Proceedings of the 6th Nordic Signal Processing Symposium - NORSIG 2004
June 9 - 11, 2004
Espoo, Finland

Decomposition and Modification of Musical Instrument Sounds Using a Fractional Delay Allpass Filter

Vesa Välimäki[1†], *Minna Ilmoniemi*[1, 2], *and Minna Huotilainen*[2, 3]

[1]Helsinki University of Technology
Lab. of Acoustics and Audio Signal Processing
P.O. Box 3000, FIN-02015 HUT
Espoo, Finland
[†]E-mail: vesa.valimaki@hut.fi
[†]URL: http://www.acoustics.hut.fi/~vpv/

[2]University of Helsinki
Dept. of Psychology, Cognitive Brain Research Unit
P.O. Box 9, FIN-00014 University of Helsinki
Helsinki, Finland
[3]Collegium of Advanced Studies
P.O. Box 4, FIN-00014 University of Helsinki
Helsinki, Finland

ABSTRACT

A method to decompose a harmonic musical signal into individual harmonic components and a noise component is proposed. The background noise can be extracted by subtracting from the original signal a delayed copy where the delay is the period length of the musical tone. A highly accurate approximation of the delay is implemented as a high-order allpass filter. Extracting individual harmonic components involves the cancellation of one of the transmission zeros of the system by a pole. The method is easy and fast to use, and it allows accurate analysis and several useful modifications of musical tones. As an application example, three types of timbre modifications for recorded cello tones are demonstrated: change of the ratio of even and odd harmonic frequencies, of the spectral brightness, and of the amount of bowing noise. The resulting processed but natural-sounding signals are part of a test sound set used in a brain research task.

1. INTRODUCTION

Analysis of the frequency and the amplitude envelope of harmonic components of a musical tone is a fundamental issue in musical signal processing. In this paper we propose a method to decompose a harmonic musical signal into its harmonic components plus a noisy part using digital filtering. The proposed 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 that involves phase locking.

Another method that can be used for this kind of decomposition is sinusoidal modeling [1, 2]. In this method the signal is analyzed using the windowed FFT and frequency and amplitude tracks are obtained by connecting data in neighboring analysis frames. For periodic or pseudo-periodic musical tones it is unnecessary to get down to an overly generic analysis method, because the frequencies of harmonic components are known after fundamental frequency estimation. Advantages of the proposed filter-based analysis method – compared with more generic FFT-based techniques – are simplicity, which follows mainly from the small number of parameters, and the possibility of designing filter coefficients in closed form. It is unnecessary to optimize time-frequency resolution by investigating difference window types or other issues related to short-term spectral analysis. Additionally, the resulting decomposition is obtained directly as a set of time-domain signals, and a synthesis stage is not required.

This paper is organized as follows. Section 2 introduces the filter structure and design techniques. In Section 3, examples of a test case, the modification of cello tones, are presented. Section 4 concludes the paper.

2. INVERSE COMB FILTER STRUCTURES

2.1 Inverse Comb Filter with Fractional Delay

The inverse comb filter[1] (ICF) is an FIR filter where the input signal is delayed by L samples and is 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 on 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 integer, the accuracy of positioning the notches at desired frequencies is 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 [4, 5, 6]. Alternatively, an FIR notch filter can be designed and used, see, e.g., [7].

[1] Following the convention of [3], the term 'inverse comb filter' is used for the feedforward system. The 'comb filter' has a delay line inside a feedback loop.

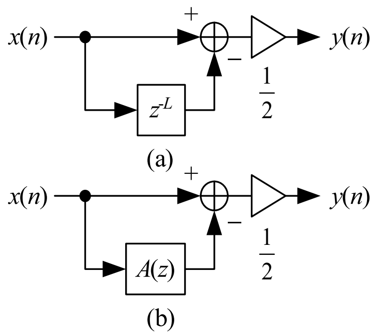

Fig. 1. (a) Conventional ICF and (b) a fractional delay ICF, where the delay is implemented with an allpass filter.

