Proceedings of the Institute of Acoustics

DIGITAL RECORDING TECHNIQUES IN ACOUSTICS

John Watkinson

Consultant

INTRODUCTION

One of the most challenging aspects of acoustics is the extremely wide range of applications which it embraces. Sound is used for entertainment, dramatic effect, communication, warnings and echo location and we prize these applications highly and research ways of making them more effective. Unwanted sound is also a by-product of industry, means of transport or simply bad plumbing and we then strive in the opposite direction to eliminate it or the unsuppressed vibrations which often cause it. There is possibly no other discipline in which as much effort is devoted to eliminating the product as to creating it.

By definition sound is a transient phenomenon, and except in rare cases such as airframe fatigue, sound does not leave any evidence. Our understanding of acoustics was enhanced dramatically by the development of means to record sound and vibration waveforms so that they could be repeated at will for study. Very few branches of acoustics have escaped the requirement to record such waveforms and as our analysis techniques have become more sophisticated this has placed greater demands on the quality of reproduction of our recording technology.

Whilst modern analog recording equipment may look sleeker than its ancestors, the principles employed remain the same, but it is now a mature technology. All of the great breakthroughs have been made, and the state of the art advances ever more slowly following a law of diminishing returns.

The main weakness of analog recording is that within the allowable bandwidth, any waveform is valid. If the speed of the medium is not constant, one valid waveform is changed into another valid waveform; a timebase error cannot be detected in an analog system. In addition, a voltage error simply changes one valid voltage into another; noise cannot be detected in an analog system. It is a characteristic of analog systems that degradations cannot be separated from the original signal, so nothing can be done about them. Conventional biased analog recording has poor performance at low frequencies, and frequency modulation (FM) was developed to allow a response down to DC. However, the speed instabilities of the tape are demodulated and restrict the signal to noise ratio achievable.

WHAT IS DIGITAL RECORDING?

For the digital recording of analog waveforms there is one system, known as Pulse Code Modulation (PCM) which is in virtually universal use. Figure 1 shows how PCM works. The time axis is represented in a discrete, or stepwise manner and the waveform is carried by measurement at regular intervals. This process is called sampling and the frequency with which samples are taken is called the sampling rate or sampling frequency Fs. The sampling rate is generally fixed and every effort is made to rid the sampling clock of jitter so that every sample will be made at an exactly even time step. If there is any subsequent timebase error, the instants at which samples arrive will be changed but the effect can be eliminated by storing the samples temporarily in a memory and reading them out using a stable, locally generated clock. This process is called timebase correction and all properly engineered digital systems must use it. Clearly timebase error is not reduced; it is totally eliminated. As a result there is little point measuring the wow and flutter of a digital recorder; it doesn't have any.

All analog signal sources from microphones, hydrophones, accelerometers and so on have a frequency response limit, as indeed do our ears. When a signal has finite bandwidth, the rate at which it can change is limited, and the way in which it changes becomes predictable. When a waveform can only change between samples in one way, it is then only necessary to carry the samples and the original waveform can be perfectly reconstructed from them.

Figure 1 also shows that each sample is also discrete, or represented in a stepwise manner. The length of the sample, which will be proportional to the voltage of the input waveform, is represented by a whole number. This process is known as quantizing and results in an approximation, but the size of the error can be controlled until it is negligible. The advantage of using whole numbers is that they are not prone to drift. If a whole number can be carried from one place to another without numerical error, it has not changed at all. By describing instrumentation waveforms numerically, the original information has been expressed in a way which is better able to resist unwanted changes.

Essentially, digital recording carries the original waveform numerically. The number of the sample is an analog of time, and the magnitude of the sample is an analog of the signal voltage. As both axes of the digitally represented waveform are discrete, the waveform can be accurately restored from numbers as if it were being drawn on graph paper. If we require greater accuracy, we simply choose paper with smaller squares. Clearly more numbers are then required and each one could change over a larger range.

