

REAL-TIME TIME-VARYING FREQUENCY WARPING VIA SHORT-TIME LAGUERRE TRANSFORM

Gianpaolo Evangelista

Audio-Visual Communication Laboratory
 Swiss Federal Institute of Technology,
 Lausanne, Switzerland
 gianpaolo.evangelista@epfl.ch

Department of Physical Sciences
 University "Federico II",
 Naples, Italy

ABSTRACT

In this paper we address the problem of the real-time implementation of time-varying frequency warping. Frequency warping based on a one-parameter family of one-to-one warping maps can be realized by means of the Laguerre transform and implemented in a non-causal structure. This structure is not directly suited for real-time implementation since each output sample is formed by combining all of the input samples. Similarly, the recently proposed time-varying Laguerre transform has the same drawback. Furthermore, long frequency dependent delays destroy the time organization or macrostructure of the sound event.

Recently, the author has introduced the Short-Time Laguerre Transform for the approximate real-time implementation of frequency warping. In this transform the short-time spectrum rather than the overall frequency spectrum is frequency warped. The input is subdivided into frames that are tapered by a suitably selected window. By careful design, the output frames correspond to warped versions of the input frames modulated by a stretched version of the window. It is then possible to overlap-add these frames without introducing audible distortion.

The overlap-add technique can be generalized to time-varying warping. However, several issues concerning the design of the window and the selection of the overlap parameters need to be addressed. In this paper we discuss solutions for the overlap of the frames when the Laguerre parameter is kept constant but distinct in each frame and solutions for the computation of full time-varying frequency warping when the Laguerre parameter is changing within each frame.

1. INTRODUCTION

Frequency warping is a general signal processing technique consisting in remapping the frequency axis to obtain a signal with desired characteristics. The spectral content of the signal is modified by displacing bands or partials to other frequency supports. Recently, a renewed interest in this technique led to interesting applications in audio signal processing [6, 8] and to new representations of sounds. Transforms based on frequency warping methods are the Frequency Warped Wavelet Transform and its pitch-synchronous version [3, 4, 7, 10], which allow for a flexible design of the basis elements that can be based on perceptual scales [5, 9].

Frequency warping is per se an interesting effect, which can be employed in sound morphing, detuning of partials, pitch-shifting in inharmonic or quasi-harmonic sounds. While general maps can be in principle designed, the exact implementation is possible for the one-parameter family of Laguerre warping curves. Laguerre warping leads to a unitary signal representation in terms of an expansion in orthogonal bases [1,2]. The implementation of this transform can be given in terms of a sampled dispersive delay line, shown in Fig. 1, in which the delay elements are allpass filters with frequency dependent characteristics. In recent papers the author generalized the Laguerre transform to a time-varying version implemented in a space-varying dispersive delay line formed by non-uniform allpass filter sections [8,11]. An interesting class of modulation effects and expression controls, such as vibrato, Flatterzunge, flanging, etc., can be implemented by means of this biorthogonal transform. While the Laguerre transform and its generalized time-varying versions can be implemented in standard DSP operations, their structure does not allow for real-time implementation since, in principle, all the signal samples are involved in the computation of the output samples. Another drawback of the frequency-warping algorithm, particularly critical in the ordinary Laguerre transform, is that the macrostructure of the signal is modified due to long frequency dependent delays. Thus, events are spectrally decomposed and different spectral regions “travel” at different speed along the line. According to the parameter of the transform, high frequency bands are perceived long before or long after low frequency bands pertaining to the same event. This effect is reduced in the time-varying Laguerre transform if the parameters vary in an oscillatory fashion. In order to circumvent this problem and, at the same time, obtain a warping algorithm suitable for real-time implementation, the author has recently proposed an overlap-add method on a running window version of the Laguerre transform [12].

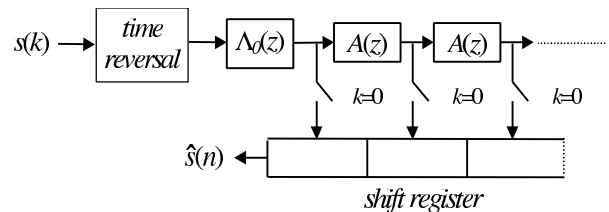


Figure 1. Dispersive delay-line structure for frequency warping via the Laguerre transform.

