A Dual-Tree Rational-Dilation Complex Wavelet Transform

İlker Bayram and Ivan W. Selesnick

***Abstract*—In this correspondence, we introduce a dual-tree rational-dilation complex wavelet transform for oscillatory signal processing. Like the short-time Fourier transform and the dyadic dual-tree complex wavelet transform, the introduced transform employs quadrature pairs of time-frequency atoms which allow to work with the analytic signal. The introduced wavelet transform is a constant-Q transform, a property lacked by the short-time Fourier transform, which in turn makes the introduced transform more suitable for models that depend on scale. Also, the frequency resolution can be as high as desired, a property lacked by the dyadic dual-tree complex wavelet transform, which makes the introduced transform more suitable for processing oscillatory signals like speech, audio and various biomedical signals.**

***Index Terms*—Rational-dilation wavelet transform, dual-tree complex wavelet transform, short-time Fourier transform, analytic signal, instantaneous freuency estimation.**

I. Introduction

Despite its usefulness for piecewise smooth signals, the dyadic wavelet transform is less effective for the processing of oscillatory signals (like speech, audio, various biomedical signals) because of its poor frequency resolution. In this regard, the rational-dilation wavelet transform (RADWT), which provides a finer frequency analysis, is more suitable [1] (see Fig. 1). However, even though the RADWT has better frequency resolution, it is not simple to carry out certain operations with the RADWT, such as the Hilbert transform, envelope detection, instantaneous frequency estimation, which are relevant for oscillatory-signal processing [2]. For such tasks, the short-time Fourier transform (STFT) is more useful because its atoms are analytic, which in turn allow for the easy construction of analytic signals. However, the time-frequency structure of the STFT is quite different from that of a wavelet transform, it being a constant bandwidth transform rather than a constant-Q transform[1]. In this paper, we introduce a dual-tree rational-dilation wavelet transform (DT-RADWT) that inherits the good frequency resolution and constant-Q property of the RADWT[2] and whose atoms form quadrature pairs (see Fig. 1a,b).

In brief, a DT-RADWT frame consists of the union of a RADWT frame and its Hilbert transform[3]. The DT-RADWT introduced in this paper is realized by two filter banks (FB) operating in parallel on the input as shown in Fig. 2. Essentially, we require that the i^{th} channel of the two filter banks, say FB_1 and FB_2, to compute inner products of the input $x(n)$ with shifts of a bandpass filter $h_i(n)$ and $-\mathcal{H}_d\{h_i(n)\}$ respectively. Here, $\mathcal{H}_d\{\cdot\}$ denotes the *discrete-time* Hilbert transform [6]. For the derivation of the equivalent bandpass filters $h_i(n)$, see Section IV in [1].

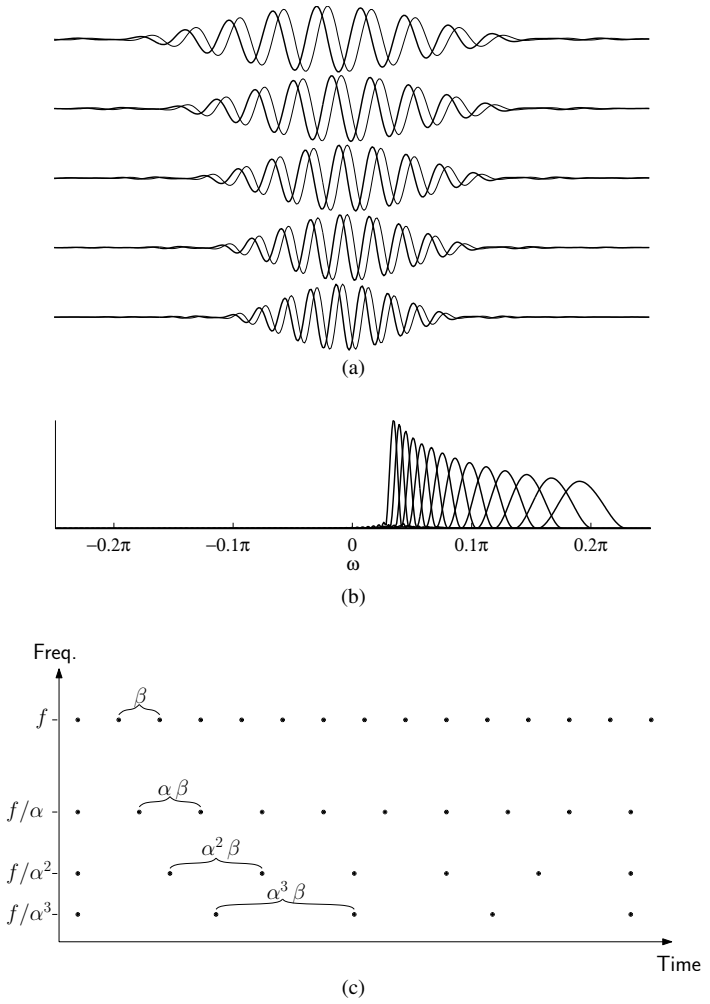

Fig. 1. (a) A number of atoms from different scales of a RADWT (thick lines). DT-RADWT employs these atoms as well as their Hilbert transforms (thin lines), (b) The frequency responses of a number of analytic (discrete-time) wavelets, (c) Sampling of the time-frequency plane by a typical wavelet frame. The dilation factor α, and the shift parameter β determine the lattice. For RADWT and DT-RADWT, α and β can be selected by changing the sampling factors of the filter bank.

