Videoconferencing Audio & Video Equipment

Author: Jim Sheach, MALTS, University of Edinburgh
Version: 1

1. OVERVIEW

The aim of this document is to give some guidance on the selection and installation of audio and video equipment that will interface with a standalone CODEC for use in videoconferencing. It is intended primarily for technical staff, and outlines the basic requirements and highlights some of the problems that can be encountered. It is strongly recommended that audiovisual specialists should be employed for any installation work.

Where recommendations are given for particular manufacturers of audio, video and distribution equipment this does not imply that other manufacturers' equipment not mentioned is unsuitable. Those suppliers listed are, however, known through the practical experience of the author to produce reliable products that meet specification. Prices indicated should be used only as a rough guide, but were reasonably accurate at the time of writing.

Most purchasers of a videoconferencing system to be used over ISDN or IP networks will choose a complete and ready to operate room-based or portable system. Another approach, common in higher/further education is to custom design a videoconferencing facility around a standalone CODEC. In this case items of audio and video equipment, i.e. microphones, loudspeakers, television cameras, picture monitors, video recorders, video/data projectors, room control systems etc., are purchased separately and connected to the CODEC to complete the system. The advantages of this approach are that the system can be customised to meet the actual needs of a specific application or made flexible in operation, (e.g. to interface a single CODEC with more than one conferencing room), or style of use (such as traditional meeting room videoconferencing or remote lecture and tutorial presentation). The disadvantages are that interfacing of audio/video equipment is not straightforward and, without specialised knowledge, problems can easily occur. In addition the installation will take much longer and probably be more expensive than a complete package system bought 'off the shelf'.

The traditional method of room system design is to use the CODEC as a simple Coding and Decoding device with single audio and video I/P (input) and O/P (output) connections. Video source selection and audio mixing of microphone and other audio sources are carried out using video and audio equipment external to the CODEC. In such a scenario the CODEC plays little part in the operation of the system once the videoconference call has been established.

Recent developments in CODEC technology have had an impact on the design of room-based systems. Modern CODECs are available with multiple inputs for remote pan and tilt cameras, document cameras, VCRs and PCs. In addition the CODEC may alter the video transmission protocol used to transmit each image depending on the type of video input which is selected. When the document camera or PC is selected the CODEC maximises image resolution to 4CIF at the expense of frame rate, as these are predominantly still images. In the case of a camera or VCR, frame rate is optimised and resolution reduced as these signals contain significant movement. In addition some CODECs allow the transmission of two simultaneous independent video signals (Duo Video or Dual Image). In order to take advantage of these advanced CODEC features it is necessary to connect each video source directly into the CODEC and carry out source selection within the CODEC. Many CODECs support multiple microphone inputs or have facilities to 'daisy chain' CODEC manufacturers proprietary microphones into a single microphone input. Utilising multiple proprietary microphones can provide additional facilities including mute buttons and voice activated camera tracking. An alternative in room based system design is to fully utilise the facilities within the CODEC for source selection and remote camera control, in addition to proprietary microphones. A room control system is still required in this case to simplify the overall control of the system including the CODEC, cameras, video/data projector, VCR, room audio system etc. but in this case the overall video and audio installation is simplified while utilising the advanced features of the CODEC.

Whilst the remainder of this document deals primarily with the first of the two scenarios outlined above many of the principles apply equally to the simpler system utilising the extensive facilities available on a modern CODEC.

Other VTAS documents cover related videoconferencing areas and are available on the VTAS web.

Videoconferencing Rooms deals with the conference environment. Evaluation of ISDN/IP Videoconferencing Equipment is concerned with ISDN and IP rollabout, desktop, portable and separate CODEC systems.

---

2. HEALTH AND SAFETY

Statutory Regulations: The advice given in this document is offered in good faith, but it is the responsibility of individuals/organisations following any or all of this advice to ensure that they comply with all of the relevant statutory regulations and safety guidelines. All electrical wiring and installation will need to comply with the IEE wiring regulations[1] and be carried out by qualified staff.

[1] The Institution of Electrical Engineers (IEE) Wiring Regulations 16th Edition (with amendments)
available from: The IEE, P.O. Box 96, Stevenage, Hertfordshire, SG1 2SD

3. SIGNAL QUALITY

To establish an effective videoconferencing link between sites, the sound and vision signals being transmitted and received must be of sufficiently high quality to enable clear, accurate communication without distracting the participants. This signal quality will depend on:

- the physical, electrical and environmental characteristics of the conference room;
- the quality of the audio visual and videoconferencing equipment in use;
- the correct adjustment of the audio visual and videoconferencing equipment;
- the type, and satisfactory adjustment, of the network linking the conferencing sites.

4. INSTALLATION OF AUDIO AND VIDEO EQUIPMENT

The installation of audio and video equipment is a specialised area that demands the skills of qualified and trained staff. While this document hopefully points prospective purchasers in the right direction, it must be emphasised that whenever possible specialist audiovisual staff should carry out the installation so as to minimise the potential technical problems. This is particularly true for the interconnection of audio equipment, where problems are commonplace and by no means always straightforward to rectify.

As a very general rule, video equipment is complex but the interconnections for videoconferencing are usually straightforward. Audio equipment, on the other hand, is normally much less complex but the interconnections require more in-depth knowledge to avoid problems.

5. THE VIDEOCONFERENCING ENVIRONMENT

These recommendations are intended for organisations planning the installation of a small to medium sized conference room, accommodating up to 20 persons. The principles apply to most videoconferencing situations, but teaching to large groups of students (more than 100) in a lecture theatre brings particular problems associated with the delivery of satisfactory sound and television pictures.

The participants at the local site need to see and hear clearly the participants at the remote site (and vice versa), so the important equipment items are: television cameras, picture monitors, microphones and loudspeaker/amplifier units. The size and layout of the room are also important factors.

For a small group, e.g. six persons, one large screen monitor (70 cm diagonal) will produce an image large enough to be viewed comfortably by everyone. For 20 persons a much larger image or several smaller viewing monitors will be needed. The size of the room will dictate the size and power output capacity required for the loudspeaker/amplifier units. Camera position and the type of camera lens are also dependent on room layout, as is the placement of microphones.

Videoconferencing from large lecture theatres poses a much greater challenge than videoconferencing from a meeting room. Sound amplification is normally needed to provide adequate coverage for the theatre microphones and sound sources (e.g. video replay) in addition to that required for the remote conference sound. The reverberation time (i.e. the amount of echo in the room) is usually also much greater than in a small room. These two factors militate against effective conferencing by increasing the tendency for acoustic feedback (i.e. howl-around) and introducing an amount of echo that normal videoconference echo cancellers cannot handle.