THE ADVANTAGES OF DIGITAL RECORDING

There are two key advantages offered by digital instrumentation recording, but it is not possible to say which is the most important, as it will depend on one's standpoint:

a) The quality of reproduction of a well engineered digital recording system is independent of the medium and depends only on the quality of the conversion processes.

b) The conversion of waveforms to the digital domain allows tremendous opportunities which were denied to analog signals.

Someone who is only interested in technical quality will judge the former the most relevant. If good quality convertors can be obtained, all of the shortcomings of analog recording can be eliminated to great advantage. One's greatest effort is expended in the design of convertors, whereas those parts of the system which handle data need only be workmanlike. Wow, flutter, particulate noise, print-through, dropouts, modulation noise, HF squashing, azimuth error, interchannel phase errors are all eliminated. When a digital recording is copied, the same numbers appear on the copy: it is not a dub, it is a clone. If the copy is indistinguishable from the original, there has been no generation loss. Digital recordings can be copied indefinitely without loss of quality.

Once a waveform is expressed in the digital domain, it becomes data, and as such is indistinguishable from any other type of data. Systems and techniques developed in other industries for other purposes can be used for audio. Computer equipment is available at low cost because the volume of production is far greater than that of instrumentation recorders. Disk drives and memories developed for computers can be put to use in new applications. A word processor adapted to handle samples of a sound or vibration waveform becomes a workstation. There seems to be little point in waiting for a tape to wind when a disk head can access data in milliseconds. The difficulty of locating a significant event by spooling tape is eliminated when it can be located by viewing the waveform on a screen or, in the case of audio, by listening at any speed to samples retrieved from a memory.

USES OF RAM

Figure 2 shows the outline of a RAM (Random Access Memory) recorder. What the device does is determined by the way in which the RAM address is controlled. If the RAM address increases by one every time a sample from the ADC is stored in the RAM, a recording can be made for a short period until the RAM is full. The recording can be played back by repeating the address sequence at the same clock rate but reading the memory into the DAC. The result is generally called a sampler. By running the replay clock at various rates, the pitch and duration of the reproduced sound can be altered. Samplers will be restricted to a fairly short playing time, although this can be extended using data reduction where appropriate. Certain oscilloscopes have waveform samplers incorporated which allows a conventional CRT to act like a storage oscilloscope.

If the RAM is used in a different way, it can be written and read at the same time. The device then becomes a delay. Controlling the relationship between the addresses then changes the delay. The addresses are generated by counters which overflow to zero after they have reached a maximum count. As a result the memory space appears to be circular as shown in Figure 3. The read and write addresses are driven by a common clock and chase one another around the circle. If the read address follows close behind the write address, the delay is short. If it just stays ahead of the write address, the maximum delay is reached. Programmable delays are useful in auditoria to align the sound from various loudspeakers.

When samples are converted, the ADC must run at a constant clock rate and it outputs an unbroken stream of samples. Time compression allows the sample stream to be broken into blocks for convenient handling. Figure 4 shows an ADC feeding a pair of RAMS When one is being written by the ADC, the other can be read, and vice-versa. As soon as the first RAM is full, the ADC output switched to the input of the other RAM so that there is no loss of samples. The first RAM can then be read at a higher clock rate than the sampling rate. As a result the RAM is read in less time than it took to write it, and the output from the system then pauses until the second RAM is full. The samples are now time compressed. Instead of being an unbroken stream which is difficult to handle, the samples are now arranged in blocks with convenient pauses in between them. In these pauses numerous processes can take place. A rotary head recorder might switch heads; a hard disk might move to another track. In all types of recording, the time compression of the signal allows space for synchronising patterns, subcode and error correction words to be recorded.

Subsequently, any time compression can be reversed by time expansion. Samples are written intermittently into a RAM at the incoming clock rate, but read out continuously at the standard sampling rate. The time expansion stage can be combined with the timebase correction stage so that speed variations in the medium can be eliminated at the same time. The use of time compression is universal in digital instrumentation recording. In general the instantaneous data rate at the medium is not the same as the rate at the convertors, although clearly the average rate must be the same.

