Real-time convolution and correlation using non-uniform spectral analysis and synthesis

John Garas[1] and *Piet C.W. Sommen*[2]

Eindhoven University of Technology
Building E-Hoog 6.34
P.O.Box 513, 5600 MB Eindhoven, The Netherlands.

tel. +31 40 2475742 fax. +31 40 2455674
email: j.garas@ele.tue.nl[1] p.c.w.sommen@ele.tue.nl[2]

Abstract

In many real-time audio applications, convolution and correlation operations are performed in frequency domain for efficiency reasons. The FFT algorithm is used to transform the time signals back and forth to frequency domain. The FFT decomposes a short-time signal into equally spaced frequency components. For audio signals, this tends to give too little spectral resolution at low frequencies, and too much spectral resolution at high frequencies; therefore a multi-resolution transform is preferred. This paper introduces a new non-uniform transformation method, that is based on the non-uniform sampling theorem and the scaling property of the Fourier transform. The speed of the new method is comparable to that of the FFT and in some cases is even better. The modification of the overlap-save method to accommodate the non-uniform decomposition is also shown.

1 INTRODUCTION

The human auditory system performs a non-uniform spectral analysis on sound waves collected by the pinnae. Constant-bandwidth analysis is used for frequencies below about 1kHz, while constant-Q (percentage bandwidth) analysis for frequencies above. Therefore, a non-uniform spectral resolution resembling that of the cochlea is preferred in audio applications. Excessive high-frequency resolution, resulting from using constant bandwidth analysis methods like the FFT, has several undesired effects. It can reveal spectral fine structure that is undetectable by the human ear, and it reduces system robustness to spatial variations in some applications such as phantom sound sources and room equalization. In this paper, we introduce a new method to perform such a non-uniform analysis which is suitable for real-time filtering and correlation applications.

2 NONUNIFORM SPECTRAL ANALYSIS AND SYNTHESIS

Recognising the importance of using nonuniform spectral analysis in audio applications, researchers began in the early seventies to look for a reliable method to perform such analysis. In subsection 2.1, we clearly define the requirements for a multi-resolution transform to be suitable for real-time convolution. A brief summary of the published work in this field is given in subsection 2.2. This shows that there is no known method that meets the specifications and our new method is introduced in subsection 2.3.

2.1 Specifications

To be of any use in real-time applications, the non-uniform transform has to satisfy the following requirements:

- As a general requirement, the transform should be a linear, unitary and invertible (nonsingular).
- It should perform multi-resolution analysis.
- It should support the convolution property.
- The discovery of the FFT algorithm made it possible to perform the convolution operation faster in frequency domain. The non-uniform transform has to maintain this speed advantage.

2.2 Existing Techniques

In 1971 Oppenheim et al. [1] introduced the digital frequency warping method. They showed that taking the DFT of a predistorted time sequence gives samples of the spectrum on a non-linear frequency scale. Although this method uses the speed of the FFT to transform the distorted sequence to the frequency domain, it requires $N(3M-1)$ real operations to calculate M distorted samples from N input time samples. This is proportional to N^2 for $M = N$ which makes it less attractive for real-time applications.

Several authors have published methods to perform non-uniform spectral analysis using block processing algorithms. Harris [2] processed the FFT output with spectral windows of constant time duration but adjustable bandwidths centered at the nearest FFT bin to the required analysis frequency. This method does not alter the time resolution of the signal and therefore information destructive and has no inverse transformation.

Another group of authors concentrated on calculating a constant-Q (proportional bandwidth) spectral analysis and synthesis. Youngberg and Boll [3] gave a constant-Q integral transform, which is a generalization of the Short Time Fourier Transform (STFT), by letting the analysis window be a variable in the product of time and frequency. Gambardella [4] showed that if a signal undergoes a short time spectral analysis via a continuous set of constant-Q bandpass filters, this process can be mathematically represented through an integral transform that can be inverted by means of the Mellin transform. Kates [5] used an exponentially decaying window whose argument is a constant times the product of time and frequency to calculate the constant-Q integral. For this specific window, the integral can be evaluated using the chirp z-transform. And finally, Brown [6] has developed a discrete version of the constant-Q integral transform but, unfortunately, this method has no inverse. Thus, all these methods can be viewed as a constant-Q filter bank and involve a high degree of complexity to be used in real-time filtering applications.

Mitra et al. [8] have recently introduced the Nonuniform Discrete Fourier Transform (NDFT). The NDFT evaluates the z-transform at arbitrary located distinct points in the complex z-plane. They showed that the NDFT is invertible and can be used in filtering operations. But since the NDFT lacks the symmetry of the Fourier transform, it can not be calculated by a fast method like the FFT and therefore requires an order of N^2 complex operations.

