Digital Audio Effects and Simulations

By Francis Rumsey
Staff Technical Writer

It is increasingly possible either to emulate legacy audio devices and effects or to create new ones using digital signal processing. Often these are implemented as plug-ins to digital audio workstation packages, using one of the proprietary systems such as VST, DirectX, Audio Suite, or Audio Units. Papers from a session chaired by David Berners at the AES 127th Convention last year in New York covered a wide range of recent innovations in this field, including Leslie speaker, plate reverb, and guitar amplifier emulation.

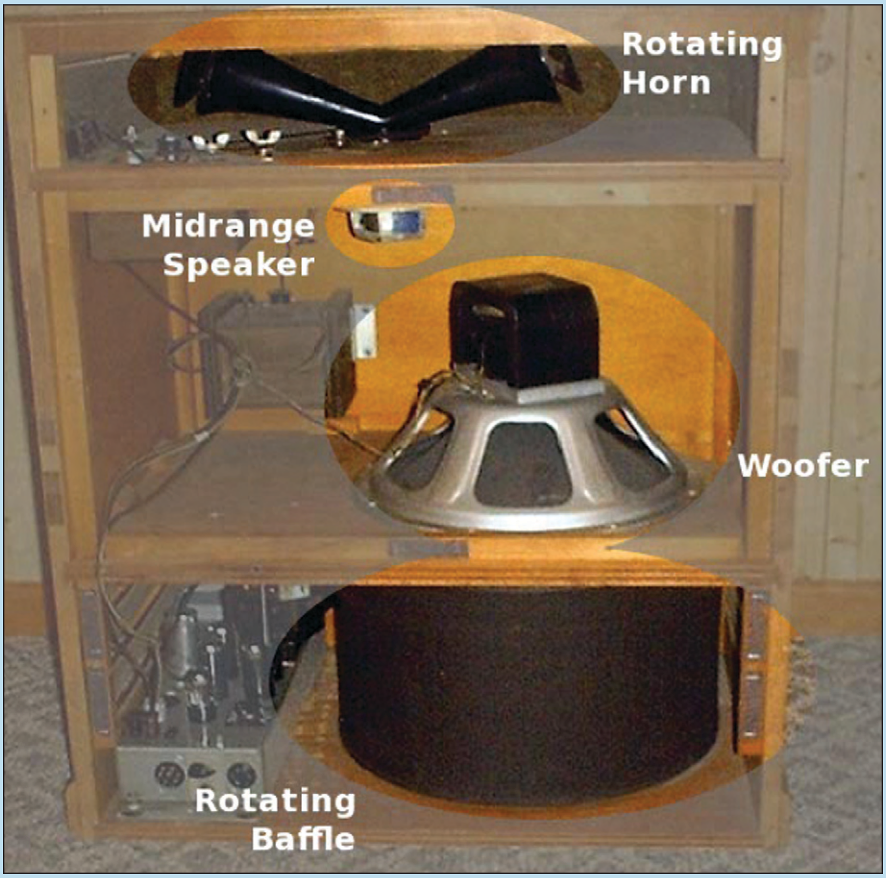


Fig. 1. The Leslie 44W loudspeaker system (Figs. 1–3 courtesy Herrera et al.)

ROLL ME A LESLIE SPEAKER

Invented by Don Leslie in 1949, the classic Leslie speaker has created an effect that defines the sound of certain bands, having been used to add a characteristic tremolo or chorus effect to the Hammond organ for many years. The whole process is achieved mechanically, which makes for an interesting challenge to emulate it in the digital domain. The effect results from rotating loudspeaker elements that have amplitude-, spectrum-, and phase-modulating influences on the audio radiated by the system. A small amount of varying Doppler shift is also said to be introduced by the system. While the mid-range horns actually rotate, the bass loudspeaker remains physically static, and a rotating baffle is used to modulate the acoustic output. A picture of a Leslie 44W model is shown in Fig. 1, in which the loudspeaker and baffle elements can clearly be seen. The two rotating elements go around at different rates to create a more interesting result, and there is a slow chorale effect as well as a fast tremolo effect.

