Reproduced from
Studio Sound, May 1990

The Gentle Art of Digital Squashing

Michael Gerzon takes us through the various methods of data compression and their feasibility for future applications

One of the problems with digital audio is the large amount of data it requires. Ignoring error-correction overheads, which can add about 30% to the data rate, the CD standard of 16 bit stereo at 44.1kHz sampling rate transmits 1,411,200 bit/s, which is around 10 Mbytes/min or 600 Mbytes/hr.

This very high data rate uses up a lot of expensive bandwidth when broadcast, sent down telephone channels or by satellite. When stored on tape, in RAM or ROM or on hard disk, an awful lot of memory is easily used up – witness, for example, the limited sampling times available on samplers and the high cost of hard disk memory in digital editors. If one could 'compress' this data rate to, say, 4 bits/sample without losing quality, one could get practical terrestrial digital broadcasting, extra long play CDs and quadruple hard disk storage or sample memory length.

The philosopher's stone of top-quality audio in as few bits as possible has been pursued for several years, based on lower-quality systems of audio data compression developed in the 1960s for telephone network applications. Some of the current systems now claim CD-indistinguishable quality at less than 2 bits per sample, and others on the market use 4 bits per sample. Clearly this technology is a coming thing, and we can expect to see many systems become commercially available. Solid State Logic's *Apt-X 100* system (a 4 bit system) is the first of this newer generation, although earlier systems such as those of dbx, Dolby and the BBC's NICAM system have been around for some years.

To non-specialists, audio data compression appears almost akin to black magic. The technical literature describing such systems is full of esoteric technical jargon on Rate-Distortion theory, Transform Coding, Adaptive Differential Pulse Code Modulation, Entropy Coding and so forth. Since such systems are going to become commonplace, and because their use is going to require some understanding of their strengths and weaknesses, there is a crying need for a straightforward description of how they work. And the fact is, that although detailed engineering design of such systems requires a lot of theory, their basic principles are surprisingly simple and understandable.

How do these systems work, do they really give results indistinguishable from 16 bits and what advantages and disadvantages do they have?

A word of caution at the beginning. All the systems giving a large reduction in bit rate *do* alter the audio signal, and what comes out is not what goes in. The trick in designing a good system is twofold: to make sure that the difference between the output and the input is as small as possible; and to design the nature of the errors in the output to be subjectively difficult to hear in the presence of the signal, *ie* to fool the ears by psychoacoustics into not noticing the error.

Before we get bogged down with the details let's look at systems that do not introduce any error in the output. These systems, known as entropy coding systems, use information theory to spot systematic patterns in the signal, and to rearrange the information in the signal to exploit these patterns to reduce the data rate. No information in the signal is lost by entropy coding. By entropy coding, 16 bits can typically be reduced to 13 or 14 bits. This is not a huge improvement, although a useful one. Why not, then, use entropy coding as a matter of course, since it loses no quality?

There are other disadvantages. First, the data rate depends on the input signal. A very random signal, like full-amplitude white noise, has very little systematic pattern, so can hardly be reduced in data rate at all by entropy coding. Also, entropy coding systems optimised for specific common types of pattern in audio signals are liable to *increase* the data rate if they encounter a very uncommon type of audio signal. Thus entropy coding is virtually useless for applications like broadcasting and constant-speed tape or CD recording where the data rate must be fixed in advance.

Second, by removing all the systematic patterns in the signal, errors become harder to spot and conceal, so entropy coding can only be used if the transmission channel has very good error protection. The tiniest error can cause huge changes in the output signal. The trouble is that extremely good error protection requires the transmission of extra data, partly nullifying the advantages of entropy coding.

Apart from a very modest rate reduction in one version of the Compusonics system, I know of no commercial high quality audio data rate reduction system that relies mainly on entropy coding. All systems giving a useful reduction in bit rate introduce signal errors that, hopefully, are subjectively masked by the signal itself.

Just like noise reduction

There is a strong conceptual similarity between analogue noise reduction systems and digital data compression. Indeed, using an analogue noise reduction system around a digital channel with fewer bits (*eg* a Dolby *SR* noise reduction around a 12 bit channel) may be considered to be a system of digital data compression. However, the term 'digital data compression' is usually reserved nowadays for systems in which all the signal processing is done digitally although earlier hybrid systems of digital data rate reduction (such as the satellite transmission systems of Dolby and dbx) used digitally controlled analogue signal processing.

