MUSICAL SIGNAL ANALYSIS USING FRACTIONAL-DELAY INVERSE COMB FILTERS

Vesa Välimäki, Heidi-Maria Lehtonen

Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing
Espoo, Finland
vesa.valimaki@tkk.fi
heidi-maria.lehtonen@tkk.fi

Timo I. Laakso

National Board of Patents and Registration of Finland, Patents and Innovations Line
Helsinki, Finland
timo.laakso@prh.fi

ABSTRACT

A novel filter configuration for the analysis of harmonic musical signals is proposed. The method is based on inverse comb filtering that allows for the extraction of selected harmonic components or the background noise component between the harmonic spectral components. A highly accurate delay required in the inverse comb filter is implemented with a high-order allpass filter. The paper shows that the filter is easy to design, efficient to implement, and it enables accurate low-level feature analysis of musical tones. We describe several case studies to demonstrate the effectiveness of the proposed approach: isolating a single partial from a synthetic signal, analyzing the even-to-odd ratio of harmonics in a clarinet tone, and extracting the residual from a bowed string tone.

1. INTRODUCTION

Analysis of the amplitude envelope of harmonic components of a musical tone is a fundamental operation in musical signal processing. We discuss the harmonic extraction using the digital filtering approach. This is an old technique that has been proposed in different forms by Moorer in the 1970s for pitch detection of speech signals [1] and for analyzing music data for additive synthesis [2]. The basic idea is to use a multi-notch filter to extract individual harmonic components as signals. The filter structure may be obtained as the inverse transfer function of a comb filter (i.e., a delay line in a feedback loop).

In this paper we expand on a recently proposed idea that the delay line can be replaced with a high-order allpass fractional-delay filter to obtain very accurate cancellation of neighboring harmonics to extract a single harmonic [3]. The proposed signal analysis method is useful for many practical cases. Numerous musical instruments, including all woodwind, brass, and bowed string instruments, produce a sound signal that is inherently harmonic, i.e., the spectral components are integral multiples of a fundamental frequency. This follows from the sound-production mechanism of these self-excited systems, which involves mode locking in the time domain [4]. It forces the sustained tones of such instruments to be periodic. There is often a noise component in these musical tones making them pseudo-periodic in practice.

Another method for this kind of signal decomposition is sinusoidal modeling [5], [6], [7]. In this method the signal is analyzed using the windowed FFT, and the frequency and amplitude tracks are obtained by connecting data in the neighboring analysis frames. This approach has its roots in the phase vocoder technique and its efficient transform-domain implementation. For periodic or pseudo-periodic musical tones it is unnecessary to get down to an overly generic analysis method, because the frequencies of the harmonic components are known after the estimation of the fundamental frequency. Advantages of the proposed filter-based analysis method – compared with the more general FFT-based techniques – are simplicity, which follows mainly from the small number of parameters, and the possibility of designing filter coefficients in closed form. Additionally, the resulting decomposition is obtained directly as a set of time-domain signals, and no separate synthesis stage is required.

Other signal processing methods proposed for analyzing the harmonic structure of musical signals include wavelets [8] and high-resolution tracking methods [9], [10]. These methods provide excellent frequency accuracy at the expense of a complicated algorithm and a high computational cost. The method proposed in this paper can also provide amplitude and frequency accuracy that is sufficient for musical signal analysis but at the same time the analysis method remains easy to apply.

This paper is organized as follows. Section 2 discusses the filter structure for canceling harmonics of a musical signal, and Section 3 introduces a filter structure for extracting a single harmonic component and another structure for separating even and odd harmonics. In Section 4, three test cases are presented to demonstrate the power of this approach in musical signal analysis.

2. FRACTIONAL-DELAY INVERSE COMB FILTERS

The inverse comb filter[1] (ICF) is an FIR filter where the input signal is delayed by L samples and is then subtracted from the original input signal, see Fig. 1(a). The corresponding transfer function is $H(z) = (1 - z^{-L})/2$, where the scaling factor ½ sets the gain to unity in the passband (i.e., between the notches). The magnitude response of this filter features periodic notches at the multiples of f_s/L, where f_s is the sampling rate (Hz) and L is the delay line length in samples, or multiples of the sampling interval.

When the delay line length is restricted to be an integral multiple of the sample interval, the accuracy of the notch frequencies can be poor. An example is shown in Fig. 2 where the fundamental frequency is 4186 Hz and the corresponding period length is 10.5351 samples. Practical ICF implementations employ a fractional-delay filter that replaces the delay line [11], [13], [14]. Alternatively, an FIR [15], [16] or an IIR notch filter [17], [18] can be designed to approximate the overall ICF characteristics.

Figure 1(b) shows the block diagram of a fractional-delay ICF, where the delay line is replaced with an allpass filter, as pro-

[1] Following the convention of [11], the term 'inverse comb filter' is used for the feedforward system with a delay line. The 'comb filter' has a delay line inside a feedback loop.

posed previously [3]. The transfer function of this system can be written as $H_{fd}(z) = [1 - A(z)]/2$, where $A(z)$ is the transfer function of the allpass filter used for delay approximation. A magnitude response of this structure with an 11th-order allpass filter that approximates the delay of 10.5351 sampling intervals is displayed in Fig. 2 (solid line).

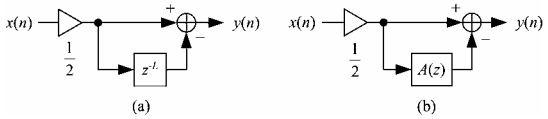

Figure 1: *(a) Conventional ICF and (b) a fractional-delay allpass-filter based ICF (after [3]).*

The transfer function of a digital allpass filter is

$$A(z) = \frac{z^{-N} D(z^{-1})}{D(z)} \tag{1}$$

where N is the order of the filter and $D(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \ldots + a_N z^{-N}$ is the denominator polynomial with real-valued coefficients a_k, and the numerator polynomial is a reversed version of the denominator. The symmetry of the numerator and denominator coefficients guarantees the exact allpass property even for rounded coefficients. In this application, the allpass filter order is typically N = round(L), which is also approximately the period length (in samples) to be cancelled. Therefore, the filter order N can be very high, such as $N = 1000$ for a low fundamental frequency of 44.1 Hz when the sampling rate is 44.1 kHz. Evidently, a method is needed that allows for the design of high-order filters.