In practical recorders some RAM will always be present for the purposes of time compression and also to provide delay on replay to allow the error correction circuits time to operate. In logging recorders used to measure noise levels over extended periods, only the low-powered RAM need be active most of the time. When the RAM is nearly full of data, these can be transferred to a disk drive which is only powered for the duration of the transfer.

SYNCHRONISATION

It is frequently necessary for a digital recorder to be able to play back locked to an external sampling rate reference so that the data rate can be determined by the capabilities of the analysis system.

Figure 5 shows how this mechanism works. The timebase expansion is controlled by the external reference which becomes the read clock for the RAM and so determines the rate at which the RAM address changes. In the case of a digital tape deck, the write clock for the RAM would be proportional to the tape speed. If the tape is going too fast, the write address will catch up with the read address in the memory, whereas if the tape is going too slow the read address will catch up with the write address. The tape speed is controlled by subtracting the read address from the write address. The address difference is used to control the tape speed. Thus if the tape speed is too high, the memory will fill faster than it is being emptied, and the address difference will grow larger than normal. This slows down the tape.

In multitrack recorders, the various tracks can be synchronised to sample accuracy so that no timing or phase errors can exist between the tracks. This is particularly advantageous in the case of recordings from, for example, hydrophone arrays. Extra transports can be slaved to the first to the same degree of accuracy if more tracks are required. In stereo audio recorders image shift due to phase errors is eliminated.

In order to replay without a reference, perhaps to provide an analog output, a digital recorder generates a sampling clock locally by means of a crystal oscillator. Provision will be made on professional machines to switch between internal and external references.

ERROR CORRECTION AND CONCEALMENT

In a recording of binary data, a bit is either correct or wrong, with no intermediate stage. Small amounts of noise are rejected but, inevitably, infrequent noise impulses cause some individual bits to be in error. Dropouts cause a larger number of bits in one place to be in error. An error of this kind is called a burst error. Whatever the medium and whatever the nature of the mechanism responsible, data are either recovered correctly, or suffer some combination of bit errors and burst errors. The severity of a bit error depends upon which bit of the sample is involved. If the LSB of one sample was in error in a high amplitude waveform, the effect would be totally masked and of no consequence. Conversely, if the MSB of one sample was in error in a quiet passage, the resulting transient would be unacceptable. Clearly a means is needed to correct errors from the recording medium.

In binary, a bit has only two states. If it is wrong, it is only necessary to reverse the state and it must be right. Thus the correction process is trivial and perfect. The main difficulty is in identifying the bits which are in error. This is done by coding the data by adding redundant bits. Adding redundancy is not confined to digital technology, airliners have several engines and cars have twin braking systems. Clearly the more failures which have to be handled, the more redundancy is needed. If a four engined airliner is designed to fly normally with one engine failed, three of the engines have enough power to reach cruise speed, and the fourth one is redundant. The amount of redundancy is equal to the amount of failure which can be handled. In the case of the failure of two engines, the plane can still fly, but it must slow down; this is graceful degradation. Clearly the chances of a two-engine failure on the same flight are remote.

The amount of error which can be corrected is proportional to the amount of redundancy. Within this limit, the samples are returned to exactly their original value. Consequently corrected samples are inaudible. If the amount of error exceeds the amount of redundancy, correction is not possible, and, in order to allow graceful degradation, concealment will be used. Concealment is a process where the value of a missing sample is estimated from those nearby. The estimated sample value is not necessarily exactly the same as the original, and so under some circumstances concealment can be audible, especially if it is frequent. However, in a well designed system, concealments occur with negligible frequency unless there is an actual fault or problem. Concealment is made possible by re-arranging or shuffling the sample sequence prior to recording. This is shown in Figure 6 where odd-numbered samples are separated from even-numbered samples prior to recording. The odd and even sets of samples may be recorded in different places, so that an uncorrectable burst error only affects one set. On replay, the samples are recombined into their natural sequence, and the error is now split up so that it results in every other sample being lost. The waveform is now described half as often, but can still be reproduced with some loss of accuracy. This is better than not being reproduced at all even if it is not perfect. Almost all digital waveform recorders use such an odd/even shuffle for concealment. Clearly if any errors are fully correctable, the shuffle is a waste of time; it is only needed if correction is not possible.