Effective sound pickup from a lecture theatre audience is extremely difficult unless a roving or highly directional microphone is used, both of which require an operator. A large audience will also require a very large picture image to view the remote site, such as that provided by a video projector. As the image produced from such projection is significantly less intense than that produced from a conventional picture monitor, the ambient room light will need to be dimmed for acceptable viewing. Unfortunately the lecture theatre video camera (providing audience images to the remote site) requires a fairly high level of lighting to produce good quality pictures. The introduction of lecture theatre video/data projectors with a light output of 3000 ANSI Lumens may allow the auditorium lights to be operated at a suitable level to provide adequate pictures from the lecture theatre. In addition automatic cameras will not normally produce an accurately framed image of members of the audience, so the cameras will need operators if close up images of audience members are required. Sometimes videoconferencing sessions need to be relayed to large lecture theatres where the audience are passive observers. In this case, only sound and vision signals from the remote site need to be transmitted to the lecture theatre for reproduction and display. This is much easier to achieve: it is when interaction (i.e. return sound and vision) is required from the lecture theatre environment that the main problems arise.

6. ROOM CONTROL SYSTEMS

6.1 Introduction

Where a CODEC is to be integrated with audio and video equipment to form a room system the control of such a system can become complex. It is possible to use a number of individual infra red remote controls and cabled control panels to operate such a system but this will be difficult and confusing for the occasional user. An alternative approach is to utilise a room control system that brings together the control of all AV equipment in the system onto one control panel. Merging the control of all the AV equipment in such a way can ensure that the most complex of systems may be 'self operated' by a conference participant without technical support.

6.2 Methods of Remote Control

All professional audio visual equipment and an increasing number of domestic products incorporate remote control facilities. In the case of professional equipment this is normally provided by a serial data connection complying with the RS232 protocol. This not only provides remote control but also reports the status of the equipment back to the control system. Where professional equipment does not support RS232 communications or domestic products utilising infra red (IR) control are used then the control is only one-way i.e. there is no facility to report back equipment status.

6.3 Control Systems

Room control systems normally comprise an equipment rack, which contains the main system processor and a number of control cards. The exact number and type of control ports will depend on the complexity of the system but will normally include RS232 ports, Infra Red (IR) ports and relay switch contacts. The control software is custom developed to take into account the particular AV equipment installed as part of the system. Control of all AV functions is available on either a touch panel LCD screen or a push button panel. LCD touch panels as shown in Figure1 have the added advantage that they may have multilayered menu structures, which facilitates increased levels of system complexity. In the illustration below the room control system interfaced with the following AV equipment:

- Two cameras with integral pan and tilt units
- SVHS VCR
- Video Data Projector
- Video and Audio Matrix
- Stereo Audio Amplifier
- Volume/Microphone Mute Control Box

In this particular example the room control system was not interfaced to the CODEC, this decision was to simplify operation for the user.


Figure 1. Room Control System

Manufacturers	• Panja AMX • Crestron
Price Guide	• Control System including monochrome touchpanel and custom programming, £4000 - £5000

7. VIDEO EQUIPMENT

7.1 Introduction

To achieve the high quality of pictures transmitted by the broadcasters requires a considerable level of expertise in the areas of lighting, scenery, etc., and operation and alignment of equipment. It is, however, relatively easy to produce acceptable pictures in a fixed setting (e.g. a conference room) with inexpensive cameras that are self-adjusting, provided that the recommendations outlined here and in the paper 'Videoconferencing Rooms' are followed.

7.2 The Transmission of Video Information

Video signals may be transmitted in a number of different ways.

7.2.1 Composite Video

Analogue television signals transmitted to domestic television receivers carry four components: the black and white signal, the colour signal, synchronisation information and the sound signal. These four signals are carried by a radio frequency carrier to enable transmission through the atmosphere.

After tuning in the receiver (to select the appropriate channel) the received signal is separated into audio (the sound signal) and video (vision signal); at this point the video signal is termed composite video. This is the most common video format used for distribution of vision signals.

The term `composite' is used because three signals: the luminance (black and white) information, the chrominance (colour) information and the synchronisation information, are combined (coded) into a single signal. As only one signal is involved, composite video is a very convenient method for processing, distributing and recording television pictures. A disadvantage is that the coding and decoding processes necessary to combine and separate the three signals also introduce artefacts (or distortions) that degrade the final images.

Composite video signals are carried on coaxial cable with a central signal core and a surrounding screen that is normally earthed through the terminal equipment. Connectors associated with composite video signals fall into two categories, BNC connectors on high-end professional equipment and Phono or RCA connectors on some professional and most domestic equipment.

7.2.2 S-video

To reduce some of the coding artefacts found in composite video the signal is split into two signals: luminance (Y), and chrominance (C). This is termed S-video (or Y/C) and is capable of producing higher quality images, provided that all the equipment in the chain (i.e. camera through to picture monitor) is designed to handle it. The disadvantage is that two signals require distribution, processing and recording, which increases the cost. If equipment is fitted with S-video connections this should be used in preference to composite to improve picture quality.

S-video or Y/C cable contains two independent screened cables in a single casing and is normally terminated in a four pin Mini DIN connector. These connectors lack the robustness of BNC or Phono connectors so care must be exercised when making S-video connections.

7.2.3 Component Video

If the colour information is separated further into two independent signals, even higher image quality is possible as the coding/decoding processes are now reduced to a minimum. However this means that, together with the luminance signal, there are now three component signals demanding three signal paths to be provided throughout the transmission chain. Equipment using Component Video is normally equipped with multiple BNC connectors for the three component signals.

7.2.4 PAL Composite Video

It has been mentioned that a composite video signal has black and white, colour, and synchronising information combined or coded together to form the signal. The way in which the colour signals are coded is determined by the colour system in use. In the United Kingdom, Phase Alternate Line (PAL) coding is used. Other colour coding systems are in use around the world, e.g. National Television Systems Committee (NTSC) in the United States, and Sequential Colour with Memory (SECAM) in France.

7.3 Video Sources

A videoconferencing room may contain several video sources:

- a camera giving a wide angle general view;
- a second camera with remote pan and tilt, providing a close up of the participant speaking;
- a document camera or 'Visualiser' to transmit hard copy, three dimensional objects or overhead transparencies;
- a video playback machine to replay pre-recorded sequences.

