Newsgroups: rec.audio.high-end
From: prls!max@uwm.UUCP (Max Hauser)
Subject: MASH, Bitstream, oversampling, etc.
Date: 17 May 91 13:53:09 GMT

In article <12047@uwm.edu>, smithjh@math.orst.edu (Jeremy Smith) inquires:

(1) what is the MASH 1-bit D/A conversion?

(2) Is this better than oversampling?

Intelligent questions deserve an intelligent answer. In fact these and related questions arise regularly, so I will try a more general response. I last did so on the "MASH" topic in (I think) summer 1989, on the group rec.audio. I will limit myself to the definitions and the technical context and thus leave the field open to others to comment intelligently on personal experience and impressions of specific products containing these techniques.

"MASH" is a commercial acronym to describe a class of oversampling data converters (either A-to-D or D-to-A) introduced and popularized by T. Hayashi of Nippon Telephone and Telegraph in February 1986 at an integrated-circuits conference [Note 1]. Thus "MASH" denotes a subset of "oversampling" rather than a competing technique. The acronym derives from and means "MultistAge noise SHaping." While the technique was popularized by Hayashi, he did not use the actual acronym "MASH," which came somewhat later, from K. Uchimura et al. [Note 2]. Also, the underlying idea is much older still [Note 3].

The technique now dubbed "MASH" is a series of small one-bit oversampling data converters each of which does a partial, and noise-shaped, conversion and then passes a residue (conversion error) to the next stage. The individual outputs of all of the stages are then combined (this is nontrivial) to form a composite output. That output is passed to a digital filter (for an A-to-D function) or an analog filter (for a D-to-A function). (Yes, that's right -- I've proofread it.) This brutally short explanation does not do the subject justice and may even leave unfamiliar readers with the mistaken impression that MASH resembles successive-approximation or pipelined data conversion (well-defined in the art), which it does not, differing deeply in philosophy. However it's an accurate nutshell description in that it will endure deeper scrutiny. For further details you may wish to look in one of my survey papers, such as the AES Journal January/February 1991.

The technique dubbed "MASH" by NTT is also called, less commercially, "feedforward" or (less precisely) "cascaded" higher-order noise-shaping [Note 3]. It is one of several different topologies that can implement the "noise-shaping modulator" portion of an oversampling data converter. Other such topologies include multibit and one-bit pure-feedback modulators. The latter (one-bit) class are called delta-sigma in the research literature, and "Bitstream" by NV Philips in digital-audio products (these techniques also have many applications outside of audio). "MASH" data converters are themselves not truly "one-bit" data converters in a meaningful sense, as I have explained in more depth in print, although they are made up of one-bit subsections and this often causes confusion.

Each of these competing modulator topologies has technical strengths and weaknesses that are very involved and do not lend themselves to reductive explanation. The signal fidelity in the "MASH" technique can be very good but depends on a different set of circuit elements than in the one-bit schemes. It is all a matter of "second-order" electrical effects; if the components are all perfect (as they invariably are assumed to be, in popular "explanations" of this subject matter), then all the techniques work well.

When Hayashi introduced the "MASH" technique to the international community in 1986 (I was present), it was suggested from the floor that this method, although a viable alternative to one-bit oversampling, did introduce certain additional vulnerabilities to circuit imperfections, and moreover that the real justification for the method was not clear (Hayashi's actual motivating statement at the time was dubious). It has since been suggested, perhaps insightfully, that the sheer novelty of the "MASH" method might afford a commercial, rather than a performance, advantage since the other basic oversampling methods occupy expired patents and are therefore in the public domain. (My use of passive verbs in this paragraph is not accidental.)

But I think it even more important, in the context of this newsgroup and the usual questions, that none of these considerations need correlate at all with audible differences among finished audio products that employ these data converters! Offhand I would suggest that audible quality could easily be affected far more by such prosaic outboard factors as the choice of output buffer amplifiers, and the degree of analog-digital isolation in the conversion-reconstruction circuitry, than by whether the internal modulator is realized via "MASH," or "Bitstream," or some other topology. This has not, of course, prevented legions of advertising copywriters and cult audio pundits from pontificating about the "revolutionary" aspects of this or that topology and deploying buzzwords like "noise-shaping" and "one-bit" with impressive vacuity. They apparently prefer this to focusing on the sonically and technically more vital, but less glamorous, issues.

