COMPUTER SIMULATION FOR DIGITAL AUDIO SYSTEMS

1653 (K-4)

Bart N. Locanthi and Mitsuya Komamura
Pioneer North America Development Laboratory
Pasadena, CA

Presented at the 66th Convention 1980 May 6 through 9 Los Angeles

AES

AN AUDIO ENGINEERING SOCIETY PREPRINT

COMPUTER SIMULATION FOR DIGITAL AUDIO SYSTEMS

Bart N. Locanthi
Mitsuya Komamura
Pioneer North America Development Lab
Pasadena, California 91101

ABSTRACT

Computer simulation programs will be discussed for psychoacoustical evaluation of various error correction and/or concealment schemes which eventually fail, and bit reduction proposals. Sixteen bit linear recordings will be used to demonstrate some of our results.

INTRODUCTION

Although digital audio systems date back at least to the time when people used drum signals for communication over moderate distances, digital audio for high fidelity came into practical use in the early 1970's, many hundreds of years later.

At the present time many different digital mastering systems are used in recording companies in Europe, the USA and Japan. Most of these systems use 16 bit linear encoding at sampling frequencies from 40 kHz to 50 kHz. Collectively, these record companies now have well over 800 digital master tape albums. The principal benefits of digital audio to the consumer at this time appear in the form of recordings on standard phonograph records which have been mastered digitally. The remaining and perhaps most important benefits await an economically viable means available to the consumer for reproducing digital audio from software which is already in the digital format. The next item on the agenda is an economically practical device which will record and play digital audio.

A great deal of activity in Japan has caused about 10 manufacturers to deliver to the marketplace digital converters which produce compatible signals in the video format for 2 channel digital recording on home VTR's. A signal recorded on tape from any one encoder may be played through any other encoder. This is true of any encoder of recent manufacture which satisfies the EIAJ standards for such machines. One company which introduced its encoder to the USA supplies a digitally recorded album in a Betamax cassette with each encoder. These digital encoders vary in price in Japan from $2300 to $6000. All of them encode the data as 14 bits linear or equivalent at a sampling rate of 44.056 kHz per word.

The principal reason for the early appearance of digital encoders for home VTR's is their widespread availability and the fact that VTR's possess the necessary bandwidth for digital audio.

Other machines in the consumer marketplace which can handle the necessary bandwidth for digital audio are video disc players. Several formats and disc sizes are being considered. A few companies have demonstrated video disc players reproducing digital audio recordings with extremely high fidelity.

With the exception of the EIAJ standard for PCM audio encoders for home VTR's, no other standards exist for home or professional use machines with respect to sampling frequencies, error correction and/or concealment schemes, or bits/sample. Furthermore, we do not believe that enough psychoacoustical testing has been done to determine what the best and most economical means are for the required recording formats.

In all PCM audio systems, one bit is used to indicate the polarity of a sample and the remaining bits are used to indicate its magnitude. In a 16 bit system the accuracy of the magnitude is one part in 2^{15} (one part in 32,768). At a sampling rate of 50 kHz, the raw data rate for 2 channel stereo is 1.6 M bits per second.

Depending on many factors -

1. The recording bit density in the recording medium,
2. The noisiness of the recording medium,
3. Dropout characteristics, and
4. Electronic noise,

additional data must be added to the "raw" data to permit error correction and/or concealment. Some proposals have led to overheads up to 150%.

COMPUTER SIMULATION

In order to avoid proliferation of a multitude of hardware designs and to arrive quickly at an optimum solution, we decided to use a high speed digital computer for digital audio simulation.

We selected a DEC PDP 11/60 computer with a 178M byte disc pack for our laboratory because:

1. It permitted us to make real time digital audio recordings onto the disc at bit rates over 3M bits per second.

2. The disc capacity permitted 6 minutes of real time recording at 3M bits per second with enough reserve capacity for programs and other housekeeping functions.

3. A broad spectrum of peripheral devices easily interface with this machine.

After laboring for several months developing programs under the DEC RSX 11M operating system our staff recommended strongly that we change over to the UNIX operating system. This system was developed at the Bell Telephone Laboratories for use on DEC computers to facilitate research in digital signal processing. After the change to UNIX, a large number of programs were developed in a relatively short time as compared to previous performance under the original operating system.

PSYCHOACOUSTICAL TESTING

For listening tests, we have determined that the most critical signal source was a single frequency sine wave. The next most critical source was a sine wave plus harmonics with amplitudes decreasing as $\frac{1}{\sqrt{n}}$. The most critical program source was the solo piano. The least critical program source located thus far is the organ.

