SOUND PROCESSING USING COMPLEX DYNAMIC REPRESENTATION

Ewa Hermanowicz and Miroslaw Rojewski
Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology
ul. Narutowicza 11/12, 80-952 Gdansk, Poland
Tel. and fax: +48 58 3472870, E-mail: hewa@eti.pg.gda.pl

ABSTRACT

In this paper we show that a complex dynamic representation (logenvelope and instantaneous frequency) can be used efficiently to compress the spectrum and band-shift a discrete-time signal. The original method and algorithm conceived here do not violate the amplitude-phase relationships typical of natural sound signals. This is achieved by simultaneous modification via transmapping of the instantaneous amplitude and frequency of a signal.

1. INTRODUCTION

Let us consider a real-valued continuous-time arbitrarily modulated band limited band pass signal

$$x(t) = a(t)\cos\phi(t) = a(t)\cos[2\pi F_0 t + \varphi(t)] \qquad (1)$$

where $a(t)$ is the instantaneous amplitude (IA) of $x(t)$, $\phi(t)$ is the instantaneous phase, $\varphi(t)$ is the modulated component of $\phi(t)$ and F_0 is a carrier frequency in Hz. This signal can be otherwise expressed as

$$\begin{aligned} x(t) &= a(t)[\cos\varphi(t)\cos 2\pi F_0 t - \sin\varphi(t)\sin 2\pi F_0 t] \\ &= x_I(t)\cos 2\pi F_0 t - x_Q(t)\sin 2\pi F_0 t \end{aligned} \qquad (2)$$

where

$$x_I(t) \stackrel{\Delta}{=} a(t)\cos\varphi(t) \text{ and } x_Q(t) \stackrel{\Delta}{=} a(t)\sin\varphi(t) \qquad (3)$$

stand for the in-phase and quadrature components, respectively. It is further assumed that the quadratures $x_I(t)$ and $x_Q(t)$ are band limited to $|F| < F_0$ [1]. In (3) $\varphi(t)$ is the instantaneous phase of each of the quadrature components. In order to define the instantaneous frequency (IF) of the signal $x(t)$ one has to create a complex analytic signal (AS) [2], [3], otherwise called the Hilbertian equivalent of $x(t)$

$$x_H(t) \stackrel{\Delta}{=} x(t) + j\tilde{x}(t) \qquad (4)$$

whose real part is $x(t)$. The imaginary part of $x_H(t)$

$$\tilde{x}(t) \stackrel{\Delta}{=} H_H\{x(t)\} \qquad (5)$$

is the Hilbert transform of $x(t)$. The linear operator H_H in (5) is known under the name of Hilbert transformer (HT). Its frequency response is defined as

$$H_H(j2\pi F) = -j\,\mathrm{sgn}(F), \quad -\infty < F < \infty \qquad (6)$$

Due to band limitedness of the quadratures, the products: $x_I(t)\cos 2\pi F_0 t$ and $x_Q(t)\sin 2\pi F_0 t$ in (2), both have a low-frequency factor: $x_I(t)$ and $x_Q(t)$, respectively, and a high-frequency factor: $\cos 2\pi F_0 t$ and $\sin 2\pi F_0 t$, respectively. Consequently, on the grounds of the Bedrosian theorem [2], [3], the signal $\tilde{x}(t)$ in (4) can be written as

$$\tilde{x}(t) = a(t)\sin[2\pi F_0 t + \varphi(t)] \qquad (7)$$

and the complex AS, $x_H(t)$ (4), for $x(t)$ is given by

$$x_H(t) = u(t)\exp(j2\pi F_0 t) \qquad (8)$$

where

$$u(t) \stackrel{\Delta}{=} a(t)\exp[j\varphi(t)] = x_I(t) + jx_Q(t) \qquad (9)$$

is the complex envelope of $x_H(t)$. The complex envelope $u(t)$, which generally is not an AS, has the IA $a(t)$ and the instantaneous phase $\varphi(t)$ as polar components, and the in-phase component $x_I(t)$, and the quadrature component $x_Q(t)$ as Cartesian components.

