Proceedings of The Institute of Acoustics

A DIGITAL SOUND STORE SYSTEM

Stephen P. Jones and Peter Smith

Electrosonic Ltd, 815, Woolwich Road, London. SE7 8LT.

INTRODUCTION

Until now the accepted methods of recording sound have always involved moving parts - a revolving disç, a magnetic tape. Now it is possible to think in terms of "solid state" sound recording, where sound can be recorded directly into EPROMS'S, resulting in a playback system that has no moving parts.

The technical realisation is now possible due to considerable densities of EPROM'S that are now available at a low cost. For example, it is currently possible to buy an EPROM that can hold 256000 bits of data for under £10.00.

The product was developed in order to give sound effects to a series of "cars" at a new ride at the National Motor Museum at Beaulieu. This application had specific requirements of multiple and delayed outputs and did not require a particularly great frequency response. However, the development programme allowed for considerable expansion of memory which could either be utilised for extended frequency response or increased record/play time or a mixture of the two.

It is envisaged that further development would include stereo and ROM "carts" for use in jingle or public announcement applications.

SYNTHESISED OR NATURAL SOUND REPRODUCTION

Synthesised sound reproduction relies on instructions obtained from memory to make up the various sound elements which in speech terms are known as "phonems". When a particular word or message is reproduced, the synthesiser or controller fetches and organises these sound elements to form words and sentences. The result, although intelligible, is unnatural and perhaps unfriendly. Examples of the use of this method are to be found in automobiles and educational toys such as "speak and spell".

Storing the Sound Signal

In order to maximise the audio storage efficiency, there are some well establised techniques available. Some of them were evaluated in the stage of development, but none of them were appropriate for the sound store. What was needed was a simple method for obtaining a percieved 10-bit a sample quality from an 8-bit wide data memory. By adopting a simple method, a microprocessor could be employed as a central audio processor/output controller for a number of audio channels. Figure 1 shows the recording and playback arrangements.

The digital compressor operates on the following principle:
During the recording process a decision is made whereby audio levels between 0 and - 8dB (high level) are stored with a coarse 7-bit resolution and below -18dB (low level) are stored with a fine 10-bit resolution. In each case there are 7 databits and 1 control bit stored, which amount to 1 byte of memory per audio sample. The control bit is used to position the 7 data bits to either the 7 MCB's for coarse resolution, or 7 LSB's for fine resolution at the memory input. In this way, there is abenefit of 6dB. (See figures 2 and 3)

Audio Memory Capacity

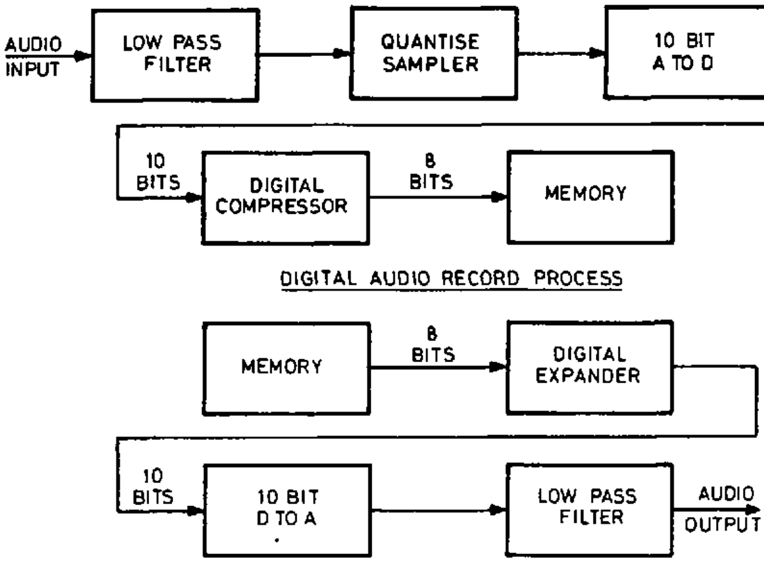
The amount of memory for a given audio message time is a function of the dynamic range and bandwidth. This can be shown to be a product of memory size/second = n bits/sample x Y samples/seconds.

In this system:
Memory size/second = 8 bits/sample x 8000 samples/seconds

= 64000 bits/seconds.

