

Audio Effects Based on Biorthogonal Time-Varying Frequency Warping

Gianpaolo Evangelista

École Polytechnique Fédérale de Lausanne, Switzerland
Email: gianpaolo.evangelista@epfl.ch

Sergio Cavaliere

Department of Physical Sciences, University of Naples, Italy
Email: cavaliere@na.infn.it

Received 31 March 2000 and in revised form 23 January 2001

We illustrate the mathematical background and musical use of a class of audio effects based on frequency warping. These effects alter the frequency content of a signal via spectral mapping. They can be implemented in dispersive tapped delay lines based on a chain of all-pass filters. In a homogeneous line with first-order all-pass sections, the signal formed by the output samples at a given time is related to the input via the Laguerre transform. However, most musical signals require a time-varying frequency modification in order to be properly processed. Vibrato in musical instruments or voice intonation in the case of vocal sounds may be modeled as small and slow pitch variations. Simulation of these effects requires techniques for time-varying pitch and/or brightness modification that are very useful for sound processing. The basis for time-varying frequency warping is a time-varying version of the Laguerre transformation. The corresponding implementation structure is obtained as a dispersive tapped delay line, where each of the frequency dependent delay element has its own phase response. Thus, time-varying warping results in a space-varying, inhomogeneous, propagation structure. We show that time-varying frequency warping is associated to an expansion over biorthogonal sets generalizing the discrete Laguerre basis. Slow time-varying characteristics lead to slowly varying parameter sequences. The corresponding sound transformation does not suffer from discontinuities typical of delay lines based on unit delays.

Keywords and phrases: signal transformations, frequency warping, Laguerre transform, Kautz functions.

1. INTRODUCTION

Frequency warping is an interesting technique in sound processing. Given a map of the frequency axis into itself, one can transform the spectral content of any sound by mapping the original set of frequencies into other frequencies. A simple example is given by transforming a periodic sound via a linear map. In that case the original set of harmonics is scaled by the angular coefficient. One obtains another set of harmonic frequencies multiple of the transformed fundamental frequency. When warping discrete-time signals care must be taken since the map may alter the periodicity of the Fourier transform, i.e., the bandwidth of the signal. Classical upsampling and downsampling operators may be written in terms of warping operators. By frequency warping a periodic sound via a nonlinear map, one obtains an inharmonic set of partials as shown in [Figure 1](#). Voiced sounds of natural instruments do not always possess a harmonic structure. For example, piano tones in the low register show inharmonicity. This is due to the stiffness of the string, which results in dispersive wave propagation. Waves travel with a velocity that depends on

their frequency. The distance of the frequencies of the partials increases with frequency. This effect can be simulated by frequency warping a periodic (harmonic) signal according to a suitable law that can be derived from the physical model or from experimental data. More generally, one can obtain sound morphing by continuously transforming the frequency content via warping.

The idea of frequency warping is not new. Broome [1] described a general class of discrete-time orthogonal transforms based on Kautz sequences. Oppenheim and Johnson [2] considered a biorthogonal set related to the Laguerre transform. In [3] Baraniuk and Jones considered frequency warping in the general framework of unitary transformations. In recent papers [4, 5, 6, 7, 8, 9, 10], the authors considered frequency warping by means of the orthogonal Laguerre transform as a building block of algorithms for sound manipulation.

Frequency warping adds flexibility in the design of orthogonal bases for signal representation. Furthermore, the computational scheme associated with the Laguerre transform has all the prerequisites for digital realizations. The authors showed that the definition of orthogonal warping set

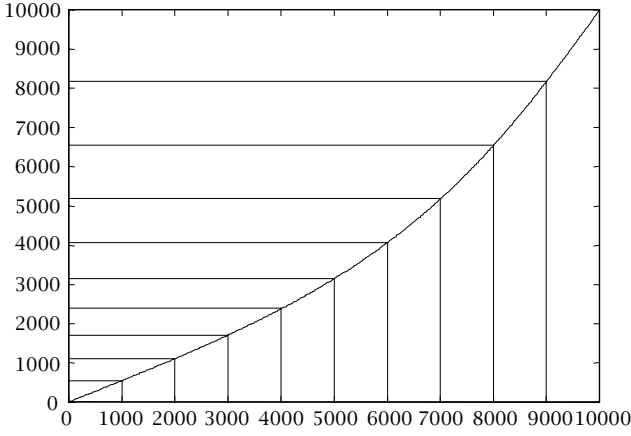


FIGURE 1: Transformation of harmonics to nonuniformly spaced partials.

based on a rational filter structure is useful for the construction of wavelet bases with arbitrary frequency band allocation [5]. The new combined transform leads to fine applications such as orthogonal and complete perceptual filter banks [6, 8].

The Laguerre transform may also be used for adapting quasi-periodic sounds to pitch-synchronous schemes [7]. In particular, by combining this transform with the pitch-synchronous wavelet transform [11, 12], one can achieve transient and noise separation from resonant components by means of a unitary transformation where resonant and noise components are projected onto orthogonal subspaces [9, 10]. In order to achieve this separation in signals whose partials are not equally spaced in the frequency domain, one needs to determine a warping map bringing partials onto harmonics. By combining inharmonic and harmonic components of different instruments one can obtain interesting cross-synthesis examples.

