

SOUND SYNTHESIS USING AN ALLPASS FILTER CHAIN WITH AUDIO-RATE COEFFICIENT MODULATION

*Jari Kleimola, Jussi Pekonen, Henri Penttinen, Vesa Välimäki**

Dept. of Signal Processing and Acoustics

Helsinki University of Technology (TKK), Espoo, Finland

{jari.kleimola, jussi.pekonen, henri.penttinen, vesa.valimaki}@tkk.fi

Jonathan S. Abel

CCRMA, Dept. of Music

Stanford University, Stanford, CA, USA

abel@ccrma.stanford.edu

ABSTRACT

This paper describes a sound synthesis technique that modulates the coefficients of allpass filter chains using audio-rate frequencies. It was found that modulating a single allpass filter section produces a feedback AM-like spectrum, and that its bandwidth is extended and further processed by non-sinusoidal FM when the sections are cascaded. The cascade length parameter provides dynamic bandwidth control to prevent upper range aliasing artifacts, and the amount of spectral content within that band can be controlled using a modulation index parameter. The technique is capable of synthesizing rich and evolving timbres, including those resembling classic virtual analog waveforms. It can also be used as an audio effect with pitch-tracked input sources. Software and sound examples are available at

<http://www.acoustics.hut.fi/publications/papers/dafx09-cm/>

1. INTRODUCTION

Allpass filters are versatile building blocks that have many uses in audio signal processing [1-4]. They have been utilized also for sound synthesis purposes, but only recent research has started to consider them as primary synthesis techniques [5,6]. This is not surprising, because allpass filters do not have a profound timbral effect on the signals they process. However, the authors found that allpass filters can also be used as rich and evolving sound sources: it is even possible to build a virtual analog synthesizer from two sinusoidal oscillators and a chain of first-order allpass filter sections – an uncommon toolkit for subtractive synthesis implementations. This paper proposes a synthesis technique that is based on modulating the coefficients of many cascaded allpass filters using audio-rate frequencies, as shown in Fig. 1. The authors call this method coefficient modulation (CM) synthesis.

This work stems from recent research on spectral delay filters, which consist of a chain of allpass filters and an equalizing filter [4]. Ideas from the authors' earlier work on a distortion algorithm based on coefficient modulation [7], dispersion modulation in plucked string models [8], and guitar body mode modulation using warped IIR filters [9] are employed.

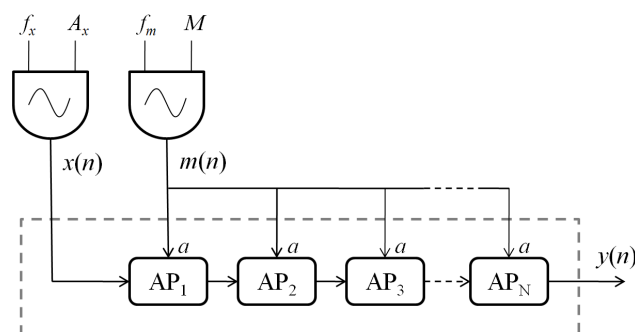


Figure 1: Block diagram of the new synthesis technique.

Additionally, this work is related to the audio-rate pick-up point modulation of waveguide models by Van Duyne and Smith [10], distributed string tension modulation by Pakarinen et al. [11], the adaptive FM technique described by Lazzarini, Timoney and Lysaght [12], and the feedback AM synthesis technique briefly discussed in [13]. Other related research on modulated filters includes the work by Greenfield on dithered digital filters, which focused on increasing numerical accuracy [14].

Section 2 of this paper discusses coefficient modulation in the case of one and many allpass filters, and analyzes how aliasing appears in the output signal. The relationship of CM synthesis technique to previous amplitude modulation (AM), frequency modulation (FM), and ring modulation (RM) methods is also discussed, followed by a description of the synthesis parameters of the proposed method. Section 3 presents implementations of CM synthesis on two platforms, the PC and mobile phone. Application examples are introduced in Section 4. The sampling rate is 44.1 kHz in all examples of this paper.

2. AUDIO-RATE COEFFICIENT MODULATION

2.1. Single First-Order Stage

The transfer function of a time-invariant first-order allpass filter is

* research performed while at CCRMA, Stanford University.