Behind the apparently very different terminologies and technologies, the similarities between analogue noise reduction and digital data compression are far greater than their differences.

Both types of system try to get a subjectively error- and noise-free signal from a channel that on its own would give a high noise level. Both are based on the same idea of reducing noise and error by increasing the signal level and 'spectral occupancy' (*ie* the range of frequencies present at a high level) of the signal so the channel is always fully modulated by the signal. The decoding that reverses the data compression or noise reduction encoding restores the original signal levels by pulling the boosted frequency components back down again, at the same time reducing the background noise level by a corresponding amount.

These principles are common to analogue noise reduction and to digital data reduction systems (other than entropy coding). The differences between the two lie in the different natures of the typical analogue channel (*eg* tape, FM broadcasting) the typical digital channel (*eg* digital tape, CD-I, ROM or hard disk storage) and, to a lesser extent, the different things that can be done most easily with analogue and digital circuitry.

The typical analogue channel suffers from an unpredictable degradation other than noise. The output of tape may fluctuate due to variations in tape coating thickness, the frequency and phase responses may have ripples and fluctuations that may vary according to the tape used, the tape machine, tape bias and head contamination. The tape medium also suffers from level- and frequency-dependent non-linear distortion and wow and flutter, as well as slight errors of tape speed. Any analogue noise reduction system must give reasonably good results in the presence of all these degradations. Additionally, when the noise reduction is applied, there is no way of knowing what the errors produced by the recording channel will be.

Digital predictability

With digital systems, on the other hand, provided error protection is doing its job (or if one is using a system such as ROM storage not subject to significant error) one can predict *exactly* at the time of coding what the errors caused by a limited number of bits will be (for example, by adding a decoder to the encoder and taking the difference of the output from the input).

This has two consequences. First, one need not design the data compression system to be subjectively tolerant of 'small' signal degradations – hopefully there will be none – which means that some of the design compromises necessary in analogue noise reduction are not necessary in digital. One can change the gain of a digital signal by 24dB between successive moments of a signal without the risk of getting the wrong gain, whereas with analogue signals one would risk getting huge gain errors for a short while. To avoid such mistracking in analogue noise reduction systems, it is necessary to make any gain changes fairly slow ones.

Second, one can predict at the time of encoding exactly what the ultimate error will be in the final decoded output due to the quantisation errors of using only a few bits. One can use this knowledge to modify the error to have minimum audibility by feeding the error information back into the coding process (see Fig 1). This process of feeding the coding error back into the coding process is very much the same idea as negative feedback in amplifiers to reduce distortion errors. The theory used is very similar.

These two features of digital data rate reduction mean that the 'noise reduction' achieved can be very much more powerful for a given number of bits than for analogue noise reduction round a channel with a similar signal-to-noise ratio. A 4 bit digital channel has a signal-to-noise ratio of around 24dB and a 4 bit digital data reduction system such as *Apt-X 100* can sound very listenable, whereas an analogue channel with a 24dB signal-to-noise ratio would sound pretty appalling, however sophisticated an analogue noise reduction system used.

Designer mistakes

Although in principle digital is capable of much better results than analogue noise reduction, it is in practice much easier for audibly bad design mistakes to be made in a digital data compression system if the designer is not very careful. This is due to the nature of digital signals and cheap digital signal processing. The potentially horrendous sound of 'quantisation noise in digital systems is, by now, familiar as is the fact that this can be turned into a nice-sounding 'analogue-type' noise by adding a carefully controlled noise' signal (dither) before quantising.

In digital data compression systems, after one has processed a signal to increase its level over a wide range of frequencies, at some stage one has to reduce its data rate to fit the limited data rate available in the channel used. In other words one has to quantise the signal to a fewer number of bits. This requantisation process can produce subjectively nasty side effects just like ordinary undithered quantisation. Even when some of the techniques described later are used to mask the quantisation error, it is still liable to produce subtly disturbing side effects. Possibly designers of data compression systems should investigate the use of dither when requantising the processed signal to reduce some of these potentially nasty effects. By a technique known as subtractive dither, whereby the dither noise signal added during encoding is subtracted again during decoding, it is possible to get the benefits of dither with relatively little noise increase. I know of no commercial data reduction system, however, that uses dither in the coding process.