The authors showed that inharmonic sounds, such as those produced by stiff strings, plates, etc., may be conveniently modeled by means of waveguides based on simple delay lines followed by frequency warping elements [7]. The warping characteristic of the Laguerre family is particularly accurate in modeling inharmonicity of piano tones, as comparison with the characteristics derived from the physical model and direct analysis of the tones shows [5, 13]. When pitch shifting sample piano tones, this concept allows us to take into account stiffness increase as we move to the lower tones.

Frequency warping generates interesting sound effects such as sound morphing, phasing, chorusing, flanging, pitch-shifting, and new effects not yet in the catalogue. However, in order to be able to capture the full range of possibilities, one needs to consider dynamic variations of the warping parameters.

In this paper, we approach the problem of time-varying frequency warping by means of generalized Laguerre transform. In this context, we show that time-varying frequency warping may be implemented in a space-varying sampled delay-line. The generalized transform reverts to the Laguerre transform if all the parameters are kept constant. Moreover,

the transform may be embedded into an invertible operation via an associated biorthogonal and complete set. This is useful for building effects that can easily be undone without degradation of the original sound. The use of the dynamic transform is demonstrated by means of examples.

The paper is organized as follows. In Section 2, we recall basic properties of time-invariant frequency warping and Laguerre transform. In Section 3, we consider space-varying sampled delay lines and relate them to a biorthogonal set generalizing the Laguerre transform. In Section 4, we show that time-varying frequency warping is obtained by projection over the biorthogonal set introduced in Section 3. In Section 5, we discuss some applications of the warping audio effect to sound editing and music and provide examples. Finally, in Section 6 we draw our conclusions.

2. FREQUENCY WARPING AND THE LAGUERRE TRANSFORM

In this section we review the basic concepts of frequency warping. Frequency warping is a transformation of the frequency axis via a frequency map $\theta(\omega)$. As a result of warping, a signal $s[k]$ is transformed into another signal $\hat{s}[k]$. The DTFT of the two signals are related as follows:

$$\hat{S}(\theta(\omega)) = S(\omega). \quad (1)$$

In other words, the frequency content of $\hat{s}[k]$ at frequency $\Omega = \theta(\omega)$ is the same as the frequency content of $s[k]$ at frequency ω [2]. Frequency warping may be performed with arbitrary maps. However, if the warping map is one-to-one and onto $[-\pi, \pi]$, for example, if $\theta(\omega)$ is monotonically increasing fixing the points 0 and π , warping is reversible and gives more predictable results. In that case, the warping frequency spectrum of a discrete-time signal $s[k]$ is

$$\hat{S}(\omega) = S(\theta^{-1}(\omega)) = \sum_k s[k] e^{-jk\theta^{-1}(\omega)}, \quad (2)$$

and the warped signal is

$$\hat{s}[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{S}(\omega) e^{jn\omega} d\omega = \sum_k s[k] h_n[k], \quad (3)$$

where

$$h_n[k] = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j[n\omega - k\theta^{-1}(\omega)]} d\omega. \quad (4)$$

Warping the signal by invertible maps is equivalent to orthogonally project the signal over the set of sequences $h_n[k]$.

If the map is odd and differentiable, then the sequences

$$h_n[k] = \text{IDFT} \left[e^{-jn\theta(\omega)} \frac{d\theta}{d\omega} \right] \quad (5)$$

form a complete set, biorthogonal to the set

$$g_n[k] = \text{IDFT} \left[e^{-jn\theta(\omega)} \right], \quad (6)$$

so that

$$s[k] = \sum_n \hat{s}[n] g_n[k]. \quad (7)$$

By factoring the (positive) derivative

$$\frac{d\theta}{d\omega} = |F_0(\omega)|^2, \quad (8)$$

one can obtain the orthogonal and complete set

$$f_n[k] = \text{IDFT} \left[e^{-jn\theta(\omega)} F_0(\omega) \right]. \quad (9)$$

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

$$F_0(\omega) \hat{S}(\theta(\omega)) = S(\omega). \quad (10)$$

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 $[B_0, B_1]$

$$\int_{B_0}^{B_1} |S(\omega)|^2 d\omega = \int_{\theta(B_0)}^{\theta(B_1)} |\hat{S}(\omega)|^2 d\omega. \quad (11)$$

A simplification is possible if the input signal is casual: in this case only the casual part of $f_n[k]$ gives nonzero contributions to the expansion.

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

$$A(z) = \frac{z^{-1} - b}{1 - bz^{-1}} \quad \text{with } -1 < b < 1. \quad (12)$$

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

$$A(e^{j\omega}) = e^{-j\theta(\omega)}, \quad (13)$$

where

$$\theta(\omega) = -\arg A(e^{j\omega}) = \omega + 2 \tan^{-1} \left(\frac{b \sin \omega}{1 - b \cos \omega} \right) \quad (14)$$

is the associated frequency warping map, shown in Figure 2 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 (14) 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 all-pass filters in order to compute the frequency warped version of the signal.