The difference between the introduced DT-RADWT and the STFT is the time-frequency (TF) distribution of their atoms. The STFT samples the TF plane uniformly wheras the TF sampling pattern for the DT-RADWT is non-uniform along the frequency axis (see Fig. 1c). For ease of notation, let us consider signals defined on the real line for now. Given a window function $h(\cdot)$, STFT atoms are comprised of time shifts and modulates of this function, i.e.,

$$\text{STFT Atoms}: \ \{h(\cdot - n\Delta x)\, \exp(jk\Delta\omega\, \cdot)\}_{k,n\in\mathbb{Z}}\,. \tag{1}$$

A number of STFT atoms are shown in Fig. 3. We remark that, for a fixed window function $h(\cdot)$, the number of oscillations in the STFT atoms

İ. Bayram was with Biomedical Imaging Group, EPFL, Switzerland. He is now with the Department of Electronics and Communication Engineering, Istanbul Technical University, Maslak, 34469, Istanbul, Turkey.

I. W. Selesnick is with the Department of Electrical and Computer Engineering, Polytechnic Institute of New York University, Brooklyn, NY, 11201, USA.

E-mail: ilker.bayram@itu.edu.tr, selesi@poly.edu

[1]The Q-factor of a function is defined to be the ratio of its center frequency to its bandwidth (i.e. $f_c/\Delta f$). A transform is said to be constant-Q if it employs atoms that have the same Q-factor. See [3] and Section IV for an argument of the utility of constant-Q transforms in audio processing.

[2]The introduced DT-RADWT is also μ-shift invariant like the RADWT (see [4] for the definition of μ-shift invariance) thanks to the specific filter frequency responses described in [1] which are also used here.

[3]To be precise, the Hilbert relation holds only for the bandpass functions. The situation for lowpass functions is similar to that in dyadic DT-ℂWT – see [5].

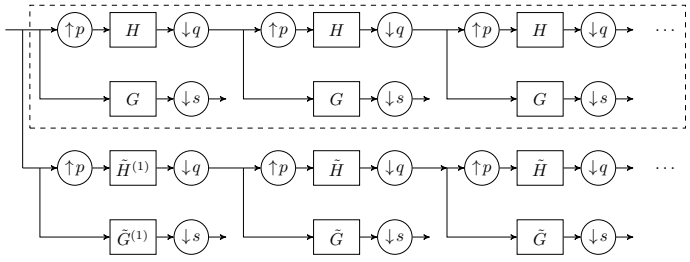

Fig. 2. DT-RADWT consists of two filter banks with rational sampling factors operating in parallel on the input. The system in the dashed box, which is referred to as the 'rational-dilation DWT' (RADWT), was discussed in [1] – here, the only requirement on p, q, s is that the FB have as many output coefficients as input coefficients, i.e. $p/q + 1/s \geq 1$.

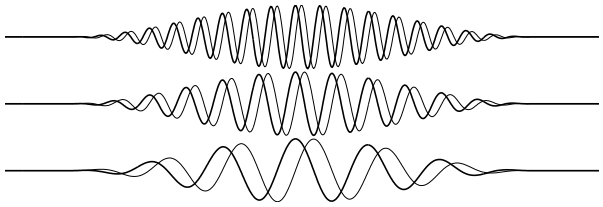

Fig. 3. (a) Real (thick lines) and imaginary parts (thin lines) of a few STFT atoms.

depend on the frequency parameter k in (1). This implies a non-uniform treatment of different *scales* – how succesfully STFT can separate two frequencies f_0 and f_1 depends not only on their ratio f_1/f_0 but their values as well. On the other hand, the introduced DT-RADWT employs time shifts and dilations of the real and imaginary part of an analytic wavelet function $\psi(\cdot)$, i.e.,

$$\text{DT-RADWT Atoms} : \left\{\alpha^{n/2}\,\mathrm{Re}\{\psi(\alpha^n\,\cdot - k\beta)\}\right\}_{k,n\in\mathbb{Z}} \cup \left\{\alpha^{n/2}\,\mathrm{Im}\{\psi(\alpha^n\,\cdot - k\beta)\}\right\}_{k,n\in\mathbb{Z}} \quad (2)$$

where α and β can be adjusted by changing the sampling factors of the FBs in Fig. 2. Because the frequency tiling is achieved by scaling the wavelet function, the number of oscillations (in fact the shape) of the atoms remains the same (see Fig. 1a). Likewise, these atoms all have the same Q-factor. The resulting time-frequency distribution with respect to these parameters is shown in Fig. 1c. The TF sampling patterns of the STFT and the DT-RADWT are different in mainly two aspects. First, the subbands are distributed logarithmically for the DT-RADWT, whereas they are uniform for the STFT. Second, the sampling period of each subband is different for DT-RADWT to account for the subband-dependent bandwidths. This way, the DT-RADWT attains a modest redundancy as well as a stable analysis/synthesis implementation, which is related to its being a tight frame. We note however that the STFT has a faster implementation than the DT-RADWT because it makes more efficient use of the FFT (due to its uniform sampling with respect to frequency).

Contribution

Our main contribution in this correspondence is the construction of an analytic rational dilation wavelet transform using the dual-tree framework. In particular, we provide a generalization of the sufficient condition derived in [7] (which was also shown to be necessary in [8]). The discussions are valid for the underlying wavelets (defined on the real line), as well as the discrete-time functions which play a more direct role in the computations.

Previous Work

The dual-tree complex wavelet transform (DT-$\mathbb{C}$WT) which employs a quadrature pair of dyadic wavelet frames (i.e. $\alpha = 2$ in (2)) was introduced by Kingsbury [9] (see also [5]). In comparison to DT-$\mathbb{C}$WT, the DT-RADWT is based on rational sampling factors (α, β in 2 and Fig.2) and can therefore attain higher frequency resolution than the DT-$\mathbb{C}$WT.