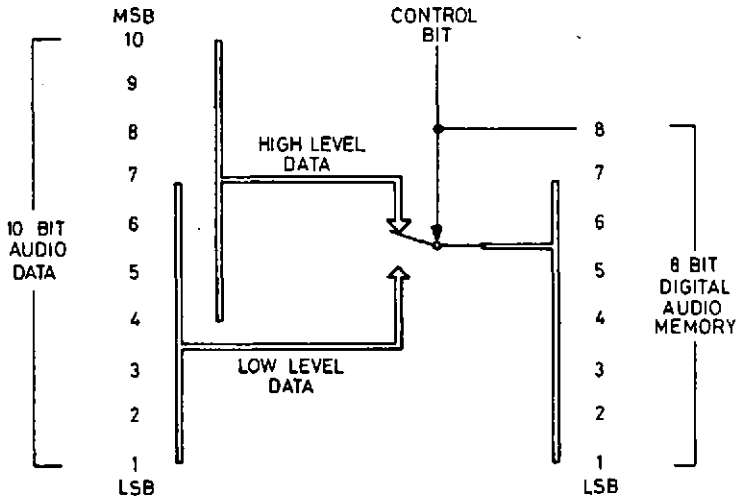
where 8000 samples/seconds is equivalent to a maximum audio bandwidth

The maximum possible audio message time is, there, a function of memory capacity which in this unit has been designed for up to 24 x 256K EPROM's giving a total message time of 96 seconds without memory expansion.



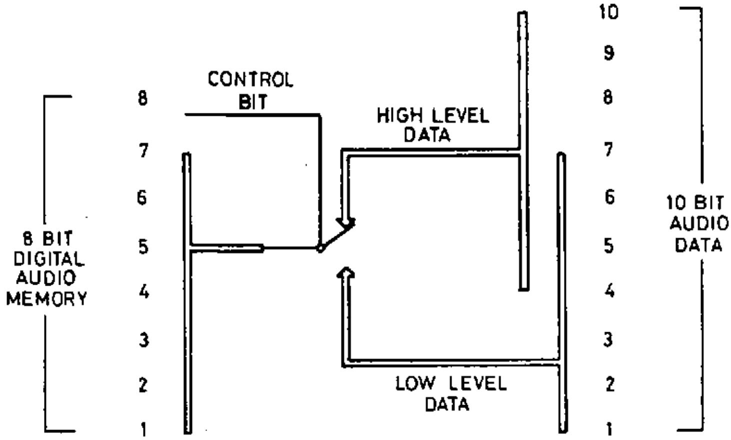
DIGITAL AUDIO PLAYBACK PROCESS

FIG. 1



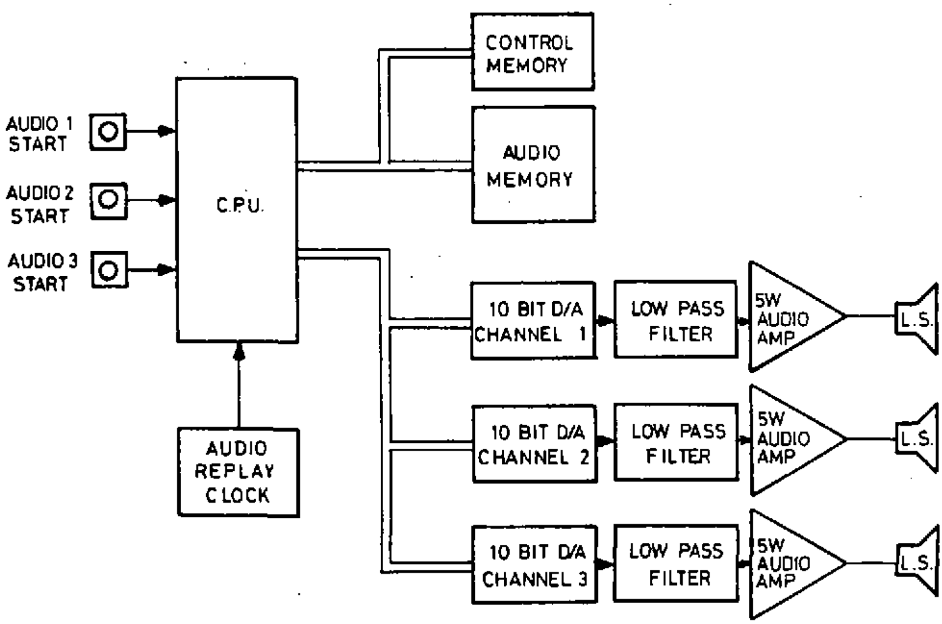
DIGITAL AUDIO RECORDING COMPRESSION

FIG. 2



DIGITAL AUDIO PLAYBACK EXPANSION

FIG. 3



DIGITAL AUDIO PLAYBACK FROM MEMORY

FIG. 4

A DIGITAL SOUND STORE SYSTEM

Processing Three Audio Output Channels

Figure 4 shows a Block Diagram of the 3-channel audio output playback process from memory.

