AES 76th Convention, NYC

Title: Strategies for the Representation and Data Reduction of Digital Music Signals

Author: John P. Stautner
CompuSonics Corp.
1 Arnold Circle
Cambridge, MA 02139
(617) 354-0050

6/20/84

Work performed and methods employed:

This paper describes research conducted and sponsored by the CompuSonics Corporation to investigate the data reduction of digitized musical signals through the application of perceptually-based filterbank algorithms. Such processing is intended to reduce the data required to represent musical signals while maintaining the original perceptible signal quality.

The algorithms are based on the auditory transform, an audio analysis algorithm modeled after observed behavior of the auditory system. The auditory transform essentially consists of a short-time Fourier transform log-magnitude analysis of the musical waveform, where the parameters of the analysis are chosen to approximate the critical-band and masking sensitivities observed in psychoacoustic studies of human hearing, as well as the physiologically observed filtering and detection nature of peripheral auditory processing.

After analysis, the musical signal is in a time and frequency dependent format which provides a convenient representation for visual display and feature detection, and forms a foundation for data reduction using a variety of waveform compression algorithms to code the individual analysis channels. These compression algorithms code both the global tendencies of the auditory transform surface, as well as the deviations or signal residue in a particular channel. The particular algorithms employed depend on the nature of the signal and on whether or not real-time operation is required. They range from adaptive, logarithmic companding, and thresholding type algorithms which operate in the real-time mode, to batch mode statistical analysis algorithms which require knowledge of the entire signal prior to coding. This combination of techniques allows a musical data rate ranging from one half to one tenth that of conventionally sampled high fidelity audio signals.

Resynthesis of the musical waveform is achieved by first expanding the coded auditory transform representation, followed by signal reconstruction using a short-time Fourier transform modified optimum least-squares synthesis algorithm.

Significance:

The results reported in this paper are significant because they demonstrate the capability to significantly reduce the amount of data required to represent conventionally sampled digital music signals (50 Khz sampling, 16-bit quantization), with little or no degradation in signal quality. Furthermore, the compression ratio may be selected subject to computational constraints and the operational mode (real-time and/or batch). These results make it possible to store large amounts of musical information on a limited media, and also provide the ability to transmit high fidelity audio signals economically over limited bandwidth transmission channels.

Conclusions:

It has been shown that the representation and data reduction of musical signals may be achieved through the application of perceptually-based signal analysis combined with waveform compression algorithms. Since the coding procedure occurs in several stages, one or several of these stages may be selected depending on data compression requirements and real-time vs. batch-mode operation. These algorithms may be applied for the representation of musical features, data reduction, the economical storage and transmission of music.

Architecture of a Real Time Digital Filterbank Processor for Tempered, Auditory, and Critical-Band Analysis/Synthesis

Gary W. Schwede, Ph.D.
CompuSonics Corporation
337 Calcaterra Court
Palo Alto, CA 94306
415 - 493 - 1623

ABSTRACT

Several Short-time spectrum analysis/synthesis patterns of interest in music are unsuitable for implementation by block-processing algorithms, e.g., the FFT. A tempered analysis with channel centers at semi-tone frequencies and response times which scale with wavelength is such a pattern. Analysis/synthesis schemes which incorporate known characteristics of the auditory system require finer time resolution at high frequencies than is consistent with low frequency selectivity. Block processing depends upon a single time window length for its efficiency, and is therefore not suited to such applications. Unfortunately, naive implementations of these filterbanks by direct convolution are too expensive for real time applications at present.

Incorporating new algorithms for constant-Q spectrum analysis and matching the structure of a multirate filterbank processor to them, a cost-effective real-time music signal analyzer/synthesizer has been designed. The data path and control structures are described, and examples of the performance of the processor in real-time music signal processing are given.

Architecture of a Real-Time Digital Filterbank Processor for Tempered, Auditory, and Critical-Band Analysis/Synthesis

Gary W. Schwede

This paper exemplifies the need to create signal processing algorithms and architectures which complement one another. Most of the computations of interest in audio are so demanding that no general purpose machine can achieve real-time rates. I describe a specialized processor which reflects the structure of spectrum analysis and synthesis patterns appropriate for music signals.

