SOUND SYNTHESIS AND COMPOSITION WITH COMPRESSION-CONTROLLED FEEDBACK

Bret Battey
Music, Technology and Innovation Research Centre
De Montfort University
Leicester UK
bbattey@dmu.ac.uk

ABSTRACT

This paper introduces a method of sound synthesis that is based on the use of automatic gain control (AGC) in a time-delayed feedback loop. The approach, which the author calls "Compressed Feedback Synthesis" (CFS), can be conceptualised as a special expansion of a generalised comb filter, where feedback gain can be unity or greater. The system can be expanded with additional processing in the feedback loop to create a highly flexible and sensually engaging sound materials. The use of CFS in the author's audiovisual composition *Sinus Aestum* will be discussed, including specific solutions to the challenging of controlling such a system compositionally.

1. INTRODUCTION

A feedback loop around a delay is a fundamental construct in digital signal processing and is commonly known as a recursive comb filter. The feedback gain (g) is constrained to the range 0<=g<1 in order to avoid runaway. The fundamental pitch of the filter (the reciprocal of the delay time) and its harmonics are emphasised in the resulting comb-like frequency response. [1]

Expanding this model with the placement of a low-pass filter in the loop provides is the essence of the classic Karplus-Strong plucked string algorithm. [2] This in turn can be seen as a foundational model in the domain of physical modelling via digital waveguide synthesis. [1] In other words, the feedback delay, though simple, is very powerful.

Seeking simple means for creating powerful sound design, in the mid-1990s the author began exploring how one might use a computer to control electroacoustic feedback between a speaker and a microphone in musically useful ways. In the resulting experiments, automatic gain control (AGC, also referred to as amplitude compression or just compression) was used to ensure that the feedback of the system did not enter into runaway. The distance of the microphone from the speaker provided control over delay length. In a sense, the system was a type of electroacoustic comb filter, in which gain could be set at unity or greater, and hence could be continuously self-sustaining once energy entered the system.

Christopher Burns and Matthew Burtner have pursued similar lines of thought, conceiving of electroacoustic feedback networks between speakers and microphones in terms of digital waveguides. Instead of using AGC to control the feedback gain, they used waveshaping with nonlinear functions to provide soft clipping, inspired in part by Charles Sullivan's physical models for the electric guitar. In this case, the spectral alterations provided by the waveshaping were seen as useful, contributing to unusual sonic results. [3,4]

The author's own explorations ultimately turned away from the electroacoustic domain towards experimentation directly in digital synthesis. AGC was inserted directly into a comb filter, and then other filters and signal processing were added to the loop. The overall aim was to create a richness and beauty of sound that transcended the conventional signatures of feedback systems and enabled distinctive compositional results.

2. CORE COMB+AGC MODEL

This simplest CFS model places the AGC in the comb filter loop and adds a DC block filter to prevent offsets from accumulating in the feedback loop (Figure 1).

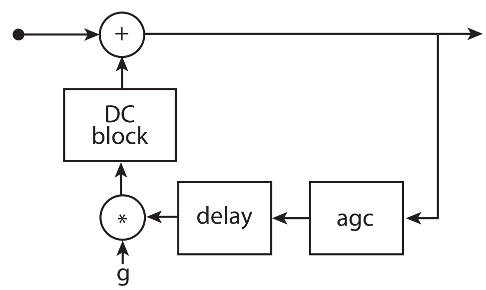

Figure 1. Addition of automatic gain control to a recursive comb filter.

