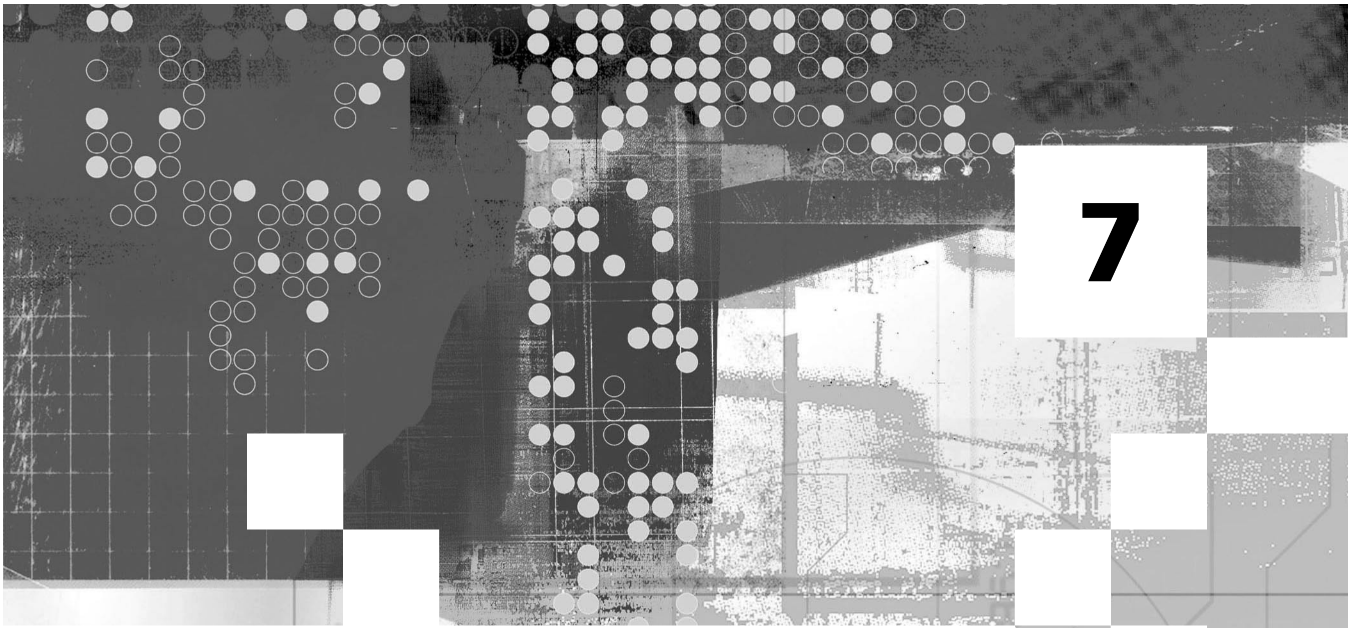


7 Video asset production

Project manager's responsibilities

- To book crews, help cast actors, presenters and subjects for recording interviews if necessary
- To explain to clients what quality of result to expect on the Web and desktop
- To make sure that the recording session is successful
- To make sure tapes are logged and prepared for editing
- To ensure that the material is prepared to the correct specification, recommending or selecting the appropriate formats
- To choose the most appropriate way of compressing video for the application and make sure it is done to an appropriate standard
- To understand the processes involved in producing this kind of asset

Managing asset production

The previous chapter has explored the production of sound in multimedia, paying special attention to the use of external facilities. For video, similar arrangements are likely to be made. It is possible that the video content of a multimedia application will be shot using a consumer camcorder and edited on the desktop, but in other cases the budget will allow a professional facility to be used and the production values will require it. As with audio, websites currently make less use of video than CD-ROMs did, but as bandwidth increases this will change.

One crucial aspect of video production is the amount of preparation required. For sound it is quite feasible for the producer to turn up to interview someone after having made the arrangements by telephone a few hours before. But for video there is much more that needs to be done and probably more people to involve. This chapter concentrates on the set-up for an interview because that is the most straightforward and arguably the most common use of video in multimedia. You might wish to shoot drama, or work on location, but unless you already have experience in these areas you are advised to hire specialized help to direct the shoot and therefore to manage it. Exterior location work is particularly difficult because of the range of permissions you need from people such as the police and local authority and because of the vagaries of the weather.

One managerial issue that rears its head in video more than in audio is the question of whether the client should attend the shoot or the edit. There is no simple answer to this but your relationship with the client should influence you. The main reason why the client should attend is to avoid arguments about changes later, since changes in a video edited at an outside facility are probably the most expensive to fix. This, of course, is a strong argument for making an offline edit to inexpensively preview the result and letting the client see it, or for editing on your desktop. If you are completely confident that the client respects your abilities then you might not need to involve them, but if you think that later changes are likely, and if you have agreed a fixed price for the project, then invite them. You should emphasize the costs associated with changing video after the edit. Of course, with a computer-based non-linear edit it is much easier to change things than it was in the old-fashioned assembly method of videotape editing, but it still takes time and costs money.

Moving pictures are a recent entrant into computing, although computers have long been a part of video, at least on the professional and broadcast television side. Engineers have for a long time been devising ways to manipulate and create images for broadcast television, starting with graphics and control of equipment. Today computers play a part in every aspect of broadcast television production.

Just as video is starting to take over as the dominant medium in the multimedia mix as inexpensive computers become more powerful, so multimedia is beginning tentatively to make its way into television. Video-on-demand is essentially a multimedia database system for the home; prime-time television programmes are using desktop computers to generate graphics and many television programmes are edited using desktop computers; and digital television is now being broadcast by cable, satellite and terrestrial transmitters to home consumers.

Television itself dates back to the 1920s, and moving film goes back into the nineteenth century. The Scottish television pioneer, John Logie Baird, devised a mechanical system of sending moving pictures down a wire and through the airwaves, at a time when radio broadcasting itself was in its infancy. Electronic scanning took over from his wheels and cogs, and the BBC launched the world's first regular scheduled television service in 1936 to a handful of very rich viewers in the southern part of England. A television set then cost more than a family car, and not many people had either.

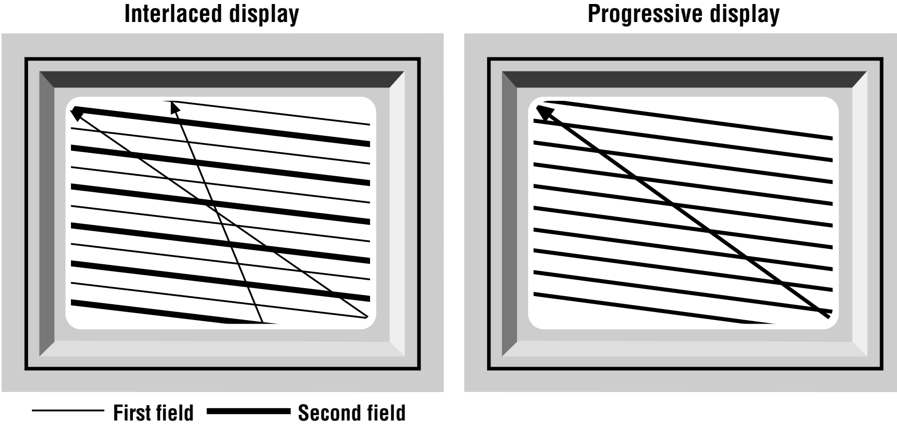
This chapter deals with the way in which a television picture is made and how you might record and manipulate it. Some background on the way a video signal is built up, and the differences between the various formats and standards, will help you work with video. As with the last chapter on audio, it will concentrate on the kind of equipment and techniques that you, as a new media professional and practitioner, will come across when you make use of video.

Basic principles of video

Because our brains are easily fooled, we see a rapid succession of still images as a continuous stream, and, with the right images, we see movement. The movies work like this, by showing us tiny slices of reality 24 times per second. The frames of a movie are those slices of reality. Television slices the reality a slightly different way because each of the frames in a TV signal is made up of hundreds of lines.

As a compromise between resolution (in pixels), the rate of change of the picture (in frames per second), and the amount of radio waves that a television signal would occupy (bandwidth), the engineers who designed electronic television made each frame out of two fields. The first field covers only half the picture by missing out alternate lines of the image as it scans down the screen. The second field fills in the gaps.

