APPROXIMATE CONTINUOUS WAVELET TRANSFORM WITH AN APPLICATION TO NOISE REDUCTION

James M. Lewis and C. Sidney Burrus

Electrical and Computer Engineering Department
Rice University, Houston, TX 77251–1892
lewi@rice.edu, csb@rice.edu

ABSTRACT

We describe a generalized scale-redundant wavelet transform which approximates a dense sampling of the continuous wavelet transform (CWT) in both time and scale. The dyadic scaling requirement of the usual wavelet transform is relaxed in favor of an approximate scaling relationship which in the case of a Gaussian scaling function is known to be asymptotically exact and irrational. This scheme yields an arbitrarily dense sampling of the scale axis in the limit. Similar behavior is observed for other scaling functions with no explicit analytic form. We investigate characteristics of the family of Lagrange interpolating filters (related to the Daubechies family of compactly-supported orthonormal wavelets), and finally present applications of the transform to denoising and edge detection.

1. INTRODUCTION

The continuous wavelet transform (CWT) is useful for many applications in signal analysis. Computer evaluation of the CWT, however, requires an efficient discretization of the transform, which has been the subject of much research. Most existing fast approximation techniques involve the use of a dyadic scaling relationship ([1, 2], and others), and/or require that the scaling functions and wavelets used have a closed-form expression [1, 2, 3]. We present a novel approach with an approximate scaling relationship and $\mathcal{O}(N)$ complexity per scale ($\mathcal{O}(N^2)$ overall), which expands the class of scaling functions and wavelets available for CWT-based analysis.

First, we define the CWT of a funcion $f(t)$ as

$$W_\psi f(a,b) = \langle f, \psi_{a,b} \rangle = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} \psi\left(\frac{t-b}{a}\right) f(t)dt, \quad (1)$$

where a and b are continuous scaling and translation parameters, respectively, and $\psi(t)$ is the real-valued mother wavelet. A *wavelet series* (WS) is simply a sampling of the CWT defined such that $WS_\psi(j,k) = W_\psi f(2^j, k2^j)$ for $j,k \in \mathbf{Z}$, where admissible wavelets ψ are restricted to those satisfying a dyadic scaling relationship

$$\psi(t) = \sum_k g(k)\sqrt{2}\phi(2t-k), \quad (2)$$

where $g(k)$ are the *wavelet filter* coefficients and ϕ is a *scaling function* which serves to allow accurate expansion of a signal at a finite number of scales, and to dilate the wavelet at dyadic scales. It must satisfy

$$\phi(t) = \sum_k h(k)\sqrt{2}\phi(2t-k), \quad (3)$$

where $h(k)$ are the *scaling filter* coefficients.

Mallat [4] derived a discrete orthogonal filter bank implementation of the wavelet series, which we will refer to here as the discrete wavelete transform (DWT). The DWT approximates a sampling of the WS (and hence the CWT), and is exact for certain classes of input signals and signal discretization methods [5]. The DWT can be interpreted as an expansion in terms of discrete basis sequences which approximate continuous basis functions [6]. The basis sequences correspond to the overall impulse response at the output of each stage of the DWT filter bank. The DWT is not invariant to time shifts of the input signal; the *redundant* wavelet transform (RDWT), also known as the *à trous* algorithm [7, 5] is a shift-invariant extension to the DWT which provides integer sampling of the time axis at dyadic scales.

In this paper, we propose a generalization of the RDWT which approximates (but does not sample) a CWT. It is shift-invariant and allows for various scale sampling densities, all of which will be more dense than the dyadic scales of the DWT or RDWT. The formulation differs from classical multiresolution analysis (MRA) [8], in that it does not require a specific scaling relation between vector spaces (or stages of a DWT), but still utilizes the idea of nested scaling spaces. A multitude of perfect reconstruction implementations are possible using only finite impulse response (FIR) filters. Applications to denoising and edge detection will be examined, with results at least comparable to existing wavelet-based methods.

2. SCALE

For a continuous analytic function, the concept of scale is clear; every point in the rescaled function is defined by the relation

$$f_{a,b}(x) = \frac{1}{\sqrt{a}} f\left(\frac{x-b}{a}\right), \quad (4)$$

where a and b are the continuous scaling and shift parameters, and the factor of $1/\sqrt{a}$ preserves the norm of f.