The IF [5] of the AS (8) is defined as

$$F_{iH}(t) \stackrel{\Delta}{=} \frac{1}{2\pi}\frac{\mathrm{d}\phi(t)}{\mathrm{d}t} = F_0 + \frac{1}{2\pi}\frac{\mathrm{d}\varphi(t)}{\mathrm{d}t} \qquad (10)$$

where $\phi(t)$ is the instantaneous phase of the signal $x(t)$ and $\varphi(t)$ is the instantaneous phase of the complex envelope $u(t)$. The second term in (10) is the IF of $u(t)$, further denoted as $F_i(t)$. Thus

$$F_i(t) \stackrel{\Delta}{=} \frac{1}{2\pi}\frac{\mathrm{d}\varphi(t)}{\mathrm{d}t} \qquad (11)$$

where the instantaneous phase $\varphi(t)$ is defined as [1], [4]

$$\varphi(t) \stackrel{\Delta}{=} \mathrm{Im}\{\ln[u(t)]\} \qquad (12)$$

Then the IF (11) of the AS $x_H(t)$ can be rewritten shortly

$$F_{iH}(t) = F_0 + F_i(t) \qquad (13)$$

The fundamentals of the AS application to sound processing were formulated in [6]. In [7] and [8] examples of shifting the spectrum of an analytical audio signal were reported where the IA remained unaffected. Opposite to that the original method and algorithm conceived here do not violate the amplitude-phase relationships typical of natural signals (speech sounds, animal voices, etc.). This is achieved by simultaneous modification via transmapping of both: the IA and IF.

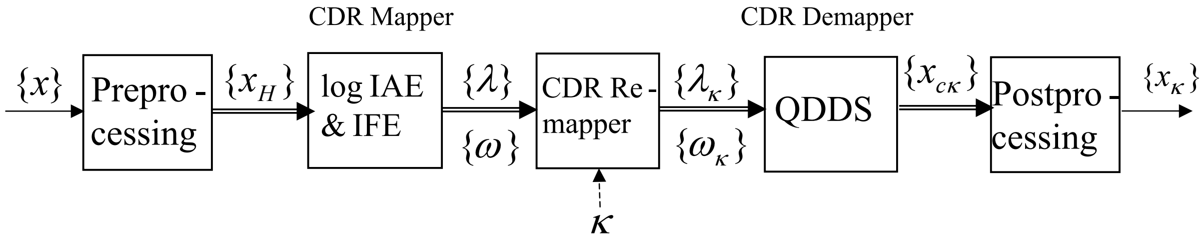

Fig.1. Block-scheme for the CDR processing with a factor of κ .

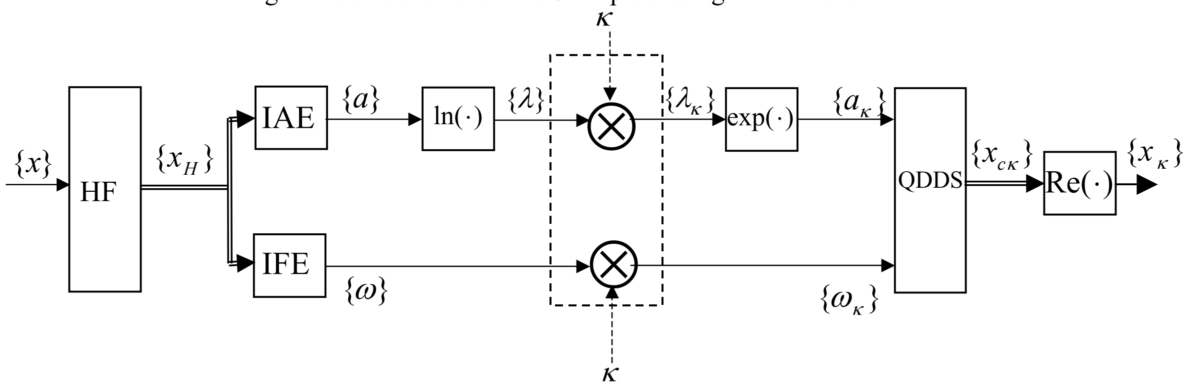

Fig.2. The CDR of the real signal record $\{x\}$ on-line processing - spectrum compression and band shifting with a factor κ .