By introducing the orthogonalizing factor

$$\Lambda_0(z) = \frac{\sqrt{1-b^2}}{1-bz^{-1}}, \quad (15)$$

one can show that the z -transforms of the basis set are

$$H_r(z) = \Lambda_0(z) A(z)^r, \quad r = 0, 1, \dots \quad (16)$$

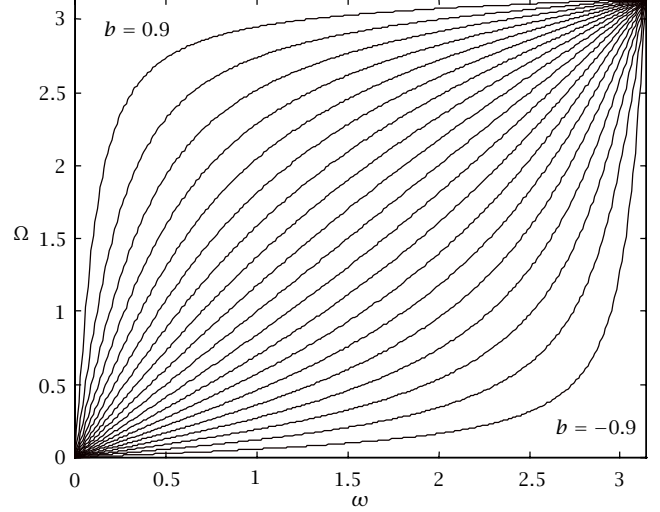


FIGURE 2: Family of Laguerre warping maps.

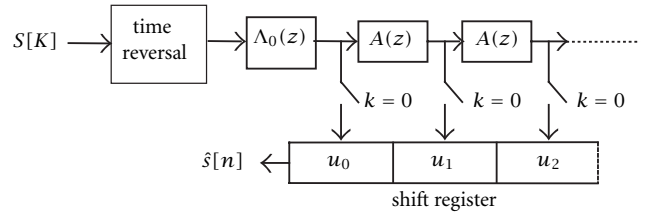


FIGURE 3: Digital structure implementing the Laguerre transform.

The operations involved in (3), with the recurrence (16) 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 3.

The inverse Laguerre transform structure is obtained simply by reversing the sign of the parameter b .

3. SPACE-VARYING DISPERSIVE DELAY LINES AND BIORTHOGONAL EXPANSIONS

Time-varying frequency warping is necessarily a time-frequency operation since we dynamically alter the spectral content of a signal. More intriguing, when implemented by means of dispersive delays, this operation requires a space-varying, that is, inhomogeneous line.

Consider the sampled dispersive delay line, shown in Figure 4, consisting of a chain of real first-order all-pass filters

$$A_n(z) = \frac{z^{-1} - b_n}{1 - b_n z^{-1}} \quad \text{with } -1 < b_n < 1, \quad (17)$$

a sampling device closing at time $k = 0$ and a shift-register loaded at $k = 0$ with the outputs of the filters and outputting the sequence of samples x_n at regular clock intervals. The dispersive line reverts to a linear delay line when all the parameters b_n are zero. It is easy to see that, upon time reversal

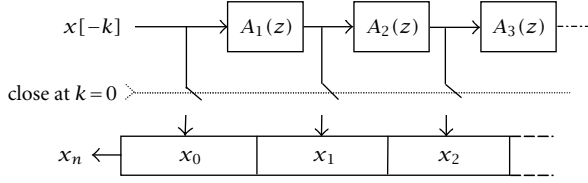


FIGURE 4: Sampled dispersive delay line.

of the input sequence, the line implements the scalar product

$$x_n = \langle \varphi_n, x \rangle = \sum_k x[k] \varphi_n[k], \quad (18)$$

where

$$\varphi_n[k] = a_1[k] * a_2[k] * \dots * a_n[k]. \quad (19)$$

Hence, the z -transform of the sequence $\varphi_n[k]$ is

$$\Phi_n(z) = \begin{cases} 1 & \text{if } n = 0, \\ \prod_{k=1}^n \frac{z^{-1} - b_k}{1 - b_k z^{-1}} & \text{if } n > 0, \end{cases} \quad (20)$$

which is the transfer function of an order n all-pass corresponding to a frequency dependent (dispersive) delay

$$\Phi_n(\omega) = e^{-j\Omega_n(\omega)}, \quad (21)$$

where

$$\Omega_n(\omega) = \sum_{k=1}^n \theta_k(\omega), \quad (22)$$

with

$$\begin{aligned} \theta_k(\omega) &= -\arg A_k(e^{j\omega}) \\ &= \omega + 2 \tan^{-1} \left(\frac{b_k \sin \omega}{1 - b_k \cos \omega} \right). \end{aligned} \quad (23)$$

The output sequence x_n may be interpreted as the coefficients of a suitable signal expansion. In fact, the set of sequences $\psi_n[k]$ whose z -transforms are

$$\Psi_n(z) = \begin{cases} \frac{1}{1 - b_1 z^{-1}} & \text{if } n = 0, \\ \frac{1 - b_n b_{n+1}}{(1 - b_n z)(1 - b_{n+1} z^{-1})} \Phi_n(z) & \text{if } n > 0, \end{cases} \quad (24)$$

can be shown to be biorthogonal to the set $\varphi_n[k]$, that is,

$$\langle \psi_n, \varphi_m \rangle = \sum_{k=0}^{\infty} \psi_n[k] \varphi_m[k] = \delta_{n,m} u[n], \quad (25)$$

$$\sum_{n=0}^{\infty} \psi_n[k] \varphi_n[m] = \delta_{k,m} u[k], \quad (26)$$

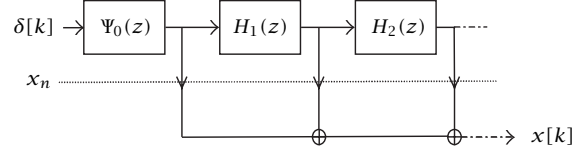


FIGURE 5: Structure implementing the inverse transform.