The beam that 'paints' the picture on the television screen starts at the top and scans from left to right at a very slight angle downwards. For the first line only, the beam starts in the middle rather than at the left. When it reaches the right it flies back to the left-hand side and starts again, a bit lower down. Eventually, after doing this a few hundred times, the beam reaches the bottom right of the screen and promptly flies back up to the top left and starts all over again. This time it will be scanning in between the lines it laid down last time; eventually it will reach the middle of the bottom line and fly back up to the middle of the top line, and the whole process starts again.



So there are two comb-like scans, interleaved with each other, making up each frame, with each consisting of half the number of lines in the whole picture. Each of these scans is called a **field**. The complete picture will be refreshed only 30 or 25 times per second, which keeps the bandwidth down, while the apparent flicker of the picture is at the rate of the fields, which is twice as high and so is less noticeable. This is a neat trick; but, as will be shown later, this system of interlaced fields causes no end of problems when we get onto the computer because computers use a non-interlaced display, in which the lines are written in turn, left to right, from top to bottom of the screen. This is called progressive or sequential scan and one of the fiercest debates going on between broadcasters and the computing fraternity is about whether future television systems should have interlaced or progressive displays.

Composite video television standards

These are the three main systems of analogue broadcast colour television. They are known as composite video (or just composite) because the brightness and colour information are mixed together into a single signal. The oldest colour system still in use is NTSC, which is used mostly in North America and Japan. The main European system is PAL, also used in places such as Australia and South Africa, and there is also SECAM, used in France (from whence it came), Eastern Europe and the Middle East.

For NTSC a constant-frequency signal at 3.58 MHz, called the colour subcarrier, is superimposed on the picture to carry the colour information. At the start of each line of the picture there is a short burst of the colour subcarrier. The colour at any point in the line (this is analogue so the concept of a pixel is not strictly valid) is determined by the phase relationship between the superimposed subcarrier at that point and the reference burst. The phase relationship is the difference between where one signal is in its positive/negative cycle and where the other one is.

There are 525 lines per frame in NTSC. Interlaced television pictures always have an odd number of lines so that one field can start in the middle of the screen (horizontally) and the other can start at the left-hand end so each field includes a half-line. In broadcast NTSC 487 lines have picture on them – half in each field – and these are called the active lines. You don't have to have an odd number of active lines and in multimedia practice, NTSC is considered to have 480 active lines rather than 487. In its earlier black-and-white incarnation, US television used 30 frames per second, but for colour this changed to 29.97. It was supposed to be 30, but obscure technical reasons to do with the subcarrier frequency changed all that.

PAL has a subtle, but significant, difference from NTSC in the way that the colour information in the subcarrier is encoded. NTSC suffers from colour shifts due to minute timing changes during a line which affect that phase relationship between the colour subcarrier in the picture and the reference burst: this causes changes in the colour. PAL compensates for this by reversing the phase of the colour encoding between alternate lines and so 'averaging' out the errors. PAL also has a higher-frequency subcarrier, at 4.43 MHz, and has 625 lines in a frame, of which 576 have picture on them: half in each field. PAL shows 25 frames per second exactly.

SECAM, the French-devised third system, uses a totally different method for transmitting the colour (it sends the two colour components alternately with each line) but is otherwise the same as PAL, so if you fed a SECAM signal into a PAL monitor you would see a black-and-white picture and vice versa. SECAM means 'Système en couleur avec mémoire' which literally means 'Colour System with Memory' although it is jokingly reputed to really mean 'System essentially contrary to the American method'. Although SECAM is still used as a transmission system, and gives very good results, it has problems for production, mainly because you cannot mix or fade a SECAM signal out because of the way the colour is encoded in the signal. Incidentally, NTSC is reputed to stand for 'Never twice the same colour', but actually it is National Television Standards Committee. (The digital equivalent in the USA is ATSC, Advanced Television Standards Committee but I have yet to hear the 'alternative' definition for this.)

There are some variations, particularly a South American variation of PAL called PAL-M, which is basically NTSC timing with PAL colour. You will sometimes see PAL-I and PAL-BG referred to, but these are different only in the way they are transmitted; the videotapes are compatible even though the televisions and VCRs are not, mainly because the sound is transmitted on a slightly different frequency, and that is why a European PAL television will not usually work in the UK and vice versa.

On European videocassette machines you may see a form of NTSC called NTSC 4.43 or Modified NTSC. This is a special version of NTSC, and it happens because the colour information on the videocassette is stored in a different way from the original composite signal: when it plays an NTSC tape the machine puts out what is basically a PAL signal but with NTSC timings. The subcarrier is at the PAL frequency. This saves having two sets of circuits for the colour in the VCR and the monitor. However, this modified NTSC will appear as black and white if you feed it into a real NTSC monitor, or a digitizer card that expects real NTSC. This can also happen with formats such as S-VHS and Hi8, which, although they separate the brightness and colour information, still use a subcarrier to encode the colour information. You will not come across this problem with broadcast videotape machines, and it is not applicable to the component systems we discuss in the next section, but it is the most likely cause of mystery colour disappearances when using NTSC videocassettes on multi-standard VCRs.

A single video cable can be used to carry a composite video signal since the colour information is encoded in with the luminance and the synchronization pulses are usually kept together with the picture information. Some equipment labels composite video, CVBS.

Component video

Professional colour videotape is now either digital or component, or both. Component video (often just called component) means that the colour (chrominance or chroma) and brightness (luminance) information are kept separate, having been mathematically derived from the red, green and blue (RGB) signals received from the camera imaging tube or chip.

The luminance part of a component signal is referred to as Y. The true colour components are red-Y and blue-Y. Red and blue signals are smaller in value than green so they make for larger (and therefore less noisy) components. Often the components are referred to as YIQ (in NTSC) and YUV (in PAL), but these are not strictly equivalent and the terms are often misused. Strictly speaking there is no such thing as NTSC or PAL component, but people do sometimes refer to component systems by their related composite system names in this way as they do with digital component systems. When it comes to components, PAL and SECAM tapes are the same.

Using colour components this way is useful for TV for two reasons. First, the luminance signal is exactly the same as the old black-and-white signal, and so the video is backwards compatible and easy to view on a monochrome monitor. Second, because we do not see as much detail in colour as we do in brightness, there is no need to have as high a definition for the chroma as for the luminance. This saves bandwidth, and is one of the two forms of information compression used in analogue TV: interlace is the other.

Component video offers better quality than composite because the colour information is kept separate from the luminance. It is possible for a colour television decoder to mistake some of the fine detail in the luminance information in a composite signal for colour, causing spurious coloured patterning. A presenter wearing a narrow-striped jacket would cause problems on a composite system. This also means that luminance information of similar frequency to the colour subcarrier will be missing from the composite picture as it will be filtered out.

Connectors for component analogue video are the same as for professional composite, only there are three of them: one for luminance (Y) and two for colour. The sync pulses are usually carried on the Y line.

Half-way between composite and component video is S-Video as used in Hi8 and SVHS. This has two signals, one being the Y and the other a combination of U and V usually referred to as C. So a connector for S-Video might have Y/C written on it. An S-Video connector is made either from two BNC or phono connectors or from a four-pin mini-D. Sync pulses are carried on the Y line. There are only two forms of S-Video since, once again, PAL and SECAM are the same at this level.

Blanking and time code

Besides the image information, making up the visible part of the lines, there are elements of the analogue television signal that tell a television monitor, and other equipment, where the lines, fields and frames begin and end, and provide reference information about the colour. The places in the television signal where the picture does not exist, but these signals do, are called horizontal (for lines) and vertical (for fields) blanking. This is because they are blank, although the vertical blanking interval (VBI) has become home to such things as Teletext, vertical interval time code (VITC, pronounced vit-see) and test signals.

Time code is very likely to be useful to you. When you look at a videotape to select extracts for digitizing, time code is the way to specify the sections you want. Every frame of a television signal can be allocated a unique number divided into hours, minutes, seconds and frames (of the form hh:mm:ss:ff, for example 05:46:35:19). This time code signal is recorded on the tape along with the video and audio.

Time code is sometimes referred to as SMPTE (Society of Motion Picture and Television Engineers, pronounced 'sempty') although, strictly speaking, SMPTE time code is for NTSC video only, and the PAL/SECAM version is EBU (European Broadcasting Union).

Burned in time code (BITC or bit-see) or time code in vision is a system whereby a character generator superimposes the numbers of the time code on the frames to which it refers. This is invaluable in choosing extracts and editing. Since a time code number can refer to either field in a frame, some readers will add a field indication, such as an asterisk, to the number you see. You can often look at the two fields of the picture separately because most videotape machines show you a field rather than the whole frame when the tape is in still 'frame'.