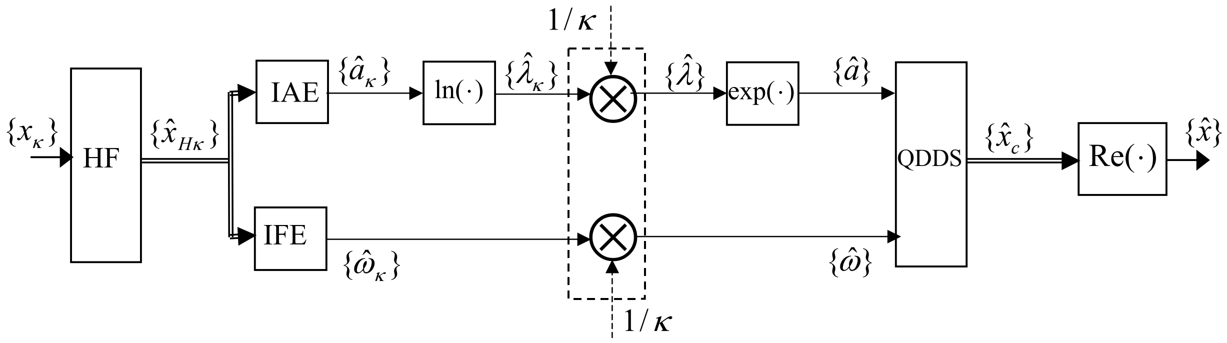

Fig.3. The CDR inverse processing of the signal record $\{x_\kappa\}$ - decompression and band deshifting with a factor $1/\kappa$.

2. COMPLEX DYNAMIC REPRESENTATION

Further on we call the pair:

$$\Lambda(t) \overset{\Delta}{=} \ln a(t) \text{ and } \Omega(t) \overset{\Delta}{=} 2\pi F_{iH}(t) \qquad (14)$$

the complex dynamic representation (CDR) of $x(t)$. The CDR has two real-valued components. The first of them, $\Lambda(t)$, is the logenvelope of the Hilbertian signal $x_H(t)$ and the second, $\Omega(t)$, is the angular IF of $x_H(t)$. The CDR components represent uniquely the AS corresponding to the real-valued signal $x(t)$ [2], [3].

In practice it is convenient to deal with IF estimators (IFE) in the discrete-time domain.

By sampling the signal $x(t)$ (1) with the sample period T we obtain a complex discrete-time signal

$$x[n] \overset{\Delta}{=} x(t)\big|_{t=nT}, \quad n = 0, \pm 1, ... \qquad (15)$$

where n is the time index of the samples. (For simplicity a unit-sampling rate has been adopted.) The CDR of the complex-valued Hilbertian signal $x_H[n] = a[n]\exp(j\phi[n])$ corresponding to the given real-valued signal $x[n] = a[n]\cos\phi[n]$ consists of the following two real-valued components:

$$\lambda[n] \overset{\Delta}{=} \ln a[n] \quad \text{and} \quad \omega[n] \overset{\Delta}{=} 2\pi F_{iH}[n] \qquad (16)$$

(cf. (13) and (10)). Having in hand the CDR (16) one can perform a number of simultaneous manipulations on the log-envelope $\lambda[n]$ and instantaneous angular frequency $\omega[n]$ of the discrete-time complex-valued Hilbertian signal $x_H[n]$ representing the real-valued signal $x[n]$ (15).

3. CDR PROCESSING

By the following remapping of the CDR of a given signal $x[n]$

$$\lambda[n], \omega[n] \Rightarrow \lambda_\kappa[n] = \kappa\lambda[n], \quad \omega_\kappa[n] = \kappa\omega[n] \qquad (17)$$

one obtains a new CDR having the components: $\lambda_\kappa[n]$ and

$\omega_\kappa[n]$. This remapping is aimed at simultaneous compression and band shifting with a factor $\kappa > 0$ of the real-valued signal $x[n]$. On this basis a new complex-valued signal $x_{c\kappa}[n]$ can be formed by using a quadrature direct digital synthesizer (QDDS) shown in Fig.1. The real-valued counterpart of this signal is readily obtained as

$$x_\kappa[n] = \mathrm{Re}(x_{c\kappa}[n]) \qquad (18)$$

We use the same parameter κ for remapping (or scaling) the logenvelope and the IF. This is because the dependence of the signal bandwidth on both these components: $\lambda[n]$ and $\omega[n]$, has the same character (see [5]). Hence κ can be treated as a pitch modification factor of sound signals. Also note that if $\kappa = 1/2^\rho$ with ρ a positive integer, the remapping results in pitch down shifting by ρ octaves.