There is my terse explanation of the relationship between "MASH," "one-bit," "Bitstream," "delta-sigma," and "oversampling" data converters. Be skeptical of anyone who asserts glibly that one of these techniques is "better" than the others. You might be able to hear differences among CD players using the competing schemes, but almost certainly not for the reasons everybody talks about. Such differences could easily be due to any of the numerous important factors other than the particular choice of data-converter topology.

Notes from text:

Note 1: Hayashi et al., ISSCC 1986. Printed version in that year's ISSCC digest, pages 182-183.

Note 2: Uchimura et al., ICASSP 1986. Printed version in 1986 ICASSP Proceedings pp. 1545-1548. Research papers quite regularly misattribute the modern origin of these circuits to this ICASSP paper rather than to the earlier ISSCC paper [Note 1] by co-workers.

Note 3: The technique has existed in various forms, including a small paper in "Electronics Letters" in 1969. Candy and Temes have told me that they would cite this in their forthcoming broad research overview of oversampling. I find it still more intriguing that the technique is also a data-conversion implementation of Black's multistage linear amplifier (US patent 1,686,792, issued 1928), a classic invention actually predating the invention and patent of negative feedback, also by Black.

Max Hauser {mips,philabs,pyramid}!prls!max prls!max@mips.com

"A lot of things can be solved by the use of jargon -- for example, the effort of thinking, or the danger of saying something that someone else may not like. You don't have to be clever, and you're always on the side of whoever has the money or power ..."
-- Stanislav Andreski, author of "Social Sciences as Sorcery"

Newsgroups: rec.audio
From: exspes@gdr.bath.ac.uk (P E Smee)

Subject: Re: Oversampling for kids
Date: 6 Jun 91 14:06:37 GMT

In article <1991Jun4.174051.594@cs.sfu.ca> rosen@cs.sfu.ca (Wilf Rosen) writes:
>
>Could somebody please explain what oversampling really means with regard
>to CD players. The only information I have been getting is from friends
>who claim to know, or from (ulp) salesmen. No explanation I have seen yet
>is really believable. (The most common explanation is that e.g. 4x
oversampling
>means the disk gets read 4 times. Somehow this seems silly).

Layman's explanation follows. I guarantee it to be sufficiently accurate for your expressed needs, but to keep it simple I'm going to skip some of the technical complexities at a detail level where they shouldn't matter.

First, the problem with non-oversampling players: The mechanics of the sampling process mean that you also get a spurious signal at half the sampling frequency (22.05 kHz for CD 44.1 kHz sampling); and that you also get a sort of 'mirror image' of the real audio signal, reflected around this 1/2-sampling freq. E.g. a 20KhZ note will also create an artifact at 25 kHz. These artifacts ARE beyond the range of human hearing, but they can cause follow-on side effects which impinge into the audible range, so they need to be filtered out. However, it is also very difficult (and or expensive) to build a sharp-cutoff filter which will not do nasty things to the bit of the signal which is passed through.

So, how do you fix this? You increase the sampling rate. However, there are only 44.1 khZ worth of samples on the disk. What an oversampling player does is to take the samples it reads (reading the disk only once) and by use of various 'curve-fitting' types of algorithms, works out intermediate values. Like extrapolation, only fancier. A 2x oversampler works out one intermediate value between each pair of read values. A 4x works out 3 intermediates.

The effect is to increase the apparent sampling rate, at which (or reflected around which) the artifacts I first mentioned occur. Thus, they can be filtered out by a much less aggressive filter, which will not mess up the desired part of the signal. (Some very high x oversampling players can actually dispense with this filter entirely, on the grounds that the artifacts have been moved to such a high frequency that they won't interfere with the final result even if they are left in.)

If you're going to try to find references, a likely keyword would be 'digital filtering'.

--
Paul Smee, Computing Service, University of Bristol, Bristol BS8 1UD, UK
P.Smee@bristol.ac.uk - ..!uunet!ukc!bsmail!p.smee - Tel +44 272 303132

Newsgroups: rec.audio
From: jpb@calmasd.Prime.COM (Jan Bielawski)
Subject: Re: Oversampling for kids
Date: 7 Jun 91 19:45:48 GMT

In article <1991Jun4.174051.594@cs.sfu.ca> rosen@cs.sfu.ca (Wilf Rosen) writes:
<
<Could somebody please explain what oversampling really means with regard
<to CD players. The only information I have been getting is from friends
<who claim to know, or from (ulp) salesmen. No explanation I have seen yet

<is really believable. (The most common explanation is that e.g. 4x oversampling
<means the disk gets read 4 times. Somehow this seems silly). I realize this
<may well be a FAQ, but I haven't seen it discussed for the past 2 months.
<Even a pointer to a good book on the matter would be real swell. Thanks
<for reading up to this point.

