Making digital filters sound "analog"

Dave Rossum

E-mu Systems, Inc., 1600 Green Hills Rd. Scotts Valley, CA 95067-0015

0. ABSTRACT

Vintage electronic music synthesizers had a characteristic "analog" sound, largely created by the distortion characteristics of their voltage controlled filters. Computer music synthesis techniques frequently use the traditional accumulator saturation method of overflow cycle prevention in time varying filters which results in harsh and unnatural sound during filter overload. While simulation or physical modelling techniques could imitate the nonlinear behavior of a particular analog filter with arbitrary accuracy, computational limits prevent their use for purposes other than those of academic interest. However, some very simple modifications to familiar digital filter flow diagrams provide remarkably "analog" behavior with very little computational load. This paper investigates some of these simple modifications and the reasons behind their influence on the overload and distortion characteristics of the audio output. Specific audio examples will be provided which illustrate such behaviors as soft limiting and time variant "pull" and recovery of resonances.

1. INTRODUCTION

The process of dynamically filtering sounds to produce articulation of musical notes was the basis of the modern music synthesizer. The famous "Moog" sound and those of many other vintage instruments were based around subtractive synthesis as implemented by the analog voltage controlled lowpass filter. Musicians today frequently bemoan the loss of these warm and expressive timbres in today's music, and often complain that digital synthesizers just don't sound "analog." It is certainly true that the design of adequate digital filters for audio is a problem that most students of digital signal processing are not equipped to handle. Thus, the problem of making a digital filter that adequately simulated an analog implementation at the musical level is an excellent engineering, as well as musical, goal.

2. SMALL SIGNAL BEHAVIOR

The major problem in the small signal behavior of digital filters is generally termed quantization "noise". However, from the point of view of comparison with analog filters, the term "noise" would mislead us with regard to this specification. Analog voltage controlled filters are frequently quite noisy, yet the noise itself is not terribly objectionable. In contrast, the effects of quantization of digital audio signals can be very objectionable. Hence, I prefer the term "quantization distortion", which accurately reflects the fact that quantization produces signal-dependent audio artifacts.

Quantization distortion is produced when the mantissa of the signal representation is inadequate. The most common forms of digital filters maintain as their state variables sequential delays of the filter output (see figure 1). To derive the state of the filter, the differences between these delays must be taken, which causes the difference to always be similar or smaller in magnitude than the mantissa of these values. An important point to note is that a floating point implementation does not improve filter quantization distortion over a fixed point scheme with the same size mantissa.
One can estimate the required size of mantissa by considering a highly resonant, low frequency filter, which has been excited and is resonating with no signal input. In this case the output signal will be an exponentially decaying sine waveform, which must be produced entirely from the state variable data available in the two delays. Considering a 20 Hz resonance at a 44.1 kHz sample rate, the worst case occurs when at the crest of the waveform, wherein the difference between two samples is of a full amplitude signal is:

$$\cos(20/44100) \approx 1\text{-}10^{-7} \approx 1\text{-}2^{-23}$$

Since this difference must encode the entire amplitude and frequency state of the decaying signal, well more than 23 bits are required to be able to accurately reproduce this signal as it decays. At E-mu, we have found that a 32 bit mantissa provides inaudible quantization distortion for most musical filters.

Limit cycles are the other problem in the small signal specification of digital filters. Limit cycles result from the quantization errors as well, and can be made arbitrarily small by the use of floating point arithmetic. However, when a filter exhibiting limit cycle behavior is applied to a finite precision D/A converter, the limit cycle will be toggling the sign (MSB) of the converter and produce an audible signal regardless of the magnitude of the untruncated signal.

Limit cycles are generally what their name implies - cyclical sequences which result in low level audibly pitched signals. The use of dither (Vanderkooy 1989) can whiten the limit cycles to noise of a similar amplitude by "breaking up" the limit cycles. Alternatively, the limit cycles can be eliminated by the use of truncation towards zero at the filter output where the large mantissa is reduced to fewer bits for further processing. It is convenient that when 32 bit mantissa arithmetic is used, the limit cycles are already small enough that floating point arithmetic is not required to make them well below the significance of further processing when truncated towards zero. While truncation towards zero causes higher errors than rounding, the errors can be made small enough in the case of large mantissa filters that the differences are insignificant.

Whether dither or truncation towards zero is used to reduce limit cycle artifacts is largely a matter of choice. Dither, when properly applied throughout a signal path with adequate word size, is a beneficial practice. Unfortunately, it is usually the case that the signal path is not properly designed throughout, and the dither appears to introduce unnecessary noise. While truncation towards zero produces more distortion in theory, we have found that it is often perceived as the superior solution in practice.