We propose two new structures, which are presented in Fig. 3. These filter structures offer freedom in the selection of the allpass filter order, which was related to the fundamental period in a previous work [3]. We have found experimentally that the order of $A(z)$ may be kept constant (e.g., $N = 80$), so when the fundamental period ($T_0 = f_s/f_0$) is longer than N samples, L extra samples of delay are required in the lower signal path in Fig. 3(a). However, when the fundamental period is shorter than N samples, K extra samples are required in the upper signal path to synchronize signals for subtraction, see Fig. 3(b). Thus, we propose to use the transfer function

$$H_{\text{low}}(z) = \frac{1}{2}\left[1 - z^{-L} A(z)\right] \tag{2}$$

when the fundamental period T_0 is larger than (or about the same as) the allpass filter order N, and the transfer function

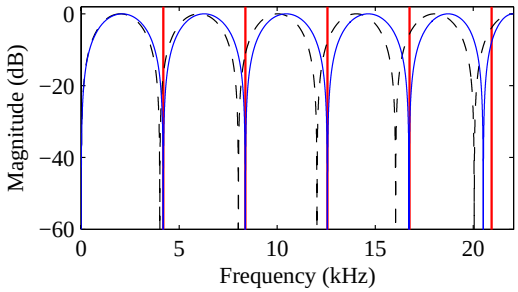

Figure 2: *Magnitude response of the conventional (dashed line) and the allpass-based (solid line) ICF. The thick vertical lines indicate the harmonic frequencies to be cancelled (f_0 = 4186 Hz).*

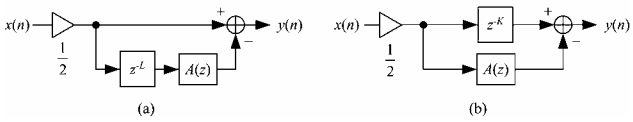

Figure 3: *Fractional-delay ICF structures that allow the use of an allpass filter A(z) of arbitrary order for signals with both (a) low and (b) high fundamental frequency.*

$$H_{\text{high}}(z) = \frac{1}{2}\left[z^{-K} - A(z)\right] \tag{3}$$

when the fundamental period T_0 is smaller than the allpass filter order N. The delay-line lengths L and K are determined as follows:

$$L = T_0 - N - d \text{ , when } T_0 \geq N \tag{4}$$

$$K = N + d - T_0 \text{ , when } T_0 < N \tag{5}$$

where $-1 < d < 1$ is the fractional-delay parameter used in designing the allpass filter.

2.1. Properties of allpass fractional-delay inverse comb filters

Let us consider the properties of the allpass-based ICF structure of Fig. 3. For simplicity, we will consider the case where $K = L = 0$, i.e., no additional delay is present in either branch, as in Fig. 1(b). We can express the transfer function (2) in the form

$$H(z) = \frac{1}{2}\left[1 - A(z)\right] = \frac{1}{2}\left[\frac{D(z) - z^{-N} D(z^{-1})}{D(z)}\right] \tag{6}$$

where the numerator polynomial can be written as

$$B(z) = D(z) - z^{-N} D(z^{-1}) = b_0 + b_1 z^{-1} + \ldots + b_N z^{-N} \tag{7}$$

and

$$b_k = a_k - a_{N-k}, \quad k = 0,1,\ldots,N \text{ .} \tag{8}$$

It is easy to verify that $B(z)$ is an *antisymmetric* polynomial, i.e., $b_k = -b_{N-k}$. In fact, the ICF is a special case of the parallel connection of two allpass filters discussed by Saramäki in [18]. As shown in this seminal paper, under general assumptions (stable allpass functions) the overall numerator polynomial of this structure is either symmetric or antisymmetric (mirror-image or anti-mirror-image polynomial, respectively). These polynomials are known to have their zeros exactly on the unit circle. Hence, our filter is known to have accurate zeros also in the fractional-delay ICF case. Note that the allpass filters are exactly allpass (unity magnitude at all frequencies) even if the phase (or phase delay, or group delay) is only approximately as desired.

The zeros of the conventional ICF (integer delay $L = L_0$) are known to be uniformly distributed on the unit circle:

$$H(z) = \frac{1}{2}(1 - z^{-L}) = 0 \quad \Leftrightarrow \quad z_n = e^{jn2\pi/L}, n = 0,1,\ldots,L-1 \tag{9}$$

In addition, there are L poles at $z = 0$. The frequency response of the ICF is obtained in the form

$$H(e^{j\omega}) = j e^{-j\omega L/2} \sin(\omega L/2) \equiv H_0(e^{j\omega}) \tag{10}$$

which will be used as a reference when comparing against another ICF structure.

For the fractional-delay ICF ($L = L_0 + d$, d real-valued) the zeros are difficult to express in general, as they depend on the approximating allpass filter of (6). However, the following notation is useful:

$$H(z) = \frac{1}{2}\left[1 - A(z)\right] = \frac{1}{2}\left(1 - z^{-L}\right) + \frac{1}{2}\left[z^{-L} - A(z)\right] \qquad (11)$$

where the latter term represents the error due to the allpass filter approximation. Note that in the z-transform formulation (11), the term z^{-L} with a noninteger power is non-realizable. A more practical expression is the frequency-domain for

$$\begin{aligned} H(e^{j\omega}) &= je^{-j\omega L/2}\sin(\omega L/2) + \frac{1}{2}\left[e^{-j\omega L} - A(e^{j\omega})\right] \\ &= H_0(e^{j\omega}) + \frac{1}{2}\left[e^{-j\omega L} - e^{j\phi_A(\omega)}\right] \end{aligned} \qquad (12)$$

The latter term utilizes the exact unit magnitude property of the allpass functions, which enables the expression with the corresponding phase function only. Hence, (12) illustrates the error term in the fractional-delay ICF frequency response caused by allpass phase approximation and the deviation from the ideal linear phase. As the phase approximation errors of the allpass filter tend to accumulate with increasing frequency, also the zeros of the corresponding ICF are more off the ideal places at higher frequencies.

