DIGITAL COMPRESSION OF ANALYTIC SIGNALS USING Nth ROOT

M. Bellanger and K. Ouaissa
CNAM - Electronique
292, rue Saint-Martin
75141 Paris cedex 03 - France
Tel: 331 40272590; fax: +331 40272779
e-mail: bellang@cnam.fr

ABSTRACT

The computation of the Nth root produces a compression of the spectrum for some types of analytic signals. Once the spectrum has been shrinked in the frequency domain, the signal can be decimated in the time domain, leading to compression. The Nth root calculation raises some problems which are discussed, like phase jumps related with zero crossings of the amplitude. Results of simulations with sinusoids and speech signals are reported. More work is needed to make the approach practical.

1 INTRODUCTION

Many existing techniques for the compression of signal are based on linear operations, like linear prediction or the Discrete Cosine Transform (DCT). These techniques have led to significant advances, but, also, they have shown their limitations. Now, to achieve further progress, it is necessary to consider non-linear approaches.

The present paper is an attempt in that direction and its aim is to draw the attention on a non-linear method which was considered more than three decades ago for the compression of speech with analog means [1, 2]. It performed a frequency compression by a factor N through the computation of the Nth root of the analytic signal and was successfully demonstrated in telephone transmission experiments. However, at that time, digital techniques were arriving and the analog compression method was abandoned. Now, in view of the progress in signal processing and, particularly, multirate techniques, it might be worthwhile to reconsider the frequency compression scheme and its implementation with digital means, and assess its potential, as well as its inherent difficulties.

The paper is organized as follows. The principle of frequency compression is presented in the next section. An implementation scheme is described in section 3 and the operations are discussed. Section 4 gives simulation results obtained with a sum of sinusoids and a speech sequence as input signals. The merits of the approach are discussed in the last section, as well as possible tracks for future work.

2 FREQUENCY COMPRESSION

Let us consider the following analytic signal

$$x(t) = a(t)e^{j\phi(t)} \quad (1)$$

where a(t) is the amplitude and $\phi(t)$ is the phase. In its spectrum, the signal has no component in the negative frequency domain [3].

The following Nth root

$$x^{\frac{1}{N}}(t) = a^{\frac{1}{N}}(t)e^{j\frac{\phi(t)}{N}} \quad (2)$$

is a signal whose instantaneous frequency, namely $d\phi/dt$, has been divided by N. In general, the instantaneous frequency is different from the signal spectrum. However, for some signals it can be equal or closely related to the spectrum. In those cases and in a digital context, one can figure out that the sampling rate can be reduced by the factor N, which leads to bit rate reduction and signal compression.

Obviously, the concept is fully valid for a single sinusoid. Assuming $x(t) = e^{j\omega t}$, the following Nth root

$$x^{\frac{1}{N}}(t) = e^{j\frac{\omega t}{N}} \quad (3)$$

has the frequency ω/N. However, if the signal is sampled, the roots of the sequence x(n) have to be computed and there are N determinations for the phase at each time n. The determination to choose at a given time n is the one which leads to the smallest variations, because the difference in phase between two consecutive samples must be ω/N.

For more complicated signals, the selection of the desired determination for the phase is more difficult for two reasons : the amplitude may become zero and the instantaneous frequency may become negative. For example, the amplitude of the signal

$$x(n) = e^{jn\omega_1} + e^{jn\omega_2} = 2cos(n\frac{\omega_1 - \omega_2}{2})e^{jn\frac{\omega_1+\omega_2}{2}} \quad (4)$$

crosses the zero axis, which results in discontinuities of the phase. Now, the amplitude of the signal

$$x(n) = e^{jn\omega_1} + \alpha e^{jn\omega_2}; \alpha < 1 \qquad (5)$$

is always greater than 1- α, while the instantaneous frequency can be negative, as shown in Figure 1.

Clearly, for an arbitrary analytic signal x(n), it is necessary to detect any zero crossing of the amplitude and add $\frac{\pi}{N}$ to the phase of the Nth root.

The detection of zero crossings is a classical technique, which finds applications in spectral estimation for example [4, 5]. It can be performed by more or less complicated algorithms. Here, just elementary approaches are considered.

Zero crossings of the amplitude can be detected from the analysis of the amplitude sequence itself and its derivatives or from the resulting jumps in the phase. Both approaches are considered below.