The AGC was implemented with SuperCollider's `Compander` object. `LocalOut` and `DelayL` (delay line with linear interpolation) were used to implement the delay. (In this case, one must subtract the signal vector duration from the delay time, since one vector is used in passing the signal from `LocalOut` back to `DelayL`.) The `LeakDC` ugen was used, which implements a DC Blocker filter described by Julius O. Smith. [5]

The relevant AGC parameters are amplitude threshold, compression ratio, attack time, and release time. The single `R` parameter of LeakDC has significant impact on the behavior of the system. As `R` approaches 1, the notch at DC gets narrower and provides faster tracking of DC, while the impulse duration increases. While `R` decreases, the stop band broadens, attenuating more low frequencies, while also providing a gentle boost of the high frequencies. [6]

Figure 2 depicts the behavior of this CFS system with a continuous 100 Hz sine tone input at -12 dB FS, running at 48 kHz, 24 bits. The AGC settings are threshold -24 dB FS, compression ratio 0.25, attack 1 ms, decay 10 ms, and DC coefficient 0.995. The feedback frequency is 100 Hz and feedback gain is 4. The 1st, 3rd and 5th harmonics are present in the result.

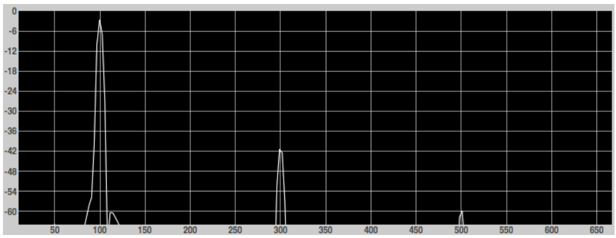

Figure 2. Sonogram of basic comb+AGC loop with continuous 100 Hz sine input.

When the input tone is removed, the system destabilizes. All of the harmonics of the fundamental begin to appear, starting with the lowest and, over time, proceeding to the highest. The amplitudes of all harmonics vary continuously, generally providing a combed spectrum with exponential decay. (Figure 3) Peak frequency modulates unpredictably from the Nyquist frequency to as low as the neighbourhood of 5000 Hz. The spacing of the combs also modulates unpredictably.

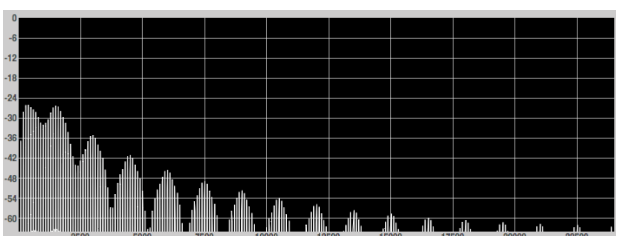

Figure 3. Sonogram of basic comb+AGC loop several seconds after removal of sine input.

The impact of the `R` parameter of `LeakDC` can be readily demonstrated at this point. With `R` set to 0.99999, the fundamental stabilizes as the strongest tone, and the combing intensity diminishes (Figure 4).

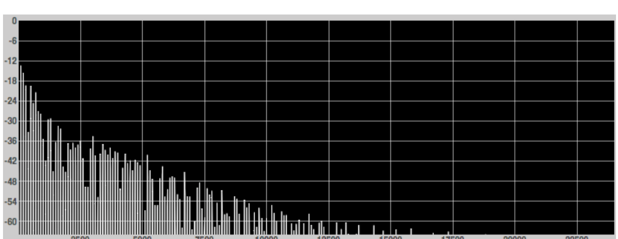

Figure 4. Sonogram of basic AGC loop after removal of sine input and shifting of the DC blocker coefficient to 0.9999.

With `R` set to 0.99, slow amplitude pulsing of comb formations across the spectrum continues. The lower end of the spectrum is significantly diminished and the upper end is emphasized (Figure 5).

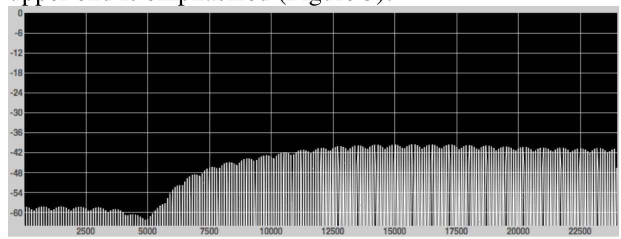

Figure 5. Sonogram of basic comb+AGC loop after removal of sine input and shifting of the DC blocker coefficient to 0.99.