With NTSC time code you will see the term 'drop frame' used. Because NTSC does not have a whole number of frames every second the time code has to be adjusted every so often to keep it in step with real time. This is just like the extra day in February in leap years. In the case of NTSC a frame is dropped, hence the term. This is fine when it is important for the time code to show time-of-day, as it is called, or to know exactly how long a programme is, but you need to remember that some time code numbers will be skipped. For everyday video editing it is more straightforward to use non-drop-frame time code. You should know which type of time code your NTSC tapes have otherwise you will miscalculate the actual duration of your video. In PAL there is no need for drop-frame time code as PAL has exactly 25 frames per second.

In the studio and on location

You might use only a small amount of video in an interactive application but its use can be very powerful and effective. One common form is known as a talking head. This is where all you see on the screen is a single person talking. If you have a famous presenter for your application then you might see him or her a few times like this. Even if most of the time you only hear the voice, it is nice to show your viewers what the face behind the voice looks like.

The talking head might be recorded in a studio or on location (meaning not in a studio), and he or she might be positioned in front of a real scene (a bookshelf is a common one for a subject expert) or against a single colour. A single flat colour is useful because, in some circumstances, you can decide on the background later and add it in the edit suite or on your computer.

Since a human face does not contain any blue, blue is commonly used for this purpose. Television people call this chroma-key or colour separation overlay (CSO – hence CSO blue to denote the colour). Film people call it a matte (if it moves it is called a travelling matte – and yes, it was used for flying carpets in Arabian Nights films – and now you know the origin of the name of that character in *Fraggle Rock*). Another name for this technique is blue screen, although other colours can be used: broadcasters often use green, and the traditional colour for movies was the yellow light of a sodium discharge lamp.

One neat method of generating the flat colour is to put a reflective cloth behind the foreground object (such as a person being interviewed) – one of those cloths that reflects light straight back at the source – and illuminate this with a low powered ring of coloured lights fixed around the lens. This creates the coloured background easily but also makes it straightforward to light the foreground subject independently.

Although in film-making a travelling matte works using totally film-based methods and some clever work at the time of printing, in television an electronic circuit detects the blue and switches the signal to another source wherever there is blue. This substitutes the other image for the blue.

Colour separation overlay.

Common problems occur due to spillage of the blue light onto the person's face, shadows on the blue backcloth, and difficulties in coping with the fine detail in hair and with shadows. The most sophisticated colour separation systems, such as Ultimatte, can solve these difficulties but, in any case, good lighting helps to avoid them.

Sometimes it pays to shoot against black. This can be very effective to isolate the person speaking and emphasize their role as a specialist, and it can be less distracting than the ubiquitous rows of books. Black does not spill onto faces and, if you are careful, you can overlay onto black just like blue or green. In this case you need to watch out for shadows, especially under the eyebrows.

Although these are film and television techniques, the use of colour separation overlay extends easily onto the desktop, and video editing and manipulation programs usually allow you to clean up and replace coloured backgrounds in this way. When the video is digitized and compressed, a static background will compress very efficiently. This is especially true if it is free of noise, so it is common to shoot the interview against black and then replace the black with 'digital' or mixer black.

The simplest ‘professional’ way of lighting a single person uses three lights: one will be a spotlight on the face, one will flood the scene to lighten the shadows, and the third will be positioned behind to both light the blue screen (if you have one) and backlight the person to help lift them from the surroundings and give the scene a three-dimensional look. The spotlight in front needs to be high enough so as not to make your speaker want to shade his or her eyes but not so high as to cast shadows in the eye sockets. As usual, good lighting will look so natural on the screen that you will wonder just what you have paid for. The lights have names, and a redhead refers to a one-kilowatt lamp while a blonde is a two-kilowatt one. The lights will need to be set up, but it is bad manners to inflict this on the speaker, who will be under the lights for long enough later. One of the crew will stand (or sit) in for the speaker while the lights are moved and adjusted. Find someone the same height as the speaker. Also, you should make sure that the speaker does not wear clothing containing the colour of the background, if you are using overlay.

For a professional and hassle-free video shoot you should avail yourself of an experienced camera operator and crew. The smallest crew will be one that consists only of a cameraman and you, the director (so get yourself a chair and a megaphone). You should seriously consider having a sound recordist, who will also help the cameraman to set things up, and a PA (production assistant), who will take notes and carry out the little administrative chores and free the director to … direct. You can reasonably expect the crew to bring their own equipment and videotape (referred to as ‘stock’) but don’t forget to check.

If your application needs more substantial shooting, such as some drama or location documentary filming, then you will probably need to bring in a specialist director. If you have the experience to direct this kind of material, you probably do not need to read this chapter anyway. Crews for drama shoots can be very large, and include workers with colourful names such as the gaffer, best boy and grips.

The kind of camera used for professional video is similar to a consumer video camcorder, but it is physically bigger with a better lens and will give a much better picture. The camera operator will probably connect up a monitor for you to see the picture. For an interview, don’t put the monitor where the speaker can see the picture or they’ll keep looking at it and not at the interviewer or camera or wherever you really want them to look.

Your speaker will need a microphone, and you will probably not want to see it. For this reason a common microphone used in filming is called a rifle mic, or hypercardioid. This is a long and thin microphone, although you will usually see it in a wind shield or wind sock. This makes it look like a big furry hotdog held on the end of a pole by the sound recordist. Alternatively a very small microphone can be clipped to the person’s tie or jacket collar. These are usually omnidirectional, and it is not unusual to position them upside down to avoid blasts of breath.

Shooting an interview

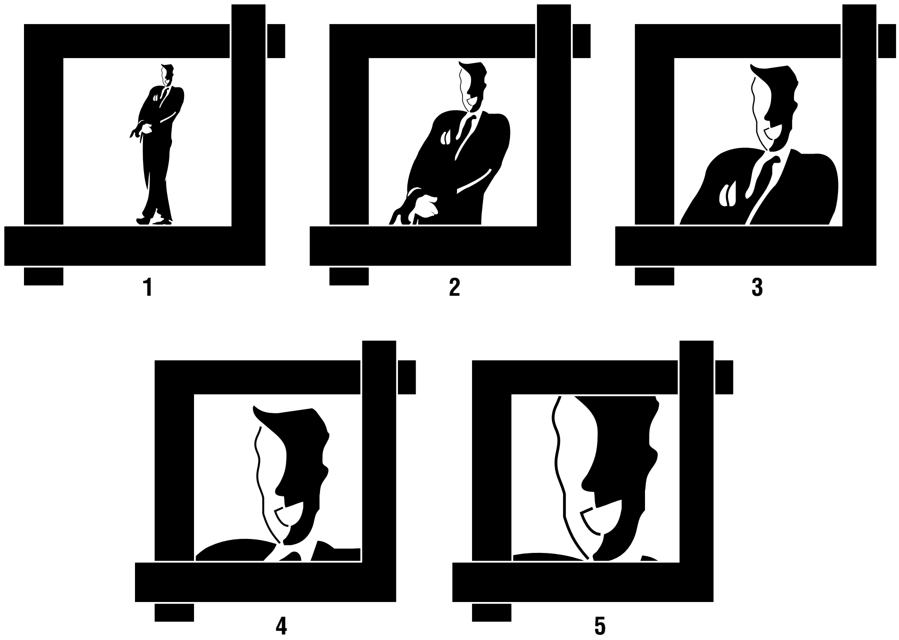
Although an interview (and similarly a recording of a single speaker at, say, a conference) seems to be one of the simplest forms of television shooting it is still difficult to direct well. Part of the problem is that you will want to edit what you record for inclusion in the finished application. With a sound interview you can cut almost anywhere as long as it makes sense, but with video you have the added difficulty of having the picture to cut along with the sound. It is highly unlikely that two similar pictures will cut together. In the case of a person speaking you will see the join as a jump, which is quite disconcerting. One tactic is to make a feature out of this 'bug' and just show the jumps, or do a quick mix across them. You might do this if you want to clearly show that the interview has been edited. To hide the jumps you can do one of two things but you need to have recorded the material to do so. First, you can cut to a different view of the person who is speaking, known as changing the angle, or zoom or move closer (in) or further away (out). To disguise a cut, a zoom will work better than a simple change of angle, but ideally you should combine both. Sometimes two cameras can be used to shoot the interview, set up with different shots so you can cut between them. It's not unusual for a broadcast camera to be used for one and a consumer camcorder for the other; just for the quirky effect. It all depends on how obvious you think the cutting should be.