where

$$u[k] = \begin{cases} 1 & \text{if } k \geq 0, \\ 0 & \text{otherwise,} \end{cases} \quad (27)$$

is the unit step sequence, required since the set is complete over causal sequences (although it can be easily extended to noncausal sequences). We remark that although at first sight the sequences $\psi_n[k]$ may seem noncausal, by substituting (20) in (24) we obtain, for $n > 0$,

$$\Psi_n(z) = \frac{z^{-1}(1 - b_n b_{n+1})}{(1 - b_n z^{-1})(1 - b_{n+1} z^{-1})} \Phi_{n-1}(z), \quad (28)$$

which clearly denotes a causal sequence. Property (25) is easily shown by writing the scalar product in the z -transform domain

$$\langle \psi_n, \varphi_m \rangle = \frac{1}{2\pi j} \oint \Psi_n(z) \Phi_m(z^{-1}) z^{-1} dz \quad (29)$$

and by observing that the integrand is a rational function. For $n \neq m$ the degree of the denominator exceeds by 2 that of the numerator, hence the integral is zero, while for $n = m$ there is a single pole inside the unit circle whose residue is 1. Property (26) requires some technical conditions on the asymptotic behavior of the parameters b_n . However, any finite selection of them within the specified range $-1 < b_n < 1$ leads to a set that can be embedded in a biorthogonal complete set. Correspondingly, the signal $x[k]$ is expanded onto the set $\psi_n[k]$ as follows

$$x[k] = \sum_{n=0}^{\infty} x_n \psi_n[k], \quad (30)$$

where the coefficients are given by (18).

There are several equivalent structures for implementing the inverse transform. The one shown in Figure 5 is based on the following recurrence:

$$\Psi_n(z) = H_n(z) \Psi_{n-1}(z), \quad n \geq 1, \quad (31)$$

where

$$H_n(z) = \frac{1 - b_n b_{n+1}}{1 - b_{n-1} b_n} \frac{z^{-1} - b_{n-1}}{1 - b_{n+1} z^{-1}} \quad (32)$$

and we used the convention that $b_0 = 0$. The analysis coefficients x_n are used as weights for the dispersive tapped delay line in a structure that generalizes Laguerre filters [1, 2]. The first 30 analysis and synthesis basis elements generated

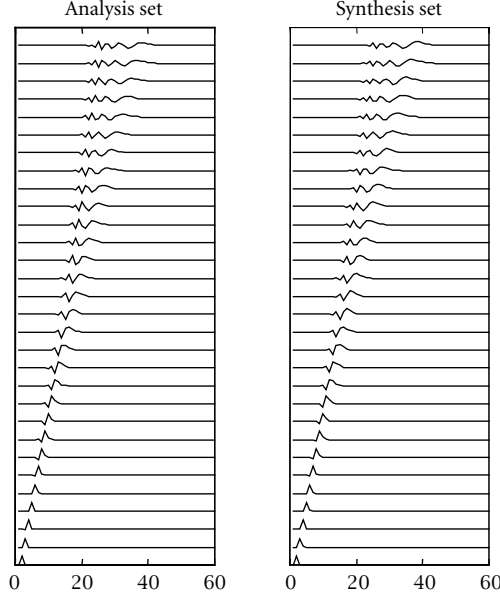


FIGURE 6: Analysis and synthesis sets of the biorthogonal with sinusoidal variation of parameters.

by a sinusoidal variation of the parameters are plotted in Figure 6.

For slow variations of the parameters the set is nearly orthogonal. As a final remark we note that the biorthogonal sequences $\psi_n[k]$ and $\varphi_n[k]$ may be used interchangeably for the analysis or for the synthesis. If the sequence of parameters $b_n = b$ is constant, one obtains the biorthogonal sets described in Section 2. We showed that this set can be orthogonalized without affecting the rational filter structure. Unfortunately, orthogonalization of space-varying dispersive delay lines yields nonrational transfer functions.

4. TIME-VARYING FREQUENCY WARPING

Time-varying frequency warping is obtained by means of the analysis structure shown in Figure 4. In previous papers [4, 7] and in Section 2 we analyzed the behavior of a homogeneous line, in which case the maps $\mathfrak{g}_k(\omega) = \mathfrak{g}(\omega)$ are identical and

$$\Omega_n(\omega) = n\mathfrak{g}(\omega). \quad (33)$$

One can show that

$$X(e^{j\omega}) = \Lambda_0(e^{j\omega})\hat{X}(e^{j\mathfrak{g}(\omega)}), \quad (34)$$

where

$$\hat{X}(e^{j\omega}) = \text{DTFT}[x_n] = \sum_{k=0}^{\infty} x_n e^{-jn\omega}. \quad (35)$$

Except for the first-order filter $\Lambda_0(e^{j\omega})$, equation (34) characterizes a pure frequency warping operation in that an input sinusoid with angular frequency ω_0 is displaced to angular

frequency $\mathfrak{g}^{-1}(\omega_0)$ in the output signal $y[n] = x_n$. The corresponding result when the sequence b_n is not constant has the following form:

$$X(e^{j\omega}) = \sum_{k=0}^{\infty} x_n e^{-j\Omega_n(\omega)}, \quad (36)$$

where we used the set $\varphi_n[k]$ for the synthesis. As expected, in time-varying warping time and frequency are mixed and not simply factored as in (33). In order to gain intuition on the features of this algorithm, suppose that the parameters b_n are periodically updated with rate $1/N$, then

$$\Omega_{q+kN}(\omega) = q\mathfrak{g}_{k+1}(\omega) + \Omega_{kN}(\omega), \quad q = 0, \dots, N-1. \quad (37)$$

