SPECTRAL DELAY FILTERS WITH FEEDBACK AND TIME-VARYING COEFFICIENTS

Jussi Pekonen and Vesa Välimäki

Department of Signal Processing and Acoustics
TKK Helsinki University of Technology
Espoo, Finland
Jussi.Pekonen@tkk.fi
Vesa.Valimaki@tkk.fi

Jonathan S. Abel and Julius O. Smith

CCRMA
Stanford University
Stanford, CA 94305, USA
abel@ccrma.stanford.edu
jos@ccrma.stanford.edu

ABSTRACT

A recently introduced structure to implement a continuously smooth spectral delay, based on a cascade of first-order allpass filters and an equalizing filter, is described and the properties of this spectral delay filter are reviewed. A new amplitude envelope equalizing filter for the spectral delay filter is proposed and the properties of structures utilizing feedback and/or time-varying filter coefficients are discussed. In addition, the stability conditions for the feedback and the time-varying structures are derived. A spectral delay filter can be used for synthesizing chirp-like sounds or for modifying the timbre of arbitrary audio signals. Sound examples on the use of the spectral delay filters utilizing the structures discussed in this paper can be found at http://www.acoustics.hut.fi/publications/papers/dafx09-sdf/.

1. INTRODUCTION

Spectral delay filtering (SDF) is an audio processing method in which different frequencies of a signal are delayed by different amounts. For example, an impulse can be transformed to a "chirp" by such a filter. SDFs can be implemented by delaying the desired frequencies of the short-time Fourier transform (STFT) of each signal frame to one or more frames [1, 2]. However, such a frequency-domain approach does not easily provide a delay that is continuously variable as a function of frequency when the delay is not introduced as an integer multiple of the STFT hop size. Obviously, interpolation between successive STFT frames could be performed, but this would increase the number of required operations. In addition, when a large delay is desired, the frequency resolution of the STFT should be high, which trades off with increased memory requirements of the algorithm. Recently, an alternative implementation utilizing a cascade of first-order allpass filters was proposed [3], and in that approach the delay produced by the filters is inherently smoothly continuous as a function of the frequency. This paper discusses extensions to this approach by introducing a spectral delay filter structure with a feedback path and time-varying filter coefficients.

Usually, filtering an audio signal with a time-invariant allpass filter does not have a major effect on the timbre because it does not change the magnitude response of the signal. Yet, if a high-order allpass filter is constructed by cascading several low-order allpass filters, each introducing a mild phase shift, the overall filter has a phase shift that is the sum of the phase shifts of the individual low-order filters. When the number of low-order allpass filters is high enough their cascade can have a long, chirp-like impulse response. Now, when audio and music signals are processed with such a filter impressive changes maybe obtained that are similar to the STFT SDF methods of [1, 2].

Allpass filters are commonly used in audio and music processing, and their applications range from effects processing to the simulation of physical phenomena in musical instruments and physical environments. In effects processing, the allpass filter has been used in the simulation of reverberation [4] and a spring reverberator [5], in digital phasers [6], in shelving filters and equalizers [7], and recently also as a distortion effect [8]. The allpass filters can be used as a fractional delay, which is often needed in audio processing to fine-tune the length of a delay-line [9, 10] or in a frequency-warped filter aimed at imitating human hearing or modeling acoustic responses at low frequencies [11, 12, 13, 14]. Kautz filters, which practically produce the same result as warped filters but with a smaller model order, use a cascade of non-identical second-order allpass filters [15].

Due to the highly dispersive, i.e. frequency-dependent, delay introduced by an allpass filter, they have been used in the inharmonic waveguide synthesis of piano tones [16, 17, 18, 19], in the simulation of inharmonic responses of spherical resonators [20], and the synthesis carillons [21]. Interesting effects are obtained when the the allpass filter is modulated at audio rate [22]. Additionally, high-order allpass filters can be used to implement the accurate time delay needed in a multi-notch filter, which cancels harmonics of a musical signal [23, 24].

The remainder of this paper is organized as follows. In Section 2, the properties of a cascade of first-order allpass filters and its relation to the chirp-like impulse response of the spectral delay filter are reviewed. In addition, a new IIR filter design for the amplitude envelope equalization of the impulse response is presented. Section 3 discusses the SDF structure utilizing feedback, and Section 4 presents the properties of a time-varying SDF. The stability conditions for both structures are also derived. Finally, Section 5 concludes this paper.

2. CASCADED FIRST-ORDER ALLPASS FILTERS

The transfer function of a first-order allpass filter is given by

$$A(z) = \frac{a_1 + z^{-1}}{1 + a_1 z^{-1}}, \tag{1}$$

where a_1 is the allpass filter coefficient. Filter stability requires that $|a_1| < 1$. The transfer function of a cascade of M identical

The work of Vesa Välimäki was performed as a visiting scholar at CCRMA, Stanford University. This work was partly supported by the Academy of Finland, project numbers 126310 and 122815.

first-order allpass filters can be expressed as

$$H(z) = A^M(z) = \left(\frac{a_1 + z^{-1}}{1 + a_1 z^{-1}} \right)^M . \tag{2}$$

It can be shown that the cascade of allpass filters is also allpass. The authors call the cascade of many identical allpass filters and an optional equalizing filter a spectral delay filter (SDF). Figure 1 shows the block diagram of a such filter.

2.1. Group Delay and Chirp-Like Impulse Response

The group delay of a filter is defined as the negative derivative of the filter's phase response $\phi(\omega)$ with respect to frequency, i.e.

$$\tau_{\mathrm{g}}(\omega) = -\frac{\mathrm{d}\phi(\omega)}{\mathrm{d}\omega}. \tag{3}$$

For the first-order allpass filter the phase response is given by

$$\phi(\omega) = -\omega + 2\arctan\left(\frac{a_1 \sin(\omega)}{1 + a_1 \cos(\omega)} \right). \tag{4}$$

Now, the group delay of the first-order allpass filter can be expressed in closed form, as

$$\tau_{\mathrm{g}}(\omega) = \frac{1 - a_1^2}{1 + 2a_1 \cos(\omega) + a_1^2}. \tag{5}$$

Figure 2 shows the group delay with various values for the coefficient a_1. The plot is unconventional in that the horizontal and vertical axes have been interchanged. That is, frequency lies along the vertical axis while group delay is plotted along the horizontal axis, instead of the other way around as is typical. For monotonic group-delay functions, as occurring in Figure 2, such a plot can be interpreted as a plot of the inverse function. This type of group-delay plot facilitates visualization of the SDF output signal spectrum and will be used in what follows. As will be seen, the plot can be interpreted as an idealized spectrogram of the impulse response of an allpass filter having that group delay.