In "Discrete Time Emulation of the Leslie Speaker" (AES 127th paper 7925), Herrera et al. describe an attempt to emulate the classic design using a novel method. In a previous attempt at emulation, they included the modeling of direct sound and time-varying first-order reflections from the inside cabinet walls, as picked up by two microphones a short distance away (see Fig. 2). Virtual source images of the reflections were added to the direct sound using filtered versions passed through a delay line, for different rotational states of the horn. While this was computationally efficient, it did not necessarily capture all the required features of the sound. In the current attempt, impulse responses picked up from the horn and baffle arrangement by a spaced pair of B&K 4006 microphones were captured for numerous rotation angles, using small manual rotations of the loudspeaker system. A sine-sweep method was used during this process. Rotation increments of about 0.9 degrees were used for the horn and 2.7 degrees for the bass baffle. As can be seen from Fig. 3, a typical impulse response contains a dense tail of reflections. The signal was split into two frequency regions using a crossover, so that the bass and mid- ➥

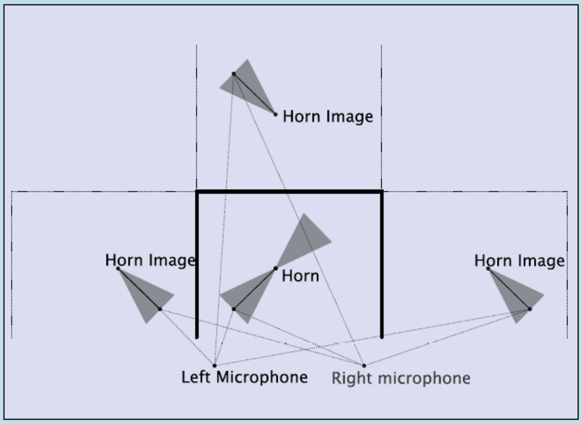
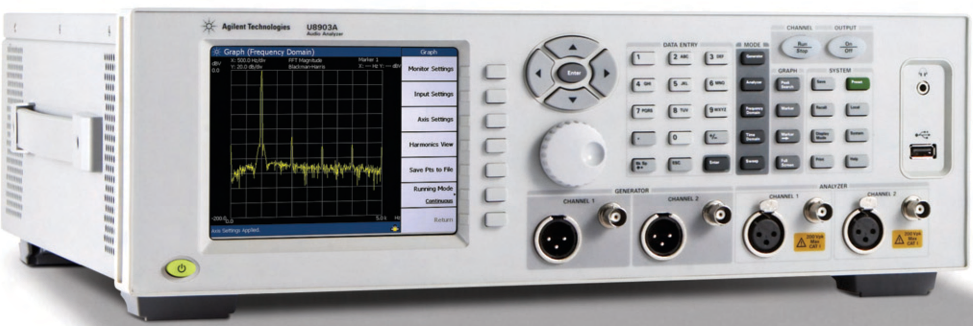


Fig. 2. Picking up first-order reflections from the Leslie horn–cabinet combination using a pair of microphones.

“I need a faster, more accurate audio analyzer.”



We knew you’d say that.

The new Agilent U8903A audio analyzer is four times more accurate and over seven times faster than its popular HP 8903B ancestor. In fact, it is among the fastest and most accurate currently in its class. It’s also a more fully capable audio analyzer, letting you make crosstalk measurements and phase measurements, FFT analysis, and graph sweep. And it will soon be upgradable to eight channels. You can’t find that anywhere else. Go online to learn more about this remarkable new audio analyzer.

Save $2,400
when you trade in your old HP 8903B for a new Agilent U8903A audio analyzer*
www.agilent.com/find/U8903Apromo

Trade in offer good for a limited time.

u.s. 1-800-829-4444 canada: 1-877-894-4414

Agilent Technologies

range components could be processed separately. The acceleration and deceleration dynamics of the rotating loudspeaker were carefully measured and implemented using separate filters in order to be able to emulate the effects of speeding up and slowing down, such as during the transition between chorus and tremolo modes.

Three different FIR filtering techniques used for processing the relevant impulse responses were compared. The first used a scaled and interpolated version of the input signal based on measurements made at numerous rotation angles, which was then added to the output. The second worked by cycling through taps in an FIR filter that accorded to the relevant rotation angle. The third worked in a similar way to an HRTF filtering technique used in binaural processing, whereby one of a small number of fixed filters was selected depending on the rotation angle, with a form of weighted mixing (or panning) between the adjacent filter outputs according to the rotation angle. The authors found that approximately 20 filters were sufficient to be able to implement convincing reproductions of the original impulse responses. (This last method is said to be very efficient in terms of memory requirements and can render multiple inputs at different angles.)