3. SPECTRAL BEHAVIOR

The spectral behavior of digital filters is easily controlled, and is largely the reason for which digital filters are frequently employed. Interestingly, our research has shown that the spectral behavior of filters is not tremendously important. While the difference between a 2nd order and 4th order rolloff is fairly audible, the precise relative placement of the poles seems to be not very critical.

Vintage analog filters used all-pole filters (typically two or four pole lowpass). Similarly, many digital implementations use the all-pole or AR model for dynamic filters. Digital filters, however, operate at a fixed sample rate and as a result do not approach an asymptotic rolloff of 6dB/octave per pole at the high end of the audio band. Instead, due to the fact that the rolloff is akin to a cosine function in frequency, the rolloff of a pole levels out at high frequency. This, fortunately, is easily compensated using a tracking zero (see figure 2a and 2b).

4. DYNAMIC BEHAVIOR

The most interesting quality of classic analog filters was that they were voltage controlled - in other words their cutoff was dynamically swept throughout the evolution of the musical note. It was this behavior that became the "synthesizer" sound known to the public. And as a result, much of making a digital filter sound analog comes down to controlling they dynamics correctly.

Controlling the filter with the proper contour is important. Because most digital filters have their coefficients fairly directly related to cutoff frequency, the natural control function is quite dissimilar to the exponential (volt per octave) control of analog filters. Consequently, a transformation must be performed so that a linear sweep of a control variable (such as a MIDI controller) produces an exponential sweep of filter cutoff frequency. This can be done either by a mapping function or by a

direct encoding of filter coefficients (Rossum 1991).

The coefficients of the filter are themselves filtered. Consider a typical feedback filter, as in figure 1. When a coefficient changes, it will immediately affect the amplitude of the feedback signal, causing a discontinuity in that signal. However, this is equivalent to a similar discontinuity in the input signal because the feedback signal and input signal are summed. Consequently, the filter output will be affected in the same way as for a signal change, thus the effect of the coefficient discontinuity will be filtered. Note, however, that if the changes in coefficients are not band-limited, they will be equivalent to aliased input signals.

To band limit the coefficient changes, the coefficients must be updated at the sample rate. To update them at any lower rate will cause audio images of the update rate to appear in the signal path, which may be somewhat attenuated by filtering.

5. LARGE SIGNAL BEHAVIOR

The biggest contributor to the "analog" sound of the filter is the large signal distortion behavior. In the 1970's, the characteristic "Moog filter" sound was recognized by many musicians. Moog had a patented voltage controlled filter structure (Moog 1969), which had a unique distortion characteristic, quite different from the voltage controlled filter implementations done with transconductance amplifiers (Colin 1974). This overload distortion was the primary source of the "Moog" sound.

Analog voltage controlled filters, whether in the Moog "ladder" implementation or the conventional transconductance amplifier form, depended on the "S" shaped response characteristic of a transistor differential pair. The key characteristic of this circuit is a soft clipping action. Furthermore, the clipping action occurred at the input to the filter, so that the result of the clipping was itself filtered.

In digital filters, it is necessary to provide some form of protection from overload because the circuit provides gain. Consequently, certain input signals which have energy near the filter resonance will attempt to produce an output larger than the numerical range of the arithmetic (or of the D/A converter). Saturating arithmetic is typically used to solve this problem, most commonly at the accumulator of the filter. This forms an extremely sharp clipping boundary - the filter is perfectly linear until saturation is reached, and then the output signal instantly becomes heavily distorted. This sound gives common digital filters a characteristic raspy behavior when they are overloaded.

Fortunately, a first order approximation of the behavior of analog filters is available with very little overhead. If we add some headroom to the accumulator portion of the filter, and instead saturate only when delaying the signal prior to the multiplier inputs, we have caused the saturation to occur at a point equivalent to the input of the filter (see figure 3). This will be analogous to the analog filter's differential pair, and although the analog filter's clipping is an "S" curve and the digital filters a sharper "Z", the similarity is audibly striking.

Nonlinear systems have received much recent attention, and have been found to produce interesting behaviors which are similar to complex natural phenomena. Interestingly, the above scheme for digital filter clipping results in some phenomena which are also exhibited by non-ideal analog filters.
The first phenomenon to be noticed when experimenting with these filters is distinctly reminiscent from the days of analog synthesis - when a filter was excited into large signal resonance, an input signal could "drag" the resonance away from its nominal frequency. In a complementary way, when the input signal driving the resonance was removed, the filter would ring down with an audible shift in the pitch of the resonance. Hearing these phenomena, the "analog" nature of the filter is very convincing.