It can be seen that the group delay of first-order allpass filters is quite small at all frequencies. For instance, when $a_1 = -0.6$ the delay at low frequencies is only four samples. This would correspond to about 0.1 ms when a sampling frequency of 44.1 kHz is used. Filtering with this filter would produce typically no audible effect. Even when the same allpass filtering is applied a second time, the effect is still inaudible. However, when this filtering is repeated very many times, such as 100 times, the low and high frequencies in the output signal become separated in time and a chirp-like effect is heard. One can observe this by listening to the impulse response of the filter itself.

This phenomenon is illustrated in Figure 3, where the impulse response and the spectrogram of a cascade of 64 first-order allpass filters with $a_1 = 0.6$ are plotted. Now, the impulse response is quite long, about 6 ms. This due to the fact that cascading M identical allpass filters produces a high-order allpass filter having phase and group delay M times the original. When an impulse, a broadband signal, is input to a such filter, it is smeared into a chirp with frequencies appearing at times specified by the group delay.

The chirps that can be generated with a cascade of M identical allpass filters can be understood by recalling the group delay plots of Figure 2 and multiplying the time axis by M. When a great many impulse responses are convolved, a chirp going downwards

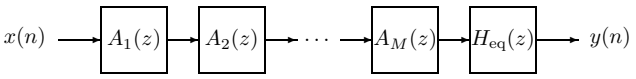
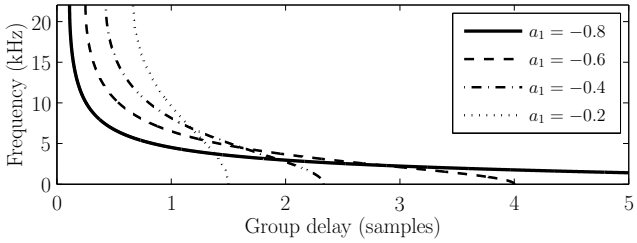
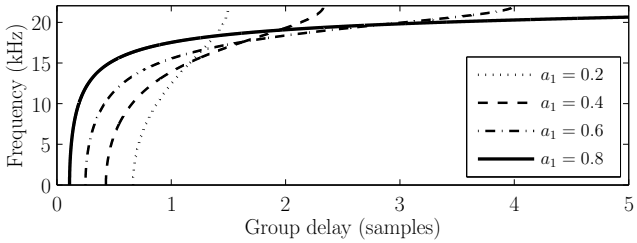


Figure 1: Block diagram of the basic spectral delay filter consisting of M allpass filters and an equalization filter [3].



(a)



(b)

Figure 2: Group delay of the first-order allpass filter with (a) non-positive ($a_1 \leq 0$) and (b) non-negative ($a_1 \geq 0$) coefficient values plotted against frequency. The sampling rate is $f_{\mathrm{s}} = 44.1$ kHz.

in frequency is heard when the filter coefficient a_1 is negative (see Figure 2(a)). When the coefficient a_1 is positive (see Figure 2(b)), a chirp going up is produced. The most useful allpass filter designs bring about a large time delay difference between bass frequencies (100 Hz or so) and high treble (well below 10 kHz). This is not effectively obtained using allpass filters with poles close to the unit circle. For instance, when coefficient a_1 is positive and close to 1.0, such as $a_1 = 0.9$, it is seen that the largest delay occurs at very high frequencies, which will be inaudible to humans at a standard sampling rate of 44.1 kHz.

It is helpful to be able to estimate the impulse response length of first-order allpass filters, because this tells how long the chirp will be. The maximum value of the group delay is related to the length of the chirp. It can be obtained by evaluating the group delay (5) at $\omega = 0$ for non-positive a_1 and at $\omega = \pi$ for positive a_1, i.e.

$$\tau_{\mathrm{g,max}} = \begin{cases} \tau_{\mathrm{g}}(0) = \dfrac{1 - a_1}{1 + a_1}, & a_1 \leq 0, \\ \tau_{\mathrm{g}}(\pi) = \dfrac{1 + a_1}{1 - a_1}, & a_1 > 0. \end{cases} \tag{6}$$

Alternatively, it is possible to estimate the length of the allpass chirp from theory for the impulse-response length of IIR filters [25]. The sufficient length of the impulse response of a first-order allpass filter that contains P% of the total energy is given by

$$N_P = \frac{\log(1 - P/100) - \log(1 - a_1^2)}{\log(a_1^2)} - 1, \tag{7}$$

where P is the percentage of energy, e.g., 99.9. Notice that here the

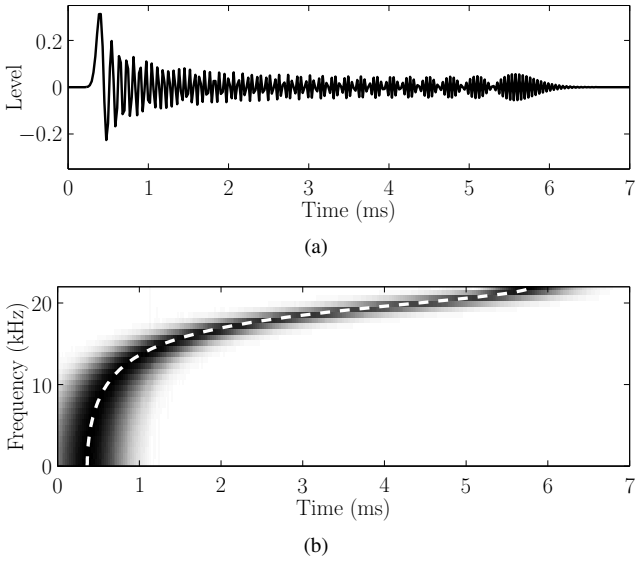


Figure 3: (a) Impulse response of the cascade of 64 first-order allpass filters ($a_1 = 0.6$) and (b) its spectrogram (linear magnitude of a 64-point sliding FFT, hop size 1 sample, Hamming window), which shows an upwards chirp. The dashed line is the group delay of one allpass filter multiplied by 64.

ceiling operation used in [25] is skipped to avoid rounding errors when many allpass filters are cascaded.

