T m -V rying Frequency Warp ng:
Results and Exp rim nts

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

Abstract

Dispersive tapped delay lines are attractive structures for altering the frequency content of a signal. In previous papers we showed that in the case of a homogeneous line with first order all-pass sections the signal formed by the output samples of the chain of delays at a given time is equivalent to compute the Laguerre transform of the input signal. 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. Simulations of these effects require techniques for time-varying pitch and/or brightness modification that are very useful for sound processing. In our experiments 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 may be associated to 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: signal transformations, frequency warping

1 Introduction

In recent papers [3-9] the authors considered frequency warping by means of 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 and, at the same time, 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 based on a rational filter structure is useful for the construction of wavelet bases with arbitrary frequency band allocation [4]. The new combined transform leads to fine applications such as orthogonal and complete perceptual filter banks [5,7].

The Laguerre transform may also be used for adapting quasi-periodic sounds to pitch-synchronous schemes [6]. 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 [8,9]. 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 [6]. 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 [10,4]. 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.

2 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, i.e., inhomogeneous line.

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 x_n at regular clock intervals. The dispersive line reverts to a linear delay line when all the parameters b_n are zero.

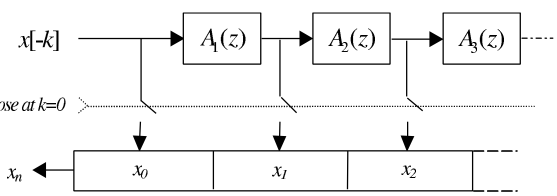

Fig. 1 *Sampled dispersive delay line.*

It is easy to see that, upon time reversal of the input sequence, the line implements the scalar product

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

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

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

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

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

where

$$\Omega_n(\omega) = \sum_{k=1}^{n} \vartheta_k(\omega),$$

with

$$\theta_k(\omega) = -\arg A_k(e^{j\omega}) = \omega + 2\tan^{-1}\left(\frac{b_k \sin\omega}{1 - b_k \cos\omega}\right)$$

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

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

can be shown 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

$$u[k] = \begin{cases} 1 & \text{if } k \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

is the unit step sequence, required since the set is complete over causal sequences (although it can be easily extended to non-causal sequences). We remark that although at first sight the sequences $\psi_n[k]$ may seem non-causal, by substituting (2) in (3) we obtain for n>0:

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

which clearly denotes a causal sequence. Property (4) is easily shown by writing the scalar product in the z-transform domain:

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

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 (5) 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

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

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:

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

where

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

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].

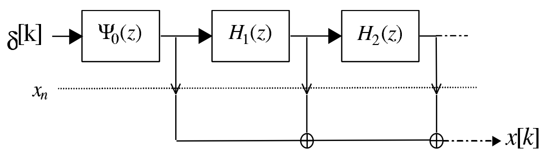

Fig. 2 *Structure implementing the inverse transform*

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 a biorthogonal sets that can be orthogonalized without affecting the rational filter structure. In fact, one obtains the set $\lambda_r(k)$ whose z-transforms are

$$\bar{\Lambda}_0(z) = \frac{\sqrt{1-b^2}}{1-bz^{-1}}$$

$$\bar{\Lambda}_r(z) = \bar{\Lambda}_0(z) A(z)^r \qquad r > 0 .$$

Unfortunately, orthogonalization of space-varying dispersive delay lines yields non-rational transfer functions.

3 Time-Varying Frequency Warping

Time-varying frequency warping is obtained by means of the analysis structure shown in Fig. 1. In previous papers [3,6] we analyzed the behavior of a constant parameter line, in which case the maps $\vartheta_k(\omega) = \vartheta(\omega)$ are identical and

$$\Omega_n(\omega) = n\vartheta(\omega) . \qquad (6)$$

One can show that

$$X(e^{j\omega}) = \bar{\Lambda}_0(e^{j\omega}) \hat{X}(e^{j\vartheta(\omega)}) \qquad (7)$$

where

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

Except for the first-order filter $\bar{\Lambda}_0(e^{j\omega})$, equation (7) characterizes a pure frequency warping operation in that an input sinusoid with angular frequency ω_0 is displaced to angular frequency $\vartheta^{-1}(\omega_0)$ in the output signal $y[n] = x_n$. The corresponding result when the sequence b_n is not constant has the following form