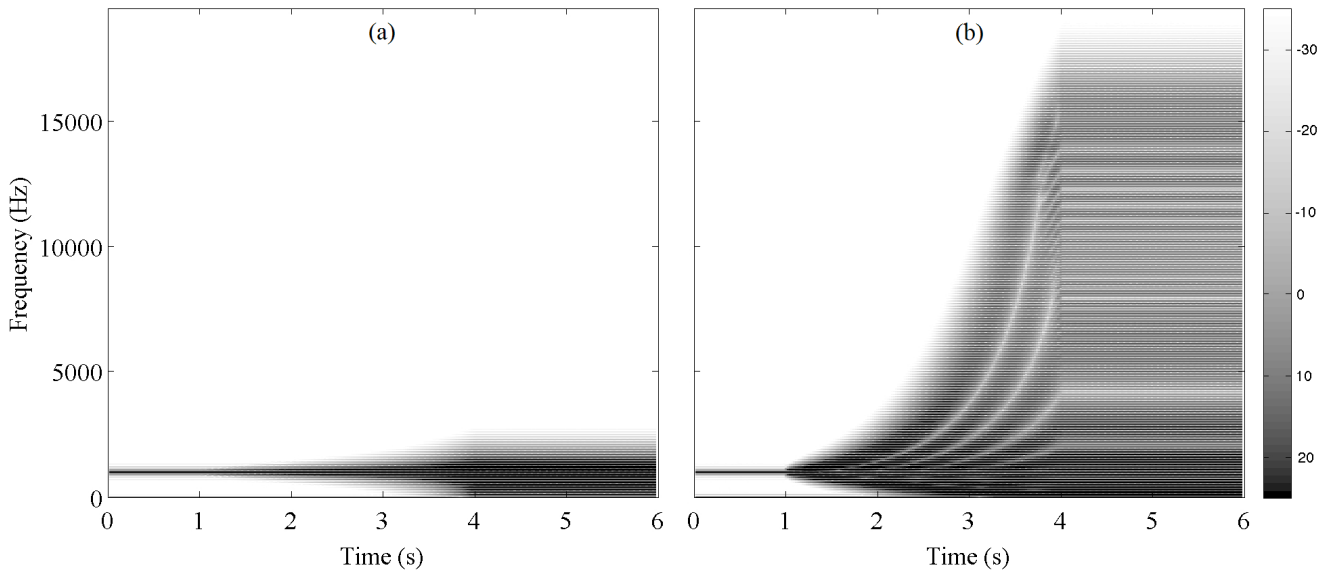


Figure 2: First-order allpass filter coefficient modulation at audio rate. Sinusoidal carrier (1 kHz) and modulator (100 Hz) signals, linearly increasing modulation index (0...0.99) between (1...4) seconds. The chain length determines the bandwidth, while the modulation index controls the amount of higher end spectral content. (a) Single allpass stage, (b) chain of 70 allpass stages.

$$H(z) = \frac{z^{-1} + a_1}{1 + a_1 z^{-1}}. \quad (1)$$

If the filter coefficient a_1 is made time varying according to the modulation function $m(n)$, the output of the allpass filter is given by the difference equation

$$y(n) = x(n-1) + x(n)m(n) - y(n-1)m(n-1), \quad (2)$$

where $y(n)$ and $x(n)$ are the output and input signals at discrete time step n , and $y(n-1)$, $x(n-1)$ and $m(n-1)$ are their delayed representations at time step $n-1$ [7].

Related research has shown that the output given by (2) may contain transients when used with sharp-edged modulation signals, requiring a smoothing function to even out the sharp discontinuities in the modulator [5]. The authors found that the transient problem is due to the one sample difference in the modulation signal terms of (2). Consequently, when $m(n-1)$ is replaced with $m(n)$, the transient problem disappears. Furthermore, because the effect of the approximation is practically negligible for continuous modulation signals, this work rewrites (2) as

$$y(n) = x(n-1) + x(n)m(n) - y(n-1)m(n). \quad (3)$$

By noticing that multiplication of two digital signals yields amplitude or ring modulation (AM/RM), it is possible to interpret the three terms on the right hand side of (3) as the delayed input, amplitude modulated input, and amplitude modulated feedback signals. Rearranging (3),

$$y(n) + m(n)y(n-1) = x(n-1) + m(n)x(n) \quad (4)$$

and noting that the weighted sum of successive samples approximates a time shift, it is seen that the time-shifted output is equal to the input time shifted in a complementary manner – the closer the absolute allpass coefficient $|m(n)|$ is to one, the larger the discrepancy between the input and output time shifts. Thus, modulation of the allpass coefficient can be seen to act as a frequency

modulation resulting from a changing time delay acting on the input signal.

In order to find the spectral structure of (3) and (4), let us first consider a sinusoidal carrier signal

$$x(n) = C \cos(\omega_x n + \theta_x), \quad (5)$$

where C is the maximum amplitude, $\omega_x = 2\pi f_x / f_s$ the angular frequency, and θ_x the initial phase offset of the carrier. Setting $C = 1$ and $\theta_x = 0$, and modulating the amplitude of the carrier with another bipolar sinusoid $m(n) = M \cos(\omega_m n + \theta)$ yields the classic RM equation

$$r(n) = M \cos(\omega_m n + \theta_m) \cos(\omega_x n), \quad (6)$$

where M is the modulation index, $\omega_m = 2\pi f_m / f_s$ is the angular frequency, and θ_m is the initial phase offset of the modulator. Using trigonometric identities and omitting the initial phase offset for clarity, we get

$$r(n) = \frac{M}{2} \cos((\omega_x \pm \omega_m)n). \quad (7)$$

Replacing the delayed input term $x(n-1)$ of Eq. (3) with the delayed carrier signal $\cos(\omega_x(n-1))$, and the amplitude modulated input term $x(n)m(n)$ with $r(n)$, one notices that the first two terms of (3) contribute partials that are located at f_x and $f_x \pm f_m$. This is equivalent to the classic double-sideband full-carrier AM spectrum. The amplitude modulated feedback term $y(n-1)m(n)$ processes this spectrum recursively, extending the width of both sidebands (see [7] for time domain recursion elimination). Negative frequencies fold back into the positive domain at DC.

The end result is that when the coefficient of a first-order allpass filter is modulated with an audio-rate sinusoid, the output for a sinusoidal input signal can be written as

$$y(n) = b_0 \cos(\omega_x n) + \sum_{k=1}^{\infty} b_k \cos((\omega_x \pm k\omega_m)n), \quad (8)$$

where b_k is the amplitude of partial k . Thus, a sinusoidal input signal is scattered into upper and lower sidebands that are located at each side of the input signal frequency, as in feedback AM [15] (see Fig. 2a). The three terms of (3) contribute partials that are located at f_x , $f_x \pm f_m$, and $f_x \pm kf_m$ ($k > 1$), respectively.