7.4 Genlocking/Timing/Colour Phasing

With multiple video sources, a means to select (i.e. switch) between sources must be introduced. The vision mixer or switcher achieves this.

In order to provide a clean switch between sources all the video signals fed into the vision mixer or switcher must be locked together. Where the video signals are not locked there will be a disturbance on the signal at the O/P of the switcher during transitions between sources. This normally induces the picture on a monitor to roll following the switch as the monitor relocks to the new signal. In order to avoid such picture disturbances it is necessary to use Genlock techniques to lock all video sources.

Before entering the vision switcher each video signal must meet certain conditions.
It must be:

- synchronous (i.e. running on the same time base);
- timed at the input to the mixer (i.e. all synchronising pulses for the separate sources arriving at the same time, i.e. with identical signal delay);
- colour phased at the input to the mixer (i.e. all colour reference signals arriving in phase).

Video sources are made synchronous by using a Genlock unit to lock the source time base to a reference time base. Good quality cameras, document cameras etc. will have the option of a Genlock unit. A means of adjusting timing and colour phase is also normally provided on these Genlock units. An additional problem is caused by low cost VCRs such as SVHS and VHS that add a considerable amount of timing jitter to the vision signal. This is usually unacceptable to videoconferencing CODECs. To overcome the problem a Time-Base Corrector (TBC) is required to remove the jitter and stabilise the off tape signal. Genlock and TBCs are normally associated with professional and broadcast VCRs and not domestic equipment.

7.5 Categories of Video Equipment

The market breaks down into three convenient categories:

- **domestic** - low cost and mass produced, but the quality may be acceptable;
- **industrial/commercial** - medium price, more durable and to higher specifications, generally the preferred choice;
- **broadcast** - very high specification and price; may be the only choice for very specialised applications, e.g. wide bandwidth conferencing where high resolution signals are a necessity.

7.6 Digital/Analogue Signals

The network which transports the sound and vision between the sites is normally digital. Within the conference room, the connections between audio and video equipment will, at present, almost certainly be analogue; this will gradually change to digital as more digital processes are introduced into these areas. The CODEC is the interface: all inputs to the CODEC will normally be analogue; all outputs from the CODEC into the network will be digital.

Digital techniques are used extensively in broadcasting studios as they introduce very little degradation into signal processing, distribution and recording. However, at present this is significantly more expensive than analogue methods. For other than high bandwidth videoconferencing the higher quality of digital equipment would be masked by the limitations of the intervening network so, until costs fall to the level of analogue, there are few incentives to change.

When digital studio products achieve more market penetration it is quite feasible that a substantial part of the compression and signal processing presently taking place within the videoconferencing CODEC will be incorporated in the cameras, etc., thus reducing the complexity and cost of the CODECs. Currently, for videoconferencing over normal capacity networks, analogue studio products have a large price advantage; however this situation is likely to change over the next few years and needs to be watched closely.

7.7 Video Cameras

Basically, a television camera converts light images into an electrical signal so that the images may be distributed, recorded and displayed. The camera lens focuses the images on a photosensitive area, normally a Charge-Coupled Device (CCD). Inexpensive cameras use one CCD or chip, sections of which (usually stripes) are arranged to be sensitive to either the red, green or blue components of light. As only one third of the chip area is sensitive to a single colour, there are some technical limitations.

More expensive cameras use three separate chips, one for each colour component; the whole chip area can then be dedicated to the single colour. This produces significantly better results, especially in terms of resolution and signal to noise ratio. These cameras are also less likely than a single chip camera to introduce spurious artefacts.

For most videoconferencing applications, especially where data transmission rates are below 2 Mbit/s, the higher quality of three chip cameras is unlikely to be noticed. For high data rate networks where high resolution is important, e.g. the transmission of medical radiographs (the images generated from an x-ray scanner), then a three chip camera will bring significant improvements in image quality. For videoconferences using ISDN or IP networks, data rates are normally below 2 Mbit/s and in such cases single chip cameras are adequate.

The introduction of cameras with pan and tilt heads integrated into a single unit has had a significant impact on this type of installation. This combined unit can simplify the installation as both camera and pan and tilt functions are controlled from a single IR remote control or a single RS232 port on the room control system. However it should be noted that cameras of this type do not normally include Genlock facilities.

Important Factors	• Horizontal resolution in excess of 400 lines (at the centre). • Signal to noise ratio better than 45 dB. • Sensitivity 1400 Lux at a lens stop of f-5.6 or greater (no camera gain in circuit). • Choosing the appropriate lens for the application. • A Genlock facility (to synchronise the camera to a reference source). • An S-video output (if this option is being used).
Manufacturers	• JVC • Panasonic • Sony
Price Guide	• Single chip CCD cameras, £500-£1,500 • Three chip cameras, £2,500-£25,000

7.8 Document Viewers

A document viewer, sometimes called an imager or visualiser, enables documents, three dimensional objects, radiographs or overhead transparencies to be introduced into a videoconference (some models have provision for single 35 mm slides). The device consists of an illuminated baseboard, lit either from above (incident light) for documents or from below (transmitted light) for transparencies, radiographs, etc., with a camera pointing vertically down and focused onto the baseboard.

Usually a simple switching arrangement allows two further video sources (e.g. cameras) to be selected as an alternative to the visualiser camera signal. In low budget rooms this facility can act as a simple vision switcher, allowing the presenter to select camera sources at will, thus avoiding the need to purchase a separate vision mixer/switcher. As documents, radiographs and overhead transparencies vary greatly in contrast it is essential that an auto iris lens be fitted, together with auto camera gain, so that these wide variations in intensity can be accommodated with minimum user intervention.

Most document viewers or visualisers use a single chip video camera and output both composite video and S-video as standard. An alternative approach is to use a camera with a progressive scanning technique, which can provide higher resolution images and output an XGA signal in addition to composite and S-video. These products are excellent for still images and documents but due to the scanning nature of the display do not cope well with image movement, eg the rotational display of a three dimensional object. Several CODECs are now equipped with a 15 pin high density D type connector to facilitate the connection and display of a PC or any other device with an SVGA or XGA output.

Important Factors	• Horizontal resolution in excess of 400 lines (centre). • Signal to noise ratio better than 45 dB. • Sensitivity 3 Lux at a lens stop of f-1.4 (with camera gain). • Lens should have - auto focus and be switchable to manual focus - minimum 10:1 power zoom range - lens stops to f-1.4. • Genlock facility. • All units will have incident (i.e. top) lighting, normally fluorescent tubes, to display documents, etc.