2. OVERLAP-ADD TECHNIQUES FOR STATIC FREQUENCY WARPING

The Short-Time Laguerre Transform (STLT) is computed by Laguerre transforming windowed frames of the signal $s(n)$. This gives rise to a 2D representation of a 1D signal obtained by orthogonal projection over a local Laguerre basis. Therefore, the STLT is the collection of the frequency-warped versions of the windowed input frames. Given a finite-length N window sequence $w(n)$ and an integer $L \leq N$, we define

$$q_r(n) = STLT[s]_{r,n} = \sum_{k=0}^{N-1} w(k)s(k+rL)\lambda_n(k) \quad (1)$$

where $\lambda_n(k)$ are the Laguerre sequences, whose z-transforms are given by

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

$$\Lambda_r(z) = \Lambda_0(z)A(z)^r, \quad r=0,1, \dots, \quad (3)$$

with

$$A(z) = \frac{z^{-1}-b}{1-bz^{-1}} \quad (4)$$

for $-1 < b < 1$. Our goal is to obtain:

$$\sum_r q_r(n-rM) \approx \hat{s}(n)$$

$$s(n) = LT[s]_n = \sum_k s(k)\lambda_n(k)$$

$s(n)$, i.e., a normalized frequency warped version of the input signal. In the frequency domain this corresponds to

$$\Lambda_0(\omega)\hat{S}(\theta(\omega)) = S(\omega), \quad (7)$$

where

$$\theta(\omega) = -\arg A(e^{j\omega}) = \omega + 2 \tan^{-1} \left(\frac{b \sin \omega}{1 - b \cos \omega} \right) \quad (8)$$

is the sign-reversed phase of the allpass filter (4) and $\Lambda_0(\omega)$ is the causal square root of the first derivative of $\theta(\omega)$. Equation (5) represents an approximate way of computing frequency warping by overlap-adding several short-time warped frames. In principle, for any fixed r , each sequence $q_r(n)$ has infinite length. However, by a minimum group delay argument one can show that at most the first

$$P \approx \frac{1+|b|}{1-|b|} N \quad (9)$$

samples are essentially non-zero, independently on the signal. Therefore, for computational purposes, the warped frames may be considered to have finite length. However, the length of the output frame is not conformal to that of the input frame. For this

reason the frame spacing M in (5) is generally different from the input frame spacing L in (1). Our goal is achieved if we are able to show that

$$q_r(n) \approx h(n)\hat{s}(n+rM), \quad (10)$$

where $h(n)$ is a finite-length window satisfying

$$\sum_r h(n-rM) = 1 \quad \text{for some integer } M > 0$$

$$\sum_n q_r(n)\lambda_n(k) = w(k)s(k+rL)$$

$$DIFT[w(k)s(k+rL)](\omega) = \int_{-\pi}^{+\pi} W(\omega-\Omega)S_r(\Omega)d\Omega, \quad (13)$$

where

$$S_r(\omega) = DIFT[s(k+rL)](\omega) = e^{j\omega rL} S(\omega). \quad (14)$$

Hence

$$\Lambda_0(\omega)Q_r(\theta(\omega)) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} W(\omega-\Omega)S_r(\Omega)d\Omega, \quad (15)$$

and

$$Q_r(\omega) = \Lambda_0^T(\omega) \frac{1}{2\pi} \int_{-\pi}^{+\pi} W(\theta^{-1}(\omega)-\Omega)S_r(\Omega)d\Omega, \quad (16)$$

where

$$\Lambda_0^T(\omega) = [\Lambda_0(\theta^{-1}(\omega))]^{-1} \quad (17)$$

is obtained from $\Lambda_0(\omega)$ by reversing the sign of the parameter b and it corresponds to the causal square root of the derivative of $\theta^{-1}(\omega)$. By performing the change of variable