The maximum of the group delay according to (6) together with the number of impulse response samples that contain 99% and 99.9% of the total energy are plotted in Figure 4. It can be seen that the maximum value of the group delay is very close to the N_{99} value when $0.2 < |a_1| < 0.7$. It approaches $N_{99.9}$ as $|a_1|$ is increased.

For M cascaded allpass filters the effective length is MN_P. Equation (6) gives for the example of Figure 3 ($M = 64$, $a_1 = 0.6$) the length estimate of 256 samples (4.00 samples for one allpass section), which corresponds to 5.8 ms. This estimate is slightly too small in comparison to what is seen in Figure 3. Using the energy-based formula with $P = 99.9\%$, the estimated impulse response length is 261 samples ($N_P = 4.07$ samples), which corresponds to 5.9 ms. $N_{99.9}$ gives a conservative estimate of 404 samples (9.2 ms, 6.32 samples per allpass filter), which seems too large in this case.

2.2. Allpass Chirp Envelope

The envelope of the allpass chirp is not constant over time. This is clearly seen in Figure 3(a). The resulting amplitude envelope is due to the fact that the impulse response contains an equal amount of energy at all frequencies, and therefore the amplitude is necessarily smaller at frequencies where the chirp lingers. Next, the envelope $\alpha(\omega)$ of an allpass chirp is derived. Previously, this derivation has appeared in [26].

Denote by ω_+ and ω_- two nearby frequencies having difference Δ and mean ω, i.e.

$$\Delta = \omega_+ - \omega_-, \quad \text{and} \tag{8}$$

$$\omega = (\omega_+ + \omega_-)/2. \tag{9}$$

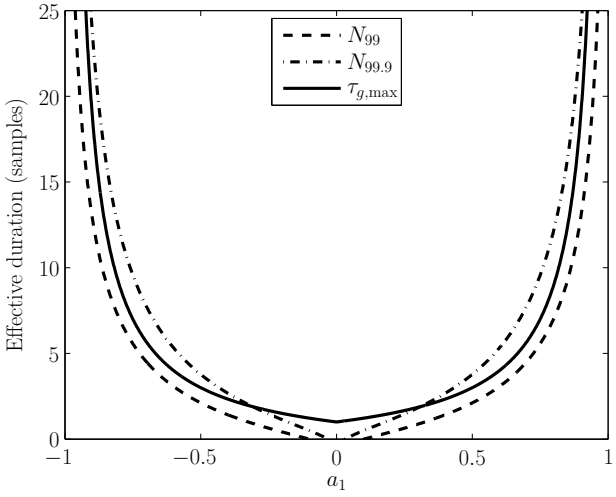


Figure 4: Three different estimates for the effective length of a first-order allpass filter impulse response as a function of coefficient a_1: energy-based effective length for 99.0% and 99.9% of the total energy and the group delay maximum.

In the frequency interval $[\omega_+, \omega_-]$ the allpass chirp has energy Δ/π. This energy is approximately equal to the signal energy in the time interval $[\tau_{\mathrm{g}}(\omega_+), \tau_{\mathrm{g}}(\omega_-)]$ determined by the group delay of the allpass filter,

$$\Delta/\pi \approx |\tau_{\mathrm{g}}(\omega_-) - \tau_{\mathrm{g}}(\omega_+)|\, \alpha^2(\omega)/2. \tag{10}$$

Taking the limit $\Delta \to 0$ gives

$$\alpha(\omega) = \left[\frac{\pi}{2}\left|\frac{\mathrm{d}\tau_{\mathrm{g}}(\omega)}{\mathrm{d}\omega}\right|\right]^{-\frac{1}{2}}. \tag{11}$$

Now, an allpass chirp may be filtered by an invertible normalization $\nu(\omega)$ with a magnitude chosen to approximate the inverse of its envelope $\alpha(\omega)$:

$$\nu(\omega) \approx \frac{1}{\alpha(\omega)} = \sqrt{\frac{\pi}{2}\left|\frac{\mathrm{d}\tau_{\mathrm{g}}(\omega)}{\mathrm{d}\omega}\right|}. \tag{12}$$

Provided the equalizing filter does not smear the phase of the signal too much, the resulting sequence will have a nearly constant envelope.

In the case of a first-order allpass, the inverse envelope $\nu(\omega)$ can be expressed in closed form as

$$\frac{1}{\alpha(\omega)} = \sqrt{\frac{\pi}{2}\left|\frac{\mathrm{d}\tau_{\mathrm{g}}(\omega)}{\mathrm{d}\omega}\right|} = \frac{\sqrt{\pi|a_1(1-a_1^2)\sin(\omega)|}}{1 + 2a_1\cos(\omega) + a_1^2}. \tag{13}$$

2.3. Allpass Chirp Equalization

An accurate equalization (EQ) filter for a cascade of first-order allpass filters can be designed based on (13). The design can be decomposed into three parts as

$$H_{\mathrm{eq}}(z) = \sqrt{M\pi|a_1(1-a_1^2)|} \cdot H_{\mathrm{eq,d}}(z) \cdot H_{\mathrm{eq,n}}(z), \tag{14}$$

corresponding to a scale factor, an all-pole "denominator" section $H_{\mathrm{eq,d}}(z)$, and a "numerator" section $H_{\mathrm{eq,n}}(z)$. The scale factor

contains M because the cascading of M indentical allpass filters multiplies the group delay of one filter by M, see Section 2.1.

It can be seen in (13) that the denominator can be implemented directly as a cascade of two one-pole filters having the pole at $-a_1$, i.e.

$$H_{\mathrm{eq,d}}(z) = \left(\frac{1}{1+a_1 z^{-1}}\right)^2 = \frac{1}{1+2a_1 z^{-1}+a_1^2 z^{-2}}. \quad (15)$$

The magnitude response of this filter matches the denominator part of the amplitude envelope.

