Proceedings of the Institute of Acoustics

AN APPLICATION OF DIGITAL TECHNOLOGY TO STEREO BROADCAST POST PRODUCTION

Paul Roberts

Television South West, Derry's Cross, Plymouth, PL1 2SP

1. HISTORY

FILM SOUND:

The post-production of sound, known as "dubbing", was from the sixties until the early seventies, mixed using magnetic film sound stock in a film dubbing theatre. Still very successful and very much in use, the film dubbing theatre is equipped with several sprocketed replay channels , anywhere between three to thirty five machines. They feed the sound through the mixing desk on to one or two sprocketed magnetic recorders. These machines, 35mm, 17.5mm, or 16mm, are linked either electrically via three-phase selsyn motors or electronically through stepping motors.
The ability to run the tapes backwards and forwards, "rock and roll", and drop back into record, "punch in", allows the dubbing mixer to work through the programme, stopping as each mistake in the mix is made, roll back, and pick up again. The modern film dubbing theatre can rock and roll at very high speeds, up to thirty times replay speed, in either direction and maintain almost perfect lock at all times when running at these speeds.
The sound tracks are laid up before the dubbing process on comparatively simple editing tables. This allows the dubbing editor to try different music, sound effects, or dialogue against the picture without using expensive dubbing time. The tracks can be moved easily backwards and forwards against the picture, and each other, either by editing in, or out lengths of film spacing. They can also be unlocked and physically moved in the dubbing theatre.
16mm magnetic film, the main medium used in television, has a rather limited band width, about 50Hz to 14KHz within one or two dB. The signal-to-noise ratio is about 45dB without noise reduction. Drop-out frequency is also extremely high as the tape base is rather thick and the magnetic coating relatively crude. None the less film dubbing is a very fast and powerful post-production sound process.

1. HISTORY CONT.

VIDEO SOUND:

As video tape became widely used in television, and was recorded and then edited, it became obvious that the two sound tracks available on the tape were not enough for sweetening the sound. To increase the number of tracks the first machines locked to video were multitrack tape recorders. Designed for the music industry they were, and still are, superior in sound quality owing to the use of better magnetic tape and, in the case of 16mm, higher tape speeds.
They also offered the advantage of a near perfect phase relationship between tracks on one piece of tape. When dubbing from one set of tracks to another it is possible to have near silent "punch ins", always a problem with separately locked film tracks.
Totally silent punch ins are not possible with magnetic tape. The record bias is usually set somewhere between two to seven dB over bias at 10KHz to reduce distortion. Therefore, on record, the bias will always pass through its peak as it ramps up causing a momentary increase in the high frequency response of the recorded material.

DISADVANTAGES:

The "advantage" of having tracks locked safely together is also the multitrack's main disadvantage.
Unlike film, moving sound tracks around is not at all easy or cheap. All the work is done in the dubbing theatre, where small sound delays are possible with delay lines, but for anything to be moved earlier or significantly later the sound must be lifted on to another tape machine, repositioned and relayed.
The other problem for a multitrack, when rocking and rolling, is maintaining a quick relocking speed with the picture. The timecode, recorded on to one of the sound tracks, maintains a very firm lock when running at normal speed, but as soon as the picture runs back at speed the multitrack, with a different spooling speed, loses lock. Often the synchronizer has to rely on tacho information alone. The whole system has to park up, and wait for the slowest machine, before setting off again.

It is possible to lock film machines to video but neither system is perfect.

2. STEREO TELEVISION

It has now become the norm to offer programmes to the television networks in stereo. Stereo adds another dimension to programmes and another dimension to the problems.
All stereo programmes have the left and right legs summed together, at the transmitter, to provide a mono feed. It is therefore imperative that the signal is phase coherent.
Analogue systems can introduce major phase errors. These mainly arise during the recording process. Misaligned heads or tape weave causing azimuth and tracking errors.
Analogue mixing desks can also add subtle problems of their own. Stereo equalisers and filters add small time delays, and if their parameters do not track properly they can introduce a lack of definition to the signal. A mixing bus will introduce small amounts of cross talk also impairing definition.
Most broadcast equipment is built and maintained to very high standards to minimise the possibility of any errors, even so they can and do happen.

3. DIGITAL ANSWER

AMS Industries were one of the first companies to realise the possibilities of using computer hard disks for storing digital sound. They developed a basic system that has grown from constant user feedback. The software has matured over a period of about six years, each update refining the process one stage further.

