressed Bit-rate Mbps			
					10 frames/s		30 frames/s	
					Grey	Colour	Grey	Colour
SQCIF	128	96	Yes		1.0	1.5	30 0	4.4
QCIF	176	144	Yes	Yes	2.0	3.0	6.1	9.1
CIF	352	288	Optional	Optional	8.1	12.2	24.3	36.5
4CIF	704	576		Optional	32.4	48.7	97.3	146.0
16CIF	1408	1152		Optional	129.8	194.6	389.3	583.9

2.7.2.3. MPEG-1

This standard is currently in 5 parts:-

1. Systems - addresses the problem of combining one or more data streams from the video and audio parts of the MPEG-1 standard with timing information to forma single stream. Thus once combined into a single stream the data are well suited to digital storage and transmission.
2. Video - specifies a code representation that can be used for video sequences to bit-rates around 1.5Mbps. this was developed to operate principally from storage media offering a continuous transfer rate of 1.5Mbps, but it can be used more widely, because the approach is generic.
3. Audio - Specifies a coded representation for compressing audio sequences (mono and stereo). A psycho-acoustic model creates a set of data to control the quantifier and coding.
4. Conformance testing - specifies how tests can be designed to verify whether bit-streams and decoders meet the requirements specified in parts 1, 2 and 3 of the MPEG-1 standard.

   These tests can be used by the manufacturers of encoders to verify whether the encoder produces valid bit-streams.

   These test can be used by manufacturers of decoders to verify whether the decoder meets the requirements set out in parts 1, 2 and 3.

   These test can be used to test application to verify whether the characteristics of a given bit-stream meet the application requirements.
5. Software simulation - is technically not a standard, but a technical report, give a full software implementation of the first 3 parts of the MPEG-1 standard.

2.7.2.4. MPEG-2

This standard is currently available in 9 parts:-

1. Systems - Addresses the combining of one or more elementary streams of video and audio, as well as data into single and multiple streams suitable for storage or transmission. This is specified in two forms; the Programme Stream and the Transport Stream, each of which is optimised for a different set of applications.
2. Video - Builds on the powerful compression capabilities of MPEG-1 to offer a wide range of coding tools.
3. Audio - is a backwards compatible multi-channel extension of the MPEG-1 Audio standard.
4. Compliance testing - corresponds to Part 4 of MPEG-1.
5. Software simulation - correspond to Part 5 of MPEG-1.
6. Extensions for DSO-CC (Digital Storage media Command and Control) is the specification of a set of protocols which provide control function and operations specific to managing MPEG-1 and MPEG-2 bit-streams. These protocols may be used to support both stand-alone and networked environments. software implementation.
7. Multi-channel Audio coding - this is not constrained to be backwards-compatible with MPEG-1 Audio.
8. 10-bit Video - this was withdrawn when there was insufficient interest.
9. Extension for real-time interface for system decoders.

2.7.2.5. MPEG-4

This standard has been developed in a period when the technology is changing rapidly, and hence it has been difficult to have a clear view of its scope. It is anticipated that MPEG-4 will :-

- address a range of applications that represent future markets
- take into account developments in the standardisation field
- address new trends in human computer interaction
- recognise the need to support communication

MPEG-4 can supply an answer to the emerging needs of applications ranging from interactive audio-visual services to remote monitoring and control. It is the first standard that is trying to take account of interactivity rather than simply determining how moving and still pictures should be transmitted and displayed.

MPEG-4 is essentially a multimedia standard integrating various communications media such as:-

- mono, stereo and multi-channel audio
- mono, stereo and multi-channel video
- 2D and 3D presentation etc.

This global approach to the use of audio and video, seeking to merge three worlds, namely the mobile communications domain, the digital television and film domain, and the interactive computer and human interaction domains. The existing standards do not cover these areas because their bit-rates are too high or the audio-visual standards do not exist. This standard seeks to cover a number of new functionalities which are divided into three groups:-

<u>1. Content based interactivity:</u>

MPEG-4 should provide access and organisation tools for audio-visual content, which may be indexing, hyperlinking, querying, browsing, uploading and down loading and deleting.

MPEG-4 should provide a syntax and coding schemes to support bit-stream editing and content manipulation to select one specific object.

MPEG-4 should provide methods fro combining synthetic and natural scenes, and is the first step towards the integration of all kinds of audio-visual information.

MPEG-4 should provide methods to randomly access, within a limited time and with high resolution parts form an audio-visual sequence.

<u>2. Compression:</u>

MPEG-4 should provide subjectively better audio-visual quality compared with existing standards and associated bit-rates.

MPEG-4 should provide ability to code multiple views/soundtracks of a scene as well as synchronisation between the elementary streams. This would handle stereoscopic and multiple view of the same scene, and take into account virtual reality and 3D requirements.

<u>3. Universal Access:</u>

MPEG-4 should provide an error robustness capability, particularly for low bit-rate applications.

MPEG-4 should achieve scalability with a fine granularity in content, spatial resolution, temporal resolution, quality and complexity.

If MPEG-4 can achieve this functionality it will provide better audio-visual quality at comparable bit-rates compared to existing standards. As mentioned above MPEG-4 seeks to take into account current and future audio-visual applications; these applications can be considered under three criteria:-

1. Timing Constraints

Applications should be classified as real-time on non-real-time. In real-time applications the information is acquired simultaneously, processed, transmitted and used in the receiver.

2. Symmetry of Transmission facilities

Applications are classified as symmetric or asymmetric. Symmetric applications are those where equivalent transmission facilities are available at both side of the communications link.

3. Interactivity

Applications are classified as interactive or non-interactive. Interactive applications are those where the user has individual presentation control, either controlling the digital storage media or also on the scheduling sequence of the information flow.

These criteria lead to eight classes of application as follows:-

<u>Applications Class 1</u>

The user has presentational control, either at the level of user control of digital storage, or also on scheduling and the sequence of information flow.

1. Video-telephpony - conversations person-to -person without any limitation in the scene environment and the network used. It may include some data transfer capabilities e.g. from auxiliary still pictures, document etc.
2. Multi-point video-telephony - inter personal communication between more than two people, each on in a different place, and the control may be either in the network or in one of the terminals.
3. Videoconferencing - interpersonal communication involving more than one person in to or more connected places. Videoconferencing often takes place in a room environment or a private office setting.
4. Co-operative Working - involves simultaneous inter-personal interaction (video-telephony) and data communications, at least as important as the audio-visual information.
5. Symmetric Remote Classroom - which is very close to video-telephony or Videoconferencing, but it differs in that the audio-visual content may include scenes without a speaker, and there is usually a fixed central control point which in multi-point connections sends it scene to all sites, and selects to receive one from everyone.
6. Symmetric Remote Expertise - experts are consulted from a remote location (e.g. telemedicine).

<u>Applications Class 2</u>

These applications are interactive. The interactivity may be user-to-user, user-to-machine or machine-to-user. The interactivity may also be conversational and user remote control through a return data channel.

1. Asymmetric remote expertise - Experts are consulted from a remote location, where the links are asymmetric in that there is one audio-visual channel and one Audio-only channel. Discussion can take place and the expert shows an illustration, or vice-versa the expert is able

to view conditions at the remote site and discussion takes place over return Audio channel. Both conditions would be applicable in telemedicine.

2. Remote Monitoring and Control - here audio-visual data is collected in real-time from a remote location, typically through a machine-user communication interface. There is typically an audio-visual channel from the remote location, and an audio or control channel in return. This would include remote control of cameras and/or microphones and would include traffic or building monitoring.
3. New gathering - new is collected from remote places where it is difficult to establish quickly a good quality connection. The interaction level is low and typically limited to a return audio channel, e.g. a direct broadcast from a remote site.
4. Asymmetric Remote Classroom - this is the situation where there is a central site sending audio-visual information and receiving sites, usually less expensive to set up, have the possibility of participating through an Audio channel.

Applications Class 3

These application are interactive in that the user can control the data flow through a control channel.

1. Multimedia messaging - messages containing text, audio, video, graphics are sent through a network to a mailbox location. Typical applications are e-mail and video answering machines.

Applications Class 4

This includes asymmetric non-real-time applications which are interactive. The user has individual presentation control, either at the level of user control of digital storage media, or also on the scheduling and the sequence of the data flow, depending on user decisions and choices - e.g. WWW or CAL decision systems.

1. Information base retrieval - audio-visual information retrieved from a remote Knowledge base on an individual basis; e.g. in entertainment, teleshopping, encyclopaedias etc. and the users may interactively browse through the audio-visual content.
2. Games - interactive audio-visual games are played with a remote computer/server or with other people through a computer /server.

Applications Class 5

In these applications the user has no individual presentation control.

1. Multimedia Presentation - local or remote multimedia presentations where no interactivity exists, and the user has no control on the scheduling or sequence of information flow.
2. Multimedia broadcasting for portable and mobile receivers - broadcasting of multimedia programmes (low resolution) for portable and mobile receivers, e.g. game-boy or watch-like terminals.

The development of MPEG-4 demonstrate the need for low-bit-rate coding algorithms and standards, and associated with this is the handling of interactivity and robust error free systems. The development of H.263 coding standard has shown the Discrete Cosine Transform motion compensation coding schemes are still able to improve compression performance.

2.8 REAL-TIME or OFF-LINE

Multimedia information can be presented in teaching both in real-time and off-line. The ability to present in real-time over a network is determined by the bandwidth of the network transporting information to the teaching space.

The advantages of real-time presentation are that live interaction is possible which is much more powerful tool for teaching. Video-conferencing is inherently real-time and much can be done with bit-rates as low as 128Kbps. Teaching at this bit-rate however is very limited, for two main reasons, firstly lips synchronisation is not possible and any movement involving a substantial part of the picture causes serious deterioration in resolution. Increasing the bit-rate to 384Kbps overcomes these problems and considerable teaching is taking place in various places throughout UK using these services. It is interesting that this speed should be acceptable to teachers in a variety of different subjects and it has been adopted in the JISC strategy for videoconferencing as the optimal bit-rate.

Increasing the bit-rate above this figure when using the ISDN services does not yield much advantage as the communications costs increase linearly with bit-rate, but the resolution does not improve significantly. This is due to the characteristics of the coding algorithm, which is trying to encode every movement within the picture, and as the bandwidth increases so the algorithm devotes its attention to the white noise in the picture.

In the SuperJANET ATM video experiment, the ATM channel was carrying primary rate ISDN at 2Mbps, and the picture quality was generally agreed to be of the same quality as good VHS video, or the reception quality of normal broadcast video in a domestic television set.

The resolution and picture quality could be improved further by using different technology, and this has been done on the Scottish MANs where K-NET’s Cell-stack CODECs use Motion-JPEG coding algorithms. Here the bit-rate has now risen to more than 15Mbps. The MPEG-2 algorithm can also provide high picture quality but there are few implementation of this technology at the present time on wide area networks. Interest in the use of the Scottish MANs is high and a significant number of teaching projects have now been instituted.

The audio quality associated with these various video compression methods uses relatively small bandwidth and high quality should always be available except at the lower ISDN bandwidths. It is essential that audio quality is available in teaching situations because if the student cannot hear what is being said then the video component is useless.

The problems with audio in real-time interactive situation can be divided into two areas as follows:-

1. Echo, which is a technical problem related to the fact that the compression and de-compression of the audio and video signals take time and so a delay is introduced into the signal. When this information is reflected back to the transmitting site it is apparent as echo. There are technical methods of suppressing echo and the equipment performance in this areas has improved markedly since the early days of the SuperJANET ATM video experiment.

   Echo suppression is now included in the specification of most videoconferencing equipment and is no longer a problem unless it is used in large rooms. The testing protocol development by the Videoconferencing Advisory Service (VCAS) has shown that in the cases of the equipment being tested, the echo cancellation provision has in a number of cases not coped with a large lecture theatre, and then additional echo cancellation equipment has had to be employed.

2. Noise interference does occur and this is commonly due to lack of good audio practice, and in most cases can be cured or dramatically reduced by proper siting of microphones and loudspeakers. It is particularly important to ensure that feedback does not take place between the loudspeaker and the microphones.

Precautions do have to be taken to ensure that the acoustic properties of the room being used do not cause resonance, and any local echo is absorbed.

More detailed information on room design and precautions to provide good quality audio are given at the VCAS Web site, URL is http://video.ja.net

Thus the essential requirements, of high quality video and audio are available, and with proper management the technology is capable of supporting a real-time interactive teaching environment. One of the important lesson learnt from the SuperJANET ATM video experiment was that good audio-visual practice was essential in the teaching space, where the audio and video signals were handled in their analogue form, and where there was good collaboration between the audio-visual staff and the networking staff then good and reliable operating conditions could prevail. The experience was that where audio-visual support was not present then the picture quality was inferior and the systems was not so reliable.

One of the current problems with the application of IP technology is that in many cases the desk-top equipment is used for communications between individuals and little attention is paid to the audio and video surroundings. The cameras are small and low-quality, attention is not given to ensure there is good lighting of the subject, and so the cameras are often attempting to work in conditions which are too stringent. As the locations, e.g. an office or laboratory, are often used by a number of people there can easily be background interference on the audio.

Although audio and video controls are provided on the display screen they are crude and do not offer the refinements required. This is unfortunate and often portrays the IP system as a low quality system, which is not the case. With proper quality control the IP system is capable of producing good real-time pictures. This has been demonstrated in the MICE project, where a link was set up to Sweden 2 years ago and a clinical operation transmitted over the INTERNET, this was done again recently and the improvements in the quality of the pictures has improved significantly.

3 The RELATIONSHIP between PEDAGOGY and TECHNOLOGY.

3.1 INTRODUCTION.

There are a number of solutions to taking multimedia into the teaching space; information can be transmitted over networks from local multimedia servers or from the World Wide Web, CD-ROM can be used with computers located in the teaching space, or video networks can be used to transmit real-time information either over legacy analogue networks or digital networks such as ISDN or the INTERNET. These possibilities open new opportunities to current thinking in Higher Education and the route to a wider acceptance of these new technologies may be through the encouragement of distance and open learning teaching in our existing institutions, especially in the higher degree and diploma courses. The interest in this technology is giving rise to a considerable wealth of experience in UK of using in using multimedia in the teaching space in a wide variety of applications.

The Dearing Report and the Government Green Paper on Lifelong Learning indicate that there will need to be changes in the pedagogical approach to the subjects which are being taught. The most common idea is to integrate the old and new experience in the hope that we can produce something which is both workable and pragmatic. There is strong resistance to change by some teachers because it threatens security, and for this reason the use of multimedia communications technology in the teaching space is often resisted. It means new relationships with the students, new ways of presenting the material, and perhaps a new role for the teacher. The teacher can now take on the role of manager and controller of streams of information or knowledge transmitted over networks into the teacher space.

The best way of managing this change is through enriching the experience of both the teacher and the student. The cycle of designing and building, prototyping and piloting courseware followed by evaluation and modification is one in which the experience is passed on in each cycle and one in which the confidence and awareness of both the teacher and the student can be built up. It is important that new technology is considered in hand with the pedagogical approach; there is little point in producing, transmitting and displaying super high-resolution images if there is no consideration of the way these images will be used in the teaching and learning. In some regions the introduction of new high speed networks has included the provision and connection of teaching space, and significant numbers of new teaching projects are developing to exploit these opportunities. In other areas no provision has been made to link teaching space to these networks and the high additional expense of making and equipping these links for the teaching departments has resulted in very little activity.

It is vital that where new technologies are brought into use that both the student and the teacher are aware of their potential - they are both users of the new systems and this implies that interactivity has a high priority in the performance characteristics of these systems. Interactivity plays an important role in all types of teaching and learning and in particular in higher education. The importance and the variety of modes of interactivity is recognised in the new MPEG-4 standard discussed elsewhere. The historic traditions of learning have always included discussion and an exchange of experience between the student and the teacher. The apprenticeship system is an expression of this process in the same way as the groups of students taking part in their discussion with Socrates. When evaluating video network teaching the students themselves have pointed out that the reason why they have come to university is to have the opportunity to hear first hand the opinions of their teachers and to be able to question them directly. They are prepared to accept that there is no reason why with sufficient care and ingenuity technology cannot extend that facility to a wider and more remote audience.

It is vital to realise the enormous step which has to be taken to convert the results of a feasibility study into a reliable service which can support teaching and learning. The original SuperJANET ATM video service and the London University LIVENET succeeded in providing a reliable service such that mainstream teaching could take place. The provision of a reliable working network involves not only the technological performance, but also the collaboration of a number of technical staff from different departments being at the right place at the right time. Video networks have exacting demands as the service is real-time.