Let me try, see how much I don't know. A digital oversampling filter is a piece of software that accepts digital data (representing the 44.1 kHz samples from the CD), does some number crunching on them and outputs recalculated digital data that represent identical musical information as the original but because of certain reformatting the new data is easier to transform into analog ("easier" meaning mostly "using less expensive electronics" although it's probably true that this oversampling approach is inherently better). The algorithm does look like sticking in extra samples between the existing ones. One even hears of sticking in *zeros* in between but this seems to have come from the fact that the oversampling program initializes the variables representing the extra samples to zero before processing -- something any decent program does (well, usually...).

The algorithm is described, for example, in John Watkinson's "Art of Digital Audio". A bit drier and theoretical discussion is in Teukolsky's et al. "Numerical Recipes".

Jan Bielawski
Computervision, San Diego
jpb@calmasd.prime.com

Newsgroups: rec.audio.high-end
From: cs.utexas.edu!wotan!prls!max@uwm.UUCP (Max Hauser)
Subject: Oversampling for the curious, the furious, and the damned
Date: 24 Jun 91 12:47:18 GMT

You may wish to save this article either for interested friends or for future reposting as necessary since my readership of these newsgroups is irregular.

Introduction:

This is a broad and not-very-technical online summary of CD oversampling, antidotal to the lies and nonsense served up in consumer-audio retailing and an alternative to the well-intentioned but misleading or fractional explanations often seen in the popular audio press (including the so-called serious or high-end publications). This synopsis is less detailed but much broader than my earlier "technical summary," written for engineers and posted to rec.audio in 1987 and occasionally since.

If you want more depth of information on these topics, my colleague Prasanna Shah recently published a dense technical overview of audio oversampling in the popular magazine Audio in January. I published a long technical tutorial (not specific to CD or DAT products) in the Journal of the Audio Engineering Society, January/February 1991 (vol. 39 no. 1/2 pp. 3-26). A lighter and shorter overview of mine is also available as Preprint #2973 from the Audio Engineering Society, 60 East 42nd Street, New York, New York 10165 USA, Telephone 212-661-8528, FAX 212-682-0477. This preprint was recently recommended by the popular magazine Stereophile. AES charges $5 for such preprints ($4 for members), $3 for journal-article copies and $10 for back issues, and is pretty efficient about getting them into your hands if you FAX them a request with V. or MC. credit-card information (I did so recently and the copies arrived in the mail in about three days). These papers contain many further references. Some of these sources emphasize the A/D rather than D/A path but the core principles are identical (the circuit implementation of each section interchanges between analog and digital).

Do not be alarmed if the following summary takes an approach different from what you read elsewhere. There are details here not usually mentioned in popular summaries (or even in the research literature).

A. Thumbnail Sketch of Oversampling

Signals such as audio, stored digitally, entail a finite *sampling rate* (44.1 kilosamples/sec for the 12-cm CD) whereas in their natural (analog) form they are continuous-time waveforms (you can think of this usefully as an "infinite" sampling rate). The circuitry that regenerates the continuous-time analog output in a CD player has two major tasks: translating a stream of digital numbers into analog values ("conversion") and also bridging between the finite sampling rate of the digital sequence and the "infinite" sampling rate of the outside world (that is, restoring a correct continuous waveform from discrete samples) -- "reconstruction."

Non-oversampling conversion-reconstruction (C-R) systems make the transition from finite to "infinite" sampling rate in one step, while oversampling systems do it through one or more intermediate sampling rates (higher than the original, but still finite). Although the details may not be obvious at this point, producing these intermediate signals with elevated sampling rates is a purely digital process and can thus be performed predictably and repeatably (although it does require that you have technologies where digital logic is very cheap, and therefore it was unattractive until recent years, although the basic techniques have been known since the 1950s and in embryonic forms since the time of the second world war).