Audio-visual requirements:

Overhead projector

AES 76th Convention, NYC

Title: A High Speed Telecommunications Interface for Digital Audio Transmission and Reception

Author: Hyun Heinz Sohn 6/20/84
CompuSonics Corporation
1 Arnold Circle
Cambridge, MA 02139
(617) 354-0050

Work performed and methods employed:

The paper describes the design, specification, and implementation of the High Speed Telecommunications Interface. The design philosophy included the following considerations:

- compatibility with existing Multibus microcomputer components, including processors

- simplicity in developing software drivers for specific computer configurations

- transparency of interface operation to the user; the interface buffer appears to all software and hardware on the Multibus as Random Access Memory (RAM)

- high bandwidth; the transmission speed of 56,000 bits per second is limited by the Accunet, the interface is capable of up to 2 million bits per second

To achieve high bandwidth, the interface is controlled by a custom micro-sequencer running on a 10 MHz clock. The micro-sequencer is implemented with discrete TTL (transistor-transistor logic) components. The digital signals transmitted and received are fully buffered 32 kilobytes of RAM. All the buffering is controlled by the micro-sequencer and totally transparent to the host system. The correct buffer is automatically mapped into the predesignated memory locations on the Multibus.

Significance:

The interface will be tested at AT&T Bell Laboratories in July, 1984. The successful transmission and reception of digital audio signals over phone line will demonstrate the feasibility of distributing digitally encoded music through databases over phone lines. Benefits of using this technique include:

- reduction or elimination of transportation and inventory problems associated with conventional records and tapes

- accountability of each transmission/sale potentially reduces software pirating

- expansion of market through increased sales from impulse buying; combined with services such as MTV (Music Television), the gap between media exposure to an actual sale of a recording may be decreased

- superior sound quality; since the music stays in digital form, no noise or distortion is introduced between the studio recording and the final purchased recording

Conclusions:

A high speed digital telecommunications interface that will allow the transmission and reception of digital audio signals over the AT&T Accunet has been designed and implemented. This will make feasible the distribution of studio quality digital recordings from databases over conventional phone lines.

AES 76th Convention, NYC

Title: Specifications and Implementation of a Computer Audio Console for Digital Mixing and Recording

Author: David M. Schwartz 6/20/84
CompuSonics Corporation
323 Acoma St.
Denver, CO 80223
303-698-0060

Work performed and methods employed:

This paper describes the specifications and the means of their implementation for the CompuSonics DSP-2000 Series Professional Digital Audio Mixer/Recorder. The DSP-2000 was developed in a two year effort aimed at meeting the needs for digital sound production facilities in both commercial and institutional environments. Among the functional specifications are:

* High resolution analog to digital and digital to analog converters at variable sampling rates
* Digital mixing
* Digital equalization
* Digital recording with virtual multi-tracking
* Digital routing for sends, track assigns and effects
* Gigabyte modules of on-line magnetic disk storage
* Mix-down to various digital stereo formats; EIJA, Sony and CompuSonics
* Real-time data compression for increased storage density
* SMPTE time code master and slave modes
* Color graphics with semi-animated 3D representations of the audio signal
* Music file editing facility
* Unix Version 7 System III compatability
* Telecommunications capability

The implementation of the functional specifications includes off-the-shelf board level components as well as custom circuitry. Topics addressed in the paper with respect to the machine design are:

* Architectural overview
* Ergonomic considerations and execution
* The Motorola MC68000 as system master
* The disk drives and controller
* Texas Instruments TMS320 DSP chips as co-processors
* The NEC 7220 + MC68000 co-processing graphics subsystem
* High speed synchronous data bus and the Intel Multibus
* AT&T Accunet 56,000 bps telecommunications interface

Significance:

The realization of an integrated mixer/recorder/editor in the form of a single user microcomputer specifically designed for audio engineering is expected to facilitate the noise-free production of digital software such as compact discs. Physical handling of tape for splicing may be eliminated. Trial mix-downs and edits may be made at the time of recording without post-processing. Data compression for storage efficiency and reduced transmission bandwidth may be accomplished in either real-time or in batch mode. Graphical display in multi-color and 3D may be used to facilitate signal analysis.

Conclusions:

A special purpose micro-computer optimized for audio software creation has been specified, designed, and put into production. Live music performances may now be mixed, recorded, analysed and edited in a single uniform digital environment by recording engineers with little or no training in the use of computers.