In addition, it can be seen that the numerator has a function $\sqrt{|\sin(\omega)|}$ that is symmetric with respect to $\pi/2$, i.e. one-fourth of the sampling frequency. Thus, it makes sense to model the first half of the function and replace z^{-1} by z^{-2} in the resulting filter. The warped Prony design is used here to concentrate on matching the desired low-frequency magnitude of the square root of the sine. This yields the fourth-order IIR filter given by

$$\begin{aligned} H_{\mathrm{eq,n}}(z) \;=\; & g\frac{(1-b_{\mathrm{eq},1}z^{-2})(1-b_{\mathrm{eq},2}z^{-2})}{(1-a_{\mathrm{eq},1}z^{-2})(1-a_{\mathrm{eq},2}z^{-2})} \\ & \cdot\frac{(1-b_{\mathrm{eq},3}z^{-2})(1-b_{\mathrm{eq},4}z^{-2})}{(1-a_{\mathrm{eq},3}z^{-2})(1-a_{\mathrm{eq},4}z^{-2})}, \end{aligned} \quad (16)$$

where $g = 0.7079$ and the rest of the coefficient values are given in Table 1.

Figure 5 shows an example where the allpass filter chain's impulse response obtained with the same parameters as in Figure 3 is equalized using the filters above. It is seen in Figure 5(a) that the magnitude response of the target and model for a chain containing one allpass filter ($M = 1$) match quite well. In order to make a difference between the EQ filter magnitude response and the inverse envelope the scaling has been omitted. If it is included, no difference between the responses can be seen. Figure 5(b) shows the equalized impulse response, which has a practically constant level. Figure 5(c) presents the group delay of the EQ filter. It can be seen that the EQ filter adds only a small delay to the signal. Especially at low and middle frequencies, where human hearing is most sensitive, the EQ filter adds practically no additional delay. Therefore, the frequency-dependent delay of the SDF can be understood to be set by the allpass filter chain.

2.4. Stretched Spectral Delay Filter

A multirate method was also introduced in [3] to obtain a more dramatic spectral delay effect with a small number of allpass filters. By replacing the unit delay in each allpass filter with two or more delay elements, the chirp becomes slower. We call this the stretched SDF. At the same time, image chirps are generated, which can be filtered out or used as part of the effect.

Figure 6(a) shows the impulse response of a stretched SDF having 64 allpass sections in which each allpass section contains three unit delays instead of one. The impulse response is three times as long as in Figure 3(a), but two samples out of three are zero. This method produces image chirps at high frequencies. Every second chirp propagates in the opposite direction as compared to the original one. This is seen in the spectrogram display in Figure 6(b). Since one of the chirps now occurs in the lower third of the audio band (between 0 Hz and about 7 kHz at 44.1 kHz), it is heard more easily than in the case of a single full-band chirp.

For the amplitude envelope equalization of a stretched SDF, an interpolated version of the EQ filter of (16) suffices, i.e.

$$H_{\mathrm{eqi}}(z) = H_{\mathrm{eq}}(z^K), \quad (17)$$

Table 1: Coefficients of the fourth-order numerator equalizing filter.

k	$b_{\mathrm{eq},k}$	$a_{\mathrm{eq},k}$
1	0.3525	0.9797
2	0.9979	0.1103
3	0.9425	0.8750
4	0.7628	0.5892

where K is the number of unit delays in each allpass filter.

3. SPECTRAL DELAY FILTERS WITH FEEDBACK

As originally suggested by Kim-Boyle [1], the spectral delay processing could be extended by feeding some of the filter output back to its input. This approach could generate a series of chirps modified by the possible filtering in the feedback path. Here, this approach is considered and the condition for the system stability is derived.

If the output of a basic SDF $H(z)$ is fed back to its input via a unit delay and filter $B(z)$, the system illustrated in Figure 7 is obtained. Both filters are assumed to be causal. Therefore, their impulse responses may be denoted $h(n)$ and $b(n)$, respectively, nonzero for $n = 0, 1, \ldots$. The unit delay is needed in the feedback path when $b(0)h(0) \neq 0$ to avoid a delay-free loop.

This system can be described with a set of equations given by

$$w(n) \;=\; x(n) + b(n) * y(n-1) \quad \text{and} \quad (18)$$

$$y(n) \;=\; h(n) * w(n), \quad (19)$$

where $w(n)$ is the signal after the first summation and $*$ denotes convolution. Now, by applying the z-transform to the equation set above, the system can be expressed as

$$W(z) \;=\; X(z) + z^{-1}B(z)Y(z) \quad \text{and} \quad (20)$$

$$Y(z) \;=\; H(z)W(z), \quad (21)$$

from which the system transfer function is obtained,

$$G(z) = \frac{Y(z)}{X(z)} = \frac{H(z)}{1 - z^{-1}B(z)H(z)}. \quad (22)$$

Now, the stability condition for the modified SDF utilizing feedback can be derived. First of all, the basic SDF $H(z)$ and the feedback filter $B(z)$ are assumed to be stable. Since

$$\frac{H(z)}{1 - z^{-1}B(z)H(z)} = H(z)\sum_{k=1}^{\infty} z^{-k}B^k(z)H^k(z), \quad (23)$$

a sufficient condition for stability is that $B(z)$ and $H(z)$ be causal and stable (as assumed), and that $|B(e^{j\omega})H(e^{j\omega})| < 1$ for all ω. These conditions provide that $G(e^{j\omega})$ is both causal and convergent everywhere on the unit circle. The loop gain condition $|B(e^{j\omega})H(e^{j\omega})| < 1$ is also *necessary* at all frequencies for which $H(e^{j\omega}) \neq 0$. Since the magnitude response of the basic SDF depends solely on the magnitude response of the EQ filter, which can be larger than one for some frequencies (see Figure 5), the feedback filter should apply attenuation that makes the feedback gain less than one, i.e.

$$|B(e^{j\omega})H_{\mathrm{eq}}(e^{j\omega})| < 1 \Rightarrow |B(e^{j\omega})| < \frac{1}{|H_{\mathrm{eq}}(e^{j\omega})|}\ \forall\,\omega. \quad (24)$$