Although detailed formulation of the spectral structure is left for future work, it was observed that the partial amplitudes b_k depend on the modulation index M and decay exponentially with increasing partial numbers k . However, the slope of the decay is slightly steeper in the upper sideband, giving rise to asymmetric spectra. For sinusoidal carrier and modulator signals, the produced spectrum has a maximum bandwidth of

$$B_w = 2Mk_t f_m, \quad (9)$$

where k_t defines the threshold partial number ($k_t = -L_t/20 \log(b_k)$, in which L_t is the threshold level below the reference amplitude in decibels). Note that the bandwidth is narrower than the maximum value given by (9) when $f_c < f_m k_t$, because negative frequencies are mirrored at DC into the positive domain. The maximum frequency component f_{\max} of the output spectrum is

$$f_{\max} = f_x + \frac{B_w}{2} f_m. \quad (10)$$

2.2. Chain of First-Order Stages

The bandwidth given by (9) can be increased by cascading several first-order allpass stages together (see Fig. 2b). The transfer function of the cascade is

$$H(z) = \left(\frac{z^{-1} + a_1}{1 + a_1 z^{-1}} \right)^N, \quad (11)$$

where N is the number of allpass stages in the cascade [4]. Using identical approximation for the delayed modulation signal as was done earlier in (3) gives the difference equation

$$y(n) = \sum_{i=0}^N \binom{N}{i} m^i(n) [x(n - (N - i)) - y(n - i)], \quad (12)$$

where $y(n - i) = 0$ when $i = 0$. For example, setting $N = 1$ will expand the summation into (3), while setting $N = 2$ will give the special case of

$$y(n) = x(n - 2) + 2[x(n - 1) - y(n - 1)]m(n) + [x(n) - y(n - 2)]m(n)m(n). \quad (13)$$

Comparing this with (3), the upper line of (13) is noted to generate a similar spectral structure as the single stage modulation instance, but with increased energy in the sidebands. The lower line of (13) produces a ring modulated version of the sidebands. With sinusoidal modulation signals, this shifts the frequency of the sidebands by an amount equal to the modulation frequency f_m , thus retaining the structure of the spectrum. It is expected that the net effect of the terms on the right side of (13) will be that the bandwidth of the generated spectrum is increased. Observations showed that this was indeed the case and that the frequency-shifted terms, i.e. the higher powers of $m(n)$, balanced the increased sideband energy indicated by the coefficient of the first order term $m(n)$. Furthermore, because the modulation signal is for stability reasons [7] constrained to amplitude values between $[-1 \dots 1]$, the amplitudes of the frequency-shifted sidebands

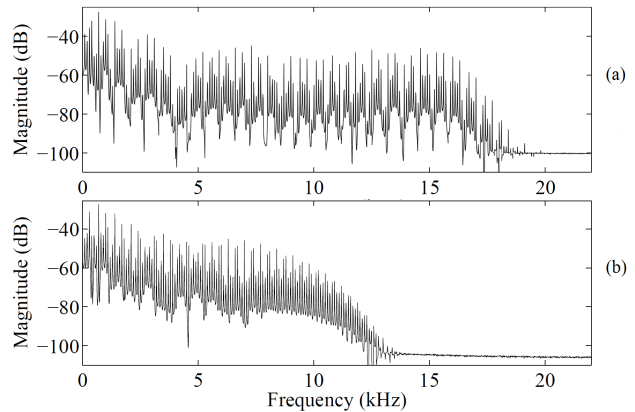


Figure 3: Effect of the modulation index M : (a) $M=0.99$, (b) $M=0.64$. Other parameters are as in Fig. 2(b).

decrease with higher values of the chain length. However, large modulation index M values generate a resonance region in the upper end of the spectrum (see Fig. 3a). Lower M values reduce the amount of high frequency partials as shown in Fig. 3b.

The maximum bandwidth of the produced spectrum for a sinusoidal carrier and modulator signals is

$$B_{w\text{chain}} = 2Mk_t f_m Nc, \quad (14)$$

where c is the bandwidth increment of a single added stage. The chain length N thus allows dynamic control over the maximum bandwidth, and provides a mechanism for high frequency aliasing prevention. The maximum frequency component of the output spectrum is given by (10) by replacing B_w with $B_{w\text{chain}}$.

Another way of looking at the effect of cascading several first-order stages together is to consider the allpass chain as a dispersive variable length delay line. As indicated by (4), each allpass stage in the chain constitutes a variable time delay between the input and output signals. Because this time delay is changing in accordance with the modulation signal, the frequency of the output signal, with respect to the input signal, is changing as well. The instantaneous output signal frequency ω_y can be approximated as

$$\omega_y = \omega_x + \frac{d\phi}{dt}, \quad (15)$$

where ω_x is the angular frequency of the input signal, and ϕ is the phase of the allpass chain. The phase term can be calculated from the angle between the imaginary and real parts of the allpass filter frequency response $H(e^{j\omega})$. Thus, when the allpass filter coefficient a_1 is modulated by $m(t) = M\cos(\omega_m t)$, the instantaneous phase of the chain is given by

$$\phi(t) = N \arctan \left(\frac{(m^2(t) - 1) \sin(\omega_x t)}{2m(t) + (m^2(t) + 1) \cos(\omega_x t)} \right), \quad (16)$$

where N is the length of the chain. Taking the time derivative of (16) and substituting the second term of (15) gives the approximate output frequency

$$\omega_y = \omega_x - \frac{2N\omega_m M \sin(\omega_m) \sin(\omega_x)}{1 + M^2 \cos^2(\omega_m) + 2M \cos(\omega_m) \cos(\omega_x)}. \quad (17)$$

We notice that the chain length N controls the maximum amount of frequency deviation, as shown in Fig. 4, while the modulation index M controls the deviation within the limits imposed by N . M also controls the shape of the applied frequency modulation, which is sinusoidal with low M values, but with higher M values it develops towards the non-sinusoidal forms shown in Fig. 4. It should be noted that the maximum amount of frequency deviation depends also on the modulation frequency ω_m , as indicated by (17).