The cause of this behavior can be understood by considering the consequence of saturating the inputs to the multipliers. When this occurs, one could either say that the signal had been saturated (thus producing a smaller product at the output of the multiplier), or alternatively that the coefficient had

been reduced in such a manner as to give the same smaller product. Viewed in the latter way, it is obvious that the onset of distortion is equivalent to changing the coefficients of the filter. As mentioned above, the coefficients are themselves filtered, so the fact that the coefficients are varying depending on the instantaneous saturation of the signal will not result in hideous distortion. In a sense, one could say that the consequence of this scheme of overload protection is to alter the coefficients such that the input signal will no longer overload the filter.

Clearly, additional techniques could be applied to more precisely match the distortion and even the frequency characteristics of a particular vintage analog synthesizer filter, all the way up to a digital "SPICE" model of the filter. But the techniques described above show that through careful design, very little modification to the normal forms of digital filters is necessary to produce adequate analog-like behavior, including some very interesting non-linear effects. We feel that the proof of our success comes from a review of the EMAX II product (our first to use our digital filter technology) in Sound On Sound (Wrightson, 1989), which praised the "warmth" of E-mu's "analogue" filters.

6. REFERENCES

D. Colin, U.S. Patent #3,805,091, 1974
J. Dattorro, "The Implementation of Recursive Digital Filters for High-Fidelity Audio," JAES, vol. 35 #11, pp.851-878, Nov. 1988
R. Moog, U.S. Patent #3,475,623
D. Rossum, "The ARMAdillo Coefficient Encoding Scheme for Digital Audio Filters," IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, October 1991
J. Vanderkooy and S. P. Lipshitz, "Digital Dither: Signal Processing with Resolution Far Below the Least Significant Bit," Proc. AES 7th Int. Conf: Audio in Digital Times, May 1989
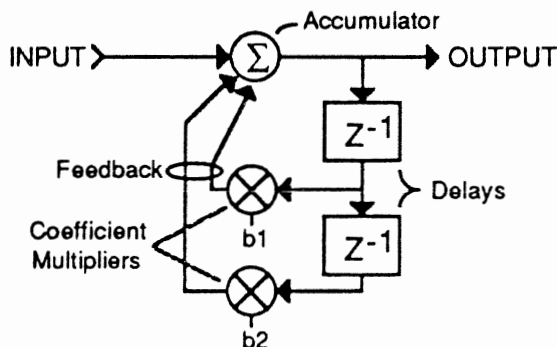**K. Wrightson**, "EMAX II," Sound On Sound, vol. 5 #2, Dec. 1989
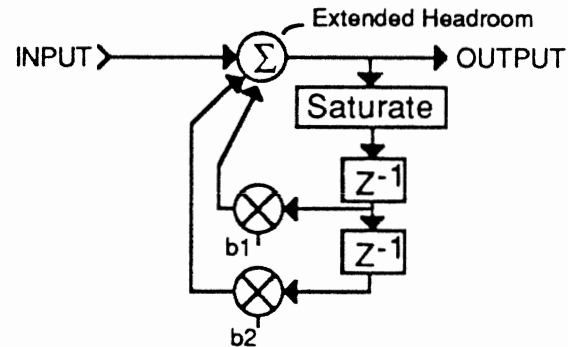
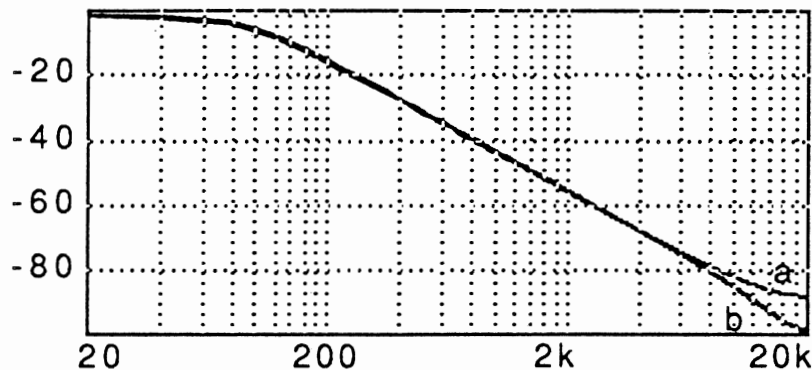Figure 1 - Typical Digital Filter

Figure 3 - Improved Overload Behavior

Figure 2 - Lowpass Filter Response: Poles Only (a), with Zeroes (b)