The role of telecommunications and networks is to move information between different sites and this implies that there is a need to share information and experience, which is the essence of collaboration. Many teachers need to be convinced that this need exists, and that there are educational and organisational advantages in collaboration. Those people who have taken part in collaborative projects have in many cases found it to be a positive experience.

For a long time students have gone to the seat of learning and the idea that in distance learning we can take the teaching and learning to the student is new. Not surprisingly this will alter the relationship between the student and the teacher in a number of ways. The student is now able to see their role as the user and the consumer, and therefore they expect to have a greater say in how the course is run. In the cases where the student is employed their employer also see himself as a consumer, especially if they are funding the study. The student requires 3 components to facilitate the process of learning;

1. The information or subject content related to the subject being considered.
2. The tools by which they can access and organise this subject material
3. The guidance on how to structure their learning to understand the relationship between the subject material.

New technologies rarely provide all the functions and features that people would like and their successful use will depend upon a collaboration between teachers, students and technologist to reach a compromise between the advantages and limitations of the technology as it is implemented in education. There is little doubt that a solution can be found to most problems if money is not limited. In education money is not freely available to introduce new technologies and compromises have to be made. To maintain that educational processes should not be technology driven is a sensible comment; but to say that educational needs should be solely determined by the pedagogues, and the technology must be at their command is likely to waste much time and money. When the SuperJANET ATM video network was set up, the users learned that collaboration was required; the expertise of the teachers to design courses and to teach, the network engineers to transmit the information to the teaching space and the audio-visual engineers to provide good the visual images and good quality audio. All these groups worked well together, and the result was a system which was able to provide a flexible environment supporting a variety of teaching styles.

3.2 THE BASIC GENERIC REQUIREMENTS:

Ideally it is desirable to teach a subject by a generic approach, but the analysis of the requirements of different teaching applications, shows that it is difficult to separate the requirements of the subject from the requirements of the individual. It is perhaps worthwhile trying identify the basic scenarios which occur in the teaching and learning processes. These processes involve a meeting between student and teacher which can be called a dialogue and these are listed as follows:-

The meeting between the teacher and student.

The purpose of these meeting is to

a) to exchange knowledge and information

b) to exchange and discuss ideas derived from this knowledge

c) to measure the progress of these exchanges.

<u>a).The exchange of Information and Knowledge.</u>

We need to consider the various ways in which this knowledge and information is available and can be accessed.

i) knowledge and information which has been filtered by the teacher and is given in a “digested” form to the student during a ”meeting”. This filtered knowledge includes demonstrations, designed and/or modified by the student for the purpose of understanding the structure of that knowledge, to demonstrate a concept, process or procedure.

ii) knowledge and information specifically created for educational purposes by a third party and which the student can access

iii) knowledge which has been filtered

iv) knowledge not created for educational purposes which places demands on the student to access and extract that knowledge.

This knowledge and information can be stored on multimedia servers or on CD-ROM and the student can manipulate this in a workstation. The use of video-on-demand servers permits several students or groups to be working from the same information but they are not constrained to work at the same pace. Thus each student or group can work in a manner and at a pace most suited to their own requirements.

<u>b).The exchange of ideas.</u>

The exchange and testing of ideas is a vital part of the educational process in Higher Education and it takes place primarily by interaction between the teacher and student. It is for this reason that it is very important for the technology to be able to support interaction. One could assert that it is only because interaction can be supported by telecommunication technology at a variety of levels and in different environments, that innovation in teaching and learning has become possible.

Networks support two-way communications which is essential for interaction. Often in teaching multi-way communications are required to allow students at several different sites to take part in discussions, and it is under these circumstances that multi-point and multi-cast network configurations are important.

<u>c).The means to measure the quality and success of these exchanges.</u>

The teacher must assess the ability of the student to absorb the knowledge imparted and to do this feedback is necessary. The interactive procedures that can be supported by network communications, both real-time and off-line are necessary if distance and open learning is to be feasible.

The interaction between the teacher and student has then to be assessed as part of the quality assurance processes that are required in all teaching and learning. This is an area where the new technologies can make new and original contributions. The explosion in ideas on how to use the World Wide Web in teaching has been totally unexpected for many people. Many people still do not realise the potential of the INTERNET, and few people would like to predict exactly what we will be doing on the WWW in a few years time. As the bandwidth becomes available video is possible and is likely to become a very pervading methods of communications between people, whether on the basis of one-to-one links or groups for teaching and tutoring. Already questionnaire and MCQs are distributed over the INTERNET, shared working space meeting are becoming more prevalent and interactive discussion between teacher and students, and between student and student is possible.

Both real-time and off-line methods need to be investigated to obtain feedback on the use of new technologies.

It is valuable to also consider the ways in which the student and teacher meet each other, e.g.

The didactic lecture.

A formal lecture without questions and answers. If questions are permitted then only for a short duration at the end of the lecture. This situation has little interaction and is typical with large audiences.

The technology requirements for interaction are limited as usually the teacher will have designed and delivered his/her presentation in a manner that does not depend for its progress upon interaction.

The interactive lecture.

A less formal lecture where question and answers are exchanged between the teacher and students, which works best with small audiences.

The technology requirements to support interaction are essential, and the limitations must be respected. In many cases this type of lecture is best restricted to small number of student at each site, but the total number involved in the session can be quite large if a number of sites are involved. This has brought about technological solutions such as multi-point working, multi-casting etc. and had considerable influence on the methods of control of multi-site conferences, e.g. voice-activated switching.

The seminar.

A session where oral dissertations are delivered and then discussed between the participants.

The technology has to support interactivity because interaction will take place, but also there must be provision for a variety of methods of presenting information; e.g. overhead projection, slides etc.. These problems are probably the concern of the audio-visual engineer who must make provision for the display of whatever information may be shown on the network terminal.

The debate.

Where discussions is formalised and structured.

The demands on the technology have probably been met in the previous scenarios, except that there may be a need for more camera work than in the other circumstances.

The tutorial.

A discussion where the teacher/tutor is in control and directs the discussion.

The technology demands here are most close to those for the conventional business meeting; a limited number of people with a high level of interaction, but much of the success of the session will depend upon the teacher/tutors command of the technology to include the remote students in the discussion.

The tutor naturally exercises a control over the discussion and so he can determine which students have the opportunity to express their views.

The discussion group.

Where a free ranging discussion takes place probably under the direction of a chairperson to ensure some cohesion.

Here the requirements on the technology can be high as frequently several people may wish to speak at once, and voice activation methods of switching between sites may break down. Other means have to be developed to see more easily what is happening at a number of different sites simultaneously.

One of the conclusions from examining the scenarios listed above is that in most cases, provided the communications links between student and teacher are in place, then their actual location in the same room is not essential. This does however reinforce the condition that unless the parties can hear each other well, their means of communication are severely hampered. The visual link is secondary to the audio link in the communications process. The manner in which these links can function is however very dependent upon the ingenuity of the engineer.

One role of the current developments taking place in the new and innovative approaches to education is the development of tools which make it easier for the teacher to control the above processes. The following list give examples of these tools:-

1. Authoring tools to speed up and simplify the production of courseware
2. Authoring tools for the updating of a courseware
3. Communications tools for group-ware or real-time sessions
4. Multimedia production tools for handling multimedia including the production of video-clips
5. Questionnaire tools for the production and assessment of questionnaire
6 Modelling and simulation tools
7. Tools for building knowledge pools and for extracting information from these knowledge pools.
8. Management tools linking course management to student information databases.
9. Tools maintaining student progress

3.3 THE TEACHING AND LEARNING FRAMEWORK

Many development programmes emphasise student learning, but this is an incomplete picture. The student’s learning cannot progress positively unless there is guidance from the teacher. There is a need for teacher and student to work together and for both to be aware of the functionality that the technology can deliver. It is the co-operative role of teacher and student which is probably the biggest change that it taking place in the teaching and learning field. Not many years ago the teacher considered their opinion was paramount and the student lacked the experience or knowledge to challenge the teacher. Discussion and interaction did take place, but the teacher had the final say in the content and organisation of the teaching.

Many students come into higher education with computing experience they have gained at school which means they are already computer literate, and are familiar with methods of searching information on the WWW. The problem is more likely to be the volume of information available to the student to search and the time to do it; this is where the guidance of the teacher is so important.

Much study involves three stages, introduction of an idea or concept, the development of that concept and finally placing it in context.

- Conceptualisation where the dialogue between teacher and student is to seek to clarify the issues, i.e. the explanation.
- Construction or building upon the concept where the dialogue between teacher and student seeks to develop collaboration to expand the learning experience under guidance.
- The testing stage where the dialogue is with the teacher and with fellow students to debate and consider the consequences and what follows from the concepts when they are applied.

In all these stages there has been a dialogue and there is a need to consider the role this dialogue plays in both the real-time and off-line modes. In other words there is a need to understand the relative contribution of different computer communications technologies and what are their drawbacks. Also we need to consider whether the desk-top computer can satisfy these different requirements - are we expecting too many solutions to be possible through the desk-top interface? Is the use of the desk-top computer for learning and teaching no longer liberating, but constraining?

Introducing an idea or concept to students is primarily done by teaching and a number of technologies are available. Most of these are directed at channelling information and knowledge into the teaching space. Some people maintain these are superior to the conventional methods of books and lectures. It is difficult to substantiate this claim, and it is better to argue that a greater variety of methods are available to transfer information and knowledge, and the role of the teacher is to chose the ones which work most effectively in the set of circumstances. Many of these methods use networks to provide datastreams of audio, video and textual information into teaching space in real-time, i.e. multimedia communications which are capable of both synchronous and simultaneous provision of information.

Associated methods, such as CD-ROMs permit the student to have this information for their own use at their own convenience, i.e. for off-line and asynchronous access. It is interesting that so many teachers do not wish to use these potentially more flexible methods, and that is probably related to their desire to use tried methods and not to be involved in experimentation in the mainstream teaching tasks. Innovation is demanding of time and methods, and the prospect of rewriting a whole lecture course does not encourage teachers to associate with technologists. In universities which are research orientated and the teachers are themselves research workers, there is a desire to control teaching effort so that it has a minimal effect upon the research programmes.

Let us look at the relationship between teaching functions in conventional teaching and technologies;

- Books, lectures, TV, radio and CDs can be compared with datastreams in video, text and audio, and the WWW can also complements this role.
- Laboratories, workshops and field studies can be compared with JAVA, and remote sensing systems
- Tutorials and informal discussion are a form of dialogue and can be compared with shared working space, conferencing and whiteboards.

3.4 THE CONCLUSION OF A TLTP VIDEO-TEACHING PROJECT

INSURRECT (Interactive Teaching of Surgery between Remote Centres).

3.4.1. Administration

Administration of the project divides into the following main activities:-

1. General administration
2. Arrangement and co-ordination of teaching
3. Arrangement and co-ordination of network and teaching space.

General administration

Administration of the project involves a number of issues. An internal project contract was drawn up at Bristol University to define the responsibilities within the project. Drawing up a project contract was an interesting exercise as there are few models available to use as a basis for such an agreement. In practice there has been no contentious issues between project partners. This agreement is one of the deliverables of the INSURRECT Project.

Initially a steering committee was set up under the chairmanship of a clinician, Prof. M Hobsley. In practice the has only functioned on a few occasions, and a number of those meetings used video conferencing facilities on SuperJANET. At the end of the first year of teaching it became apparent that there was a need to discuss the progress of the project and a meeting was organised at Trinity Hall, Cambridge for a weekend conference. This meeting proved very successful and significant progress was made as a result. The main outcome was the decision that each session should be structured in the same way, but still permitting the individual style of teaching relating to each teacher.

Communications between sites was carried out using four methods:-

1. Facsimile - this enabled all partners to be reached but was relatively time consuming.
2. Electronic mail - this was the most convenient method but unfortunately a number of medical partners were not on e-mail, and consequently facsimile had to be used.
3. As the project continued use was made of the Internet to send out information on timetables, MCQs and general matters.
4. Video conferencing - this was used for a number of meeting between partners and for discussions, including visits from TLTP secretariat.

Arrangement and co-ordination of teaching

The notification of timetables and the list of topic for each session was organised between the medical partners. This required one project staff member with a large number of man-hours contacting and circulating information to each site. Depending on the activity and local organisation at each site, it was advisable for one person to be in touch with all teachers involved in the project. It was not practical to have a single point of contact at each site, because medical staff did not always come in frequent contact and one could not be certain messages were passed on. The medical secretaries were very helpful but for reliability it was better to have one person in the project responsible for the liaison.

At the beginning of the project there was liaison to establish the curriculum to be followed and who was to do what. This was not as difficult a problem as was anticipated as the medical curriculum is under review in many centre. The course was intended to be a basic course in surgery for undergraduate students and the General Medical Council was interested for the Medical School to develop their own core curriculum. This project gave the opportunity for the 6 Medical Schools to discuss this and come up with their version. This proved relatively easy to do. Following on agreement of the curriculum there was the decision concerning who was to teach what. Again this was divided between the sites easily, the guiding principle being utilise the expertise at each centre as far as possible.

The third issue was to produce a common timetable between all sites. This proved very difficult to get all 6 sites together at a specific time. This was complicated by the fact that one medical school was introducing its new curriculum and its basic design was very different to the existing systems. Interestingly the first time most sites were not flexible, about the second time there was much more flexibility.

Arrangement and co-ordination of network and teaching space.

The communications aspect is divided into two areas, firstly the action which takes place in the analogue domain, i.e. the production and reception of analogue audio and video signals and their display; and secondly that which takes place in the digital domain when the audio and video signals are digitised and transmitted over the network. The network can be divided between the local and long distance networks. The CODEC (compression de-compression equipment) is located at the boundary between the analogue and digital domains. It is not likely to be in the teaching space as the CODEC feeds directly into the ATM switch. Various technologies have been used to link the teaching space to the CODEC, e.g. LIVENET at UCL, Microwave links at Newcastle, SMDS from Bristol to London.

The natural division of responsibility has been between the Audio-Visual staff who are responsible for the analogue signal and the Computer Centre Network Staff who have been responsible for the digital signals. It has been important for these two groups to work well together. In the initial stages the Network Staff gave invaluable assistance in installing the links between the Computer Centre and the teaching space. The AV staff have been responsible for the quality control of the AV signal being fed to the CODECs.

In general it has been apparent to the project team that those centres which could count on the support of the AV staff in the teaching space have been able to produce the better quality images and have had greater reliability of activity within the teaching space. Those centres that have had a poor level of co-operation or even none at all have had greater difficulties, especially in the early stage when the system was being set up and commissioned for routine working. In the case of Bristol this problem was more acute as the technology was at a developmental stage.

A continuing problem in this area has been booking of resources. The network bookings, made through Edinburgh have been relatively simple to establish. It has been more difficult to ensure that the local bookings are satisfactory because each centre has its own booking procedures for local resources, and it is extremely difficult to establish a global system. This subject has been of serious concern to UKERNA and a special project has been set up to look into this matter directed by University of Swansea.

There have been a few situations where local errors have resulted in difficulties in teaching space being available.

3.4.2. Network Performance

The network teaching can fail at a number of levels:-

Network Failure	1. Catastrophic failure of the whole network
	2. Catastrophic failure of a section of the network
Site Failure	3. Failure of the internal network facilities within a site
	4. Failure of AV facilities in the teaching space
	5. Non-availability of teaching space
Teaching failure	6. No show of teacher affecting the whole session
	7. No show of support staff

3.4.3. Initial Setting Up Facilities

Each centre had to make some investment to convert teaching space to be connected to SuperJANET.

One centre, Bristol, had to have special facilities and used a different technology to the other centres. The SMDS network was used and this is the subject of a separate report.

At each site connectivity had to be made between the Computer Centre, where the ATM switch is situated, and the teaching space. In all cases very significant help was provided by the Computer Centre Network staff. In all cases an optical fibre extension was provided, and connection made to the SuperJANET CODEC, except for Newcastle which used their project funds to purchase their CODEC.

Also it was necessary to equip a room, lecture theatre or seminar room as the teaching room. The equipping of the room was usually carried out with the assistance of the local Audio-Visual staff, although in cases where there was no AV support the situation was more difficult. Manchester depended upon local expertise in the Royal Infirmary and Computer Centre, Bristol brought in a commercial company.

3.4.4. Multi-media Presentations

Equipment required and design of facilities (See http:// www.ukerna.tech.ac.uk) Sites were advised to use the standard arrangement of equipment, i.e. three cameras, with monitors or video projector, a fixed microphone for the teacher and a roving microphone for the students and appropriate loudspeakers. The three cameras included one to view the speaker, one to view the students and a document camera. A further camera was also used on occasions when a site was broadcasting and patients or other people were present.