	• Transmitted (bottom) lighting is also essential for overhead transparencies, radiographs, etc.
Manufacturers	• Barco • Elmo • JVC • Wolfvision • Samsung
Price Guide	• £2,500-£12,000

7.9 Slide Viewers

These enable 35 mm transparencies to be introduced as illustrations during a conference. Basically they consist of a slide projector optically coupled to a video camera. In general the degree of adjustment of iris, black level and colour balance required to produce an acceptable image from each slide mitigates against the use of such equipment. The display of still images from a PC and shared with remote sites using data sharing may be a more efficient and effective way of dealing with still images.

Important Factors	• Some models have provision for only a few slides; to load/unload/adjust slides during a live conference can prove uncontrollable and, at the least, inconvenient. • A zoom facility can be useful for a poorly framed slide. • Colour balance adjustment can also improve a poor slide image.
Manufacturers	• Elmo • Tamron
Price Guide	• ~£2,000

7.10 Vision Switcher/Mixer

When multiple vision sources are installed then a vision switcher/mixer will be necessary to select the intended camera, document camera, VCR, etc. for transmission. For the transition between cameras to be flawless (i.e. without loss of synchronism, or disturbance to the picture) all sources must be timed, Genlocked and colour phased at the input to the vision switcher/mixer (see 7.4 above), i.e. they must be fully synchronous.

Only simple mixers or more likely simple video switchers are necessary for videoconferencing. Sophisticated broadcast types of switcher/mixer, which offer a host of special effects, are unsuitable for most videoconferencing applications as these special effects can produce artefacts during the compression/decompression processes in the CODEC.

Camera sources, being purely electronic, are relatively stable in operation, i.e. timing fluctuations in or jitter of their vision signals are very small. Thus, provided the signals are Genlocked, timed and colour phased, they pass through the vision mixer with ease and are termed stable sources.

Video cassette players have an electro-mechanical tape path, which introduces a great deal of instability to the vision signals. This instability is too great to allow the signals to be Genlocked and colour phased, so the signals are termed asynchronous. Some CODECs are unable to process these unstable signals but several will play unstable video from a domestic VHS VCR successfully. To reduce this instability to an acceptable level requires an expensive device called a time base corrector (TBC), which electrically reduces the signal jitter and synchronises the player to a reference signal.

Another approach to handling unstable signal sources is to design a vision mixer around two digital television picture stores; these picture stores can memorise a full television picture. Essentially, an asynchronous or unstable source is digitised and written into one picture store with clock pulses derived from the unstable source. This information is then read out of the store with clock pulses generated from a stable (camera type) reference source. If this process is repeated for a second source (a camera or another asynchronous source), then the two memorised digital images can be read out together using a stable clock pulse, and the sources may be cut, mixed or whatever in complete synchronisation. There is no need to Genlock sources if this frame store type of mixer is used.

To minimise any potential switching artefacts during vision switcher transitions (i.e. cuts or mixes), the transitions can be delayed to occur always in the vertical interval between consecutive television frames. This facility is termed vertical interval switching: a very desirable feature. However an undesirable outcome of the use of a frame store mixer is that it delays the video signal by one frame when compared to the audio signal presented to the CODEC. In all videoconferencing systems video and audio signal synchronisation, also called lip sync, is very important. The lack of lip sync can be very disturbing to view and impact on the flow of the conference. As such any process that contributes to a lack of audio and video synchronisation should be avoided.

Simple switchers in addition to having source selection may have RS232 remote control. This will allow the switcher to be controlled from a room control system.

In summary there are three options available:

Frame Store Mixer:
Complex in operation and adds delay to the video but provides flawless transitions.

Simple Switcher with Genlocked Sources:
Simple to operate and integrate with room control system and provides flawless transitions, but requires expensive TBC on VCR sources.

Simple Switcher with non Genlocked Sources:
Simple to operate and integrate with room control system but there will be picture disturbances on each video switch. The performance of this system is similar to that of a rollabout CODEC where multiple cameras and a VCR are connected directly to the CODEC with source selection taking place within the CODEC.

Important Factors	• Frequency response 0 - 5.5 MHz ± 0.5 dB. • Signal to noise ratio better than 65 dB. • Differential phase, differential gain less than 2°, 2% respectively (JESSICA). • Vertical interval switching. • S-video inputs. • RS232 Remote Control.
Manufacturers	• Panasonic - frame store mixer/switcher (no Genlock necessary). • Panasonic - conventional switchers

	(sources require Genlocking for flawless switching). • Autopatch - conventional switchers (sources require Genlocking for flawless switching). • Kramer - conventional switchers (sources require Genlocking for flawless switching).
Price Guide	• £400-£25,000

7.11 Picture Monitors

These are categorised into three grades:

1. Broadcast quality, with a high specification and price. These include special facilities for high quality signal monitoring, with such features as under-scanning the image to see all the edges of the picture.
2. Good quality monitors with defined specifications; may have under-scanning.
3. Domestic receivers, perhaps with a more robust metal case.

The signal compression during videoconferencing reduces the overall picture quality to a level that is normally exceeded by the potential picture quality of a domestic receiver (grade i). Unless the special facilities provided by a professional or broadcast quality monitor are considered necessary there is little point in paying considerably more money for grades ii and iii.

Important Factors	• Ensure the monitor is of the appropriate size for the intended viewing distance (see Videoconferencing Rooms, section 2.7). • Ensure that the appropriate video interfaces are fitted (i.e. S-video if this video format is being used). • Most domestic televisions from mainstream manufacturers are of such high quality that the picture will be adequate for videoconferencing.
Manufacturers	• Barco

	• JVC • Panasonic • Philips • Sony
Price Guide	• Domestic £300-£1,500 • Grade 2 monitors £1,000+ • Broadcast monitors £3,000+

7.12 Camera Pan and Tilt Heads

The introduction of cameras with an integrated pan and tilt unit has had a major impact on small videoconferencing installations. Where systems require cameras with Genlock or special lenses, for example in large lecture theatre installations, separate pan and tilt units are required.

Important Factors	• In videoconferencing both horizontal movement (pan) and vertical movement (tilt) are very important. • Ensure that both the pan and tilt movements are smooth and controllable. • Many remote pan and tilt heads are designed for outdoor industrial surveillance/security, where quiet and smooth operation are not necessary requirements. These are too noisy, and the movement too jerky for conference use.
Manufacturers	• Dennard • Molyneux • Videmech
Price Guide	• £500-£1,000

7.13 Video Players/Recorders

Currently most conferencing rooms will use either VHS or S-VHS analogue format cassette machines for recording and playing sequences, as these are the most popular. Less popular domestic formats, 8 mm or Hi-8, will also be seen. Digital formats, that offer many advantages but at present cost rather more, are gradually establishing themselves on the market, in particular MiniDV.