Alteration of the AGC attack and decay times can impact the relative spectral distribution of energy and the rate of modulation of the combing effects. In general, a faster attack time will result in faster modulation of the combing and a steeper rolloff of the higher frequencies — and a lower amplitude overall. A longer release time slows the rate of the modulation but increases the emphasis on the higher end of the spectrum. Therefore, one often has to balance a desire to avoid overt amplitude pulsing with the desire to maintain a spectral balance that avoids excessive high-end emphasis.

3. COMB+AGC WITH PITCH SHIFTERS

By inserting pitch shifters into the loop, the sonic potential expands dramatically (Figure 6).

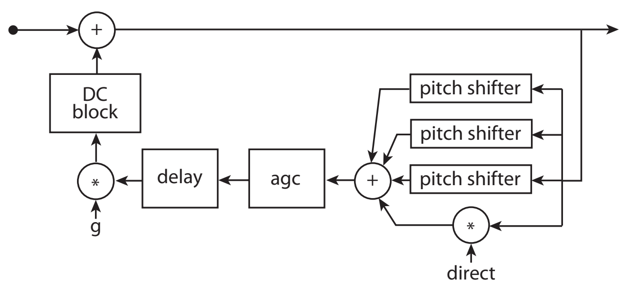

Figure 6. Addition of parallel pitch shifters to the comb+AGC loop.

The pitch shifting was implemented with the SuperCollider `PitchShift` class, a time-domain granular pitch shifter. The parameters are resampling ratio, window size, randomization of resampling ratio, and randomization of grain start time. Each of three parallel pitch shifters has its own gain control, plus the `direct` parameter controls the gain of a signal path that bypasses the pitch shifters.

The following setup demonstrates some core aspects of the system: feedback gain 4.0, feedback frequency 100 Hz, threshold -24 dB FS, compression ratio 0.25, attack time 0.001 sec, decay time 0.5 sec, and DC coefficient 0.995. The pitch shifting ratios are set to 2.0, 1.01, and 0.5, with a window size of 1.5 sec and pitch and time randomization turned off. The direct gain is set to 1.0 Feeding the system a continuous 100 Hz sine at -

12 dB FS, the pitch shifters provide upward and downward shifts which echo through the feedback loop like an audio hall of mirrors. The system gradually exhibits higher harmonics and settles into a relatively stable state, with only minor, and consistent, pulsation of the harmonics, sounding something like an electronic organ — with particularly strong peaks at octave displacement series built from 50 Hz and 150 Hz (Figure 7).

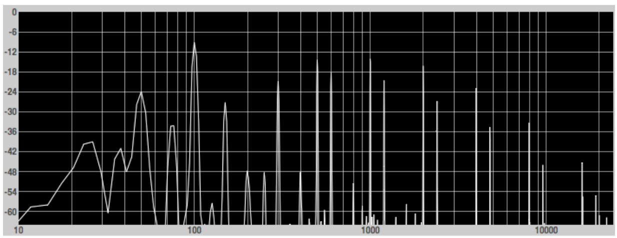

Figure 7. Stable state of an example setup with pitch shifters, with full direct gain and no randomization (log scale).

The sound blossoms when we introduce some randomness to the pitch shifting (pitch dispersion = 0.005 and time disperson = 0.05). The remaining dominant peaks are now more clearly octaves above and below 100 Hz, and the upper frequency rolloff is much more rapid (and comfortable to the ear). The amplitude of the harmonics modulates more rapidly, imparting a dynamic liveliness to the sound. If running in stereo, the decorrelation of the stereo field generated by having two randomized copies provides a spacious soundfield.

Removing the sine input and setting the direct gain to zero (leaving only pitch-shifted signals) plus slightly detuning the unity pitch shifter (to a resampling ratio of 1.01) dramatically transforms the sound. The space between the harmonics begins to be filled. (Figure 8) We hear multiple pitch centers inhabiting a field of noise. The line the between pitch and noise has been dramatically blurred, to surprisingly sensuous effect. Due to the upward weighting of the detuned pitch shifter, the pitch trails seems to form an eternal upward climb. The continual transformation of the sound field is remarkably engaging. We are certainly far from the gritty and harsh sonic signatures normally associated with feedback systems.