4. DIGITAL HARD DISKS

The Audiofile is a computer that handles sound in much the same way as a word processor handles words.
The analogue sound is sampled at either 44.1khz or 48khz. The sampling frequency must be at least double the required frequency response. Unfortunately the worlds of commercial recording and broadcast sound have allowed several separate standards to blight the recording chain.

4. DIGITAL HARD DISKS CONT (1).

The convenience for digital sound to be frame referenced with the video, led the broadcast world to choose a sampling rate of 48khz.
When a video frame is edited an integer number of sound samples is needed in as few frames as possible, this will prevent a gradual build up of uneven samples causing a glitch.
There are four video frame rate standards, 30, 29.97, 25, 24 frames per second.
48khz allows an integer number of samples for each of these frame rates to be achieved, in most cases, in only one frame. At 29.97 it can be achieved in five frames, which is still acceptable.

SAMPLING FREQUENCY KHz	FRAME RATE Hz			
	30	29.97	25	24
48	1	5	1	1

The sampled sound is quantitized, measured into 16 bits, and recorded as a digital signal on to computer hard disks. A standard Winchester drive of 360 Mbytes can store just under one hour of sound at 48khz sampling.
Each block of sound, anything up to the maximum storage available, can be recorded against timecode, if not it will have a timecode starting at zero allocated to it. The stored sound can then be trimmed with, if necessary, a fade or ramp at the top and tail. This information is stored onto the hard disk as well.
The digital information is buffered from the hard disk into RAM. The RAM is configured as a "FIFO", standing for "first in first out". With current AMS hard disk technology on offer there are a maximum of 16 FIFO's available. These can be split between 8 inputs and 16 outputs.
The stored sound on the disk is now ready to be used against the vision. The eight track version of the Audiofile has a piece of graphically presented "software" tape that is twenty four hours long.
The stored sound with timecode will synchronize on any track chosen to the correct part of this tape, or it can be repositioned. All other stored sound can be placed at will.
It is also possible to mark the sound into segments. These can be called up for use if and when needed. Each sound, or part thereof, can be started or finished with a ramp or fade. These are linear and can be set in steps from 2 msecs to 10 secs.

4. DIGITAL HARD DISKS CONT (2).

Now that the sound is in the digital domain "time squeezing" or "expanding", at the same time as pitch correcting, is possible. All these editing processes are non-destructive to the original sampled sound, allowing a flexibility in editing not possible with analogue tape.

It should be noted that with systems using eight FIFO's, only eight outputs are available at one time, not eight tracks! So if a stereo sound is edited on a pair of tracks the resultant edit will use up four FIFO's not two! Each FIFO requires about eleven frames to reload.

5. DIGITAL MIXING

The next obvious step towards the fully digital studio is the mixing desk. The approach taken by AMS has been to develop a desk that is not only digital but also totally automated; the Logic series.
The amount of computing power required for these desks has only recently become practically available in the form of transputers.

With analogue desks fader and mute automation has been around since the early seventies. Although an extremely useful tool it has limited use in the dubbing theatre.
In sound dubbing, unlike other mixing processes, most of the time only a few faders are in use. A lot of time and effort though is put into equalizing, dynamic control, auxiliary buses, and now with stereo, panning. With full automation of all these controls the Logic desk is a very powerful tool.

At the desk's heart are the transputer controlled signal processing cards, "TSP's". They process all the digital information, enabling filtering, equalization, and dynamic control of the sound. As with the Audiofile the system is 16 bit and works at either 44.1khz or 48khz sampling frequency.
The automated channel faders are moved by linear motors, and all the switch covers and control knobs are made from conductive plastic. By touch the computer knows when to start updating any automation data.

The stereo path through the channels in a digital system is phase correct and accurately tracked. As the equalization and dynamic networks are software driven they can act upon the signal without any analogue delays, phase errors or the usual analogue restrictions.
There are up to four fully parametric EQ bands for each channel or group. Each one can be swept from 20hz to 20khz, with the "Q" variable from 0.1 to 10, the shape can be peaking or shelving in either direction, and the boost or cut is 20dB.

5. DIGITAL MIXING CONT.

In addition one or two filters can be set from 6dB per octave up to 24dB per octave, either as highpass or lowpass, with continuously variable frequency from 12hz to 20khz.
The compressors, expanders, and gates can be set in channel, or groups, with or without EQ side chains. As with the EQ-parameters the settings are beyond conventional analogue dynamics.

The Logic desk is available in two basic versions. Logic 1: a stereo twelve channel, two group, one main output desk, or Logic 2: a stereo twenty four channel, four group, one main output desk.