The time base instability of video players has been mentioned in 7.10 above. Some CODECs are unable to handle unstable vision sources however most modern CODECs will successfully transmit video from a domestic VCR.

Manufacturers	• JVC • Panasonic • Philips • Sony
Price Guide	• Domestic VHS £200-£700 • S-VHS £700+ • Digital £1,000+

7.14 Scan Converters for Computers

Within videoconferences there is a requirement to use PC presentations and applications, which can be transmitted to other sites in the videoconference. This may be achieved in a number of ways.

Where a videoconferencing system is T.120 capable a PC may be connected to the CODEC at either site using a serial data connection. The data transmission between PCs will occur over a data channel contained within the overall videoconference transmission signal pass band. For example in an ISDN conference the available bandwidth will be divided between video, audio and data. An alternative is to use a data connection over the wide area network or the Internet to connect PCs within the videoconferencing suites and separate the data connection from the videoconference channel. This is generally known as a data sharing conference. Where T120 or data sharing is not available the signal from the PC may be converted into video format and sent across the normal video channel. This is achieved by converting the computer (PC) signal into a composite video (or S-video) signal. This composite signal is then processed as any other video input to the vision switcher/mixer.

Several CODECs also include the ability to connect the SVGA O/P of a PC or laptop directly into the CODEC thus avoiding the requirement for an additional scan converter, however this function may only operate between CODECs from the same manufacturer and may not operate in MCU multisite conferences.

Important Factors	• There are a large number of scan converters on the market, ranging from software solutions to comprehensive digital picture store devices. The performance of these products is very variable, so an evaluation prior to purchase is strongly recommended. From our experience the manufacturers listed below can be relied on. • As only one video signal is normally sent at a time the receiving videoconference suite will see only the presenter OR the PC presentation, T120 or Data Sharing allows the display of the presenter AND the presentation at the same time.
Manufacturers	• Analog Way • Vine • Extron • Sony • Tektronix
Price Guide	• Software versions from £150 • Hardware versions £300-£2000

7.15 Data Projectors

Due to the poor performance of scan converters as described in section 7.14 and the desire to share video and computer presentations simultaneously, the installation of a data projector and screen in medium to large videoconference suites is becoming the norm. The main criteria for projector selection in the videoconferencing environment are that it should be bright, (better than 2000 ANSI-Lumens) and quiet in operation (less than 45 dBA).

Normal operation of data projection during presentations would require dimming the room lights however it is a requirement for the videoconferencing cameras that the participants are adequately lit. Careful consideration of screen position, lighting, and projector selection can provide high quality projection whilst not impacting on the transmitted images from the videoconferencing cameras.

The introduction into the videoconferencing environment of any unwanted noise will have two distinct effects. If the room microphones pick up the noise then it will be a distraction to the other sites in the conference and may have an impact on effective video switching during a multisite conference. In addition any unwanted noise will act as a mask making it difficult for participants in the room with the projector to hear the audio from other sites. It is essential that the selected projector is as quiet as possible, in addition housing the projector in an acoustic baffle may be necessary.

Manufacturers	• Epson • A+K • Sanyo • NEC • Philips
Price Guide	• £2000 upwards

7.16 Video equipment connections

A basic connection diagram for a typical video system is illustrated in Figure 2. The video feed from the remote site (via the CODEC) provides the principal conferencing image on the main picture monitor (remote monitor). A switching preview facility avoids the need to install a picture monitor for each local video source. The document camera provides the synchronising reference signal for the other cameras' Genlock units. Distribution amplifiers are used extensively to distribute the signals.

For larger budget installations a separate picture monitor can be provided for each vision source and a separate Synchronising Pulse Generator (SPG) can provide the reference signal for Genlocking.

A separate SPG provides a dedicated reference signal for Genlocking the vision sources and is a preferred and more reliable method. If the reference is a document camera and it is inadvertently switched off, then synchronism between the vision sources would be lost.

The room control system simplifies the operation of the installation by bringing together control of all the equipment, cameras, document camera, VCR, video switcher, CODEC, room lighting, etc. onto one control panel. When specifying equipment consideration should be given to the required number of video switcher inputs. The ability to connect temporary video equipment into an additional auxiliary video input may prove useful in specific conferences. Video sources such as DVD Players, video microscopes and PC scan converters may all be connected in this way.

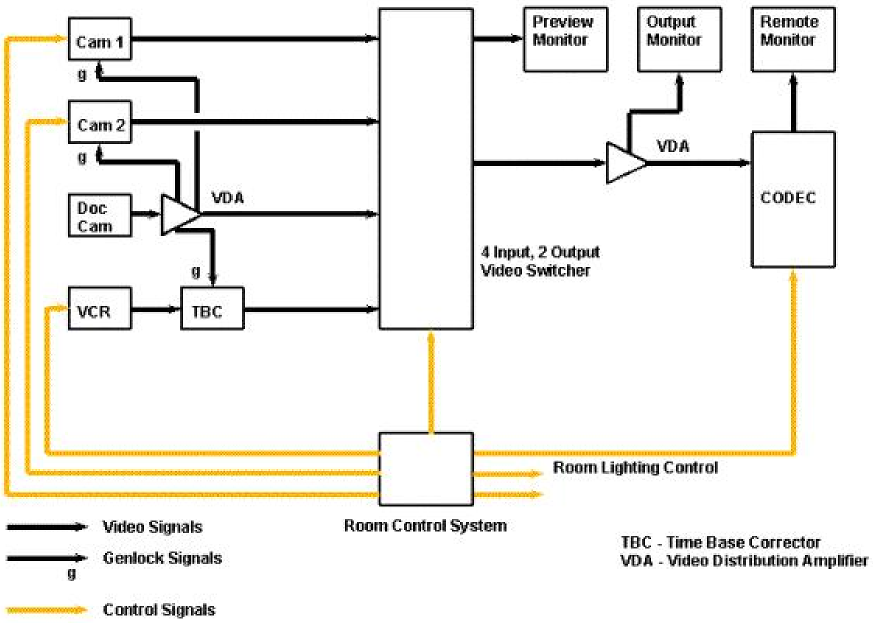

Figure 2. Video Equipment Connections

8. AUDIO EQUIPMENT

8.1 Introduction

The production of reasonable quality audio is very much more difficult than the production of acceptable quality video. Even broadcasting professionals are stretched to produce acceptable results in some situations.