Consider the STFT of the output when the signal is analyzed using the set $\varphi_n[k]$:

$$\hat{X}_k(m) = \sum_n x_n w[n - kN] e^{-j(2\pi/N)mn}, \quad (38)$$

where $w[n]$ is the rectangular window of length N , k is the time index and $(2\pi/N)m$ is the frequency. After some routine manipulation, we obtain

$$\begin{aligned} \hat{X}_k(m) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\mathfrak{g}_{k+1}^{-1}(\omega)}) e^{j\Omega_{kN}(\mathfrak{g}_{k+1}^{-1}(\omega))} \\ &\quad \times \frac{d\mathfrak{g}_{k+1}^{-1}(\omega)}{d\omega} W_m(\omega) d\omega, \end{aligned} \quad (39)$$

where

$$W_0(\omega) = e^{j(N-1/2)\omega} \frac{\sin(N\omega/2)}{\sin(\omega/2)} \quad (40)$$

is the Dirichlet kernel,

$$\begin{aligned} W_m(\omega) &= W_0\left(\omega - \frac{2\pi}{N}m\right), \\ \frac{d\mathfrak{g}_k^{-1}}{d\omega} &= \frac{1 - b_k^2}{1 + 2b_k \cos \omega + b_k^2}. \end{aligned} \quad (41)$$

Except for smearing produced by the finite window and filtering due to the derivative term, the STFT of the output signal is approximately:

$$\hat{X}_k(m) \approx NX(e^{j\mathfrak{g}_{k+1}^{-1}(2\pi m/N)}) e^{j\Omega_{kN}(\mathfrak{g}_{k+1}^{-1}(2\pi m/N))}, \quad (42)$$

which includes a warped version of the input with map $\mathfrak{g}_{k+1}(\omega)$ and a phase term due to the frequency dependent characteristic of the basis. We observe that slow variations of the parameters induce a time-dependent frequency warping of the signal and introduce frequency distortion of the envelopes.

5. APPLICATIONS AND EXAMPLES

Time-dependent frequency warping is an interesting effect and a powerful sound-editing tool. Algorithms based on this

concept can be devised in order to reduce a given real-life signal, showing pseudoperiodic features, to a nearly perfectly periodic one. By means of this technique we are able to compensate for slow frequency shifts such as vibrato in instrumental sounds or intonation in spoken or sung vowels. Vice versa, we can use this technique to artificially introduce these features as special effects.

5.1. Pitch modulation removal

In order to appreciate the power of the algorithm, we produced an example in which a spoken vowel pronounced with relevant intonation results in a time-varying pitch characteristic, as shown in the spectrogram of Figure 7(a). By using the inverse frequency law in time-varying frequency warping we were able to “regularize” the sound, reverting it to its almost constant pitch version shown in Figure 7(b). This transformation may be used in order to detect and track sound features such as formant shapes for both analysis and resynthesis purposes. Moreover, after reducing the signal to its periodic version, any pitch synchronous technique will work with constant pitch. In particular, this technique improves noise extraction in comb or multiplexed wavelet transforms. In this case, after the time-dependent frequency law of the signal is compensated for, a resonant comb filter removes noise on the resulting periodic signal, for example, by attenuating the part of the signal spectrum which falls far from the fundamental or its harmonics. The inverse transform will recover the source signal after denoising. In this case completeness or the possibility to revert the transform demonstrated in the above are invaluable results.

In Figure 8 we show an example of editing a flute sound in order to remove unwanted vibrato. Pitch detection of both the original and the time-varying warped signal was performed by means of a zero-crossing algorithm acting on the fundamental frequency. It is clear that the vibrato is completely removed by projection on the biorthogonal set. It should be noted that in flute sounds, vibrato is coupled with amplitude modulation. This effect can still be heard in the edited sound but it can be easily removed by means of time-domain techniques.

5.2. Detection

Another application is in the field of signal detection. Suppose that we have a signal showing a time-varying pitch characteristic and buried in high level noise.

If the signal is locally monochromatic or harmonic, we are able to compensate for a known frequency law. While detection of the source signal requires a time-frequency representation, by reducing the signal to the constant pitch case we may detect the signal by means of FFT. This will lower the variance of the detection error. Due to coherent averaging, the narrow band in the Fourier domain containing all the energy of our signal will stand clearly against the noise. This allows for the detection of signals in noise even at very low SNR. The example reported in Figure 9 shows detection of a

