REPRESENTATION AND MODIFICATION OF TIME-VARYING SOUND SIGNALS

Gianpaolo Evangelista

Audiovisual Communications Laboratory,
Swiss Federal Institute of Technology,
Lausanne, Switzerland.

e-mail: Gianpaolo.Evangelista@epfl.ch

Sergio Cavaliere

ACEL, Dipartimento di Scienze Fisiche,
Università "Federico II" di Napoli,
Complesso Universitario di M.S.Angelo,
Via Cinzia 80126 Napoli
e-mail: cavaliere@na.infn.it

Abstract

In recent papers the authors introduced in the range of acoustical signals the use of a powerful instrument for the analysis and modification of signals: the Laguerre Transform, mapping a signal space into another one whose frequency axis is warped in a controlled way. While altering the overall frequency content of specified signals is very useful in many applications, most real world signals show time-varying features, e.g., in both their amplitude and frequency content. A second step, therefore, has been that of extending the principles and the architecture of the Discrete Laguerre Transform to the time varying case so that the frequency content of a signal could be displaced over time to different values, realizing therefore a time varying warping transform. This transform has nice features and high regularity, it allows perfect reconstruction and it can effectively succeed in regularizing real world signals or in modifying in a controlled way some of their relevant parameters. These features well match those of ordinary sound signals whose frequency content is slowly varying with time, such as intonation in speech, glissando and vibrato in music. It also meets many processing needs for a wide range of sweep signals. In the paper, after briefly recalling the relevant features of the recently introduced transform, many examples are given, demonstrating the wide range of applications for which it seems to be well suited.

1 Introduction

Most real time signals depart significantly from the ideal model of a constant-pitch, oscillating signal. Due to many different physical reasons, the underlying oscillatory features are mostly time varying. Also, some relevant information is bound to these features, such as the perception of a pleasant timbre, the internal texture of a sound, interpretation cues and communicative or expressive contents. A transform able to compensate against these effects or, in turn, to add these features would be without doubt a strong tool for the analysis, the representation and the modification of sound signals.

2 Frequency Warping and the Discrete Laguerre Transform

Starting point of our work is the Discrete Laguerre Transform. It can be shown that by projecting a discrete time signal onto the orthonormal set of Laguerre sequences [1][2] we obtain a new signal whose spectrum $\hat{X}(e^{j\omega})$ is simply the frequency-warped version of the source spectrum $X(e^{j\omega})$:

$$X(e^{j\omega}) = \Lambda_0(e^{j\omega})\hat{X}(e^{j\vartheta(\omega)}).$$

Here $\Lambda_0(e^{j\omega})$ is a normalizing factor, which, due to the unitary property of the transform, preserves energy while passing from the source to the destination domain. The warping law $\vartheta(\omega)$ is a function of a single parameter, the real pole of a first order all-pass filter section. The transformation may be controlled by this parameter -- the Laguerre parameter -- which allows a large degree of freedom in the choice of the warping characteristics. The authors used this transformation in conjunction with other transforms such as the Discrete Wavelet Transform and the Pitch Synchronous Wavelet Transform, introducing a new class of orthogonal transforms having the advantage of arbitrarily allocating the analysis bands [3][4][11]. They have also pointed out an entire new range of applications for the introduced transforms [8][9][12].
Finally, the Laguerre Transform can be computed by conventional DSP methods, by means of the all-pass cascade structure depicted in Fig. 1. The time-varying structure generalizes that of the constant case. In the constant case all the filters $A_i(z)$ are equal to a single all pass function whose phase $\vartheta(\omega)$ actually turns out to be the warping law described in the above.

2 The Generalized Discrete Time Laguerre Transform

The Laguerre transform may be extended to its time varying version. For this purpose, however, we must introduce a new parametric class of sequences, namely the Time-Varying Discrete Laguerre Sequences [10][11][12]. These sequences generalize the Laguerre sequences in that they allow for modification of the

frequency content of the signal by means of a time-varying frequency warping law.
Consider the sampled dispersive delay line shown in Fig. 1, consisting of a chain of real first-order all-pass filters

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

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 $\hat{x}[k]$ at regular clock intervals. The dispersive line reverts to a linear delay line when all the parameters b_n are zero, and in the case of b_n constant from section to section, it reverts to the structure for computing the ordinary Laguerre Transform.