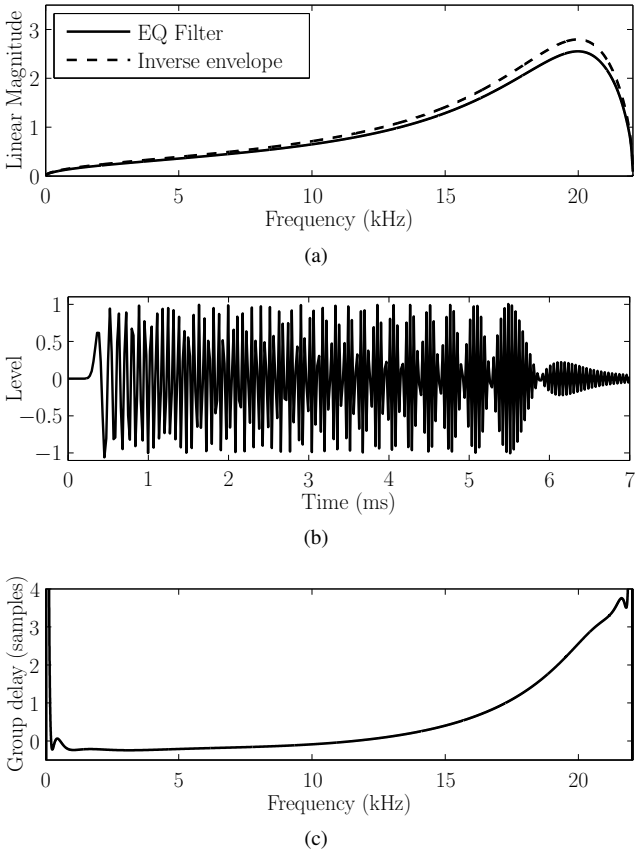


Figure 5: (a) Inverse envelope of a one-stage allpass chirp and an EQ filter's magnitude response with the scaling omitted. (b) The equalized impulse response (cf. Figure 3(a)). (c) The group delay of the equalizing filter used in this example.

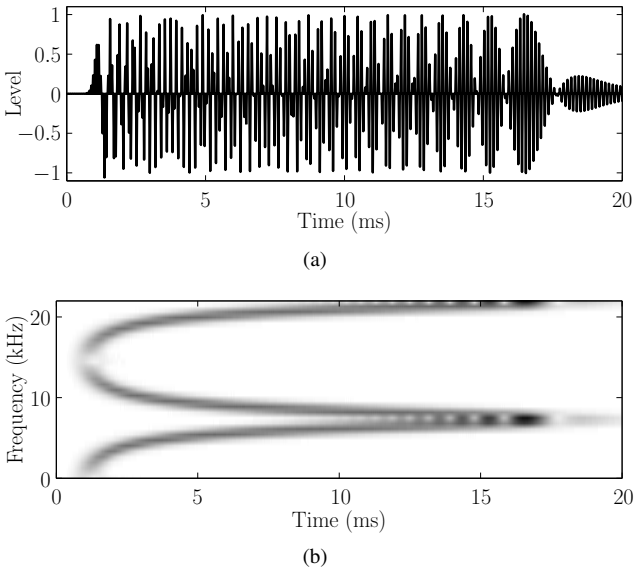


Figure 6: (a) Impulse response of a cascade of 64 first-order allpass filters ($a_1 = 0.6$) having the unit delay replaced with three unit delays ($K = 3$) and (b) its spectrogram computed with the same parameters as in Figure 3. Now three long chirps are seen instead of one. The interpolated EQ filter has been applied to the impulse response.

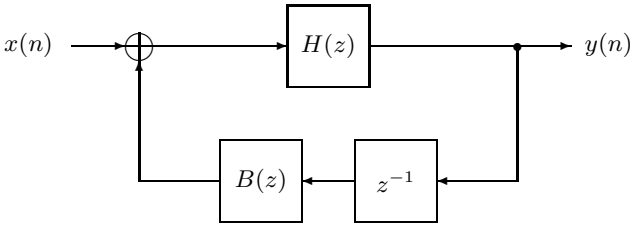


Figure 7: A modified spectral delay filter utilizing a basic SDF $H(z)$ and a time-invariant filter $B(z)$ in the feedback path.

Otherwise the signal fed back to the input of the basic SDF will be amplified over and over again as it travels through the loop.

Now, if the feedback filter is designed to approximate the amplitude envelope of the chirp and to satisfy the stability condition given above, maximum amplification of the feedback is then obtained. However, this leads to quite sophisticated design criteria (cf. the design of the EQ filter) and unsophisticated first-order or second-order filters cannot produce clearly different results due to the attenuation required for the stability. Yet, if the EQ filter is omitted, the feedback filter can be more freely designed, as long as the stability criterion holds.

An example of the use of a SDF with feedback is given in Figure 8, where the impulse response and the spectrogram of a basic SDF having 64 first-order allpass filters with coefficient $a_1 = 0.6$, an EQ filter, and a feedback filter

$$B(z) = \frac{1}{23}\left(1 + z^{-1}\right) \tag{25}$$

are plotted. As can be seen, the impulse response (Figure 8(a)) has now small amplitude variations not present in the impulse response of a basic SDF without feedback (cf. Figure 3). These variations are due to an additional low-amplitude chirp at high frequencies, which can be observed from the spectrogram (Figure 8(b)). Since the low frequencies are passed by the basic SDF almost instantly and the feedback filter is a highpass filter, they are attenuated as can be seen in the spectrogram. In addition, for this reason they are not present in the additional chirp.

4. TIME-VARYING SPECTRAL DELAY FILTER STRUCTURES

If the coefficients of the allpass filters in the allpass chain are allowed to be time-varying, more lively effects are obtained. Now, the spectral delay generated by the chain varies over time, and the output will no longer be a pure chirp. With an appropriate choice of the coefficient modulation, the desired dynamic effect can be obtained.

However, when the coefficients of the allpass filters are time-varying, the stability of the system becomes a more crucial design criterion. It is easily proved that if two filters are stable, their cascade is also stable. Now, if all filters in the allpass filter chain are first-order, the same stability condition as for a time-invariant first-order allpass filter holds, i.e. the filter coefficient must be in

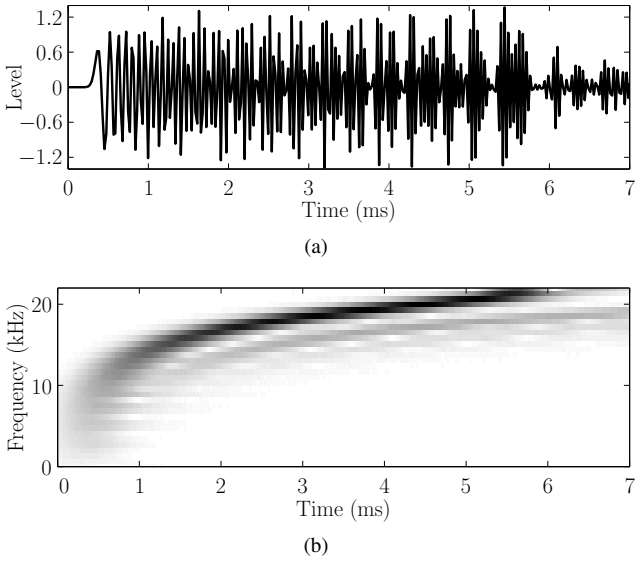


Figure 8: (a) Impulse response of a SDF having 64 first-order allpass filters with coefficient $a_1 = 0.6$, an EQ filter, and a feedback filter $B(z) = \frac{1}{23}\left(1 + z^{-1}\right)$, and (b) its spectrogram (obtained with the same parameters as in Figure 3).

the range $[-1, 1]$, end points included [8]. Here, the stability condition for a time-varying stretched allpass filter is derived.

