Dispersive and Pitch-Synchronous Processing of Sounds

Gianpaolo Evangelista (1)
Sergio Cavaliere (2)
(1) École Polytechnique Fédérale de Lausanne, Switzerland
evangelista@na.infn.it
(2) Department of Physical Sciences, Univ. of Naples, Italy
cavaliere@na.infn.it http://www.na.infn.it/Gener/acust/

Abstract

The aim of this paper is to present results on digital processing of sounds by means of both dispersive delay lines and pitch-synchronous transforms in a unified framework. The background on frequency warping is detailed and applications of this technique are pointed out with reference to the existing literature. These include transient extraction, pitch shifting, harmonic detuning and auditory modeling.

1 Introduction

Frequency warping is a unitary transformation [4] that can be interesting from the point of view of sound manipulation and synthesis. Novel applications, some of them recently published by the authors, make use of dispersive delay lines, i.e., delay lines whose elements are simple all-pass filters, for implementing frequency warping operators. Digital delay lines have received a considerable interest in both synthesis (digital waveguide models) and processing (reverberation, flanging, phasing) of sound. A particular type of dispersive delay line is given by the Laguerre network, which leads to orthogonal transforms offering a concerned method of warping that preserves the signal energy in any arbitrary frequency band. The shape of the corresponding frequency warping function is obtained from a one-parameter family of mappings. The Laguerre transform may be considered as the building block of several new representations, useful for the synthesis, analysis and processing of sound. A large class of sounds, such as those produced by stiff strings, plates, etc., shows an inharmonic distribution of the partials. As an example, the waveguide physical model of piano tones in the low-register is accurately described in terms of orthogonal Laguerre sets. It can be shown that the stiff-string model is equivalent to a flexible string model with frequency warped output.

The authors recently defined the Pitch-Synchronous Warped Wavelet Transform and applied it to the separation of transients and noise from the resonant part of quasi-harmonic sounds. Frequency warping by means of Laguerre expansion allows for adaptation of the pitch-synchronous comb basis set to unequally spaced partials. Adaptation requires the optimization of the Laguerre parameter, for which two methods are discussed in the paper. Once the partials are framed by the non-uniform comb scaling set, the wavelet components represent the fluctuations at several scales and are appreciably non-zero on transients. With these methods we were able to resolve the piano hammer noise from the resonant component.

Another application of frequency warping embedded in orthogonal transforms is in critically sampled filter banks with arbitrary band allocation. By combining a two band quadrature mirror filter bank with Laguerre expansions one can build a two band filter bank with arbitrary cutoff frequency. By iterating this filter bank one obtains wavelets whose bandwidths are no longer constrained to be one octave. The cutoff frequencies may be selected by choosing a set of Laguerre parameters, e.g., by adapting the frequency bands to a perceptual model, such as Bark scale. A procedure for obtaining these parameters from a closed-form relationship is provided. Some authors considered warping a uniform filter bank for obtaining an approximate Bark scale filter bank. In contrast, our method consists of iteratively warping the signal and allows for the design of a perfect reconstruction structure that exactly meets any specification of cutoff frequencies. Since the structure is based on iterating a two-band filter bank, there is no need for designing a uniform filter bank with a large number of bands. However, the computational cost of the new transform is higher since several chains of real first-order all-pass filters are included in our scheme.

Frequency warping by means of Laguerre expansion has interesting applications to sound transformations. Among the others, we experimented with microdetuning. By warping a signal by means of projection on a Laguerre set with a small value for the parameter one obtains a natural sounding approach to pitch shifting. In piano tones tuning is due to both the thickness of the string and tension. Frequency warping allows for changing the stiffness parameter.

Larger amounts of warping introduce interesting effects on harmonic signals where both phase relationships and spacing of the partials are non-uniformly changed, creating, for example, gong-like sounds from simple periodic signals, such as pulse trains. Moreover, by using the pitch-synchronous warped wavelet transform for noise and transient extraction, interesting examples of cross-synthesis of different instruments may be generated.

2 Frequency Warping

Suppose that we want to modify the frequency spectrum of a signal $s(k)$ so that its frequency content is displaced to other frequencies [7]:

$$\tilde{S}\big(\theta(\omega)\big) = S(\omega) \quad (1)$$

One needs to provide a mapping $\Omega = \theta(\omega)$ transforming each frequency into a new one. The warped frequency spectrum of a discrete-time signal $s(k)$ is

$$\tilde{S}(\omega) = S(\theta^{-1}(\omega)) = \sum_k s(k) e^{-jk\theta^{-1}(\omega)}$$

The warped signal is

$$\tilde{s}(n) = \frac{1}{2\pi}\int_{-\pi}^{+\pi} \tilde{S}(\omega) e^{jn\omega} d\omega = \sum_k s(k) h_n(k) \quad , \quad (2)$$

where

$$h_n(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{j\left[n\omega - k\theta^{-1}(\omega)\right]} d\omega \quad .$$

Warping the signal is equivalent to orthogonally project it onto the set $h_n(k)$.