2.2. Allpass fractional-delay filter design

Three closed-form design methods are known for fractional-delay allpass filters: the Thiran allpass filter design [20], [19], [11], the truncated Thiran allpass filter [21], and the Pei-Wang method [22]. Such methods are needed to increase the allpass filter order to be large enough for good wideband approximation in audio applications. Both the standard and the truncated Thiran methods allow the filter order to be increased up to $N = 1029$ (when $d = -0.5$) using 64-bit double floating-point computing.

The Thiran design formula can be expressed as

$$a_k = (-1)^k \binom{N}{k} \prod_{n=0}^{N} \frac{d+n}{d+k+n} \quad \text{for } k = 1, 2, 3, ..., N, \qquad (13)$$

where N is the filter order and d is the fractional delay parameter ($-0.5 < d \leq 0.5$). At low frequencies, this filter has the phase delay of $N + d$ samples. This design method was used to produce Fig. 2 and Fig. 4(b) with parameter values $N = 11$ and $d = -0.4649$.

The truncated Thiran design is obtained by modifying (13):

$$a_k = (-1)^k \binom{M}{k} \prod_{n=0}^{M} \frac{d+n}{d+k+n} \quad \text{for } k = 1, 2, 3, ..., N, \qquad (14)$$

where M is the prototype filter order ($M > N$) [21]. By using a value for M that is much larger than N in (14), it is possible to extend the bandwidth of good approximation. This comes at the expense of losing quality at low frequencies: the approximation error is larger than in the original allpass filter. This design technique allows a useful tradeoff between approximation accuracy and bandwidth, as discussed in [21] and [3].

Figure 4 compares the standard and truncated Thiran allpass filters. The design parameters are $N = 80$ and $d = -0.5$ for both filters, and the prototype filter order for the truncated Thiran filter is $M = 9N = 720$. The magnitude response, which is exactly flat in both cases, the phase delay (i.e., the negative phase function divided by angular frequency), and the frequency-response error (i.e., difference between frequency responses of the allpass filter and the ideal fractional delay element $e^{-j\omega D}$) are displayed. It is seen that the difference between the fractional-delay approximation of the two filters is microscopic below about 17 kHz, see Fig. 4(b), but the relaxed accuracy allows for the truncated Thiran filter to perform significantly better above 17 kHz, see Fig. 4(c).

A comparison of two ICFs based on allpass filters is shown in Fig. 5. The same design parameters were used as in Fig. 4, and the delay-line lengths were chosen to be $L = K = 0$. The impulse responses of the two ICFs are very similar, but not identical (see Fig. 5(a)). The magnitude responses in Fig. 5(b) are also nearly identical except at frequencies close to the Nyquist limit. In Fig. 5(c) it is seen that below 17 kHz the ICF using the truncated allpass filter is worse than the ICF using the standard Thiran filter, but both are sufficiently good, because the attenuation is more than 140 dB. Above 17 kHz the performance of the Thiran ICF collapses, but with the truncated version of the allpass filter the ICF offers an attenuation of 140 dB up to 20 kHz.

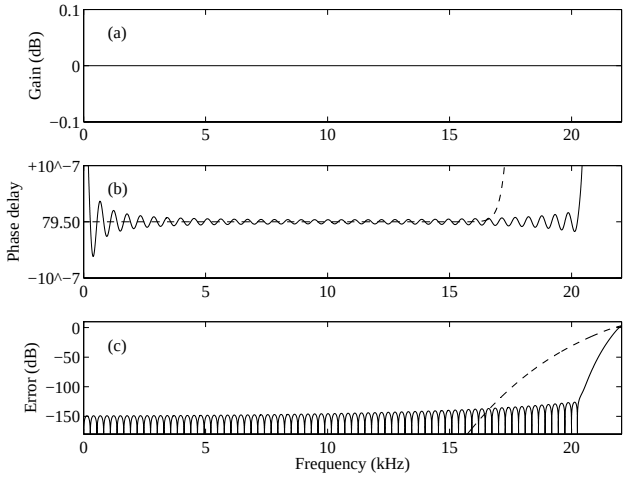

Figure 4: *(a) Magnitude response, (b) phase delay (in samples), and (c) frequency-response error of the Thiran allpass filter (dashed line) and the truncated Thiran allpass filter (solid line). The parameter values are N = 80, M = 720, and d = –0.5.*

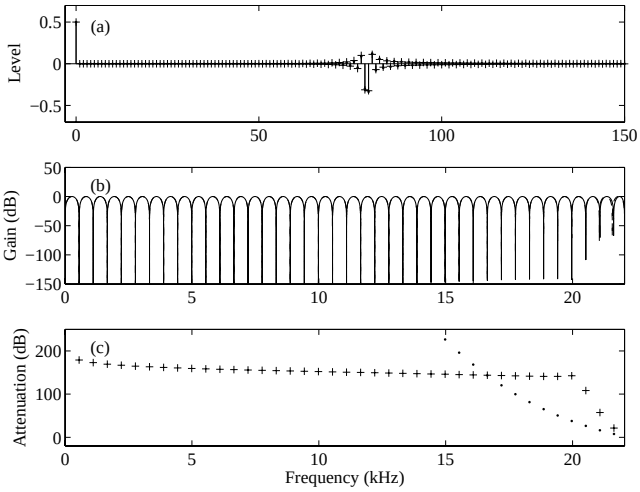

Figure 5: *(a) Impulse responses, (b) magnitude responses, and (c) harmonic attenuation of ICFs with the Thiran ('.', dashed line in (b)) and truncated Thiran allpass filters ('+', solid line in (b)).*

3. EXTRACTING HARMONIC COMPONENTS

Instead of canceling all the harmonic components, single harmonics can be extracted. This is achieved by cascading with an ICF a second-order all-pole filter that cancels a zero at a given harmonic frequency. This section describes the design of such a filter, which we call the *harmonic extraction filter* (HEF).

3.1. Harmonic extraction filter

It is not recommended to place a pole exactly on the unit circle in the z plane, because the resulting second-order filter is marginally stable and the hidden pole may cause numerical problems. A better approach is to move the zeros of the ICF slightly inside the unit circle by defining the radius of all zeros of the transfer function to be $r = 1 - \varepsilon$, where ε is a very small non-negative constant. Consequently, the pole of the second-order filter can also have the same radius, so that the stability of the recursive filter can be assured.