Changing the size of the shot has a curious side-effect. Zooming in on a person speaking, either as a real zoom or as a cut to a closer shot, makes what he or she says seem very dramatic, as if confiding in the viewer. Zooming out has the opposite effect.

The second method of covering an edit is known as the cutaway. You can cut away to anything relevant. This might be some footage of whatever the person is describing. Alternatively the cutaway can be to something as simple as a shot of the person's hands as they move to emphasize a point. Of course, while you are cutting away from the speaker's face you can edit the sound to your heart's content.

It is possible for the speaker to work from a script but not appear to be reading. This kind of technique with half-silvered mirrors has been of benefit to politicians everywhere since Ronald Reagan showed it to Margaret Thatcher, but in television it has been in use for decades. It is used extensively in news and current affairs programmes where the presenter has little opportunity to learn a script. Almost as useful is to photocopy the script, blow it up to a large size, and tape it just below the camera lens. This obviously works best with short scripts.

Prior to the interview you should run through the topics you want to cover, and you should have given the speaker time to research the answers, even if it is his or her specialist field. Rehearsing can reduce spontaneity, but retakes are always a possibility. You can change the shot for retakes to make editing between them possible.



The shorthand for describing the amount of the speaker you can see in a shot is roughly like this:

1. **Long shot** – you can see the whole person and maybe enough of the surroundings to see where you are.
2. **Mid-shot** – you can see most of the person speaking, including his or her hands.
3. **Medium close-up (MCU)** – you can see the top of the person, from the middle of the chest upwards. This is sometimes called a 'head and shoulders' shot, and is the basic shot for an interview or statement to camera.
4. **Close-up** – you can see little more than the head.
5. **Big close-up** – very close (tight) on the head, cutting off the forehead.

Incidentally, women who wear strapless dresses are difficult in interviews because the MCUs and close-ups make them look naked. This may not be the effect you want.

A very useful technique for interviews is to turn them into statements. This is used a lot in television documentaries, and basically means that the interviewer is neither seen nor heard in the edited result. This means, of course, that one place to which you cannot cut away is a picture of the interviewer nodding sagely (these shots are known as noddies). In this case it is important that the speaker never actually answers a question directly, and you should ask the interviewer to say things like 'Tell me about your adventures up the Limpopo river' rather than saying 'Was it fun up the

Limpopo?' A common tactic is to ask the interviewee to repeat the question as the first thing he or she says in reply.

It is possible that you might be conducting the interview yourself, especially if you are not including the interviewer in the final edited footage. Part of your role is therefore as much to put the speaker at his or her ease as it is to ask anything meaningful. You will nod encouragingly, but never say anything while the interviewee is speaking. Some subject experts are so used to appearing in front of a camera that they will respond accurately to a request like 'tell me about Jupiter's red spot in 30 seconds' with consummate ease. Subject experts like that can be worth their weight in gold.

If you have a PA with you, he or she can make notes of questions to set against the time code, which can usually be read on the side of the camcorder. You should resist the temptation to use time-of-day time code or drop-frame time code in NTSC, and you should start each new roll at a new hour on the time code. This way you can easily judge which roll is which (they are still called rolls even though they will be videocassettes) from the time code numbers. So the time code 04:23:01:10 is on roll 4 and 13:21:10:00 is on roll 13. It is vitally important that the time code always goes forwards as you work through a tape and that the code never passes through midnight. This is simply to make it possible for the editing system to know which way to spool a tape to find a number, and unfortunately 23:59:59:24 reads as being almost 24 hours after 00:00:00:00 instead of one second before it. When you digitize the footage for editing, the software might report a time code discontinuity as dropped or missing frames and may not tell you where the problem is. This means you will have to review the file to make sure it is OK. Better not to have any time-code discontinuities if at all possible.

Do it yourself

So far we've looked at video production as being something that you hire people to do, but you don't have to. DV camcorders can produce quality that is eminently suited to multimedia and this begs the question of how much can you just go out and shoot yourself?

The answer is, as much as you are able to do. The key to shooting good video lies as much in the camera operator as in the equipment – more so perhaps. Lighting is important as well, partly because normal room lighting will not necessarily work well with a camcorder but also because inexpensive cameras generally take better pictures if the scene is well lit.

Bright sunlight is good for general shots, but you might find it difficult to shoot a face, for example. If the sun is shining directly on the person then they will probably want to shade their eyes to compensate and frontal lighting does not usually look as nice as an offset lighting, which 'models' the shape a bit by producing shadows. If the light is behind the subject then the camera will probably show it as a silhouette. If the shadows are too dark then that will look wrong and you may have to use a white board to reflect some light into the shadows. Alternatively you can put a light on the camera but you need to be careful about the balance of the brightness and colour of the light so that the scene doesn't become unnatural.

Since many video compression techniques compare frames to help reduce data, it is a good idea to keep the camera steady and this is especially important with the high compression used for Internet video. In any case, a shaky camera is usually regarded as being a bit amateurish (or possibly the current fashion for access television). If you don't want to use a tripod and you can't afford a small Steadicam (and there are versions designed for camcorders) then you might get away with the shot steadying or shake compensation feature in the camera, especially if you shoot at the wide end of the zoom. Optical ones work best, but the electronic ones are surprisingly good as well, and only slightly reduce the quality. You should not use a shot steadier when on a tripod. This is because the smooth movements from the tripod will slightly confuse the system and it will counteract the start of, say, a pan and then have to quickly adjust. So, ironically, the steady shot can look jerky!

Beware of automatic focus, white balance and exposure on a camcorder. A camera's auto-focus might try to reset itself once in a while, so if you are shooting an interview and the person you are shooting is not moving, turn the auto-focus off. If you don't then there's a good chance that every so often the shot will go out of focus and then back into focus quickly, as the system resets. If you have your centre of interest at the edge of the frame the auto-focus may well try to focus on something else.

Auto white balance will be taking an average of the colours in the whole scene and adjusting the colour balance so that this average matches a preset average. If you take a shot of a face against a reddish wall then the colour will slowly drift away from red towards blue, to try to reach this preset goal, which will be something close to grey. Any strong background colour will affect the foreground. It is best either to use a colour balance fixed setting that is close enough – usually the choice is daylight or artificial light – and correct in the edit or to take a white card and set the colour balance on this, assuming the camcorder lets you. Fluorescent lights usually have a very strange colour spectrum and quite often faces will have a greenish tinge. Since these kinds of light do not produce a smooth continuous spectrum of colours to make up their 'white' light, unlike tungsten lamps or the sun, it may be impossible to completely correct the colour.

Auto-exposure will again be taking an average of the scene, or a weighted version of this, perhaps weighted to take more account of the centre of the picture (called centre-weighted). So if you have a brightly lit figure on a stage with the rest of the scene black then the camera will probably over-expose the figure. In these cases it is a good idea to zoom in so that the figure fills the frame, lock the exposure, and then zoom out again. That's a quick way of avoiding the problem if you don't actually have manual exposure.

There is one thing that always caused headaches for video cameras, and that is computer monitors. In this business you are very likely to be taking shots which include a computer monitor. Here you may be out of luck. Some camcorders allow you to change the shutter speed which might help things, but with most of them the computer monitor will just flicker. Using electronic shot steadying will make this worse, so it helps to use a tripod and turn it off.

Preparing for editing

The tapes from your recording are known as rushes, from the time when they were shot on film and were rush-processed so that you could see the results the next morning. They will be dubbed onto videocassette, usually VHS, with burned-in time code (BITC) so that you can equate the time code to the parts of the material you are viewing. A time-coded transcript of the rushes is a useful first step in editing, and can help you to choose which parts of the interview you want to use.

There are usually three stages of videotape editing, although in the new media world you may not actually do any of them in a dedicated facility since more and more video processes can be carried out on the desktop. The first stage of the process is to prepare for the edit. This will involve looking through the recording in the cold light of day and choosing which parts of it to use. The odds are that you will have recorded much more than you need for the final application. The result of this stage is a list of extracts, probably marked on the transcript with their time code numbers.

Offline

The second stage is the offline edit. Offline editing was originally introduced because of the high cost of editing equipment. Early offline systems used videocassettes, but offline editing was the first application of non-linear desktop video. The offline session is like a rehearsal for the online. Edits would be tried out, again and again, until the editor and the director settled on the ones that worked. The result of this stage was a list of the edit decisions.