2.3 Time/Frequency Scaling (TFS)

In this subsection we introduce a new non-uniform spectral analysis and synthesis method that requires only an FFT and a resampler. The resampling (or interpolation) operation can be performed by hardware A/D and D/A converters, reducing the complexity to that of an FFT, which makes it very attractive for real-time applications. In cases where all signal processing is performed on blocks of data, the resampler is not needed and the complexity can be less than that of the FFT.

2.3.1 TFS Formulation

The new method, which we refer to as Time/Frequency Scaling (TFS), is based on Clark's one-dimensional non-uniform sampling theorem [7] and the scaling property of the Fourier transform. The non-uniform sampling theorem states that a band limited function $f(t)$ of one variable sampled at sampling moments t_n that are not necessarily equally spaced (fig. 1(a)) can be completely reconstructed from its samples provided that there exists a one-to-one continuous stretching/compressing transformation $\tau = \gamma(t)$ that maps $f(t)$ to $g(\tau)$ (fig. 1(b)) with sampling moments $nT = \gamma(t_n)$. The signal $g(\tau)$ is uniformly sampled on the time axis τ and Shannon sampling theorem holds. Therefore $f(t)$ can be reconstructed from its non-uniform samples $f(t_n)$ using the following formula

$$f(t) = \sum_{n=-\infty}^{\infty} f(t_n) \frac{sin\left[\frac{\pi}{T}(\gamma(t) - nT)\right]}{\left[\frac{\pi}{T}(\gamma(t) - nT)\right]} \quad (1)$$

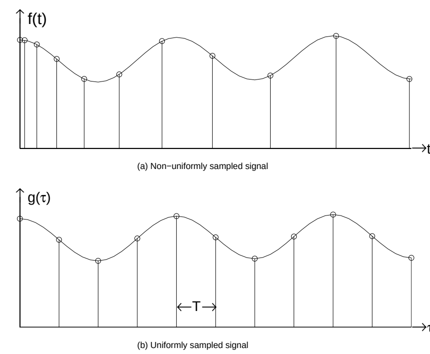

Figure 1: A function $f(t)$ sampled at sampling moments t_n not equally spaced can be transformed to a uniformly sampled function $g(\tau)$ by applying a compression/expansion mapping.

The relation between the spectra of $f(t)$ and $g(\tau)$ is given by the scaling property of the Fourier transform which is, for a constant scaler α, is given by

$$f(\alpha t) \rightleftharpoons \frac{1}{\alpha} F(\frac{\omega}{\alpha}) \quad (2)$$

It states that if the time axis is stretched (compressed) by a factor α, the frequency axis is compressed (stretched) by the same factor so that the product of time and frequency is always constant. Applying the scaling property to the non-uniform sampling case shows that the scaling factor γ is a function of time and thus, the frequency axis will be stretched/compressed nonlinearly. To see that this is true, let the frequency variable corresponding to t be ω and that corresponding to τ be Ω. Then $g(\tau)$ is given by

$$g(\tau) = \frac{1}{2\pi} \int_{-\frac{\pi}{T}}^{+\frac{\pi}{T}} G(\Omega)\, e^{j\Omega\tau}\, d\Omega \quad (3)$$

noting that $g(\tau) = g(\gamma(t)) = f(t)$, (3) can be written as

$$f(t) = \frac{1}{2\pi} \int_{-\frac{\pi}{T}}^{+\frac{\pi}{T}} G(\Omega)\, e^{\frac{j\Omega\tau t}{t}}\, d\Omega \tag{4}$$

substituting $\omega = \frac{\Omega\tau}{t}$, (4) can be written in the form

$$f(t) = \frac{1}{2\pi} \int_{(-\frac{\pi}{T})(\frac{\tau}{t})}^{(+\frac{\pi}{T})(\frac{\tau}{t})} \frac{1}{\tau/t}\, G\left(\frac{\omega}{\tau/t}\right) e^{j\omega t}\, d\omega \tag{5}$$