$$\alpha = \theta(\Omega - \theta^{-1}(\omega)) + \omega \quad (18)$$

in the integral (16) we obtain

$$Q_r(\omega) = \Lambda_0^T(\omega) \frac{1}{2\pi} \int_{-\pi}^{+\pi} \hat{W}(\omega-\alpha)S_r(\theta^{-1}(\omega) + \theta^{-1}(\alpha-\omega))d\alpha, \quad (19)$$

where

$$\hat{W}(\omega) = |\Lambda_0^T(\omega)|^2 W(\theta^{-1}(\omega)) \quad (20)$$

is a warped version of the input frame window. If $W(\omega)$ is narrow band then also $\hat{W}(\omega)$ is narrow band and the largest contributions to the integral are for $\alpha \approx \omega$. Hence

$$Q_r(\omega) \approx \frac{1}{2\pi} \int_{-\pi}^{+\pi} \hat{W}(\omega-\alpha)\Lambda_0^T(\alpha)S_r(\theta^{-1}(\alpha))d\alpha. \quad (21)$$

In this approximation

$$q_r(n) \approx \hat{w}(n)\hat{s}_r(n), \quad (22)$$

where

$$\hat{s}_r(n) = \text{IDIFT}[\Lambda_0^T(\omega)S_r(\theta^{-1}(\omega))](n). \quad (23)$$

Comparing equation (22) with (10), we can identify $h(n)$ with $\hat{w}(n)$. If $W(\omega)$ is narrow-band then the map $\theta^{-1}(\omega)$ can be linearized in a neighborhood of $\omega = 0$ to obtain

$$H(\omega) \approx |\Lambda_0^T(0)|^2 W\left(\frac{d\theta^{-1}}{d\omega}\bigg|_{\omega=0} \omega\right) = \frac{1-b}{1+b} W\left(\frac{1-b}{1+b} \omega\right). \quad (24)$$

In this case, the window $h(n)$ is approximately a shrunk or dilated version of the input frame window $w(n)$ by a factor

$$\beta = \frac{1-b}{1+b}. \quad (25)$$

In principle, the window $h(n)$ has infinite length. However, following our considerations in (9), the window has a finite effective length. Furthermore, for narrow-band lowpass windows, the effective length is approximately equal to the delay in $\omega = 0$ times the length of the input window, i.e., to

$$P = \text{round}(\beta N). \quad (26)$$

This result is reported in Fig. 2 where the difference error between a stretched version of the Hanning window

$$h_p(n) = \frac{1}{2} \text{rect}_P(n) \left(1 - \cos\left(\frac{2\pi n}{P}\right)\right) \quad (27)$$

and its warped version is displayed. One can show that among the class of windows satisfying (11) the Hanning window is optimal for our frequency warping technique. Intuitively, this is explained by the fact that, except for finite-length effects, this window contains just a constant term and a sinusoidal one. Warping preserves the zero frequency, while the frequency of the sinusoid is altered, resulting in pure stretching.

Concerning the term $\hat{s}_r(n)$ in (22), notice that

$$\hat{s}_r(n) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \Lambda_0^T(\omega) S(\theta^{-1}(\omega)) e^{j(n\omega + rI\theta^{-1}(\omega))} d\omega, \quad (28)$$

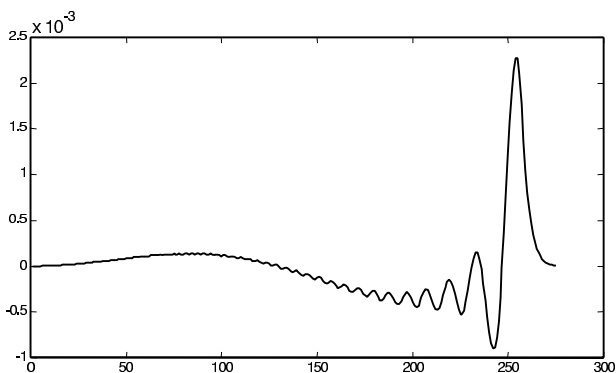


Figure 2. Difference error between warped and scaled 256-samples Hanning window for a warping parameter $b = -\frac{1}{3}$.