An interesting approach to devise a constant-Q transform was formulated by Gambardella [10]. His idea is to generalize the STFT by using window functions that depend on the center frequency of interest. As a special case, if one scales the window function with the center frequency, the resulting transform becomes,

$$I_f(\omega, t) = \int f(x)\, h\big(\omega\,(x-t)\big)\, e^{j\omega\, x}\, dx \quad (3)$$

where $f(\cdot)$ is the input signal and $h(\cdot)$ is the window function. If we set $\psi(x) = h(x)\, e^{j\,x}$, this can be written as,

$$I_f(\omega, t) = e^{j\,\omega t} \int f(x)\, \psi\big(\omega(x-t)\big)\, dx. \quad (4)$$

Upto a phase term (and scaling), (3) is equivalent to a continuous wavelet transform. This form of the transform was studied by Youngberg and Boll [11], Petersen and Boll [12]. In [12] the authors also show how to sample this constant-Q transform (i.e. sampling the parameters ω and t), deriving the pattern in Fig. 1c. This particular constant-Q transform was also utilized for time-scaling of audio [13], an application which we will discuss in Section IV. Brown [3] proposed a similar constant-Q transform by utilizing channel-dependent window functions and non-uniform sampling in the frequency domain (also see [14]).

Frequency-warping [15] (also see [16] for a comprehensive tutorial and further references) is another interesting approach to obtain nonuniform frequency analysis. Here, the idea is to transform the input signal $x(n)$ to another signal $g(n)$ so that their DTFTs satisfy

$$X(e^{j\omega}) = G(e^{j\theta(\omega)}) \quad (5)$$

where $\theta(\omega)$ is a one-to-one function on the unit circle. By appropriately selecting $\theta(\omega)$, FFT of $g(n)$ allows one to nonuniformly sample $X(e^{j\omega})$ in the frequency domain. In general, the mapping that transforms $x(n)$ to $g(n)$ is not orthonormal. Noting the relation to Laguerre sequences, Evangelista and Cavaliere [17] prefilter $x(n)$ so as to ensure that this mapping is orthonormal. Following this orthonormalized frequency-warping, they apply a dyadic DWT to the frequency-warped signal to obtain an orthonormal time-frequency analysis similar to that of a rational-dilation DWT.

Organization

In Section II, we discuss some basic operations that can be carried out naturally with the DT-RADWT. Following this, Section III provides a sufficient condition which ensures that the two filter banks employ quadrature pairs of discrete-time atoms. We will also show that the two filter banks are related to quadrature pairs of wavelets which can be combined to form an analytic wavelet. In Section IV, we discuss an application that requires both the analysis and synthesis capabilities of DT-RADWT to demonstrate the potential of the introduced transform. Section V is the conclusion.

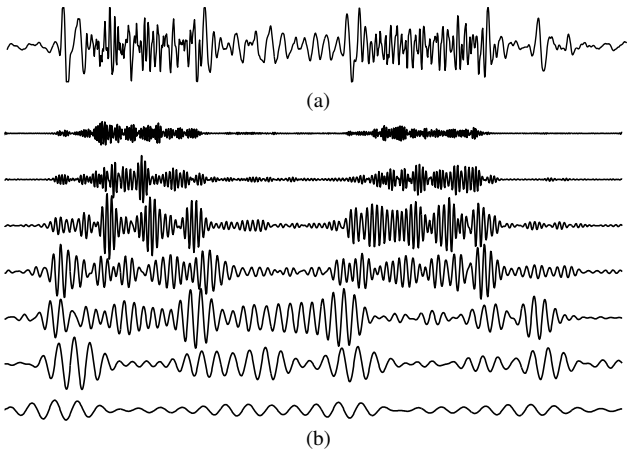

Fig. 4. Decomposition of a phonocardiogram (PCG) signal using the RADWT. (a) PCG signal, (b) Several channels of the RADWT of the PCG signal. (The RADWT parameters are $p = 2$, $q = 3$, $s = 2$ with reference to Fig. 2.)

II. Basic Operations with the DT-RADWT

Before going into the details of the construction, we first briefly discuss some basic operations for which the DT-RADWT is useful.

Consider the phonocardiogram (PCG) signal $x(n)$ shown in Fig. 4a. RADWT gives us a decomposition of $x(n)$ as

$$x(n) = \sum_{i=1}^{\#\text{Chn.}} \sum_{k\in\mathbb{Z}} c_i(k)\, h_i(n - k\, s_i) \tag{6}$$

where $h_i(n)$ is the equivalent filter employed by the i^{th} stage of RADWT, s_i is a channel-dependent shift parameter (see [1] for details) and

$$c_i(k) = \langle x(\cdot), h_i(\cdot - k\, s_i)\rangle \tag{7}$$

We can isolate the separate channels $x_i(n)$ as

$$x_i(n) = \sum_{k\in\mathbb{Z}} c_i(k)\, h_i(n - k\, s_i). \tag{8}$$

A number of $x_i(n)$'s are shown in Fig. 4b for a particular RADWT (with parameters $p = 2$, $q = 3$, $s = 2$ – we use the the filters described in Section II.B of [1]). The Hilbert transform of $x(n)$ can be computed by swapping the coefficients as illustrated in Fig. 5. The Hilbert transform of a single channel $x_i(n)$ can be similarly computed by setting to zero the coefficients of the rest of the channels. A particular channel and its Hilbert transform is shown in Fig. 6a. Now that we have $x_i(n)$ and $\mathcal{H}_d\{x_i(n)\}$, we can form the analytic signal for $x_i(n)$ via