There is some ambiguity, however, when considering discrete signals or sampled continuous signals. Only a finite number of samples are available to represent all scalings of the signal. The finest scale available is set by the sampling rate; to achieve a finer

This work was supported by Texas ATP grant and TI.

scale, one must know or assume an underlying continuous function and resample at a higher rate. Likewise, to move to coarser scales requires some decision about how to combine information from several pixels. In practice, scale is generally defined by the characteristics of the smoothing, or *scaling* function used to produce a coarser scale. This is usually a well-defined continuous function, or in the case of wavelets, a function that can be defined by its self-similarity over a dyadic grid of scales. Then, instead of rescaling the signal for each level of analysis, one can simply rescale the analyzing function; the effect is the same as if a single analyzing function had been applied to different scales of the signal.

In what follows, we will not achieve an exact scale relationship between basis functions, as in the dyadic wavelet case. Instead, we will consider discrete basis sequences which converge to a continuous function as the scale becomes infinitely large. Implemented in a filter bank, the ouput of each low-pass stage can be considered to be an appoximately scaled version of that continuous function, with the approximation becoming more accurate as the number of iterations increases.

3. A SCALE-REDUNDANT TRANSFORM

The scale-redundant discrete wavelet transform (SRDWT) is a completely undecimated filter bank implementation which results in a very fine (approximate) scale sampling. It is not to be confused with the RDWT, which is often referred to as an undecimated wavelet transform. Since the RDWT is a dyadic transform, it must include either downsampling or upsampling in its implementation. Shensa [5] describes two equivalent implementations of the RDWT; one involves averaging decimated DWTs and the other (the à trous algorithm) upsamples the filters at each stage of the filter bank. Figure 1 shows the general form for the SRDWT, which is simply a cascaded filter bank with (potentially) different scaling filters at each stage.

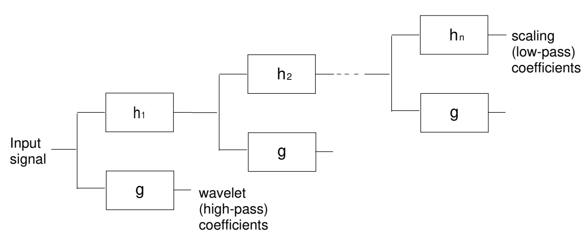

Figure 1: General filter bank for the SRDWT.

The simplest implementation of the SRDWT in Figure 1 for a given scaling filter h would require $h_{2n+1} = h$ and $h_{2n} = h(-n)$ for $n = 0, 1, 2 \ldots$ where the time-reversed version of h yields symmetric impulse responses at the output of each even stage. In practice, however, a logarithmic sampling of scales is often desirable and more computationally feasible. In this paper, we will investigate a scheme which requires

$$h_1 = h, \;\; h_2 = h(-n), \text{ and } h_n = h^{(2^{n-2})} \qquad (5)$$

for $n = 3, 4, 5 \ldots$ where $h^{(n)}$ denotes the filter with a Fourier transform of $|H(\omega)|^n$. The second stage is time-reversed to produce a symmetric basis sequence, and each remaining filter is simply chosen to match the overall impulse response of the system at the output of the previous stage. Thus, the basis sequence at each scale is the convolutional square of the previous scale's sequence.

We will now analyze the characteristics of the SRDWT when h and g are chosen to be members of the Daubechies family of compactly-supported orthonormal wavelets [9]. We will use the notation D_{2n} for each wavelet system, where $2n$ is the wavelet/scaling function filter length ($D2$ is the well-known Haar system). Related to the Daubechies family, and perhaps more fundamental, are the family of Lagrange à trous interpolating filters, which are the convolutional square of the Daubechies filters [5]. The term "à trous" implies that the filter satisfies $f_{2n} = \delta(n)$ within a constant. This property allows the filter to interpolate an upsampled sequence without disturbing the original values of the sequence. It can be shown that a DWT which uses an à trous scaling filter exactly samples the CWT [5], a useful property for our purposes. The scaling functions associated with the Lagrange filters in a DWT are known as the Deslauriers and Dubuc limiting functions [10] and, like the Daubechies functions, have no explicit analytic form.

Figure 2 shows the limiting shapes of the discrete basis sequences in our scheme. The Haar ($D2$) case converges to a Gaussian function, a result which is fairly well-known and will be examined in the next section. The $D4$ through $D8$ cases converge to functions which are remarkably close to scaled versions of the Deslauriers and Dubuc limiting functions, one of which is shown as the dashed curve overlaying the $D4$ SRDWT scaling function. This suggests that the basis sequences in our scheme are converging rapidly to an accurate scale relationship. The Deslauriers and Dubuc limiting functions are produced by exact dyadic scalings of the original Lagrange filter, since it interpolates itself perfectly at each stage of a DWT filter bank.