In Fig. 1 the pre-processing and post processing consists in Hilbertian filtering (HF). The Hilbert filter has an ideal frequency response defined as

$$H(e^{j\omega}) \overset{\Delta}{=} \begin{cases} 2, & 0 < \omega < \pi \\ 0, & -\pi < \omega < 0 \end{cases} \qquad (19)$$

where $\omega = \Omega T = 2\pi FT$ stands for the normalised „digital" angular frequency in radians per sample. Further on, in Fig.1, IFE stands for the IF estimator.

Fig.2 gives a closer insight into the realization of our concept of simultaneous compression and band shifting with a factor of $\kappa \in (0,1)$ of the real-valued signal record $\{x\}$. Firstly, the input signal is filtered by the HF. Next, this signal is mapped into its CDR $\{\lambda, \omega\}$. The CDR components are extracted using the IAE, $\ln(\cdot)$ and IFE blocks. Further on both CDR components are multiplied by the same coefficient $\kappa \in (0,1)$. After this remapping, performed in accordance with (17), the new CDR: $\{\lambda_\kappa, \omega_\kappa\}$ is demapped into the complex signal representation $\{x_{c\kappa}\}$ using the QDDS. The target real-valued signal $\{x_\kappa\}$ shown at the output in Fig. 2 is simply the real part of $\{x_{c\kappa}\}$ (18). Fig.3 shows the scheme of inverse processing, with deremapping. It is aimed at verification of the CDR inverse processing.

4. EXPERIMENTS

In our experiments performed in the MATLAB environment we have used the following estimators of the CDR of $x[n]$:

$$\lambda[n] = \ln|x_H[n]| \qquad (20)$$

for the log-envelope (IAE and $\ln(\cdot)$ blocks) and

$$\omega[n] = Arg(x_H[n] x_H^*[n-1]) \qquad (21)$$

for the instantaneous angular frequency (IFE block), where the asterisk stands for the complex conjugate. The QDDS applied as the CDR demapper is shown in Fig.4. Figs. 5 and 6 present the results of on-line CDR processing of a recording of a canary song, by a cascade depicted in Figs. 2 and 3. The processing is aimed at pitch down shifting using a pitch modification factor κ=1/10. (Other experimental conditions are specified in these figures.) It results in simultaneous spectrum compression and down shifting, both by a factor of κ, as shown in Fig. 6.