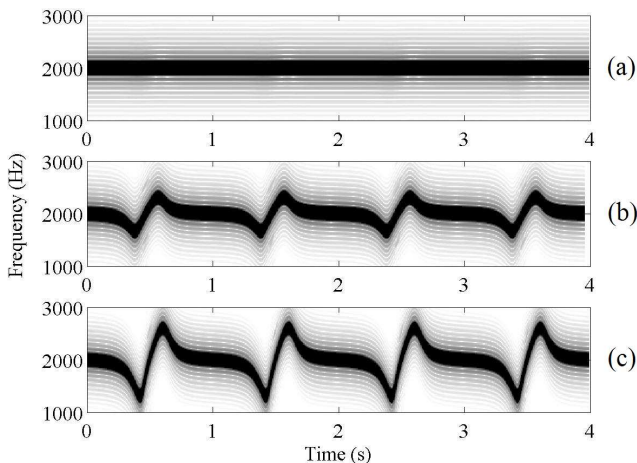


Figure 4: Frequency deviation and the effect of the chain length N . (a) $N=1$: no FM. (b) $N=100$: FM frequency deviation 1 kHz. (c) $N=200$: deviation 1.6 kHz. $M=0.99$, $f_c = 2000$ Hz, $f_m = 1$ Hz.

Because the delay modulation is performed at audio-rate frequencies, the output signal frequency changes rapidly, resulting in FM-like timbres. A similar modulation principle has been used previously in varying the pick-up position of a 1D waveguide [10] and in changing the output tap position of an interpolated variable length delay line [12]. These methods employ a cascade of unit delays and an interpolating read pointer that is modulated at audio-rate for FM-like results. In contrast, FM is an inherent property of CM synthesis, because the delay is modulated directly without a separate read pointer.

To summarize, the output spectrum of CM synthesis consists of components that are produced by a combination of feedback AM and FM processing. A sinusoidal input signal is first scattered into the upper and lower sidebands in the front-most stage of the chain because of the feedback AM connection inside the allpass filter. Succeeding stages then reinforce this effect by processing each previously scattered partial, thus widening the bandwidth of the generated spectrum. On the other hand, because the allpass chain can also be considered as a dispersive variable length delay line, the output signal frequency is changing rapidly with respect to the input signal. This phenomenon works in parallel with the feedback AM processing, adding FM synthesis-like effects to the composite spectrum.

2.3. Aliasing Analysis

Since the proposed structure produces a spectrum using a combination of feedback AM and FM, it is not truly bandlimited, and hence it produces some aliasing distortion. Therefore, the amount

of aliasing distortion generated by the proposed structure must be evaluated somehow. Here, the amount of aliasing distortion is evaluated by computing the noise-to-mask ratio (NMR) [16] of a sawtooth waveform generated by the allpass chain. The NMR utilizes a perceptual model and therefore provides an appropriate measure for the evaluation of the audibility of aliasing distortion. The NMR, originally developed for the evaluation of audio codecs, outputs a ratio of the aliasing distortion to the masking threshold of the non-aliased signal, i.e. the sawtooth waveform obtained by additive synthesis. With NMR, smaller values correspond to lower audible aliasing distortion, and in audio coding a threshold of $\text{NMR} = -10$ dB is usually considered to be the limit of inaudible difference.

Since the aliasing distortion produced by the allpass chain depends on both the length of the chain and the modulation index, it can be effectively set by varying these two parameters. In order to find out how these parameters affect the amount of aliasing, the analysis is split into two parts. First, the varying of the aliasing distortion as a function of the allpass chain length with the modulation index kept constant is analyzed. Then, the effect of the modulation index for the aliasing distortion is analyzed while keeping the allpass chain length constant. The analysis is performed at two frequencies, 120 Hz and 1.2 kHz, in order to get an idea how the aliasing distortion varies at different frequency ranges.

In the first part of the analysis, the modulation index M is set to a constant 0.85 and the length of the allpass chain N is varied from 60 to 110 in steps of 10 in the case of 120 Hz and from 1 to 6 in steps of 1 in the case of 1.2 kHz test frequency. The NMR values of these tests are plotted in Fig. 5a. It can be seen, that the NMR of the 120 Hz test case remains almost constant as the allpass chain length is varied, whereas the NMR of the 1.2 kHz test case increases somewhat linearly as the length of the allpass chain is increased. This is an expected result, as a longer chain increases the bandwidth of the output signal, thus producing more aliasing distortion. In addition, at low frequencies the aliasing folded around the Nyquist frequency components is already at quite a low level, and therefore the increase of the chain length, even a rather large increase, does not increase the aliasing distortion as much as at higher frequencies where the aliasing is already more pronounced.

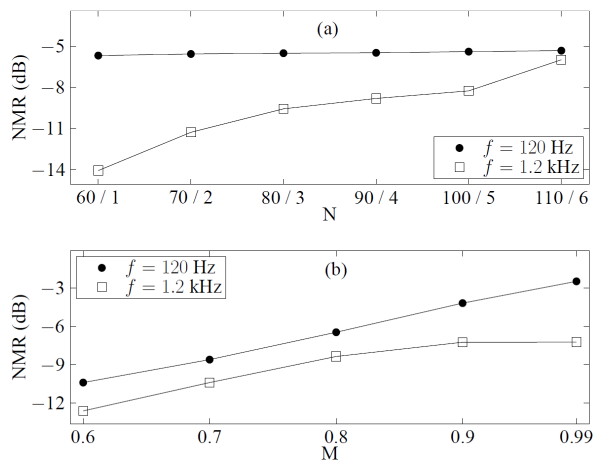


Figure 5: Aliasing effects of the (a) chain length N , and the (b) modulation index M .