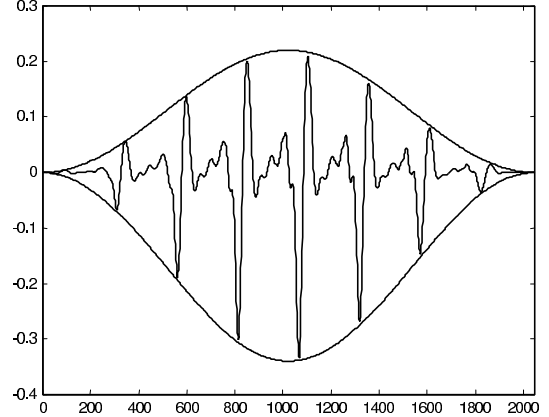


Figure 3. Warped frame of trumpet sound with superimposed stretched Hanning window.

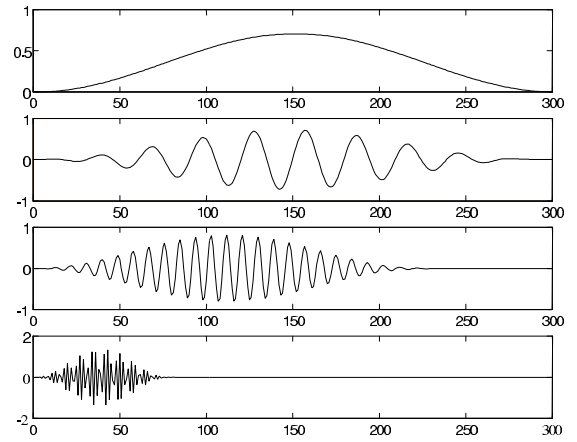


Figure 4. Short-time warping of sinusoids: frequency-dependent effects on the effective window length.

while for the term $\hat{s}(n + rM)$ in (10) we have:

$$\hat{s}(n + rM) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \Lambda_0^T(\omega) S(\theta^{-1}(\omega)) e^{j(n+rM)\omega} d\omega. \quad (29)$$

At low frequencies, these two terms are equal if we let, in the same approximation as in (24),

$$M = \text{round}(\beta L). \quad (30)$$

Within these approximations, equation (5) represents an alternate way of computing the Laguerre transform by overlap-add of the STLT (1). The overall behavior can be estimated from Fig. 3, where a warped trumpet frame and the stretched Hanning window are shown. Actually, from a perceptual point of view, this new form of frequency warping can be preferred over the pure Laguerre transform since in the latter the frequency dependent term in $I\theta^{-1}(\omega)$ introduces a strong time spreading of the signal components pertaining to different areas of the frequency spectrum. On long signals, this effect destroys the time organization of the sound event. In the overlap-add STLT method

this term is replaced by a linear phase term. In this way one can frequency warp the short-time frequency spectrum while preserving the long-term evolution of the signal.

Each output frame $q_r(n)$ is approximately equal to a warped version of the input frame modulated by a scaled version of the window. Property (11) ensures that the shape of the amplitude envelope of the signal is preserved, except for a natural dilation or shrinkage produced by warping. The effective length of the output frame depends on the spectral region occupied by the input signal. This is shown in Fig. 4, where, the length of the output frame varies according to the frequency of the sinewave. This effect can cause unbalancing in the amplitude of the different frequency components. Its extent can be controlled by the window length and frame spacing, which are parameters of the algorithm.

3. OVERLAP-ADD TECHNIQUES FOR DYNAMIC FREQUENCY WARPING

In this paper we approach the problem of generalizing the overlap-add STLT algorithm presented in the previous section to time-varying frequency warping. Two main issues are considered. The first one concerns the frame boundary transition when the Laguerre parameter is constant within each frame but different in adjoining frames. The other issue concerns the design of the window and the corresponding output overlap parameter when the Laguerre parameter is varying within each frame. In this case the overlap parameter is selected in terms of the average window-stretching factor.