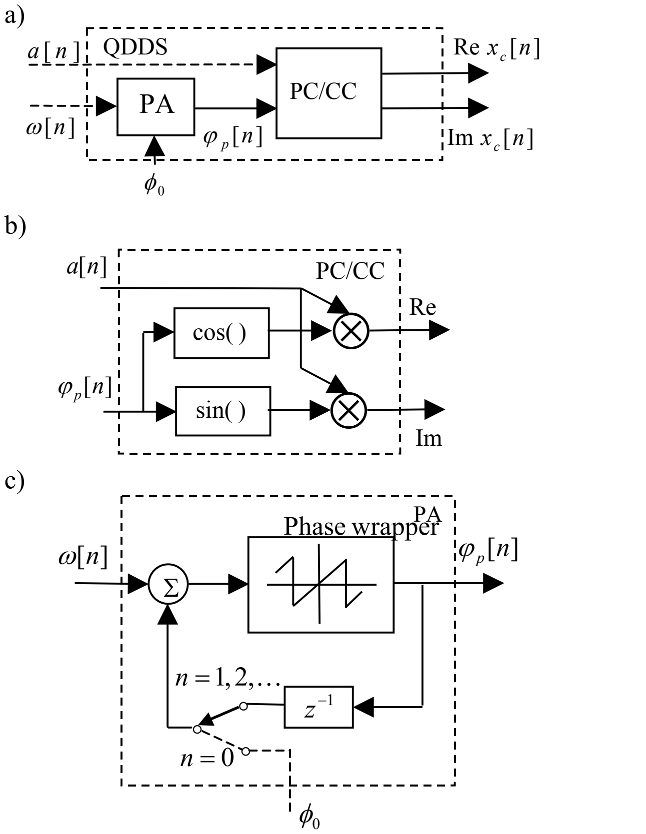

Fig. 4. The QDDS basic architecture (a) with functioning of internal blocks (b) and (c) revealed; PA stands for the phase accumulator, PC/CC is the polar to Cartesian coordinates converter.

5. CONCLUSIONS

The conclusions are the following.

1. The proposed CDR processing appears as a powerful means for pitch shifting of sound signals.
2. The quality of the CDR processing depends strongly on the quality of the Hilbertian filter. The role of this filter in Figs. 1 and 2 is twofold. Firstly, it creates an AS from a given real-valued signal. Secondly it serves as an antialiasing filter.
3. The CDR remapper introduced in this paper is capable of performing a variety of useful and invertible operations such as spectrum rotation and shifting, compression and expansion, and inverse, as well as signal level dynamic matching. All this can be achieved on-line by manipulating on the values of the log-envelope and IF remapping coefficients.
4. The CDR remapper as a pitch shifter can also serve for entertainment. It exhibits an excellent performance as a generator of different melodies of ring signals, where the sound of each note in a melody is derived from a short recording of a voice of an arbitrarily chosen creature via CDR processing, e.g. a nightingale or canary chirp.

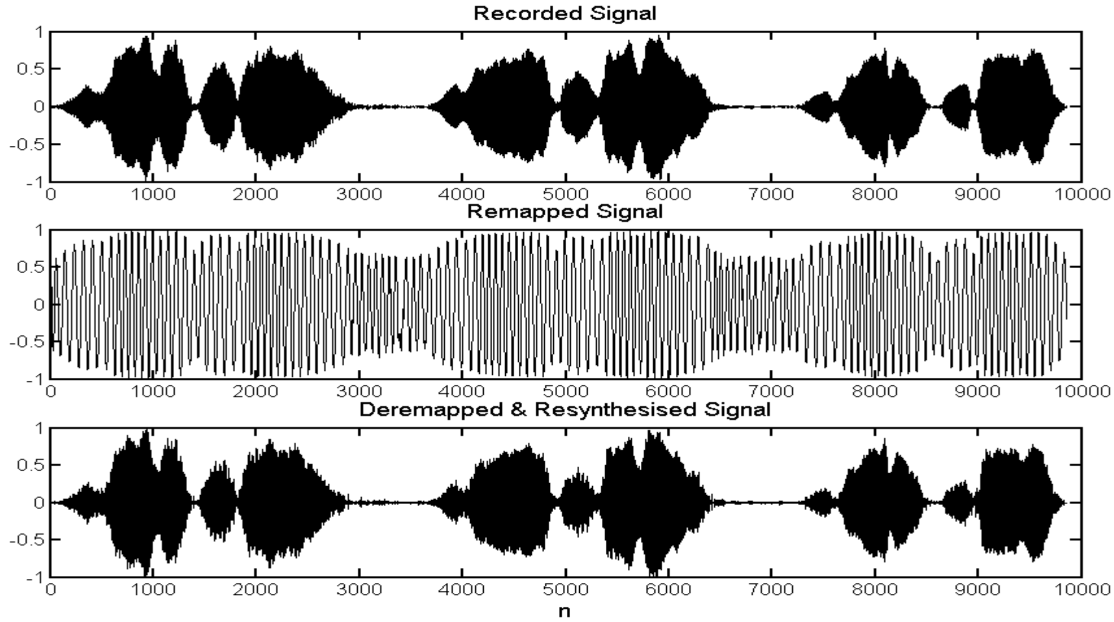

Fig.5. Exemplary waveforms of a canary song processed via CDR remapping aimed at pitch down shifting with κ =1/10.