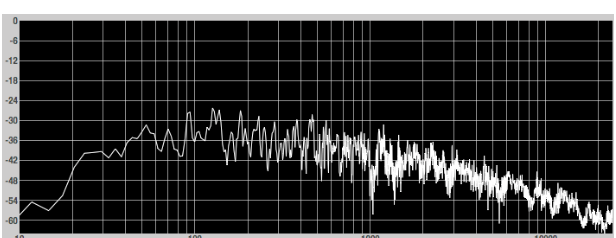

Figure 8. Pitch shifter setup with no direct signal path or sine input, and detuned unity pitch shifter (log scale). "We aren't in Kansas anymore."

Manipulations of the full available parameter set for this configuration yields a surprisingly wide range of perceptually engaging potentials beyond those demonstrated here.

4. MODEL USED IN *SINUS AESTUM* (2009)

Figure 9 shows the block diagram of the synthesis configuration used in the author's audio-visual composition *Sinus Aestum* (2009). The low pass filter is SuperCollider's LPF, a second order filter. It provides an initial rolloff of the high end to help produce a less strident and more naturalistic sound. However, the cutoff can also be modulated for other specific compositional purposes.

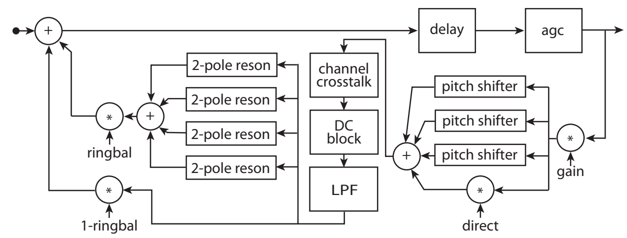

Figure 9. *Sinus Aestum* synthesis configuration.

A bank of four SuperCollider Ringz filters (two pole resonators) can be engaged to create "centers of gravity" in the spectrum. The ringbal control determines how much of this filtered signal is used in the feedback loop.

To create a quadraphonic audio field, four copies of the configuration ran in parallel. The channel crosstalk component allows a varying amount of signal to be fed from one copy of the synthesis to its neighbor. With full crosstalk enabled, the soundfield becomes monophonic. With crosstalk disabled, a fully decorrelated soundfield opens up around the listener. The full spectrum between these extremes was utilised compositionally.

Not pictured is a high pass filter on the final output, used to manage some of the problematic sub-bass fluctuations that can arise in the system, particularly when making rapid changes in parameter settings (as was explored in the middle section of the piece).

The initial system input, or stimulus, was a sine tone with frequency set to the fundamental of the feedback (reciprocal of the delay time) plus a second harmonic at ¼ amplitude.

5. CONTROL METHOD

The Figure 9 configuration has 37 control parameters. While some of these rarely require active manipulation, the majority can have significant impact on the sound. Controlling this many parameters towards compositional ends is a challenge, particularly in a context where the system behavior is complex, unpredictable, and may take time to respond to or stabilize after any change.

The author chose to compose *Sinus Aestum* through continual manipulation of the parameters of CFS, so that the sound proceeds as one process from the beginning to the end, without cuts or edits. The system can replicate essentially the same behavior on repeated passes when fed the same inputs and parameters controls.

MAX/MSP was configured to serve as the compositional controller, sending OSC messages to trigger and manipulate the CFS running in SuperCollider. Then a series of routines containing algorithms and control calls were setup in MAX/MSP, called one after another to generate the whole piece. (SuperCollider could have been used for the control, but the author chose to take advantage of his greater familiarity with the relevant mechanisms in MAX/MSP.)

So, though the CFS runs in realtime, the process of composing was very much a non-realtime process. Since the piece starts with a stimulus at the beginning and continues as one process to the end, composing after getting past the beginning of the piece can be problematic: in theory, to properly hear a change made at (for example) two minutes into the piece, one has to restart the piece from the beginning. Since this is too time consuming to be practical, it proves necessary to snapshot the parameters at certain key points in the piece. One can then call up that snapshot, stimulate the system, wait for it to stabilize, and then start the routine sequence at the appropriate point. This generally provides a sufficiently close approximation to be compositionally useful.