Figure 1(b) shows the block diagram of a fractional delay inverse comb filter, where the delay is replaced with an allpass filter. 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 11^{th}-order allpass filter that approximates the delay of 10.5351 sampling intervals is displayed in Fig. 2 (solid line).

The transfer function of a digital allpass filter is

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

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. In this application, the allpass filter order is typically N = round(L), which is also approximately the period length (in samples) to be cancelled. In general, order N can thus 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 design method is needed that allows design of high-order filters.

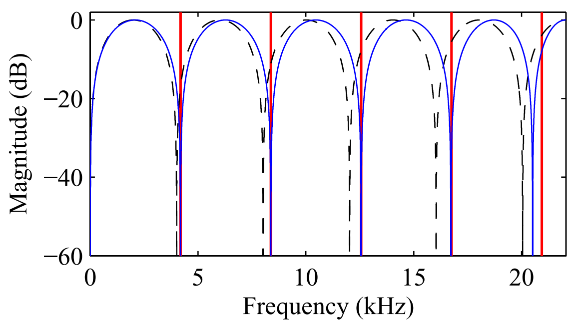

Fig. 2. Magnitude response of the conventional (dashed line) and the allpass-based (solid line) inverse comb filter. The thick vertical lines indicate the harmonic frequencies to be cancelled. In this case they are multiples of 4186 Hz.

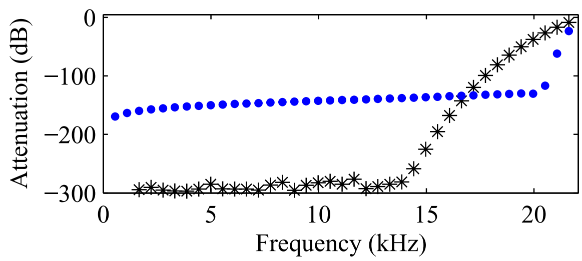

Fig. 3. Attenuation of harmonics using the Thiran (stars) and the truncated Thiran (circles) allpass filter. The fundamental frequency and period are 554.4 Hz and 79.55 samples, respectively.

2.2 Allpass Filter Design

Two design methods that allow allpass filter order to be increased up to N =1000 are the Thiran allpass filter [8, 9, 4] and the truncated Thiran allpass filter [10]. Both methods compute the filter coefficients based on a closed form formula. The Thiran design formula can be expressed as

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

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

The truncated Thiran design is a modification of Eq. (2):

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

where M is the prototype filter order ($M > N$) [10]. By using a value for M that is larger than N in computing the allpass filter coefficients, it is possible to extend the bandwidth of good approximation. This comes at the expense of loosing quality at low frequencies: the approximation error is larger than in the original allpass filter. Nevertheless, this technique allows a useful tradeoff between approximation accuracy and bandwidth, as discussed in [10]. A comparison of two ICFs based on 80^{th}-order allpass filters is given in Fig. 3. Prototype filter order for the truncated filter was $M = 800$. It is seen that at frequencies below 17 kHz the truncated allpass filter is worse but still sufficiently good, because harmonics are attenuated more than 130 dB. Above 17 kHz the performance of the Thiran filter collapses, but its truncated version offers an attenuation of 130 dB up to 20 kHz.

2.3 Extracting Single Harmonic Components

Instead of canceling all harmonic components, single harmonics can be extracted. This is achieved by cascading with the system of Fig. 1(b) a second-order filter that cancels a transmission zero at a given harmonic frequency. The new transfer function is

$$H_{\mathrm{single}}(z) = R(z)[1 - A(z)]/2, \quad (4)$$

where

$$R(z) = \frac{b_R}{1 - a_R z^{-1} + z^{-2}} \quad (5)$$

and b_R is a gain factor, $a_R = 2\cos(2\pi m f_0/f_s)$, and m is the index of the harmonic to be retained.

It was observed in practical tests that for some musical instrument tones with strong low-indexed harmonics, the filtering of the signal with $H_{\mathrm{single}}(z)$ is insufficient: listening of the filtered signal reveals that the fundamental is still perceived although one of the high-frequency partials is strongly emphasized. Filtering the signal twice with transfer function of Eq. (4) seems to adequately attenuate the rest of the harmonics.