To place all the zeros at radius r, the coefficient of an ICF with a delay-line of length L must have a filter coefficient r^L [23]. Consequently, a scaling coefficient

$$g_0 = 1/(1+r^L) \tag{15}$$

must be used to ensure the maximum gain of the filter to be unity (i.e., 0 dB). Then, the minimum gain of the filter, which occurs at the bottom of the notches at harmonic frequencies, is $g_0(1 - r^L)$, which we call A. We can now solve for the required g_0, and consequently the required r, when the gain A is set to a given value. From

$$A = g_0(1-r^L) = (1-r^L)/(1+r^L) \tag{16}$$

it follows that

$$r^L = (1-A)/(1+A)\,. \tag{17}$$

Since the radius of all zeros is r, the pole radius must also be selected to be r. Based on (17) the radius can be determined to be

$$r = \sqrt[L]{(1-A)/(1+A)}\,. \tag{18}$$

The HEF transfer function can be written as

$$H_{\mathrm{HEF}}(z) = g_1 R(z)[1 - r^L A(z)] \tag{19}$$

where the scaling coefficient g_1 that sets the maximum gain at the bottom of the notches (without the resonator) to be unity is

$$g_1 = 1/\left[r(1-r^L)\right] \tag{20}$$

and the transfer function of the resonant filter is

$$R(z) = b_0/(1 + a_1 z^{-1} + a_2 z^{-2}) \tag{21}$$

with coefficients $b_0 = (1 - r^2)\sin(2\pi f_{\mathrm{res}}/f_{\mathrm{s}})$, which scales the maximum gain of the resonant filter to be unity (see, e.g., [11]), $a_1 = -2r\cos(2\pi f_{\mathrm{res}}/f_{\mathrm{s}})$, $a_2 = r^2$, and f_{res} is the resonance frequency that determines which harmonic component is retained. The filter that has the transfer function (19) with the given scaling coefficients has a maximum gain of 0 dB at the peak of the passband.

When the delay-line length is an integer, the allpass filter $A(z)$ is reduced to a delay line. The zeros of this integer-delay HEF are inside the unit circle, on a smaller circle with radius r:

$$\begin{aligned} &H_{\mathrm{HEF}}(z) = g_1 R(z)\left[1 - r^L z^{-L}\right] = 0 \\ &\Leftrightarrow \quad z_n = r e^{jn2\pi/L},\ n = 0,1,\ldots,L-1 \end{aligned} \tag{22}$$

The transfer function can be given in the form

$$H_{\mathrm{HEF}}(z) = g_1 R(z)\left[r^L(1-z^{-L}) + (1-r^L)\right] \tag{23}$$

which enables the frequency-domain expression

$$H_{\mathrm{HEF}}(e^{j\omega}) = g_1 R(e^{j\omega})\left[2r^L H_0(e^{j\omega}) + (1-r^L)\right] \tag{24}$$

This illustrates the constant (frequency-independent) term due to the zeros being placed inside the unit circle. Since all the zeros are inside the unit circle, the overall magnitude never reaches zero exactly.

Finally, we obtain the expressions

$$H_{\mathrm{HEF}}(z) = g_1 R(z)\left[r^L(1-z^{-L}) + (1-r^L) + r^L(z^{-L} - A(z))\right] \tag{25}$$

and

$$\begin{aligned} &H_{\mathrm{HEF}}(e^{j\omega}) \\ &= g_1 R(e^{j\omega})\left[2r^L H_0(e^{j\omega}) + (1-r^L) + r^L(e^{-j\omega L} - e^{j\phi_A(\omega)})\right] \end{aligned} \tag{26}$$

where both the constant and allpass phase dependent errors in the fractional-delay HEF response are visible.

3.2. Design of parameter values

For good attenuation, it is required that A is sufficiently small and that the resonant filter $R(z)$ accurately cancels one of the zeros of the ICF. For example, when it is required that the inverse comb filter attenuates the harmonic frequencies by 100 dB, the value of A must be set to 10^{-5}, since $20\log_{10}(10^{-5}) = -100$ dB.

In practice, the high-order allpass filter does not provide a perfect phase approximation. Thus, it may be necessary to set A to a smaller value, such as 10^{-6}. However, when the filter structure for selecting a single harmonic is used, the resonant filter provides additional attenuation at frequencies away from the resonance, which further improves the attenuation at the notches.

It was reported in [3] that for some musical instrument tones with strong low-indexed harmonics, the filtering of the signal with the transfer function $H_{\mathrm{HEF}}(z)$ is insufficient. Listening to the filtered signal reveals that the fundamental frequency is still perceived although one of the high-frequency partials is strongly emphasized. Filtering the signal twice with transfer function (19) adequately attenuates the rest of the harmonics in this case.

There is a minor mismatch in the cancellation of the mth transfer function zero with the pole of the resonant filter with the resonance frequency $f_{\mathrm{res}} = mf_0$, because the frequency of the mth zero is offset by the inaccuracy of the phase approximation of the allpass filter. In practice, this mismatch produces a kink around the main lobe of the bandpass filter, and the gain at the resonance frequency becomes larger than 0 dB. A correction to the pole frequency is required to reduce this error.

One way to correct the resonance frequency of the all-pole filter is to search for the minimum of the ICF's magnitude response around the mth notch. For example, computing the magnitude response at 10,000 points between $0.999990mf_0$ and $1.000010mf_0$, and selecting f_{res} as the frequency, where the minimum occurs, reduces the mismatch sufficiently. To reduce the number of magnitude response evaluations, the local minimum can be estimated by using interpolation. After a local minimum on a coarse grid of spectral points has been found, two straight lines can be fit through it and its neighboring points. It may be assumed that the slopes of the notch are symmetrical, so that the frequency of the actual minimum can be found where the two lines cross. This method gives accurate enough results at low computational costs, and considerably improves the performance of the algorithm.