The multi-media information used in the lectures could be slides, transparencies and video. On occasions special facilities were used such as an ISDN link to a General Practice surgery, and links to the clinical operating theatres. When a site was the transmitting site then video switches would be used to select the appropriate camera view for transmission.

On some occasions an analogue video server was used, the control being with the teacher, and slides and video was played over the network from the server located at University College London.

- The multi-media presentations used the following media:-
- 35mm slides
- PAL video tape
- Laser Video Recorder analogue video server - still and moving images and
- animation.
- Overhead transparencies - pre-prepared
- Overhead transparencies - writing and drawing in real-time
- Patients
- ISDN links to other sites e.g. General Practice surgery
- Audio/Telephone links to Australia and New Zealand.

3.4.5. Pedagogic Experience

Each course lasted 8 weeks, a period designed to correspond with the time spent by students on any one clinical attachment (although this ideal was not attained except in certain centres). A course consisted of 15 or 16 presentations, delivered twice a week between 9 am and 10 am. The necessity to book exact periods on SuperJANET meant that presentations had to be strictly tailored to fit the allotted time. After some early lapses of discipline, presenters realised the necessity and by the end were conforming precisely.This itself was considered by many to be a significant advantage of the system.

Structure of Presentation

At first these took the form of formal lectures, and the speakers brought along their usual visual aids, almost always slides. However, during the first year of teaching (i.e., the second year of the project) it gradually became apparent that the formal lectures translated very poorly into this new format and were considered by the students to be dull and didactic. It was also apparent that the real strengths of the multimedia facilities were not being exploited. A full conference of the INSURRECT participants was held at Cambridge in late summer 1995, these problems (together with the problems of failing to promote interaction (see below) were discussed, and remedies identified.

Subsequent presentations have since conformed more or less to the following formula (see Document: The Way Ahead for further details):-

i. The start is a patient (live or videoclip)

ii. Management of that patient is discussed

iii. The problem presented by that patient is then set in the context of all the other possible problems he might have had despite presenting in much the same manner. The procedure of dissecting the clinical presentation to determine the correct management for each individual situation is provided in the form of a decision-tree.

iv. Throughout the presentation special efforts (detailed below) are made to encourage interaction.

Interaction

Most of the teaching was delivered to groups of students of 10-15 persons. Students were usually just starting the clinical part of their course;,they had previously, in the pre-clinical course, been lectured in large lecture theatres with little interaction and they took some time to adjust to the more interactive manner of the small group teaching. These small groups are consistent with the normal format for clinical teaching, but despite the small numbers in the group, the response to a request for questions at the end of each session was meeting a nil or miniscule response. The Document The Way Ahead details the procedures that were decided upon to promote interaction.

They may be summarised as follows:-.

i. Preparation Presenters are encouraged to build pre-determined periods into their hour for specific episodes of interaction

ii. Student Chairmen Students are reluctant to make comments or answer questions in case they make mistakes and are ridiculed for them. The appointment of a student chairman at each site results in a spokesperson who can focus suggestions from his colleagues without the opprobrium of being the author of a wrong answer if it is wrong.

iii. Competition: The speaker can produce an element of competition between the sites to encourage healthy rivalry and a desire to respond rapidly.

iv. Restriction of material / greater use of handouts: Because the various features above limit the amount of information that can be conveyed, it becomes more important to use handouts so as to increase the flow of information. It is important that handouts expand on what has been said, rather than paraphrasing it. Algorithms ('decision-trees, biii above) are especially valuable as they can be built into an important repository of useful information.

v. MCQ A procedure was developed using a few MCQ to test the knowledge of the students on a subject before the presentation. The students then recorded their answers to the questions for a second time during the presentation, thereby ensuring that they were immediately using the new information they had been given and facilitating its introduction into their long-term memory stores. At the end of the presentation, the speaker tooks his audiences through the questions, indicating the correct and incorrect answers. This procedure gave the opportunity for more student/teacher interaction.

Self-Learning

The format of the decision tree proved a useful approach to the construction of CAL based upon a particular presentation. The decision tree is presented, with suitable explanatory notes, for those students who are at an early stage of their studies of a subject. They then proceed to a number of clinical presentations, making full use of the multimedia facilities available, to teach themselves how to manage each presenting problem. Later in the course, they can revise by testing their knowledge of how to deal with the problems, and if they make any mistakes the programme refers them back to the relevant point of the decision tree.

3.4.6. User Reaction

Student reaction

Student reaction was assessed in a number of ways; by questionnaire, by interview (video recording) and by special session at the end of course to discuss the programme. Some students will always prefer face-to-face teaching and did not like the presence of the network. The students were neutral about whether they felt the session could have been given better without the network.\

Teacher reaction

Teacher reaction was assessed by discussion at meetings and by individual communications by telephone, facsimile, e-mail or letter. With only one exception, the teachers gradually came round to using the principles embodied in “THE WAY AHEAD”. The exception was not rejected, because despite his format of presentation he was an effective teacher whom the students well appreciated. We recognised that a few specially gifted individuals may be able to get away with breaking the pedagogical priciples, but most of us could not.

The third issue was to produce a common timetable between all sites. This proved very difficult to get all 6 sites together at a specific time. This was complicated by the fact that one medical school was introducing its new curriculum and its basic design was very different to the existing systems. Interestingly the first time most sites were not flexible, about the second time there was much more flexibility.

3.5 CONCLUSIONS.

One conclusion from examining the results of the educational development programmes is that there is too great an emphasis on the products of applying technology to education at the expense of investigating the mode of the dialogue between students and teacher, and with other students. The challenge is the nature of the engagement between the partners of the educational process and consideration of how this process can be developed utilising the technology to provide opportunities which are not present in conventional teaching and learning.

Where products are produced they should be flexible and not require significant time and effort producing and updating. There should be more interest in the use of the technology to facilitate delivery of the educational material, and as this implies the transport of information across boundaries we should think more freely in terms of the movement of information and knowledge. A comparison can be made with the commercial world where wealth is directly related to the ability to move information and knowledge to the place where it is most useful.

The educational world should be thinking about moving its “merchandise” to the user rather than insisting the user comes to them. The relationship between teacher and student should be collaborative and providing guidance rather than didactic.

4 ROLE OF NETWORKS WHEN USING MULTIMEDIA INFORMATION.

4.1 PRESENTING MULTIMEDIA INFORMATION

Most university campus will be limited in the bandwidth available in their networks, often about 100Mbps Ethernets, which means the maximum available to individual users is likely to be about 10Mbps. With time further bandwidth will become available and then the ability to handle multimedia information will be more assured.

Two scenarios will be common; first the presentation of multimedia material in the teaching space which is being relayed over the local area network from a network server and secondly by the use of standalone systems. When the networks are used, there are two principle methods in use; either the material may be multicast to a number of sites simultaneously over the MBone, or multipooint ISDN configurations can be used. In either case teaching can be carried out to distribution of sites collaborating in a shared teaching programme.

The MBone operates on the INTERNET, and is designed to link a number of sites but uses pathways which minimise the traffic on the INTERNET, by sharing routes between sites as much as possible. The MBONE can be used interactively, permitting each site to communicate with every other in the session.

This is different to multipoint ISDN conferences between a number of sites simultaneously, where all sites are connected to a central Multipoint Control Unit (MCU). The switching of the MCU is commonly by voice activation, so that the site where the audio is active, is transmitted to all receiving sites. This system until recently required that all sites were operating at the same bandwidth (e.g. 128Kbps for ISDN-2), which resulted in the whole network having to operate at the speed of the slowest partner. Recent advances will permit mixed working thus permitting some sites to use their equipment capable of operating at 384Kbps and higher resolution, and at the same time allowing smaller sites to be part of the conference with their equipment only operating at 128Kbps.

The multimedia material used in teaching applications can be delivered to the user over a variety of types of networks, which can be internal departmental or campus LANs or they can be wide area networks (WANs) linked to various regions and institutions, supported by the academic networks (JANET and MANs) or by public service provider such as ISDN. Each of these technologies will be discussed in this paper.

It is assumed that the higher the sophistication of the supporting information in teaching and learning the better it is for the student. This greater sophistication requires more effort in the production, and more sophisticated equipment for the presentation, so the teacher should always ask what is the simplest and clearest way of presenting information consistent with the use of the material. A high resolution picture is not any advantage to the student if the information could be presented perfectly adequately as a line drawing. As multimedia is expensive to produce and to present it’s use must be justified. High resolution images are good to look at but the added detail must contribute something positive to the student’s knowledge.

The advantage of multimedia is the ability to bring together a number of different media in a way which permits synchronisation of information, e.g. vision and sound, and which overall enhances the information. It should be remembered that the “people” themselves can provide information in several media naturally, e.g. the patient demonstrating a particular clinical situation provides visual, tactile and even olfactory information of signs and symptoms, as well as being able to describe and be questioned about the effects of the disease. Another example is the ability to link into the teaching space people over remote video links which can provide visual and audio information from a distance, e.g. the application of open and distance learning techniques.

In practice the good use of multimedia gives the teacher opportunities to provide a number of stimuli to the students, which can make the teaching much more comprehensive and interesting. The ability to provide more varied stimuli to the teaching and learning process can also include the student working on his/her own; the ability to present on the computer screen in a variety of coloured images, and sound effects can enhance the information being presented to the student. The eye is undoubtedly attracted by colours and shapes and these are most frequently used to attract the interest of the student.

4.2 NETWORKED SYSTEMS

If legacy network, such as optical fibres carrying analogue audio-visual signals still exist in an institution they can be used to good purpose. For example in the TLTP INSURRECT project a SONY LVR videodisc server was used to store still and moving images and the LIVENET optical fibres, were used to take these analogue images into the teaching space. Also they carried the images to CODECs for transmission to remote teaching sites over the SuperJANET ATM video network. The SONY video server stored up to 54,000 image frames as either still images or short video sequences. This SONY system could be controlled using RS232A commands, transmitted either over the LAN or over the INTERNET from remote sites. Thus it was possible to control slide-shows easily from a laptop in the lecture theatre, either within the same institution or remotely. Video sequences could also be controlled in an interactive manner; stopping, starting and playing backwards and forwards. In the future the digital network will increasingly pervade the whole campus, but it must be remembered that the analogue network is likely to provide a higher bandwidth and therefore higher resolution images.

In the INSURRECT project; LIVENET analogue circuits were used to transmit video and audio signals around the UCL campus, and also were fed into CODECs to be transmitted over the SuperJANET ATM video network to the remote sites. One site (Bristol) was not on SuperJANET, and TCP/IP was transmitted over a 100Mbps SMDS link very successfully. In the early part of the experiment there was some difficulty synchronising the voice channel with the video but this was improved during the lifetime of the project.

Ethernet LANs are common in university campus which are capable of providing good resolution signals using IP into the teaching space but there is no bandwidth reserved for the video signal, and thus the quality of the signal will vary according to the level of usage of the Ethernet. It is notable that current LAN video conferencing systems are using IP for the campus networks, but use an ISDN gateway to connect with remote sites because the guaranteed bandwidth available on ISDN ensures that the picture quality can be maintained.

ATM networks can be used with typical bandwidths of 25Mbps to 155Mbps, both in LANs and WANs. ATM networks are often used as a transport medium carrying a number of protocols such as IP, SMDS and LAN Emulation. In some applications native ATM technology is utilised to transmit the audio and video information. In many campus networks there is a mixture of different technologies in use; a typical scenario would be an ATM backbone with Ethernets connected. As far as possible the ATM backbone should be routed close to major teaching areas.

When teaching there is frequently a need to link more than two sites, and then Mbone and multipoint configuration are necessary. A multipoint conference uses a star network configuration, and is capable of being interactive as each port on the MCU will support two-way video and audio. The audio signal is used to control the switching in voice switching systems.

Multi-casting is particularly relevant to IP networks, and the Mbone has been developed to economise/minimise the traffic on the network by choosing routes between all sites so that as much as possible of the traffic can share the same routes and branching out in the final stages of the links.

Ideally the multi-cast network will link all sites to each other and further this network can support interactive traffic. This will be considered in more detail in section 4.4.

4.3 WHAT IS ISDN?

ISDN -- Integrated Services Digital Network -- is a telephone service that enables you to have high-speed data connections through your phone line. ISDN is basically the telephone network turned completely digital, using existing wiring. ISDN is capable of transmitting large amounts of data rapidly between sites, and it is possible to support compressed video transmission because the bandwidth is guaranteed and therefore the audio and video synchrony of the signal is maintained.

ISDN is much cheaper than many other methods of moving data at high speeds, but it is still expensive relative to a normal phone line. Normal phone lines -- the kind that work reliably with your 14.4Kbps modem, fax, answering machine, are known in some circles at POTS-- plain old telephone service. There are serious attempts to provide video on these circuits as exemplified by video-telephones, but these in general do not provide the resolution necessary for teaching.

A major drawback to ISDN is that because it moves digital data instead of analogue data, it doesn't work with your regular modem, or answering machine. You need special, expensive equipment to perform those functions at ISDN speeds. However, ISDN is faster than a standard modem, and is available in various parts of the world, including Australia, Western Europe, Japan, Singapore, France, and portions of the U.S.

There is listed below three web sites which give very comprehensive information about ISDN, and include reference to the operation of ISDN in a number of different countries.

http://www.cis.ohio-state.edu/~fine/ISDN/references.html

http://alumni.caltech.edu:80/~dank/isdn

http://www.pacbell.com/ISDNbook/

4.4 THE MBONE and MULTICASTING

This section contains extracts from two major articles and books on the Mbone and multicasting. This information is contained on the Web and an investigation of the web sites will provide much more detailed and comprehensive information. The book MBone: Multicasting Tomorrow's INTERNET by Kevin Savetz, Neil Randall, and Yves Lepage (ISBN: 1-56884-723-8) and published by IDG Books Worldwide, Inc., gives comprehensive information. An extract from this book is shown in section 4.4.1

An extracts from a second book, **-** MBONE, the Multicast BackbONE, written by Mike Macedonia and Don Brutzman of the Naval Postgraduate School is described in section 4.4.2.

4.4.1. Multicasting - Tomorrow's INTERNET Today, (by Kevin Savetz, Neil Randall, and Yves Lepage).

The MBONE is a critical piece of the technology that's needed to make multiple-person data, voice, and video conferencing on the INTERNET -- in fact, sharing any digital information -- cheap and convenient.

What Is Multicasting?

Multicasting is a technical term that means that you can send a piece of data (a packet) to multiple sites at the same time. (How big a packet is depends on the protocols involved-it may range from a few bytes to a few thousand). The usual way of moving information around the INTERNET is by using unicast protocols -- tools that send packets to one site at a time.

You can think of multicasting as the INTERNET's version of broadcasting. A site that multicasts information is similar in many ways to a television station that broadcasts its signal. The signal originates from one source, but it can reach everyone in the station's signal area. The signal takes up some of the finite available bandwidth, and anyone who has the right equipment can tune in. The information passes on by those who don't want to catch the signal or don't have the right equipment.

On a multicast network, you can send a single packet of information from one computer for distribution to several other computers, instead of having to send that packet once for every destination. Because 5, 10, or 100 machines can receive the same packet, bandwidth is conserved. Also, when you use multicasting to send a packet, you don't need to know the address of everyone who wants to receive the multicast; instead, you simply "broadcast" it for anyone who is interested. (In addition, you can find out who is receiving the multicast -- something television executives undoubtedly wish they had the capability to do.)

How is the MBONE different from Multicasting?

Unfortunately, the majority of the routers on the INTERNET today don't know how to handle multicasting. Most routers are set up to move traditional INTERNET Protocol (IP) unicast packets -- information that has a single, specific destination. Although the number of routers that know how to deal with multicast are growing, those products are still in the minority.

Router manufacturers have been reluctant to create equipment that can do multicasting until there is a proven need for such equipment. But, as you might expect, it's difficult for users to try out a technology until they have a way to use it. Without the right routers, there's no multicasting. Without multicasting, there won't be the right routers. Catch-22.

What is a router? A router is a device that connects a local area network -- such as an inter-office LAN -- to a wide area network -- such as the INTERNET. The router's job is to move information between the two networks.

Many routers today are unicast routers: They are designed to move information from a specific place to another specific place. However, routers that include multicasting capabilities are becoming more common. In 1992, some bright fellows on the INTERNET Engineering Task Force (IETF) decided that what no one would do in hardware, they could do in software. So they created a "virtual network" -- a network that runs on top of the INTERNET -- and wrote software that allows multicast packets to traverse the Net. Armed with the custom software, these folks could send data to not just one INTERNET node, but to 2, 10, or 100 nodes. Thus, the MBONE was born.