Frequency warping may be performed with the help of arbitrary maps. However, there are criteria for selecting the suitable map for the specific application, some of them are listed below.

1. In order to guarantee reversibility the map should be one-to-one and onto the interval $[-\pi, \pi)$, characterized by an increasing or decreasing function. If one needs to preserve frequency ordering one needs the map to be increasing. The effect of altering the frequency ordering is similar to aliasing. In creating musical effects, however, one may be also interested in non-invertible maps.

2. From a psychoacoustic point of view it may be desired to equalize the warped spectrum in order to preserve power in any frequency band, since, as a result of warping, by stretching a given band one increases the associated power. Equalization is fully justified when one deals with unvoiced sounds. However, in pseudo-periodic sounds one may desire to preserve, as much as possible, the envelopes of the partials. In this case a criterion could be that of moving the partial frequency by means of frequency warping while preserving the bandwidths of the partials. The overall warping map must be piecewise linear with 45° local slopes, i.e., the map is a sum of a (non-decreasing) staircase function and the identity (45° line) map.

3. If one wants to transform a real signal into another real signal then the map must be an odd function of ω.

In order to define signal representations or transforms based on frequency warping one needs to constrain the map to be one-to-one and onto. If the map is odd and differentiable then

$$h_n(k) = IDFT\left[e^{-jn\theta(\omega)} \frac{d\theta}{d\omega}\right], \quad (3)$$

forms a complete set of sequences, biorthogonal to the set

$$g_n(k) = IDFT\left[e^{-jn\theta(\omega)}\right],$$

so that

$$s(k) = \sum_n \tilde{s}(n) g_n(k) \; .$$

By factoring the (positive) derivative

$$\frac{d\theta}{d\omega} = \left|F_0(\omega)\right|^2$$

one can obtain the orthogonal and complete set

$$f_n(k) = IDFT\left[e^{-jn\theta(\omega)} F_0(\omega)\right].$$

Projection on this set obtains warping (see (1)) combined with spectrum scaling:

$$F_0(\omega)\tilde{S}(\omega) = S(\omega).$$

This form of scaling implies energy preservation. The warped spectrum has the same energy in the transformed frequency band $[\theta(B_0), \theta(B_1)]$ as the original spectrum in any arbitrary band $[B_0, B_1]$:

$$\int_{B_0}^{B_1} \left|S(\omega)\right|^2 d\omega = \int_{\theta(B_0)}^{\theta(B_1)} \left|\tilde{S}(\omega)\right|^2 d\omega \quad .$$

A simplification is possible if the input signal is causal: in this case only the causal part of $f_n(k)$ gives non-zero contributions to the expansion.

A remarkable case is based on the map generated by the phase of the first order all-pass filter [1]:

$$A(z) = \frac{z^{-1} - b}{1 - bz^{-1}} \quad \text{with } -1 < b < 1 \ .$$

$A(z)$ maps the unit disk in itself. On the unit circle

$$A(e^{j\omega}) = e^{-j\theta(\omega)} \ ,$$

where

$$\theta(\omega) = - \ \arg \ A(e^{j\omega}) = \omega + 2\tan^{-1}\left(\frac{b\sin\omega}{1 - b\cos\omega}\right) \quad (4)$$

is the associated frequency warping map, shown in *Figure 1* for several values of the parameter, whose inverse $\theta^{-1}(\omega)$ is obtained by reversing the sign of the parameter *b*. One can show that (4) is the unique one-to-one warping map generated by rational functions. This is important in digital realizations, where one has to implement a chain of allpass filters in order to compute the frequency warped version of the signal.

By introducing the orthogonalizing factor $\Lambda_0(z) = \frac{\sqrt{1-b^2}}{1-bz^{-1}}$, one can show that the z-transforms of the basis set are

$$H_r(z) = \Lambda_0(z) A(z)^{r+1} \ , \qquad r=0,1,\ \dots \ . \qquad (5)$$

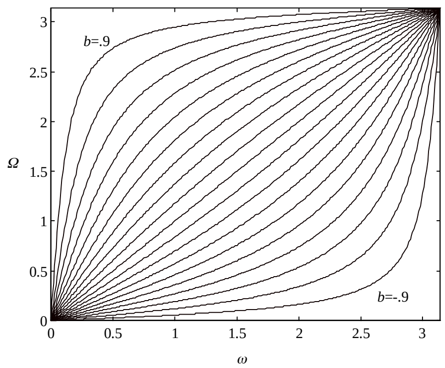

Figure 1. Family of warping maps generated by allpass

The operations involved in (2), with the recurrence (5) taken into account, are equivalent to time-reversing the signal, filtering by the orthogonalizing factor and evaluating at time lag 0 the iterated convolution by the first order all-pass impulse response, as shown in *Figure 2*. The inverse Laguerre transform structure is obtained simply by reversing the sign of the parameter *b*.