In high density recorders, more data are lost in a given sized dropout. Adding redundancy equal to the size of a dropout to every code is inefficient. Figure 7 shows that the efficiency of the system can be raised using interleaving. Sequential samples from the ADC are assembled into codes, but these are not recorded in their natural sequence. A number of sequential codes are assembled along rows in a memory. When the memory is full, it is copied to the medium by reading down columns. On replay, the samples need to be de-interleaved to return them to their natural sequence. This is done by writing samples from tape into a memory in columns, and when it is full, the memory is read in rows. Samples read from the memory are now in their original sequence so there is no effect on the recording. However, if a burst error occurs on the medium, it will damage sequential samples in a vertical direction in the de-interleave memory. When the memory is read, a single large error is broken down into a number of small errors whose size is exactly equal to the correcting power of the codes and the correction is performed with maximum efficiency.

The interleave, de-interleave, time compression and timebase correction processes cause delay and this is evident in the time taken before audio emerges after starting a digital machine.

The presence of an error correction system means that the audio quality is independent of the recording medium/head quality within limits. There is no point in trying to assess the health of a machine by listening to it, as this will not reveal whether the error rate is normal or within a whisker of failure. The only useful procedure is to monitor the frequency with which errors are being corrected, and to compare it with normal figures. Professional digital recording equipment should have an error rate display.

CHANNEL CODING

In most recorders used for storing digital information, the medium carries a track which reproduces a single waveform. Clearly data words representing audio samples contain many bits and so they have to be recorded serially, a bit at a time. Some media, such as hard disks only have one track active at a time, so it must be totally self contained. Other media have many parallel tracks. At high recording densities, physical tolerances cause phase shifts, or timing errors, between parallel tracks and so each track must still be self contained until the replayed signal has been timebase corrected.

Recording data serially is not as simple as connecting the serial output of a shift register to the head. In the two's complement scheme used for bipolar waveform coding, a common sample value is all zeros, as this corresponds to silence. If a shift register is loaded with all zeros and shifted out serially, the output stays at a constant low level, and constant magnetisation is recorded on the track. On replay there is nothing to indicate how many zeros were present, or even how fast to move the medium. Clearly serialised raw data cannot be recorded directly, they have to be modulated in to a waveform which contains an embedded clock irrespective of the values of the bits in the samples. On replay a circuit called a data separator can lock to the embedded clock and use it to separate strings of identical bits.

The process of modulating serial data to make it self-clocking is called channel coding. Channel coding also shapes the spectrum of the serialised waveform to make it more efficient. With a good channel code, more data can be stored on a given medium. Spectrum shaping is used in DAT to allow re-recording without erase heads.

HARD DISK RECORDERS

The hard disk recorder stores data on concentric tracks which it accesses by moving the head radially. Clearly while the head is moving it cannot transfer data. Using ADCs and RAM for time compression, a hard disk drive can be made into a waveform recorder.

Figure 8 shows the principle. The instantaneous data rate of the disk drive is far in excess of the sampling rate at the convertor, and so a large time compression factor can be used. The disk drive can read a block of data from disk, and place it in the timebase corrector in a fraction of the real time it represents in the audio waveform. As the timebase corrector steadily advances through the memory, the disk drive has time to move the heads to another track before the memory runs out of data. When there is sufficient space in the memory for another block, the drive is commanded to read, and fills up the space. Although the data transfer at the medium is highly discontinuous, the buffer memory provides an unbroken stream of samples to the DAC and so a continuous output waveform is obtained.

Recording is performed by using the memory to assemble samples until the contents of one disk block is available. This is then transferred to disk at high data rate. The drive can then reposition the head before the next block is available in memory.

An advantage of hard disks is that access to the recorded material is much quicker than with tape, as all of the data are available within the time taken to move the head.

ROTARY HEAD DIGITAL RECORDERS

The rotary head recorder borrows technology from videorecorders. Rotary heads have extremely high packing density: the number of data bits which can be recorded in a given space. In a digital recorder packing density directly translates into the playing time available for a given size of the medium.

