INFURIATING NONLINEAR REVERBERATOR

Miller Puckette

Music Department and CRCA, UCSD

reprinted from *proceedings*, ICMC 2011

ABSTRACT

A ceramic tile or other small rigid object is attached to a contact microphone, and any resonances of the physical system are filtered out using linear predictive analysis. The resulting audio signal is fed into a nonlinear reverberator to simulate a variety of real or fanciful percussion instruments. The result is a highly expressive and playable electronic percussion instrument that is both easy and inexpensive to build. The effects of various design parameters of the nonlinear reverberator on the resulting sound are discussed.

1. INTRODUCTION

This is a report on an ongoing project to build a collection of electronic percussion instruments in which the player energizes one or more nonlinear reverberation network by tapping on small objects with attached contact microphones. The algorithm has two steps: first, removing any resonant frequencies of the physical system (in rough terms, "deconvolving" the impulse response of the real object), and second, using the resulting audio signal as an excitation for the nonlinear reverberator. Depending on the design of the reverberator, a nice variety of sounds resembling hand drums, snare drums, cymbals, and gongs can be designed, not to mention a variety of more fanciful percussion instruments. The instruments are easy to build and are musically expressive.

The following section summarizes related prior work. In section 3 we describe the preparation of the input signal (the deconvolution step) and in the following section we describe the nonlinear reverberator design.

2. RELATED PRIOR WORK

The use of percussion instruments that incorporate contact microphones has such an exceedingly rich history that it is hopeless to review it here. The most famous example might be John Cage's *Child of Tree* (1975), which has inspired generations of musicians to attach contact microphones to cactus plants and amplify the sound made by plucking the needles. Many artists have built instruments using contact microphones to amplify the (sometimes processed) sounds collected from solid vibrating objects using contact microphones.

In contrast to that approach, the aim here will be to deliberately remove any artifacts of the sounding object itself in order to harvest the most neutral possible signal as an excitation signal for the artificial resonance provided by our nonlinear reverberator.

The work reported here is also distinct from the well-known practice of triggering synthetic or sampled drum sounds from an almost silent drum pad with a microphone or other captor. Such solutions are widely available commercially. They are far less expressive than instruments that transmit or process the vibrations themselves; for instance, sliding a brush over a drum trigger isn't likely to produce anything useful, whereas doing the same thing on an instrument that operates directly on the audio signal from the contact microphone (as we do here) has the possibility to create a wide range of useful musical sounds.

The idea of using a nonlinear reverberator as a sound-generating device is not at all novel. Essentially any network involving delays and obeying energy conservation rules (as a stable nonlinear reverberator normally should) can be regarded as a waveguide network[3]. What is interesting here is that we identify a particularly useful class of such networks and a way of analyzing their behavior that can lead to useful percussion sounds by imitating the way the skin of a drum stretches or the way the vibrational modes of a metal object can couple (passing energy back and forth among themselves) as a result of acoustical excitation.

3. PREPARATION OF THE INPUT SIGNAL

The physical system (including the contact microphone) should be as acoustically transparent as possible in order to provide the greatest flexibility in designing the sound of the instrument. Two possible approaches suggest themselves: (1) the instrument could itself be acoustically neutral, somehow recording the force of the the player's fingers or sticks directly as an audio signal; or (2) one could try to model the vibrating object faithfully enough to be able to infer the force of activation from the vibrations at the site of the microphone.

After some experimentation, the second approach was found to be the more successful one. Common ceramic and stone tiles and scraps of steel, ranging from 4 to 6 inches on a side, turn out to vibrate in a way that can be successfully modelled as a linear, all-pole system. Each object is glued to a piezoelectric contact microphone and placed on a piece of cloth so that it rings when struck or otherwise set vibrating.

The modes of vibration to be cancelled may be quite numerous. The modes of vibration of a three-dimensional solid object below a given frequency are roughly proportional to the frequency cubed, so even if the lowest mode rings at, say, 1000 Hz. there might well be 10,000 of them in the audible range. In practice the tiles and metal scraps were selected to maximize their lowest-frequency modes, which ranged from about 1000 to 2500 Hz. Tiles proved the most natural choice since they are easier to find than steel scraps of specified shapes.

Each mode can be thought of as a pole pair in the tile's transfer function, assuming the tile is hit at a fixed location. (Although the tile obviously will be hit at various locations, not a fixed one, the poles do not change depending on that location; what changes is the strength of coupling from the physical impetus to the various poles).