During testing it was found that the first two time-varying filter methods produced a low-level "hashy" sound at rotation speeds above 6 Hz, which was thought to be due to spatial aliasing and linear interpolation between irregularly distributed impulse responses. By projecting the impulse responses onto a sinusoidal rotation pattern, then representing them as a set of uniformly spaced (in angle) fixed filters with spatially band-limited responses, the authors were able to eliminate this unpleasant effect even at high rotation rates.

A TRIODE GUITAR AMP

In "Simulation of a Guitar Amplifier Stage for Several Triode Models: Examination of Some Relevant Phenomena and Choice of Adapted Numerical Schemes" (AES 127th paper 7929), Cohen and Hélie look into the matter of simulating the high-gain triode stage in a guitar preamplifier, in particular considering the nonlinear distortions that can give a vacuum tube-based guitar amp its unique sound. They choose to investigate the commonly used 12AX7 triode valve, also known as the ECC83, which has certain saturation distortion characteristics that are said to be valued by musicians and audiophiles alike.

Mathematical models aiming to simulate static triode behavior were tried and compared with those found on available datasheets for the valve in question. A simulation model by Norman Koren was found to be more accurate than one by Leach, as it modeled the triode behavior better for high grid and plate voltages, as well as limiting the behavior of the model to more realistic values in the case of extreme grid voltages. A couple of dynamic aspects of triode modeling were also tried, including the parasitic capacitances exhibited by the valve in question that tend to affect the frequency response of the amplifier (see Fig. 4). The combined result appeared to be successful both from a measured point of view and perceptually, and it was possible to implement it within a VST plug-in architecture.

The authors briefly discuss the effect of sampling rate on their simulations, suggesting that oversampling (at 96 kHz for example) is generally considered desirable in nonlinear simulations. Oversampling in such situations reduces the likelihood of aliasing effects becoming audible. It also seems to improve the accuracy of the various filters and the speed of convergence of an algorithm designed to solve nonlinear and differential equations.

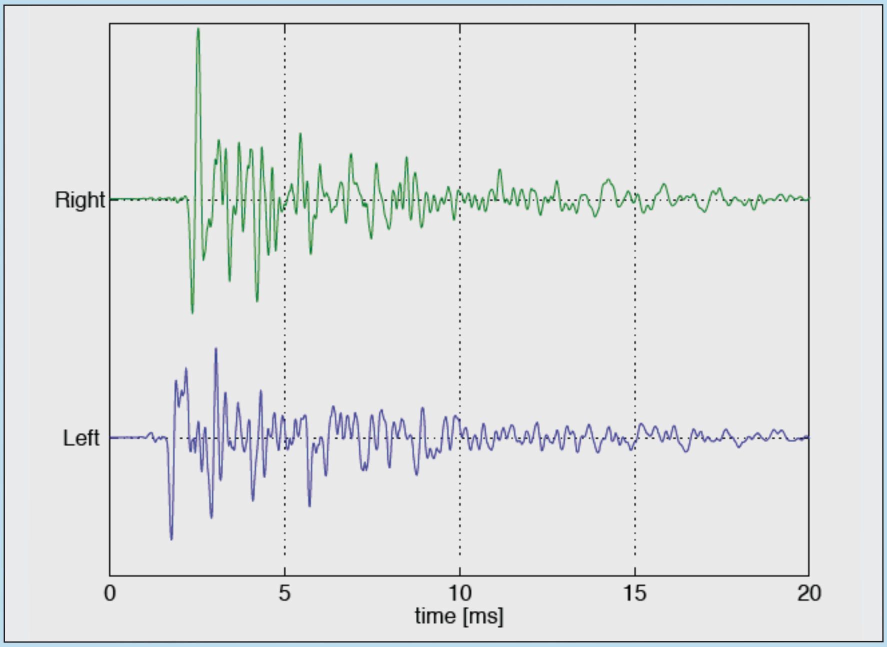


Fig. 3. Example of a stereo impulse response picked up from the Leslie horn

OVER-THRESHOLD FEEDBACK DISTORTION SYNTHESIS

Staying with the "triode sound," Tom Rutt describes an algorithm he developed to synthesize nonlinear distortion using over-threshold power function feedback (AES paper 7930). This is said to emulate closely the effect of vacuum tube triode grid limit distortion. The process acts as an instantaneous soft limiter for both the positive and negative half cycles of the signal and never clips the output signal peaks, even at the maximum peak input level.