In the absence of dither, it is still possible to improve subjective results by very careful design of the quantising process but this is still a poorly understood topic among designers, especially at the very low bit rates of some recent systems.

Signal errors

Although data compression systems vary widely, the general principle of all systems is to raise the signal to near peak level (either overall or in several separate frequency bands), so the signal-to-noise ratio is more or less constant the whole time, and then to take the signal level back down again during decoding, taking the noise down with it.

The effect of this process is that the noise level goes up and down with the signal, causing what is termed modulation noise. One can measure modulation noise by comparing how far (in dB) the error-signal level is below the wanted signal.

Modulation noise is already familiar with analogue noise reduction systems. Certain signals – notably piano – are exceptionally good at showing up modulation noise subjectively. dbx noise reduction used with poor tape channels (*eg* cassette tape) is well known often to produce audible modulation noise with some sounds, and personal sensitivity to this fault varies from acceptable to totally intolerable. Although no noise reduction system can totally eliminate modulation noise, they do differ markedly from each other in the degree to which they subjectively mask modulation noise.

Masking is a psychoacoustic phenomenon, whereby the presence of a low level sound in one frequency band is masked or hidden by a much higher level sound in another frequency band. In general (and with some important exceptions), low level errors in any frequency band are well masked by much higher level sounds in the same frequency band (which at mid frequencies can be $\frac{1}{3}$ - or $\frac{1}{4}$ -octave wide). The degree of masking reduces as the frequencies of the wanted high level signal and unwanted low level error get further apart. The worst case normally encountered for masking is a low frequency signal accompanied by a high frequency noise – in some cases, the noise has to be up to 100dB below the signal before it becomes inaudible! In other cases, where the frequencies of signal and noise are similar, a noise well under 40dB down can be completely masked by the signal.

The more advanced analogue noise reduction systems (such as *telcom* and Dolby) make extensive use of masking to reduce the audibility of modulation noise. They all make sure that high level, low frequency signals are not accompanied by a high level of high frequency noise. The multiband systems *{telcom,* Dolby *A* and Dolby *SR)* additionally control the precise relative levels of signal and noise in adjacent frequency bands.

The crudest digital data compression systems, like the BBC 14/10 bit NICAM system, the 16/12 bit DAT 'long-play' system and the 10/8 bit system used in Video 8 digital sound, are all wideband companding systems (analogous to systems like dbx) and so have relatively poor masking of noise by low frequency signals. As a result, such systems have to be designed to use a relatively large number of bits, with only a modest degree of data compression, if the modulation noise is not to become too audible. To get the most efficient noise reduction and data compression, more elaborate systems that take into account the masking properties of different frequencies and adapt to the instantaneous frequency content of a signal are necessary. There are several different ways of doing this.

ADPCM

One approach is to use a single-band system of increasing the level of signals but to vary the frequency response of the signal according to its frequency content. Thus, if a signal has very little treble, it is encoded with the treble boosted more than the bass. On decoding, the treble content is reduced back again, taking down the level of treble noise to a point where it is masked by the bass. This, of course, is the well-known principle behind Dolby *B* noise reduction. The audio data compression system used on CD-I (CD-Interactive) allows a choice of four different equalisations in encoding, which may be varied as the signal varies.

In digital systems, one can predict the exact noise error at the time of encoding (as in Fig 1), so one can try to cancel out the noise error by subtracting it from the input, *ie* by negative feedback. Because digital systems are sampled only at discrete moments of time, such feedback can only operate if the feedback signal is delayed at least one sample. Such feedback turns out to alter the frequency spectrum of the quantisation noise. In general, this frequency spectrum can be adjusted by putting a digital filter in the feedback path as shown in Fig 2. This 'noise shaping' process can shape the frequency spectrum of the noise so that it is masked as well as possible by the frequency spectrum of the signal, possibly by varying the noise shaping from moment-to-moment to match the signal's spectrum. The effect of the filtered 'error-feedback' system of Fig 2 is not to alter the spectrum of the signal at all but to alter the spectrum of the noise by, in effect, passing the noise through the filter shown in Fig 3.