The MBONE is called a virtual network because it shares the same physical media -- wires, routers and other equipment -- as the INTERNET.

The MBONE allows multicast packets to travel through routers that are set up to handle only unicast traffic. Software that utilises the MBONE hides the multicast packets in traditional unicast packets so that unicast routers can handle the information. The scheme of moving multicast packets by putting them in regular unicast packets is called tunnelling. In the future, most commercial routers will support multicasting, eliminating the headaches of tunnelling information through unicast routers.

When the multicast packets that are hidden in unicast packets reach a router that understands multicast packets, or a workstation that's running the right software, the packets are recognised and processed as the multicast packets they really are. Machines (workstations or routers) that are equipped to support multicast IP are called mrouters (multicast routers). Mrouters are either commercial routers that can handle multicasting or (more commonly) dedicated workstations running special software that works in conjunction with standard routers.

So, what's the difference between multicasting and the MBONE? Multicasting is a network routing facility -- a method of sending packets to more than one site at a time. The MBONE is a loose confederation of sites that currently implement IP multicasting.

The MBONE -- or multicast backbone -- is a fancy kludge, a hack. It is at best a temporary utility that will eventually become obsolete when multicasting is a standard feature in INTERNET routers. By then there will be an established base of MBONE users (which should make the router manufacturers happy). The utilities and programs that work on today's MBONE will undoubtedly work on the multicast backbone of tomorrow.

Perhaps the most sought-after function that the MBONE provides is videoconferencing. The MBONE originated from the INTERNET Engineering Task Force's attempts to multicast its meetings as INTERNET videoconferences. MBONE video is nowhere close to television quality, but at a few frames a second, video quality is good enough for many purposes.

The MBONE's capability to carry remote audio and video makes it a wonderful tool for seminars, lectures, and other forms of "distance education." Imagine sitting in on a lecture that's being given live thousands of miles away and even asking questions or contributing to a panel discussion. Indeed, much of what happens today on the MBONE is of a technical nature, information that most of us would find dull. However, the nerds don't get to keep the MBONE to themselves. Besides esoteric engineering events, the MBONE is home to more exciting fare, such as multicasts of concerts, a do-it-yourself-radio station, and even poetry readings.

Can Your Computer Handle the MBONE?

Although anyone who has the right equipment can use the MBONE, the hardware and connectivity requirements for using the MBONE are much greater than what's available on the equipment that most INTERNET users have in their homes. A PC or Macintosh system coupled with a standard modem doesn't have enough computing power or bandwidth to send or receive MBONE transmissions. You need a good deal of power to handle multicast IP. Today, multicasting software -- the behind-the-scenes tools for moving, encoding, decompressing, and manipulating multicast packets -- is available only for high-end UNIX workstations, but the situation is changing, it's not too much to imagine that MBONE tools will soon be available for home computers -- PCs that are running Microsoft Windows and Macintosh computers. It will probably take the most powerful home computers (with Pentium and PowerPC chips), but it seems to be a likely eventuality. The software tools are being built: PC/TCP Version 2.3 from FTP Software Inc. supports multicasting for PCs, as does Windows 95, and it is rumoured that the next version of MacTCP will support multicasting.

How Much Bandwidth Is Necessary?

Even if users had the hardware to do multicasting today, another huge hurdle would prevent the MBONE from taking over the INTERNET: Most users don't have enough bandwidth. A multicast video stream of 1 to 4 frames per second eats about 128Kbps of bandwidth (ISDN) and gives you slow, grainy, bandwidth-hogging video. (By comparison, television-quality video scans at about 24 frames per second). Remember though, that a video stream uses the same bandwidth whether it is received by 1 workstation or 100.

In their paper, "MBONE Provides Audio and Video across the INTERNET," authors Michael Macedonia and Donald Brutzman (see below section 4.4.2) write, "Only a few years ago, transmitting video across the INTERNET was considered impossible. Development of effective multicast protocols disproved that widespread opinion.

> "Until recently, experts believed that the MBONE could not be used for transmission of simultaneous video, audio, and data because of limited bandwidth," notes Professor Don Brutzman. "This effort to push the envelope of computing technology has provided valuable data to computer scientists and has shown that methods can be employed to work around the bandwidth problem."

There's a ceiling to the amount of information that can move around on the MBONE as a whole: 500Mbps. At full tilt, the MBONE itself can handle no more than four simultaneous videoconferencing sessions or eight audio sessions.

4.4.2. MBONE, the Multicast BackbONE, (by Mike Macedonia and Don Brutzman).

Introduction.

MBONE stands for Multicast Backbone, a virtual network that has been in existence for about three years. The network originated from an effort to multicast audio and video from the INTERNET Engineering Task Force (IETF) meetings. MBONE today is used by several hundred researchers for developing protocols and applications for group communication. Multicast is used because it provides one-to-many and many-to-many network delivery services for applications such as videoconferencing and audio that need to communicate with several other hosts simultaneously.

Multicast networks.

Multicasting has existed for several years on local area networks such at Ethernet and FDDI. However, with INTERNET Protocol (IP) multicast addressing at the network layer the service group communication can be established across the INTERNET. The reason that MBONE is a virtual network is that it shares the same physical media as the INTERNET, though it must use a parallel system of routers that can support multicast (e.g. dedicated workstations running with modified kernels and multiple interfaces) augmented with "tunnels". Tunnelling is a scheme to forward multicast packets among the islands of MBONE subnets through INTERNET IP routers which typically do not support IP multicast. This is done by encapsulating the multicast packets inside regular IP packets.

Bandwidth.

The key to understanding the constraints of MBONE is thinking about bandwidth. The reason why a multicast stream is bandwidth-efficient is that one packet can touch all workstations on a network. Thus a 125Kbps video stream (1 frame/second) uses the same bandwidth whether it is received by one workstation or twenty. That is good. However that same multicast packet is prevented from crossing network boundaries such as routers or bridges. The reasons for this restriction are religious and obvious: if a multicast stream which can touch every workstation could jump from local network to local network, then the entire INTERNET would quickly become saturated by such streams. That is very bad! Thus the MBONE scheme encapsulates multicast streams into unicast packets which can be passed as regular INTERNET protocol packets along a virtual network of dedicated multicast routers (mrouters) until they reach the various destination local area networks. The use of dedicated mrouters segregates MBONE packet delivery, protecting standard network communications such as mail and TelNet from MBONE experiments and failures. Once properly established, an mrouter needs little or no attention. Given this robust distribution scheme, responsible daily use of the MBONE network consists only of making sure you don't overload your local or regional bandwidth capacity.

Networking details.

When a host on an MBONE-equipped subnet establishes or joins a group it announces that event via the INTERNET Group Management Protocol (IGMP). The multicast router on the subnet forwards it the other routers in the network. MBONE sessions use a tool developed by Van Jacobson of Lawrence Berkeley Laboratories called sd (session directory) to display the announcements by multicast groups. sd is also used for launching multicast applications and for automatically selecting an unused address for any new groups.

Groups are disestablished when everyone leaves, freeing-up the IP multicast address for reuse. The routers occasionally poll hosts on the subnets to determine if any are still group members. If there is no reply by a host, the router stops advertising that hosts group membership to the other multicast routers.

Protocols.

The magic of MBONE is that teleconferencing can be done in the hostile world of the INTERNET where variable packet delivery delays and limited bandwidth play havoc with applications that require some real-time guarantees. It is worth noting that only a few years ago putting audio and video across the INTERNET was considered impossible. Development of effective multicast protocols disproved that widespread opinion.

In addition to the multicast protocols, MBONE applications are using the Real Time Protocol (RTP) on top of User Datagram Protocol (UDP) and IP. RTP is being developed by the Audio-Video Transport Working Group within the IETF. RTP provides timing and sequencing services; permitting the application to adapt and smooth out network-induced latencies and errors. The end result is that even with a time-critical application like an audio tool, participants normally perceive conversations as if they are in real-time, even though there is actually a small buffering delay to synchronise and sequence the arriving voice packets. Protocol development continues. Although operation is usually robust, many aspects of MBONE are still considered experimental.

Data Compression.

Another aspect of this research is the need to compress a variety of media and to provide privacy through encryption. Several techniques to reduce bandwidth include Joint Photographic Experts Group (JPEG) compression and the ISO standard H.261 for video. Visually this translates to velocity compression: rapidly changing screen blocks are updated much more frequently than slowly changing blocks. Encoding for audio include Pulse Coded Modulation (PCM) and GSM. Outside of the concerns for real-time delivery, audio is a difficult media for the MBONE and teleconferencing in general because of the need to balance signal levels for all the parties who may have different audio processing hardware (e.g. microphones and amplifiers). Audio also generates lots of relatively small packets, which are the bane of network routers.

Application Tools.

Besides basic networking technology, MBONE researchers are developing new applications that typify many of the goals associated with an "information superhighway." Video, audio, and a shared drawing whiteboard are the principal applications, provided by software packages called nv (net video), vat (visual audio tool) and wb (whiteboard).

Groupwork on groupware.

The MBONE community is active and open. Work on tools, protocols, standards, applications and events is very much a co-operative and international effort. Feedback and suggestions are often relayed to the entire MBONE mailing list (as an example, this article was proofed by that group). Co-operation is essential due to the limited bandwidth of many networks, in particular transoceanic links. So far no hierarchical scheme has been necessary for resolving potentially contentious issues such as topology changes or event scheduling. Distributed problem solving and decision making has worked on a human level just as successfully as on the network protocol level. Hopefully this decentralised approach will continue to be successful even in the face of rapid addition of new users.

The future.

It is not every day that someone says to you "Here is a multimedia television station that you can use to broadcast from your desktop to the world." These are powerful concepts and powerful tools that tremendously extend our ability to communicate and collaborate. These tools are already changing the way people work and interact on the net. See you later!

4.5 THE ROLE OF ATM NETWORKS - THE METROPOLITAN AREA NETWORKS

4.5.1. The development of the London MAN

What is a MAN?

The definition of a Metropolitan Area Network (MAN) cannot be precise. It falls between a Local Area Network (LAN) and a Wide Area Network (WAN) and has some of the features of each. A set of computer interconnections becomes a "Network" when they have some form of common management. A LAN operates over a restricted geographical areas and almost always operates within one organisation. A number of LAN technologies have evolved offering very high performance at low cost over distances of under 10Km. The WAN on the other hand uses technologies that do not limit the geographical reach and expects to connect nodes under different management domains.

A MAN aims to serve a geographic area beyond the scope of LAN technologies, yet is restricted by some well defined community of interest, often a city. A MAN will often provide interconnection between sites of the same organisation as well as interconnecting organisations. This means that MANs have to provide performance close to that obtained on LANs yet cope with the interaction of multiple management domains. As a consequence the establishment of a successful MAN requires the correct balancing of a number of technical and political factors.

Historical background

London is a natural area for the development of an academic MAN, with more than 20% of the Higher Education Institutions (HEI) in the UK located within a relatively compact area approximately 50Km in diameter. However, the establishment of a co-operative venture between a group of independent organisations does not often happen without some particular stimulus to overcome inertia. This did not arrive until 1994, but a review of the academic networking arrangements in London before then will set the scene.

Two changes in the late 1980s led to the break up of the University of London network. Firstly, within the federation of the University of London a greater degree of autonomy and control was given to individual colleges. A number of colleges merged into larger units and, for the purposes of computing services, the Schools and Institutes were grouped into seven Clusters, each led by a large computer centre. The second factor was the replacement of the Computer Board of the universities and research councils with the Information Systems Committee, which had significantly less central control over IT spending in universities. The result of these two changes was that the components of the University of London wanted to exercise their new found independence by providing their own networking. The federal network was replaced with a number of independent networks based on the Clusters, each with a single link to ULCC for connection to JANET. Throughout this period networking in the polytechnics was substantially underdeveloped due to the absence of the top sliced funding that had assisted the development of university networking.

A separate analogue video network (LIVE-NET) (see section 2.3 Legacy networks) within the University of London provided a number of useful inputs into the development of national networking and the London MAN. It demonstrated that co-operation across distinct management domains was possible - LIVE-NET thrived at the time when the federal X.25 network was being dismantled. It was an important component of "Shoestring", the national pilot IP network, operating a bridged IP network over analogue channels of the video network. Many lessons learnt from the use of the video conferencing services have provided direct input in the plans for the SuperJANET video conferencing network.

The geographical scope of the London MAN was agreed to be approximately the area contained within the M25 motorway, but this is not seen as a rigid boundary.

The sites connecting to the network in the first phase are:

Birkbeck College

City University

University of East London

University of Greenwich (Woolwich, Avery Hill and Dartford)

Goldsmiths College

Imperial College of Science, Technology and Medicine

Kingston University

London Business School

The London Institute

London School of Economics and Political Science

London Guildhall University

Middlesex University (Bounds Green and Hendon)

University of North London

Ravensbourne College of Design and Communications

Roehampton Institute London

Royal College of Art

Royal College of Music

South Bank University (Main Campus and Wandsworth Road)

United Medical and Dental School of Guy's and St Thomas's Hospitals

University College London

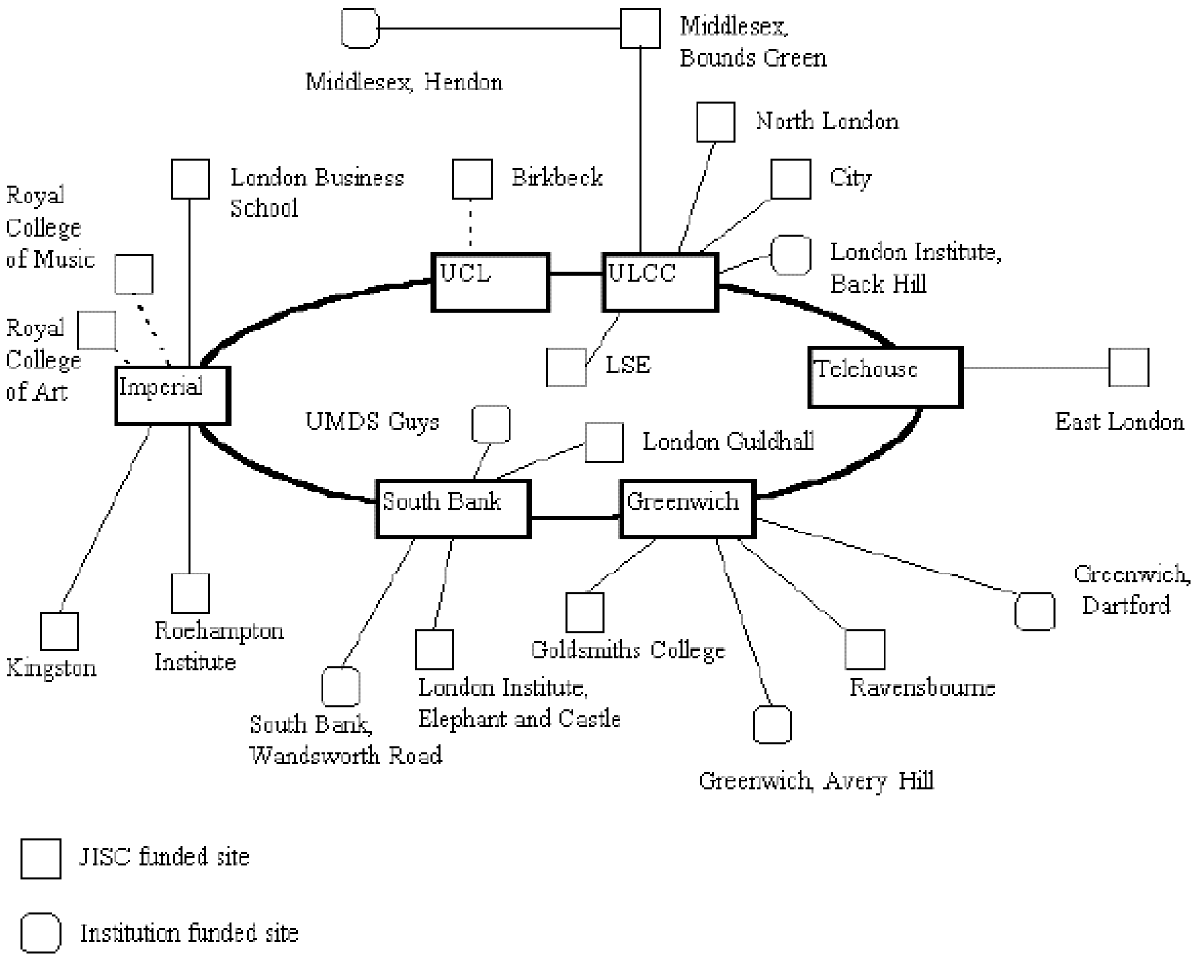
University of London Computer Centre (ULCC and Telehouse)

Services and technology

The primary requirement on day one of operation of the MAN was for a routed IP service. However, there is a clear requirement, within a short timescale, for a multi-service network that will support voice, video and data. Consequently, an Asynchronous Transport Service (ATM) network has been provided to all sites from the start. ATM is a technology specifically designed to carry delay critical traffic such as voice and video as well as more tolerant data traffic. Many of the features of the full service, such as dynamic allocation of bandwidth are not yet fully implemented, but ATM is now sufficiently stable for service deployment in a well managed environment. As well as the full IP service, the London MAN will offer a pilot ATM service to facilitate service and application development.