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

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 (6). In order to gain intuition on the features of this algorithm, suppose that the parameters b_n are periodically updated with rate $\frac{1}{N}$, then

$$\Omega_{q+kN}(\omega) = q\vartheta_{k+1}(\omega) + \Omega_{kN}(\omega) , \quad q = 0,..,N-1 .$$

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

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

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

$$\hat{X}_k(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\vartheta_{k+1}^{-1}(\omega)}) e^{j\Omega_{kN}(\vartheta_{k+1}^{-1}(\omega))} \frac{d\vartheta_{k+1}^{-1}}{d\omega} W_m(\omega) d\omega$$

where

$$W_0(\omega) = e^{j\frac{N-1}{2}\omega} \frac{\sin\frac{N\omega}{2}}{\sin\frac{\omega}{2}} , \qquad W_m(\omega) = W_0(\omega - \tfrac{2\pi}{N}m)$$

is the Dirichlet kernel, and

$\frac{d\vartheta_k^{-1}}{d\omega} = \frac{1-b_k^2}{1+2b_k\cos\omega + b_k^2}$. Except for smearing produced by the finite window and filtering due to the derivative term, the STFT of the output signal is approximately:

$$\hat{X}_k(m) \approx NX(e^{j\vartheta_{k+1}^{-1}(\frac{2\pi m}{N})}) e^{j\Omega_{kN}(\vartheta_{k+1}^{-1}(\frac{2\pi m}{N}))} ,$$

which includes a warped version of the input with map $\vartheta_{k+1}(\omega)$ 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.

4 Applications

Time-dependent frequency warping may be invaluable 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.

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 Fig. 3(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 Fig. 3(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, e.g., 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.

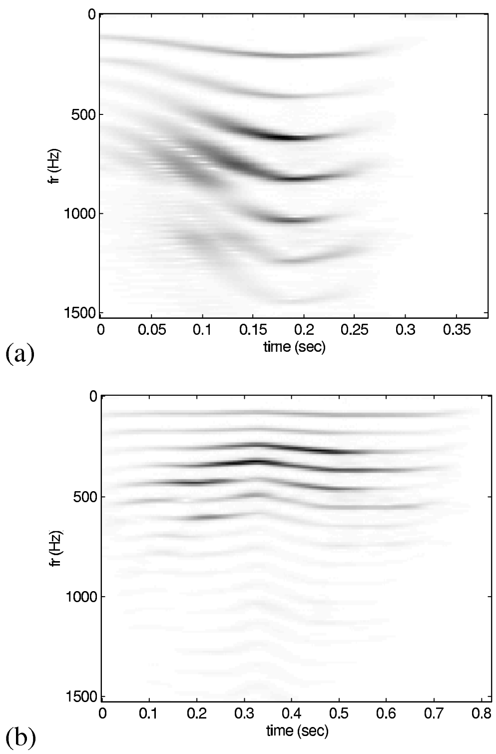

Fig. 3 *Spectrogram of (a) spoken /a/ with relevant intonation and (b) pitch-compensated signal obtained by means of time-varying frequency warping.*

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 reverting the signal to the constant pitch case we may identify the signal by means of the periodogram, which, by averaging will reduce the variance of the estimate. Due to coherent averaging, the narrow band in the periodogram containing all the energy of our signal will stand clearly against noise. This allows for the detection of signals in noise even at very low SNR, e.g., -15 dB.

Finally, for musical purposes, microdetuning, i.e., a slight transposition of the frequency content of the signal, proved very efficient in the constant parameter case. 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. Acoustical results are in some cases very impressive.

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] G. Evangelista, S. Cavaliere, "Auditory Modeling via Frequency Warped Wavelet Transform," *Proc. EUSIPCO 98,* vol. I, pp. 117-120, Sept. 1998.

[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] I. Testa, S. Cavaliere and G. Evangelista, "A Physical Model of Stiff Strings," *Proc. ISMA'97*, pp. 219-224, Aug. 1997.

[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, Feb. 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, Dec. 1993.