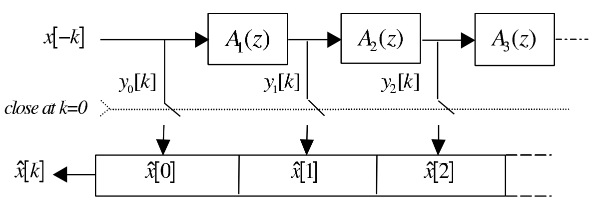

Figure 1. Sampled dispersive delay line.

Since the input sequence is time-reversed, the line implements the scalar product

$$\hat{x}[n] = \langle \varphi_n, x \rangle = \sum_k x[k]\, \varphi_n[k], \tag{1}$$

where

$$\varphi_n[k] = a_1[k] * a_2[k] * \ldots * a_n[k] .$$

Hence, the z-transform of the sequence $\varphi_n[k]$ is an order n all-pass filter with transfer function

$$\Phi_n(z) = \begin{cases} 1 & \text{if } n = 0 \\ \prod_{k=1}^{n} \frac{z^{-1} - b_k}{1 - b_k z^{-1}} & \text{if } n > 0 \end{cases} . \tag{2}$$

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

$$\Psi_n(z) = \begin{cases} \frac{1}{1 - b_1 z^{-1}} & \text{if } n = 0 \\ \frac{1 - b_n b_{n+1}}{(1 - b_n z)(1 - b_{n+1} z^{-1})} \Phi_n(z) & \text{if } n > 0 \end{cases} , \tag{3}$$

can be shown ([11][12]) to be biorthogonal to the set $\varphi_n[k]$, i.e.,

$$\langle \psi_n, \varphi_m \rangle = \sum_{k=0}^{\infty} \psi_n[k] \varphi_m[k] = \delta_{n,m} u[n] \tag{4}$$

and

$$\sum_{n=0}^{\infty} \psi_n[k] \varphi_n[m] = \delta_{k,m} u[k], \tag{5}$$

where $u[k]$ is the unit step sequence. The set is complete over causal sequences. Property (5) requires some tecnical 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; this is in fact the practical case where the signal to transform is causal and also has finite duration. Correspondingly, the signal $x[k]$ is expanded onto the set $\psi_n[k]$ as follows

$$x[k] = \sum_{n=0}^{\infty} \hat{x}[k] \psi_n[k], \tag{6}$$

where the coefficients are given by (1).
There are several equivalent structures for implementing the inverse transform. The one shown in Fig. 2 is based on the following recurrence:

$$\Psi_n(z) = V_n(z) \Psi_{n-1}(z), \qquad n \geq 1$$

where

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

and we used the convention that $b_0 = 0$. The analysis coefficients $\hat{x}[k]$ are used as weights for the dispersive tapped delay line in a structure that generalizes Laguerre filters [1][2].
If the sequence of parameters $b_n = b$ is constant, the resulting transform, as expected, reverts to a biorthogonal variant of the Laguerre transform. The biorthogonal sequences $\psi_n[k]$ and $\varphi_n[k]$ may be used interchangeably for the analysis or for the synthesis, with obvious modifications of the structures.