Not only the reconstruction process but also separately the conversion process (bits into volts) benefits from the use of an intermediate sampling rate on the way to continuous time. Designers can orchestrate eloquent mathematical tricks to trade a higher deliberate sampling rate for lower required resolution in internal digital-to-analog converter (DAC) circuitry. This in turn tends to render the analog part of the C-R chain simpler and more tolerant of component fluctuations. But moreover, in practice oversampling C-R systems blend the two tasks of conversion and reconstruction so that they overlap in actual hardware, unlike a classical, non-oversampling system. The subjects of this paragraph are extremely complex and seductively counterintuitive even to well-trained engineers, and they habitually garner the most imaginative misinterpretations in popular-press writing.

An oversampling conversion-reconstruction (C-R) system in practice normally contains a series of four major blocks. The first is a sampling-rate-increasing digital filter, the second a digital quantization-management subsystem or "noise-shaping modulator," the third a DAC circuit _per se_ and the fourth an analog lowpass filter. A classical, or non-oversampling, system lacks the first two blocks, but is far more demanding of the last two blocks, which are analog circuits that largely determine the performance and subtler behaviors of the signal path. (That's the whole reason for oversampling, in a nutshell.)

By the way, these four blocks reflect a combination of traditionally separate specialties in electrical engineering, each with a different intuition and set of assumptions about what is technically difficult or important. This is why you will find many different explanations of oversampling (some of them seemingly in conflict) even from competent specialists. The first of the four blocks is generically a digital filter, the second a quantizer (or quantized feedback system), the third a precision analog circuit and the fourth an analog filter, and most or all are realized in integrated circuitry. Thus, for example, someone familiar with digital filtering will usually focus on the first of the four blocks, and when asked for more information will instinctively steer you to the general digital-filter literature (which unfortunately is extremely dilute on this subject). In reality an oversampling C-R chain entails intimate concert between all four blocks and between multiple technical specialties -- none alone is sufficient to explain what is going on.

B. Interpolator

The first block, the sampling-rate-increasing digital filter, in an oversampling C-R system is commonly nicknamed an "interpolator." This jargon is triply unfortunate. First, almost everybody unfamiliar with multirate digital filtering assumes from the name, incorrectly, that this block performs "interpolation" in the common mathematical sense (such as linear or polynomial interpolation between data points). Actually the name is a specialized digital-filtering coinage subtly but crucially different. Second and third, as if that weren't trouble enough, the term "interpolative" is sometimes applied in two further ways to oversampling C-R systems (one of these usages is a subset of the other). More details about this and other glorious terminological pitfalls are in my recent AES Journal paper.

Here is the briefest sketch of how the rate-increasing filter works. The objective is to convert a signal at a sampling rate like 44.1 ks/s to a signal at a higher sampling rate *without* changing the information content. Mathematically this is a well-defined and tractable problem. If you just take the original sequence of samples and insert after each of them, for example, three new samples (with value zero, or holding the last old-sample value, or almost anything else intelligent) then you will obtain a new sequence at four times the original rate. In frequency spectrum this new sequence will however include new high-frequency replicas (images) of the original signal's spectrum. A digital lowpass filter will remove these images and leave a signal spectrally identical to the original. In the time domain, you will now see a higher-rate sequence that will look like the original but with the "right" new samples smoothly inserted between old. (In actual practice the "insertion" of new samples is NOT a separate step as above, but is incorporated into the digital-filter arithmetic.)

C. Multibit vs. single-bit vs. MASH etc.

All four of the major blocks of an oversampling C-R system, outlined in Section A, admit endless variations, opportunities for design cleverness, and high/low quality choices, and account for practical performance and manufacturing-cost differences among CD players. (All of which players, incidentally, appear more or less indistinguishable in the rudimentary and unrevealing standard specifications -- peak SNR, frequency response, step response -- commonly published.) A lot of fuss and advertising copy are however devoted to one particular design difference, the organization of the noise-shaping modulator (and consequently the format of the internal DAC circuit(s)). The common organizations are:

-- Multibit feedback noise shaping. The modulator properly predistorts (noise-shapes) the oversampled digital signal sent to a lower-resolution multibit DAC so that when properly analog-postfiltered its output will yield the full 16-bit resolution stored on the CD. This is the oldest scheme common in consumer products, widely popularized by the NV Philips SAA 7030 / TDA 1540 chip set (1983) with a 14-bit internal DAC and 4:1 oversampling yielding 16-bit final resolution. (For more details on how 4:1 relates to two additional bits, see either of my AES papers above.)