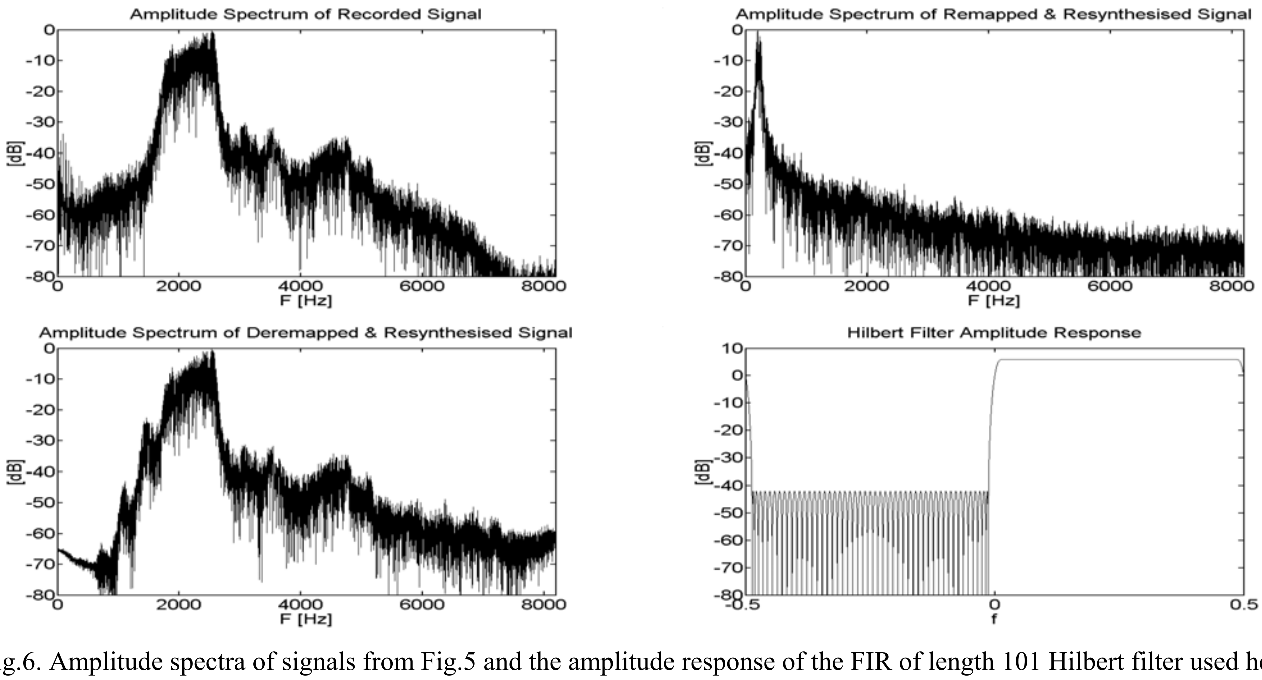

Fig.6. Amplitude spectra of signals from Fig.5 and the amplitude response of the FIR of length 101 Hilbert filter used here.

REFERENCES

[1] E. Hermanowicz, M. Rojewski, G.D. Cain and A. Tarczynski, "On an instantaneous frequency estimator with FIR filters having maximally flat frequency response error magnitude", *Signal Processing*, 2001, vol.81, pp.1491-1504.

[2] S.L. Hahn, *Hilbert Transforms in Signal Processing*, Artech House, 1996.

[3] D. Vakman, *Signals, Oscillations and Waves. A Modern Approach*, Artech House 1998.

[4] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal - Part 1: Fundamentals and Part 2: Algorithms and applications", *Proceedings of the IEEE*, Vol. 80, No. 4, April 1992, pp. 520-568.

[5] L. Cohen, *Time-frequency analysis*, Prentice Hall, 1995, Chapter 1, p. 16.

[6] J.H. Justice, "Analytic signal processing in music computation", *IEEE Trans., v*ol. ASSP-27, No. 6, 1979.

[7] S. Wardle, "A Hilbert transformer frequency shifter for audio", *First Workshop on Digital Audio Effects DAFx*, Barcelona, Spain, 1998.

[8] S. Disch and U. Zolzer, „Modulation and delay line based digital audio effects", *2nd Workshop on Digital Audio Effects DAFx*, Trondheim, Norway, 1999.