When one chooses to make significant changes to an earlier part of the piece, this can shift the state of the system to such a degree that the later control segments no longer create the expected results. One does develop a feel for what kind of system states are reliably reproduced, and those can be considered key states. It is relatively safe to make changes to routines prior to such key states.

6. EXAMPLE PASSAGES

We will now consider some specific examples from *Sinus Aestum*. An initial subpatcher turns on the synthesis and sets the initial parameters. This includes setting the DC coefficient to 0.99, ratio to 0.25, attack time to 0.001, and decay to 0.5. The resonators are tuned to approximately C1, G1, C2, and A4. Pitch randomness is 0.0075 and time randomness 0.05 seconds, with a grain size of 1.5 seconds. The pitch shifter ratios are 2.83, 5.65, and 8.49. The system starts the stimulus tone and begins manipulating parameters. The initial routine turns up the gain on the stimulus and modulates the sine and feedback frequency up exponentially from 200 Hz to 225 Hz over eight seconds, moves ringbal slowly to 0.023 to just begin engaging the resonators, and brings feedback gain up to 3.5 over 12 seconds. At that point the low pass filter center frequency drops from 5 kHz to 3 kHz over 8 seconds, and the feedback frequency changes slowly from 225 Hz to 1478 Hz. Three seconds later, this routine is completed, and control is passed to the next routine.

The second routine starts 300 ms later. The pitch shift ratios are modulated by a pattern generated from Lehmer's Linear Congruence formula (an iterated map) [7]. The resonators are set to a series of multiples of 60 Hz. The feedback frequency is changed smoothly between presequenced settings, using Bézier spline curves for naturalistic shaping [8]. On top of the basic generative structure, the author inserted numerous additional event controls to create specific effects. The result is a series of waves formed through flow between relatively pitched and relatively noisy plateaus, building up to the first climax of the work.

This CFS system can sometimes provide uncomfortably strong high-frequency emphasis and relatively little bass, resulting in listener fatigue. The author finalized the mix with a multiband compressor to address these issues.

7. CONCLUSIONS

Though difficult to analyze with rigor, synthesis systems involving AGN in a feedback delay system can provide highly flexible, diverse, and sensuous sound palettes. Using computer algorithms to control the many parameters of such as system is one effective approach to enable distinctive compositional results, though it offers its own distinct challenges.

8. REFERENCES

[1] Roads, C. *The Computer Music Tutorial*. MIT Press, Cambridge, MA, 1996.

[2] Karplus, K. and Strong, A. "Digital Synthesis of Plucked String and Drum Timbres". *Computer Music Journal*, v.7 n.2: 43–55, 1983.

[3] Burns, C. "Emergent Behavior from Idiosyncratic Feedback Networks". *Proceedings of the International Computer Music Conference, Singapore*, 2003.

[4] Burns, C. and Burtner, M. "Recursive Audio Systems: Acoustic Feedback in Composition". *Leonardo Electronic Almanac*, v.12, n.2, 2004.

[5] Smith, J. O. "DC Blocker". *Introduction to Digital Filters with Audio Applications*, https://ccrma.stanford.edu/~jos/filters/DC_Blocker.html, 2007.

[6] Smith, J. O. "DC Blocker Frequency Response". *Introduction to Digital Filters with Audio Applications*, https://ccrma.stanford.edu/~jos/filters/DC_Blocker_Frequency_Response.html, 2007.

[7] Battey, B. "Musical Pattern Generation with Variable-Coupled Iterated Map Networks". *Organised Sound,* v.9 n.2: 137–150, 2004.

[8] Battey, B. "Bézier Spline Modeling of Pitch-continuous Melodic Expression and Ornamentation". *Computer Music Journal*, v. 28 n.4: 25–39, 2004