which clearly shows that the analysis frequencies of the non-uniformly sampled signal is scaled by a factor that is dependent on the ratio between the new and old time scales. Therefore, if this ratio is a non-linear function, as it would normally be the case, the frequency axis will be warped non-linearly.
From the time/frequency symmetry of the Fourier transform, it can be shown that the above theorem also holds for sampling spectra. This suggests that a uniformly sampled spectrum of a *time limited* function (the output of an FFT say) can be mapped to a non-uniformly sampled one by simply taking the Fourier transform of a non-uniformly sampled version of the time function. If the original spectrum is sampled at frequencies $\Omega_k = \frac{2\pi k}{NT} = k\Delta$, where N is the length of the corresponding uniformly sampled time sequence of sampling period T, and the required sampling frequencies are $\omega_k = \beta^{-1}(k\Delta)$, then the new sampling moments are $t_n = nT\frac{k\Delta}{\omega_k}$. The non-uniform sampling theorem ensures that no aliasing occurs, and the original signal can be recovered by first taking the inverse FFT, and then resampling from the non-uniform to the uniform domain.

2.3.2 Example

To demonstrate the non-uniform properties of the TFS method, consider the logarithmic mapping function $\gamma(\cdot) = log_b(\cdot)$ where b is any suitable base. The non-uniform sampling moments are $t_n = \gamma^{-1}(nT) = b^{nT}$ and the corresponding uniform moments $\tau_n = log_b(t_n) = nT$. Taking the Discrete Fourier Transform (DFT) of the non-uniformly sampled signal is then equivalent to evaluating the DFT at analysis frequencies $\omega_k = \Omega_k \frac{nT}{t_n}$. This mapping function is applied to a time signal composed of four sine waves as shown in fig. 2 with $b = e$ and frequency components 400, 600, 18000 and 18200 Hz. The time signal is first sampled exponentially at sampling moments $t_n = e^{nT}$ with $T = 1/44100$. The FFT of 1024 samples is shown in fig. 2(b). The time signal is then resampled uniformly in the same time period as in the exponential case (thus obtaining more samples). The FFT of this linearly sampled time signal is shown in fig. 2(c). Comparing the spectra in fig. 2 (b) and (c) shows that the resolution in case (b) decreases as the frequency increases while the resolution is constant in case (c).

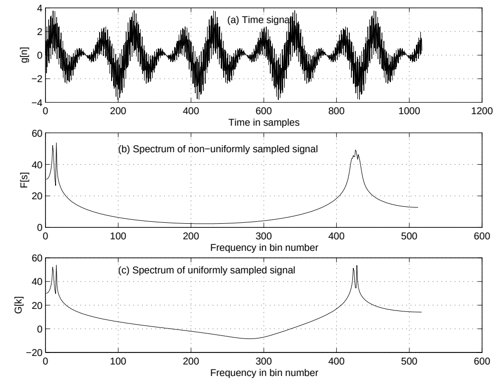

Figure 2: Spectra of a time signal sampled exponentially (b) and uniformly (c).

2.3.3 Hardware Resampler

The resampling operation can be performed using any of the well known interpolation algorithms. In this case, the transform complexity depends on the complexity of the interpolation algorithm in use; which ranges from $O(N)$ for cubic interpolation to $O(N^2)$ for spline algorithms. Simulations have shown that good results can be obtained using cubic interpolation when the signal is oversampled by a factor four. A more attractive solution is using a hardware resampler consisting of an A/D converter connected back to back with an D/A converter. To transform a signal from a uniform sampling domain to a non-uniform one an D/A converter clocked at the uniform sampling moments is first used to transform the signal to the continuous domain which is then resampled by an A/D converter clocked at the non-uniform sampling moments. Similarly, to transform a signal from the non-uniform sampling domain to the uniform domain, the D/A is clocked at the non-uniform sampling moments while the A/D is clocked at the uniform sampling moments as given by (1).
In cases where all processing is performed on blocks of data, and it is required to perform non-uniform spectral analysis in the whole system, it is wiser to sample the signals from the continuous time domain directly to the non-uniform domain. This has to be performed such that the non-uniform sampling pattern is repeated every block. In this case the resampling operation is not required and the system complexity can be less than that of a uniform sampled system since less number of samples are needed each block when the mapping function γ is a time compressing function such as that in subsection 2.3.2.