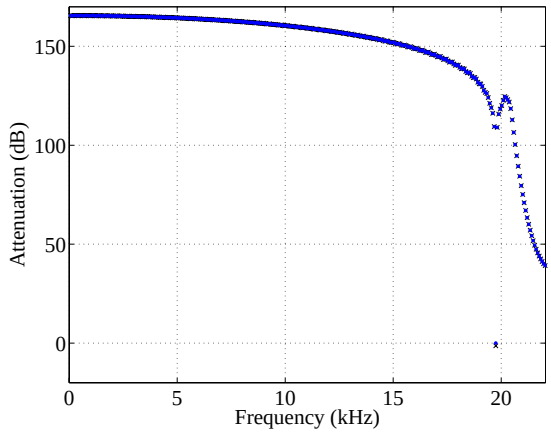

Figure 6: *Attenuation of harmonic partials using the single-harmonic canceling filter when the resonance frequency is the nominal mf_0 ('x'), and the corrected one ('.'). Notice that the largest difference between these data occurs at the frequency of the harmonic #285 at 19.7 kHz.*

Fig. 6 gives an example of the attenuation obtained without and with the proposed correction of the resonance frequency when the fundamental frequency is 69.2957 Hz, the harmonic #285 at 19749.2 Hz is selected, the allpass filter orders used are $N = 80$ and $M = 720$, and attenuation is $A = 10^{-5}$. In this case, the pole radius is $r = 1 - 31.4 \times 10^{-9} = 0.999999969$. The difference between the nominal (mf_0) and the corrected resonance frequency is 13.5×10^{-3} Hz, but the attenuation of the partial is 1.4 dB without and 0.0051 dB with the correction. This difference is enough to make the correction worth the effort, since it makes the analysis filter an accurate tool for signal analysis.

For comparison, we designed a linear-phase FIR bandpass filter that imitates the obtained magnitude response. The filter was designed by using the Chebyshev window with a sidelobe level of –100 dB. To extract the harmonic #285 and to obtain an attenuation of more than 100 dB for all other harmonics, the smallest filter order is 4657. The proposed allpass-filter based ICF of order 80 is computationally much less expensive. Its number of filter coefficient is 1.7% of that of the FIR filter.

3.3. Separation of odd and even partials

While it is possible to cancel the harmonic components one by one by applying the above HEF structure multiple times, alternatively even and odd harmonics may be separated using a single filtering operation, as suggested in [3]. The odd and even harmonics can be separated by first canceling the even harmonics using the fractional-delay inverse comb filter and then subtracting the resulting signal from the original.

The structure of Fig. 3 can be used, but the delay to be approximated is half of that used in canceling all harmonics, i.e., $f_s/2f_0$ samples. With this delay, the notches are located at the multiples of the second harmonic, and the filter now cancels the even harmonics and preserves the odd harmonics. The signal containing even harmonics is then obtained by subtracting the estimated odd harmonics from the original signal, as shown in Fig. 7:

$$s_{\text{even}}(n) = s_{\text{orig}}(n) - s_{\text{odd}}(n). \tag{27}$$

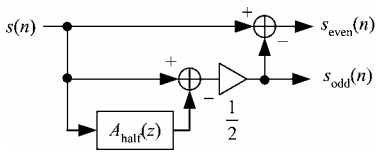

Figure 7: *Structure for separating the even and odd harmonics of a musical signal using one allpass filter.*

4. CASE STUDIES

This section presents how the proposed filtering algorithms perform with synthetic signals and recorded instrument tones. In addition, the proposed filter is compared against two techniques: a fractional-delay FIR filter using the well known Lagrange interpolation [11] and sinusoidal modeling [5], [6], [7]. The sampling frequency is 44.1 kHz in all the test cases.

4.1. Harmonic extraction from a synthetic test signal

The following example illustrates how the algorithm works with a synthetic test signal. The signal is determined to be the sum of sinusoids:

$$x(n) = A_{\text{sc}}(n)\sum_{k=1}^{K}\sin\left(\frac{2\pi nkf_0}{f_s} + \varphi_k\right), \tag{28}$$

where $A_{\text{sc}}(n)$ is an envelope function, K is the number of harmonics present in the signal, f_0 is the fundamental frequency of the signal, and φ_k is the phase of the kth harmonic. In this case, the parameters were chosen as follows: $K = 84$, $f_0 = 261.626$ Hz (C4), which corresponds to the cycle length of 168.562 samples. The initial phases φ_k are uniformly distributed random numbers in the range $[0, 2\pi]$. In order to examine the temporal smearing resulting from the harmonic component extraction, the envelope of the signal is chosen to be rectangular, containing sharp transitions.

As an example, two components, the fundamental frequency component and the 76th overtone, have been extracted from the signal (28). This is done with the HEF structure (19). The order of the truncated Thiran filter was chosen to be $N = 80$ and the order of the prototype filter was set to $M = 9N$. The attenuation coefficient A was determined to be equal to 10^{-6}. A comparison against a Lagrange FIR filter and sinusoidal modeling technique was carried out. The order of the Lagrange FIR filter N_{L} was set to 80 so that the amount of coefficients is the same as that of the proposed filter. In sinusoidal modeling, the following parameters were used. The short-time FFT was computed using a Blackman window of length $4f_s/f_0$, and the FFT size was 2048. The hop size was set to be one-fourth of the window length.

The results are presented in Figs. 8 and 9. Figs. 8(a) and (b) present the spectrum of the original signal (28), and Figs. 8(c) and (d) present the spectrum of the partials #1 and #76 that are extracted with the proposed method. Figs. 8(e) and (f) present the corresponding result obtained with the Lagrange FIR filter, and Figs. 8(g) and (h) represent the result obtained with sinusoidal modeling. All spectra were calculated from a 0.5 s excerpt taken between 0.1 s and 0.6 s of the signals. The Hamming window was used, and the spectra were computed using the discrete-time Fourier transform at 851 equally spaced points so that every 10th point matched one harmonic. This choice of parameters yields a clear visual representation of the sharp spectral peaks.