An electronic circuit implementation of this algorithm is described, as well as a DirectX software plug-in version known as Tuboid (www.tuboid.com). In the latter a mapping array is used that relates certain input signal values to modified (distorted) output signal values, in an attempt to emulate the characteristics of the electronic circuit. While a 1:1 mapping function would require an array of 65,536 sixteen-bit integers, in fact the input signal is scaled by 256, and an array of only 256 values is used. The intervening

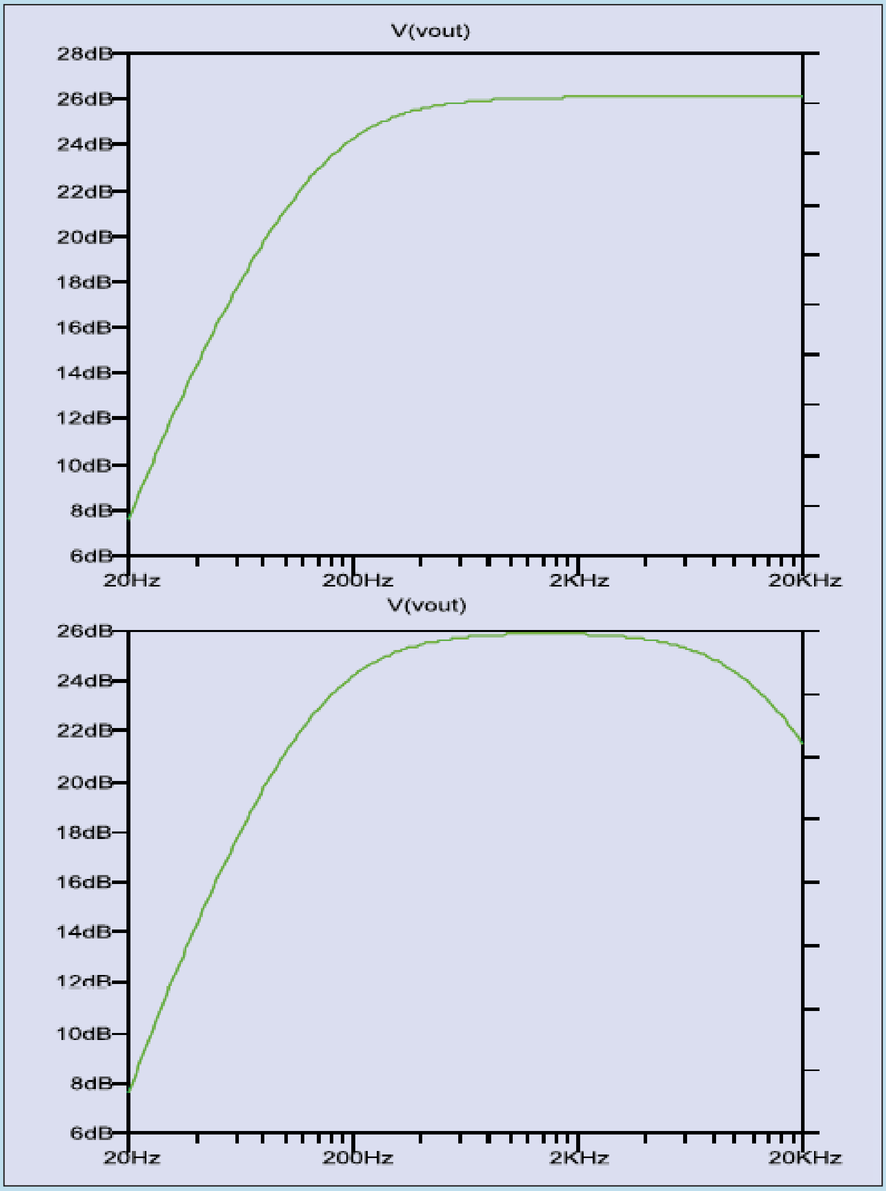


Fig. 4. Frequency response of a triode valve simulation either with (top) or without (bottom) modeling of parasitic capacitances. The roll-off in HF response can be seen in the lower plot (courtesy Cohen and Hélie).

output values are linearly interpolated. The positive- and negative-going parts of the signal are separately processed to enable the plug-in to be adjusted in a flexible fashion to emulate a variety of different kinds of both symmetric and asymmetric tube distortion. The author has made a number of audio examples available for audition at www.coastenterprises.com/OTPF/demoClips.