3 FAST CONVOLUTION AND CORRELATION USING THE TFS NON-UNIFORM SPECTRAL ANALYSIS

Convolution and correlation of long sequences are best performed in frequency domain; where they map to element-wise multiplications. When one or both signals to be convolved or correlated is infinitely long, as in the case of real-time (adaptive) filtering, an overlap method, such as overlap-add or overlap-save [9] has to be used. In this section the modifications to accommodate the TFS non-uniform spectral analysis in the overlap-save method are presented. Similar modifications can be applied to the overlap-add procedure. Fig. 3 shows the block diagram of the

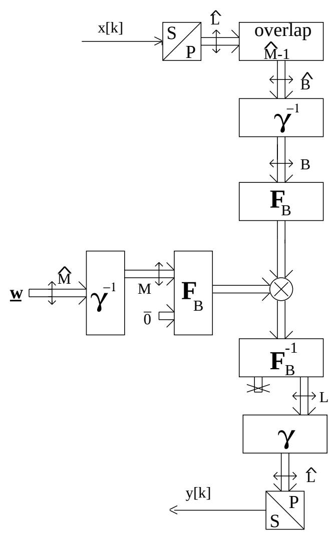

Figure 3: Block diagram of the overlap save method employing the TFS non-uniform spectral analysis.

overlap-save employing the TFS for an infinitely long input signal $x[k]$ and a finite length filter coefficients $\underline{\mathbf{w}}$. In this figure, $\widehat{M}$ is the filter length, $\widehat{L}$ is the number of new input samples per block and $\widehat{B} = \widehat{M}+\widehat{L}-1$ is the block length in the uniform sampling domain. M, L and $B = M + L - 1$ are the corresponding variables in the non-uniform sampling domain. The uniformly sampled signal $x[k]$ is stored in an input buffer of length $\widehat{B}$ from which $\widehat{L}$ samples are new and the rest $(\widehat{M} - 1)$ samples are taken from the previous block. Every $\widehat{L}$ samples, the input buffer is resampled (block γ^{-1}) producing a non-uniformly sampled sequence of length B. This length B vector is transformed to frequency domain using an FFT algorithm (block $\mathbf{F}_B$). The $\widehat{M}$ filter coefficients are also resampled, padded with zeros and transformed to frequency domain as shown by the left part of fig. 3. The convolution operation is then performed by element-wise multiplication of the two frequency vectors. In case of correlation, the complex conjugate of one of these vectors is taken before multiplication. The product is transformed back to time domain, and the first $M - 1$ samples are discarded since they represent cyclic convolution result. The last L samples are resampled back to the uniform sampling domain giving $\widehat{L}$ samples (block γ). This vector is finally unbuffered (block P/S) and sent out one sample every sampling period giving the linear convolution (correlation) result $y[k]$.

4 CONCLUSIONS

The Time/Frequency Scaling method (TFS) for performing non-uniform spectral analysis and synthesis proved to be effective, fast and simple. TFS is based on the non-uniform sampling theorem and the scaling property of the Fourier transform. It only requires an FFT and a resampler. The resampler can be implemented in hardware, thus reducing the complexity to that of the FFT. It is also shown to be readily incorporated in real-time frequency domain filters.

References

[1] Alan V. Oppenheim, Don Johnson and Kenneth Steiglitz, *Computation of Spectra with Unequal Resolution Using the Fast Fourier Transform*, Proc. IEEE, Vol. 59, pp. 299-301, Feb. 1971.

[2] Fredric J. Harris, *High-resolution Spectral Analysis with Arbitrary Spectral Centers and Arbitrary Spectral Resolutions*, J. Comput. & Elect. Eng. Vol. 3, pp. 171-191, 1976.

[3] James E. Youngberg and Steven F. Boll, *Constant-Q Signal Analysis and Synthesis*, Proc. ICASSP, pp. 375-378 April 1978.

[4] G. Gambardella, *The Mellin Transforms and Constant-Q Spectral Analysis*, J. Acoust. Soc. of Am., Vol. 66, No. 3, pp. 913-915, Sept. 1979.

[5] James M. Kates, *Constant-Q Analysis Using the Chirp Z-Transform*, Proc. ICASSP, CH-1379, pp. 314-317, 1978.

[6] Judith C. Brown, *Calculation of a Constant Q spectral Transform*, J. Acoust. Soc. of Am., Vol. 89, No. 1, pp. 425-434, Jan. 1991.

[7] James J. Clark, Matthew R. Palmer, and Peter D. Lawrence, *A Transformation Method for the Reconstruction of Functions from Nonuniformly Spaced Samples*, IEEE trans. Vol. ASSP-33, No. 4, pp. 1151-1165, Oct. 1985.

[8] S. K. Mitra, S. Chakrabarti and E. Abreu, *Nonuniform Discrete Fourier Transform and Its Application in Signal Processing*, Proc. EUSIPCO '92, Vol.2, pp. 909-912, Aug. 1992.

[9] Alan V. Oppenheim and Ronald W. Schafer, *Discrete-time Signal Processing*, Prentice-Hall, Englewood Cliffs, 1989, ISBN 0-13-216-771-9.