-- One-bit feedback noise shaping (called "Bitstream" by Philips and "delta-sigma" by the research community [Note 1]). Same as the previous but taken to an extreme: the internal DAC has only one bit of resolution and the 16-bit net D-to-A resolution is accomplished by the oversampling, noise-shaping and postfiltering process. This approach requires a higher oversampling factor, such as 128 or 256, other things being equal.

-- Feedforward or multistage noise shaping (abbreviated "MASH" [Note 2] by Nippon Telephone and Telegraph). A series of small one-bit feedback noise shapers, each of which operates on a quantization-error (residue) output from the previous stage, while simultaneously the quantized outputs are combined properly to form the main output. "MASH" data converters are definitely not "one-bit" data converters in a meaningful sense, as I've explained in more depth in print, although they are commonly made up of one-bit subsections and this sometimes causes confusion.

Each of these competing modulator topologies has technical strengths and weaknesses that are very involved and do not lend themselves to summary. The signal fidelity in each of them can be excellent but depends on different sets of circuit elements. It is all a matter of "second-order" electrical effects; if the components are all perfect (as they invariably are assumed to be, in popular explanations of this subject matter), then all the techniques work equally well. Very broadly, however, I would say that the one-bit designs have the fewest subtle distortion vulnerabilities.

D. What does it mean for sound

The electrical specifications of an oversampling C-R system depend on innumerable component values and design choices and are in no way simply predictable from whether the internal modulator uses, for example, MASH or Bitstream or some other topology. Still less predictable are perceptual fidelity measures, which of course are the ultimate figures of merit in audio and other human-interface electronics (a point taken for granted not only by musicians but also, of course, by competent engineering researchers and by the graduate communication-theory texts, such as Jayant and Noll).

This does not, of course, prevent glib advertising copywriters and cult audio pundits from directly linking this or that audible property to the glamorous _au courant_ labels like MASH and delta-sigma (just as it is thought very fashionable to talk about Fast Fourier Transforms, no matter how ignorantly or irrelevantly, in stock-price analysis, or about chaos theory in management -- I could go on and on, and I do). Such writers might even be right, but they almost certainly don't know it -- the audible differences are much more likely due to the oblique dependence of the C-R topology on other design choices in the player, or to the quality of analog-digital ground isolation, or to the choice of output-filter op amps.

Notes from the text:

Note 1: "Delta-sigma" modulation and data conversion (the inventors' term) was unintentionally rechristened "sigma-delta" at the Bell Telephone Laboratories in 1963 and this reversal has propagated through many paper titles, so you will see both names in use. No difference is intended. I have made efforts to redress this reversal and the principals are now in accord. My recent JAES paper mentions this and I have further details if anyone professionally interested sends a mailing address.

Note 2: Some people dislike the acronym MASH for MultistAge noise SHaping, though there certainly are endless well-known precedents (UNIted nations ChildrEn's Fund; GEheime STAatsPOlizei). When its coiners introduced "MASH" in the US in 1986 a colleague remarked to me that MUSH was better on acronym style. I think however that MUSH would have less audio-marketing cachet.

The technique now dubbed "MASH" by NTT has existed in various forms since long before its recent popularization by Toshio Hiyashi et al. in February 1986 (this origin itself is usually misattributed to a later paper by Uchimura et al.). I have antecedents going back at least to 1969.

Multibit feedback noise shaping is even older, due to Cutler in 1954.

Max W. Hauser   {mips,philabs,pyramid}!prls!max   prls!max@mips.com

"In graduate school, books are called `Introduction to ...' while in high school, they are called `All About ...'   This seminar is `All About' high-speed vertical-amplifier design."   -- Einar Traa, Tektronix, 1978

(PS: Experience with the Usenet compels the following cautions. This summary is a terse sketch of a complicated and counterintuitive subject. Abundant clarification and amplification of details is available to anyone who will take the trouble to obtain and study the references cited at the beginning, and if necessary take the further trouble to acquire, or else discuss them with someone having, the technical background on which the entire subject is built. If you cannot be bothered to take these steps then please do not ask me to do them for you. Also, I regret that I cannot offer audio-equipment recommendations.)