$$x_i^A(n) = x_i(n) + j\,\mathcal{H}_d\{x_i(n)\}. \tag{9}$$

$x_i^A(n)$ can be used to compute the envelope as $\text{env}_i(n) = |x_i^A(n)|$ (see the dashed line in Fig. 6a). We can also estimate the instantaneous frequency of $x_i(n)$ by tracking the phase of $x_i^A(n)$. For the particular filters that we use in DT-RADWT, it can be shown that $x_i^A(n)$ is obtained by LTI filtering $x(n)$ with a bandpass filter $f_i^A(n)$. Since we know the center frequency of $f_i^A(n)$, we can write $x_i^A(n)$ as,

$$x_i^A(n) = \text{env}_i(n)\, \exp(j\,(\omega_i\, n + \phi(n))) \tag{10}$$

where ω_i denotes the center frequency of $f_i^A(n)$ and $\phi(n)$ is the cumulative phase due to the deviation of the instantaneous frequency of $x_i^A(n)$ from ω_i. From this expression, we take the instantaneous frequency as $\phi'(n) + \omega_i$ where $\phi'(n)$ is the principal value of $\phi(n+1) - \phi(n)$. We remark that $\angle x_i^A(n+1) - \angle x_i^A(n)$ could have been used to directly obtain the instantaneous frequency, but this can lead to phase-wrapping problems. Taking the principal value of $\phi(n+1) - \phi(n)$ does not guarantee to avoid phase-wrapping problems but it is less likely to encounter such problems. The instantaneous frequency obtained through $\phi'(n) + \omega_i$ is shown in Fig. 6b.

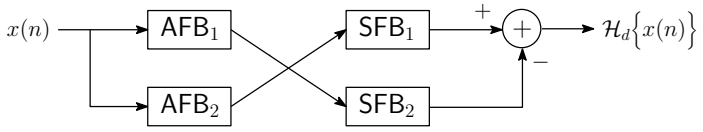

Fig. 5. Using DT-RADWT, the analytic signal transform of a signal $x(n)$ can be performed by swapping the outputs of the analysis filter banks and taking the difference. This follows because if $x(n) = \sum_i c_i f_i(n)$ for a collection of atoms $f_i(n)$, the Hilbert transform of $x(n)$ is $\sum_i c_i \mathcal{H}\{f_i(n)\}$. The analysis part in the figure compute c_i's for f_i, swapping in the synthesis stage replaces f_i by $\mathcal{H}\{f_i\}$. In the figure, AFB and SFB stand for analysis filter bank and synthesis filter bank respectively.

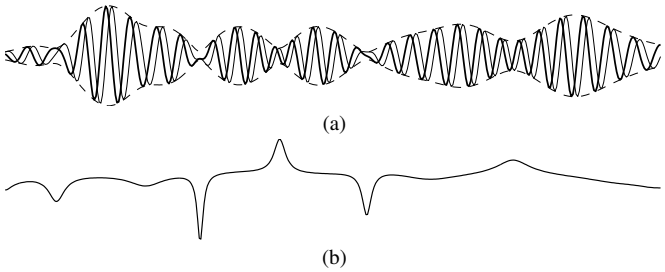

Fig. 6. (a) A particular channel of the PCG in Fig. 4a (thick line), its Hilbert transform (thin line) and the envelope signal (dashed line), (b) instantaneous frequency of the signal in (a).

In brief, we think that the DT-RADWT will be useful in situations where one needs to perform time-frequency operations on logarithmically distributed frequency subbands.

III. Dual-Tree RADWT : A Sufficient Phase Condition

The RADWT consists of an iterated FB as shown in Fig. 2 (see the system inside the dashed rectangle). For the i^{th} bandpass channel, the analysis FB essentially computes the inner products of the input with shifts of a discrete-time bandpass function $h_i(n)$. We will construct another FB, with the same sampling factors, such that $\tilde{h}_i(n)$, the function whose shifts form the i^{th} bandpass channel of this new iterated FB, satisfies $\tilde{h}_i(n) = -\mathcal{H}_d\{h_i(n)\}$.

We remark that both of the iterated FBs are associated with unique wavelet functions $\psi(x)$, $\tilde{\psi}(x)$ [1] (this is not the case in general – see [18] for a detailed discussion) similarly as in dyadic wavelet frames. We will also show that $\tilde{\psi}(x) = \mathcal{H}\{\psi(x)\}$, where $\mathcal{H}\{\cdot\}$ denotes the Hilbert transform for $L_2(\mathbb{R})$.

In the rest of this section, we provide sufficient phase conditions on the filters for realizing a DT-RADWT. For this, let us first recall the (sufficient) perfect reconstruction conditions for the RADWT (for a more detailed treatment, we refer to [1]).

Perfect Reconstruction Conditions for RADWT

The RADWT is self-inverting (or forms a tight frame) if the analysis and synthesis FBs shown in Fig. 7 have the perfect reconstruction (PR) property. A set of sufficient conditions on the filters $H(\omega)$ and $G(\omega)$ ensuring PR can be derived[4] as (notice that $H(\omega)$, $G(\omega)$ are 2π-periodic

[4]For general sampling factors, FIR filters cannot yield a RADWT that is also a tight frame [1].

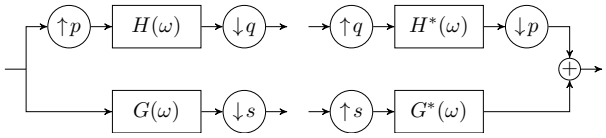

Fig. 7. Rational-dilation wavelet transform is a tight frame provided this filter bank has the perfect reconstruction property.