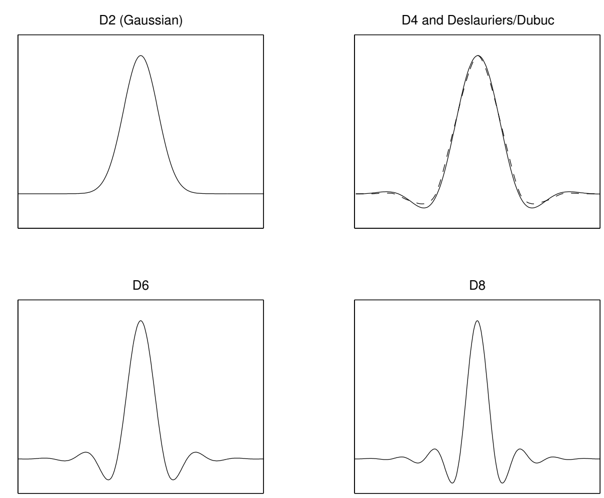

Figure 2: Scaling functions approximated by the SRDWT. $D4$ is compared to the Deslauriers and Dubuc limiting function.

3.1. Scaling Behavior

The Daubechies filters exhibit an interesting scaling behavior when applied to the SRDWT with the method in (5). Figure 3 shows the RDWT scaling filter frequency responses on the left for D2, D4, and D6 at the outputs of the first three filter bank stages. The same filters applied to the SRDWT for several stages are shown on the right. The bold responses roughly match the three stages of the RDWT in cutoff frequency. The $D2$ SRDWT appears to contain two stages for every RDWT stage, the $D4$ four stages, and so on. This scaling behavior was also observed on the time axis, suggesting that the scaling relation satisfied in the limit is of the form

$$\phi_{j+1,k}(t) = \sum_k h_j(k) 2^{-j/2L} \phi_{j,k}(2^{-j/L}t - k), \qquad (6)$$

where L is the even length of the corresponding Daubechies filter. The amplitude scaling behavior does not converge as rapidly as the time axis scaling, so the SRDWT must be re-normalized at each stage. The effect on computational complexity is not significant since both the filter coefficients and normalization constants can be pre-calculated and stored.

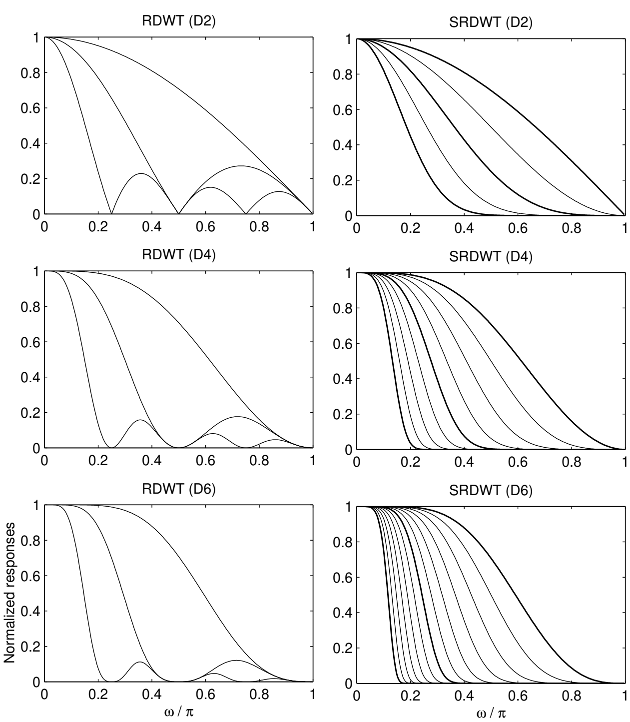

Figure 3: Frequency responses of first three stages of RDWT and corresponding SRDWT stages.

3.2. Inversion

The RDWT (and SRDWT) allow great freedom in the design of analysis and synthesis filters, which is often not fully exploited. For a given set of analysis filters, there is an infinite number of possible synthesis filters, which can be designed to optimize some criterion such as regularity or smoothness.