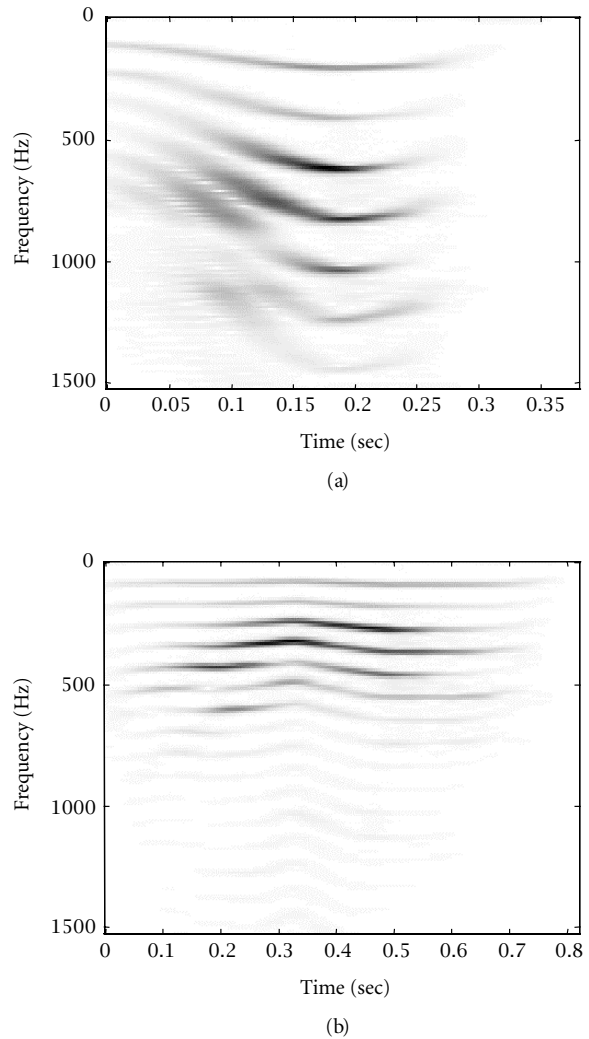


FIGURE 7: Spectrogram of (a) spoken /a/ with relevant intonation and (b) pitch-compensated signal obtained by means of time-varying frequency warping.

frequency sweep buried in noise with -15 dB SNR. The magnitude FT of the warped signal shows the line typical of a sine wave while the energy of the original signal was distributed in the frequency domain.

5.3. Pitch modulation effects

In the constant parameter case, the Laguerre transform proved very efficient for detuning a given signal. In the piano case this method is compatible with the physical model [13, 14] and produces a very natural pitch-shift of the piano sound with an increase of the degree of inharmonicity as the pitch lowers. In other sounds, this technique proved very efficient for microdetuning, that is, for introducing a slight transposition of the frequency content of the signal. Due to the unfavorable rational approximation requiring large factors, it is difficult to obtain this effect by a combination of upsampling and downsampling.

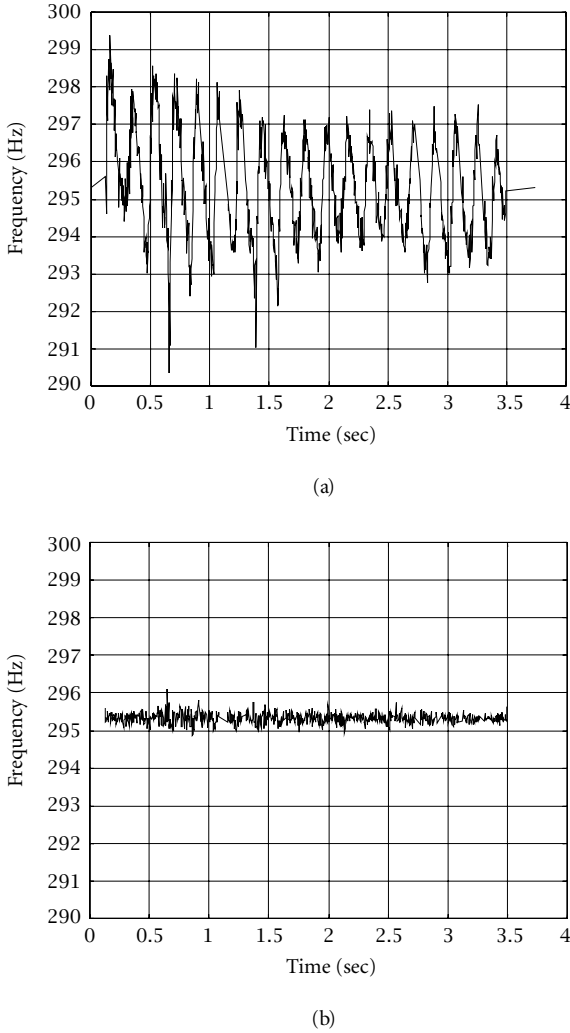


FIGURE 8: Pitch of flute sound: (a) with vibrato and (b) with vibrato removed by time-varying frequency warping.

In the time-varying case, a natural or synthesized sound may be modified according to a specified frequency law for vibrato or other effects. By adding this signal to the original one can introduce flanging, phasing, chorus and more general effects. The acoustical results are in most cases very impressive.

As a first example we introduced a sinusoidal fluctuation of the pitch of a flute sound. The frequency of the modulating signal was in turn modulated in order to achieve a realistic vibrato sound. The result is shown in Figure 10 where we report the spectrogram of the sound transformed by means of the biorthogonal set described in Section 3 with sinusoidal variation of the parameters. We notice that the depth of the pitch modulation increases with the frequency of the partial. This effect is due to the nonlinear warping map and it produces a more natural vibrato than equally modulating each partial.