Fig. 8. The perfect reconstruction conditions given in (11) imply certain restrictions on the frequency responses of $H(\omega)$ and $G(\omega)$. In addition to a band limit, there is a constant frequency response band.

functions of ω),

$$H(\omega) = 0, \quad \text{for } |\omega| \in \left[\frac{\pi}{q}, \pi\right], \tag{11a}$$

$$G(\omega) = 0, \quad \text{for } |\omega| \in \left[0, \left(1 - \frac{1}{s}\right)\pi\right], \tag{11b}$$

$$\frac{1}{pq}\left|H\left(\frac{\omega}{p}\right)\right|^2 + \frac{1}{s}\left|G(\omega)\right|^2 = 1 \quad \text{for } \omega \in [-\pi, \pi]. \tag{11c}$$

These conditions imply a transition band as well as a constant-response band (shown upto a possible phase factor) for $H(\omega)$ and $G(\omega)$ as depicted in Fig. 8.

Phase Condition for the Second FB

The RADWT essentially computes the inner products of the input with a set of discrete-time functions $\{h_{i,k}\}_{i\in I, k\in\mathbb{Z}}$ indexed by subband i and position k. We will construct an additional FB using filters $\tilde{H}$, $\tilde{G}$ such that the resulting set of functions $\{\tilde{h}_{i,k}\}_{i\in I, k\in\mathbb{Z}}$ are the discrete-time Hilbert transforms of $\{-h_{i,k}\}_{i\in I, k\in\mathbb{Z}}$, save for the functions in the lowpass subband.

To derive a sufficient condition, we consider the system shown in Fig. 9, that upsamples, filters, downsamples the input and then computes the inner product of the output with a discrete-time function $u(n)$. The system in Fig. 9 maps the input discrete-time function to a number. Since it is also linear, it can be regarded as a correlator with a discrete-time function $y(n)$. Here, the correlator with $u(n)$ may be regarded as a device that maps the input signal to a particular sample of one of the bandpass channels. In other words, we think of $u(n)$ as one of $h_{i,k_0}(n)$ described above (i.e. the discrete-time function at subband i, position k_0). Given this, $y(n)$ will be equal to $h_{i+1,k_1}(n)$ (i.e. the discrete-time function at subband $i+1$, position k_1). We will first investigate how $u(n)$ and $y(n)$ are related. Next, we will replace $u(n)$ and $H(\omega)$ with their counterparts in the second FB, denoted by $\tilde{u}(n)$, $\tilde{H}(\omega)$. This new system will allow us to compare the functions in the second FB, namely $\tilde{h}_{i,k_0}$ (that is, $\tilde{u}(n)$) and $\tilde{h}_{i+1,k_1}$ (that is, $\tilde{y}(n)$) with the functions in the first FB, namely h_{i,k_0} and h_{i+1,k_1}. In particular, our intent is to show that, if h_{i,k_0} and $\tilde{h}_{i,k_0}$ are related by a discrete-Hilbert transform, then h_{i+1,k_1} and $\tilde{h}_{i+1,k_1}$ are also related by a discrete-Hilbert transform, thereby implying the Hilbert relation between the two FBs, by an induction argument.

Let us return to the system in Fig. 9. We remark that $y(n)$ is implicitly determined by p, q, $H(\omega)$ and $u(n)$. We now take a look at this relation.

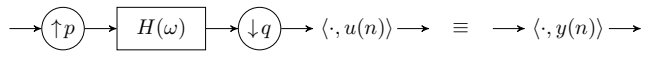

Fig. 9. Upsampling, filtering, downsampling followed by correlator. Under (11)a, this is equivalent to computing the inner product of the input with $y(n)$, whose DTFT, $Y(\omega)$, is given in (14).

We take H to be bandlimited to π/q and also $p < q$. In this setting, the output of this system for a given input signal $x(n)$ is,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{q}X\left(\frac{p}{q}\omega\right)H\left(\frac{\omega}{q}\right)U^*(\omega)\,d\omega. \tag{12}$$

Using the bandlimitedness of H (the expression is not correct otherwise), we can change variables to write this as,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{p}X(\omega)H\left(\frac{\omega}{p}\right)U^*\left(\frac{q}{p}\omega\right)d\omega. \tag{13}$$

Therefore, we see that

$$Y(\omega) = \frac{1}{p}H^*\left(\frac{\omega}{p}\right)U\left(\frac{q}{p}\omega\right) \qquad \text{for } \omega \in [-\pi, \pi]. \tag{14}$$

Now consider another system, that utilizes a different filter $\tilde{H}$ and a different function $\tilde{u}(n)$. Suppose further that

$$\tilde{U}(\omega) = U(\omega)\,e^{-j\Theta(\omega)} \qquad \text{for } \omega \in [-\pi, \pi], \tag{15}$$

where

$$\Theta(\omega) = -\operatorname{sign}(\omega)\frac{\pi}{2} + \frac{\omega}{a} \qquad \text{for } \omega \in [-\pi, \pi]. \tag{16}$$

We will try to find $\tilde{H}(\omega)$ such that $\tilde{Y}(\omega) = p^{-1}\tilde{H}^*(\omega/p)\tilde{U}(q\omega/p)$, satisfies

$$\tilde{Y}(\omega) = Y(\omega)\,e^{-j\Theta(\omega)} \qquad \text{for } \omega \in [-\pi, \pi]. \tag{17}$$

This will imply, by induction, the Hilbert relation we are after – we will make this claim clearer below.