In a rotary head recorder, the heads are mounted in a revolving drum and the tape is wrapped around the surface of the drum in a helix as can be seen in Figure 9. The helical tape path results in the heads traversing the tape in a series of diagonal or slanting tracks. The space between the tracks is controlled not by head design but by the speed of the tape and in modern recorders this space is reduced to zero with corresponding improvement in packing density.

The added complexity of the rotating heads and the circuitry necessary to control them is offset by the improvement in density. The discontinuous tracks of the rotary head recorder are naturally compatible with time compressed data. As Figure 9 illustrates, the audio samples are time compressed into blocks each of which can be contained in one slant track.

In a machine such as DAT (Rotary-head Digital Audio Tape) there are two heads mounted on opposite sides of the drum. One rotation of the drum lays down two tracks. Effective concealment can be had by recording odd numbered samples on one track of the pair and even numbered samples on the other.

A rotary head recorder contains the same basic steps as any digital audio recorder. The record side needs ADCs, time compression, the addition of redundancy for error correction, and channel coding. On replay the channel coding is reversed by the data separator, errors are broken up by the de-interleave process and corrected or concealed, and the time compression and any fluctuations from the transport are removed by timebase correction. The corrected, time stable, samples are then fed to the DAC.

Digital recording has considerable advantages in a large number of acoustics related applications and it naturally complements the widespread use of signal analysis in computers. With the exception of some specialist applications, digital recording will eventually become universal.

REFERENCES

Watkinson, J R , The Art of Digital Audio (1994: Focal Press ISBN 0240 51320 7) Watkinson, J R , The Art of Data Recording (1994: Focal Press ISBN 0240 51309 6) Watkinson, J R , RDAT (1991: Focal Press ISBN 0240 51306 1)

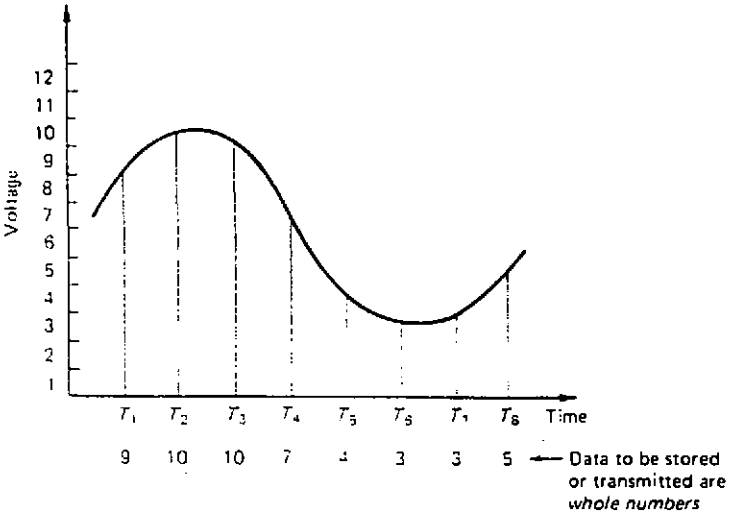

Fig. 1. In pulse code modulation (PCM) the analog waveform is measured periodically at the sampling rate. The voltage (represented here by the height) of each sample is then described by a whole number. The whole numbers are stored or transmitted rather than the waveform itself.

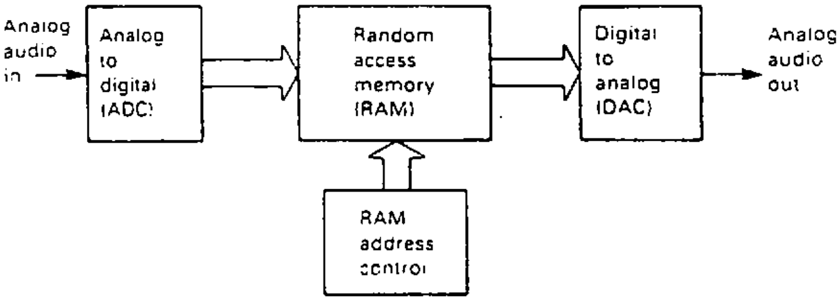

Fig. 2. In the digital sampler, the recording medium is a random access memory (RAM). Recording time available is short compared with other media, but access to the recording is immediate and flexible as it is controlled by addressing the RAM.