The edit decision list can be taken into a ‘real’ edit suite and the edits carried out on the tapes. This is called conforming. If this process is done automatically it is called auto-conforming, and this technique came back into fashion in the early days of digital video tape machines because of their high cost.

Today an offline session is likely to take place on a computer, using a system like AVID or even Premiere, although the basic principle is the same. With a non-linear system the offline would be done using lower quality digitized video and for the auto-conform the full quality footage would be loaded up as needed. Although the same equipment is often used for offline and online editing, the amount (and cost) of hard-disk storage needed to hold everything at broadcast resolution usually means that high quality is reserved for the auto-conform.

Quite a substantial industry has grown up around these non-linear systems, as they are called. Non-linear is another way of describing what a computer person is more likely to call random access and the term is used to differentiate a hard-disk editing system from one using videotape. With tape your editing basically had to start at the beginning and work forwards, linearly, through the programme. Some non-linear editing systems are designed to seem familiar to video professionals, while other systems are more computer-like and so are not as popular with video editors but have their fans in the multimedia community because they feel familiar to them. Non-linear systems started as offline edit systems, although they have migrated upwards in quality, and have passed through news broadcasting to full acceptability as a high-quality editing medium.

Online

With a non-linear system the only difference between offline and online is likely to be the sophistication of the editing software and whether it has features for compositing or video effects. Otherwise, as already outlined, at the end of the offline session you virtually just press a button to load up high-quality footage and auto-conform it and then, finally, copy it onto videotape or, possibly, produce a computer file you can take away and compress. Or the facility might compress it for you.

You might still find yourself using an edit suite with tape machines, or one configured to be like a tape suite rather than a computer. The videotape machines, with their whirring motors and fans, were kept in a separate room, called the machine room. This is sometimes still the case, but with the quieter videocassette-based systems the machines are now often kept in a rack next to the desk. In some facilities it is common practice to have a central pool of machines and allocate them to suites as required, rather than having machines dedicated to a particular suite. So you may not even see the machines being used for your session. You may not even know whether you are using tape or non-linear until the edit starts.

You, as the client, will be in the control room of the edit suite. There will be a bank of television monitors in front of you, with a control desk and a vision mixer. It is one of the standing jokes of the video industry that in *Star Wars*, when the denizens of the Death Star zapped a planet, they controlled the destructor beam with a Grass Valley vision mixer, common to many videotape editing suites around the world. (Computer people get their laughs from Scottie trying to use a Mac Plus in the *Star Trek* movie about saving the whale.) The staff of the suite will probably be two people: the editor and an assistant, also known as the tape-op if there are tapes. Almost without fail a character known as a runner will come in from time to time and offer you coffee or toast.

Computer-based non-linear edit suite.
(Photo courtesy of Sightline, Godalming)

The editor's skill is in working out how to cut your material, to your specification, so that it flows and so that the edits do not jar. There are little tricks, such as cutting just before a movement in the incoming picture, which distract the viewer's attention momentarily and can be used to disguise the cut. With an edit, the material you have before the edit is called the outgoing shot, and that after the edit is called the incoming shot. The edit is adjusted by tightening it (making it happen earlier) or loosening it (making it happen later).

The editor will be editing sound and video. Often the sound will be mixed across an edit even if the video is cut. This is done to soften the impact of the cut. If the sound is heard before the vision cut it is said to be leading it. An edit where sound and video are cut at different points is said to be split (a split A/V). Sometimes you will continue the sound and come back to drop a short cutaway shot to replace some of the video.

Online videotape editing is a linear process. You started at the beginning of the programme material, edit shot onto shot, and finally reach the end. This meant that you needed to be satisfied with each edit as you did it, since you do not have much option for changing it. Hard-disk systems allow you to edit in any order, and you can go back to an earlier edit to change it if you want to. Despite this versatility it is important that you are satisfied with the edited material when you leave the edit. The last thing you do should be to view the material all the way through to make sure it is cut to your satisfaction. You might do this while the videotape is produced, if you are confident that it is OK. To be honest you don't have anything to lose by doing this. You should always have two copies of the finished programme, one can then be a safety for the other in case of problems. You could also ask for a viewing cassette to show, perhaps, to your client. Do not transport the main and safety copies together.

Just as audio studios have signal processing equipment for echo and delay, so a videotape edit suite will have its equivalent either built-in to the editing software or as self-contained boxes of equipment. One very common piece of equipment is the digital video effects unit, or DVE. With a DVE attached to the vision mixer, the video can be processed to make it change size, appear to move with 3-D perspective, and even be wrapped around shapes in real time. Such devices have a variety of names and manufacturers, but one legendary name was for a piece of BBC hardware, never marketed, which was called TIPSE. This stood for Technical Investigations Picture Shuffling Engine.

One important thing to remember about digital processing of television pictures using external boxes is that, without exception, the picture will be delayed as it passes through the device. Usually it will be delayed by one frame, which in NTSC is just over 33 milliseconds and in PAL is exactly 40 milliseconds. This delaying of the picture will eventually make the sound lose its synchronism, so it is usual to delay the audio to match, and edit suites have audio delay lines for this purpose. There are many opportunities for sound and vision to lose synchronization so this is something you should watch for.

Analogue videotape machines

It is possible that you will come across a variety of videotape formats, depending on the kind of work you are doing. Some of them will only be used now for archive footage, but it may be useful to know what the terms mean. Unless otherwise stated these machines can be found in NTSC and PAL formats, although component formats should be referred to as 525/60 and 625/50 (for lines and fields) or simply 525 or 625. SECAM and PAL are the same at a component level so there is no difference between a Betacam-SP or Digi-Beta tape recorded in France and one recorded in Germany or the UK. Similarly SECAM SVHS is the same as PAL SVHS as far as the recording on the tape is concerned.

The oldest tape format still used, even if only for archive material, is 2-inch Quad. Originally the American company Ampex developed a system called Quadraplex, which used 2-inch tape, in 1956. Other companies also made 2-inch machines for this format. Long-established broadcasters such as the BBC have archive material on 2-inch, and it was in regular use until the 1970s, so it includes colour. Some stock deteriorated over time because of physical problems with the tape itself, and the Quadraplex system was relatively violent, spinning four heads rapidly across the width of the tape, so it is possible when playing a 2-inch tape to damage the tape, especially if the edge of the tape is already damaged a little. If the owners will still let you use this material, and they allow it to be played despite the risk, then it should be copied onto a more recent format as it is first played. Oxide from the tape can easily clog up the tape heads, and sometimes engineers will use a piece of card or even their thumbnail to remove oxide from the heads as the tape plays. The basic rule is to leave 2-inch to the specialists, and, in any case, very few facilities will handle it these days.

There was a 2-inch helical scan system that was used by a few broadcasters. This is incompatible with the 2-inch Quadraplex machines. A well-aligned 2-inch recording can be of extremely high quality with a sharper signal than later analogue formats.

After 2-inch came 1-inch. This was a helical scan format, which means that the tape heads are on a drum that moves in almost the same plane as the tape path. Videocassettes are helical scan as well. The main manufacturers of 1-inch machines were Ampex and Sony (who made C format machines) and Bosch (who made B format machines). Almost all 1-inch tapes in the UK and USA will be C format but some countries, such as Brazil, Germany and Austria, used B format. The essential thing to know is that the two are incompatible.

The 1-inch tape format was in use from the early 1980s until the early 1990s, by which time cassette-based formats like Betacam-SP had replaced it. Sony developed Betacam as a derivative of their ill-fated Betamax VCR system. There is a difference between Betacam and Betacam-SP, and although most engineers will use the words interchangeably (or just say 'Beta') you should always specify Betacam-SP if you mean Betacam-SP just in case the video library also handles the older Betacam format. This is mainly a problem in the USA rather than Europe, where PAL Betacam never really caught on.

Betacam-SP has been very successful. It is a very handy system to use, with small cassettes being used in camcorders (with 20 or 30 minutes of tape time) and larger tapes in console machines. Video quality is good, and since it is a component system unlike 1-inch or 2-inch, the pictures are better suited to digitization. There is another component cassette format that competed with Betacam-SP, and that is M-II (M-Two). Beta machines come from Sony, Ampex and Thomson and M-II from JVC and Panasonic.