For simplicity, we have constrained the synthesis scaling filters to be identical to their analysis counterparts. The synthesis wavelet filters are then given by solving the frequency-domain equation

$$H(\omega)\tilde{H}(\omega) + G(\omega)\tilde{G}(\omega) = 1, \qquad (7)$$

where $\tilde{H}$ and $\tilde{G}$ are the frequency responses of the synthesis scaling and wavelet filters.

The resulting filters are all FIR, but get very long after several stages. Due to the small scaling factors, however, the significant portion of the filter grows quite slowly so they can be truncated with very little error.

3.3. An Edge Detection Application

First we will consider the $D2$ case which, according to Bernoulli's theorem, converges to a Gaussian envelope given by

$$\phi(t) = \frac{1}{\sqrt{2\pi}\sqrt{N}/2}\, e^{-\frac{1}{2}\left(\frac{t-N/2}{\sqrt{N}/2}\right)^2} \qquad (8)$$

for a length $N+1$ scaling sequence, assuming a unity sampling rate. The intermediate basis sequences are coefficients of B-splines of increasing order [11].

The Gaussian kernel has been proven to be optimal for some applications in the fields of scale-space theory and edge detection [12, 13]. Berkner has implemented a fast scale-redundant transform [14] for edge detection which uses $h_n = D2$ for all stages. We have adopted the same name for our generalized application.

Using our logarithmic scale sampling scheme of (5) and the Gaussian function (8), we observe that the scale parameter a in the CWT is equivalent to the factor $\sqrt{N}/2$ in (8). In our scheme, N is doubled at every stage, so the scaling factor (asymptotically) given by (6) is $\sqrt{2}$, instead of the usual factor of two.

3.4. A Denoising Example

In this example, we show promising results from the use of the SRDWT in Donoho's denoising scheme [15]. Earlier work has shown a significant improvement in denoising performance by using the RDWT instead of the DWT [16]. This improvement suggested that a more redundant transform might yield further improvements.

We compared the performance of the D4 SRDWT using a hard-thresholding scheme with D4, D6, and D8 RDWTs also using hard-thresholding. Figure 4 shows a realization where the SRDWT compares favorably with the best (in mean squared error) RDWT result.[1] Figure 5 shows mean squared and maximum errors averaged over 40 realizations with about 10 dB signal to noise ratio (SNR). The thresholds were the only parameter optimized here; the intent is only to illustrate the potential of the SRDWT for denoising applications.

4. CONCLUSION

Many of the constraints associated with the DWT are required to ensure an orthogonal basis for signal analysis. We have shown here that the dyadic scaling relationship, while convenient for exact scaling of discrete signals, can also be approximated effectively. The framework of the SRDWT allows great flexibility in designing

[1] Test signal was generated by Donoho's MATLAB routine `MakeSignal` from his software package `WaveLab`.

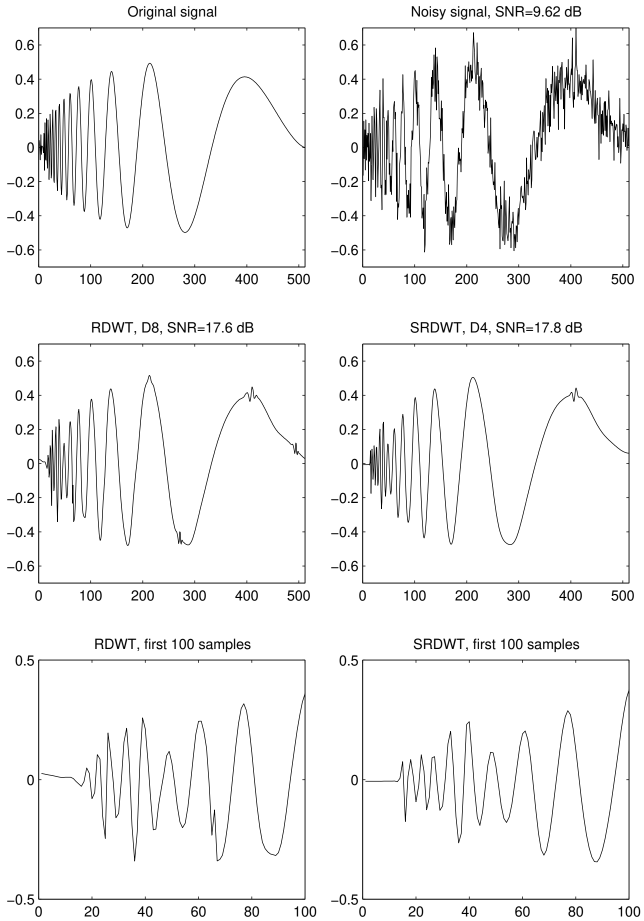

Figure 4: Denoising comparison for "Doppler" signal with optimal thresholds determined experimentally. RDWT uses a full decomposition with hard-thresholding; SRDWT uses 16 stages with hard-thresholding.