CLASSIC EMT 140 PLATE REVERBERATION

The EMT 140 was a classic plate reverberation system introduced in 1957 that was used widely before the days of digital reverb units. Its sound is still sought by connoisseurs of a certain type of reverb that can be useful for vocals in a mix. In "An Emulation of the EMT 140 Plate Reverberator Using a Hybrid Reverberator Structure" (AES 127th paper 7928), Greenblatt et al. consider how its characteristics can be emulated digitally, so as to avoid the need to keep and maintain a large physical unit such as this in a studio.

Essentially plate reverbs consist of a suspended, thin steel plate to which a driving transducer and one or more pickup transducers are attached. Decaying vibrations are set up in the plate that behave to some extent like room reverberation, while a damping plate located close to the reverberation plate controls the rate at which energy is dissipated (and hence the reverb time). According to the authors' analysis, the EMT 140 unit behaves in an almost linear and time-invariant fashion. It has a very fast onset transient response and a high echo density, little influenced by the damping plate, while the low-frequency decay can be controlled to range from around half a second to over eight seconds. When considering how to implement a digital emulation it is suggested that, although an obvious suggestion, the idea of storing, interpolating, and convolving impulse responses of the plate up to eight seconds is computationally very expensive. Physical models have been tried in which the mechanical characteristics of the plate are modeled numerically using partial differential equations, but so far these have been unable to emulate the rapid onset response of a real plate. They are, however, good for emulating the high-level nonlinear behavior of such a plate.

In the emulation described here, the authors explain that they have employed a hybrid reverberator that combines two elements: a short convolution stage to deal with the onset transient and a computationally efficient feedback-delay network (FDN) to emulate the tail, with access to a range of control parameters. Because the transition to a "late field" (the time after the initial impulse at which the echo density becomes perceptually dense) occurs after only a few milliseconds with the EMT plate, the initial convolution period can be short. It depends to some extent on the setting of the damping plate. In fact, in this implementation a single boundary point between the initial convolution and FDN parts is introduced at 300 ms, regardless of the damping setting. As the energy in the FDN builds up during the first few hundred milliseconds, its output can then be crossfaded into the end of the initial convolution section using a crossfade with energy matching. Measured and modeled decay traces and spectrograms for this reverb emulation are shown in Figs. 5 and 6.

A stereo emulation of the EMT plate can be implemented by introducing a second FDN signal path for the late ➥


129TH
AES CONVENTION
CONFERENCE
November 4-7, 2010
EXHIBITS
November 5-7, 2010
Moscone Center
San Francisco, CA
www.aesshow.com
Your Bridge to the Future
Audio
Engineering
Society

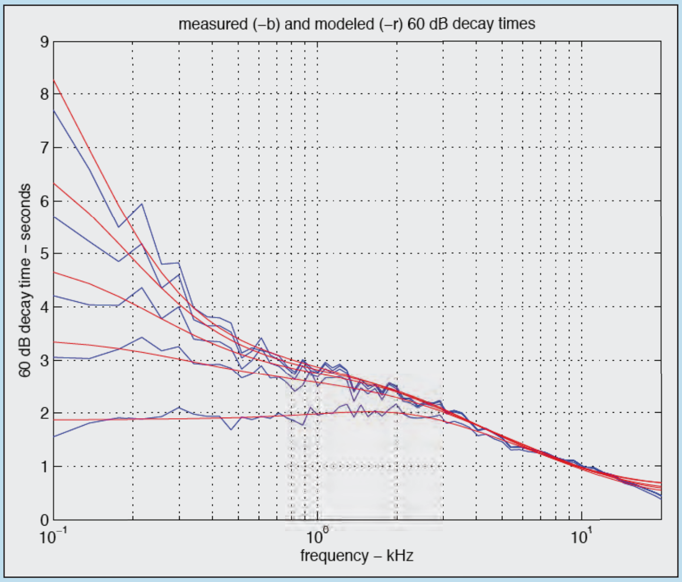


Fig. 5. Measured (blue) and modeled (red) decay traces for emulation of EMT 140 reverberator (Figs. 5 and 6 courtesy Greenblatt et al.)