A stretched allpass filter, i.e., a first-order allpass filter where the unit delay is replaced with two or more unit delays, having a stretching factor $K = 2, 3, \ldots$, and a time-varying coefficient $m(n)$, can be described by the following relations:

$$w_1(n+1) = m(n)y(n) + x(n) \quad (26)$$
$$w_k(n+1) = w_{k-1}(n), \qquad k = 2, \ldots, K, \quad (27)$$
$$y(n) = w_K(n) - m(n)x(n), \quad (28)$$

where $w_k(n)$ is the signal exiting the kth delay of the filter. This system can be given in matrix form

$$X_{n+1} = P_n X_n + Q_n x(n), \quad (29)$$
$$y(n) = R_n X_n + S_n x(n), \quad (30)$$

where $X_n = [w_1(n), w_2(n), \ldots, w_K(n)]^t$, $S_n = -m(n)$, $Q_n = 1 - m^2(n)$, R_n is a vector containing the value one at index K and zero elsewhere, and P_n is a $K \times K$ matrix given by

$$P_n(i,j) = \begin{cases} m(n), & i = 1, \text{and } j = K \\ 1, & i = 2, \ldots, K, \text{and } j = i-1 \\ 0, & \text{elsewhere,} \end{cases} \quad (31)$$

where $P_n(i,j)$ denotes the element in row i and in column j. In other words, the matrix P_n has ones in the first diagonal below the main diagonal and $m(n)$ in the last element of the first row, i.e.

$$P_n = \begin{bmatrix} 0 & 0 & \cdots & 0 & m(n) \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}. \quad (32)$$

Now, the stability condition can be given as [27]

$$G = |S_n| + \sum_{i=n-1}^{\infty} \left| R_n \prod_{k=i+1}^{n} P_k \, Q_i \right| < \infty \; \forall \, n. \quad (33)$$

Applying the Cauchy-Schwartz inequality gives

$$G \leq |S_n| + \sum_{i=n-1}^{\infty} ||R_n|| \prod_{k=i+1}^{n} ||P_k|| \, |Q_i|. \quad (34)$$

Now, the norm of vector R_n is one, and the absolute values of scalars S_n and Q_i are $|m(n)|$ and $|1 - m^2(i)|$, respectively. The norm of matrix P_k is defined as $||P_k|| = \max_{||x||=1} ||P_k x||$, where $x \in \mathbb{R}^K$, and it can be shown to be the positive square root of the largest eigenvalue of matrix $P_k^t P_k$. In this case,

$$P_k^t P_k(i,j) = P_k(:,i)^t P_k(:,j) = \begin{cases} 1, & i = j \neq K, \\ m^2(k), & i = j = K, \\ 0, & i \neq j, \end{cases} \quad (35)$$

where $P_k(:,i)$ denotes the ith column of the matrix P_k. In other words, the matrix $P_k^t P_k$ is an identity matrix with the last diagonal value replaced with $m^2(k)$. The eigenvalues of this matrix are now easily obtained, and the positive square root of its largest eigenvalue is $\max(1, |m(k)|)$.

The stability condition is now expressed as

$$G \leq |m(n)| + \sum_{i=n-1}^{\infty} \prod_{k=i-1}^{n} \max(1, |m(k)|) |1 - m^2(i)|. \quad (36)$$

Now, if $m(n) = 0 \; \forall \, n$, the filter is stable as it then reduces to a pure delay of K samples. If $0 < |m(n)| \leq 1$, $\max(1, |m(k)|) = 1$ and $|1 - m^2(i)| < 1$, and thus the sum term becomes

$$\prod_{k=i-1}^{n} 1^k |1 - m^2(i)| = |1 - m^2(i)| < 1, \quad (37)$$

which makes sum term of the condition convergent, i.e the filter is stable. If $|m(n)| > 1$, then $\max(1, |m(k)|) = |m(k)|$, and the sum term increases continuously as i decreases, effectively making the sum term infinite contradicting the stability condition. Therefore, the stretched allpass filter is stable if and only if $|m(n)| \leq 1 \; \forall \, n$.

In the case of a time-varying SDF, the equalizing filter must then also be time-varying. Now, the scaling and the denominator filter depend on the time-varying coefficient, and the stability is determined by the stability of the denominator filter. Since the denominator filter can be composed of two one-pole filters, it suffices to analyze the stability of one of its stages. For a time-varying one-pole filter modulated by the signal $m(n)$, the parameters for (33) are $S_n = R_n = Q_n = 1$ and $P_n = m(n)$. Substituting these in (33) leads to the condition $|m(n)| < 1$. However, if $|m(n)| = 1$, the scaling of the EQ filter makes the output become zero, so the EQ filter is also stable when $|m(n)| \leq 1 \; \forall \, n$.

If a time-varying SDF includes a feedback path, the observations given in Section 3 hold. However, since in such cases the magnitude response of the time-varying SDF depends on both the modulation and the input signal (via filter states), it is not well defined. Therefore, it is recommended that only feedback filters with a quite large attenuation at all frequencies should be used. Again,

this would lead to a filter structure that does not differ much from a basic time-varying SDF. If the EQ filter is omitted, the feedback filter can be more freely designed and more interesting effects are obtained, as illustrated next.

In Figure 9, the impulse response and the spectrogram of a SDF having 64 first-order allpass filter with the coefficients modulated by an 8 Hz sine having amplitude 0.9 and a feedback path with a constant multiplier $B(z) = 0.99$ are plotted. Now, the SDF does not contain an EQ filter. The impulse response (Figure 9(a)) is noise like and more dynamic than the impulse response of the time-invariant spectral delay filter utilizing feedback (cf. Figure 8). When looking at the spectrogram (Figure 9(b)), the dynamic characteristics of this filter are more clearly visible. Now, there are several simultaneous chirps going up and down in a sine-like fashion with a period corresponding to the modulation period. An audio-rate coefficient modulation of the SDFs has some interesting practical applications and some of them are presented in [28].