To exacerbate this, most videoconferencing systems introduce appreciable signal delays that produce an unacceptable echo. Echo cancellation circuits are then introduced to reduce this echo to a tolerable level. Effective echo cancellation is dependent on reasonable room acoustics and, whilst for simple conference rooms most systems cope adequately, for more complex installations (e.g. lecture theatres) more comprehensive echo cancellation may be required.

Generally more attention is concentrated on the visual part of videoconferencing, but significantly more problems tend to occur with the aural.

8.2 Audio Installation

Audio equipment, particularly microphone circuits, are very prone to picking up interference in the form of background hum, buzz or noise. As the signal levels generated by microphones are very small (of the order of millivolts) any potential electrical or electromagnetic interference need not be very large to cause a very noticeable effect.

To minimise this interference it is recommended that only balanced, low impedance audio connections are used. This applies to the whole audio signal chain and includes microphones, mixers, amplifier, etc. All connecting cables should be of high quality, screened, twisted pair.

Most audio equipment intended for the domestic market will not be balanced or low impedance and will use unbalanced inputs and outputs and single screened cabling.

Unbalanced equipment using single screened connecting cables is not recommended for videoconferencing installations.

Audio cabling (particularly microphone cables) should be routed away from sources of interference and should not be run parallel to mains supply cables. All signal and mains supply cables must be installed according to IEE regulations.

Room lighting dimmers are particularly troublesome due to the high frequency nature of the interference signals radiated, so are best avoided in the conference room.

8.3 Echo Cancellation

Unless headphones are used to monitor the remote sound then acoustic echo cancellation will be needed to enable good quality, two way communications. The compression of the vision signal in the CODEC takes an appreciable time to execute (200-300 milliseconds). The sound signals must be delayed by the same amount to maintain lip synchronisation. This delay would introduce intolerable echo if ignored, so echo cancellers are used to reduce the echo to an acceptable level. Echo cancellers function by sampling a proportion of the remote site's sound within the local conference room, picked up by the local microphone, and generating a correction signal to minimise any remote sound which would normally return to the remote site as an echo. To enable this, one microphone must be fixed in the conference room in relation to the loudspeaker that is radiating the remote site's sound. This allows the echo canceller to monitor the remote sound in the acoustic setting of the local conference room and align itself to minimise the echo. This process is sometimes referred to as training. The internal echo cancellers within most CODECs are adequate for small to medium sized locations. However in lecture theatres where voice reinforcement PA systems are in use, or in difficult acoustic environments, e.g. rooms with a lively internal echo due to hard surfaces on walls and floors, an external echo canceller with a wider window of correction may be required. In addition to providing echo cancellation to each microphone input, stand alone echo cancellers can provide sophisticated audio processing including automatic gain control (AGC), microphone gating and equalisation. Facilities can also be provided to receive audio signals from sources which do not require echo cancellation i.e. PCs, VCRs, DVD Players, CD and audio tape players. External echo cancellers that have been set up with the assistance of an experienced audio engineer can produce excellent results in the most difficult of audio environments.

Manufacturers	• Gentner • ASPI Digital

8.4 Microphones

To achieve high quality sound, generally the microphones should be close to the participating speakers. Many microphones are now available with battery power. Whilst these can be attractively priced, for convenience and reliability microphones powered from the audio mixer or through a mains supply power unit will prove a better choice. Many different types of microphone are available, in two main categories: capacitor (condenser) and magnetic (dynamic).

Capacitor microphones incorporate a signal amplifier that requires power; this may be supplied from a battery or from an AC mains power supply. Magnetic microphones do not require any power but are susceptible to electromagnetic fields. Capacitor microphones have a higher sensitivity, picking up sounds from a greater distance, and are thus normally preferred in videoconferencing.

Capacitor microphones can be subdivided into:

- **tie clip microphones** - worn on the lapel to ensure close proximity to the speaker;
- **desk/stand microphones** - convenient as participants are not constrained by microphone cables;
- **gun microphones** - highly directional; used where it is difficult to get close to the speaker (useful in lecture theatres).

Microphones can have different response patterns:

- **omnidirectional** - i.e. equally sensitive in all directions;
- **cardioid (heart shaped)** - i.e. more sensitive to the front of the microphone than to the back and sides, so is useful for minimising extraneous noise;
- **super cardioid** - a highly directional microphone for special applications (e.g. gun microphone);
- **hemispherical response** - this is characteristic of the `boundary layer' type of microphone that depends on a hard surface, e.g. a table top, for its performance. These microphones can be very effective in videoconferencing.

Important Factors	• Top quality microphones are very expensive, but excellent results can be obtained from inexpensive ones.
Manufacturers	• AKG • Audio Technica • Beyer • Sennheisser • Shure • Sony
Suppliers	• Canford Audio • R-S Components

(inexpensive)	• Tandy
Price Guide	• £50-£200

8.5 Audio Mixers

There are many inexpensive mixers on sale which offer only unbalanced inputs and outputs: these should be avoided. For a conference room intended for meetings and small groups, four microphone inputs should be sufficient. In addition a number of high level inputs to connect a VCR, PC or Auxiliary Audio I/P from temporary equipment is required. While the input to the CODEC will be mono the output of many of the devices mentioned above will be stereo. The mixer will be required to derive a mono feed from each stereo source. A disadvantage of this single audio input to the CODEC is that echo cancellation will operate on the signal from the VCR and all non microphone inputs, this will reduce the quality of such audio sources. Many CODECS have a separate audio input for VCRs and non microphone inputs on which echo cancellation does not operate. Consideration should be given to the use of such inputs. Automatic mixer/microphone systems known in the UK generically as 'Smartmixers' can be very effective in audio conference situations, as only the microphone closest to the person speaking will be activated. This enables extraneous noise to be minimised and the voice signals to be optimised. In videoconferencing, however, care must be taken in deploying such automatic systems with an echo canceller in the chain. Effective echo cancellation relies on applying a dynamic correction signal to the audio from the local room microphones (see Section 8.3 Echo Cancellation). This correction signal removes any remote audio, which may be picked up on the local room microphones. The echo canceller takes a finite time to create this correction signal following any change to the room audio environment and during this time, known as training, the echo may be intrusive. If a smartmixer is deployed and the active microphone continually changes position the echo canceller will retrain on every microphone switch and may produce poor audio results. An alternative approach is to ensure that one of the room microphones is continually live, perhaps at the chairmans position with additional microphones controlled by the smartmixer. In such a scenario two microphones will always be live, one fixed and one switching to the conference participant currently speaking, the fixed microphone enables the echo canceller to train and operate effectively while the smartmixer will reduce the background noise as the number of active microphones is reduced.