A listening test was designed to supply paired signals, each of 10 seconds duration, to listeners for their judgment as to whether the pairs were identical or different. A computer program was developed for random selection of each element of the pair. Twenty such randomly determined pairs were presented to listeners for their evaluation. At the end of each test, the computer printed out the actual pairs for grading the listeners.

Figure 1 shows the percentages of correct scores vs the number of bits used in floating point representations for bit reduction. In the floating point schemes tried, the accuracy of the magnitude was variable but the dynamic range was always the same (92dB). The four sources referenced above, sine wave, complex tone, piano and organ are indicated in the Figure.

The following demonstration through the sound system represents an example of the tests used for obtaining the data shown in Figure 1. The program sources were transferred to a Sony Umatic VTR recorder through a PCM 1600. Because of the limited time for the demonstration, these paired tests will always start with the original signal and end with the floating point comparisons. Three successive pairs for each program source and floating point scheme will be demonstrated. Our floating point representation m/n simply means a mantissa of (m-1) bits having 2^n levels. Figure 2 lists the program source and floating point representation for each test:

Test No.	Program Source	Original Representation	Floating Point
1	Organ	15 bit linear	8/3
2	Piano	15 bit linear	8/3
3	Piano	15 bit linear	10/3
4	Sine Wave	15 bit linear	10/3
5	Sine Wave	15 bit linear	12/2

(FIGURE 2)

For studying the noise immunities of several error correction and concealment schemes, we developed a dropout noise generator according to the Gilbert model where the probability P is the probability of the digital data stream to change from the good state to the bad state and p is the probability of changing from the bad state to the good state. The quantity $\frac{hP}{P+p}$ represents the bit error rate where h is typically 0.5. We studied two types of noise models. In one, the data in the "bad" state were represented by random bits, and in the other, data in the bad state were all zeros.

Theoetically, a 16/22 Hamming encoder followed by a 1024 word interleaver should produce about eight uncorrected errors per hour for a dropout generator having $P = 6.23 \times 10^{-7}$ and $p = 2.95 \times 10^{-2}$ after decoding. These Gilbert noise constants are reported as typical for a home VTR recorder.

In actual tests over two minute intervals such an encoded data stream subjected to the above referenced dropout generator produced no uncorrected data after decoding.

We then proceeded to a noise representation where $P = 4.9 \times 10^{-5}$ and $p = 6.29 \times 10^{-3}$, for which the bit error rate is 3.9×10^{-3}. Such a dropout generator would produce over 100,000 uncorrected dropouts per hour using a Hamming 16/22 encoder and 1024 word interleaver after decoding. This was used for further noise immunity investigations.

The demonstration will show test results in pairs, the first being the original signal followed by the results of the encoding, interleaving, dropouts and decoding, sometimes followed by concealment. Figure 3 is a list of the tests to be demonstrated. The piano was the program source.

Test No.	Noise	Encoding	Interleaving	Concealment
1	Random bits	None	None	None
2	Zero bits	None	None	None
3	Random bits	Hamming 16/22	1024 word single	None
4	Random bits	Hamming 16/22	1024 word double	None
5	Random bits	Hamming 16/22	1024 word single	*
	(* Previous good sample holding for double bit errors)			
6	Random bits	Hamming 16/22	1024 word double	*
	(* Continuing particular bit correction for double bit errors and correcting remaining bit)			
7	Random bits	Hamming 16/22	1024 word double	*
	(* Previous good slope holding)			
8	Random bits	← 1024 word FFT →		
9	Zero bits	← 1024 word FFT →		

(FIGURE 3)

In the last two tests the FFT concealment scheme is the analog of an optical holographic recording. In test No. 8 random bits for dropouts is the eqüivalent of a portion of the desired hologram with many random holograms interleaved. Test No. 9 is the original hologram with a portion removed by dropouts. This corresponds to the well known optical holographic phenomenon in which the pertinent data for a complete image is spread over the whole recording and any portion thereof may be reproduced showing the complete scene but with lower resolution.