The core sites are connected by a dedicated 155Mbit/s ATM ring, and a router at each core site provides the IP service. The connections to leaf sites are 34Mbit/s PDH circuits. At the leaf site they terminate in ATM switches and there is also a router for IP routing.

The MAN supports a wide range of services. The IP service supports all standard INTERNET services, such as remote terminal access, file transfer, electronic mail and information access including the World Wide Web. With the bandwidth available on the MAN the IP service will also provide good support for the voice and video based services that are emerging. The ATM service will provide a guaranteed high quality video and voice services as well as an ability to allocate specific portions of the MAN bandwidth to particular applications or groups of users.



The initial topology of the London MAN

Teaching Applications

The basic services provide users of the network with a set of tools that will assist them in their learning, teaching and research activities. In the early days of computer networking the largest users were from the science and technology faculties, but now use of computers has spread to all disciplines. Some of the most demanding applications are in the arts subjects, where the network was of no use until it was capable of transferring high quality images at reasonable speeds. The use of electronic mail in the academic community is particularly widespread and, though it is relatively undemanding on network performance, it is an application that is heavily relied on.

Access to information servers, and particular the World Wide Web, is the fastest growing area of network usage. The availability and quality of information is still very variable from subject to subject, but something can be found in most areas. The traditional libraries are evolving into learning resource centres and rapidly embracing new forms of electronic delivery of information. Libraries and their customers will be major users of the MAN, which will enable learning resource service to collaborate effectively across institutional boundaries.

A major revolution can be expected in the way that teaching is delivered over the next ten years. Students will no longer be required to sit in the same lecture theatre as the lecturer or even to be participating at the same time. Remote delivery of teaching material will take a substantial time to develop and in the longer term will probably not even require the student to physically attend the university, with lectures being delivered to the home. The London MAN will provide video and associated services that will allow these teaching and learning methods to be investigated and developed. One benefit would be to reduce the amount of travelling between campuses of multi-site institutions, which is a particularly unpleasant activity in London. The MAN is also a mechanism to enable the collaborative development of course material involving multiple institutions.

Video conferencing has applications in most areas of HEIs activities. The technology is still evolving and there is also much to learn about the human factors involved. Pilot video conferencing facilities on the MAN will allow development of the technology at the same time as providing an initial service to those sites with immediate requirements. Conferencing facilities can be linked to the national SuperJANET service and commercial services in order to communicate with a larger community.

4.5.2. Connectivity between MANs

It is possible to link with other MANs in UK, and a selection is given in the list below. Other MANs are proposed so that in due course there will be a distribution of MANs around the whole of UK, connected together by a CONNECT highspeed network.

Aberdeen - AbMAN

Bristol

Edinburgh - EaStMAN

Fife and Tayside - FaTMAN

Glasgow - ClydeNET

Manchester - G-MING

South Wales

4.5.3. The Scottish MANs

This section describes the activities on one of the Scottish MANs, namely EastMAN, the Edinburgh and Stirling Metropolitan Area Network.

There is considerable interest in both public and private sectors in the construction of MANs in order to integrate networking provision, achieve high speed connectivity both locally and to external networks (such as the INTERNET) and to make cost savings over self provision. Indeed, early access to and acquisition of the networking infrastructure supporting high bandwidth is perceived as a pre-requisite for the exploitation of multi-media and multi-service applications which are now appearing in the marketplace.

The following paragraphs provide some informative detail on the initiative by the four Universities: Edinburgh, Heriot-Watt, Napier, Stirling and the three Higher Education Institutions: Edinburgh

College of Art, Moray House Institute of Education and Queen Margaret College to create EaStMAN, the Edinburgh and Stirling Metropolitan Area Network.

The network must be able to support new generation applications, for example, video conferencing, video on demand and multi-media for teaching, self directed learning and research as well as supporting the normal data and voice requirements of the MAN partners.

At this time, there are various MAN awareness programmes running to stimulate the user community into consideration and deployment of new and familiar applications which can be supported on the FDDI and ATM high-speed network.

Access to or connection to SuperJANET is an important service in the UK Academic Community. Connection to SuperJANET through high speed interfaces will allow access to and delivery of new services, e.g. electronic publishing, voice, and video, throughout the highly connected UK Academic Community.

Links to Further Education colleges and schools are seen as important in the short term, both in pursuit of our aim to increase access and develop schools liaison, and to establish and expand distance learning interests.

EaStMAN partners place great emphasis on the extensive use of technology in teaching and learning, and in the associated management and administration. Areas of particular interest include multi-media, computer-based learning environments and distance learning. The support of these areas requires access to the highest affordable bandwith and ATM (and is successors) is/are seen as the medium to long term bandwith development strategy.

All HEIs are aware of the changing nature of user applications towards support of multimedia capability, sharing of data structures, conferencing and so on. Of particular importance to the Universities is the potential for interactive distance learning through video in conjunction with image databases as well as the enhancement of research tools using the same types of applications. This change is accelerating. Naturally, such applications can require large amounts of bandwidth, if not to support a single application then to support the sheer numbers of its users. Indeed, the first network applications on SuperJANET all concern multimedia or visualisation experiments e.g., surgery teaching, supercomputer data visualisation and interaction, special datasets (rare document access), electronic journal testbed, remote consultation (pathology), remote sensing data (browsing earth image files), group communication (Pandora at Cambridge - networked advanced desktop) and so on.

We will also prepare for and stimulate multi-media applications such as computer-based teaching/learning and video-conferencing. Innovative teaching and learning systems are likely to be of increasing significance.

4.6 VIDEOCONFERENCING

Video conferencing will be broadly divided into ISDN and IP techniques. The difference is essentially one of guaranteeing the bandwidth available, and therefore the reliability of the quality of the picture. When using ISDN the user is able to obtain from the Telecom service provider a guaranteed bandwidth on the network on demand. The quality of the pictures is consistent with that bandwidth and any deterioration is more likely to be related to other factors in the network operation or the equipment performance which is being used. When IP is used then the bandwidth available to the link will vary with the number of users at any particular time on the network, and this cannot be guaranteed. Further if a new user comes onto the network during the session there is increased competition for the bandwidth. Thus the variability of the bandwidth leads to variation in the quality of the picture. As video-links are real-time, any reduction in the bandwidth available results in less information coming over the link and the picture quality falls. On the whole the ability to transmit video on the UK academic network, SuperJANET, is not impossible, but once one wishes to connect to sites outside UK there are major difficulties. If European connections are made over JAMES or the TransEuropean Networks then the situation is better.

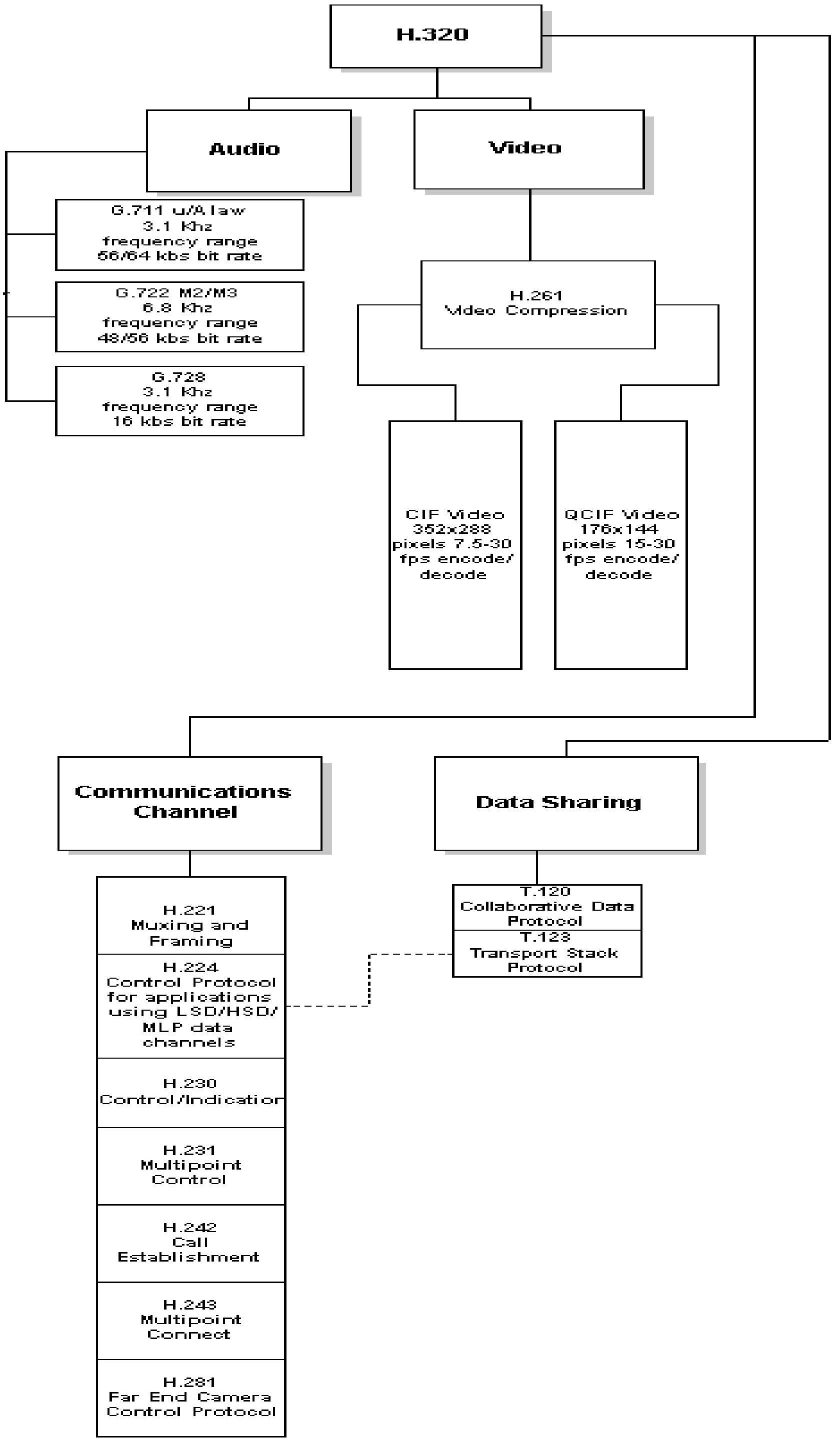
There are efforts being made to develop software tools which will guarantee a limited bandwidth over LANs and the INTERNET, and these will improve matters. The real solution is making the bandwidth available on the INTERNET much larger, but then the problem may be that the use will expand to fill the available bandwidth with no overall improvement, in a manner analogous to the way road traffic increases when better roads are provided. The position with IP link is that much work is being carried out to improve the service and this will continue into the future. The service currently available is probably tolerable for point-to-point operations and for individuals communicating, but the service is not sufficiently reliable for important teaching application in a general way.

Videoconferencing depends upon the analogue to digital conversion and compression of video and audio signals for their transmission between sites and then the decompression and digital to analogue conversion of the signals for their presentation at the remote site. There are several encoding/decoding protocols which are currently in common use (see section 2.7). There are a number of proprietary systems also in use and the user has to make a decision whether to use the accepted standards. If one is sure that always and forever you will only be communicating with equipment from a specific manufacturer then you can use their proprietary algorithms, but this will preclude communicating with the systems of another manufacturers; e.g. the standard algorithm for encoding video to be transmitted over ISDN is H.261 and all manufacturers implement this capability in their equipment, hence compatibility problems rarely exist currently between the equipment of the best known manufacturers.

Other encoding algorithms are MPEG-1 and MPEG-2. MPEG-1 was developed for CD-ROM applications and works in the region of 1-2Mbps. This produces video of a quality comparable with good domestic television, and at the upper level of the bandwidth range there is little delay in the movement in the picture. MPEG-1 visually is very similar to H.261. MPEG-1 systems are easily available and the encoding and decoding equipment is not expensive.

H.261 is normally used in ISDN transmissions and cover the range 128Kbps to 2Mbps. One finds that at the lower end of the range the picture is delayed and there is a severe lack of synchronisation between the movement of the lips of a person speaking and hearing the audio. As the bandwidth increases to 384Kbps this synchronisation problem largely disappears. This bandwidth corresponds to the ISDN-6 service, and the delay in movement is also very much reduced. These improvements will continue up to 768Kbps, but above that figure to the maximum of 2Mbps there is little perceptible advantage. The H.261 algorithm is mainly working on the white-noise levels of the original picture and this does not result in major improvements in the visual quality of the picture.

MPEG-1 and H.261 standards are averaging effects over several video frames and consequently the precision of control in interactive systems are more complicated.



Diagrammatic representation of the H.320 Standard

This diagram show all the individual standards that go to make the full videoconferencing standard H.320. This show that there are two fundamental streams for audio and video. The video compression algorithm is H.261, providing 2 resolutions, CIF and QCIF. In addition to these two sections, there is a communications channel and a Data-sharing channel. The communications channel deals with matters such as framing, establishing the communications links, supporting multi-point working and provision of control for cameras at the remote sites. The Data-sharing channel is concerned with the T.120 standard which supports collaborative data sharing and whiteboards.

Another algorithm is H.263 which is specially designed to improve picture quality at lower bandwidths and is applied in LAN systems. It is also used in some ISDN systems and does bring about a perceptible improvement in picture quality, see section 4.5.1.

MPEG-2 is a higher bandwidth algorithm for transmission rates above 2Mbps. The MPEG-2 standard was developed for cable systems where the encoding process required a powerful computer at the head-end of the distribution system, but the decoding was cheap and could be supplied easily in a large number of user systems without raising the costs significantly. The quality of image that can be achieved as very good but there are relatively few MPEG-2 systems in use routinely primarily because of the cost of encoding the video material.

Motion-JPEG is another algorithm which is used extensively in the Metropolitan Area Networks, e.g. it is used in the Scottish MANs and the quality of the video pictures are very high. This algorithm works primarily on each video frame in turn and with high processing power can process the video frame by frame in real-time. The high processing power means that the delay in the video is very small, but the equipment is more expensive.

4.6.1. The development of H.323 Standard for Videoconferencing:

In contrast to seeking high resolution images with high speed network, there is another field of development which concentrates on the H.263 video encoding algorithm which provides better images at the low bandwidth end of the spectrum. This algorithm was developed to encourage more efficient use in the region of 128Kbps and lower, and it also had applications on LANs in mind. Where the INTERNET was being used, and when the bandwidth had to be shared with other non-video applications it is important to have the best video possible in the restricted bandwidth.

The H.263 standard is discussed in section 2.7.2.2. H.263 is the coding algorithm underpinning the H.323 videoconferencing standard, and some people are of the opinion that in time it will replace H.320. H.323 is beginning to find its way into ISDN video conferencing systems and it produces a better picture at 128Kbps than H.320. However if it is applied at higher bandwidths the advantages are not so apparent.

This is a video coding standard and was published around 1995-6 , and it differs from H.261 in the following ways:-

> It uses half-pixel precision for motion compensation where H.261 used full pixel precision. Some parts of the hierarchical structure of the data-stream are now optional, so that the CODEC can be configured for a lower bit-rate., or better error recovery.
>
> There are four negotiable options to improve performance;
>
> 1. Unrestricted Motion Vectors
> 2. Syntax-based arithmetic coding
> 3. Advanced prediction.
> 4. Forward and backward frame prediction similar to MPEG.

H.263 is supposed to be capable of providing the same quality at half the bit rate to H.261, and also H.263 support 5 resolutions which enables it to compete with the MPEG standards.

4.6.2. Videoconferencing for Meetings.

The early experiment on the SuperJANET ATM Video network were with videoconferencing. The common concept was that all that was necessary was a camera from the local video shop and some loudspeakers as might be available with one of the PCs in the departments and all was ready. The experience of SuperJANET was that did not produce good results, the pictures were of poor definition, the audio was difficult to hear, and all kinds of things might happen, from unpleasant echo effects to howling as the microphones resonated.