Such noise shaping is not possible in analogue systems. With digital compression, one can tinker in encoding not only with the level and frequency response of the signal, but also with the frequency spectrum of the noise. Systems doing both are capable of a lower and better-masked noise than analogue noise reduction. A digital system using equalisation and noise shaping is termed a Differential Pulse Code Modulation (DPCM) system, for historical reasons we shall not go into here. Even if the equalisation is fixed for all signals (say at a 6dB/octave bass cut) such systems can give much better masking of noise by signals (by 20 or 30dB) than simple near-instantaneous companding systems like NICAM.

If the EQ and the noise shaping are made adaptive, *ie* to vary with the signal to improve masking further, the data compression system becomes known as Adaptive Differential Pulse Code Modulation (ADPCM). ADPCM was widely studied by engineers in the '60s and '70s. The data compression system on CD-I is an ADPCM system, albeit a crude one with only up to four different equalisations. The CD-I standard offers various 8 bit and 4 bit data compression options, the 4 bit options using more varieties of equalisation but having a higher modulation noise and poorer quality.

The strategy that gives the lowest objective amount of noise with ADPCM is to equalise the signal so its spectrum becomes white, and to shape the noise spectrum in such a manner that, after decoding, it becomes white. This is termed predictive coding because it attempts to predict the next sample of the signal from previous decoded samples and transmits a quantised version of the difference between the sample and its predicted value. Additional noise shaping beyond the white results of predictive coding, to maximise the subjective masking of noise by the signal, will give subjectively better results.

One special case of predictive coding is of particular interest. Many audio signals in speech and music (and in other cases such as machine noises) have periodic waveforms, *ie* waveforms that repeat over and over again almost exactly. If a coder is designed cleverly enough, it can use a period of the repetitive waveform to predict future periods. A predictive encoder of this type is equivalent to an ADPCM coder with an extremely elaborate equalisation and noise shaping, and has the advantage that it codes well a wide variety of commonly occurring signals that the ears are good at analysing critically.

Multiband systems

Although a well-designed ADPCM system with enough equalisation options (perhaps hundreds, or even a continuously variable family of equalisers, rather than the four of CD-I!) could obtain a near-optimal low level of modulation noise with good masking, most efforts to improve on crude ADPCM systems have involved splitting the audio into several frequency bands. Each band is data-compressed and quantised separately, and the bands are re-expanded and put back together again during decoding. This means that any noise produced because of the presence of a signal frequency will be fairly near that frequency and so will be well masked by it.

All the multiband systems I am aware of use a technique known as dynamic bit allocation between the bands. This means that if one frequency band has a lot more energy than another (as perceived by a listener), more of the available bits are allocated to quantising that band and less to the others. In this way, the noise behind the highest energy bands (which would otherwise be at quite a high level) is brought down in level, whereas the noise behind the low energy bands (which would be at a very low level indeed) is brought up a bit in exchange. This way, if the bit allocation is carefully done, the overall amount of noise can be substantially reduced. Bit allocation achieves a similar redistribution of noise energy with frequency to that achieved by noise shaping in wideband systems.

By dynamic bit allocation, the most energetic signal components are encoded with a higher relative accuracy, reflecting the fact that they are the most important parts of the signal.

Actually, there is nothing that dynamic bit allocation achieves in a multiband system that, in principle, cannot be achieved by dynamic equalisation and noise shaping in the ADPCM system. Both systems redistribute signal and noise energy between the different frequencies to achieve roughly similar results. One has a greater flexibility with ADPCM systems since one is not restricted to a fixed set of frequency bands with rigidly designed crossover frequencies. In particular, the multiband system has no simple method corresponding to predictive coding of periodic repetitive waveforms in the ADPCM case. It is not altogether clear to me why multiband dynamic bit allocation systems are being widely worked on in preference to ADPCM systems.

Commercial multiband systems

The *Apt-X 100* system, developed in Belfast, uses a combination of dynamic bit allocation with just four rather wide bands (not in themselves narrow enough to give effective masking) with ADPCM techniques within each band. In some ways, this gives the best of both worlds, since it allows predictive coding of repetitive waveforms within each band. However, *Apt-X 100,* to judge from the limited published information, does not permit the absolute maximum advantage to be obtained from masking on non-periodic waveforms.