As concerns the analysis of the amplitude sequence a(n), it is preferable to consider the root $u(n) = a^{\frac{1}{N}}(n)$ because it has more contrast in its values. Now, the peak of the second order derivative of the sequence $u(n)$ provides the desired information. It is given by the classical raised cosine digital filter :

$$y(n) = u(n) - 2u(n-1) + u(n-2) \qquad (6)$$

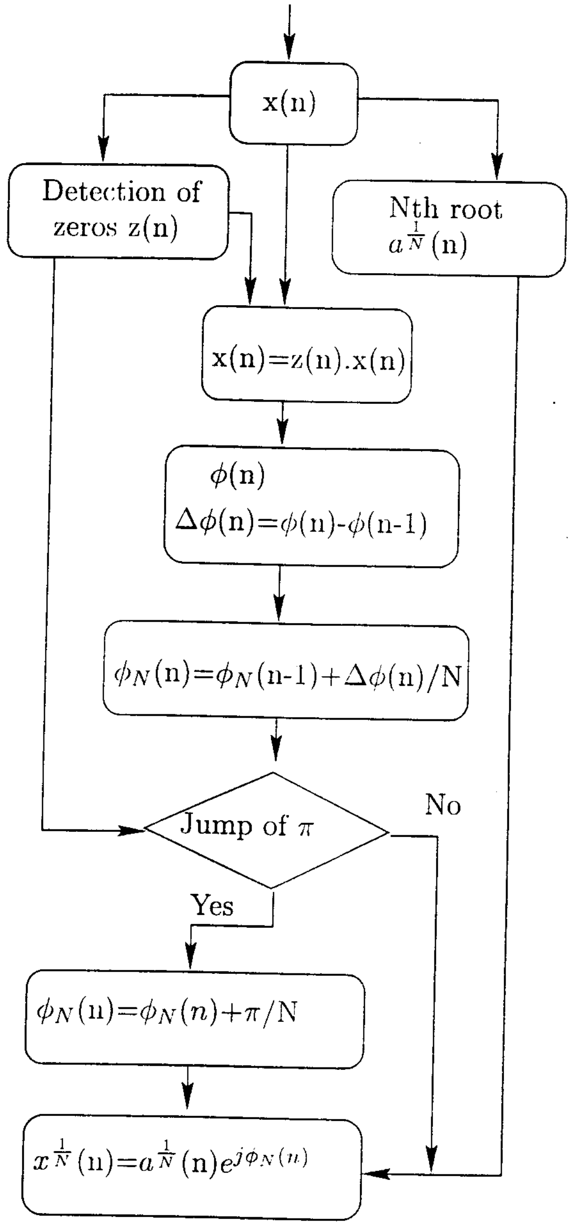
Various tricks can be employed to make the approach robust, particularly the introduction of a threshold. Here, the threshold is set to zero and negative values of y(n) are discarded ; then, the changes of sign of the derivative of y(n) are detected. The method works well with many signals, but it is not completely fault-proof.

The raised cosine digital filter can be applied to the phase ϕ(n) of the analytic signal x(n), producing a sequence e(n). However, in that case, the phase has to be unwrapped, in order to avoid the occurrence of jumps of 2π, and make sure that only the relevant jumps are detected from the signal e(n). Again, a threshold is introduced in the procedure to get rid of small and meaningless values of e(n). The control signal z(n) is a square signal whose edges coincide with the changes from plus sign to minus sign in e(n).

Both approaches have been tested on the two cisoid signal. It has been found that the method based on the phase jumps is more accurate.

3 COMPRESSION SCHEME

Starting from a real signal with sampling frequency f_s, the compression scheme is shown in Figure 2. The real input signal $x_r(n)$ is fed to an IQ filter which produces the complex signal $x(2n)$ at half the initial sampling rate. Then, the Nth root is calculated and a low-pass decimation filter produces the compressed sequence at the rate $\frac{f_s}{2N}$. Using the phase jumps to detect the zeros of the signal modulus, the operations involved in the computation of the Nth root are organized as follows.



The restitution of the initial signal consists of the reverse operations, namely interpolation, Nth power computation and complex-to-real conversion.

The specification of the complex decimation filter is a delicate operation. After the Nth root calculation, the signal spectrum is expected to be in the frequency band [0, $\frac{f_s}{2N}$]. Therefore, the filter can be computed as a real low-pass filter with cut-off frequency $\frac{f_s}{4N}$, its coefficients being afterwards multiplied by

$$e^{j2\pi\frac{n}{4N}} \qquad (7)$$

to produce a shift of $\frac{f_s}{4N}$ on the frequency axis. The stop band attenuation to be imposed depends on the signal power remaining above the frequency $\frac{f_s}{2N}$, which is susceptible to be aliased in the useful band.

4 SIMULATION RESULTS

In the simulations, the sampling frequency is 8000 Hz and a signal consisting of two sinewaves with frequencies 500 Hz and 800 Hz is considered first. The spectrum after the square root calculation is shown in Figure 3. After decimation with the factor 2, interpolation and square calculation, the spectrum of the reconstructed signal is shown in Figure 4.