The Audio-Visual departments in some of the principal SuperJANET sites in the first SuperJANET experiments were involved in setting up a teaching project and they were able to see that simply by implementing good audio-visual practice that simple remedies could be applied and these were very successful. Other problems like the malfunctioning of echo-cancellers were not so simple and it was decided to set up the Audio-Visual Consultancy based upon a collaboration of the Universities of Cambridge and Newcastle-upon Tyne with University College London.

This consultancy set in 1993 and it set about inspecting all the SuperJANET sites, to check there room design, the equipment being employed, and investigating how the audio-visual signals were being presented to the echo-canceller. Very quickly this became an investigation into the performance of the echo-cancellation equipment, and it was found that in a very significant number of cases the equipment was incorrectly installed. Once these investigations had been completed there was an important improvement in the quality of the pictures and audio on the network. There have been considerable improvement in equipment since those times, and present day equipment usually has competent echo-cancellation equipment incorporated. However if this equipment is to be sued in large room such as lecture theatres then often a free-standing echo-canceller is necessary. (Document about echo-cancellation in http://www.video.ja.net_

The lesson to be learned form this exercise was that the collaboration of the Network Groups and the Audio-Visual Groups was important to the successful running of the network and this collaboration has continued now for several years; it has led to the setting up of the Video-TAG, the publication of a number of papers on the WWW on aspects of video conferencing, the development of databases to help potential users about the equipment available, and the development of testing protocols for audio and video performance of videoconferencing equipment. This has now reached the stage that a national Videoconferencing advisory Service has been set up on 1st January 1998, run jointly by the University of Newcastle-upon-Tyne and University College London. Information about this advisory service is available at the URL: http://www.video.ja.net

4.6.3. Videoconferencing - Teaching versus Meetings:

One of the most important observation form this exercise was the realisation that videoconferencing for meetings was not the same as for teaching. This is important as increasingly people are looking to the possibilities of using video conferencing in teaching.

The following are a list of characteristics of a typical videoconferencing meeting:-

1. Most of the people taking part are of a similar status, and so the interaction takes place between “equals”
2. The session is usually required to be highly interactive.
3. In many case the meeting is point-to-point, although the number of multi-point meetings is increasing.
4. The meetings are often designed to meet a problem between people working at some distance apart and part of the reason for setting up the meeting has been to save on travel.

In teaching the participants have different roles.

1. There will be a teacher and there will be pupils who usually will regard the teacher as the authoritative figure.
2. The session will depend upon the teacher for the level of interaction, i.e. how well the interaction takes place depends upon how well the teacher stimulates the pupils and encourages them to interrupt with questions.
3. The sessions are likely to be multi-point as the use of videoconferencing is associated with a form of distance learning and teaching, and the motivation for using videoconferencing will be the ability to reach a number of small groups of students who are distributed widely.

4. The session are not usually designed to handle a specific problem but to deliver a series of tutorial/sessions over a period of time.

This list is by no means exhaustive but it give examples of the types of differences between teaching applications and meetings held on the network. There are other comparison which can be made between the use of videoconferencing in the universities and colleges over the ISDN network or the INTERNET. Currently many applications on the INTERNET involve desk-top systems, with a small personal video camera and small loudspeakers alongside the PC or UNIX machine. This type of system has developed a poor name because the same level of care has not been taken to ensure that the user works in good lighting conditions and so makes most efficient use of what is in fact a cheap video camera, and that the microphones are placed to avoid resonance with the loudspeakers and do not pick up extraneous noise from the surroundings. The quality of picture that can be obtained is directly related to the bandwidth available, and if similar bandwidths are available on the INTERNET as on ISDN the picture quality should be comparable.

At the present time the use of desk-top systems is attracting a lot of interest, and if the bandwidth available within campus networks is sufficient then one can expect increasing use. The issue is if too many people take up these facilities on the campus, then the network will suffer. People should remember that the video on these systems is often only providing 5-10 frames per second as opposed to conventional video which is 25-30 frames per second; and hence the images are jerky. It is the decision of the teacher whether that is acceptable for their purposes. In practice (and as recommended by the videoconferencing strategy document) a bandwidth of 384Kbps is necessary if ISDN is to be used between institution. At this speed the picture-movement is smooth and there is good synchronisation between the lip movements and the speech heard.

The current situation is that if the users in teaching applications are not experienced in the use of IP based videoconferencing then they should weigh up carefully whether they should attempt to use such as system for teaching. The ISDN systems provide a guaranteed bandwidth and therefore picture and speech quality, and they are supported by the large service providers such as BT. Local able companies are beginning to support ISDN. If the service is to be used to other parts of the globe then it is more reliable than the INTERNET because of the guaranteed bandwidth, although it may be difficult to obtain more than 128Kbps in the worst situations. It is not as simple to maintain the picture quality on the INTERNET and it does require people with some understanding of how the system works. No doubt in the near future this will improve, and the INTERNET is more appropriate for the research worker who can tolerate some level of inconvenience from time-to-time. Uni-cast links on the INTERNET are not difficult to establish and maintain, but multi-cast links are dependent on the routers in the network having the software to support multi-casting. This situation is rapidly changing and multi-cast will become commonplace in the near future.

A number of teaching projects have been carried out on both the ISDN and IP networks. In Scotland there are a large number of teaching projects being considered which will use the Scottish MANs. A survey and analysis of these project would be an interesting exercise to try to establish what are the issues involved in network teaching and what a are the lessons to be learned from this experience. The ISDN and MAN projects are following similar approaches in that special rooms have been designed for videoconferencing and these are used for the teaching. The room are not large and much of the teaching concerns small groups. The experience of the INSURRECT project was that this technique was more appropriate for small group teaching and in medicine could be used for the post-graduate training programme where there were large number of students distributed over a large area, but at each individual site there was only a small number of students. This type of teaching requires the facilities for the whole group to see on a large screen the video information being transmitted round the network. The Scottish MAN experiments are using IP for their teaching. Teaching on the INTERNET has been carried out using IP, but this has provided each student with a terminal as in the language training programme RELATE. There is a need fro considerable work to be done is this area, and after the current pilot for teaching with IP on the Mbone, there will be further larger scale projects funded through JISC later this year (1998).

Within the education domain the indications are that desk-top systems, working over the INTERNET can assist researchers and collaborating colleagues to communicate with each other easily and cheaply. The whiteboard is proving to be a very valuable development of this type of communications, where the document under discussion can be seen at each of the participating sites and any changes made at any of the sites can be seen by the other sites. Sometimes people using this type of communications do not use the video channel, using only the audio links and the data link (the document under discussion) - this is the basis of Microsoft's NET-MEETING, which is utilising the T.120 standard.

People should beware if they wish to transmit graphic documents between sites in a videoconferencing meeting or teaching. The ISDN bandwidths are limited and it may be necessary to have one ISDN link for graphics(i.e. data) and a second ISDN link for videoconferencing. ISDN is used extensively for transmitting data, it is not only used in videoconferencing.

4.6.4 The Technical Specification of a Videoconferencing System.

The Technical specifications of a video conferencing ystem conforming with the internationally recognized ITU H.320 standard guarantees that a system will function with any other standardized videoconferencing system:

Room systems,

Rollabout systems, studios

and multipoint videoconferencing bridges, etc.

Computer Requirements

- PC with Pentium processor or Macintosh with G3 processor (200Mhz plus)
- H.320 7" PCI format CODEC with built-in ISDN BRI (1 slot)
- Telephone handset
- Color camera with integrated microphone
- Camera connector and power supply
- Loudspeakers
- VGA or SVGA screen, or a scan convertor to television screen.

Video Requirements

- Camera features:
  - PAL color camera
  - Display resolution: 330 lines
  - Pixels: 500 (H) x 582 (L)
  - Min. sensitivity: 20 Lux F2.8
  - 6 mm lens and integrated microphone
  - Swivel mount
  - Adjustable focus
  - Retractable shutter for video secrecy
- Video input for PAL or, NTSC
- Video display:
  - 24 bits/pixel
  - 30 pictures/sec. in NTSC mode and 25 pictures/sec. in PAL mode
  - Max. size: 640 x 480
- Full-motion video compression:
  - H.261 CIF (352 x 288) and QCIF (176 x 144)

Audio Requirements

- Hands-free or via telephone handset
- Audio input:
  - microphone or camera microphone
  - telephone handset
- Audio output:
  - loud speakers or external output
  - telephone handset
- Software selection and control of audio inputs and outputs
- Sound coding 3.1KHz in G.711 (56Kbps) and G.728 (16Kbps)
- Sound coding 7KHz in G.722 (48 or 56Kbps)
- Lip synchronization with sound
- Echo cancellation:
  - 256 ms min. in G.711 and G.728
  - 256 ms min in G.722

Still pictures

- JPEG compression of local picture

Data transmission

- Dynamic allocation of H.221 frame data channels:
  - LSD 6.4 kbps, 14.4 kbps and 40 kbps
  - HSD 64 kbps
- X.25 multiplexing with routing of incoming calls

Communication protocols

- H.221 framing/deframing on two B channels and channel realignment
- H.230, H.242 protocol management
- H.231 and H.243 multipoint conferencing

Communication interfaces

- ISDN (2B+D) daughterboard: 2 x 64Kbps or 2 x 56Kbps
- MVIP connector:
  - ISDN PCI 2B card
  - ISO ENET card (isochronous Ethernet)

Application software:

The directory, whiteboard and file transfer functions can be used independent of videoconferencing mode:-

- Videoconference application:
  - Display of local or remote window
  - Picture in picture (PIP)
  - Video secrecy
  - Audio secrecy
  - Clipboard transfer
  - JPEG acquisition and compression of still picture
- Directory:
  - Automatic audio or video call set-up
  - Multi directory
- Shared whiteboard:
  - Vector drawing tools
  - Multidocuments
  - Partial screen capture
  - Messaging
  - Remote presentation
- INTERNET access: Planet PPP

5 FUTURE WORK

Two topics which have not been covered in this document are:-

Video Servers and Knowledge Pools

Document Sharing - T.120

These two topics are the subject of special studies to be carried out b y the videoconferencing Advisory Service within the next few months and so have not been presented in this study.

GLOSSARY

ADPCM (Adaptive Differential Pulse Code Modulation)

It is a compression technique which encodes the predictive residual instead of the original waveform signal so that the compression efficiency is improved by a predictive gain. Rather than transmitting PCM samples directly, the difference between the estimate of the next sample and the actual sample is transmitted. This difference is usually small and can thus be encoded in fewer bits than the sample itself.

A-LAW

A technique for encoding audio into an 8 bit word used in G.711 encoding.

ANSI (American National Standards Institute)

ANSI works with various organisations and manufacturers of telecommunications equipment to determine domestic standards not covered by ITU.

API (Application Programming Interface)

It is software from which user interfaces (e.g., pull down menus) can be created.

ARITHMETIC CODING

Perhaps the major drawback to each of the Huffman encoding techniques is their poor performance when processing texts where one symbol has a probability of occurrence approaching unity. Although the entropy associated with such symbols is extremely low, each symbol must still be encoded as a discrete value.

Arithmetic coding removes this restriction by representing messages as intervals of the real numbers between 0 and 1. Initially, the range of values for coding a text is the entire interval [0, 1]. As encoding proceeds, this range narrows while the number of bits required to represent it expands. Frequently occurring characters reduce the range less than characters occurring infrequently, and thus add fewer bits to the length of an encoded message.

ATM (Asynchronous Transfer Mode)

ATM is a switching/transmission technique where data is transmitted in small, fixed sized cells (5 byte header, 48 byte payload). The cells lend themselves both to the time-division-multiplexing characteristics of the transmission media, and the packet switching characteristics desired of data networks. At each switching node, the ATM header identifies a virtual path or virtual circuit that the cell contains data for, enabling the switch to forward the cell to the correct next-hop trunk. The virtual path is set up through the involved switches when two endpoints wish to communicate. This type of switching can be implemented in hardware, almost essential when trunk speed range from 45Mbps to 1Gbps.

The ATM Forum, a worldwide organization, aimed at promoting ATM within the industry and the end user community was formed in October 1991 and currently includes more than 500 companies representing all sectors of the communications and computer industries, as well as a number of government agencies, research laboratories.

AUDIO STANDARDS

These refer to the audio standards for the compression/decompression and transmission of P*64 audio signals. G.xxx classifications gives these standards.

AVC system

Audio/Video/Communication system.

BANDWIDTH

A measurement expressed in bits per second (bps) on the amount of information that can flow through a channel.

B-CHANNEL

Bearer Channels. A 64Kbps channel in an ISDN line.

BIT-RATE

The rate at which the compressed bit-stream is delivered from the storage medium to the input of a decoder.

BLOCK

An 8-row by 8-column matrix of pels, or 64 DCT coefficients (source, quantised or dequantised).

BRI (Basic Rate Interface)

A data rate standard for ISDN. It provides a customer with 144Kbps divided into three channels (2B channels carrying 64Kbps each and one D-channel assigned to 16Kb of signalling information).

BACKWARD MOTION VECTOR

A motion vector that is used for motion compensation from a reference picture at a later time in display order.

CAPABILITY SET

Within the H.242 standard, capability set is used to define the set of functions which the audio-visual end point supports. At the initiation of an audio-visual call, the end points exchange their respective sets of information and establish a call within the bounds of their mutual capability sets.

CCITT

Commite' Consultatif International de Telecommunications et Telegraphy A committee of the International Telecommunications Union responsible for making technical recommendations about telephone and data communication systems for PTTs and suppliers. Plenary sessions are held every four years to adopt new standards.

CIF (Common Image Format)

A format for displaying an image on a screen. CIF has an image resolution of 352 by 288 pixels at 30 frames per second. This format is optional within the H.261 standard.

The standardisation of the structure of the samples that represent the picture information of a single frame in digital HDTV, independent of frame rate and sync/blank structure. The uncompressed bit rates for transmitting CIF at 29.97 frames/sec is 36.45 Mbps.

CHI (Concentration Highway Interface)

Programmable time division multiplex communication bus used by Lucent Technologies for interconnecting telecommunication chips. Can be programmed to act as one channel of an MVIP.

CHI LOOP

A board which can connect to one or two MVIP or concentration highway connectors and emulates a common interface by supplying the clock and by connecting the CHI or MVIP back to itself (loop) or to another CHI or MVIP (cross connect). Useful for testing where actual communication lines are not available.

CHROMINANCE (component)

A matrix, block or single pel representing one of the two colour difference signals related to the primary colours in the manner defined in the bit-stream. The symbols used for the colour difference signals are Cr and Cb.

D-CHANNEL

Data Signalling Channel. A 16Kbps channel in an ISDN line.

DCT

Discrete Cosine Transform, used in Fourier Analysis

DLL (Dynamic Link Library)

DSM (Digital Storage Media)

A digital storage or transmission device or system.

DSP (Digital Signal Processing)

The technique of processing data as numbers instead of voltages.

ECHO CANCELLATION

A technique used to prevent a speaker from hearing a delayed version of his own speech. Echo cancellation is required in video telephony due to the delays required by the video.

ENTROPY

Entropy, the average amount of information represented by a symbol in a message, is a function of the model used to produce that message and can be reduced by increasing the complexity of the model so that it better reflects the actual distribution of source symbols in the original message.

Entropy is a measure of the information contained in message, it's the lower bound for compression.

FFT

Fast Fourier Transform

FIELD

For an interlaced video signal, a field is the assembly of alternate lines of a frame. Therefore an interlaced frame is composed of two fields a top field and a bottom field.

FORWARD MOTION VECTOR

A motion vector that is used for motion compensation from a reference picture at an earlier time in display order.

FRAME PERIOD

The reciprocal of the frame rate.

FRAME RATE

The rate at which frames are be output from the decoding process.

FUTURE REFERENCE PICTURE

A future reference picture is a reference picture that occurs at a later time than the current picture in display order.

G.711

A standard for compressing and decompressing audio (50 -3000Hz) into a 48, 56, or 64Kbps stream.

G.722

A standard for compressing and decompressing audio (50 - 7000 Hz) into a 48, 56, or 64Kbps stream.

G.728

A standard for compressing and decompressing audio (50 - 3000 Hz) into a 16Kbps stream.

G.SERIES

The family of audio-related ITU standards. It includes G.711, G.722, and G.728.

H.SERIES

The family of ITU standards for use of video equipment (over 64 to 1920Kbps channels) during conferencing. Frequently referred to as P*64.

H.281

The ITU-T standard for far end camera control in an H.320 conference.

H.320

The ITU recommended standard for narrow-band visual telephone systems and terminal equipment.