If there are N modes, they can theoretically be computed numerically by computing a linear predictive (LP) analysis of the captured signal of order $2N$, provided the analysis is carried out over a period of time during which the tile is ringing after being struck (but not including the strike itself). Under these conditions, the LP residual ("error") signal should theoretically be zero.

To do the analysis we record a test signal made by striking the tile a few dozen times, at various locations, with a small, hard object. We then determine the times of the strikes using the `bonk~` object (available in Pure Data or Max) and analyze the tile's ringing during the period from 15 to 25 milliseconds after each strike. The analysis is done using the covariance method, not the more popular autocorrelation method, since the analysis period is too short to compute the autocorrelation precisely.

In practice, LP analysis of order $N = 60$ points, at a sample rate of 48 kHz, gives a residual signal with about 0.9% to 1.2% of the input RMS amplitude. Further increasing the order does not decrease the error signal much, and there is an advantage to keeping the order small since the impulse response of the FIR filter we apply should be kept short in order to keep the time resolution of the whole system as fine as possible; a 60-point FIR filter already has over a 1 msec impulse response at our sample rate.

To verify the system consisting of the tile, the microphone, and the FIR filter we pass the test signal through the filter and look at the residual signal to check that each strike indeed appears as a pulse of between 1.5 and 2 msec in length, and that otherwise the residual is close to zero.

Since the existence of a mode at, say, one kHz suggest that acoustic waves take on the order of a millisecond to traverse the tile itself, a time resolution of 2 msec seems likely to be the best obtainable with the hardware used, and to improve the time resolution further will require finding objects with yet higher resonant frequencies.

Commercial tiles often have one rough side and one smooth one, with the obvious intention that the smooth surface face upward. In practice it is far better to place the tile with its rough surface facing upward. The smooth surface is better for attaching the contact microphone, and the rough surface allows for a wider variety of sounds obtained by rubbing or scraping the tile.

3.1. Can we determine the location of a strike?

At this point we may compare the LP residual signals resulting from striking the tile at different locations with an eye to finding the strike location based on the acoustic signal alone. Graphing the residuals of two strikes at the same location with the same hard object does indeed show a close resemblance, whereas strikes at distinct locations give clearly distinct residuals. If we regard the tile as an acoustic space, we are looking at the early echo pattern of the space which should indeed depend on the location of the sound source.

Unfortunately, using different strikers and/or changing the "deadness" of the strike also changes the residual, and in effect we would have to solve a blind deconvolution problem to separate the effect of different strikes from different locations; this is probably not feasible given the setup we propose here.

If we were to allow ourselves three microphones per tile instead of just one, we might intuitively expect to be able to recover both the strike location (in two dimensions) and the striking force, as three independent functions of time. Even then, however, it is far from clear in advance what the effect might be of the fact that strikes (and other forces) never occur at a single physical point, but are instead distributed over a non-zero, and variable, area.

If, indeed, we had managed to extract the location and then to estimate the force (as a function of time) applied at the point of contact—still under the considerably simplifying assumption that the contact was indeed at a single point—it might be possible to use physical modelling to make an exceedingly realistic virtual percussion instrument. But, having no direct access to the real driving force, we resort instead to designing a network that responds to it in a less literal (but hopefully still expressive) way, on the basis of desired acoustic properties, as described in the next section.

4. DESIGN OF THE REVERBERATOR

The nonlinear reverberator developed here is a particular type of a unitary delay network[1, 4], in which the unitary feedback matrix depends on the signal passing through the reverberator. Here we will restrict our consideration to the simplest possible, two-channel delay network shown in Figure 1. Here, the boxes labeled "d_1" and "d_2" are delays whose times are given in samples. The four multipliers and two adders at bottom implement a rotation matrix times a factor g:

$$R = g \cdot \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

where s and c are given by $g\cos(\theta)$ and $g\sin(\theta)$ respectively.