As a second example we introduced a pitch tremolo effect on the same flute sound. This is achieved by means of

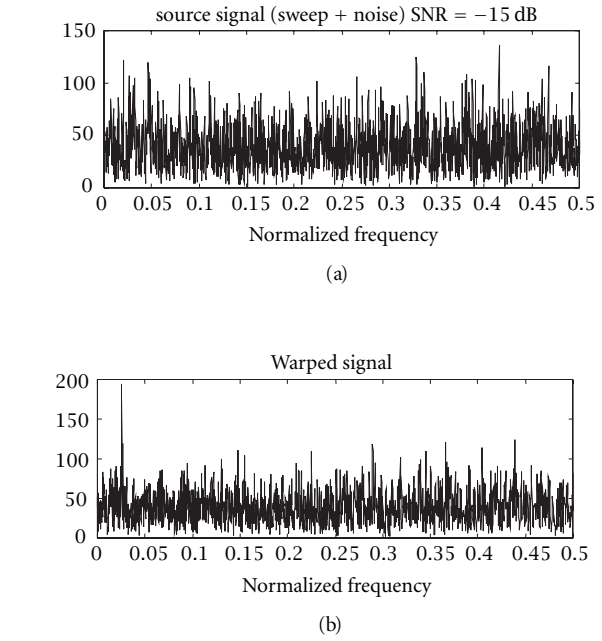


FIGURE 9: Detection of a sweep signal buried in noise (SNR = -15 dB): (a) magnitude FT of the original signal and (b) magnitude FT of the signal after pitch-regularization by means of time-varying frequency warping.

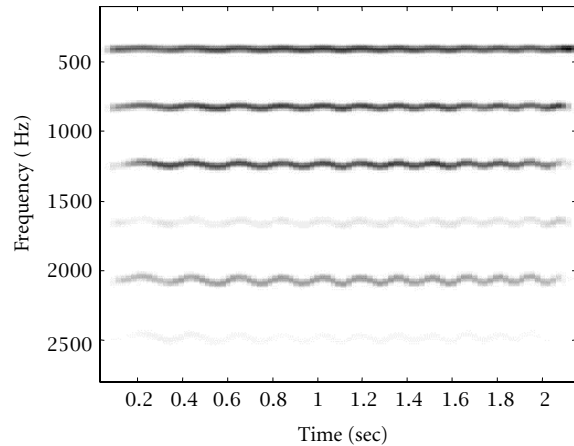


FIGURE 10: Spectrogram of flute sound with vibrato generated by means of time-varying warping with modulated sinusoidal law.

a square wave law for the generalized Laguerre parameters. The resulting spectrogram is shown in Figure 11. This example also demonstrates the ability of the transform to introduce rapid pitch fluctuations.

In Figure 12, we show the spectrogram of a Flatterzunge effect on the flute sound obtained by random variation of the parameters. Finally, in Figure 13 we show the spectrogram of a glissando effect obtained by an exponential law for the parameters.

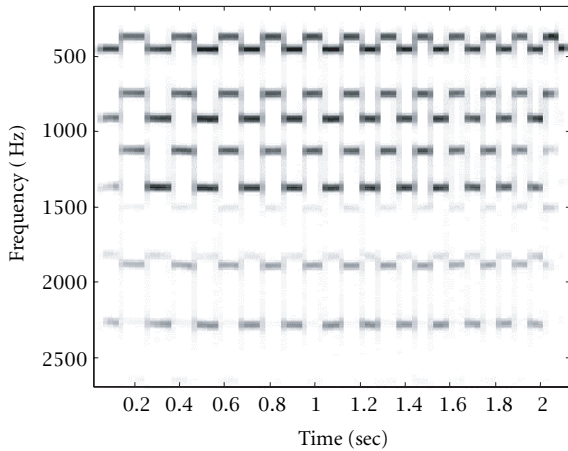


FIGURE 11: Spectrogram of flute sound with vibrato generated by means of time-varying warping with modulated square wave law.

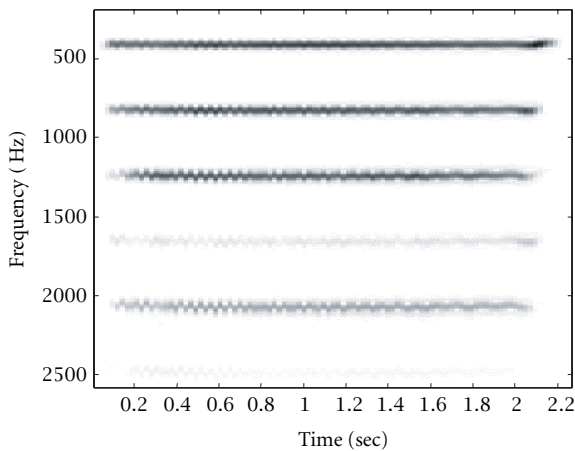


FIGURE 12: Spectrogram of flute sound with vibrato generated by means of time-varying warping with random law.

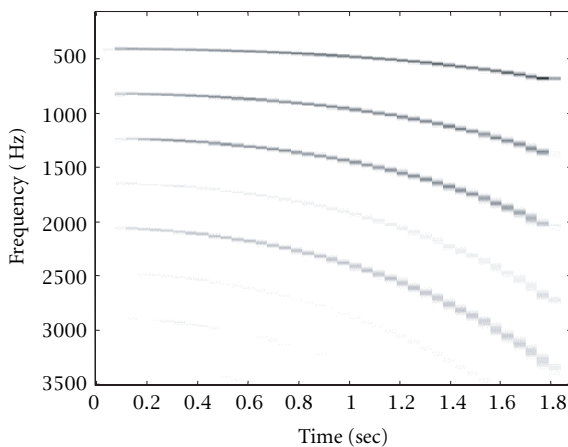


FIGURE 13: Spectrogram of flute sound with glissando generated by means of time-varying warping with exponential law.