HUFFMAN CODING

For a given character distribution, by assigning short codes to frequently occurring characters and longer codes to infrequently occurring characters, Huffman's minimum redundancy encoding minimises the average number of bytes required to represent the characters in a text.

Static Huffman encoding uses a fixed set of codes, based on a representative sample of data, for processing texts. Although encoding is achieved in a single pass, the data on which the compression is based may bear little resemblance to the actual text being compressed.

Dynamic Huffman encoding, on the other hand, reads each text twice; once to determine the frequency distribution of the characters in the text and once to encode the data. The codes used for compression are computed on the basis of the statistics gathered during the first pass with compressed texts being prefixed by a copy of the Huffman encoding table for use with the decoding process.

By using a single-pass technique, where each character is encoded on the basis of the preceding characters in a text, Gallager's adaptive Huffman encoding avoids many of the problems associated with either the static or dynamic method.

INTERLACE

The property of conventional television frames where alternating lines of the frame represent different instances in time.

INTRA CODING

Coding of a macroblock or picture that uses information only from that macroblock or picture.

I/O SPACE

Input/Output Space. The address at which the computer communicates with an add-in card.

I-PICTURE (INTRA-CODED PICTURE)

A picture coded using information only from itself.

ISDN (Integrated Services Digital Network)

ISDN is a CCITT term for a relatively new telecommunications service package. ISDN is basically the telephone network turned all-digital end to end, using existing switches and wiring (for the most part) upgraded so that the basic call is a 64Kbps end-to-end channel, with bit-diddling as needed. Packet and maybe frame modes are thrown in for good measure, too, in some places. It's offered by local telephone companies, but most readily in Australia, France, Japan, and Singapore, with the UK and Germany somewhat behind, and USA availability rather spotty.

A Basic Rate Interface (BRI) is two 64K bearer (B) channels and a single delta (D) channel. The B channels are used for voice or data, and the D channel is used for signalling and/or X.25 packet networking. This is the variety most likely to be found in residential service. Another flavour of ISDN is Primary Rate Interface (PRI). Inside the US, this consists of 24 channels, usually divided into 23 B channels and 1 D channel, and runs over the same physical interface as T1. Outside of the US then PRI has 31 user channels, usually divided into 30 B channels and 1 D channel. It is typically used for connections such as one between a PBX and a CO or IXC.

ITU

International Telecommunications Union, formerly CCITT, a body of the United Nations.

LUMINANCE (component)

A matrix, block or single pel representing a monochrome representation of the signal and related to the primary colours in the manner defined in the bitstream. The symbol used for luminance is Y.

MACROBLOCK

The four 8 by 8 blocks of luminance data and the two (for 4:2:0 chroma format), four (for 4:2:2 chroma format) or eight (for 4:4:4 chroma format) corresponding 8 by 8 blocks of chrominance data coming from a 16 by 16 section of the luminance component of the picture. Macroblock is sometimes used to refer to the pel data and sometimes to the coded representation of the pel values and other data elements defined in the macroblock header. The usage should be clear from the context.

MOTION COMPENSATION

The use of motion vectors to improve the efficiency of the prediction of pel values. The prediction uses motion vectors to provide offsets into the past and/or future reference pictures containing previously decoded pel values that are used to form the prediction error signal.

The book Motion analysis for Image Sequence Coding by G.Tziritas and C.Labit documents the technical advances made through the years in dealing with motion in image sequences.

MOTION ESTIMATION

The process of estimating motion vectors during the encoding process.

MOTION VECTOR

A two-dimensional vector used for motion compensation that provides an offset from the coordinate position in the current picture to the coordinates in a reference picture.

MVIP

Multi-Vendor Integration Protocol. An 8 channel time division multiplex communication bus which can be used to connect various digital communication boards in a PC.

NON-INTRA CODING

Coding of a macroblock or picture that uses information both from itself and from macroblocks and pictures occurring at other times.

NT1

Network Terminating Device. The ISDN telephone line from your local exchange carrier connects to your system through an NT1. An NT1 performs network performance and integrity checks. It enables loopback testing which verifies your digital line is connected and working properly. Outside the US the NT1 is considered part of the network and is installed by the telephone company. In the US, it is considered Customer Premises Equipment (CPE). It is common to change the NT1 when trying to diagnose a dead connection if you previously had a working connection. Some hardware manufacturers build the NT1 into their device; this may limit your flexibility in adding services to your ISDN line. The NT1 is microprocessor controlled and requires its own power source.

NTSC (National Television System Committee)

USA video standard with image format 4:3, 525 lines, 60 Hz and 4 Mhz video bandwidth with a total 6 Mhz of video channel width. NTSC uses YIQ NTSC-1 was set in 1948. It increased the number of scanning lines from 441 to 525, and replaced AM-modulated sound with FM.

PEL

Picture element.

PICTURE

Source, coded or reconstructed image data. A source or reconstructed picture consists of three rectangular matrices of 8-bit numbers representing the luminance and two chrominance signals.

For progressive video, a picture is identical to a frame, while for interlaced video, a picture can refer to a frame, the top field or the bottom field of the frame depending on the context.

PREDICTION

The use of a predictor to provide an estimate of the pel value or data element currently being decoded.

P-PICTURE (Predictive-coded picture)

A picture that is coded using motion compensated prediction from past reference pictures.

PREDICTION ERROR

The difference between the actual value of a pel or data element and its predictor.

PREDICTOR

A linear combination of previously decoded pel values or data elements.

PROFILE

A defined sub-set of the syntax of a specification.

QCIF RESOLUTION

Quarter Common source Intermediate Format (1/4 CIF , i.e. luminance information is coded at 144 lines and 176 pixels per line at 30 frames per second.). The uncompressed bit rates for transmitting QCIF at 29.97 frames/sec is 9.115 Mbit/s. This format is required by the H.261 standard.

QUANTISATION MATRIX

A set of sixty-four 8-bit values used by the dequantiser.

QUANTISED DCT COEFFICIENTS

DCT coefficients before dequantisation. A variable length coded representation of quantised DCT coefficients is storedas part of the compressed video bitstream.

QUANTISER SCALE

A scale factor coded in the bitstream and used by the decoding process to scale the dequantisation.

REFERENCE PICTURE

Reference pictures are the nearest adjacent I or P pictures to the current picture in display order.

SCALABILITY

Scalability is the ability of a decoder to decode an ordered set of bitstreams to produce a reconstructed sequence. Moreover, useful video is output when subsets are decoded. The minimum subset that can thus be decoded is the first bitstream in the set which is called the base layer. Each of the other bitstreams in the set is called an enhancement layer. When addressing a specific enhancement layer, lower layer refer to the bitstream which precedes the enhancement layer.

SPID

Service Profile Identifiers which are used to identify what sort of services and features the switch provides to the ISDN device. When a new subscriber is added, the service representative will allocate a SPID just as they allocate a directory number. The subscriber needs to input the SPIDs into their terminal device before they will be able to connect to the central office switch (this is referred to as initialising the device).

SUB BAND CODING

Sub-band coding for images has roots in work done in the 1950s by Bedford and on Mixed Highs image compression done by Kretzmer in 1954. Schreiber and Buckley explored general two channel coding of still pictures where the low spatial frequency channel was coarsely sampled and finely quantized and the high spatial frequency channel was finely sampled and coarsely quantized. More recently, Karlsson and Vetterli have extended this to multiple subbands. Adelson et al. have shown how a recursive subdivision called a pyramid decomposition can be used both for compression and other useful image processing tasks.

A pure sub-band coder performs a set of filtering operations on an image to divide it into spectral components. Usually, the result of the analysis phase is a set of sub-images, each of which represents some region in spatial or spatio-temporal frequency space. For example, in a still image, there might be a small sub-image that represents the low-frequency components of the input picture that is directly viewable as either a minified or blurred copy of the original. To this are added successively higher spectral bands that contain the edge information necessary to reproduce the original sharpness of the original at successively larger scales. As with DCT coder, to which it is related, much of the image energy is concentrated in the lowest frequency band.

For equal visual quality, each band need not be represented with the same signal-to-noise ratio; this is the basis for sub-band coder compression. In many coders, some bands are eliminated entirely, and others are often compressed with a vector or lattice quantizer. Succeedingly higher frequency bands are more coarsely quantized, analogous to the truncation of the high frequency coefficients of the DCT. A sub-band decomposition can be the intraframe coder in a predictive loop, thus minimizing the basic distinctions between DCT-based hybrid coders and their alternatives.

T.120

Defines a series of communication protocols for data conferencing. The protocol closest to the hardware is T.123 which provides reliable transport of data between end points, over various types of media including modems, ISDN, and LANs. T.125, Multipoint Control Service, coordinates and synchronises the various participants in a multipoint call. T.124, Generic Conference Control, provides setup and control of a conference. T.126 provides whiteboarding and graphic image annotation. T.127 provides file transfer.

TEMPORAL SCALABILITY

A type of scalability where an enhancement layer also uses predictions from pel data derived from a lower layer using motion vectors. The layers have identical frame rates size, and chroma formats, but can have different frame rates.

TOP FIELD

One of two fields that comprise a frame of interlaced video. Each line of a top field is spatially located immediately above the corresponding line of the bottom field.

VARIABLE BITRATE

Operation where the bitrate varies with time during the decoding of a compressed bitstream.

Although variable bit rate is acceptable for plain linear playback, one important consideration not to use variable bit rate is that reasonably quick random access becomes nearly impossible. There is no table of contents or index in MPEG. The only tool the play back system has for approximating the correct byte position is the requested play back time stamp and the bit rate of the MPEG stream. MPEG streams do not encode their play back time.

To approximate an intermediate position in a variable bit rate stream, the play back system must grope around near the end of the stream to calculate the playback time, and assume the stream is approximately constant bit rate. The groping around for the correct position can take several seconds.

This is not appropriate for an interactive presentation or game. This groping around is at least annoying when trying to view a portion of a movie but it's not even possible for video streams because there are no time stamps (the SMPTE time codes in video streams need not to be continuous or unique).

Audio streams are always fixed bit rate.

VIDEO COMPRESSION

A video image is compressed to minimise the amount of space or data needed to store or transmit the image.

VIDEO DECOMPRESSION

To take a compressed video image and restore it to the size and format needed to view the video image.

VIDEO OVERLAY DEVICE

This is a logical device which overlays analogue video into a window on a VIA display. It may also perform frame grabbing, including saving and loading frames to and from a disk.

YCbCr

The international standard CCIR-601-1 specifies eight-bit digital coding for component video, with black at luma code 16 and white at luma code 235, and chroma in eight-bit two's complement form centred on 128 with a peak at code 224. This coding has a slightly smaller excursion for luma than for chroma: luma has 219 risers compared to 224 for Cb and Cr. The notation CbCr distinguishes this set from PbPr where the luma and chroma excursions are identical.

For Rec. 601-1 coding in eight bits per component,

Y_8b = 16 + 219 * Y

Cb_8b = 128 + 112 * (0.5/0.886) * (Bgamma - Y)

Cr_8b = 128 + 112 * (0.5/0.701) * (Rgamma - Y)

Some computer applications place black at luma code 0 and white at luma code 255. In this case, the scaling and offsets above can be changed accordingly, although broadcast-quality video requires the accommodation for headroom and footroom provided in the CCIR-601-1 equations.

CCIR-601-1 Rec. calls for two-to-one horizontal subsampling of Cb and Cr, to achieve 2/3 the data rate of RGB with virtually no perceptible penalty. This is denoted 4:2:2. A few digital video systems have utilized horizontal subsampling by a factor of four, denoted 4:1:1. JPEG and MPEG normally subsample Cb and Cr two-to-one horizontally and also two-to-one vertically, to get 1/2 the data rate of RGB. No standard nomenclature has been adopted to describe vertical subsampling. To get good results using subsampling you should not just drop and replicate pixels, but implement proper decimation and interpolation filters.

YCbCr coding is employed by D-1 component digital video equipment.

YPbPr

If three components are to be conveyed in three separate channels with identical unity excursions, then the Pb and Pr colour difference components are used:

Pb = (0.5/0.886) * (Bgamma - Y)

Pr = (0.5/0.701) * (Rgamma - Y)

These scale factors limit the excursion of EACH colour difference component to -0.5 .. +0.5 with respect to unity Y excursion: 0.886 is just unity less the luma coefficient of blue. In the analog domain Y is usually 0 mV (black) to 700 mV (white), and Pb and Pr are usually +- 350 mV.

YPbPr is part of the CCIR Rec. 709 HDTV standard, although different luma coefficients are used, and it is denoted E'Pb and E'Pr with subscript arrangement too complicated to be written here.

YPbPr is employed by component analog video equipment such as M-II and BetaCam; Pb and Pr bandwidth is half that of luma.

YIQ

The U and V signals above must be carried with equal bandwidth, albeit less than that of luma. However, the human visual system has less spatial acuity for magenta-green transitions than it does for red-cyan. Thus, if signals I and Q are formed from a 123 degree rotation of U and V respectively [sic], the Q signal can be more severely filtered than I (to about 600 kHz, compared to about 1.3 MHz) without being perceptible to a viewer at typical TV viewing distance. YIQ is equivalent to YUV with a 33 degree rotation and an axis flip in the UV plane. The first edition of W.K. Pratt "Digital Image Processing", and presumably other authors that follow that bible, has a matrix that erroneously omits the axis flip; the second edition corrects the error.

Since an analog NTSC decoder has no way of knowing whether the encoder was encoding YUV or YIQ, it cannot detect whether the encoder was running at 0 degree or 33 degree phase. In analog usage the terms YUV and YIQ are often used somewhat interchangeably. YIQ was important in the early days of NTSC but most broadcasting equipment now encodes equiband U and V.

The D-2 composite digital DVTR (and the associated interface standard) conveys NTSC modulated on the YIQ axes in the 525-line version and PAL modulated on the YUV axes in the 625-line version.

YUV

In composite NTSC, PAL or S-Video, it is necessary to scale (B-Y) and (R-Y) so that the composite NTSC or PAL signal (luma plus modulated chroma) is contained within the range -1/3 to +4/3. These limits reflect the capability of composite signal recording or transmission channel. The scale factors are obtained by two simultaneous equations involving both B-Y and R-Y, because the limits of the composite excursion are reached at combinations of B-Y and R-Y that are intermediate to primary colours. The scale factors are as follows:

$U = 0.493 * (B - Y)$

$V = 0.877 * (R - Y)$

U and V components are typically modulated into a chroma component:

$C = U*\cos(t) + V*\sin(t)$

where t represents the ~3.58 MHz NTSC colour sub-carrier. PAL coding is similar, except that the V component switches Phase on Alternate Lines (+-1), and the sub-carrier is at a different frequency, about 4.43 MHz.

It is conventional for an NTSC luma signal in a composite environment (NTSC or S-Video) to have 7.5% setup :

$Y\_setup = (3/40) + (37/40) * Y$

A PAL signal has zero setup. The two signals Y (or Y_setup) and C can be conveyed separately across an S-Video interface, or Y and C can be combined (encoded) into composite NTSC or PAL:

$NTSC = Y\_setup + C$

$PAL = Y + C$

U and V are only appropriate for composite transmission as 1-wire NTSC or PAL, or 2-wire S-Video. The UV scaling (or the IQ set, described below) is incorrect when the signal is conveyed as three separate components. Certain component video equipment has connectors labelled YUV that in fact convey YPbPr signals.

2B+D DATA RATE

This refers to a Basic Rate Interface, control i.e., a data rate standard for ISDN. It provides a customer with 144 kilobytes divided into three channels (2B channels carrying 64 kilobytes each and one D-channel assigned to 16 kilobytes of signalling information), I.441.

APPENDIX-1 Video Coding Algorithm

Most image or video applications involving transmission or storage require some form of data compression to reduce the otherwise inordinate demand on bandwidth and storage. The principle of data compression is quite straightforward. Virtually all forms of data contains redundant elements. The data can be compressed by eliminating those redundant elements with various compression methods. However, when compressed data are received over a communications link, it must be possible to expand the data back to the original form. As long as the coding scheme is such that the code is shorter than the eliminated data, compression will still occur.

1. Video compression

Video compression is a process whereby a collection of algorithms and techniques replace the original pixel-related information with more compact mathematical descriptions. Decompression is the reverse process of decoding the mathematical descriptions back to pixels for display. At its best, video compression is transparent to the end user.

There are two types of compression techniques:

<u>Lossless</u>

A compression technique that creates compressed files that decompress into exactly the same file as the original. Lossless compression is typically used for executables applications and data files for which any change in digital make-up renders the file useless. Lossless compression typically yields only about 2:1 compression, which barely dents high-resolution uncompressed video files.