As far as sound goes, 2-inch tape carried a single mono track with reasonable quality, although it suffered from a high-pitched whine caused by the video tracks, which lie at right angles to the sound track. The orientation of the tape oxide was optimized for video and lay across the tape rather than along it, and so was in the wrong direction for audio which affected audio quality. The 1-inch C format has either three or four tracks, depending on configuration. The performance is good, but Dolby noise reduction was often used to improve the signal-to-noise ratio.

Betacam-SP has four sound tracks. Two are very high quality but cannot be used during editing since they use the video track although they can be used for straight recording. So in practice the other two lower (but still good) quality tracks are used most.

Cassettes

Although virtually all professional videotape formats are now cassette-based rather than reel to reel (also known as open reel), there are videocassette formats that are regarded as consumer or industrial. There had been domestic and industrial open-reel videotape formats, notably from Philips and Ampex, which became available in the late 1960s, but the price was far too high for widespread use.

Videocassettes first came onto the market in the 1970s with the 1500 format from Philips and U-Matic from Sony. U-Matic has high-quality (Hi-Band and SP) versions. It is still also used for digital audio as the Sony 1630/1610 format. However, you are unlikely to find U-Matic tapes used in high-quality video today since variants of Betacam-SP, digital video cassette (DV) and S-VHS/Hi8 have taken that niche. U-Matic cassettes use $\frac{3}{4}$-inch tape, and are sometimes referred to as the $\frac{3}{4}$-inch format. One serious disadvantage of U-Matic is that the maximum tape length is 74 minutes. (Coincidentally this is the supposed maximum duration of a compact disk.)

The most common videocassette format for domestic and industrial use is the irrepressible VHS. The picture quality of VHS is relatively poor, with particularly fuzzy colour performance. Hi-fi sound tracks were retrofitted to VHS, and achieve very high quality even though they are analogue. S-VHS uses cassettes of the same size as VHS but with different tape. S-VHS gives much better results with higher bandwidth, resulting in sharper pictures and better colour. S-VHS machines can usually play and record VHS but not vice versa, and both systems use half-inch tape. JVC developed a higher-quality version called Professional-S.

Video-8 is Sony's successor to Betamax, and boasts the smallest analogue cassettes in video. There is a high-quality version, analogous to S-VHS, called Hi8. These formats sometimes have digital sound. The tape width is 8 mm – hence the 8 in the names. There is a digital version of Hi8 using DV encoding on Hi8 tapes.

There are industrial versions of Betacam-SP with a lower price and a lower, but still very good, quality. The gap between S-VHS and Betacam-SP is also narrowed by the high-quality version of S-VHS available from JVC. One advantage of the industrial Beta systems is that they are compatible with the full spec. version.

Videodisks

John Logie Baird actually sold 78 rpm videodisks in the 1920s, but it was not until the 1980s that commercially viable videodisk machines came onto the domestic market. In Europe the domestic market for videodisks never really caught on, although it was successful in a niche for film fans and classical music enthusiasts. This market virtually disappeared when high-quality digital video from a DVD became available. There is more on DVD video in Chapter 3 of this volume.

Interactive video, which uses videodisks with computer control, was the precursor of multimedia, and you are unlikely to be working with videodisks (in any case that is outside the scope of this book) any more. However, you might use videodisks as a source for video and audio for multimedia, and there are a few points to bear in mind.

PAL LaserDisks can, if they are older, have analogue soundtracks. In any case PAL disks can have only two tracks, analogue or digital but not both, whereas NTSC disks can have both and often do. If you are trying to digitize from an NTSC videodisk, check whether your player outputs real or modified NTSC. If your digitizer sees black and white then this might be the cause.

Sony produced a WORM (Write Once Read Many) videodisk that records component video and is considered to be of broadcast quality. This system has been used by some television stations to hold station identification sequences because of the random access and lack of deterioration (unlike tape) but has been superseded by digital storage.

Pioneer produced a magneto-optical recordable/erasable videodisk that also had two separate head assemblies, which can be used independently for playback. This is also a component system.

Both these machines exist in PAL and NTSC versions.

Digital formats

Video is digitized in much the same way as audio: the waveform is measured many times a second and the resulting value is stored as a digital number. It is possible to digitize composite video but the highest quality systems digitize the component signal. The standard for this has a sampling rate of 13.5 MHz (for both NTSC and PAL timings) for the luminance component. The two colour components are usually sampled at half of this rate. There are several ways of sampling the luminance and colour components which are all denoted as ratios with the number 4 representing the 13.5 MHz sample rate of the luminance. In the studio, where bandwidth is not such an issue, a configuration of 4 : 2 : 2 is used meaning that the colour information is sampled at half the rate of the luminance. This is what happens along each line so the colour resolution is higher horizontally than vertically. On tape, and in MPEG-2, the colour is effectively sampled every other line, evening out the horizontal vertical resolutions and saving about 25% data. This is known as 4 : 2 : 0.

For production it is important that the colour bandwidth is great enough for chroma-keying and 4 : 2 : 2 sampling gives a chroma bandwidth of 3.37 MHz, which is high enough although there is a four channel system, 4 : 2 : 2 : 4, with a separate key signal at high resolution added to the 4 : 2 : 2 format.

Digital video is digitized with 8- (and occasionally 10-) bit resolution. To people coming from audio this seems crazy since eight bits gives a very noisy result with sound. But with a video image, noise is not perceived in the same way. The effect of noise is even reduced by the way the eye and brain are inherently turning the series of stills into a moving image so a movie will seem to have less noise (or grain in the case of film) than an individual frame will. So eight bits will suffice.

Digital video levels do not go between 0 and 255 in an 8-bit system as you might expect. Black is at 16 and white at 235. This is to minimize the effect of spikes and overshoots in the digitized waveform but it can lead to 'real' video seeming washed out when shown on a computer screen. Conversely computer-generated graphics can seem crushed at black and white levels unless the 16/235 limits are taken into account.

In the 1990s digital videotape formats began to replace analogue ones. First there were D1 and D2. D1 is a very high-quality component digital system developed by Sony, but unfortunately the cost of machinery and tape stock made it popular only with advertising agencies, film companies and video research. D2 was a derivative system, which recorded digitized composite signals (PAL or NTSC). As a result it could be easily integrated into existing systems, and was more popular. Neither D1 nor D2 compressed the data used to record the video or audio. However, it was the advent of half-inch digital systems that brought digital video within the cost range of the majority of broadcasters.

D3 is a composite digital system, D4 does not exist, and D5 is a component digital system without compression. There is a mode for D5 that supports high definition pictures with some compression. Digital Betacam is from Sony (of course) and uses a 2 : 1 compression in the signal. Some Digi-Beta machines can play Betacam-SP tapes, which makes for versatility in editing. It is also possible to get Betacam-SP machines with digital inputs and outputs, but these are not Digital Betacam.

Even though there is a digital output from these tape formats, the usual way of copying the pictures into a multimedia system is by digitizing or grabbing the analogue output. This is partly because of the high data rate of professional digital video (270 megabits/second) and partly because of incompatibilities in the shape of the pixels between computers and the international standard for digital television.

A consumer digital videocassette format using 6 mm tapes, called DV, has become popular for high-end consumer, industrial and even some broadcasting applications, especially in news. The format is not fully component but uses the Y/C format which separates luminance and colour but only has a single colour channel. Two semi-compatible professional derivatives, DVCAM and DVCPRO (which also has a derivative high-quality version), together with a long-play derivative of basic DV, are also available. DV machines are now made by all the major camcorder manufacturers, and at its best with a three-chip camera and good lighting, DV produces results that are hard to distinguish from broadcast equipment. DV has a 5 : 1 compression which does not compare succeeding frames. Sony has produced a digital version of its Hi8 system which uses DV encoding on Hi-8 tapes and gives owners a backwardly compatible path to their old recordings. As the tape is bigger than DV the Digital-8 camcorders cannot benefit from the same level of miniaturization. Most of the manufacturers provide a digital connection (using the new IEEE-1394 standard, better known as FireWire) with which video can be copied to and from a computer. This system provides a very cost-effective high-quality way of shooting and preparing video for multimedia on a website or CD-ROM. There is a limited interoperability between these formats, which mainly means that the professional formats can play back the domestic tapes. Some consumer DV machines can play back the professional tapes although the pro formats run the tape slightly faster – for example a 60-minute DV tape recorded in a DVCAM machine only lasts 40 minutes.