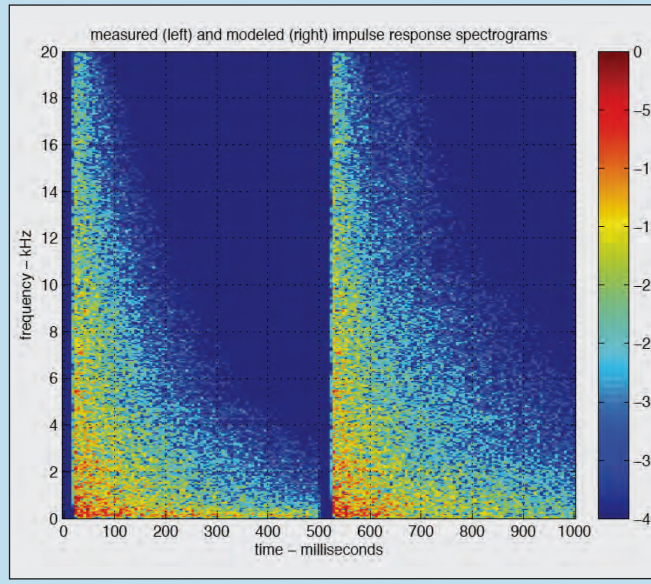


Fig. 6. Measured (left) and modeled (right) impulse response spectrograms of the EMT reverb emulation

reverberation, which will produce an output that is uncorrelated with the first. An identical window is applied during the convolution of the early part of the response for left and right channels in order to maintain the stereo image. The stereo width of the late part is adjusted to match that of the original plate by varying the panning of the two signals between the outputs so that the cross-correlation of the channels matches that of the original.

A SWITCHED CONVOLUTION REVERBERATOR

Moving into the modern age and discussing the implementation of new effects algorithms, Lee et al. continue the reverberation theme in relation to a low-cost algorithm for use in mobile devices in 'The Switched Convolution Reverberator" (AES paper 7927). Conventional convolution and FDN-based reverberators need considerable amounts of processing and memory, which makes them unsuitable for mobile devices where these resources are at a premium. Consequently the authors propose the reverberator structure shown in Fig. 7, which consists of a recursive filter (comb filter) driving a process that convolves the comb filter output with a short, exponentially decaying noise sequence. The reverberator's equalization is controlled using low-order IIR filters at the input and in the feedback path of the comb filter. The impulse responses at different points in the chain are shown in Fig. 8, where it can be seen that the output of the comb filter is a simple decaying train of pulses, whereas that of the filter is an infinitely long decaying noise sequence. The time constant of the comb filter is supposed to be the same length as the noise sequence, which in this case is only in the region of 100 ms, leading to low computation and memory requirements. This structure creates a decay that can have similar properties, both spectrally and in terms of echo density, to typical late reverberation. The authors found that this basic structure worked well for stationary signals but less well for transient signals, having audible periodicity.

Because the noise sequence is repeated in a cyclical fashion, as the authors found above, there can be audible periodicity in the perceived reverberation decay. For this reason the noise sequence has somehow to be varied. In a switched convolution reverberator a different noise sequence can be switched in at every period of the comb filter, which helps to reduce periodicity. However, a basic switched convolution reverberator was found to have the opposite problem to the simple filter structure described above, in that it performed quite well with transient signals but less well for stationary signals. It was clear that time-varying artifacts in the reverberation had to be

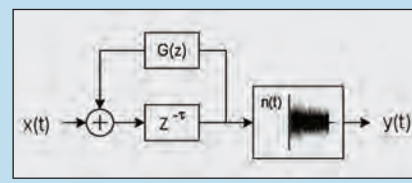


Fig. 7. Comb filter reverberator structure using a short noise sequence output filter (Figs. 7–11 courtesy Lee et al.)

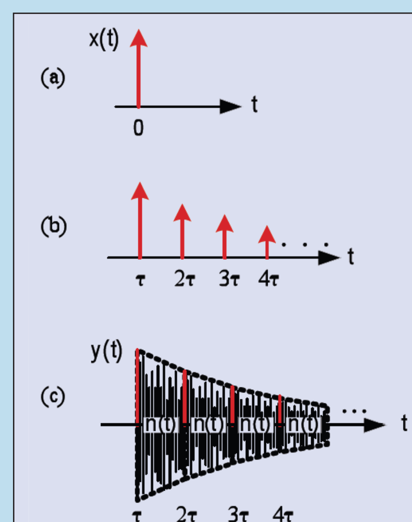


Fig. 8. Impulse response of the reverberator structure shown in Fig. 7: (a) input signal x(t), (b) comb filter output, (c) reverberator output y(t).

minimized, and this became a key goal of the authors' research. One proposed solution, as shown in Fig. 9, employed a crossfade between the noise sequences to be convolved, having a length of twice the comb filter impulse peri- ➡