<u>Lossy</u>

Lossy compression, used primarily on still image and video image files, creates compressed files that decompress into images that look similar to the original but are different in digital make up. This "loss" allows lossy compression to deliver from 2:1 to 300:1 compression. A wide range of lossy compression techniques is available for digital video.

In addition to lossy or lossless compression techniques, video compression involves the use of two other compression techniques:

<u>Interframe Compression</u>

Compression between frames (also known as temporal compression because the compression is applied along the time dimension).

<u>Intraframe Compression</u>

Compression within individual frames (also known as spatial compression).

Some video compression algorithms use both interframe and intraframe compression. For example, MPEG uses JPEG, which is an intrafame technique, and a separate interframe algorithm. Motion-JPEG uses only intraframe compression.

1.1. Interframe compression.

Interframe compression uses a system of key and delta frames to eliminate redundant information between frames. Key frames store an entire frame, and delta frames record only changes. Some implementations compress the key frames, and others don't. Either way, the key frames serve as a reference source for delta frames. Delta frames contain only pixels that are different from the key frame or from the immediately preceding delta frame. During decompression, delta frames look back to their respective reference frames to fill in missing information.

All interframe compression techniques derive their effectiveness from interframe redundancy. Low-motion video sequences, such as the head and shoulders of a person, have a high degree of redundancy, which limits the amount of compression required to reduce the video to the target bandwidth. Until recently, interframe compression has addressed only pixel blocks that remained static between the delta and the key frame. Some new CODECs increase compression by tracking moving blocks of pixels from frame to frame. This technique is called motion compensation. The data that is carried forward from key frames is dynamic.

Although dynamic carry forwards are helpful, they cannot always be implemented. In many cases, the capture board cannot scale resolution and frame rate, digitise, and hunt for dynamic carry forwards at the same time. Dynamic carry forwards typically mark the dividing line between hardware and software CODECs.

1.2. Intraframe compression.

Intraframe compression is performed solely with reference to information within a particular frame. It is performed on pixels in delta frames that remain after interframe compression and on key frames. Although intraframe techniques are often given the most attention, overall CODEC performance relates more to interframe efficiency than intraframe efficiency. The following are the principal intraframe compression techniques.

<u>Null suppression</u>

Is one of the oldest, and simplest, data compression techniques. A common occurrence in text is the presence of a long string of blanks in the character stream. The transmitter scans the data for strings of blanks and substitutes a two-character code for any string that is encountered. While null suppression is a very primitive form of data compression, it has the advantage of being simple to implement. Further more, the payoff, even from this simple technique, can be substantial (gains of between 30 and 50 percent).

<u>Run Length Encoding (RLE)</u>

A simple lossless technique originally designed for data compression and later modified for facsimile. RLE compresses an image based on "runs" of pixels. Although it works well on black and white facsimiles, RLE is not very efficient for colour video, which have few long runs of identically coloured pixels.

2. JPEG

A standard that has been adopted by two international standards organisations: the ITU (formerly CCITT) and the ISO. JPEG is most often used to compress still images using discrete cosine transform (DCT) analysis. First, DCT divides the image into 8x8 blocks and then converts the colours and pixels into frequency space by describing each block in terms of the number of colour shifts (frequency) and the extent of the change (amplitude). Because most natural images are relatively smooth, the changes that occur most often have low amplitude values, so the change is minor. In other words, images have many subtle shifts among similar colours but few dramatic shifts between very different colours. Next, quantisation and amplitude values are categorised by frequency and averaged. This is the lossy stage because the original values are permanently discarded. However, because most of the picture is categorised in the high-frequency/low-amplitude range, most of the loss occurs among subtle shifts that are largely indistinguishable to the human eye. After quantization, the values are further compressed through RLE using a special zigzag pattern designed to optimise compression of like regions within the image. At extremely high compression ratios, more high-frequency/low-amplitude changes are averaged, which can cause an entire pixel block to adopt the same colour. This causes a blockiness artefact that is characteristic of JPEG-compressed images. JPEG is used as the intraframe technique for MPEG.

3. Vector quantization (VQ)

A standard that is similar to JPEG in that it divides the image into 8x8 blocks. The difference between VQ and JPEG has to do with the quantization process. VQ is a recursive, or multi-step algorithm with inherently self-correcting features. With VQ, similar blocks are categorised and a reference block is constructed for each category. The original blocks are then discarded. During decompression, the single reference block replaces all of the original blocks in the category. After the first set of reference blocks is selected, the image is decompressed. Comparing the decompressed image to the original reveals many differences. To address the differences, an additional set of reference blocks is created that fills in the gaps created during the first estimation. This is the self-correcting part of the algorithm. The process is repeated to find a third set of reference blocks to fill in the remaining gaps. These reference blocks are posted in a lookup table to be used during decompression. The final step is to use lossless techniques, such as RLE, to further compress the remaining information. VQ compression is by its nature computationally intensive. However, decompression, which simply involves pulling values from the lookup table, is simple and fast.

4. MPEG

MPEG addresses the compression, decompression and synchronisation of video and audio signals. In most general form, an MPEG system stream is made up of two layers:

<u>System layer</u>

The system layer containing timing and other information needed to de-multiplex the audio and video streams and to synchronise audio and video during playback.

<u>Compression Layer</u>

The compression layer includes the audio and video streams.

The system decoder extracts the timing from the MPEG system stream and sends it to the other system components. The system decoder also de-multiplexes the video and audio streams from the system stream; then sends each to the appropriate decoder. The video decoder decompresses the video stream as specifies in part 2 of the MPEG standard. The audio decoder decompresses the audio stream as specifies in part 3 of the MPEG standard.

The MPEG standard defines a hierarchy of data structures in the video stream.

Video Sequence

Begins with a sequence header (may contain additional sequence headers), includes one or more groups of pictures, and ends with an end-of-sequence code.

Group of Pictures (GOP)

A header and a series of one or more pictures intended to allow random access into the sequence.

Picture

The primary coding unit of a video sequence. A picture consists of three rectangular matrices representing luminance (Y) and two chrominance (Cb and Cr) values. The Y matrix has an even number of rows and columns. The Cb and Cr matrices are one-half the size of the Y matrix in each direction (horizontal and vertical).

Slice

One or more ``contiguous'' macroblocks. The order of the macroblocks within a slice is from left-to-right and top-to-bottom. Slices are important in the handling of errors. If the bitstream contains an error, the decoder can skip to the start of the next slice. Having more slices in the bitstream allows better error concealment, but uses bits that could otherwise be used to improve picture quality.

Macroblock

A 16-pixel by 16-line section of luminance components and the corresponding 8-pixel by 8-line section of the two chrominance components.

Block

A block is an 8-pixel by 8-line set of values of a luminance or a chrominance component. Note that a luminance block corresponds to one-fourth as large a portion of the displayed image as does a chrominance block.

The MPEG audio stream consists of a series of packets. Each audio packet contains an audio packet header and one or more audio frames.

Each audio packet header contains following information:

- Packet start code Identifies the packet as being an audio packet.
- Packet length Indicates the number of bytes in the audio packet.

An audio frame contains the following information:

- Audio frame header Contains synchronisation, ID, bit rate, and sampling frequency information.
- Error-checking code Contains error-checking information.
- Audio data Contains information used to reconstruct the sampled audio data.
- Ancillary data Contains user-defined data.

5. Inter-Picture Coding

Much of the information in a picture within a video sequence is similar to information in a previous or subsequent picture. The MPEG standard takes advantage of this temporal redundancy by representing some pictures in terms of their differences from other (reference) pictures, or what is known as inter-picture coding. This section describes the types of coded pictures and explains the techniques used in this process.

Picture Types

The MPEG standard specifically defines three types of pictures: intra, predicted, and bi-directional.

Intra Pictures

Intra pictures, or I-pictures, are coded using only information present in the picture itself. I-pictures provide potential random access points into the compressed video data. I-pictures use only transform coding (as explained in the Intra-picture (Transform) Coding section) and provide moderate compression. I-pictures typically use about two bits per coded pixel.

Predicted Pictures

Predicted pictures, or P-pictures, are coded with respect to the nearest previous I- or P-picture. This technique is called forward prediction. Like I-pictures, P-pictures serve as a prediction reference for B-pictures and future P-pictures. However, P-pictures use motion compensation (see the Motion Compensation section) to provide more compression than is possible with I-pictures. Unlike I-pictures, P-pictures can propagate coding errors because P-pictures are predicted from previous reference (I- or P-) pictures.

Bi-directional Pictures

Bi-directional pictures, or B-pictures, are pictures that use both a past and future picture as a reference. This technique is called bi-directional prediction. B-pictures provide the most compression and do not propagate errors because they are never used as a reference. Bi-directional prediction also decreases the effect of noise by averaging two pictures.

6. Video Stream Composition

The MPEG algorithm allows the encoder to choose the frequency and location of I-pictures. This choice is based on the application's need for random accessibility and the location of scene cuts in the video sequence. In applications where random access is important, I-pictures are typically used two times a second.

The encoder also chooses the number of B-pictures between any pair of reference (I- or P-) pictures. This choice is based on factors such as the amount of memory in the encoder and the characteristics of the material being coded. For example, a large class of scenes have two bi-directional pictures separating successive reference pictures.

The MPEG encoder reorders pictures in the video stream to present the pictures to the decoder in the most efficient sequence. In particular, the reference pictures needed to reconstruct B-pictures are sent before the associated B-pictures.

7. Motion Compensation

Motion compensation is a technique for enhancing the compression of P- and B-pictures by eliminating temporal redundancy. Motion compensation typically improves compression by about a factor of three compared to intra-picture coding. Motion compensation algorithms work at the macroblock level.

When a macroblock is compressed by motion compensation, the compressed file contains this information:

- The spatial vector between the reference macroblock(s) and the macroblock being coded (motion vectors)
- The content differences between the reference macroblock(s) and the macroblock being coded (error terms)

Not all information in a picture can be predicted from a previous picture. Consider a scene in which a door opens: The visual details of the room behind the door cannot be predicted from a previous frame in which the door was closed. When a case such as this arises--i.e., a macroblock in a P-picture cannot be efficiently represented by motion compensation--it is coded in the same way as a macroblock in an I-picture using transform coding techniques (see Intra-picture (Transform) Coding Section).

The difference between B- and P-picture motion compensation is that macroblocks in a P-picture use the previous reference (I- or P-picture) only, while macroblocks in a B-picture are coded using any combination of a previous or future reference picture.

Four codings are therefore possible for each macroblock in a B-picture:

- Intra coding: no motion compensation
- Forward prediction: the previous reference picture is used as a reference
- Backward prediction: the next picture is used as a reference
- Bi-directional prediction: two reference pictures are used, the previous reference picture and the next reference picture.

Backward prediction can be used to predict uncovered areas that do not appear in previous pictures.

8. Intra-picture (Transform) Coding

The MPEG transform coding algorithm includes these steps:

- Discrete cosine transform (DCT)
- Quantization
- Run-length encoding

Both image blocks and prediction-error blocks have high spatial redundancy. To reduce this redundancy, the MPEG algorithm transforms 8 x 8 blocks of pixels or 8 x 8 blocks of error terms from the spatial domain to the frequency domain with the Discrete Cosine Transform (DCT).

Next, the algorithm quantises the frequency coefficients. Quantization is the process of approximating each frequency coefficient as one of a limited number of allowed values. The encoder chooses a quantization matrix that determines how each frequency coefficient in the 8 x 8 block is quantised. Human perception of quantization error is lower for high spatial frequencies, so high frequencies are typically quantized more coarsely (i.e., with fewer allowed values) than low frequencies.

The combination of DCT and quantization results in many of the frequency coefficients being zero, especially the coefficients for high spatial frequencies. To take maximum advantage of this, the coefficients are organised in a zigzag order to produce long runs of zeros. The coefficients are then converted to a series of run-amplitude pairs, each pair indicating a number of zero coefficients and the amplitude of a non-zero coefficient. These run-amplitude pairs are then coded with a variable-length code, which uses shorter codes for commonly occurring pairs and longer codes for less common pairs.

Some blocks of pixels need to be coded more accurately than others. For example, blocks with smooth intensity gradients need accurate coding to avoid visible block boundaries. To deal with this inequality between blocks, the MPEG algorithm allows the amount of quantization to be modified for each macroblock of pixels. This mechanism can also be used to provide smooth adaptation to a particular bit rate.

9. Synchronisation

The MPEG standard provides a timing mechanism that ensures synchronisation of audio and video. The standard includes two parameters: the system clock reference (SCR) and the presentation time-stamp (PTS).

The MPEG-specified ``system clock'' runs at 90KHz. System clock reference and presentation time-stamp values are coded in MPEG bitstreams using 33 bits, which can represent any clock cycle in a 24-hour period.

10. System Clock References

An SCR is a snapshot of the encoder system clock which is placed into the system layer of the bitstream. During decoding, these values are used to update the system clock counter in the CL480.

11. Presentation Time-stamps

Presentation time-stamps are samples of the encoder system clock that are associated with video or audio presentation units. A presentation unit is a decoded video picture or a decoded audio time sequence. The PTS represents the time at which the video picture is to be displayed or the starting playback time for the audio time sequence.

The decoder either skips or repeats picture displays to ensure that the PTS is within one picture's worth of 90 KHz clock tics of the SCR when a picture is displayed. If the PTS is earlier (has a smaller value) than the current SCR, the decoder discards the picture. If the PTS is later (has a larger value) than the current SCR, the decoder repeats the display of the picture.

12. Conclusion

Transmission costs are the most substantial portion of most data communications and voice communication budgets. With these compression techniques being used, communication costs are being reduced considerably.

BIBLIOGRAPHY

References on Multimedia and Videoconferncing Standards

http://www.dtsys.com/005dtva.htm (6k) - Descrition of MPEG-2 standard.

http://www.stud.ee.ethz.ch/~rggrandi/ (2k) - A comprehensive review of Video compression techniques and standards

http://www2.msstate.edu/~fhm1/ETMM2.html (24k) A review of multimedia standards.

http://www.cis.ohio-state.edu/~fine/ISDN/references.html - a reference list on ISDN which is a useful lead in.

http://www.itu.ch - website of the International Telecommunications Union , with details of all ITU standards.

References on ISDN

http://alumni.caltech.edu:80/~dank/isdn - useful information on ISDN

http://www.pacbell.com/ISDNbook - a comprehensive website on many aspects of videoc9onferencing including ISDN.

References on the Mbone and Multicasting

MBone: Multicasting Tomorrow's INTERNET by Kevin Savetz, Neil Randall, and Yves Lepage (ISBN: 1-56884-723-8)

Published by IDG Books Worldwide, Inc. An International Data Group Company

919 E. Hillsdale Blvd., Suite 400, Foster City, CA 94404

Casner, Steve. "Frequently Asked Questions (FAQ) on the Multicast Backbone," 6 May 1993, available via anonymous FTP from venera.isi.edu:/mbone/faq.txt

Casner, Steve and Schulzrinne, Henning. "RTP: A Transport Protocol for Real-Time Applications," IETF Draft, 20 October 1993.

Casner, Stephen and Deering, Stephen. "First IETF INTERNET Audiocast," ACM SIGCOMM Computer Communication Review, San Diego California, July 1992, pp. 92-97.

Comer, Douglas E. INTERNET working with TCP/IP, volume I, Prentice-Hall, New Jersey, 1991.

Deering, Stephen. "Host Extensions for IP Multicasting", RFC 1112, August 1989.

Deering, Stephen. "MBONE-The Multicast Backbone," CERFnet Seminar, 3 March 1993.

Moy, John. "Multicast Extensions to OSPF," IETF Draft, July 1993.

Perlman, Radia. Interconnections: Bridges and Routers, Addison-Wesley, New York, 1993, p. 258.

Curtis, Pavel, MBone map, available via anonymous FTP from parcftp.xerox.com:pub/net-research/mbone-map-big.ps

References on Networks

http://www.lonman.net.uk - details of the London Metropolitan Area Network

http://www.ed.ac.uk/EUCS/NSD/EastMAN-pages - details of the EASTMAN (Edinburgh and Stirling Metropolitan Area Network)

http://www.ja.net - website of JANET (Joint Academic Network)

References on Projects

http://www.video.ja.net - description of the Videoconferencing Advisory Service, which includes a number of databases and document on videoconferencing.

http://www.video ja.net, - section on IP Software tools (1998)

http://www.cu-seeme.com/pub/iw.html, - CU-SeeMe for Macintosh and Windows videoconferencing (1997)

http://www-mice.ca.ucl.ac.uk/mice/mbone-review - review and references on the Mbone and the MICE project (Multimedia Interactive Conferencing in Europe).