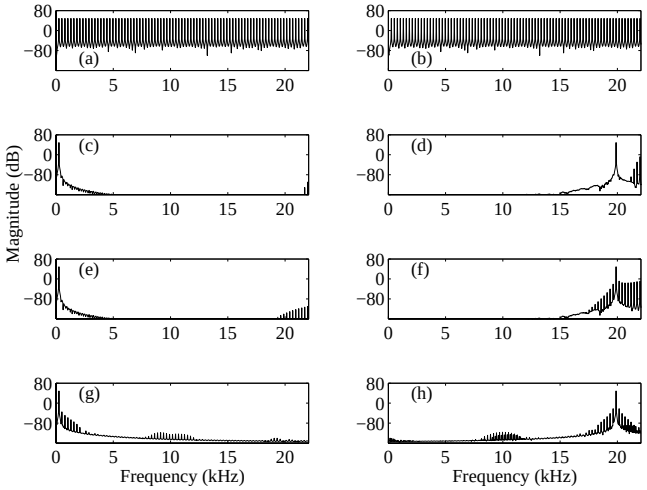

Figure 8: *Results of comparison in the frequency domain. (a), (b) Synthetic test signal and harmonic components #1 and #76 obtained with (c), (d) with the proposed method, (e), (f) with Lagrange FIR filter, and (g), (h), with sinusoidal modeling.*

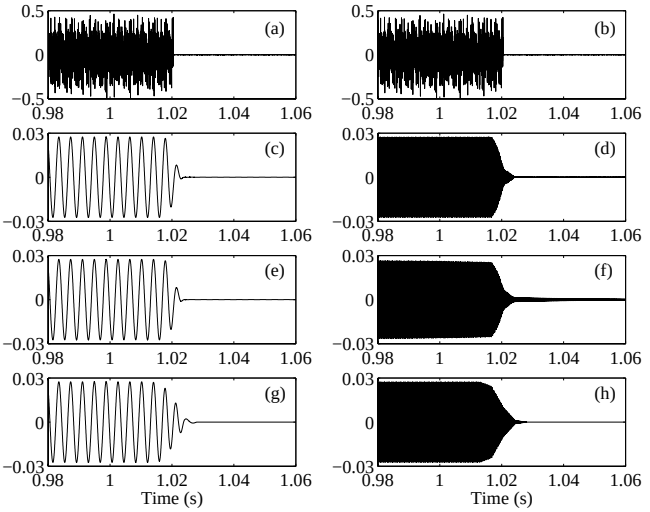

Figure 9: *The effects of temporal smearing. (a) and (b), an excerpt of the original synthetic test signal, (c) and (d) excerpts of the 1st and 76th harmonic obtained with the proposed allpass filter, (e) and (f) with the Lagrange FIR filter, and (g) and (h), with sinusoidal modeling.*

As can be seen in Figs. 8(c) and (e), the proposed filter and Lagrange filter are able to extract the first harmonic efficiently, and the other harmonics are properly attenuated. In the case of the 76th harmonic component, some of the uppermost harmonics are not properly attenuated because of the error in the phase delay near the Nyquist limit (see Fig. 4) and its effect on the attenuation of the inverse comb filter. The Lagrange FIR filter in Figs. 8(e) and (f), the error near the Nyquist limit is greater, and moreover, the lowpass nature of the magnitude response has to be taken into account in the analysis in order to maintain the level of the original signal at high frequencies. The sinusoidal modeling technique depicted in Figs. 8(g) and (h), suffers from sidelobes of the window function, and the attenuation of the neighboring harmonics is not as efficient as with the other methods.

The effects of temporal smearing are illustrated in Fig. 9. The signals are zoomed to the window 0.98 – 1.06 s. Figs. 9(a) and (b) show the original signal and Figs. 9(c) and (d) present the cases where the first harmonic and the 76th harmonic have been extracted with the proposed method, respectively. The performance of the Lagrange FIR filter is depicted in Figs. 9(e) and (f). It can be seen in Fig. 9(e) that in the case of the first harmonic the effect of temporal smearing is about same as with the proposed method. However, the smearing is greater in higher frequencies, which is visible in Fig. 9(f), since the lowpass nature of the filter complicates the usage of the resonator. In this case, the magnitudes of the resonator and the Lagrange filter do not compensate each other, which leads to improper attenuation. The temporal smearing in sinusoidal modeling technique depicted in Figs. 9(g) and (h) is slightly greater than that of the other methods.

4.2. Even-to-odd ratio calculation

In the case of the clarinet, the relation of even and odd harmonics and its effect to the timbre has been studied by Barthet *et al.* [27]. They have derived new descriptors for the clarinet timbre by constructing a simple but efficient parametric model for the clarinet to control certain parameters: the dimensionless mouth pressure γ and the embouchure parameter ζ, in addition to the fundamental frequency of the bore f_b and the reed resonance frequency f_r. The γ parameter defines the ratio between the pressure inside the player's mouth and the static beating reed pressure. The ζ parameter, in turn, takes the lip position and the section between the mouthpiece opening and the resonator into account. By varying these parameters Barthet *et al.* derived a relation between the parameters and the timbre characteristics.

We have examined the synthetic sounds generated by Barthet et al. [27] with the proposed algorithm in the case where the parameter γ varies between 0.40 and 0.50. As the relation of the odd and even harmonics has an effect on the spectum and timbre of the sound, the odd and even harmonics were separated with the structure presented in Fig. 7. The original signal and the separation results are shown in Fig. 10 for γ = 0.40. The other parameters, ζ, f_b,and f_r, are equal to 0.33, 170 Hz and 2500 Hz, respectively. The magnitude response is calculated with a 512-point FFT from an excerpt taken from 0.5 – 1.0 s. A Hamming window is used in the computation.

As Fig. 10 shows, the magnitude of the odd harmonics is greater than that of the even harmonics, which is typical for the clarinet sound. The odd-to-even ratio is determined by first calculating the envelopes of both signals. This is done by averaging the full-wave rectified signal over a window of 500 samples. The ratio between the magnitude of the odd and even harmonics is obtained by dividing the envelopes. The result seems to depend on the parameter γ. This relation is illustrated in Fig. 11 for four different values of γ (0.40, 0.42, 0.47, and 0.50).

Fig. 11 shows that the difference between the player's mouth pressure and the static beating reed pressure affects the odd-to-even ratio. That is, with larger pressure differences, the proportion of the odd harmonics in the tone is greater than in the case when the pressure difference is smaller (smaller γ value). Also the static state is reached faster with a larger pressure difference.