The same scheme has been applied to the speech sequence shown in Figure 5. Most of the spectral energy of the analytic signal is situated below 600 Hz, as shown in Figure 6. After square root calculation, most of the energy is situated below 300 Hz as shown in Figure 7. The reconstructed signal is shown in Figure 8. It appears clearly that the voiced sections of the sequence are adequately reconstructed, while the non-voiced sections are highly distorted.

5 CONCLUSION

A non linear method has been presented to process and compress signals. In its principle, it is perfectly adapted to a single sinusoid in a spectrum. With multiple sinusoids, the Nth root calculation raises the problem of the zero crossings of the signal amplitude and the resulting jumps in the phase, which have to be taken into account. An accurate and reliable zero crossing detection method is essential and two simple approaches have been suggested.

The application to a speech sequence shows that, in the compression-reconstruction process, the voiced sections of the speech are reasonably well treated, while the non-voiced sections are severely distorted.

To make the proposed scheme practical more work is needed, on both theoretical aspects and implementation. The effect of Nth root calculation on the complex signal spectrum must be investigated further, at least for a number of typical families of signal, as well as the design of the decimation and interpolation filters. The effects of finite precision arithmetic and quantization after frequency compression must also be studied in details.

Depending on the results of this efforts, the Nth root technique might find application in the compression of some types of signals or in other areas like demodulation or carrier recovery in transmission or in instrumentation and measurement.

References

[1] P. MARCOU and J. DAGUET, "Une nouvelle méthode de transmission de la parole", *Annales des Télécommunications*, Vol. 11, N. 6, Juin 1956, pp. 118-126.

[2] M. R. SCHROEDER, "Vocoders : analysis and synthesis of speech", *Proceedings of the IEEE*, Vol. 54, N. 5, May 1966, pp. 720-734.

[3] B. Picinbono, "Représentation des signaux par amplitude et phase instantanées", *Annales des Télécommunications*, Vol. 38, N. 5-6, 1983, pp. 179-190.

[4] R. W. A. Scarr, "Zero crossings as a means of obtaining spectral information in speech analysis", *IEEE Transactions*, Vol. AU.16, N. 2, June 1968, pp. 247-255.

[5] V. Friedman, "A zero crossing algorithm for the estimation of the frequency of a single sinusoid in white noise", *IEEE Transactions on Signal Processing*, Vol. 42, N. 6, June 1994, pp. 1565-1569.