3.1. Piecewise dynamic frequency warping

The overlap-add frequency warping technique is easily generalized to the case where the warping parameter is changing in time but remains constant within each input frame. Due to the different window stretching factors, overlapping introduces amplitude distortion, particularly relevant on abrupt Laguerre parameter changes. Property (11) is no longer verified. In order to circumvent this difficulty one can set constant output frame spacing and modify the input frame spacing accordingly. In practical applications, the range of the warping parameters is small so that variable size frame buffering does not constitute a problem. Often, as in vibrato, the Laguerre parameters vary on a zero-mean oscillatory curve. In this case, only a slow frame buffer length modulation is needed. If the output frame spacing is fixed to M samples, the time-varying input spacing parameter L_r is obtained by reversing equation (30):

$$L_r = \text{round}\left(\frac{M}{\beta_r}\right), \quad (31)$$

where, as usual, the integer r denotes the frame index and

$$\beta_r = \frac{1-b_r}{1+b_r}. \quad (32)$$

Since the warping parameter is frame dependent, the input frame window $w_r(n)$ depends on the frame index r as well. In our overlap-add frequency-warping algorithm, the output window can be chosen as a Hanning window of fixed length and the window

$w_r(n)$ can be computed by unwarping $h(n)$ with parameter b_r . Alternately, unwarped versions of the output window in the range of the warping parameters can be precomputed and stored in a lookup table to avoid computing an extra inverse Laguerre transform on the fly. As shown in the previous section, these windows can be approximated by a stretched version of the Hanning window, which can be simply generated by means of a sinusoidal oscillator.

The piecewise dynamic frequency-warping algorithm is suitable for applications where the rate of variation of the warping parameter is slow. It can be effectively employed for modulating the pitch of the original sound, to edit phrases and/or to add embellishments.

3.2. Dynamic frequency warping

In order to increase the flexibility of the overlap-add technique for frequency warping one can use the methods illustrated in conjunction with the time-varying version of the Laguerre transform [8,11]. Time-varying warping is obtained by means of a biorthogonal transform in which the signal is analyzed by orthogonal projection over a set of sequences $\Psi_n(k)$ whose z-transform are

$$\Psi_n(z) = \begin{cases} 1 & n=0 \\ \frac{1}{z^{-1}-b_1} & n=0 \\ \frac{z^{-1}(1-b_n b_{n+1})}{(1-b_n z^{-1})(1-b_{n+1} z^{-1})} \Phi_{n-1}(z) & n>0 \end{cases}, \quad (33)$$

where

$$\Phi_n(z) = \begin{cases} 1 & \text{if } n=0 \\ \prod_{k=1}^n \frac{z^{-1}-b_k}{1-b_k z^{-1}} & \text{if } n>0 \end{cases}. \quad (34)$$

The warped signal is the sequence of the analysis coefficients

$$\hat{s}(n) = \sum_k s(k) \Psi_n(k)$$

$$s(k) = \sum_n s(n) \varphi_n(k)$$

$\varphi_n(k)$ are the inverse z-transforms of (34). $\Phi_n(z)$ is the transfer function of a chain of allpass filters with overall frequency response

$$\Phi_n(\omega) = e^{-j \sum_{k=1}^n \theta_k(\omega)}, \quad (37)$$

where each $\theta_k(\omega)$ has the form (8) with parameter b_k changing from section to section. Consequently, (35) and (36) form a pair of time-varying warping and unwarping transformations generalizing the Laguerre transform. One can show that the functions $\Psi_n(z)$ satisfy the following recurrence:

$$\Psi_n(z) = T_n(z) \Psi_{n-1}(z), \quad n \geq 1 \quad (38)$$

where

$$T_n(z) = \frac{1 - b_n b_{n+1} z^{-1} - b_{n-1}}{1 - b_{n-1} b_n - b_{n+1} z^{-1}} \quad (39)$$

and $b_0 = 0$. Hence, the structure for computing time-varying frequency warping reported in Fig. 5 can be deduced, with simple modifications, from that of Fig. 1.