It is worth noting that while the cancellation of one of the zeros with a pole is almost perfect, there is a hidden oscillating mode in the transfer function of Eq. (4). In practice we have not faced problems because of this. One reason is that we have processed short signals only.

2.4 Separation of Odd and Even Partials

While it is possible to cancel the harmonic components one by one by applying the above technique multiple times, it is also possible to suppress all even harmonics at once. 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 signal. The structure of Fig. 1(b) 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:

$$\hat{s}_{\mathrm{even}}(n) = s_{\mathrm{orig}}(n) - \hat{s}_{\mathrm{odd}}(n). \quad (6)$$

3. DECOMPOSITION AND MODIFICATION OF A RECORDED CELLO TONE

We present application examples where the proposed filtering techniques are used. The following case studies are part of a project in which acoustic test stimuli were generated for brain research. A recorded cello tone (C5, fundamental frequency 524.0 Hz, duration 200 ms) is used in all tests. Its magnitude spectrum without any processing is shown in Fig. 4. The sampling rate used is 44.1 kHz.

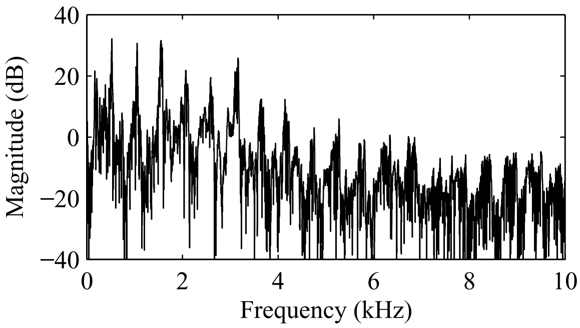

Fig. 4. Spectrum of the original cello tone.

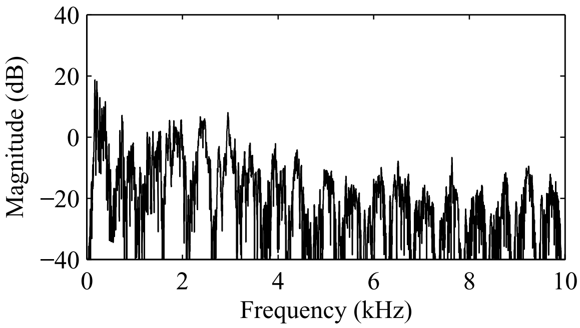

Fig. 5. Spectrum of the extracted noise of the cello tone.

3.1 Extracting the Noisy Part

The noisy part of the cello tone is extracted by using the structure of Fig. 1(b). The Thiran allpass filter with parameters N = 84 and d = 0.1611 was used. The magnitude spectrum of the extracted noisy part, which contains the attack and bowing sound, is shown in Fig. 5.

3.2 Extracting Single Harmonics

The filtering method explained in Sec. 2.3 was used to extract the 30 lowest harmonics of the cello tone. The filter was applied twice in each case. The rest of the harmonics, i.e., those above 15.7 kHz, were buried in background noise and could be neglected. As an example, Fig. 6 shows the spectrum of the signal where only the fifth harmonic is retained. Comparison against Fig. 4 demonstrates that the peak around 2620 Hz is unchanged while other parts of the spectrum are much attenuated.

3.3 Odd and Even Harmonics

The odd and even harmonics can be extracted using two alternative ways: either by filtering out each harmonic separately and then summing the odd or even ones, or by applying the filter that first cancels the even harmonics. The spectrum of the sum of even harmonics is shown in Fig. 7. This should be compared with Fig. 4 to see that all odd harmonics around frequencies (524 Hz, 1572 Hz, 2620 Hz etc.) are suppressed, as desired. This signal sounds like a cello tone played one octave higher (fundamental frequency 1048 Hz).