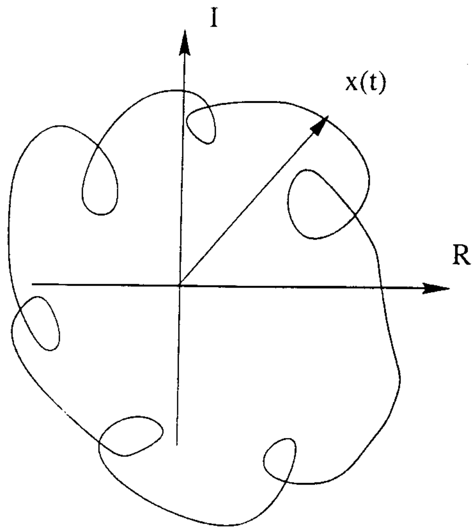


Figure 1: Analytic signal in the complex plane

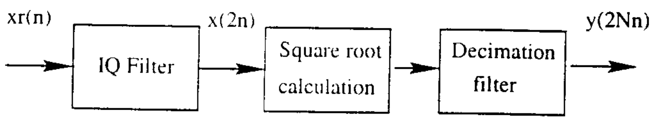


Figure 2: Compression scheme

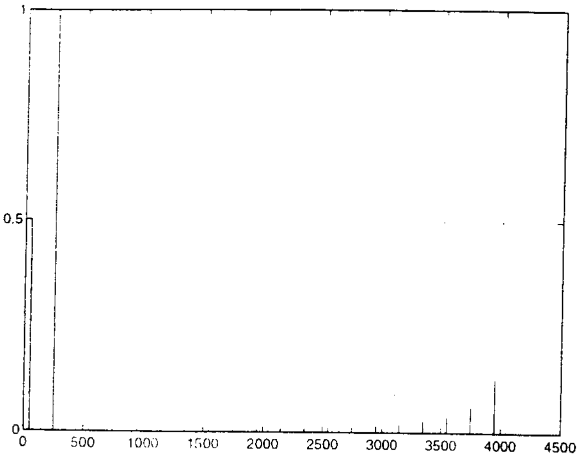


Figure 3: Spectrum of square root signal

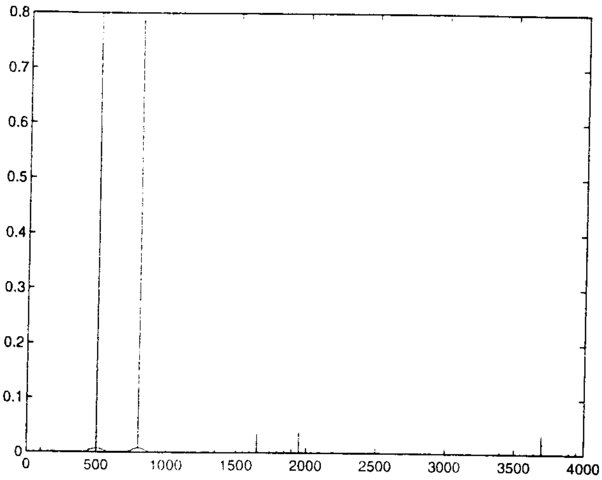


Figure 4: Spectrum of reconstructed signal

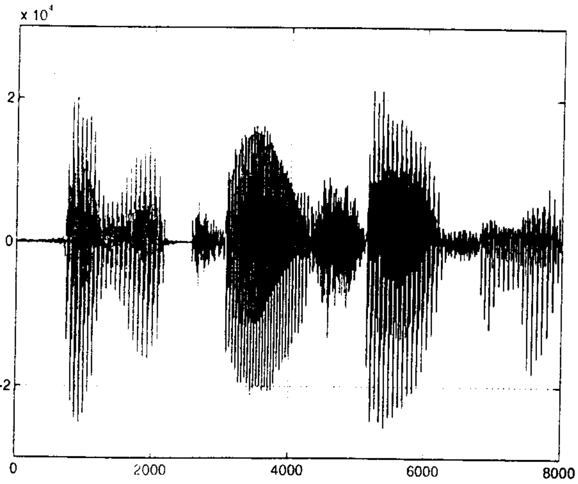


Figure 5: Original speech in the time domain

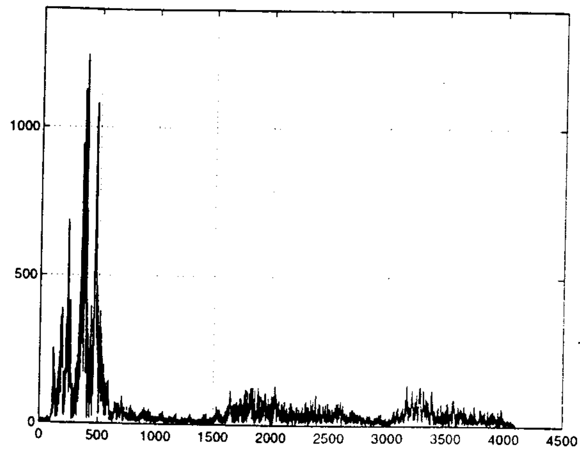


Figure 6: Spectrum of the complex signal

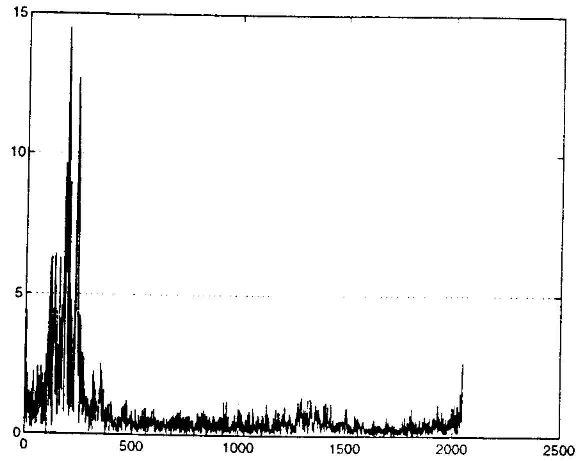


Figure 7: Spectrum of the square root signal

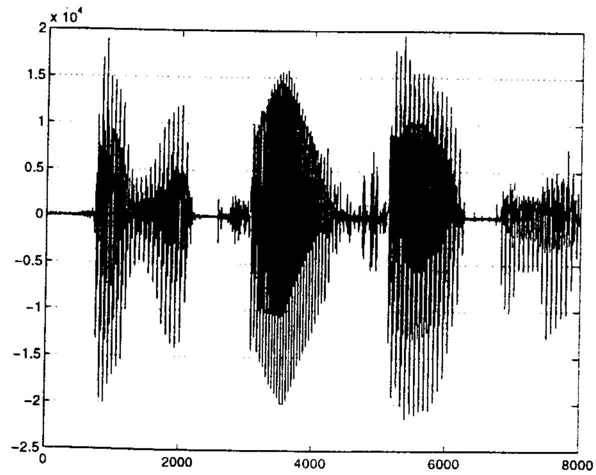


Figure 8: Reconstructed speech