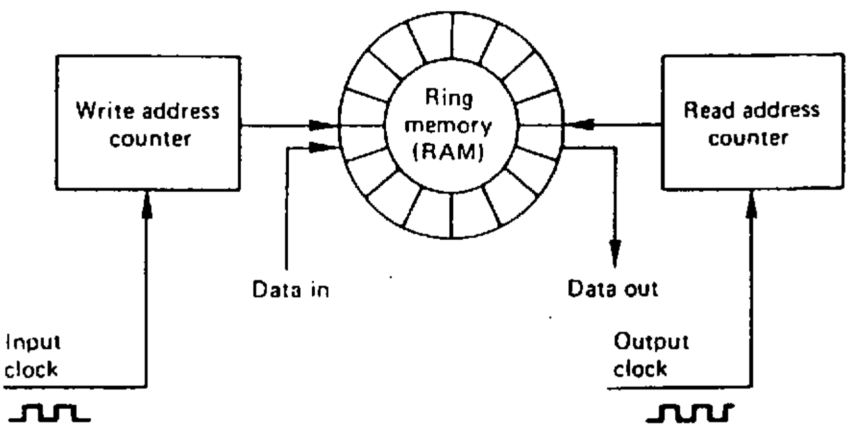

Fig. 3. If the memory address is arranged to come from a counter which overflows the memory can be made to appear circular. The write address then rotates endlessly, overwriting previous data once per revolution. The read address can follow the write address by a variable distance (not exceeding one revolution) and so a variable delay takes place between reading and writing.

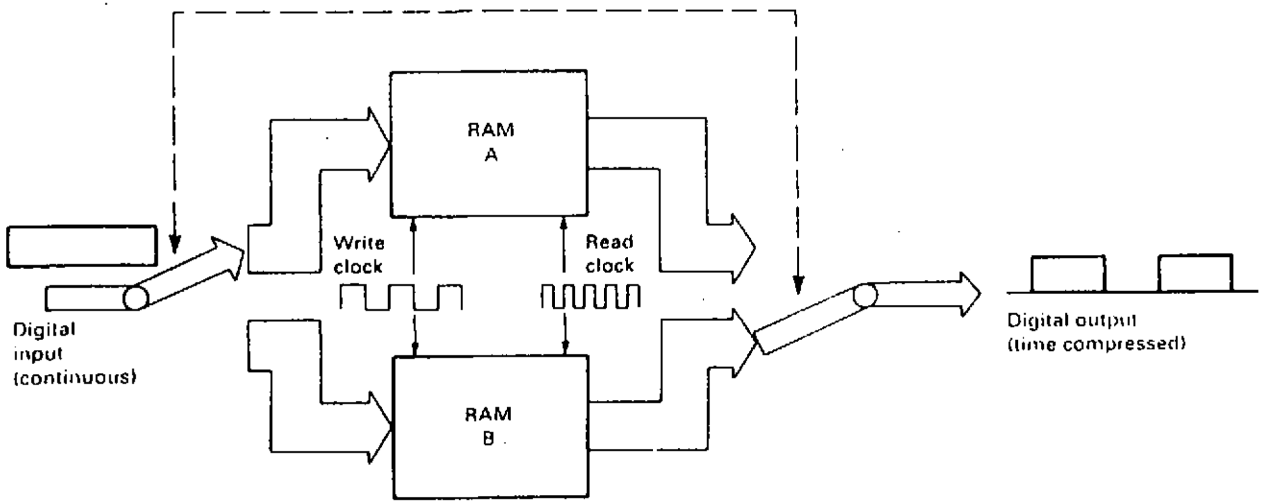

Fig. 4. In time compression, the unbroken real-time stream of samples from an ADC is broken up into discrete blocks. This is accomplished by the configuration shown here. Samples are written into one RAM at the sampling rate by the write clock. When the first RAM is full, the switches change over, and writing continues into the second RAM whilst the first is read using a higher-frequency clock. The RAM is read faster than it was written and so all the data will be output before the other RAM is full. This opens spaces in the data flow which are used as described in the text.