JVC has introduced a further broadcast digital cassette system called Digital-S, and has announced D-VHS as a time-shifting format for the home market. Sony has an industrial/news-gathering digital version of Betacam, Betacam-SX. An SX machine can play SP (i.e. analogue) tapes, and so an SX machine offers an upgrade option for anyone with a large archive of SP material. It is interesting to see how many competing digital videotape formats have been introduced in the past few years. The market is obviously now large enough for the manufacturers to try to shake out standards in a way that was not really possible 20 years ago. The most interesting aspect of these new formats for multimedia is that the inherent quality even of the consumer DV format is so high that the tape format is not really the limiting factor: the quality of the lens and camera chip (CCD) have the most effect.

Copying tapes for use in multimedia

Since it is extremely unlikely that you will be allowed direct access to a master videotape it will be useful to outline some suggestions for formats.

If the master is composite then your working copy can be either composite or component. Since, as we shall explain in a moment, it is better to digitize from a component source, in most cases you will want to ask for a component dub of the master. If you can work with a digital source then you should ask for a component digital copy. As discussed above, one very practical option for desktop video is to use a DV or DVCAM dub and work from this, especially if you can transfer the video to your computer using FireWire. The lack of tape noise on DV makes it especially attractive for compression, assuming the lighting was sufficient and did not lead to noise from the camera circuitry.

It is possible that you will be editing and possibly processing the video to make a new master tape. This might happen if you were to be encoding material for MPEG on DVD and wanted to prepare a master of your own first. This might include video noise reduction and re-editing. In this case a digital composite dub would be recommended (such as D3). In this way the decoding of the composite video to component can be done as part of your processing, and you will have more control over it.

Under no circumstances should you get a composite dub of a component tape; neither should you digitize from a dub on videocassette such as VHS, S-VHS or Hi8 unless you really have no alternative. Incidentally, you can save money by asking the facility house to dub a VHS with burned-in time code in vision, for your viewing, as they dub the master.

Finally you should not standards-convert a tape between PAL/SECAM and NTSC or vice versa. Your copy should be in the same standard as the original, assuming you can digitize from the format of course, so that you will not be encoding conversion artefacts.

A final note about DVD. DVD players have a system called Macrovision in them specifically to stop analogue copying and this is likely to interfere with any dubbing you try to do. The MPEG information on the DVD will also be encrypted so you can't just copy the file unless you legally have access to decryption. It may be that in future there will be legitimate reason to dub from a DVD because other versions no longer exist. But since you will be legally licensing the material (you will … won't you?) you have to ask for a copy in a format you can use, such as broadcast tape, or an unencrypted MPEG data file.

Digitizing video for multimedia

Digital video in multimedia can mean partial screen or full screen, full motion (that is, 25 or 30 frames per second) or partial motion (down to 10 or fewer frames per second). The details of digitizing from a video source are very much down to the particular equipment in use, but there are some general principles. Remember that here we're talking mostly about digitizing from a video source, not copying DV into the computer using FireWire.

'Recording' the video into the computer is called frame grabbing, and this grabbing should, if at all possible, be done at full frame and full frame rate. It is possible to frame grab uncompressed, but this requires very fast digitizing hardware and local storage. The kind of hard disk used in a desktop computer cannot work this fast unless several are joined together in an array. (When it comes to copying DV, the video is already compressed so most hard disks can cope with it but DV still takes up a lot of disk space: over 200 megabytes per minute for sound and vision.)

The pixels of a television picture and the pixels on a computer screen are not actually the same shape. This affects 525 and 625 line pictures differently, but essentially means that when copying a digital TV signal to the desktop, such as DV via FireWire, the aspect ratio of the image changes slightly. If you are editing on the computer and copying the finished video data back to the DV machine then this is not a big problem, but if you are going to distribute the video on a computer then you need to correct the aspect ratio change at some point. A frame grabber will sort out the pixel shape as it digitizes the video.

In order to play the video back on the computer it will have to be compressed and probably shrunk in size, but this is better done ‘at leisure’ rather than in real time. From a practical point of view high-quality JPEG can be used during the grabbing process, so that the movie becomes a series of JPEG stills, which will save disk space. The results from this will usually look very good.

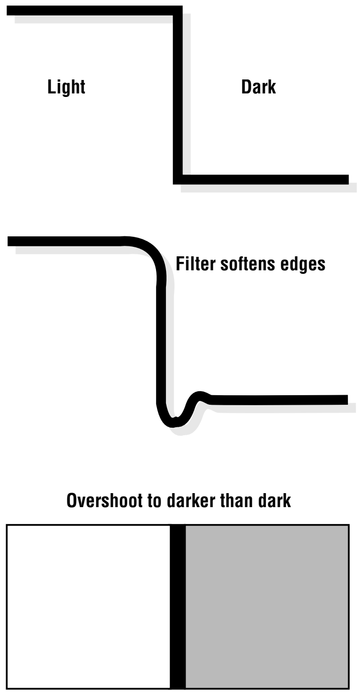
One important point about frame grabbing is that the frames you store on your hard disk will be either component or RGB; they will not be composite. This means that if you are grabbing from a composite source the grabbing board will need to convert (decode) NTSC or PAL into component or RGB. This will involve filtering the incoming signal to separate out the colour from the luminance. Unfortunately, the cost of a good filter to do this would be much higher than the cost of the rest of the equipment put together, so the filtering in a commercially viable grabber board has to compromise on quality, and some artefacts may be noticeable.

First, the luminance bandwidth may be compromised by the filtering out of the colour subcarrier if the video input is a composite one. Many filters (including those in most television sets) remove the higher frequencies of the luminance signal, those above the subcarrier, as well as the subcarrier itself.

Second, filters ‘ring’ and this has an effect on vertical edges in the picture. A sharp vertical edge has high frequencies in it: the sharper the edge, the higher the frequencies involved. Passing such a signal through a filter can have a detrimental effect on the shape of the edge, and the signal can overshoot and ring (just like a bell). What you see is a dark band between a bright and a dark part of the picture, darker than the dark part. Now if the dark part is black this results in a small part of the picture that is blacker than black, which is possible in an analogue television signal as shown overleaf. For this reason, among others, black and white in a digital television signal are not set at 0 and 255.

A good filter can avoid these kinds of problems, but a better way is to avoid filtering altogether and always digitize from a component source, such as Betacam-SP or Digi-Beta. If the signal has to be converted from composite, because it was originally recorded that way, then a high-quality broadcast filter can be used during a dub from the original tape or a digital composite copy of the original tape to a component tape.

Taking a movie film and recording it onto videotape uses a machine called a telecine and the process of putting the film onto tape is called a transfer. For digital video in multimedia this telecine transfer should be done into a component videotape, and preferably a digital one. The film should be clean to avoid dust, and for top-quality transfer the imaging gate through which the film passes in the telecine machine can be submerged in a liquid with the same refractive index as the film base. This is called a wet gate, and produces excellent results. Many telecine facilities also have

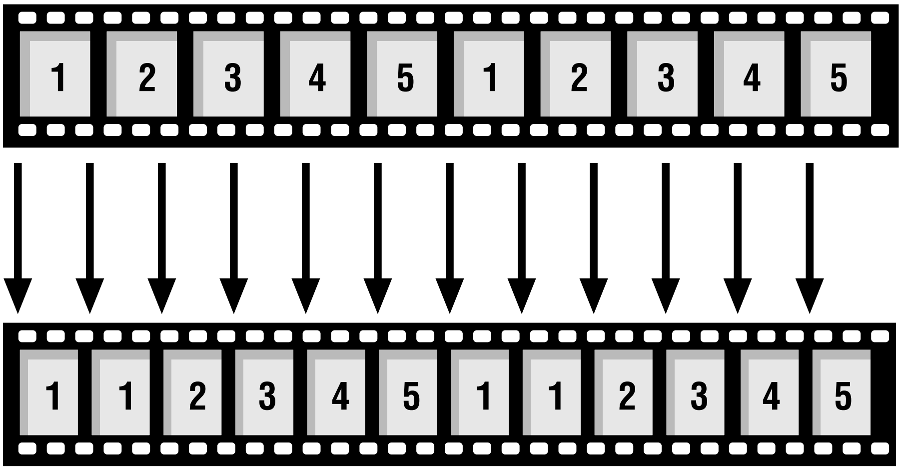


Ringing.

digital noise reducer, which make use of the similarity between frames to 'remove' dust and scratches or have infrared sensors to detect dust and so help remove it.