Time-varying warping shares the computational problems with the Laguerre transform. In order to be able to compute the transform in real time we have to resort to an overlap-add method based on a generalization of the STLT. The degree of controllability of the time-varying warping algorithm based on the biorthogonal transform (35) and (36) is however much higher than that of the version illustrated in section 3.1 since here the parameter b can vary on each sample. This makes it possible to use rapidly varying parameter modulation sequences or even white noise in order to edit effects such as Flatterzunge in flute or trumpet sounds or *al ponticello* expressions in string sounds. Furthermore, the use of mixed parallel time-varying warping sections plus the original signal allows for interesting extensions of flanging or phasing effects with full parameter control.

The major difficulty in extending our overlap-add STLT technique to time-varying warping lies in the fact that the parameters are allowed to change within the input frame length. As illustrated in section 3.1, one can fix both the output frame length P and the output frame spacing M . The parameter sequence b_n is organized in frames of length P with the same overlap as the output frames. The frame-dependent input window $w_r(n)$ is obtained by unwarping the unique output window $h(n)$ by means of the inverse time-varying warping formula (36). Clearly, the sets $\varphi_n(k)$ and $\psi_n(k)$ can be interchanged as analysis or synthesis bases since they satisfy the following biorthogonality conditions:

$$\sum_k \varphi_k(n) \psi_k(n') = \delta_{n,n'}$$

$$\sum_k \varphi_n(k) \psi_n'(k) = \delta_{n,n'}$$

$w_r(n)$ may be heavily altered. The effective length N_r of the window can be estimated from the length P of the output frame window $h(n)$ by inverting (26) and taking as scaling parameter β_r the average of the zero-frequency delay terms

$$d_k = \frac{1 - b_k}{1 + b_k}, \quad (42)$$

where k ranges over the indices of the parameters used to compute the output frame r , i.e.,

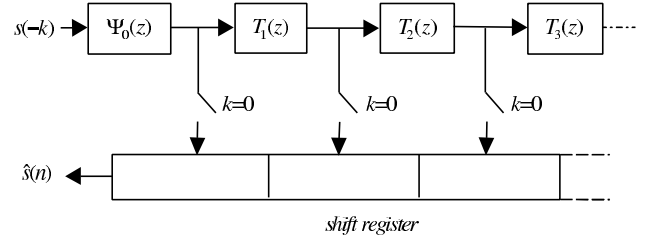


Figure 5. Structure for computing time-varying frequency warping.

$$\beta_r = \langle d_k \rangle_r, \quad (43)$$

where the symbol $\langle \bullet \rangle_r$ denotes average over the r -th frame. The input frame-spacing integer L_r is computed following the same reasoning as in (31), except that β_r is replaced by its new definition (43).

The overlap-add time-varying frequency warping algorithm can be summarized as follows:

- Initialize the output frame-length P and the output frame spacing M .
- Generate the normalized output window (Hanning).
- For any input frame of index r do the following:
 - ◊ Compute the length N_r unwrapped input window $w_r(n)$ using the inverse time-varying warping algorithm (36) on the current parameter set $\{b_k\}$ pertaining to the output frame r .
 - ◊ Multiply the current input frame by the window $w_r(n)$ and use the time-varying warping algorithm (35) on the current parameter set $\{b_k\}$ to compute the output frame.
 - ◊ Overlap-add with the previous output frames.
 - ◊ Advance both input and output frame pointers.

In order to assess the quality of the overlap-add time-varying frequency warping algorithm we compared the sounds obtained by this method with those edited by applying the pure biorthogonal transform (35) on a variety of instrument sounds and controlling sequences. In most cases, the sounds obtained by the computationally appealing short-time method were not perceptually distinct from those obtained by applying the pure method, even in fast parameter transition conditions.

An example of applying vibrato to a trumpet sound is reported in Figs. 6 and 7. In that case we used a sinewave amplitude modulated by a ramp to generate the control parameter sequence, thus inserting vibrato with increasing depth.