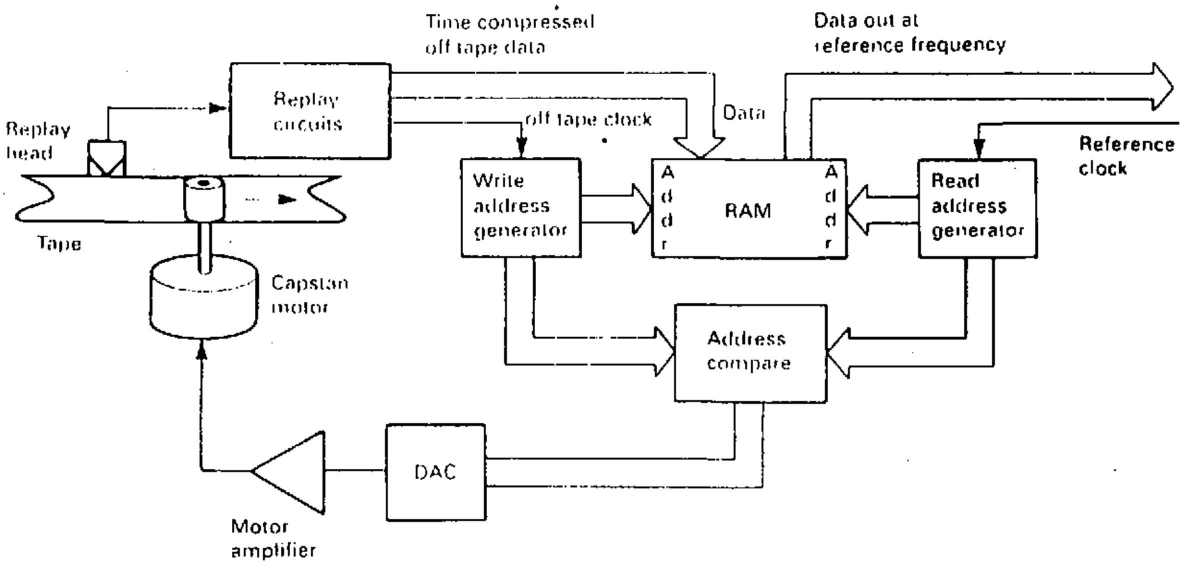

Fig. 5. In a recorder using time compression, the samples can be returned to a continuous stream using RAM as a timebase corrector (TBC). The long-term data rate has to be the same on the input and output of the TBC or it will lose data. This is accomplished by comparing the read and write addresses and using the difference to control the tape speed. In this way the tape speed will automatically adjust to provide data as fast as the reference clock takes it from the TBC.

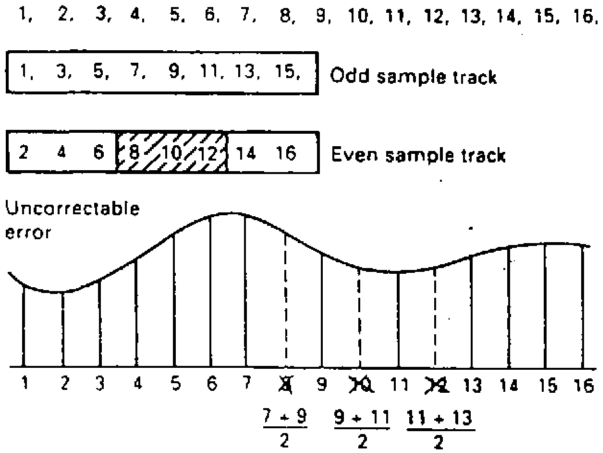

Fig. 6. In cases where the error correction is inadequate, concealment can be used provided that the samples have been ordered appropriately in the recording. Odd and even samples are recorded in different places as shown here. As a result an uncorrectable error causes incorrect samples to occur singly between correct samples. In the example shown, sample 8 is incorrect, but samples 7 and 9 are unaffected and an approximation to the value of sample 8 can be had by taking the average value of the two. This interpolated value is substituted for the incorrect value.