Video digitization should be done from the original television standard- and not from a standards-converted videotape. This is because standards conversion changes the image, even if only slightly. In any case, for digital video, the computer controlling the playback is generating frames at its own rate, and the system will cope automatically with the input frame rate. Unfortunately it will do this by attempting to read out a frame when it needs one, even if there is not a new frame available. In this case it repeats the last frame. (The opposite may occur where frames are dropped.) There is nothing you can do about this, but you would make it worse by digitizing from a standards-converted source where this process, or a sophisticated version of it, has been carried out already.

A television standards converter will interpolate new frames, or combine old frames, in order to avoid the jerkiness that this basic 5–6 conversion has, but a computer display does not do this. This may suggest that, if you have the freedom of choice, you might actually use NTSC rather than PAL for your video in Europe, because the NTSC frame rate is more like that on the computer.



To read out six frames where five are available means that one frame is repeated.

Another set of problems is caused by the television signal being interlaced. When a television camera captures motion during a frame it does so in two separate slices of time, a 60th or 50th of a second apart. A moving object will have changed position between those two slices, the fields. When this is played back on an interlaced television screen there is no problem because the output display is interlaced with two fields per frame in the same way as the input camera. This does not happen when the full frame is displayed on a computer screen because a computer screen is not interlaced, but writes the picture in one sweep from top to bottom, left to right. The result is that horizontally moving objects can appear to break up like a comb. Vertical filtering, also known as convolution, can be used to correct this fault but at the expense of resolution. It does not occur when the camera was a film camera because a film camera exposes the whole frame at once.

This situation will also only occur if the video on the computer screen is the same size (has the same number of pixels) as the original video grabbed. If it has been shrunk to half size, for example, to make a quarter-screen image of 320 by 240 pixels, then the result will be a slight horizontal blurring or a double image rather than a comb effect. Alternatively you could simply drop one of the two fields, although this might produce artefacts.

When shrinking an image (it does not matter whether it is movie or still), new pixels should be calculated from combinations of the old ones, rather than by sub-sampling (that is, only selecting some of) the old pixels. If this is not done then near-horizontal lines will ‘staircase’ and also appear and disappear with movement of the picture, as they go from ‘visible’ pixels to ‘no longer visible’ pixels and back again. This vertical filtering (to blend

Combing.

together information from the fields) is very important when dealing with television pictures in the digital domain.

Some computer systems will allow the computer display to be set up to be interlaced, but use of this is rare unless the application is going to be shown on a television monitor. Using an interlaced display with a computer causes flickering of fine detail in the picture (called twitter), and filtering this out reduces the effective resolution of the image.

Compression

Video in multimedia is, by definition, digital. For practical reasons it also has to be compressed, because it is impractical to handle the high data rate used by raw video. Broadcast technology is able to handle this, but at a price and even digital broadcast television is compressed. For multimedia we have a choice of various 'standard' methods of compression, and the choice available increases as time goes by. These standards use compression and decompression software or hardware called codecs.

Two of the international standards that are applied to video compression are JPEG and MPEG. JPEG, which is discussed in the chapter on graphics (Chapter 8), is a compression system used for still images. JPEG is attuned to work with photographic images, which have smooth gradations of tone rather than sharp lines. JPEG works in only two dimensions: the height and the width of a frame. MPEG adds further data reduction by comparing groups of successive frames. Since not all of an image changes – frame by frame – in a movie, this redundancy can be used in the compression.

With sufficient data bandwidth available, from a fast hard disk for example, JPEG images can be displayed fast enough to become a movie, and this is a pseudo-standard called M-JPEG and is often used for non-linear editing. MPEG uses less data but is more difficult to decompress (and compress). Both of these standards, when used for movies, require hardware or a fast processor. MPEG as used in video compression actually encompasses three variations, with MPEG-1 being the version designed for use on compact disc at 1.5 megabits per second. MPEG-2 is a higher-quality, higher-bandwidth version, which is used for digital television and DVD, and MPEG-4 is aimed at low-bandwidth applications and interactivity. Recently, codecs like DivX ;-) (which does include that smiley in its name), and 3ivx have appeared, based on MPEG-4, and are aimed at watching movies via broadband connections.

Several software video compression systems exist, making use of operating system architectures such as Windows Media and QuickTime. For use on the Internet the issue is how the systems will perform with a low bandwidth, and at the time of writing the most popular standard is RealVideo, which is designed to cope with modem or ISDN data rates. In this case the video is streamed across the Internet, rather like broadcasting, instead of being copied to the user's computer.

As a general rule, video codecs that give good results at one data rate will not perform as well at a very different one. This is called scalability (or lack of it), and means that increasing the data rate will not allow an increase in quality because the codec has a quality ceiling. Cinepak is a codec that behaves in this way and has been replaced by the Sorensen codec, which can also work at higher rates.

In fitting movies into an operating system environment, the manufacturers (such as Apple and Microsoft) define standard interfaces for control of the movie with VCR-like controls for play and fast forward and rewind. A progress bar, to show how far through a movie you are at any moment, is also a standard part of such an interface.

QuickTime and Windows Media both can work transparently with a range of video compression systems, since the operating systems present a consistent interface to your application. This means that you can switch from using, say, Indeo to using MPEG-1 without changing the application as long as the two videos have the same resolution and share the necessary control parameters. However, you have to be confident that your intended audience has the codec you use since few will be willing to download a new plug-in or codec especially for your video. On the Web this usually means RealVideo, although Sorensen (through QuickTime) or Windows Media's MPEG-4 codec are serious alternatives.

With the movie come sound and stills, since an audio file is basically a movie with no pictures and a still is a movie consisting of only one frame. It is MPEG that currently offers the best set of these features since MPEG audio and MPEG stills are of extremely high quality and are very efficiently compressed.

Judging quality

The points outlined in the previous chapter on quality as it applies to sound also apply to video. However, it is usually very difficult to achieve broadcast quality with video on a multimedia system (if not currently impossible on a website), and your client should be aware of this. Again, use of the term 'appropriate' to describe the quality may be the best option. Some of the quality issues are less tangible. The quality of editing and the camerawork are independent of the technical quality of the desktop video, and in these cases you should aim for a quality threshold similar to that of broadcast television. Explaining early on that the multimedia video will not look as good as TV, and showing the client some examples, will avoid this becoming an issue, but you should be prepared to explain why, especially to a client who has experience of corporate video or broadcasting.

THEORY INTO PRACTICE 7

One of the ways of building video skills is by watching. Look carefully at what you see on television. See how people are framed by the camera, and watch particularly how faces usually look into space within a frame. Look at how a movie is cut, and think about why a cut might occur in a particular place. Watch for the way the incoming shot often starts with an action that will distract your attention from the edit itself.

If you have a camcorder then you can try some of this out for yourself.

You should also look at good quality web video, such as that on the BBC site (www.bbc.co.uk) and compare how it looks with television.

Summary

- Moving pictures are a very effective way of conveying information in multimedia.
- You need to be aware of the different broadcast television standards that you might encounter and of the difference between composite and component video. NTSC and PAL are the main forms of composite video, and each has a component and digital equivalent.

- For a variety of reasons component video offers better quality than composite and is more suited to digitizing.
- Time code is a numbering system used in videotape recording that will also be useful to you. You can use it to show tape numbers as well as time.
- Shooting a video interview is one of the most likely kinds of video you will undertake. The subject can be separated from the background by shooting against blue or black.
- If you want to shoot video yourself using your camcorder then you need to take care over things like colour balance, focus and exposure.
- The stages of editing video are preparation, offline and online. New technology and non-linear editing systems using PCs have blurred the distinction between online and offline editing.
- Options for compressing video so that it can easily be handled on the desktop include moving JPEG, MPEG and software codecs. The most common equivalent for the Web is RealVideo with QuickTime and Windows Media as alternatives.
- Codecs will usually have a range of data rates and applications for which they are suited, so a web codec may not work well on CD-ROM and vice versa.

Recommended reading

Watkinson J. (2000). *Art of Digital Video*. 3rd edn, London: Focal Press

Watts H. (1998). *On Camera*. 2nd edn London: Focal Press

Quantel Ltd *The Digital Factbook*, http://www.usa.quantel.com/dfb/

Zoom Culture, a US Internet Video production company, have produced a quirky but excellent video (In Real format) showing the dos and don'ts of shooting video for the Internet. The URL is http://www.zc.tv/clips/rams/046738.ram