4. CONCLUSIONS

In this paper we presented new algorithms for time-varying frequency warping sound signals. We focused on the real-time approximation of previously presented off-line techniques. The key idea in developing these algorithms was the introduction of a Short-Time Laguerre Transform and its time-varying version in the form of a biorthogonal basis determined by the control parameters b_k .

Although frequency warping a finite-length signal yields an infinite-length signal, we showed that the essential length of the warped signal is bounded. Furthermore, in good approximation the output signal is equal the product of a warped version of the input window times the warped signal.

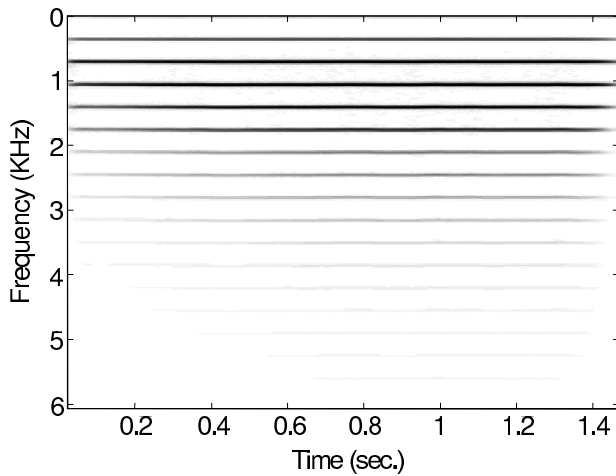


Figure 6. Spectrogram of trumpet tone.

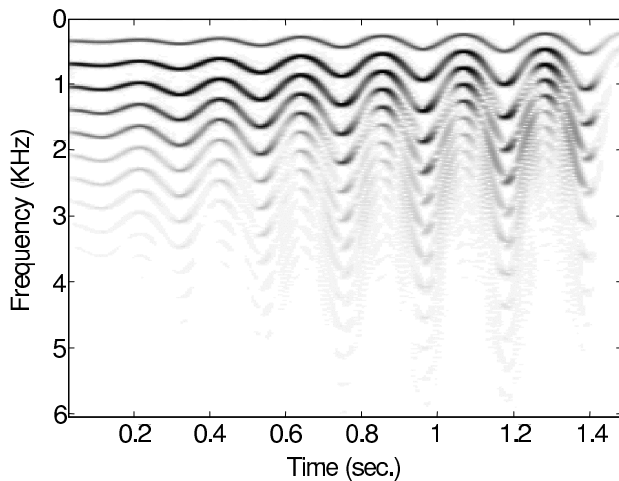


Figure 7. Spectrogram of trumpet tone with increasing depth vibrato obtained by overlap-add time-varying frequency warping.

We found relationships for transforming the input frame length and spacing parameters into the corresponding output quantities. We showed that well-behaved time-varying warping algorithms are obtained by fixing the output frame window and parameters and computing the input window by unwarping. Time-frequency warping proves to be an interesting editing tool for inserting or deleting expressive effects in musical sounds. The computationally efficient overlap-add algorithms presented in this paper allow for real-time implementation of these effects.

5. 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, "Analysis and Regularization of Inharmonic Sounds via Pitch-Synchronous Frequency Warped Wavelets," *Proc. of ICMC '97*, Thessaloniki, Greece, pp. 51-54, Sept. 1997.
- [8] G. Evangelista, S.Cavaliere, "Time-Varying Frequency Warping: Results and Experiments," *Proc. of DAFx99*, Trondheim, Norway, pp. 13-16, Dec. 1999.
- [9] G. Evangelista, S.Cavaliere, "Arbitrary Bandwidth Wavelet Sets," *Proc. ICASSP98*, vol. III, pp. 1801-1804, May 1998.
- [10] 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.
- [11] G. Evangelista, S. Cavaliere, "Audio Effects Based on Biorthogonal Time-Varying Frequency Warping," to appear on *Applied Signal Processing*, Springer-Verlag, Dec. 2000.
- [12] G. Evangelista, "The Short-Time Laguerre Transform: a New Method for Real-Time Frequency Warping of Sounds," *Proc. of ICMC 2000*, Berlin, Germany, pp. 380-383, Aug. 2000.