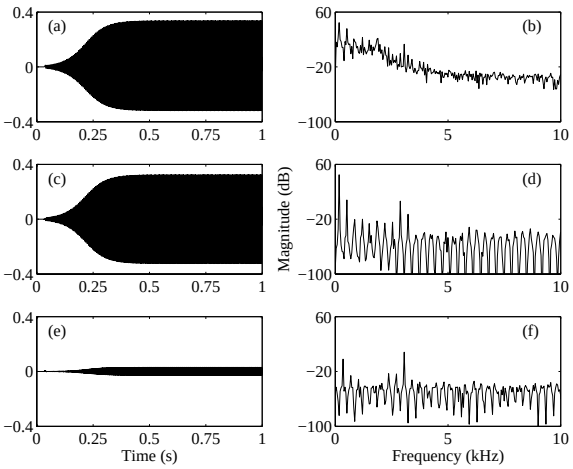

Figure 10: *(a) and (b) The synthetic clarinet tone in time and frequency domains, respectively, (c) and (d) the separated odd harmonics in time and frequency domains, respectively, and (e) and (f) even harmonics in the time and frequency domains ,respectively.*

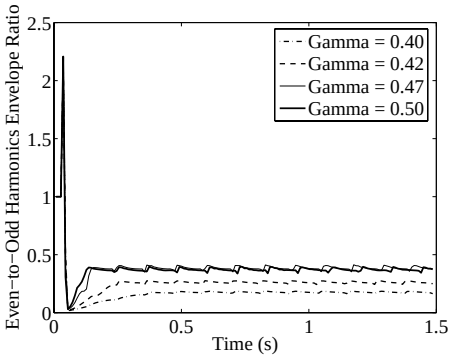

Figure 11: *The ratio of even and odd harmonics with four different γ values.*

4.3. Residual signal extraction

In order to investigate how the proposed algorithm works with recorded tones, a double bass tone was analyzed with the proposed method. In Figs. 12(a) and (b), the tone is presented in the time and frequency domains, respectively. The fundamental frequency of the tone was measured to be $f_0 = 58.2670$ Hz.

It is now desired to remove all the harmonics instead of preserving one. A modified form of the transfer functions (2) or (3) can be used, depending on the fundamental frequency,

$$H_{\mathrm{HEF}}(z) = g_1\left[1 - r^L A(z)\right], \tag{29}$$

where the coefficients g_1 and r^L are determined in the same manner as described in Sec. 3.1. $A(z)$ written is as in (1). The order of the truncated Thiran filter is $N = 80$ and the order of the prototype filter is $M = 9N$. The attenuation coefficient A is set to 10^{-6}. The filtered residual signal is presented in Figs. 12(c) and (d) in time- and frequency domains, respectively. When comparing Figs. 12(b) and (d), it is seen that the harmonic components are attenuated efficiently. Moreover, the noise between the harmonics is preserved.

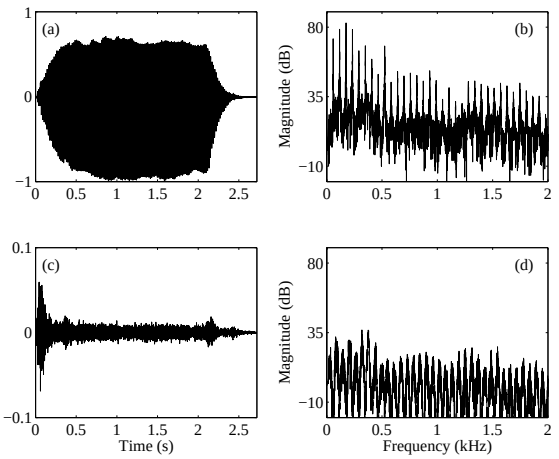

Figure 12: *Time- and frequency-domain presentations of (a), (b) the double bass tone and (c), (d) the extracted residual signal.*

5. CONCLUSIONS AND FUTURE WORK

Digital filtering techniques were proposed to obtain useful decompositions of harmonic musical signals. The basic approach taken here is to subtract a delayed copy of the signal from itself to cancel the harmonic components. A high-order digital allpass filter implements an accurate approximation of the required time delay. A harmonic extraction filter is obtained by cascading a second-order all-pole filter with the inverse comb filter. Division of a musical signal into two signals, one containing the even and the other the odd harmonics, and the extraction of the background noise or residual were also suggested as promising operations that are easy to realize using the proposed filter structures.

Case studies were presented to show how the techniques perform in the feature analysis for musical tones. Single harmonic components were extracted from a synthetic test tone. The neighboring harmonics were attenuated more than 100 dB. The harmonic even-to-odd ratio was determined for synthetic clarinet tones. Finally, the residual noise component was extracted from a bowed string tone by suppressing all the harmonics.

Future research includes developing a useful method to account for varying fundamental frequency of the signal, such as vibrato. In practice, this problem calls for a time-varying delay line to be used in the inverse comb filter. There are known methods for modulation of the delay-line length for example in effects processing, such as flanging and chorus algorithms. Which interpolation technique should be used and for how fast and wide delay-length modulation can this method be accurate? A further application of the time-varying inverse comb filter would be the separation of harmonic audio signals, such as musical tones or voiced speech, as discussed by de Cheveigné [28].

Another interesting special case is the analysis of inharmonic tones, such as piano tones or other instrument sounds with regular inharmonicity caused by dispersion. A filter-based analysis tool for such tones requires an allpass filter that approximates the dispersion in cascade with the delay line. This is a known method in digital waveguide synthesis of string sounds, see, e.g., [29], [30]. A more accurate approximation of dispersion characteristics is needed for the analysis tool than for sound synthesis.

6. ACKNOWLEDGMENTS

The work of Heidi-Maria Lehtonen has been supported by Tekniikan edistämissäätiö, the Emil Aaltonen Foundation, the Finnish Cultural Foundation, and the GETA graduate school. The authors are grateful to Ms. Minna Ilmoniemi and Dr. Minna Huotilainen for their collaboration in the early stages of this work. Special thanks go to M. Barthet, P. Guillemain, R. Kronland-Martinet, and S. Ystad for allowing us to use their synthetic clarinet tones as test signals in this work.