The heart of the audio replay system is the Central processor Unit (CPU), which for this application employs a Z80 microprocessor. In order to simultaneously control and output 3 audio channels, an audio sample replay clock is needed which interrupts the CPU at the rate of the clock frequency. At every interrupt, the next sample of each audio channel is fetched from memory, processed and outputted to the respective D/A converters. As mentioned earlier, for a 4KHz bandwidth the sample output rate is 8KHz. For 3 audio channels, the CPU data process rate is defined as the product of the number of audio channels, the number of audio samples per second and the number of bits per sample.

This is an extremely fast data rate for an 8-bit CPU running at 6MHz and was only achieved by writing carefully optimised machine code.

One of the primary virtues of this method is that it allows the control of 3 audio channels independently, unlike a multitrack tape recorder where the tracks are effectively cogged.

One of the principle advantages of employing a CPU to process and control the audio outputs, is that many variations of the system are possible which allow individual customer requirements to be met simply by modifying the control software. Also there is "off peak" spare time available which can be effectively accumulated to do other relatively non time critical functions, such as controlling relay contacts. These auxiliary functions can be time or event programmed and included in the audio memory.

Memory Expansion

A memory expansion port is fitted which allows external memory cards to be connected. The expanded memory increases the audio message time potential to maximum of 480 seconds. This playing time can be divided in any required manner between the three outputs.

Random Access of Audio

A 32-way audio message select port is fitted which allows the random access of audio messages. During the audio programming stage, any of the audio channels can be allocated to one of 32 addresses. In this way, a number of messages can be stored for instant playback. Because the 3 audio channels are independent, it is possible to allocate a group of messages to each channel for random access on a parallel basis.

This random access facility provides a substantial economy over tape-based systems where one may have one or two common messages linked numerous variable messages. For example, in a supermarket where the check-out girls ask for someone to check a price, or bring some change, the message is currently given live into a microphone in the form "Customer Service 200 checkout 14". The intelligibility of the live system and its environmental acceptability depends largely on the check-out girl's ability. By recording the words "Customer Service", "checkout", and the series of numbers 100, 200... and also 1, 2...20 the CPU can be programmed to access the whole sentence according to the button depressed. The checkout girl has 9 buttons representing 100 to 900 calls and each checkout panel can be wired to a keyboard chip which will convert the commands into the binary form necessary for the input of the CPU.

Another natural extension of this principle is to precede all calls by a ding-dong chime or other entry tone.

Variable Audio Playback Speed Without Pitch Change

The facility of being able to vary the audio playback speed is particularly useful when synchronising a sound commentary to the speed of an entertainment ride. Such a requirement was needed at the National Motor Museum for their "Wheels" Ride. The variable audio playback speed is achieved by contracting, or expanding, the silence intervals that naturally occur between phrases or sentences of a speech commentary.

During the recording process, these silence intervals are detected, measured and stored as separate quantities in the audio memory for subsequent use by the playback store. The ride speed input port of the sound store measures the deviation from the standard ride speed and forms a multiplication silence factor which is used to alter the commentary. When the ride speed is nominal, the silence factor is unity resulting in no change in silence intervals, but when slow or fast, the silence factor is correspondingly increased or reduced. In this way, the pace of the audio commentary is matched to the speed of the ride, without altering the pitch.

Dumb Messages

A "message" does not need to be a message as such, it can be a command which is either kept within the store or sent outside. For example, if an emergency message is to be repeated, the CPU can be programmed to do so to a particular message until another message is accessed. Thus a silence or "dumb" message can be used to escape from the loop. Similarly, a dumb message can be sent to a non-audio port which simply closes a relay contact. The looping principle can be used to hold that contact rather than a latching circuit if required.

Applications for these techniques could include lifts (elevators) where the normal messages are giving level status and the loop message is for saying "the lift is overloaded" until the situation is corrected. Another application could be "Son et Luminaire" where the sound effects are issued with a dry loop contact for synchronously firing off lighting effects.

Conclusions

The digital sound store has certain advantages over conventional reproduced sound. The complete absence of moving parts goes a long way to providing a sound reproducer that is immune from sound deterioration with use or age and the concept of zero maintenance. This requirement is important for applications in museums, exhibitions and automatic announcement systems, where sound sources are needed on demand, or continuously, for up to 365 days a year.