In the second part of the analysis, the allpass chain length N is set to a constant 100 in the 120 Hz test case and to a constant 6 in the 1.2 kHz test case, and the modulation index M is varied from 0.6 to 0.9 in steps of 0.1. In addition, a modulation index 0.99 was also tested. The NMR values of these tests are plotted in Fig. 5b, in which it can be seen that the NMR values of both the low and the high frequency tests increase almost linearly as the modulation index increases. Yet again, this is an expected result, as with a larger modulation index the allpass chain applies more phase distortion to the input signal which causes more aliasing. As can be seen from Fig. 5, the aliasing of the sawtooth is quite moderate with these parameters. However, as a trade-off the levels of the sawtooth partials are far from the levels of the ideal case. This is illustrated in Fig. 6, where the error of the partial amplitudes is plotted for both test cases. Now, smaller values of amplitude error indicate a more accurate match to the partial amplitudes of the ideal sawtooth. However, it should be noted that when either the allpass chain length or the modulation index is increased, the amplitude error decreases. Yet, as discussed above, this leads to more aliasing distortion.

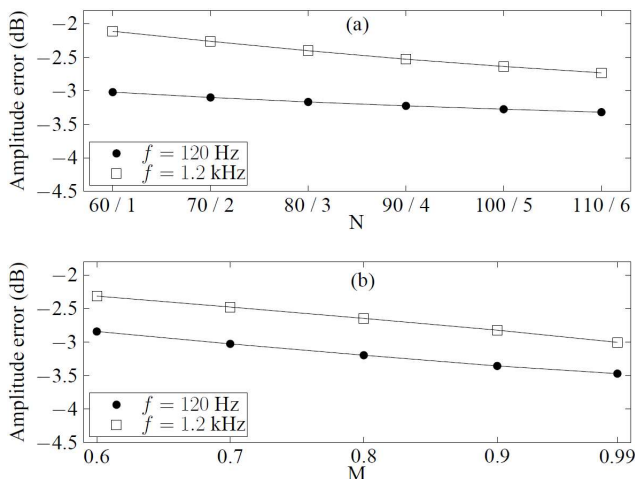


Figure 6: Comparison of the synthesized pseudo sawtooth and the ideal sawtooth partial amplitudes as a function of the (a) chain length N and the (b) modulation index M .

2.4. Relation to AM/RM and FM

This section compares the spectral properties of AM/RM, FM, and CM methods using sinusoidal carrier and modulator signals. The classic amplitude modulation spectrum has two (RM) or three (AM) partials around the carrier frequency, see Fig. 7a. Single stage CM produces a feedback AM-like spectrum, which is also centered at the carrier frequency, but has a larger bandwidth than the corresponding AM/RM spectrum. Moreover, the bandwidth of the CM spectrum can be further expanded by cascading the stages in series.

CM also applies an additional complex-modulator FM to the carrier signal, as shown in Fig. 7c. The instantaneous frequency of the carrier is thus changed in a different pattern than in sinusoidal-modulator FM, which is shown in Fig. 7b. Therefore, CM synthesis has a different spectral structure than sinusoidal FM.

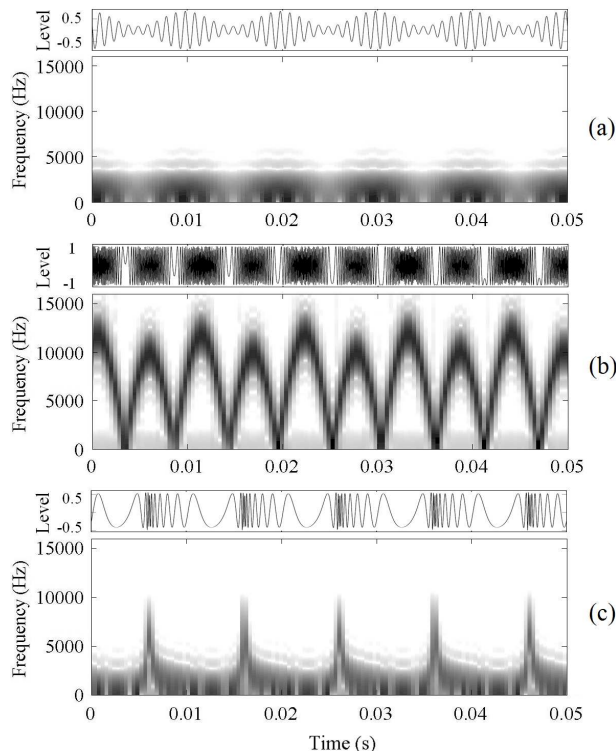


Figure 7: Comparison of (a) AM, (b) sinusoidal FM, and (c) CM synthesis. In each subfigure, the amplitude domain waveform is plotted on top of a frequency domain spectrogram – time runs on the horizontal axis. Sinusoidal carrier and modulator, $f_x = 1000$ Hz, $f_m = 100$ Hz, CM chain length $N = 70$.

Because CM synthesis is a combination of feedback AM and FM methods, the CM spectrum is often dense, and therefore well suited for thick and rich subtractive synthesis –like spectrum generation. However, because of this spectral density, partials reflected from DC are often perceived as tremolo effects rather than inharmonic components of the spectrum. For this reason, FM is more flexible for inharmonic spectrum generation.

2.5. Synthesis Parameters and Their Effects

Table 1 lists the synthesis parameters of the proposed method.