In the special case where the two delay lines have equal length $d_1 = d_2 = d$, this network can be analyzed by considering the two input channels and output channels

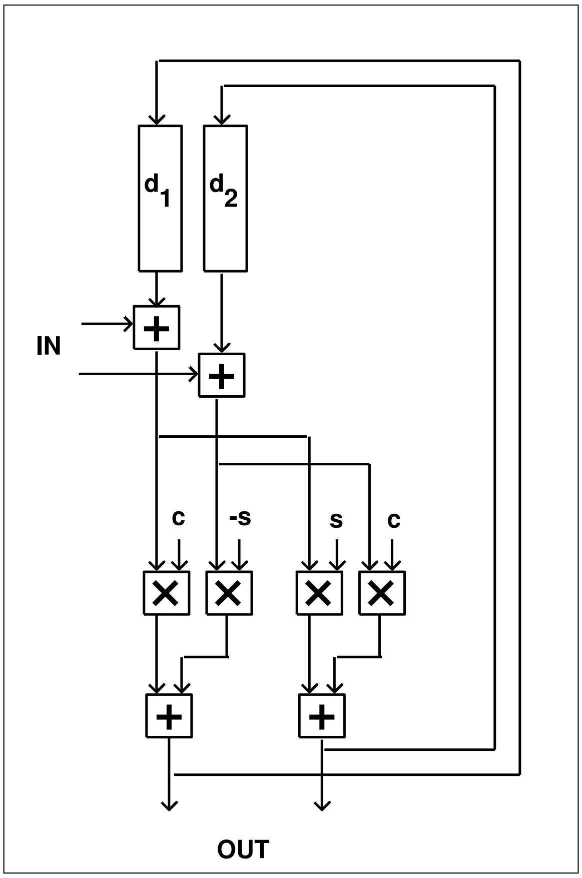

Figure 1. A two-channel unitary delay network.

as the real and imaginary part of a single complex-valued signal. The two delays become one complex-valued delay and the feedback matrix is a scalar complex multiplication by $G = g \cdot \exp(i\theta)$, and the transfer function is:

$$H(z) = \frac{1}{1 - GZ^{-d}}$$

with peaks at angular frequencies given by:

$$\omega = \frac{2\pi k + \theta}{d}, \quad k = \cdots, -1, 0, 1, 2, \cdots$$

This generalizes the recirculating comb filter (whose peaks are at $2\pi k/d$) by offsetting the peaks by θ.

If we regard one or both outputs of the network as a real monaural or stereo signal the negative-frequency peaks will fold back to frequencies at the opposite offset $-\theta$. It is useful to be able to be able to get the positively and negatively offset frequencies separately. This can be done in a way that is analogous to single-sideband modulation, using a pair of all-pass filters with transfer functions A_1, A_2 so that for positive frequencies $A_1 = iA_2$ and for negative ones, $A_1 = -iA_2$. If we multiply the output of the network (considered as a single complex number) by the value $A_1 - iA_2$, only the positive frequencies will emerge. We need only compute the real part of the product which we get by passing the left-hand side output through the filter A_1, the other by A_2, and adding.

Although this analysis only applies when the two delay times are equal, in practice this technique still works to clean up sidebands from negative-frequency components when the two delays are chosen to be different. In this case the resonant frequencies change qualitatively in the same way as for equal delay times but are no longer regularly spaced.

Returning to our virtual drum, we can now imitate the effect of stretching a drum membrane, which sometimes audibly raises the pitch of the drum as it is struck harder. To do this we simply make θ depend on the time-varying signal power in the network. Figure 2 shows a recirculating network that implements this as well as the allpass filter pair to select positive frequencies described above.

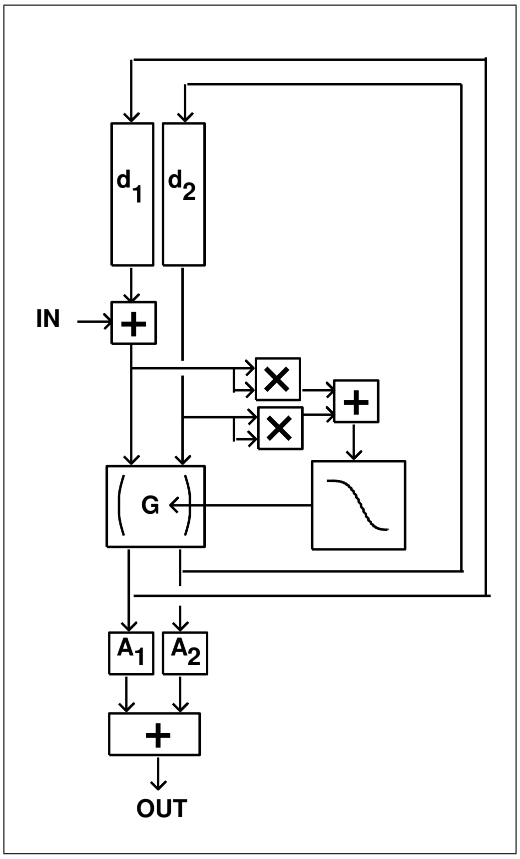

Figure 2. Measuring signal power to determine rotation angle and using allpass filters to select positive frequencies. Only one input is shown reflecting the way the network is actually used in practice.