In an attempt to satisfy (17), take

$$\tilde{H}(\omega) = H(\omega)\,e^{-j\tau(\omega)} \qquad \text{for } \omega \in [-\pi, \pi], \tag{18}$$

with

$$\tau(\omega) = \frac{\omega}{b} \qquad \text{for } \omega \in [-\pi, \pi]. \tag{19}$$

Noting that $\tilde{Y}(\omega) = Y(\omega) = 0$ for $|\omega| \in [p\pi/q, \pi]$, we have

$$\begin{aligned}\tilde{Y}(\omega) &= \frac{1}{p}\tilde{H}^*\left(\frac{\omega}{p}\right)\tilde{U}\left(\frac{q}{p}\omega\right) \qquad \text{for } \omega \in [-\pi, \pi] \\ &= Y(\omega)e^{-j\,[\Theta(q\omega/p) - \tau(\omega/p)]} \quad \text{for } \omega \in [-\pi, \pi].\end{aligned} \tag{20}$$

Therefore $\tilde{Y}(\omega) = Y(\omega)\,e^{-j\Theta(\omega)}$ holds if

$$\Theta(q\omega/p) - \tau(\omega/p) = \Theta(\omega), \tag{21}$$

i.e., if

$$-\operatorname{sign}(\omega)\frac{\pi}{2} + \frac{q\,\omega}{p\,a} - \frac{\omega}{p\,b} = -\operatorname{sign}(\omega)\frac{\pi}{2} + \frac{\omega}{a}. \tag{22}$$

This is equivalent to

$$(q - p)\,b = a. \tag{23}$$

Therefore, considering the FB structure in Fig. 2, if we set the filters in the second FB as $\tilde{G}(\omega) = G(\omega)\,e^{j\Theta(\omega)}$, $\tilde{H}(\omega) = H(\omega)\,e^{-j\tau(\omega)}$, where $\Theta(\omega)$, $\tau(\omega)$ are defined by (16), (19) and $(q-p)\,b = a$, we see, by an induction argument, that the bandpass functions of the iterated FB constructed by $H(\omega)$, $G(\omega)$, are discrete-time Hilbert transforms of the bandpass functions of the iterated FB constructed by $\tilde{H}(\omega)$, $\tilde{G}(\omega)$ upto a shift by $-1/a$. We can remove this shift by delaying the lowpass filter

in the first stage of the second FB by p/a, i.e. by replacing $\tilde{H}(\omega)$ by (see the second FB in Fig. 2)

$$\tilde{H}^{(1)}(\omega) := \tilde{H}(\omega)e^{-j\,p\,\omega/a} = H(\omega)\,e^{-j\omega(1/b+p/a)}. \tag{24}$$

If we further ask that $1/a + 1/(pb) = q/(2p)$, so as to distribute the shifts of the scaling function evenly (see [5] for an argument), we find that (recall $a = (q-p)b$),

$$a = 2, \quad b = \frac{2}{q-p}. \tag{25}$$

Given H, G, this specifies the phase terms of $\tilde{H}$ and $\tilde{G}$ completely. We remark however that it is the relation (23) that will yield Hilbert pairs of wavelets which we discuss in the following subsection.

In summary, we have shown the following proposition.

Proposition 1. *For the two FBs in Fig. 2, with sampling factors p, q, s, suppose $H(\omega)$ and $G(\omega)$ satisfy the conditions in* (11)*. Let*

$$\Theta(\omega) = -\operatorname{sign}(\omega)\,\frac{\pi}{2} + \frac{\omega}{2} \qquad \textit{for } \omega \in [-\pi, \pi]. \tag{26}$$

If

$$\tilde{H}(\omega) = H(\omega)\,e^{-j\,(q-p)\,\omega/2}, \tag{27a}$$
$$\tilde{G}(\omega) = G(\omega)\,e^{j\,\Theta(\omega)}, \tag{27b}$$
$$\tilde{H}^{(1)}(\omega) = H(\omega)\,e^{-j\,q\omega/2}, \tag{27c}$$

then provided it is not the highpass or lowpass subband, the i^{th} subband of the two FBs compute inner products of the input with discrete time sequences that are related by the discrete-time Hilbert transform.

Hilbert Pairs of Wavelets

In [1], we defined the wavelets via the infinite product formula as,

$$\hat{\psi}(\omega) = \sqrt{\frac{p}{q}}\,G\left(\frac{p}{q}\,\omega\right)\,\prod_{k=2}^{\infty}\frac{1}{\sqrt{pq}}\,H\left(\frac{\omega}{p}\left(\frac{p}{q}\right)^{k}\right). \tag{28}$$

We also recall that the wavelets are bandlimited to $q\pi/p$. Now as in the previous section, if we set

$$\tilde{G}(\omega) = G(\omega)\,e^{j\,\Theta(\omega)} \tag{29}$$
$$\tilde{H}(\omega) = H(\omega)\,e^{-j\,\tau(\omega)}, \tag{30}$$

then for $|\omega| \le q\pi/p$.

$$\hat{\tilde{\psi}}(\omega) = \sqrt{\frac{p}{q}}\,\tilde{G}\left(\frac{p}{q}\omega\right)\,\prod_{k=2}^{\infty}\frac{1}{\sqrt{pq}}\,\tilde{H}\left(\frac{\omega}{p}\left(\frac{p}{q}\right)^{k}\right) \tag{31}$$
$$= \hat{\psi}(\omega)\,\exp\left\{j\left[\Theta\left(\frac{p}{q}\,\omega\right) - \sum_{k=2}^{\infty}\tau\left(\frac{\omega}{p}\,\left(\frac{p}{q}\right)^{k}\right)\right]\right\}. \tag{32}$$

Now using

$$\Theta(\omega) = -\operatorname{sign}(\omega)\,\frac{\pi}{2} + \frac{\omega}{a}, \tag{33}$$
$$\tau(\omega) = \frac{\omega}{b} \tag{34}$$

along with the sufficient condition $(q-p)b = a$, we obtain,

$$\Theta\left(\frac{p}{q}\,\omega\right) + \sum_{k=2}^{\infty}\tau\left(\frac{\omega}{p}\,\left(\frac{p}{q}\right)^{k}\right) \tag{35}$$
$$= -\operatorname{sign}(\omega)\,\frac{\pi}{2} + \frac{p}{q}\,\frac{\omega}{a} - \sum_{k=2}^{\infty}\frac{\omega}{b}\,\frac{1}{p}\,\left(\frac{p}{q}\right)^{k} \tag{36}$$
$$= -\operatorname{sign}(\omega)\,\frac{\pi}{2}. \tag{37}$$