Table 1: CM synthesis parameters.

parameter	description	main effect
f_x/f_m	frequency ratio	spectral structure
M	modulation index	bandwidth
N	allpass chain length	bandwidth
f_x	source frequency	fundamental freq.
A_x	source amplitude	output amplitude
θ_m	phase difference	partial amplitudes

The frequency ratio f_x/f_m defines the spectral structure of the synthesized sound. Consonant ratios, i.e., when the ratio can be expressed as n/m using small integer n and m values, produce harmonic timbres. Irrational frequency ratios produce inharmonic spectra. A slight detuning between f_x and f_m results in beating.

The modulation index M is the counterpart of FM modulation index I , or the cutoff frequency parameter of a subtractive synthesis lowpass filter. It controls the amount of higher frequency partials.

The allpass chain length N controls the maximum bandwidth of the spectrum. It has also an impact on the maximum phase deviation of the FM modulation effect.

The fundamental frequency is, in most cases, defined by the source signal frequency f_x . However, certain f_x/f_m frequency ratios produce spectra that contain a more prominent partial at a lower frequency value than f_x . The overall amplitude of the synthesized sound can be controlled with the source signal amplitude A_x .

The phase difference between the source signal and the modulator can be controlled by θ_m . It has only a subtle timbral effect when manipulated at control rate, but is capable of producing a wide variety of timbres when modulated at audio rate.

3. IMPLEMENTATION

3.1. PC

Equation (3) was translated into C language and augmented with a loop that iterates through the chained allpass stages. The loop is iterated for each sample, and it contains one multiplication, two additions, and two state variable accesses per allpass stage.

The algorithm was then implemented as a Pure Data (Pd) external [17], with two audio rate inlets (source and coefficient modulation signals) and one control rate inlet for manipulating the allpass chain length N , see Fig. 8. The external was utilized in several Pd patches implementing the example instruments of Section 4.

A simplified version of the extended Karplus-Strong (EKS) algorithm [18] was implemented as another custom Pd external. The EKS algorithm was then augmented by inserting a CM allpass chain into the delay loop of the string model.

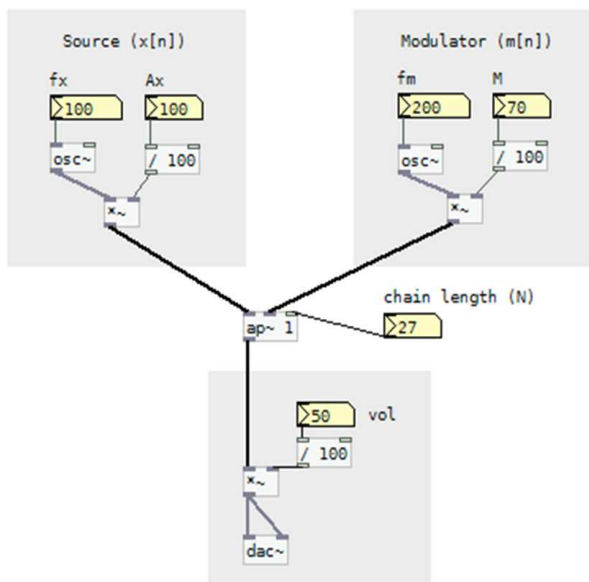


Figure 8: Pd patch used for the synthesizer sounds. The CM algorithm is implemented inside the `ap~` external, which takes two audio rate and one control rate input.

3.2. Mobile

A mobile version of the algorithm was implemented using fixed-point arithmetic. It was further optimized at C++ code level and retrofitted into a mobile audio synthesis framework developed in the SAME project [19]. A Nokia N95 smartphone was used as the hardware platform. The implementation runs on top of Symbian S60 v3.2, and does not access the DSP chip directly.

The fixed point implementation required careful signal scaling analysis. It revealed that because the signal travelling inside the chain can be over-modulated to a certain extent, it should be scaled only after it exits the last allpass stage. Otherwise, the increased number of allpass filters will eventually silence the incoming signal. This happens because the chain has a damping effect on the processed signal, as allpass filter coefficients are less than 1 in magnitude. Mobile fixed point implementation issues will be explored further in future work.

4. EXAMPLE APPLICATIONS

All examples below use sinusoidal source and modulating signals. The Pd externals, source code, patches, and sound examples are available at

<http://www.acoustics.hut.fi/publications/papers/dafx09-cm/>

4.1. Synthesizer Sounds

The described synthesis method is capable of synthesizing dense, complex, and evolving spectra by using the simple instrument of Fig. 8, or by cascading two or more allpass chains in series and providing a dedicated modulator for each chain (see Fig. 9). Parallel chain arrangements can be used, e.g., in formant synthesis. The method works with complex source waveforms, such as two oscillator FM or sawtooth timbres as well.

It is also possible to imitate classic virtual analog waveforms by using the instrument of Fig. 8 with frequency ratios 1:1 and 1:2 between the carrier and the modulator signals, and emulating the lowpass filter cutoff frequency parameter of the subtractive

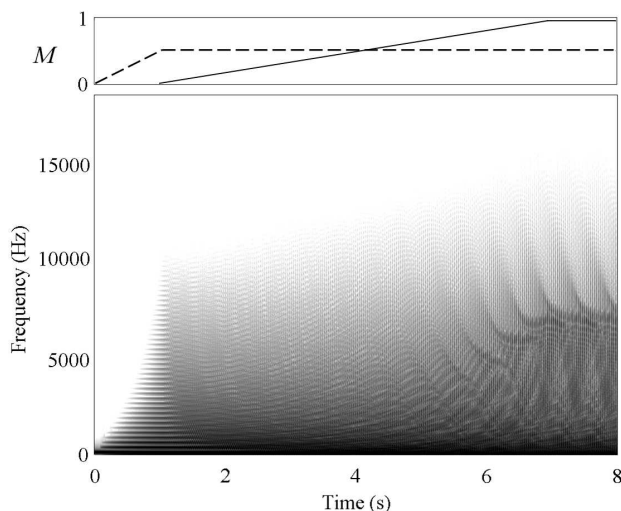


Figure 9: Two allpass chains in a cascade produce complex dynamic spectra. Modulation indices ramp from 0 to 0.5 for the first chain and from 0 to 0.99 for the second.

synthesis technique with the modulation index M (see Fig.10). More complex carrier/modulator frequency ratios produce a variety of other waveforms, while a slight detuning between the carrier and the modulator generates a beating effect between the spectral components.