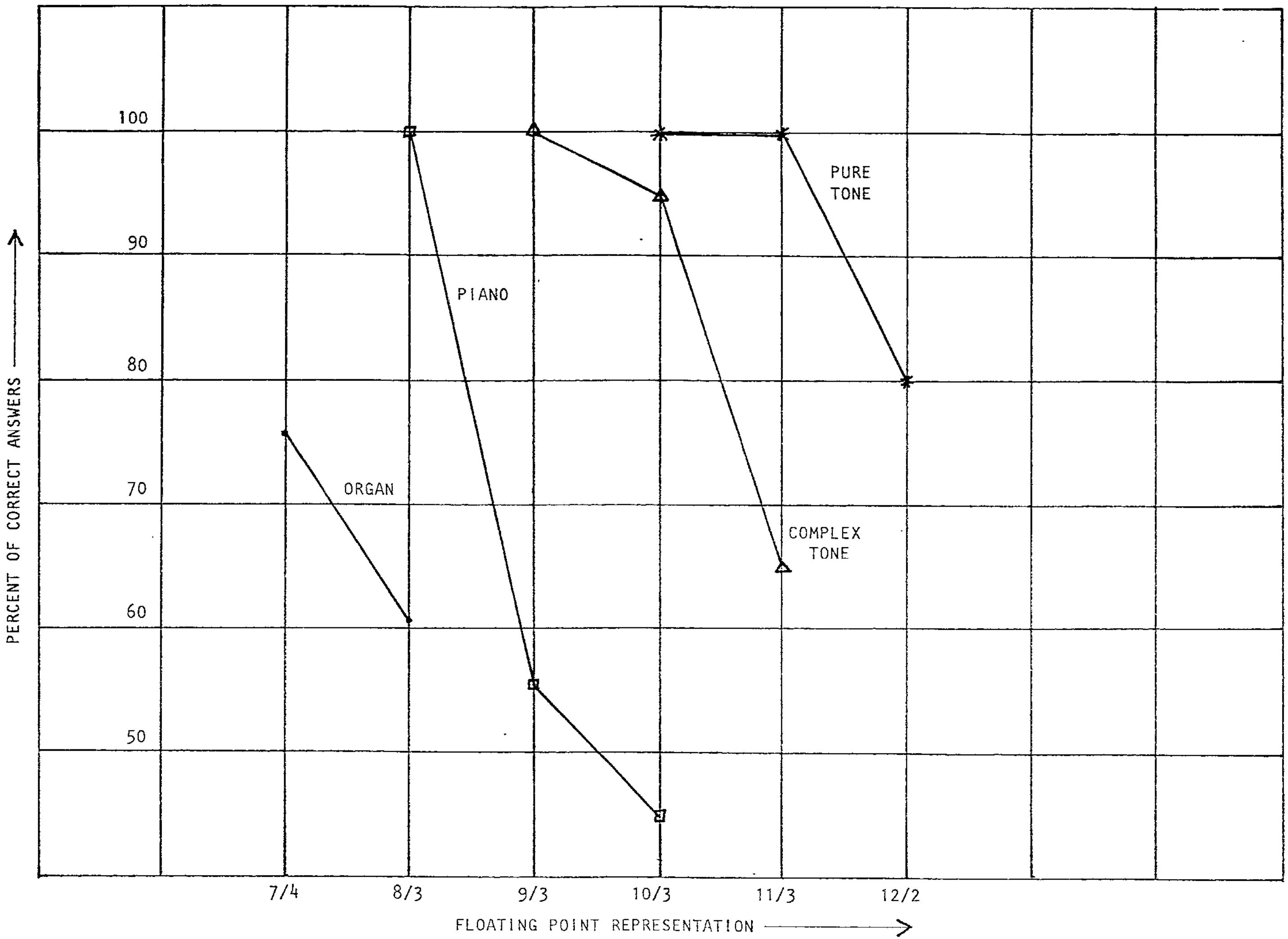
In our error correction with concealment programs, the overhead produced was only 37.5%. In the FFT concealment scheme the overhead was 100%. We plan to try more involved error correction schemes with higher overhead to permit more balanced comparisons to the present FFT concealment scheme. Our opinion is that no acceptable scheme should produce clicks or pops or obviously extraneous sounds from digital audio recordings.

CONCLUSION

Once analog program information is converted to the digital format and can be stored in the memory of a high speed general purpose computer, the efficacies of various error correction and concealment schemes may be tested (not necessarily in real time), the results stored and later played back in "real time" for critical listener evaluation. This avoids the necessity of generating a proliferation of hardware for testing ideas. Furthermore, as often is the case, the hurried implementation of an idea into hardware may result in failure not because of the idea but rather because of the limits imposed by the implementation. The use of a computer can be used to test the effectiveness of the real idea and later carefully transformed into hardware knowing full well that the idea is a good one.

ACKNOWLEDGEMENT

We wish to acknowledge the help given us on this project by Larry White and Post Doctorate Research Fellows Sally Browning and Bart Locanthi, Jr. of the Computer Science Department, California Institute of Technology. This project is sponsored by the Pioneer Electronic Corporation of Japan.


100
90
80
70
60
50
PERCENT OF CORRECT ANSWERS
PIANO
PURE TONE
ORGAN
COMPLEX TONE
7/4
8/3
9/3
10/3
11/3
12/2
FLOATING POINT REPRESENTATION

(FIGURE 1)