5. DISCUSSION

The above example can be easily and widely generalized: instead of the instantaneous power (the sum of the squares of the two recirculating signals) one can use a wide variety of functions, continuous or discontinuous. One can take the value, or the absolute value, of one of the two signals ignoring the other, or the function that returns one for the first and third quadrants and zero for the others, or functions with winding numbers of two or more, or minus one. Further, the filter may be omitted (in which case the recirculating signal modulates the matrix directly, so that the frequency content splashes quickly over the audible spectrum) or tuned for a variety of other effects.

Other sources—an oscillator for example—may change the matrix independently of the recirculating signal. Also not shown in Figure 2 but interesting is to insert a series of allpass filters into one or both sides of the recirculating path[2]. Finally, there is no reason to restrict the technique to two delay lines, except that the behavior of a network with three or more delay lines is comparatively much harder to understand; and anyway, two delay lines already suffice for getting a wide variety of useful sounds.

It is noteworthy that, in this design, only one microphone is used per instrument. If one had two or more microphones on a single tile, the extra information might make it possible to deduce the striking or rubbing position on the tile. But the desirability of physical simplicity and robustness is a powerful reason that we should restrict ourselves to a single microphone. Indeed, once there are a half dozen or more instruments in a setup (as happens quickly) the issue of cabling, pre-amplifying and digitizing all those input channels becomes important and argues for using only one microphone per tile.

Overall, the physical system is designed to be as cheap and easy to replicate as possible. A typical setup might consist of four instruments and a USB converter with 4 channels of pre-amplification; the whole setup only need weigh three or four pounds (not including the computer) and can be set up, including re-running the LP analysis (in case the tiles' transfer functions have changed in transit), in a few minutes.

Although the instruments may be played with sticks, in practice it's possible to get a much wider range of sounds using fingers. Changing the part of the finger tapping the tile and/or the part of the tile struck makes very audible (and controllable) changes in the resulting sound. Rubbing and scratching also work well and make sounds that vary nicely over the surface of the tile.

In general, it is an important aspect of the instrument's playability that the tiles' response to strikes and other impulses lasts a very short time (one or two msec depending on the tile). This allows for flams and fast rolls, and for textured effects obtained by rubbing or scraping objects on the tiles.

6. CONCLUSION

The work described here is a hybrid between a physical instrument design and a new class of synthesis techniques, neither of which would warrant much notice if it were not for the synergistic way they work together to make a practical and playable percussion instrument. The sounds picked up from the tiles are well adapted as excitation signals for the virtual resonators; and the particular resonators described here react in interesting ways to various inputs in a way that a linear model, or perhaps even a nonlinear physical model that might be more faithful to a possible physical instrument, might not.

There is no easy way to *prove* the superiority of this instrument design over that of some other instrument, as would be appropriate in an engineering forum. Instead, its utility must be judged by musicians over time. The ease and low cost of constructing the instruments should make them easily accessible to anyone wishing to try them—a necessary but admittedly not sufficient condition for getting them in peoples' hands. As of this writing the instruments have only been used in low-key, unadvertised contexts that cannot yet count as artistic output, and so this paper should be regarded as a progress report and not a description of a finished product.

7. REFERENCES

[1] M. Gerzon, "Synthetic stereo reverberation (part 1)," *Studio Sound*, vol. 13, no. 12, pp. 632–635, 1971.

[2] V. Lazzarini, J. Timoney, J. Pekonen, and Vesalimaki, "Adaptive phase distortion synthesis," in *Proceedings of the International Conference on Digital Audio Effects*. Hamburg: www.dafx.de, 2009, pp. 28–35.

[3] J. O. Smith, "Physical modeling using digital waveguides," *Computer Music Journal*, vol. 16, no. 4, pp. 74–91, 1992.

[4] J. Stautner and M. Puckette, "Designing multichannel reverberators," *Computer Music Journal*, vol. 6, no. 1, pp. 52–62, 1982.