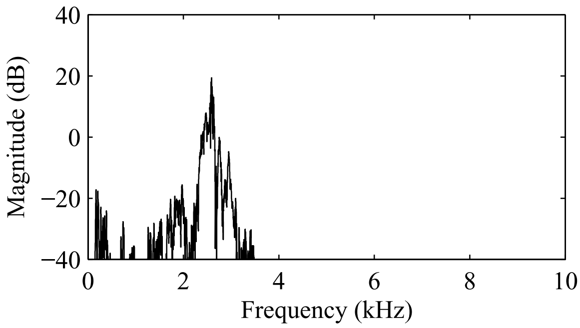

Fig. 6. Spectrum of the extracted fifth harmonic.

3.4 Modifications

Several useful modifications are possible after the decomposition of the cello tone into the noisy part and harmonics. It is known that some of the relevant features of a musical tone are its brightness, ratio of the magnitude of odd and even harmonics, and the degree of noisiness. For the test sound set, we reconstructed the original tone from the 30 extracted signals that contained the harmonics. To control the ratio of even and odd harmonics, it is possible to assign weights to even and odd harmonics to be used in the summation. The noisiness can be increased by adding the extracted noisy part with a non-zero weight. Interestingly, also decreasing the noise is possible by using a negative weight, because synchronization of the events in the signal is restored. To decrease brightness, an exponentially decaying sequence of numbers is used as weights for the harmonics. However, to increase brightness, it is unnatural to simply increase the weight of harmonics exponential. Instead, we use exponential increase up to the 15th harmonic and then saturate the weight to that value for the rest of the harmonics. This still sounds natural, while the overall brightness perception is radically modified. Fig. 8 shows the spectrum of a signal modified using this method.

4. CONCLUSIONS

Digital filtering techniques to obtain useful decompositions of harmonic musical signals were proposed. 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. Modifications of this theme include extraction of all harmonic components into separate signals, division of the signal into two signals one of which contains the even and the other the odd harmonics, and extraction of the background noise, which in the case of a cello tone consists of the attack transient and the bowing noise. An application example for these signal processing methods is a brain research task where processed musical instrument recordings are used to learn about the human cognition of musical sounds.

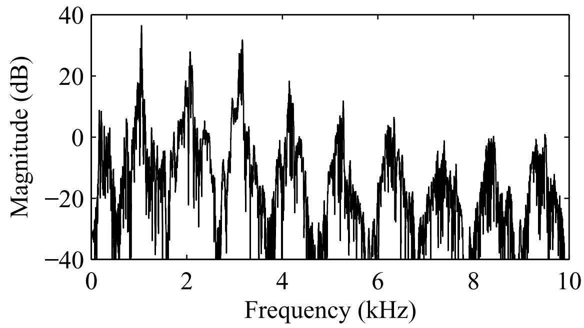

Fig. 7. Spectrum of the signal that is obtained as the sum of the extracted even harmonics.

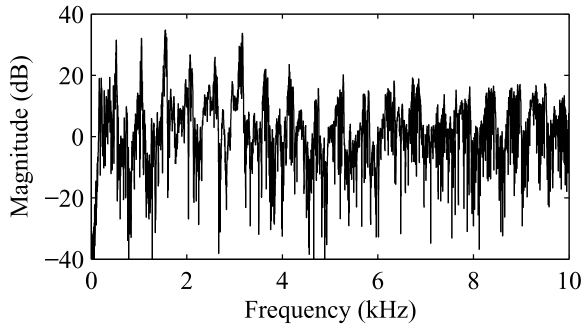

Fig. 8. Spectrum of the brightened tone.

REFERENCES

[1] 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.

[2] 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.

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

[4] T. I. Laakso, V. Välimäki, M. Karjalainen, and U. Laine, "Splitting the unit delay—Tools for fractional delay filter design," *IEEE Signal Processing Magazine*, vol. 13, no. 1, pp. 30–60, Jan. 1996. Matlab files for fractional delay filter design are available online at http://www.acoustics.hut.fi/software/fdtools/.

[5] 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.

[6] 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.

[7] 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.

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

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

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