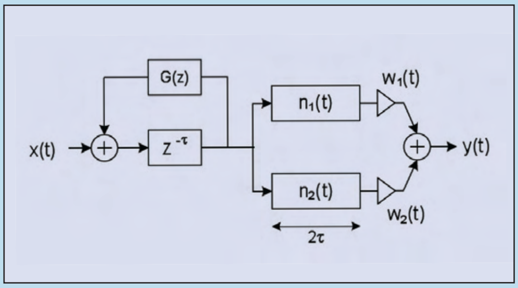


Fig. 9. Switched convolution reverberator structure with cross-faded noise sequences

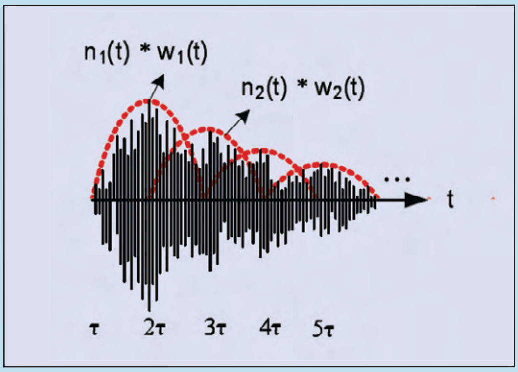


Fig. 10. Impulse response of the crossfaded switched convolution reverberator shown in Fig. 9

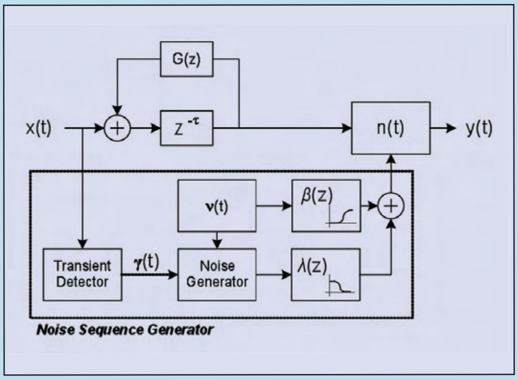


Fig. 11. Switched convolution reverberator using frequency-dependent noise update

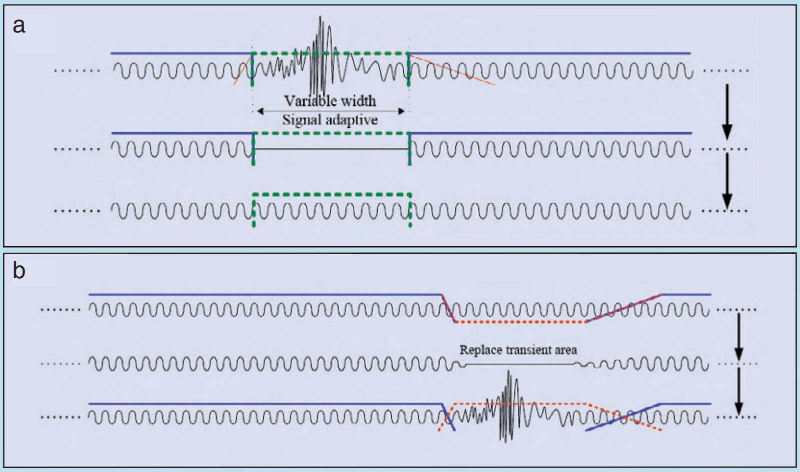


Fig. 12. Transient removal and interpolation in a novel time-stretching algorithm: (a) The transient is detected and windowed, then it is subtracted from the signal, finally the gap is interpolated; (b) upon resynthesis the stationary signal is stretched, then a gap is prepared for the transient by multiplying with the inverse window used for its original extraction, then the transient is added back in (Figs. 12 and 13 courtesy Nagel and Walther).

ods. The system's impulse response is shown in Fig. 10. The perceived result was found to represent an improvement; however, it cost more in terms of computation and memory requirements than the basic structure described earlier.

Alternative structures were tried that used only a single convolution stage, based on the idea of gradually changing the noise signal over time so as to reduce artifacts. A leaky integrator concept was used to control the rate at which the noise sequence was changed, and proposals were made to vary this rate depending on the time-varying nature of the input signal. A version based on this idea is shown in Fig. 11. A transient detection algorithm could be used to assist the process of selecting the update rate of the noise sequence. For complex music signals it was found to be helpful to make the leaky integrator's time constant frequency-dependent. Sparse noise sequences known as velvet noise were tried to good effect, these having less dense impulse responses but similar perceptual characteristics to Gaussian noise.