6. CONCLUSIONS

In this paper we introduced a biorthogonal set for time-varying frequency warping. This signal expansion is associated with an inhomogeneous dispersive delay line, which provides an efficient computational structure for the transform. We demonstrated several applications of time-varying frequency warping in sound editing and signal detection. The technique proved very effective for introducing high-quality pitch-modulation effects or for eliminating or reducing these effects in recorded sounds. Interesting variations on the theme include mixing together several frequency-warped signals with the original sound in order to obtain innovative chorus, flanging and phasing effects and the like. A feature of the biorthogonal set is that the inserted effects can be easily undone without storing the original sound. The catalogue of known effects is subjected to be largely extended by the use of time varying frequency warping.

REFERENCES

- [1] P. W. Broome, "Discrete orthonormal sequences," *J. Assoc. Comput. Machinery*, vol. 12, no. 2, pp. 151–168, 1965.
- [2] A. V. Oppenheim and D. H. Johnson, "Discrete representation of signals," *Proc. IEEE*, vol. 60, pp. 681–691, 1972.
- [3] 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, 1995.
- [4] G. Evangelista and S. Cavaliere, "Frequency warped filter banks and wavelet transforms: a discrete-time approach via Laguerre expansion," *IEEE Trans. on Signal Processing*, vol. 46, no. 10, pp. 2638–2650, 1998.
- [5] G. Evangelista and S. Cavaliere, "Discrete frequency warped wavelets: theory and applications," *IEEE Trans. on Signal Processing*, vol. 46, no. 4, pp. 874–885, 1998.
- [6] G. Evangelista and S. Cavaliere, "Auditory modeling via frequency warped wavelet transform," in *Proc. EUSIPCO 98*, 1998, vol. 1, pp. 117–120.
- [7] G. Evangelista and S. Cavaliere, "Dispersive and pitch-synchronous processing of sounds," in *Proc. DAFX98*, 1998, pp. 232–236.
- [8] G. Evangelista and S. Cavaliere, "Arbitrary bandwidth wavelet sets," in *Proc. ICASSP'98*, 1998, vol. 3, pp. 1801–1804.
- [9] G. Evangelista and S. Cavaliere, "Representation of pseudoperiodic signals by means of pitch-synchronous frequency warped wavelet transform," in *Proc. EUSIPCO 98*, 1998, vol. 2, pp. 625–628.
- [10] G. Evangelista and S. Cavaliere, "Analysis and regularization of inharmonic sounds via pitch-synchronous frequency warped wavelets," in *Proc. ICMC '97*, 1997, pp. 51–54.
- [11] 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, 1994.
- [12] G. Evangelista, "Pitch synchronous wavelet representations of speech and music signals," *IEEE Trans. on Signal Processing*, vol. 41, no. 12, pp. 3313–3330, 1993.
- [13] I. Testa, S. Cavaliere, and G. Evangelista, "A physical model of stiff strings," in *Proc. ISMA'97*, 1997, pp. 219–224.
- [14] S. A. Van Duyne and J. O. Smith, "A simplified approach to modeling dispersion caused by stiffness in strings and plates," in *ICMC Proc.*, 1994.

Gianpaolo Evangelista received the laurea in physics (*summa cum laude*) from the University of Napoli, Napoli, Italy, in 1984 and the M.Sc. and Ph.D. degrees in electrical engineering from the University of California, Irvine, in 1987 and 1990, respectively. Since 1998 he has been a Scientific Adjunct with the Laboratory for Audiovisual Communications, Swiss Federal Institute of Technology, Lausanne, Switzerland, on leave from the Department of Physical Sciences, University of



Napoli Federico II, which he joined in 1995 as a Research Associate. From 1985 to 1986, he worked at the Centre d'Etudes de Mathématique et Acoustique Musicale (CEMAMu/CNET), Paris, France, where he contributed to the development of a DSP-based sound synthesis system, and from 1991 to 1994, he was a Research Engineer at the Microgravity Advanced Research and Support (MARS) Center, Napoli, where he was engaged in research in image processing applied to fluid motion analysis and material science. His interests include speech, music, and image processing; coding; wavelets; and multirate signal processing. Dr. Evangelista was a recipient of the Fulbright fellowship.

Sergio Cavaliere received the laurea in electronic engineering (*summa cum laude*) from the University of Napoli Federico II, Napoli, Italy, in 1971. Since 1974 he has been with the Department of Physical Sciences, University of Napoli, first as a Research Associate and then as an Associate Professor. From 1972 to 1973, he was with CNR at the University of Siena. In 1986, he spent an academic year at the Media Laboratory, Massachusetts Institute of Technology, Cambridge. From 1987 to 1991, he received a



research grant for a project devoted to the design of VLSI chips for real-time sound processing and for the realization of the Musical Audio Research Station (MARS) workstation for sound manipulation, IRIS, Rome, Italy. He has also been a Research Associate with INFN for the realization of very-large systems for data acquisition from nuclear physics experiments and for the development of techniques for the detection of signals in high level noise in the VIRGO experiment. His interests include sound and music signal processing, signal transforms and representations, VLSI, and specialized computers for sound manipulation.