Important Factors	• Balanced inputs and outputs. • A meter to measure sound level (Peak Programme Meter (PPM) preferred). A Volume Unit (VU) meter is more frequently fitted but these are less easy to read. • Provision for microphones, plus at least one high level input (to enable video playback). • Phantom powering for microphone (i.e. to provide power for capacitor microphones through the signal connections - this avoids batteries or separate power supplies).
Manufacturers	• Audio Technica • Canford • Shure
Price Guide	• £300-£1,000

8.6 Audio Equipment Connections

A typical layout for the audio equipment is shown in Figure 3. The four microphones, video player and auxiliary audio source are fed to the audio mixer. The audio mixer has two outputs, a mono mix of all the mixer inputs, which will be routed to the CODEC, and a stereo mix of all the stereo sources excluding the microphones. This feed allows the local audience to hear the output of the VCR, PC, etc. The distribution amplifier routes the sound to the echo canceller and then to the CODEC for transmission to the remote site. Sound from the remote site (via the CODEC) passes through the echo canceller before being amplified and reproduced through the mono loudspeaker system. The transmitted audio signal level is monitored using the PPM.

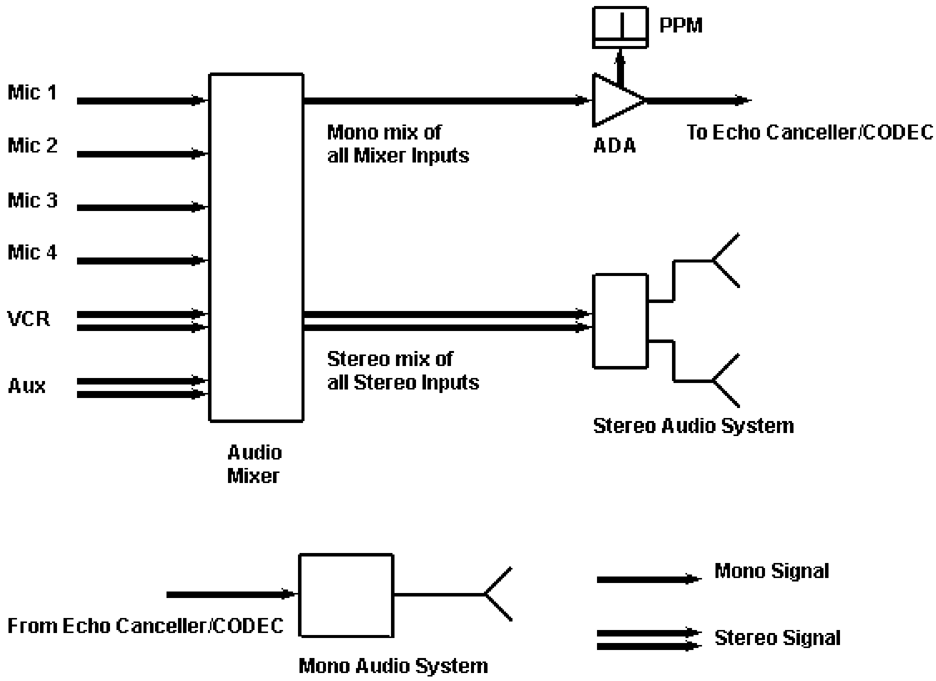

Figure 3. Audio Equipment Connections

9. DISTRIBUTION EQUIPMENT

9.1 Distribution Amplifiers (DAs)

These are amplifiers that split a single input to produce several isolated outputs identical to the input. This enables a main transmission signal to be routed to many destinations without signal degradation. Another big advantage is that the signals are isolated, so that if one output has a fault (e.g. a short circuit) none of the other outputs is affected. Distribution amplifiers are used for both audio and video signals. If the main video output of the conference room is fed into a distribution amplifier (see Figure 1) then one output can feed the CODEC, another a picture monitor, another a video recorder and another a waveform monitor. Video and audio DAs are frequently packaged together.

Specifications:

Video DA

- frequency response 0 - 5.5 MHz 0.5 dB;
- differential gain and phase < 0.5% and 0.5 respectively;
- DC clamping at black level;
- signal to noise ratio >70 dB.

Audio DA

- balanced inputs and outputs;
- frequency response 20 - 20 kHz 1 dB;
- third harmonic distortion <0.1 %;
- signal to noise ratio >80 dB;
- crosstalk better than 55 dB (when several DAs are packaged together).

Important Factors	• Where video cables exceed 15 metres the higher frequency parts of the signal, especially the colour information, can be attenuated. Some DAs incorporate correction circuitry (equalisation) to compensate for this loss. • Where long video cables (100+ metres) are used then much more compensation is required over the frequency band. Some amplifiers will provide this. • This wide band equalisation does introduce noise to the signal so has to be used with caution, especially where CODECs are in the signal chain, because of the adverse affects of noise on compression systems.
Manufacturers	• Kramer • Toa • DVC • Alice
Price Guide	• £100-£500 each, combined audio/video DAs more expensive

9.2 Fibre Drivers

Fibre drivers have been used successfully for several years with digital signals. As analogue signal processing is still used extensively in audio and video studio equipment, (e.g. television cameras, audio microphones and mixers), it can be appreciably less expensive to distribute these signals to local, and remote locations in the analogue domain, i.e. by using analogue fibre drivers. Analogue fibre drivers are available that multiplex both audio and video information over a single fibre. However, serious problems have been experienced with some products. When analogue audio and video signals are transmitted on separate fibres there does not seem to be a problem. It is when audio and video are multiplexed that trouble seems to arise: quite alarming crosstalk can occur (i.e. video modulating the audio and vice versa). Some manufacturers, however, have overcome this problem and market first class products.