Figure 10 plots the impulse response and the spectrogram of a SDF having 64 stretched allpass filter with $K = 3$ and with the coefficients modulated by the output of an envelope-following filter given by $y(n) = 0.1|w(n)|+0.9y(n-1)$, where $w(n)$ is the sum of the input and feedback signals, and a feedback path with a constant multiplier $B(z) = 0.99$. Again, the SDF does not contain an equalizing filter. The impulse response (Figure 10(a)) is again more dynamic than the impulse response of the time-invariant SDF utilizing feedback (see Figure 8) and it contains a series of chirps with decaying amplitudes. The spectrogram (Figure 10(b)) now shows more clearly that the chirp durations increase as they travel through the system.

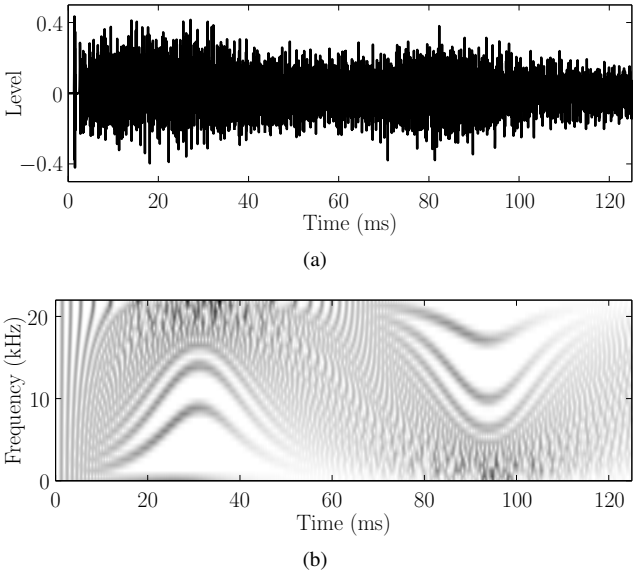


Figure 9: (a) Impulse response of a SDF having 64 first-order allpass filters with the coefficients modulated by an 8 Hz sine having amplitude 0.9 and a feedback path with a constant multiplier $B(z) = 0.99$. (b) The spectrogram of the impulse response (obtained with the same parameters as in Figure 3). The SDF does not contain an equalizing filter.

5. CONCLUSIONS

The properties of the recently introduced spectral delay filter (SDF) were reviewed. The SDF, which consists of a cascade of many first-order allpass filters and an equalizing filter, produce unusual effects on an audio signal. The SDF has a long, chirp-like impulse response, which effectively makes the various frequency components in an audio signal travel through the filter at different speeds. Since the amplitude envelope of the allpass chirp is inversely proportional to the square root of the frequency derivative of the filter group delay, the slowest portions of the chirp are also the softest. In order to equalize the variations in the chirp amplitude, a new equalizing filter, which approximates the inverse chirp amplitude envelope, was presented. In addition, properties of SDFs utilizing feedback and/or time-varying coefficients were discussed. The stability conditions for both cases were derived and examples of their use were also presented.

The SDF can be used to process arbitrary audio signals to transform their timbre. This is computationally more efficient than convolution, since the length of the impulse response of the SDF can be very long. Alternatively, the impulse responses produced with SDFs can be used as sound effects or synthetic percussion samples. In general, the SDF can be used to bring electronic or synthesizer-like characteristics to arbitrary audio signals. In this sense it is related to previous extended synthesis methods that can process arbitrary audio signals, such as the audio signal driven sound synthesis methods proposed by Poepel and Dannenberg [29] or the adaptive FM synthesis and adaptive split sideband synthesis methods introduced by Lazzarini et al. [30, 31].

Sound examples on the use of the spectral delay filters utilizing the structures discussed in this paper can be found at `http://www.acoustics.hut.fi/publications/papers/dafx09-sdf/`.

6. REFERENCES

[1] D. Kim-Boyle, "Spectral delays with frequency domain processing," in *Proc. DAFx-04*, Naples, Italy, Oct. 2004, pp. 42–45.

[2] X. Amatriain, *An Object Oriented Metamodel for Digital Signal Processing with a Focus on Audio and Music*, Ph.D. thesis, Universitat Pompeu Fabra, Barcelona, Spain, 2004, pp. 184–188. See also: `http://www.create.ucsb.edu/~xavier/Thesis/html/node127.html`.

[3] V. Välimäki, J. S. Abel, and J. O. Smith, "Spectral delay filters," *J. Audio Eng. Soc.*, vol. 57, no. 7/8, July/Aug. 2009.

[4] M. R. Schroeder, "Natural sounding artificial reverberation," *J. Audio Eng. Soc.*, vol. 10, no. 3, pp. 219–223, July 1962.

[5] J. S. Abel, D. P. Berners, S. Costello, and J. O. Smith, "Spring reverb emulation using dispersive allpass filters in a waveguide structure," presented at *the AES 121st Convention*, San Francisco, CA, Oct. 2006, paper no. 6954.

[6] J. O. Smith, "An allpass approach to digital phasing and flanging," Report STAN-M-21, CCRMA, Dept. of Music, Stanford University, Stanford, CA, 1984.

[7] P. A. Regalia and S. K. Mitra, "Tunable digital frequency response equalization filters," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 35, no. 1, pp. 118–120, Jan. 1987.

[8] J. Pekonen, "Coefficient-modulated first-order allpass filter as distortion effect," in *Proc. DAFx-08*, Espoo, Finland, Sep. 2008, pp. 83–87.