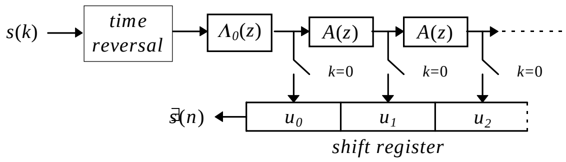

Figure 2. Digital structure implementing the Laguerre transform.

Although the Laguerre transform is the unique orthogonal frequency warping scheme that can be exactly implemented with rational transfer functions, there are other orthogonal warping schemes that can be approximated by IIR or FIR filter chains. For example the all-pass impulse response corresponding to the map $\theta(\omega) = \pi\sin(\omega/2)$, $|\omega| < \pi$, is well approximated by means of a high-order FIR filter whose coefficients are expressed in terms of Bessel functions.

3 Applications

In order to discuss applications of the results illustrated in the previous section to sound synthesis and processing, we consider an example where the role played by frequency warping is a fundamental one. This is the case of a stiff string, such as a low-register piano string, whose deformation is governed by the fourth-order partial differential equation:

$$\varepsilon \frac{\partial^4 y(x,t)}{\partial x^4} - \frac{\partial^2 y(x,t)}{\partial x^2} + \frac{1}{c^2}\frac{\partial^2 y(x,t)}{\partial t^2} = 0 \ .$$

The general solution of this equation is the superposition of four terms of which only two are oscillating and contribute to the sound. In the frequency domain (with respect to time) one can write:

$$Y(x,\omega) \approx C_1(\omega) e^{j|\alpha(\omega)|x} + C_2(\omega) e^{-j|\alpha(\omega)|x} \ .$$

Unlike the flexible string case, the solution is not a sum of purely progressive and regressive waves. Rather, if we subdivide the string into several spatial sections, at any given time the solution at one extreme of the section is obtained by all-pass filtering the solution at the other extreme of the section. This phenomenon is typical of dispersive wave propagation in which group velocity is frequency dependent. As a result, the partials of both the hinged and the clamped string are not uniformly spaced in the frequency domain. The simple waveguide model (Karplus-Strong algorithm) needs to be modified in order to be able to synthesize stiff-string sounds: it must now include frequency dependent delays [10]. This can be implemented in the dispersive delay line shown in *Figure 3*. It can be shown that dispersive waveguide model is equivalent to a non-dispersive waveguide cascaded by a frequency warping structure. Thus, by means of frequency warping one can bring the stiff-string structure to a simple waveguide form, including elementary delays.

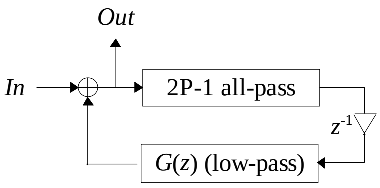

Figure 3. Waveguide model for the stiff string

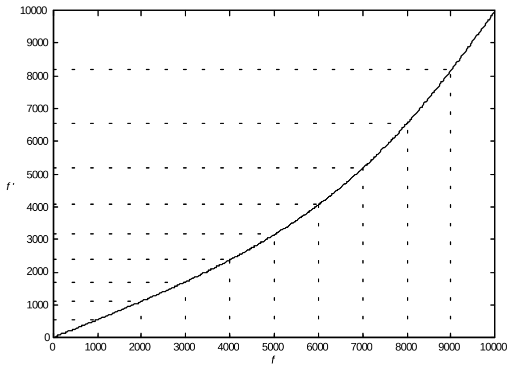

Figure 4. Transformation of harmonics to non-uniformly spaced partials.

A further result is that the dispersive piano characteristic is well approximated by the Laguerre warping map, for a wide range of the physical parameters [9]. Accurate synthesis of piano can be obtained by frequency warping a suitable pseudo-harmonic signal. This concept can be applied to arbitrary quasi-periodic signals, such as pulse trains, which may be transformed to more interesting waveforms if we detune their harmonics by means of frequency warping, as shown in *Figure 4*. The pitch of the transformed signal is not the same as the original one. In fact the fundamental frequency is transformed according to the warping map. For small values of the Laguerre parameter harmonic detuning is a marginal effect. Detuning of the fundamental frequency is perceptually more relevant. Thus, frequency warping may be successfully employed as a pitch-shift algorithm for micro-detuning, say within one half-tone. For dispersive instruments, such as piano, pitch-shifting by frequency warping is a more concerned method since the detuning of the harmonics depends on pitch.