A very different approach aimed at squeezing absolute maximum advantage from psychoacoustic masking, has been developed in Germany in association with the Eureka project. These systems are still under development and, according to reports, are continuing to improve dramatically with virtual CD results being reported at astonishingly low rates of as little as 1 bit/sample. The Eureka systems are based on dividing the audio signal into a large number of frequency bands (around 20 or 30), each typically around $\frac{1}{3}$-octave wide. Each band is quantised separately and the number of bits allocated to each band is chosen to maximise the masking of the resulting noise spectrum by the signal, by using very detailed models derived from psychoacoustic experiments on how different audio frequencies mask one another.

These systems are very extreme in that, if a particular frequency band of the signal itself is at a sufficiently low level to be (supposedly according to the models of psychoacoustic masking) completely masked by the rest of the signal, then that band is allocated 0 bits, *ie* completely gated out. The Eureka systems incorporate an adaptive multiband noise gate (using around 30 bands) to reduce the audio data rate. It is claimed that the effect of these noise gates is inaudible due to masking. I need rather a lot of convincing that this is the case, since simple psychoacoustic masking experiments on how sine wave frequencies or narrow bands of noise mask one another need not necessarily apply to complex signals having a high degree of mutual correlation, and conveying subtle cues about stereo positioning, distance, space, instrumental resonances and complex orchestrations of sound.

Subjectively, while the *Apt-X 100* system has more obvious modulation noise than early prototype Eureka systems, this audible modulation noise is far less disturbing (despite a rather 'grainy' sound) than the artefacts of the latter. To my ears the Eureka systems have a rather 'unstable' sound quality, especially in stereo, somewhat akin to the effects of slight gain mistaking and pumping in analogue noise reduction systems. Theoretical analysis of the behaviour of quantisers at very low bit rates (even at more than 0 bits!) shows that gain modulation effects are highly likely unless extraordinary design care is taken, especially if the quantiser is not accurately matched to the signal statistics. In analogue noise reduction systems, the effects of gain mistracking of less than 0.1dB can be highly audible as a loss of sense of depth, and some people have suggested that gain modulation much less than this (down to 0.001dB) might be audible.

Also, since these multiband systems do not allow full predictive coding of nearly repetitive waveforms, they are liable to produce more audible effects on such waveforms than properly designed ADPCM systems. My experience in developing a dynamic multiband ambisonic decoder in the ’70s showed that the ears seem to be exceptionally sensitive to modulation effects on signals having a narrow bandwidth (flute, cello, etc), the resulting effect sounding like a particular kind of gross non-linear distortion. Possibly because my ears are particularly tuned to this effect, I have noted similar 'narrowband' distortion effects on demonstrations of early multiband systems. Systems like *Apt-X 100* which incorporate predictive coding of repetitive waveforms such as narrowband signals, would be expected to be much better in this respect.

It cannot be denied that the multiband coding systems being developed in Germany are a remarkable technological feat, and as work proceeds, no doubt they will be improved further. Even if some of the faults mentioned remain, they will provide an extremely useful means of conveying acceptable signal quality at bit rates that would otherwise prevent audio from being conveyed at all. The main caution about these and all other audio data compression systems is that they should not be used totally uncritically and their performance should not be overclaimed. (Remember 'perfect sound forever' on early CDs?) This is the case in critical professional and state-of-the-art high quality applications.

Nothing like the input

One remarkable thing about all systems having a very low bit rate is that they sound much better than they measure! The output waveform, compared side-by-side with the input waveform on an oscilloscope, bears little resemblance to the input. It is well known that two signals can have very different waveforms and yet sound similar. For example, passing a signal through a simple all-pass network can totally mangle the shape of a square wave and yet have remarkably little audible effect.

Nevertheless, the alteration of the waveform does suggest that efficient bit rate reduction systems cannot be treated purely as a neutral transmission channel and a lot of questions need to be asked about their performance in the real world before they are used in any given application. For example, what happens to stereo effect? Stereo works through having precise amplitude and phase relationships between the two channels. If a separate bit rate reduction system is used for each of the two channels, will the stereo quality be degraded? and if so, to what degree? What happens to more subtle cues like sense of distance (on recordings that have it) or of space and ambience?