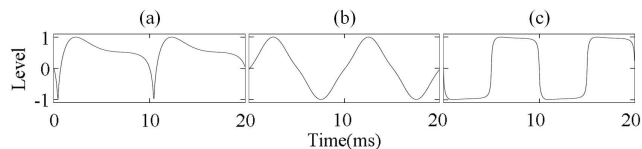


Figure 10: *Imitated virtual analog waveforms: (a) Pseudo sawtooth, (b) triangle, and (c) square.*

4.2. Physical Models

CM synthesis can be used in conjunction with physical modeling techniques either as a post-processing effect located outside the model, or inside the model itself. For example, when the coefficient modulated allpass chain is inserted into the feedback loop of a simple EKS string model, each harmonic is augmented with upper and lower sidebands. If the modulator frequency is slightly detuned from the fundamental frequency of the string, the harmonics will produce beating, resulting in chorus-like timbres (see Fig. 11). A similar beating phenomenon can be observed when the pick-up point [10], or the dispersion amount [8] parameter of a waveguide string model is modulated at audio-rate frequencies.

It was noticed that the generated sound had a similar character independent of whether the chain was located inside or outside the model. However, the produced effect was more pronounced when the chain was inserted inside the model. The allpass chain length can also be significantly shorter in the inside configuration, due to the delay element present in the model.

The implementation revealed that, in order to suppress the generation of excessive high frequency content, the allpass chain should be placed in front of the loss filter.

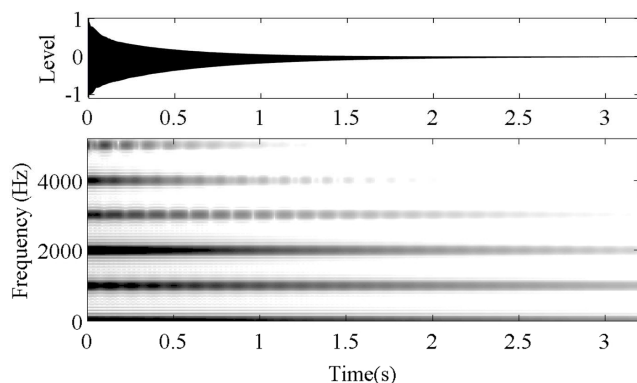


Figure 11: *Coefficient modulated allpass chain inserted into the feedback loop of a simple string model. Beating harmonics animate the produced sound. $f_0=1000$ Hz and $f_m=1010$ Hz, white noise excitation*

4.3. Effect Processing of Arbitrary Signals

For (monophonic) pitched instruments, the proposed structure can be utilized also as an approximate adaptive frequency modulator [12] which is similar to the phase distortion effect described in [5-7]. In this application, the modulator signal can be composed by driving a sinusoidal oscillator whose frequency is estimated from the input signal. There are several approaches on the fundamental frequency estimation (see, e.g., [20]), ranging from computationally efficient to more sophisticated algorithms. Unfortunately, the number of correct estimates is proportional to the complexity of the algorithm, i.e. the computationally efficient approaches produce more false estimates than the more sophisticated ones.

The fundamental frequency estimation can be avoided if a lowpass filtered version of the input signal is used as the modulator. This approach is computationally efficient, but the modulator signal will then be spectrally richer, thus applying more distortion to the input signal and causing more aliasing distortion. In addition, if the lowpass filter also suppresses the fundamental frequency of the input signal, the modulator signal may then be corrupted by the low frequency noise present in the input. However, as a good compromise, the lowpass filtering can be utilized as a preprocessing step before a computationally efficient fundamental frequency estimator. The lowpass filtering can increase the robustness of the otherwise non-robust estimator, yielding to a still computationally efficient approach and lowered distortion when compared to the lowpass filtered modulator.

The method seemed to be most useful with plucked string, organ, and monophonic wind instrument timbres, whereas inharmonic sounds, such as bells and percussion instruments, did not benefit from the effect. Synthesizer sounds can also utilize the effect, provided that there is free space in the source spectrum.

5. DISCUSSION

The strength of the CM synthesis technique lies in the generation of thick, complex, and evolving timbres, although it can also be used in building a virtual analog synthesizer from sinusoidal oscillators and allpass filters. Simple but powerful control parameters allow efficient control rate manipulation of the produced sound and the ability to control the maximum bandwidth prevents upper range aliasing artifacts.

CM synthesis has a characteristic sound, which makes it less generic than sinusoidal FM. Therefore, the method is not well suited for acoustic instrument emulation. It also requires a pitch detector to work properly as an audio effect.

The method is quite CPU intensive, because each processed sample has to be routed through all chained allpass stages. Nevertheless, it is also possible to run the synthesis algorithm in mobile smartphones.

The aliasing distortion of the proposed structure can be set by choosing the allpass chain length and the modulation index appropriately. Yet, it was noted that the aliasing distortion is more sensitive to variations in the modulation index at all frequencies whereas the allpass chain affects the amount of aliasing distortion more as the frequency of the input signal is increased.