The number of TSP cards installed governs the amount of processing power, and therefore the number of EQ, dynamics, pans, auxiliaries, etc.
The operator can choose where all these facilities are, on a "desk set up" screen, thereby configuring the desk for any particular session.

Inputs and outputs to the outside world are provided by digital to analogue converters, "DAC's", analogue to digital converters, "ADC's", and AES EBU digital stereo in-out cards.
Like the TSP's the number of these links are up to the individual user.

6. THE SYSTEM

The Audiofile control surface and Logic 1 control desk are housed in the same unit. All the main processing cards, power supplies and hard disk drives are stored, in a separate rack, in a separate room. The noise from the rack's cooling fans requires isolation from the sound control room.

An eight output system is just adequate for stereo dubbing for television when mixing on to a mastering machine. We are a present using two DAT machines locked via an ES Bus synchronizer.

6. THE SYSTEM CONT (1).

The AES EBU digital data link provides a two channel link. It is practically the same specification as IEC 958.
AES EBU is a balanced interface with a send impedance of 110 ohms, and uses 110 ohm cable. The receiver impedance is 250 ohms, this allows parallel connection of several receivers.
The link can send up to two channels of digital audio with a maximum of 24 bits per channel. The information is sent as a 64 bit frame. Each frame has two sub frames, A and B, with 32 bits per sub frame.

```
I---------------sampling period = 1 frame----------------I
I-4-I-4-I-----20-----I---4---I
```

sync	aux	audio data	V	U	C	P	sync	aux	audio data	V	U	C	P

```
I---------sub frame A--------I--------sub frame B--------I
```

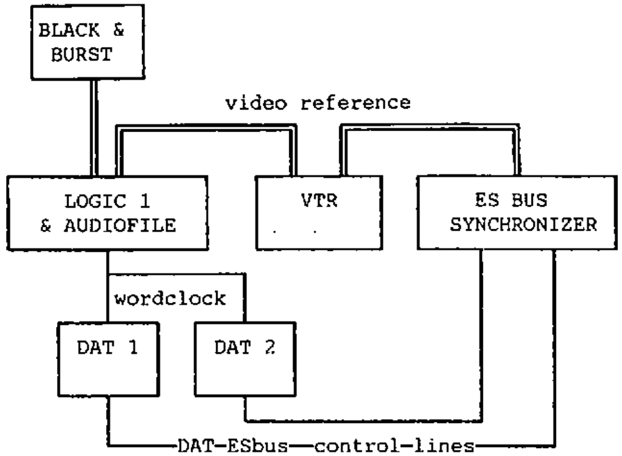
V = audio sample validity bit
U = user's bit
C = channel status bit
P = sub frame parity bit

Although when used for direct machine to machine linking AES EBU does not require any external reference, it is wise in multi machine use to provide an external clock. This is to prevent the build up of a timing jitter owing to each piece of equipment regenerating a clock from the audio signal.
The external clock used, the "word clock", is one pulse per sample of audio. (48khz)
With a system that is linked to a VTR, a video reference, known as black and burst, is usually supplied by a sync pulse generator. This master reference is fed to the Logic 1, VTR, and ES Bus synchronizer.
The Logic 1 generates a word clock signal from it, which is sent to all ancillary equipment.

The whole system is now referenced to one signal.

6. THE SYSTEM CONT (2).

SYSTEM LOCK:



7. THE OPERATION

A hard disk system will "lock up", to a video timecode feed, within four seconds. No matter where the "software" tape is, it will always take the same amount of time. With hard disks, unlike multitracks, there is no waiting for the system to park.
A digital system does not wow or flutter, multi generations do not deteriorate the signal, and no routine machine line up is necessary. Taking into account the editing power as well, the system sounds pretty perfect.
Like every thing in life there are drawbacks.
Once you have run out of storage time, there is nothing else for it, you have to start losing something. Either you can erase samples from the disks, or they can be backed-up on to a different medium.
At present most back-up systems run in real time. If you have one hour on the disk to back up it takes one hour! If you have recorded two half hour mono signals simultaneously, they will still take one hour to back up.
As the sound samples are backed up they take with them all the editing information stored on the disks. When restored back again, all the information is complete with no generation loss.

All digital sound hard disk suppliers are working on faster back up and restore systems, but until they arrive careful disk management is a priority.

After a year of use, I am not yet sure if DAT machines are the final solution for mastering the sound. With the advent of several digital video formats, it could be more elegant to use the four digital tracks they have available.