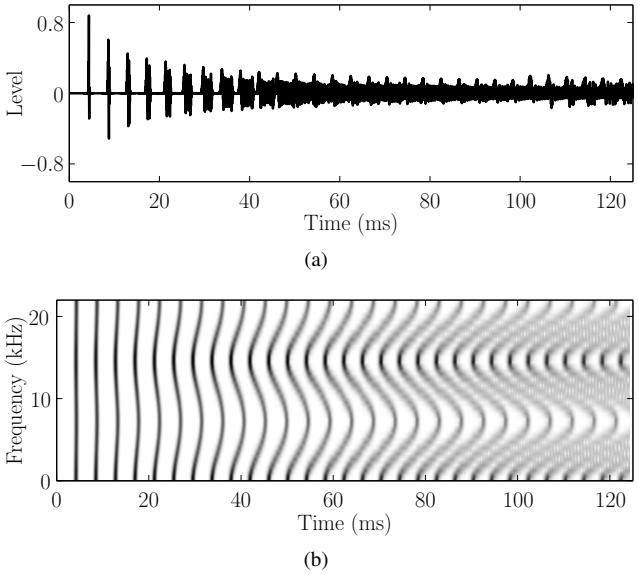


Figure 10: (a) Impulse response of a SDF having 64 stretched allpass filters with $K = 3$ and with the coefficients modulated by the output of an envelope following filter given by $y(n) = 0.1|w(n)| + 0.9y(n-1)$ where $w(n)$ is the sum of the input and feedback signals, and a feedback path with a constant multiplier $B(z) = 0.99$. (b) The spectrogram of the impulse response (obtained with the same parameters as in Figure 3). The SDF does not contain an equalizing filter.

[9] D. A. Jaffe and J. O. Smith, "Extensions of the Karplus-Strong plucked-string algorithm," *Computer Music J.*, vol. 7, no. 2, pp. 76–87, Summer 1983.

[10] T. I. Laakso, V. Välimäki, M. Karjalainen, and U. K. Laine, "Splitting the unit delay — tools for fractional delay filter design," *IEEE Signal Process. Mag.*, vol. 13, no. 1, pp. 30–60, Jan. 1996.

[11] M. Karjalainen and J. O. Smith, "Body modeling techniques for string instrument synthesis," in *Proc. ICMC'96*, Hong Kong, Aug. 1996, pp. 232–239.

[12] J. O. Smith and J. S. Abel, "Bark and ERB bilinear transforms," *IEEE Trans. Speech Audio Process,*, vol. 7, no. 6, pp. 697–708, Nov. 1999.

[13] A. Härmä, M. Karjalainen, L. Savioja, V. Välimäki, U. K. Laine, and J. Huopaniemi, "Frequency-warped signal processing for audio applications," *J. Audio Eng. Soc.*, vol. 48, no. 11, pp. 1011–1031, Nov. 2000.

[14] G. Evangelista, "Time and frequency warping musical signals," in *DAFX: Digital Audio Effects*, U. Zölzer, Ed., chapter 11, pp. 439–464. Wiley & Sons, Ltd., Chichester, UK, 2002.

[15] T. Paatero and M. Karjalainen, "Kautz filters and generalized frequency resolution: Theory and audio applications," *J. Audio Eng. Soc.*, vol. 51, no. 1, pp. 27–44, Jan. 2003.

[16] S. A. Van Duyne and J. O. Smith, "A simplified approach modeling dispersion caused by stiffness in strings and plates," in *Proc. ICMC'94*, Aarhus, Denmark, Sep. 1994, pp. 407–410.

[17] B. Bank, F. Avanzini, G. Borin, G. De Poli, F. Fontana, and D. Rocchesso, "Physically informed signal processing methods for piano sound synthesis: A research overview," *EURASIP J. Applied Signal Process.*, vol. 2003, no. 10, Sep. 2003.

[18] I. Testa, G. Evangelista, and S. Cavaliere, "Physically inspired models for the synthesis of stiff strings with dispersive waveguides," *EURASIP J. Applied Signal Process.*, vol. 2004, no. 7, pp. 964–977, 2004.

[19] J. Rauhala and V. Välimäki, "Tunable dispersion filter design for piano synthesis," *IEEE Signal Process. Lett.*, vol. 13, no. 5, pp. 253–256, May 2006.

[20] D. Rocchesso and P. Dutilleux, "Generalization of a 3-D acoustic resonator model for the simulation of spherical enclosures," *EURASIP J. Applied Signal Process.*, vol. 2001, no. 1, pp. 15–26, 2001.

[21] M. Karjalainen, V. Välimäki, and P. A. A. Esquef, "Efficient modeling and synthesis of bell-like sounds," in *Proc. DAFx-02*, Hamburg, Germany, Sep. 2002, pp. 181–186.

[22] J. Kleimola, "Dispersion modulation using allpass filters," in *Proc. DAFx-08*, Espoo, Finland, Sep. 2008, pp. 193–197.

[23] V. Välimäki, M. Ilmoniemi, and M. Huotilainen, "Decomposition and modification of musical instrument sounds using a fractional delay allpass filter," in *Proc. NORSIG2004*, Espoo, Finland, June 2004, pp. 208–211.

[24] H.-M. Lehtonen, V. Välimäki, and T. I. Laakso, "Cancelling and selecting partials from musical tones using fractional-delay filters," *Computer Music J.*, vol. 32, no. 2, pp. 43–56, Summer 2008.

[25] T. I. Laakso and V. Välimäki, "Energy-based effective length of the impulse response of a recursive filter," *IEEE Trans. Instrum. Meas.*, vol. 48, no. 1, pp. 7–17, Feb. 1999.

[26] J. S. Abel and D. P. Berners, "MUS424/EE367D: Signal processing techniques for digital audio effects," Unpublished Course Notes, CCRMA, Stanford University, Stanford, CA, Apr. 2005.

[27] J. Laroche, "On the stability of time-varying recursive filters," *J. Audio Eng. Soc.*, vol. 55, no. 6, pp. 460–471, June 2007.

[28] J. Kleimola, J. Pekonen, H. Penttinen, V. Välimäki, and J. S. Abel, "Sound synthesis using an allpass filter chain with audio-rate coefficient modulation," in *Proc. DAFx-09*, Como, Italy, Sep. 2009.

[29] C. Poepel and R. Dannenberg, "Audio signal driven sound synthesis," in *Proc. ICMC'05*, Barcelona, Spain, Sep. 2005, pp. 391–394.

[30] V. Lazzarini, J. Timoney, and T. Lysaght, "The generation of natural-synthetic spectra by means of adaptive frequency modulation," *Computer Music J.*, vol. 32, no. 2, pp. 9–22, Summer 2008.

[31] V. Lazzarini, J. Timoney, and T. Lysaght, "Non-linear distortion synthesis using the split sideband method, with applications to adaptive signal processing," *J. Audio Eng. Soc.*, vol. 56, no. 9, pp. 684–695, Sep. 2008.