6. CONCLUSION AND FURTHER WORK

CM synthesis, which is based on audio-rate modulation of allpass filter coefficients, was introduced as a new sound generation method. This idea has been inspired by the recent introduction of spectral delay filters [4] and it can be seen as its time-varying extension. CM synthesis is related to previous modulation synthesis techniques, such as AM, FM, and RM. It appears that the CM technique using one first-order allpass filter is capable of producing similar spectra to the feedback AM method. Moreover, when a cascade of several identical first-order allpass filters is employed, as in spectral delay filters, sonic effects combining the characteristics of feedback AM and complex-modulator FM arise.

It was shown that the CM synthesis is a powerful and versatile tool. It can generate FM synthesis-like sound and also imitations of classical waveforms, such as the sawtooth, triangular, and rectangular pulse waves. Additionally, it can be used for adding chorus-like effects to simple synthesis techniques, such as the Karplus-Strong string algorithm. Furthermore, the CM method can modify arbitrary monophonic sources, such as musical instrument tones, when a pitch detector is used to estimate the modulation frequency. It is thus related to recent adaptive FM and split sideband techniques [12], [21].

The spectral structure of the produced sound is not yet entirely understood and more work is needed on the derivation of the partial amplitudes. Space issues prevented also more detailed description of the mobile implementation. Interesting extensions worth further exploration include the modulation of the phase difference between the carrier and the modulator oscillators, the effect of non-uniform modulation amount in the chain, and the arrangement of multiple CM chains into various topologies.

7. ACKNOWLEDGEMENTS

The authors would like to thank Prof. J. O. Smith for pointing out inspiring related research material. This work has been co-funded by the Academy of Finland (projects no. 122815, no. 126310, and no. 128689), and by the European Union (EU) as part of the 7th Framework Programme with the SAME project (ref. 215749).

8. REFERENCES

- [1] J. O. Smith, "An allpass approach to digital phasing and flanging", *Rep. STAN-M-21*, Center for Computer Research in Music and Acoustics (CCRMA), Dept. of Music, Stanford University, Stanford, CA, 1984.
- [2] 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, November 2000.
- [3] 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, paper no. 6954, Oct. 2006.
- [4] V. Välimäki, J.S. Abel, and J.O. Smith, "Spectral Delay Filters", *J. Audio Eng. Soc.*, vol. 57, no. 7/8, July/August 2009.
- [5] J. Timoney, V. Lazzarini, J. Pekonen, and V. Välimäki, "Spectrally rich phase distortion sound synthesis using an allpass filter", in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2009)*, Taipei, Taiwan, April 19-24, 2009.
- [6] V. Lazzarini, J. Timoney, J. Pekonen, and V. Välimäki, "Adaptive Phase Distortion Synthesis", in *Proc. 12th Int. Conf. on Digital Audio Effects (DAFx-09)*, Como, Italy September 1-4, 2009.
- [7] J. Pekonen, "Coefficient-modulated first-order allpass filter as distortion effect", in *Proc. 11th Int. Conf. on Digital Audio Effects (DAFx-08)*, Espoo, Finland, September 1-4, 2008, pp. 83-87.
- [8] J. Kleimola, "Dispersion modulation using allpass filters", in *Proc. 11th Int. Conf. on Digital Audio Effects (DAFx-08)*, Espoo, Finland, September 1-4, 2008, pp. 193-197.
- [9] H. Penttinen, A. Härmä, and M. Karjalainen, "Digital Guitar Body Mode Modulation with One Driving Parameter", in *Proc. COST-G6 Conf. on Digital Audio Effects (DAFx-00)*, Verona, Italy, December 7-9, 2000, pp. 31-36.
- [10] S. A. Van Duyne and J. O. Smith, "Implementation of a variable pick-up point on a waveguide string model with FM/AM applications", in *Proc. Int. Computer Music Conf.*, San Jose, 1992.
- [11] J. Pakarinen, M. Karjalainen, and V. Välimäki, "Modeling and real-time synthesis of the kantele using distributed tension modulation", in *Proc. Stockholm Music Acoustics Conf. (SMAC'03)*, pp. 409-412, Stockholm, Sweden, Aug. 6-9, 2003.
- [12] 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, 2008.
- [13] T. Nishimoto, "Electronic musical instrument using amplitude modulation with feedback loop", U.S. Patent 4,655,115, April 1987.
- [14] R. Greenfield, "Increased precision digital filter coefficients using digital dither", *Electronics Letters*, vol. 34, no. 5, pp.444-446, Mar. 1998.
- [15] V. Lazzarini, J. Timoney, J. Kleimola, and V. Välimäki, "Five variations on a feedback theme", in *Proc. 12th Int. Conf. on Digital Audio Effects (DAFx-09)*, Como, Italy September 1-4, 2009.
- [16] K. Brandenburg, "Evaluation of Quality for Audio Encoding at Low Bit Rates", in *Proc. AES 82th Convention*, London, UK, Feb. 1987.
- [17] M. Puckette, *The Theory and Technique of Electronic Music*, World Scientific Press, 2007.
- [18] D.A. Jaffe, and J.O. Smith, "Extensions of the Karplus-Strong Plucked-String Algorithm", *Computer Music J.*, Vol. 7, No. 2 (Summer, 1983), pp. 56-69.
- [19] *SAME Project* (Sound and Music For Everyone Everyday Everywhere Everyway), Homepage at <http://www.sameproject.eu/>.
- [20] D. Gerhard, "Pitch Extraction and Fundamental Frequency: History and Current Techniques", *Technical Report TR-CS 2003-06*, Department of Computer Science, University of Regina, Regina, Canada, November 2003.
- [21] 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, 684-695, Sept. 2008.