It is possible to design audio data compression systems specifically to preserve stereo relationships (and, done properly, this is not simply a question of 'ganging' the compression parameters of the two channels) but I am unaware of any true stereo compression system under development.

There is also the problem of timing cues. Both in hearing stereo and in unraveling the relationships between many musical lines in a complex orchestration, the ears make use of the precise timings of transients down to a fraction of a millisecond. All the more efficient data compression systems tend to blur or displace such timing in a signal-dependent fashion. The German multiband systems have involved a considerable amount of empirical work optimising 'temporal masking' – the degree to which error signals need to coincide in time with the wanted signal. If the error proceeds the wanted signal too much, it becomes highly audible and masking ceases to work. However, such timing displacements and errors may also have a more subtle disturbing effect on the ears' ability to sort out complex stereo signals.

Professional use

Enough of how audio data compression works. What uses do such systems have and what kind of operational problems might they cause? Even if such systems have problems, we have learnt to live with the problems of analogue noise reduction and in appropriate applications we might learn to live with the problems of digital data compression, too.

Whether or not a data compression system is adequate for mid-fi consumer use, professional users are much more demanding. A first problem is that of processing delay in the encoding and decoding process. Suppose that one has a wonderful system that gives good CD subjective quality at 2 bits per sample. For many applications, it would nevertheless be quite useless if it has a long delay before the decoded signal finally emerges. For example, if data compression is used to store samples in ROM or RAM in a keyboard or sampler, one cannot wait half a second before the sound starts. In fact, for musical purposes, delays of more than 4ms are certainly unacceptable, and delays of under 1ms are desirable. Otherwise, the timing and feel of the music are affected.

Unfortunately, the most powerful data compression systems involve significant processing delays. A delay of 50 or 100ms may not be too important in tape playback or broadcasting applications, they might even be acceptable in digital cart applications for spinning in commercials, but in applications where timing is critical, less powerful data compression systems having shorter delays have to be used, at least for the early portions of a sound sample.

Then there is the problem of the complexity of the signal processing used. The most powerful compression systems involve very complex processing, which will involve very expensive circuitry or chips unless they are produced in huge commercial volumes. Generally, simpler systems involve cheaper processing.

For some uses (satellite links between broadcasters) this cost is not particularly important but it is important for consumer use and for professionals who may require tens or hundreds of encoder/decoder systems (*eg* for a 48-track digital recorder).

And then there is another problem in professional applications. You have just spun into your mix a 200 sec sample, which had been data compressed to 2 bits to fit into the RAM of what would, at 16 bits, be a 25 sec sampler. Fine, except that in later post-production work, you might need to recompress the mix you did back down to 2 bits again. What happens to sounds after encoding and decoding several times? Does all the modulation noise that has been so cunningly masked remain masked? Do those ever-so-subtle side effects that you are reassured cannot be heard in subjective tests remain subtle? I would be suspicious of using data compression for serious professional use in broadcasting, sampling, hard disk storage/editing/mixing or for digital tape recording unless the results of encoding and decoding (say) 10 times in succession are still highly acceptable. Moreover, this acceptability must still hold even if the signal is subjected to normal post-production operations like editing, gain changes, adding effects and mixing with other sounds, at intermediate stages.

Uses

Despite all these problems, which professional users will have to be aware of, it is likely that data compression will become an increasing part of the audio technology we all use. It is interesting to speculate about the kind of products a successful and economical bit rate reduction system would make possible.

One could envisage a suitably packaged collection of eight encoding and decoding systems for compressing 16 bit audio channels into 4 bits, and of putting the compressed channels into a conventional 16 bit stereo signal format, as a 'black box' for converting a stereo DAT recorder into an 8-channel recorder. Such a box would also need to incorporate eight A/D and D/A converters. Although such a unit would only give simultaneous recording of all eight channels at the same time, if it also incorporated means to add additional 4-bit channels to information containing less than eight channels, it could be used with two DAT machines to provide full 8-track recording facilities. Quality losses due to data compression could be minimised by using more than 4 bits/channel if less than eight tracks were used.