HANDLING TRANSIENTS IN TIME-STRETCHING ALGORITHMS

Time-stretching algorithms are used to alter the speed of audio signals without altering their pitch. Ideally this process should avoid side effects. While sustained signals can be treated with relative success using various types of overlap-and-add algorithms, transient components are not always so easy to deal with. In "A Novel Transient Handling Scheme for Time Stretching Algorithms" (AES 127th paper 7926), Nagel and Walther explain that it is usually necessary to extract the transient components and only process the sustained elements of a signal, adding the transients back in after processing. In attempting to address this problem they tackle the challenge of stretching a sustained note such as that from a pitch pipe, combined with a highly transient sound such as that from a castanet.

The authors cite a 1994 paper on time stretching by Quatieri et al. in which the algorithm attempts to preserve the overall envelope of the signal in its time-stretched version. However, this leads to the side effect that a time-dilated percussive event will decay more slowly than the original. In musical signals, they argue, it is preferable to preserve the envelope of transient events. In the authors' new method, transients are extracted from the composite signal and replaced with a stationary interpolated signal to fill the gap, as shown in Fig. 12. (The reason for this is to avoid the artifacts that can arise when stretching stationary signals that have gaps in them.) The resulting quasistationary signal is then well suited to time stretching, and the transients are subsequently used to replace parts of the interpolated signal after it has been stretched. A phase vocoding technique was used to stretch the stationary parts of the signal. As shown in Fig. 13, the signal was FFT analyzed in grains of around 10 ms and then resynthesized at the new time scale using a different hop size and overlap-add processing (in this example, a stretching factor of two is shown).

The main challenges of this method were to detect transients accurately and to achieve a perceptually satisfactory interpolation of the stationary part of the signal. In particular there was the problem of what would happen if the

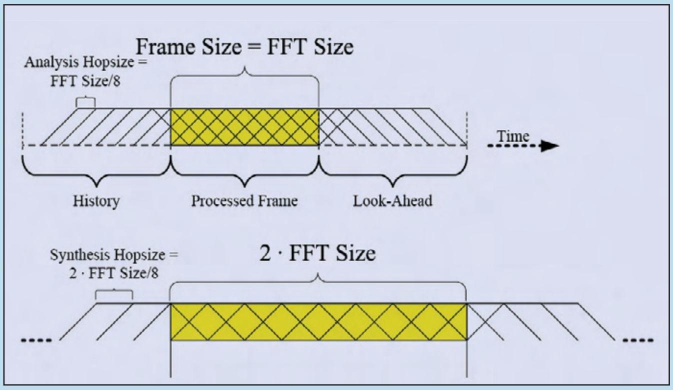


Fig. 13 The phase vocoder algorithm used for time stretching resynthesizes the original signal by using a hop size that is different to that used in the original FFT analysis (here shown as two times longer)

stationary signal's composition changed substantially over the duration of the transient, which the authors proposed to address by using forward and backward prediction with a cross-fade. Detection of the ends of transient sections was complicated by the fact that the reverberant decays of the transient part typically merge with the following stationary part, and this required some careful handling. Results were promising, but did not appear to outperform state-of-the-art software currently available. Interestingly, some listeners seemed to prefer versions of the time stretch in which the reverb of the transient was stretched along with the stationary part of the signal, which rather contradicted the authors' original assumption that it would be preferable to treat the transients and their associated decays as an unchanged and combined entity. They concluded that more insight into listener preferences are needed. The novel approach is promising, they said, and its essential principles had been demonstrated in the case of a special signal. Anyone interested in hearing the results of this approach are invited to ask the authors (Frederik.Nagel@iis.fraunhofer.de) for sample audio files.

SUMMARY

Using today's signal processing tools one can both emulate the sounds of yesteryear and implement new effects that were not possible in the days of analog processing. The sound of classic systems such as the Leslie speaker and tube guitar amp can be made available in the world of plug-ins, for use in modern digital audio production applications. Although plug-ins may not have some of the raw physical appeal or nostalgic value of the classic hardware they emulate, they can bring the sounds of yesterday to a contemporary audience in a relatively convenient fashion.

Editor's note: The papers reviewed in this article, and all AES papers, can be purchased online at www.aes.org/publications/preprints/search.cfm and www.aes.org/journal/search.cfm. AES members also have free access to past technical review articles such as this one and other tutorials from AES conventions and conferences at www.aes.org/tutorials.