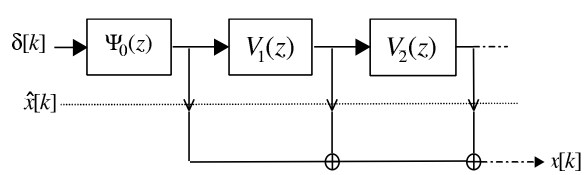

Figure 2. Structure implementing the inverse biorthogonal transform

4. Time-Varying Frequency Warping

Time-varying frequency warping is obtained by means of the analysis structure shown in Fig. 1, which, actually, may be seen as a special type of time-varying filter. Each of the y_n lines, n=0,1,..., carries a filtered version of the input signal $x(-k)$, the transfer function at index n being the cascade of the sections up to n. The signals $y_n(k)$ are therefore filtered versions of the input signal by the transfer functions

$$\Phi_n(z) = A_1(z) \cdot A_2(z) \cdots A_n(z).$$

For a real signal $x(k)$ we have in the frequency and time domains, respectively:

$$Y_n(\omega) = X^*(\omega)\cdot\Phi_n(\omega),$$

$$y_n(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} X^*(\omega)\left(\prod_{m=1}^{n} A_m(\omega)\right)e^{jk\omega}d\omega$$

The *n-th* sample of the output sequence stored in the shift register is given by the sample of the $y_n(k)$ sequences, at time k=0:

$$\hat{x}[n] = y_n(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} X^*(\omega)\left(\prod_{m=1}^{n} A_m(\omega)\right)d\omega = \langle x,\varphi_n\rangle \quad (7)$$

where,

$$\langle x,\varphi_n\rangle = \sum_k x(k)\,\varphi_n(k)$$

is the orthogonal projection coefficient of the signal over the analysis set.

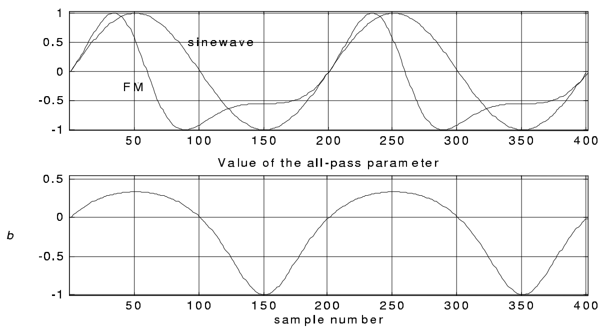

Fig. 3 Frequency modulation by means of warping

The $A_m(\omega)$ are all-pass functions characterized by a pure phase response: $A_m(\omega) = e^{-j\vartheta_m(\omega)}$, where for $|\omega| < \pi$

$$\vartheta_m(\omega) = \omega + 2\arctan\frac{b_r\sin\omega}{1-b_r\cos\omega} = 2\arctan\left(\frac{1+b_r}{1-b_r}\tan\frac{\omega}{2}\right). \quad (8)$$

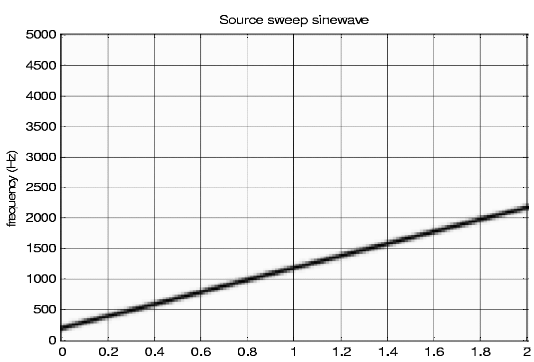

Fig. 4 A deeply swept sinusoid

At the *n*-th stage we obtain:

$$\Phi_n(\omega) = \prod_{r=1}^{n} e^{-j\vartheta_r(\omega)} = e^{-j\cdot\sum_{r=1}^{n}\theta_r(\omega)}.$$

For signals which are close to be periodic, as happens in most sound signals, the resulting frequency warping for each partial may be analyzed in the following way. For each partial we fall in the simple case of a single complex exponential tone

$$x(k) = e^{jk\omega_0},$$

with

$$X(\omega) = 2\pi\delta(\omega - \omega_0) \text{ for } |\omega| < \pi.$$

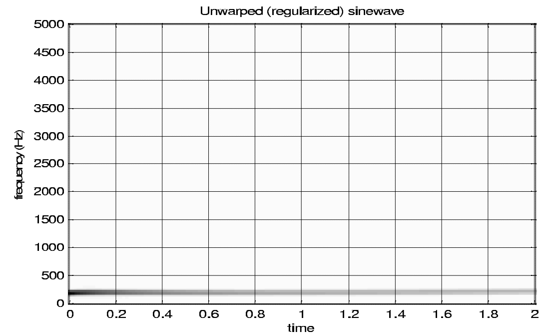

Fig. 5 The swept sinusoid is completely "regularized"