Such a unit would also allow other stereo digital media to be converted to (say) 8-track at relatively low cost. For example, one could send out library music on data-compressed CD in 8-track format, permitting the final mix to be optimised by the end user for his/her specific program use – although it would be wise to choose levels in the eight channels such that a straight equal-level mix should give the preferred standard mix for cases where the time is not available for detailed post-production work.

Similarly, the number of channels on hard disk media could be increased greatly. This would increase storage time and allow more rapid writing and reading of the hard disk (due to the lower bit rate) and more rapid loading and unloading from the hard disk system to and from tape. For the same reason, the transfer of samples via MIDI exclusive systems, which is normally very slow due to the low data rate of MIDI, could be speeded up.

Obvious applications of data compression would include terrestrial or satellite digital broadcasting using modest bandwidths and extra-long play CD or DAT for music, muzak and talking-book type applications or for low-cost archival purposes. Data compression also makes more likely the long-discussed idea of being able to access music from a central library anywhere in the world via digital phone link, since the music could be accessed at a reasonable rate via a modest capacity digital channel. Setting up links between studios in different parts of the world when artists are unable to travel to a session also becomes more economically viable without spending a fortune on the satellite link. A standard 56 kbit/s or 64 kbit/s link normally used for telephony might prove adequate for near-CD quality mono channels. Even if the quality of such a link is not up to the most critical studio standard, it would be good enough for preliminary production decisions to be taken and, providing a means of sync'ing is available, an uncompressed digital tape could be sent by mail or courier later for syncing up during post-production.

Providing its quality is good enough, data compression also makes practical methods of production hitherto ruled out by the lack of tape channels. For example, most multitrack work today is still multi-mono, mixing together say 24 or 48 monophonic tracks. It has long been known that the results could be a lot better if each of the 'tracks' were stereo, or even 4-channel B-format ambisonics, but this doubles or quadruples the required number of tape tracks, turning a 24-track machine into a 12- or even 6-track machine. However, if each track is fitted with a stereo or 4-channel data compression/expansion system, optimised to work well on stereo or B-format material, then each tape track could be allocated a stereo or ambisonic signal at no extra cost. This would mean using mixers with purpose-designed stereo or ambisonic 'channels' for best results, or else using very large mixers, but for the first time, data compression might make the use of multi-stereo production, with all its known advantages in terms of 'feel' and quality of stereoism, feasible.

Again, if data compression can be used to reduce the storage requirements per unit time in samplers and hard disk systems, it will become much more economic to incorporate sampling and spin-in facilities as parts of other studio equipment – perhaps the day is not far off when every mixer channel incorporates its own sampler? At this point, the boundary between tape recording, sampling, editing and mixing will start to get very blurred (as it has already on top-end mixer/hard disk systems), and product definitions and packaging will have to be re-evaluated.

All this, of course, presupposes two things: that the quality of data compressed audio can be upgraded to the highest professional audio standards, and that the processing chips can be made in sufficiently large quantities to bring down unit costs to a low level. The latter will be most likely if the same chips are used for both domestic and professional use, possibly with internal switching to different grades and quality levels of data compression to cope with different applications. Providing the quality and operational problems of audio data compression can be solved, its future looks assured.