Frequency warping is important for the analysis, since one can regularize a piano sound by making its partials equally spaced in the frequency domain. Once the signal is regularized, the methods of pitch-synchronous analysis by means of tuned comb filters can be applied in order to separate the regular (periodic) component from transients and noise [2, 3]. The results can be formalized in terms of the Frequency Warped Pitch-Synchronous Wavelet Transform [5, 6], whose basis sequences have the frequency spectra show in *Figure 5*. By means of this representation we were able to separate the noise of the hammer from the string vibration in a piano sound.

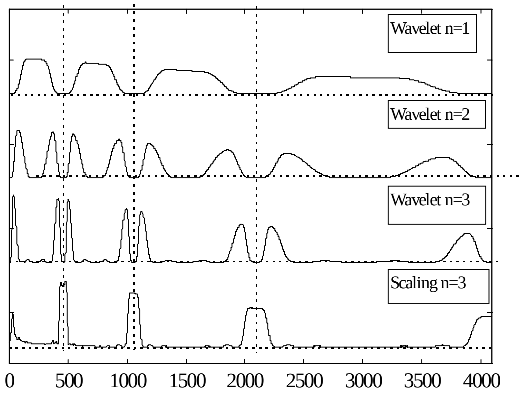

Figure 5. Fourier Transform of typical FW-PS wavelets.

Another application of frequency warping is in the design of filter banks whose frequency bands match certain perceptual criteria, e.g., the Bark scale [8]. This can be achieved by iteratively warping the signal prior to feed it to a two-band quadrature-mirror filter bank [6]. The output of the low-pass section of the filterbank is then warped and fed to the next two-band filter bank, and so on. At each stage the cutoff of the corresponding frequency band can be determined from the warping map. If Laguerre warping is employed the parameters b_k are determined by the following recurrence [6]:

$$b_1 = \tan\left(\frac{\pi - 2\omega_1}{4}\right),$$

$$b_n = \tan\left(\frac{\pi}{4} - \Omega_{n-1}(\omega_n)\right), \qquad n=2,3,... \tag{6}$$

where

$$\Omega_n(\omega) = \theta_n(2\theta_{n-1}(...2\theta_2(2\theta_1(\omega))...)) \ . \tag{7}$$

and $\omega_1 > \omega_2 > .... > \omega_N$ are the desired cutoff frequencies. This result is formalized in terms of arbitrary bandwidth wavelet sets.

References

[1] P. W. Broome, “Discrete Orthonormal Sequences,” *J. Assoc. Comput. Machinery,* vol. 12, no. 2, pp. 151-168, 1965.

[2] G. Evangelista, “Comb and Multiplexed Wavelet Transforms and Their Applications to Signal Processing,” *IEEE Trans. on Signal Processing*, vol. 42, no. 2, pp. 292-303, Feb. 1994.

[3] G. Evangelista, “Pitch Synchronous Wavelet Representations of Speech and Music Signals,” *IEEE Trans. on Signal Processing*, special issue on Wavelets and Signal Processing, vol. 41, no.12, pp. 3313-3330, Dec. 1993.

[4] R.G. Baraniuk and D.L. Jones, “Unitary Equivalence : A New Twist on Signal Processing,” *IEEE Trans. Signal Processing*, vol. 43, no. 10, pp. 2269-2282, Oct. 1995.

[5] G. Evangelista and S. Cavaliere, “Frequency Warped Filter Banks and Wavelet Transforms: A Discrete-Time Approach Via Laguerre Expansion,” *IEEE Trans. on Signal Processing*, October 1998.

[6] G. Evangelista and S. Cavaliere, “Discrete Frequency Warped Wavelets : Theory and Applications,” *IEEE Trans. on Signal Processing*, special issue on Theory and Applications of Filter Banks and Wavelets, April 1998.

[7] A.V. Oppenheim and D. H. Johnson, “Discrete Representation of Signals,” *Proc. IEEE*, vol. 60, pp. 681-691, June 1972.

[8] G. Evangelista, S. Cavaliere, "Auditory Modeling via Frequency Warped Wavelet Transform," Proc. EUSIPCO '98, pp. 117-120.

[9] I. Testa, S. Cavaliere and G. Evangelista, "A Physical Model of Stiff Strings," *Proc. of Int. Symposium on Musical Acoustics ISMA'97*, Edinburgh, August 1997.

[10] S. A. Van Duyne and J. O. Smith, "A Simplified Approach to Modeling Dispersion Caused by Stiffness in Strings and Plates," *ICMC Proc.*, 1994.