Trevor de Clercq
February 9th, 1999

Digital Audio Processing II
E85.2601, Prof. K. Peacock

Research Status Report #3

Despite having previously focused my research on oversampling and noise shaping techniques, I thought that this week I might take a closer look at Sony's Direct Stream Digital. Tom Doczi's poster session report in DAP two weeks ago on DVD touched on this topic and got me interested enough to do some outside exploration on my own.

The main reason that Sony engineered a whole new encoding method called Direct Stream Digital (DSD) was to overcome the complaints and disadvantages of the standardized pulse code modulation (PCM) system of digitization. The need for a new conversion method arose out of Sony's huge catalog of aging master tapes. Because of the concerns that PCM audio, especially 16-bit 44.1 kHz "CD quality" sound, does not capture all of the sonic information that exists on the industry standard 1/2" two-track analog tapes, Sony felt insecure to trust PCM encoding to their thousands of gold and platinum masters. "We have 300,000 master tapes, and tens of thousands of lacquers," said a Sony rep, "and they're not getting any younger." DSD is hopefully Sony's answer to accurately preserving these masters in a digital format.

The means by which DSD digitally encodes waveforms differs rather profoundly from the currently accepted PCM encoding system. With PCM audio, the exact amplitude (or voltage) of each sample is stored as a specific wordlength. This method, however, often contains much redundant information since the rate of change between the voltages/amplitude of any signal is inversely proportional to the sampling frequency. In other words, with increasingly high sample rates, the amount change that a waveform can undergo is extremely limited and predictable, especially if we consider that most of the relevant sonic information is way below half of the Nyquist frequency. This redundancy often comes in handy with error correction systems in the form of concealment when their error correction capacity has been surpassed. Ignoring the error corrective powers of redundant information, DSD boils down digitization of audio to the simplest, most important data--that of the change between samples.

This method of measuring the change between samples is commonly known as Delta-Sigma modulation. Delta, of course, is the greek mathematical symbol for change (often delta-t for changes of time in physics). Sigma, similarly, is the greek mathematical symbol for summation (much like the integral of calculus). Delta-Sigma modulation therefore refers to measuring the change from one sample to the next (delta) and then adding this to the last samples voltage to find the voltage of the next sample (sigma). Obviously, some sort of reference voltage is needed to start the whole process off, but this beginning is just a matter of calibration.

Since the digital information is only measuring the change between one sample and the next, the data can be stored as just a single bit. This single bit refers to an increase in amplitude (1 or on) or a decrease in amplitude (0 or off). Since their is no way of representing an unchanging signal level, the sampling rate of such a system has to be relatively high. Luckily, DSD operates at an extremely high sampling frequency, typically 2.8224 MB per channel. If one does the math, this works out to 64 times the commonly accepted sampling frequency of 44.1 kHz. Just a little more math will prove that a 1-bit 2.8224 MB coding system holds four times the information of a 16-bit 44.1 kHz system (1 X 2,822,400 = 4 X 16 X 44,100). The information rate, therefore, of DSD already surpass the traditional PCM system.

This extremely high sampling rate also translates into other benefits of the DSD system of PCM audio. Anti-aliasing filters can be set with an ultra-high cut-off point and a smooth slope, thus avoiding noticeable analog filter artifacts from the audible spectrum. In general, DSD has frequency range up to about 100 kHz, far above the 20 kHz theoretical ceiling to human hearing and even above the 80 kHz frequency which some people have attributed to audible transient information. Dynamic range is a whopping 120 dB, just about the exact difference between humans threshold of pain (140 dB SPL) and threshold of hearing (20 dB SPL). All information is thus captured with the DSD system.

With such specs, it seems that DSD has a promising future, at least in the field of archiving and mastering applications. Currently, it is more difficult to conduct DSP applications with DSD audio than traditional PCM audio, but it is only a matter of time before such algorithms are developed. Even if one wants to store or conduct processing with todays PCM products, the down-conversion of DSD to PCM is quite a simple task considering the high sampling frequency of DSD. It will accomodate PCM sample rates up to 96 kHz and word lengths up to 24 bit, the modern upper-limit for both time and amplitude resolution in anyone's book. In ten years, perhaps out PCM based compact discs will be replaced with a DSD based medium.

Bibliography

Borwick, John. "Inside Audio: Sony Introduces Direct Stream Digital." Gramophone. May 1996. pgs. 143-144.

Eargle, John. "DSD on DVD?" Audio. January 1998. pgs 20, 22.

Kessler, Ken. "Mondo Audio: Sony's Simpler Sampling." Audio. October 1996. pgs. 32, 34, 36.

Lehrman, Paul. "Sony Direct Stream Digital." Mix. May 1996. pg. 130+.

Verna, Paul. "Sony, Phillips at work on successor to CD." Billboard. June 21, 1997. pgs. 1, 86.