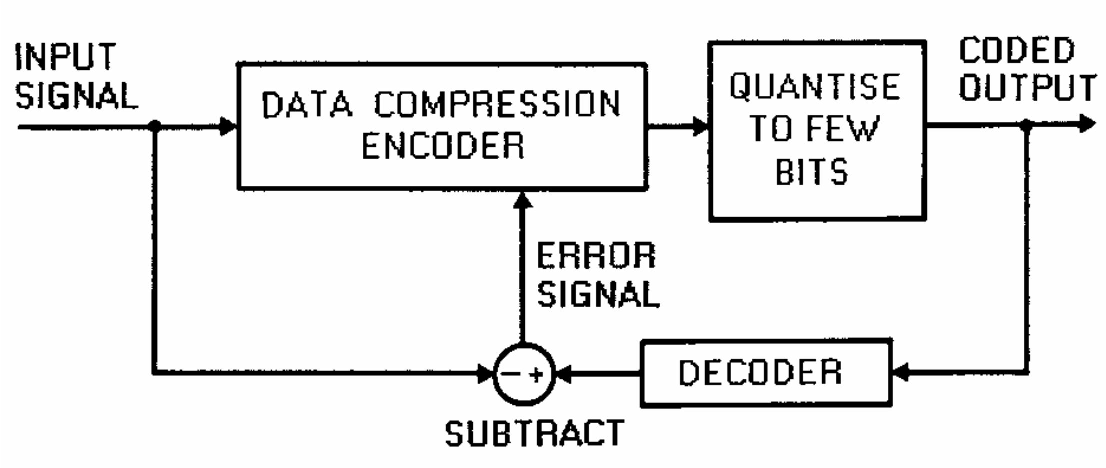

Fig 1: Feeding back the error due to coding into the encoding process to improve quality

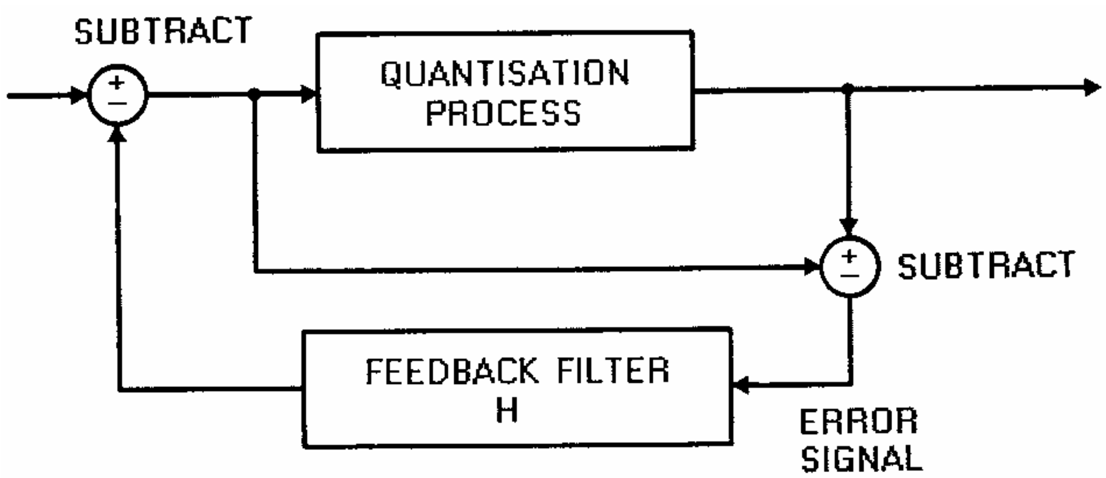

Fig 2: Noise shaping the quantisation error by negative feedback of the error via a filter

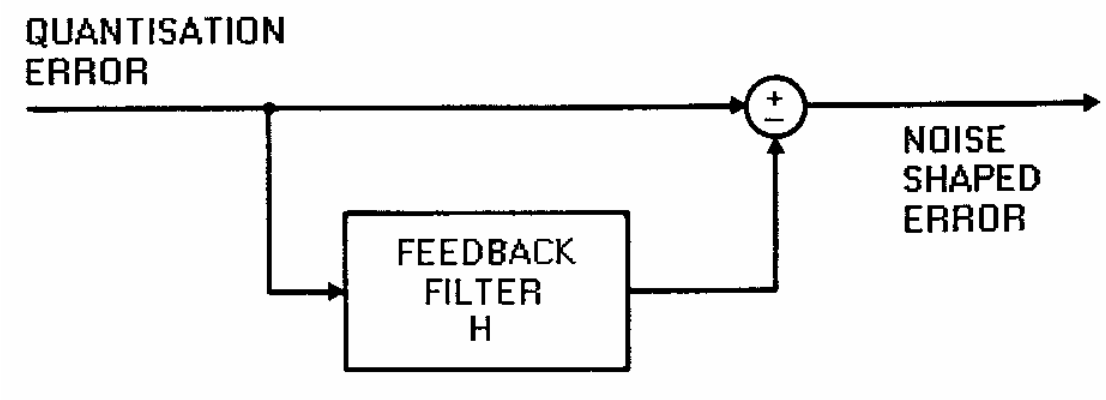

Fig 3: The effect of Fig 2 is to filter the quantisation error signal as shown here