In this case, using the above inversion formula (7), we obtain for the output sample:

$$\hat{x}[n] = e^{j\cdot\sum_{r=1}^{n}\theta_r(\omega_0)} \quad (9)$$

This formula has a simple meaning when the all-pass sections are identical, with $\theta_r(\omega)=\theta(\omega)$, *r=1,2….* In this case we have: $\hat{x}[n] = e^{j\cdot n\cdot\theta(\omega_0)}$ and the output is a complex exponential whose frequency is the image of the source frequency ω_0 via the warping map $\theta(\omega)$, as expected in this case of ordinary Laguerre Transform.

If different warping laws are applied at each stage by using distinct values of the parameters b_r, then (9) corresponds to a phase-modulated signal, whose features and spectrum depend on the $\theta_r(\omega)$ laws. This signal corresponds to a frequency modulated sinusoid, depending on the choice of b_n, as shown in Fig. 3.

Frequency warping works properly even in case of drastic modification in the source frequency as shown for the linear frequency sweep in Fig. 4, which, after warping appears perfectly reduced to the single tone as in Fig. 5.

5. Acoustical applications of the transform

A first example concerns vibrato removal. Once we have identified the law of frequency variation over time in a sound, we may compute the sequence b_n of values of the Laguerre parameters by means of which the vibrato may be removed. Actually warping the sound by use of this sequence of parameters allows complete removal of vibrato, as shown in Fig. 6. In Fig. 7 we can see that the low frequency component (amplitude envelope) is practically unmodified under the warping transform. In fact the low frequency components, at the usual small values of the Laguerre parameters, move along an approximately 45-degree straigth line under the warping law.

In other applications we may want to increase the depth of vibrato in order to add some expression to our sound. The same warping law may be added, but now the Laguerre parameters will be chose to have opposite sign and will be scaled by a proper coefficient in order to impart any desired depth of vibrato: $-kb_n$. Finally the same depth may be easily modified over time, as in fact it happens in real signals. For example, the vibrato depth may be linked in an arbitrary way to the amplitude envelope of the signal.

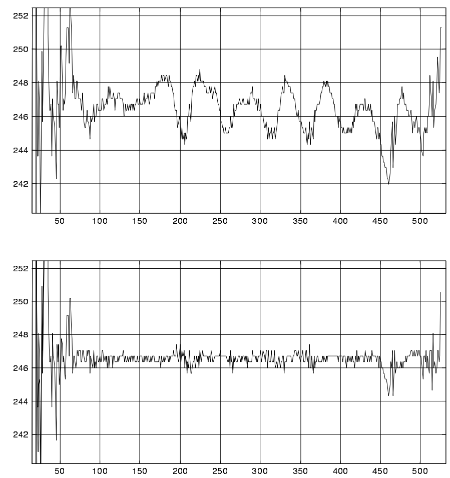

Fig. 6 The pitch of a flute sound showing vibrato and the same after removal by TVFW.