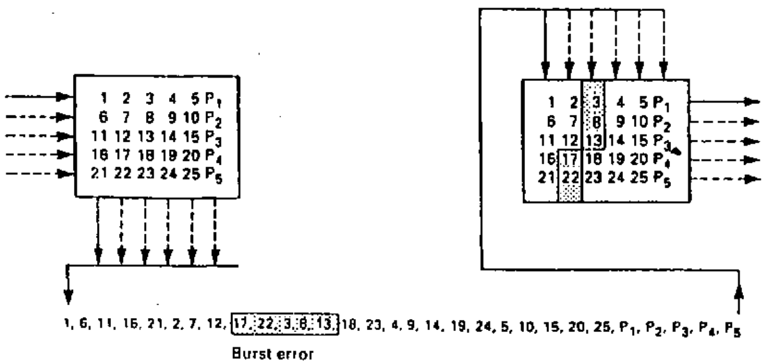

Fig. 7. Interleaving is essential to make error-correction schemes more efficient. Samples written sequentially in rows into a memory have redundancy P added to each row. The memory is then read in columns and the data sent to the recording medium. On replay the non-sequential samples from the medium are de-interleaved to return them to their normal sequence. This breaks up the burst error (shaded) into one error symbol per row in the memory, which can be corrected by the redundancy P.

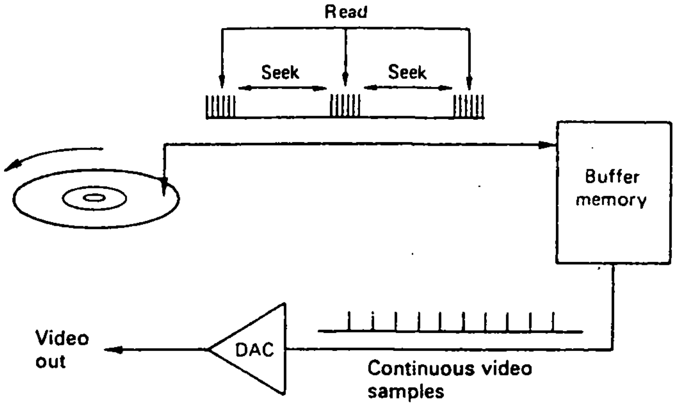

Fig. 8. In a hard-disk recorder, a large-capacity memory is used as a buffer or timebase corrector between the convertors and the disk. The memory allows the converters to run constantly despite the interruptions in disk transfer caused by the head moving between tracks.

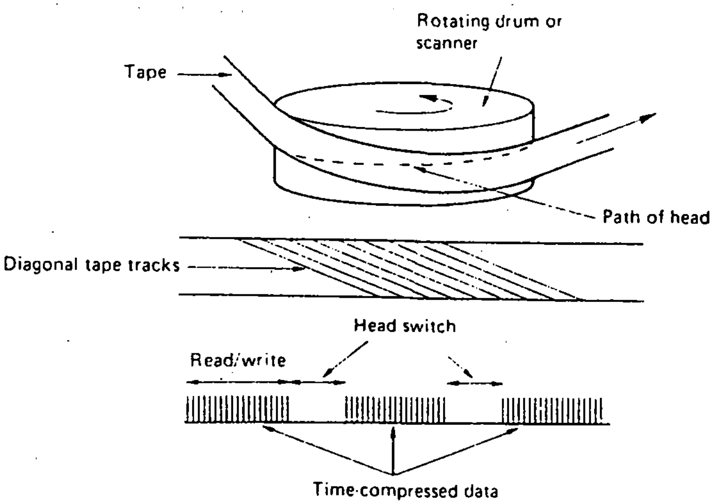

Fig. 9. In a rotary-head recorder, the helical tape path around a rotating head results in a series of diagonal or slanting tracks across the tape. Time compression is used to create gaps in the recorded data which coincide with the switching between tracks.