7. REFERENCES

[1] J. A. Moorer, "The optimum comb method of pitch period analysis of continuous digitized speech," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 22, no. 5, pp. 330-338, Oct. 1974.

[2] J. A. Moorer, "Signal processing aspects of computer music: a survey," *Proc. IEEE*, vol. 5, no, 8, pp. 1108-1137, 1977.

[3] V. Välimäki, M. Ilmoniemi, and M. Huotilainen, "Decomposition and modification of musical instrument sounds using a fractional delay allpass filter," in *Proc. Nordic Signal Processing Symp.*, pp. 208–211, Espoo, Finland, June 2004. Available online at http://wooster.hut.fi/publications/nosig2004/60_VALIM.PDF.

[4] N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*. Springer, 1991.

[5] R. J. McAulay and T. F. Quatieri, "Speech analysis/synthesis based on a sinusoidal representation," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 34, no. 4, pp. 744–754, 1986.

[6] X. Serra, *A System for Sound Analysis/Transformation/Synthesis Based on a Deterministic plus Stochastic Decomposition*. Ph.D. thesis. Report no. STAN-M-58, CCRMA, Stanford University, Stanford, CA, 1989. Available online at http://ccrma.stanford.edu/STANM/stanms/stanm58/.

[7] X. Serra and J. O. Smith, "Spectral modeling synthesis: A sound analysis/synthesis based on a deterministic plus stochastic decomposition," *Computer Music J.*, vol. 14, no. 4, pp. 12-24, 1990.

[8] G. Evangelista, "Pitch synchronous wavelet representations of speech and music signals," *IEEE Trans. Signal Processing*, vol. 41, no. 12, pp. 3313–3330, Dec. 1993.

[9] B. David, R. Badeau, and G. Richard, "HRHATRAC algorithm for spectral line tracking of musical signals," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, Toulouse, France, May 15-19, 2006, vol. 3, pp. 45–48.

[10] R. Badeau, B. David, and G. Richard, "High resolution spectral analysis of mixtures of complex exponentials modulated by polynomials," *IEEE Trans. Signal Processing*, vol. 54, no. 4, pp. 1341–1350, Apr. 2006.

[11] K. Steiglitz, *A Digital Signal Processing Primer with Applications to Digital Audio*. New York: Addison Wesley, 1996.

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

[13] V. Välimäki and T. I. Laakso, "Fractional delay filters—Design and applications," in *Nonuniform Sampling: Theory and Practice*. F. Marvasti, Ed. Kluwer Academic/Plenum Publishers: New York, 2001, Chapter 20, pp. 835–895.

[14] S. C. Pei and C.–C. Tseng, "A comb filter design using fractional-sample delay," *IEEE Trans. Circ. Syst.—Part II*, vol. 45, no. 6, pp. 649–653, June 1998.

[15] S. C. Dutta Roy, S. B. Jain, and B. Kumar, "Design of digital FIR notch filters," *IEE Proc. Vis. Image Signal Process.*, vol. 141, no. 5, pp. 334–338, Oct. 1994.

[16] C.-C. Tseng and S.-C. Pei, "Sparse FIR notch filter design and its application," *Electronics Letters*, vol. 33, no. 13, pp. 1131–1133, June 1997.

[17] S.-C. Pei and C.-C. Tseng, "IIR multiple notch filter design based on allpass filter," *IEEE Trans. Circuits and Systems Part II: Analog and Digital Signal Processing*, vol. 44, no. 2, pp. 133–136, Feb. 1997.

[18] C.-C. Tseng and S.-C. Pei, "Stable IIR notch filter design with optimal pole placement," *IEEE Trans. Signal Processing*, vol. 49, no. 11, pp. 2673–2681, Nov. 2001.

[19] A. Fettweis, "A simple design of maximally flat delay digital filters," *IEEE Trans. Audio and Electroacoust.*, vol. 20, no. 2, pp. 112–114, 1972.

[20] J.–P. Thiran, "Recursive digital filters with maximally flat group delay," *IEEE Trans. Circ. Theory*, vol. 18, no. 6, pp. 659–664, 1971.

[21] V. Välimäki, "Simple design of fractional delay allpass filters," in *Proc. European Signal Processing Conf.*, vol. 4, pp. 1881–1884, Tampere, Finland, Sept. 2000.

[22] S.-C. Pei and P.-H. Wang, "Closed-form design of all-pass fractional delay filters," *IEEE Signal Processing Letters*, vol. 11, no. 10, pp. 788–791, Oct. 2004.

[23] S. J. Orfanidis, *Introduction to Signal Processing*. Prentice Hall, 1996.

[24] T. Tolonen, "Methods for separation of harmonic sound sources using sinusoidal modeling," in *Proc. AES 106th Convention*, Preprint 4958, Munich, Germany, May 1999. Available online at http://lib.tkk.fi/Diss/2000/isbn9512251965/.

[25] P. A. A. Esquef, M. Karjalainen, and V. Välimäki, "Frequency-zooming ARMA modeling for analysis of noisy string instrument tones," *EURASIP J. Applied Signal Processing*, vol. 2003, no. 10, pp. 953–967, 2003.

[26] M. F. McKinney and J. Breebaart, "Features for audio and music classification", in *Proc. Int. Conf. Music Information Retrieval* (ISMIR), 2003. Available online at http://ismir2003.ismir.net/papers/McKinney.PDF.

[27] M. Barthet, P. Guillemain, R. Kronland-Martinet, and S. Ystad, "On the relative influence of even and odd harmonics in clarinet timbre," in *Proc. Int. Comp. Music Conf.*, Barcelona, Spain, Sept. 2005.

[28] A. de Cheveigné, "Separation of concurrent harmonic sounds: Fundamental frequency estimation and a time-domain cancellation model of auditory processing," *J. Acoust. Soc. Am.*, vol. 93, no. 6, pp. 3271–3290, June 1993.

[29] D. Rocchesso and F. Scalcon, "Accurate dispersion simulation for piano strings," in *Proc. Nordic Acoustical Meeting*, Helsinki, Finland, pp. 407–414, June 1996.

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