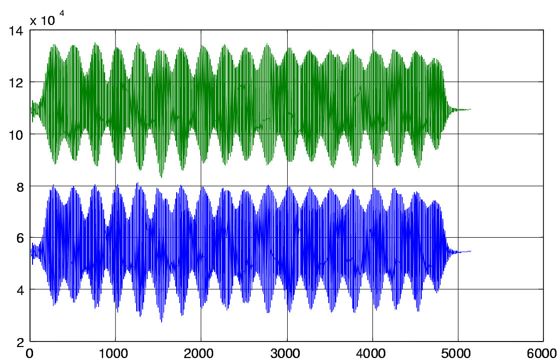

Fig. 7 Time domain flute sound before and after the removal of vibrato

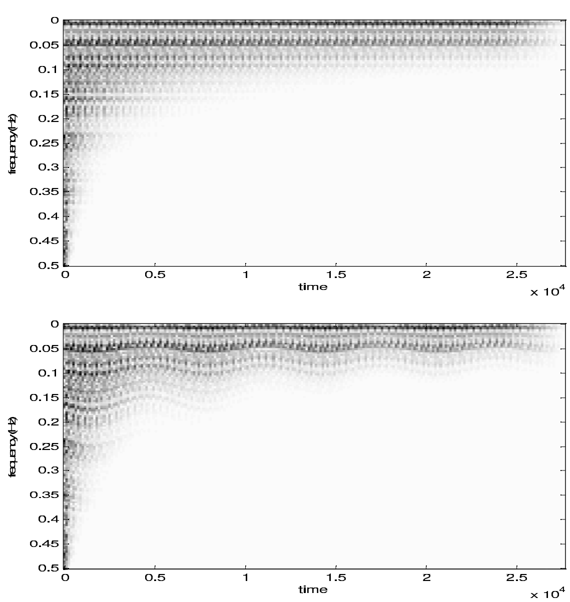

Fig. 8 Karplus-Strong plucked string original and with added vibrato.

A second example is provided, where we added to a simple sound produced by the Karplus-Strong algorithm a selected amount of vibrato (Fig. 8). In this case we may impart the desired amount of 'dispersion' to the harmonic structure (see [5]) and, in the mean time, also add vibrato, thus improving the naturalness of the sound. This is obtained by adding a non-zero mean value to the sequence of warping coefficients. In this way we obtain a mixture of effects typical of the constant and time-varying frequency warping transforms.

Other applications of time-varying frequency warping range from flanging-phasing to chorusing effects, when one mixes multiple time-varying warped versions of the same signal with the original signal, using suitable gains.

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, June 1972.

[3] 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, Oct. 1998.

[4] G. Evangelista and S. Cavaliere, “Discrete Frequency Warped Wavelets: Theory and Applications,” IEEE Trans. on Signal Processing, vol. 46, no. 4, pp.874-885, April 1998.

[5] I. Testa, G. Evangelista and S. Cavaliere, "A Physical Model of Stiff Strings," Proc. of the Internat. Symp. on Music and Acoustics, ISMA'97, pp. 219-224 Edinburgh, U.K., Aug. 1997 .

[6] G. Evangelista, S. Cavaliere, “Dispersive and Pitch-Synchronous Processing of Sounds,” Proc. DAFX98, pp. 232-236, Nov. 1998.

[7] G. Evangelista, S.Cavaliere, “Arbitrary Bandwidth Wavelet Sets,” Proc. ICASSP'98, vol. III, pp. 1801-1804, May 1998.

[8] S.Cavaliere, G. Evangelista, “Representation of Pseudoperiodic Signals by means of Pitch-Synchronous Frequency Warped Wavelet Transform,” Proc. EUSIPCO 98, vol. II, pp. 625-628, Sept. 1998.

[9] G. Evangelista, S. Cavaliere, “Analysis and Regularization of Inharmonic Sounds via Pitch-Synchronous Frequency Warped Wavelets,” Proc. ICMC '97, pp. 51-54, Sept. 1997.

[10] G. Evangelista, S. Cavaliere, “Time-Varying Frequency Warping: Results and Experiments,” Proc. DAFX99 Nov. 1999.

[11] G. Evangelista, "Dyadic Warped Wavelets," invited paper to appear on Advances in Imaging and Electron Physics, Academic Press, Dec. 2000.

[12] G. Evangelista, S. Cavaliere, "Audio Effects Based on Biorthogonal Time-Varying Frequency Warping," to appear on Applied Signal Processing, Springer-Verlag, Dec. 2000.