Manufacturers	• Meridian • Probot • SVT
Price Guide	• For a simplex link, i.e. one transmitter and one receiver, £1,000-£2,500

10. TEST AND MONITORING EQUIPMENT

The equipment needed for comprehensive video and audio testing is both complex and expensive, and requires the skills of an experienced engineer. As most videoconference centres are working to tight budgets this level of expenditure is inappropriate. Basic audio and video monitoring using test signals and measuring equipment, however, will verify that equipment and systems are operating within acceptable limits. For audio a Peak Programme Meter (PPM), as shown in Figure 2, ensures that the audio signal leaving the studio is of an acceptable level. For more comprehensive testing of the video equipment and systems, a colour bar generator and a means of measuring video level and video colour phase would be recommended; for the audio, a reference tone source together with the PPM above. One manufacturer, Hamlet, markets monitoring products particularly suited to videoconferencing. The equipment superimposes both the video waveform (for video amplitude) and colour vectors (to measure colour phase) over the image on the outgoing picture monitor, eliminating the need for expensive waveform monitors and vectorscopes. A simulated PPM is also superimposed on the picture to monitor audio level. Some high quality, wide range echo cancellers will need setting-up before they are installed into a conference room. This requires a Sound Pressure Level (SPL) to set levels.

Important Factors	• An inexpensive SPL meter quite adequate for setting up echo cancellers is available from Tandy. • There are numerous inexpensive audio tone sources on the market but we found that many did not perform to specification, the main problems being inaccuracy at the nominal output level of 0 dBm (at 1 kHz) together with variation in output level as the frequency varied.
Manufacturers	• Canford Audio - PPMs • Hamlet - picture monitor overlay of video waveform and PPM • Videotech - tone source • Tandy - SPL meter
Price Guide	• PPM £250-£500 • Hamlet measuring equipment £1,200-£2,500 • Audio tone sources £300-£1,000 • SPL meter £35-£2,000

11. GLOSSARY

Artefacts
A term used to describe the distortions added to the original signal during the coding and decoding processes.

Asynchronous
A video signal that is not synchronised to the local reference (or camera) signal.

CODEC - **CO**der **DEC**oder
The equipment that provides the compression and signal processing to convert high bandwidth analogue sound and vision signals to a form that allows them to be transmitted and received over low bandwidth digital transmission paths.

CCD - **C**harge **C**oupled **D**evice
Used in television cameras as a photo-sensitive device to convert light into an electrical signal.

Composite Video
A method of transmitting video information where the luminance, chrominance and synchronisation components of a television signal are combined into a single signal.

DA - **D**istribution **A**mplifier
Amplifiers that split a single input to produce several isolated outputs identical to the input, enabling signals to be routed to many destinations without signal degradation.

Echo Cancellation
The CODEC delays the vision signal by approximately 200 milliseconds. To maintain sound/vision coincidence the audio signals are delayed by a similar amount. This time delay produces unacceptable echo into the conference. Echo cancellation is introduced electronically to reduce this echo to a workable level. The conference environment influences the amount of echo, so echo cancellers are set up within the conference room in use.

Frame/Picture Store
A means of storing electronically one complete television picture, i.e. one frame of information.

H.320
The umbrella ITU-T standard for narrow band videoconferencing interoperability over ISDN networks.

H.323
The umbrella ITU-T standard for narrow band videoconferencing interoperability over Local Area and Wide Area Networks (LANS, WANS).

H.261
An ITU-T standard for video coding (or compressing the video signal).

H.263
An ITU-T standard for video coding, specifically designed for operating at low data rates, i.e. 64 to 128 kbit/s.

Hi-8
A modified form of the Sony 8mm video recording format that produces higher quality by splitting the video signal into black and white (Y) and colour (C) information, i.e. it records/replays S-video.

IEE - **I**nstitution of **E**lectrical **E**ngineers

ISDN - **I**ntegrated **S**ervices **D**igital **N**etwork
A non-dedicated dial up, digital service offered by world-wide telecommunication providers.
It enables digital transmission over the existing telephone infrastructure.

ITU-T
The telecommunications section of the International Telecommunications Union, dealing with videoconferencing and other standards.
A list of ITU-T standards is available at: http://www.itu.ch/publications/itu-t/itut-es.html

Jitter
When an electronic device (e.g. a television camera) generates a signal, the synchronisation signals are not absolutely stable, i.e. there is a small amount of timing jitter around a mean.
An electro-mechanical device (e.g. a video player) generates a significantly higher amount of jitter due to the mechanical tape transport mechanism.

NTSC - **N**ational **T**elevision **S**ystems **C**ommittee
The United States' system for coding colour information onto the composite video signal.

PAL - **P**hase **A**lternate **L**ine
The system used in the United Kingdom for coding colour information onto a composite video signal.

PPM - **P**eak **P**rogramme **M**eter
Used to measure audio signal amplitude. The meter characteristics are weighted

to produce a fast rise but a very slow fall of the needle movement while following sound signals. This characteristic makes the measurement of rapidly varying sounds (e.g. music) more accurate.

SECAM - **S**equential **C**olour with **M**emory
The French system for coding colour information onto the composite video signal.

SPL Meter - **S**ound **P**ressure **L**evel Meter
Used to measure ambient sound levels: required for setting-up high quality echo cancellers.

SPG - **S**ynchronising **P**ulse **G**enerator

S-VHS
A modified form of the JVC VHS video record format that produces higher quality results by recording/playing back S-video (i.e. Y,C) signals.

S-video
The luminance and chrominance information of a colour television signal are transmitted as separate components to improve quality (see Composite video).

T.120
An ITU-T standard for data transmission that enables data, application sharing, etc. to occur within the pass band of the normal videoconferencing channel.

Training
To operate effectively, echo cancellers need to monitor the echo received from a remote site within the local conference room and generate a correction signal to cancel the echo. This alignment procedure is termed `training'.
This process considers not only echo from the remote site but also the acoustic characteristics of the local conference room.
If the local room has poor acoustics, e.g. a high ambient noise level or very long term echo, then the echo canceller may not be able to achieve a satisfactory alignment. The local site will then introduce annoying echo into any conference in which it participates.

Unstable
A video signal is termed unstable when it emanates from an electro-mechanical replay device such as a video player and has not received electronic time base correction to render it stable.

Vectorscope
An instrument that displays colour vectors, and thus enables colour phase to be measured.

Vertical Interval Switching (VITS)
A method of switching within video switchers/mixers that arranges for the actual transition to occur in the time between consecutive television pictures so as to minimise disturbance to the images.

VHS
A video cassette recording format developed by JVC.

VU - **V**olume **U**nit Meter
Used to measure audio signal amplitude. For measuring constant signal levels (e.g. tone) it is adequate.
For the measurement of audio signal levels that are rapidly changing (e.g. music) it is unsatisfactory (See PPM above).

Waveform Monitor
An instrument that displays a video waveform to enable amplitude measurements to be measured.

YC Luminance Chrominance
A synonym for S-video.

Updated: 30/10/2001