Recalling $\hat{\psi}(\omega) = \hat{\tilde{\psi}}(\omega) = 0$ for $|\omega| > q\pi/p$, this implies that $\tilde{\psi}(\cdot)$ is the Hilbert transform of $\psi(\cdot)$. We remark that for the bandpass discrete-time functions, the atoms from the *first* FB were the Hilbert transforms of the atoms from the *second* FB. This difference stems from the definition of the wavelets – the wavelets are defined using the filter frequency responses wheras the discrete-time atoms come from the time-reverse of the filter impulse responses which involve the conjugates of the filter frequency responses.

IV. An Application : Time-Scaling Audio

DT-RADWT is not a tool restricted to perform analysis or synthesis only – it can be useful for doing both, thanks to its tight frame property. Therefore it can be used for *processing* signals which typically require an analysis as well as a synthesis scheme. In the following, we will briefly present an application, namely, time-scaling of audio signals, that makes use of both the analysis and the synthesis parts of the transform.

Consider a multicomponent model for an audio signal given by

$$f(t) = \sum_{i=1}^{K} A_i(t)\,\cos(\varphi_i(t)). \tag{38}$$

In principle, this signal can be scaled in time by a constant α (without varying the instantaneous frequency content) [19], using

$$f_\alpha(t) = \sum_{i=1}^{K} A_i(\alpha\,t)\,\cos(\alpha^{-1}\,\varphi_i(\alpha\,t)). \tag{39}$$

Lacking a representation as in (38), we have to address the problem of defining/separating the 'K' components, in order for the scaling formula (39) to work. One particular method is to employ the phase vocoder (or STFT – see [20]). In that case, one obtains an expansion of $f(x)$ as,

$$f(t) = \sum_{i=1}^{I} b_i(t)\,\exp\big(j(\omega_i\,t + \varphi_i(t))\big), \tag{40}$$

where i runs over the subbands of STFT (with a total of I subbands), and ω_i denotes the center frequency of the i^{th} subband. In this expansion, one can regard each subband signal as a single component signal, with known center frequency. Therefore, one can compute the deviations from the center frequency by monitoring the change in the phase term. The time-scaled version of $f(x)$ is then obtained as,

$$f_\alpha(t) = \sum_{i=1}^{I} b_i(\alpha\,t)\,\exp\left(j\big(\omega_i\,t + \int_0^t \varphi_i'(\alpha\,y)\,dy\big)\right). \tag{41}$$

Phase-vocoder based time-scaling method is very fast and performs quite well even for values of α far from unity. However, one artifact is that the time-scaled audio output sounds 'phasy' – the output is perceived as if the source is far away from the microphone (for a discussion of this effect and some remedies see [21] and the references therein).

The only requirement in the formulation (41) based on the phase vocoder is that the subbands be narrow enough in frequency, in order to be able to calculate the instantaneous frequency using the phase information. Therefore this approach is suitable also for DT-RADWT with a high enough Q-factor. We can use the DT-RADWT in place of the STFT to obtain a time-scaled version of the input. One motivation for replacing the STFT with a constant-Q transform like the DT-RADWT is the constant-Q behavior of human audio perception for frequencies above 500 Hz [22]. A particularly relevant model would therefore be a constant-Q filter bank on this frequency range, which we can approximate with an analytic wavelet

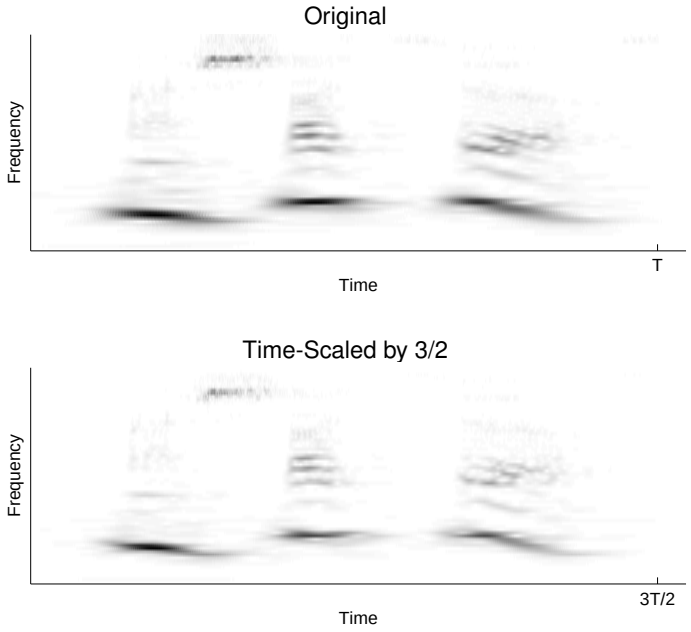

Fig. 10. The time-frequency distributions of the original and the time-scaled signals, using the DT-RADWT, are similar as expected. Ideally, we would like the two scalograms to be the same upto a scaling along the time axis.

transform[5]. We also remark that for this particular method, an analytic transform like the DT-RADWT is more suitable than the RADWT since it facilitates computation of the instantaneous frequency.

The idea of using an analytic constant-Q transform for audio time-scaling was in fact mentioned by Youngberg in [13]. However the implementation in [13] requires approximations of continuous-time functions. On the other hand, the DT-RADWT is fully discrete and provides exact reconstruction.

An example is shown in Fig. 10. The scalograms are seen to be similar, as expected from a successful time-scaling algorithm. An interesting and positive side effect when we use the DT-RADWT (for which we do not have an explanation) is the lack of phasiness that is heard clearly for the same time-scaling factor when the phase vocoder is used. However, for speech, in some of our experiments some distortion occurs with the DT-RADWT, which is not present when the phase-vocoder is utilized.[6]

V. Conclusion

We proposed a dual-tree RADWT that extends the RADWT, similarly as the DT-$\mathbb{C}$WT extends the dyadic DWT. The DT-RADWT allows us to choose the scaling factor of the wavelet frame and the Q-factors of the atoms. This in turn leads to a finer frequency resolution than the dyadic wavelet frames and facilitates the processing of signals which are known to possess quasi-periodic components. Therefore, using DT-RADWT, one can perform tasks that are typically out of the scope of DT-$\mathbb{C}$WT. Also, unlike the STFT, it is a constant-Q transform which can be a desired property for audio-processing and/or, signal processing based on 'scales'.

Here, we mainly focused on the construction and discussed an application to demonstrate the potential of DT-RADWT. In fact, one disadvantage of DT-RADWT, in comparison to an FFT implementation of STFT, is the computation time required. We hope to address this issue in the near future.

[5]The quality factor of these filters do not determine the dilation factor. It is in fact suggested that, for modelling the peripheral auditory system, one should consider the overlap in frequency to be as high as possible, leading to a dilation factor close to unity – see [22], p.159.

[6]Matlab code for time-scaling using the DT-RADWT along with some examples is available at `http://web.itu.edu.tr/ibayram/DtRadwt/`

References

[1] İ. Bayram and I. W. Selesnick, "Frequency-domain design of overcomplete rational dilation wavelet transforms," *IEEE Trans. Signal Processing*, vol. 57, no. 8, pp. 2957–2972, Aug. 2009.

[2] L. Cohen, *Time-Frequency Analysis*. Prentice Hall, 1995.

[3] J. C. Brown, "Calculation of a constant Q spectral transform," *J. Acoust. Soc. Amer.*, vol. 89, no. 1, pp. 425–434, Jan. 1991.

[4] R. Yu, "A new shift-invariance of discrete-time systems and its application to discrete wavelet transform analysis," *IEEE Trans. Signal Processing*, vol. 57, no. 7, pp. 2527–2537, Jul. 2009.

[5] I. W. Selesnick, R. G. Baraniuk, and N. G. Kingsbury, "The dual-tree complex wavelet transform - A coherent framework for multiscale signal and image processing," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 123–151, Nov. 2005.

[6] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*. Prentice-Hall, 1999.

[7] I. W. Selesnick, "Hilbert transform pairs of wavelet bases," *IEEE Signal Processing Letters*, vol. 8, no. 6, pp. 170–173, Jun. 2001.

[8] R. Yu and H. Özkaramanlı, "Hilbert transform pairs of orthogonal wavelet bases: Necessary and sufficient conditions," *IEEE Trans. Signal Processing*, vol. 53, no. 12, pp. 4723–4725, Dec. 2005.

[9] N. G. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *J. of Appl. and Comp. Harm. Analysis*, vol. 10, no. 3, pp. 234–253, May 2001.

[10] G. Gambardella, "A contribution to the theory of short-time spectral analysis with nonuniform bandwidth filters," *IEEE Trans. Circuit Theory*, vol. 18, no. 4, pp. 455–460, Jul. 1971.

[11] J. E. Youngberg and S. F. Boll, "Constant-Q signal analysis and synthesis," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Proc. (ICASSP)*, 1978.

[12] T. L. Petersen and S. F. Boll, "Critical band analysis-synthesis," *IEEE Trans. Acoust., Speech, and Signal Proc.*, vol. 31, no. 3, pp. 656–663, Jun. 1983.

[13] J. E. Youngberg, "Rate/Pitch modification using the constant-Q transform," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Proc. (ICASSP)*, 1979.

[14] J. C. Brown and M. S. Puckette, "An efficient algorithm for the calculation of a constant Q transform," *J. Acoust. Soc. Amer.*, vol. 92, no. 5, pp. 2698–2701, Nov. 1992.

[15] A. V. Oppenheim, D. H. Johnson, and K. Steiglitz, "Computation of spectra with unequal resolution using the fast Fourier transform," *Proc. IEEE*, vol. 59, no. 2, pp. 299–301, Feb. 1971.

[16] A. Harma, M. Karjalainen, L. Savioja, V. Välimäki, U. K. Laine, and J. Huopaniemi, "Frequency-warped signal processing for audio applications," *J. Audio Eng. Soc.*, vol. 48, no. 11, pp. 1011–1031, 2000.

[17] G. Evangelista and S. Cavaliere, "Frequency-warped filter banks and wavelet transforms: A discrete-time approach via Laguerre expansion," *IEEE Trans. Signal Processing*, vol. 46, no. 10, pp. 2638–2650, Oct. 1998.

[18] P. Auscher, "Wavelet bases for $L^2(R)$ with rational dilation factor," in *Wavelets and Their Applications*, M. B. R. et al., Ed. Jones and Barlett, Boston, 1992.

[19] S. Mallat, *A Wavelet Tour of Signal Processing*. Academic Press, 1998.

[20] M. Dolson, "The phase vocoder: A tutorial," *Computer Music Journal*, vol. 10, no. 4, pp. 14 – 27, 1986.

[21] J. Laroche and M. Dolson, "Improved phase vocoder time-scale modification of audio," *IEEE Trans. Signal Processing*, vol. 7, no. 3, pp. 323–332, May 1999.

[22] H. Fastl and E. Zwicker, *Psychoacoustics: Facts and Models*, Third ed. Springer, 2005.