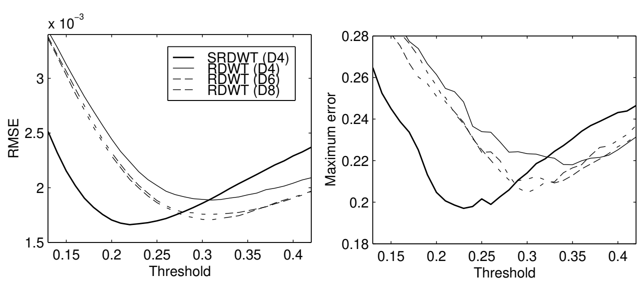

Figure 5: Comparison of RMSE and maximum error for 10 dB Doppler signal, averaged over 40 realizations.

approximate wavelet algorithms. For the case of the Lagrange interpolating filters, we have extended the scaling relationship from 2^j to $2^{j/L}$ without utilizing the dyadic scaling equation.

5. ACKNOWLEDGMENTS

The authors thank Kathrin Berkner for valuable discussions and insight into continuous wavelet transforms and the scale-redundant transform.

6. REFERENCES

[1] M. Vrhel, C. Lee, and M. Unser, "Fast continuous wavelet transform," in *IEEE Proc. Int. Conf. Acoust., Speech, Signal Processing*, (Detroit, MI), pp. 1165–1168, 1995.

[2] O. Rioul and P. Duhamel, "Fast algorithms for discrete and continuous wavelet transforms," *IEEE Trans. Inform. Theory*, vol. 38, pp. 569–586, Mar. 1992.

[3] M. Unser, A. Aldroubi, and S. J. Schiff, "Fast implementation of the continuous wavelet transform with integer scales," *IEEE Trans. Signal Processing*, vol. 42, pp. 3519–3523, Dec. 1994.

[4] S. Mallat, "A theory for multiscale signal decomposition: The wavelet representation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

[5] M. J. Shensa, "The discrete wavelet transform: Wedding the à trous and Mallat algorithms," *IEEE Trans. Inform. Theory*, vol. 40, pp. 2464–2482, 1992.

[6] O. Rioul, "A discrete-time multiresolution theory," *IEEE Trans. Signal Processing*, vol. 41, pp. 2591–2606, 1993.

[7] M. Holschneider, R. Kronland-Martinet, J. Morlet, and P. Tchamitchian, "A real-time algorithm for signal analysis with the help of the wavelet transform," in *Wavelets Time-Frequency Methods and Phase Space* (J. M. Combes, A. Grossman, and P. Tchamitchian, eds.), (Berlin Heidelberg), pp. 286–297, Springer-Verlag, 1989.

[8] S. Mallat, "Multiresolution approximation and wavelets," *Trans. of American Math. Soc.*, vol. 315, pp. 69–88, 1989.

[9] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Comm. Pure Applied Math.*, vol. XLI, pp. 909–996, Nov. 1988.

[10] G. Deslauriers and S. Dubuc, "Symmetric iterative interpolation processes," *Constr. Approx.*, vol. 5, pp. 49–68, 1995.

[11] M. Unser, A. Aldroubi, and M. Eden, "On the asymptotic convergence of B-spline wavelets to gabor functions," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 864–872, 1992.

[12] T. Lindeberg, *Scale-Space Theory in Computer Vision*. Boston, MA: Kluwer Academic Publishers, 1994.

[13] C. J. Evertsz, K. Berkner, and W. Berghorn, "A local multiscale characterization of edges applying the wavelet transform," in *Proc. Nato A.S.I., Fractal Image Encoding and Analysis*, (Trondheim), Center for Complex Systems and Visualization, University of Bremen, July 1995.

[14] K. Berkner and R. O. Wells, Jr., "A fast approximation to the discrete wavelet transform with applications," in *Proc. 31st Asilomar Conf.*, 1997.

[15] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inform. Theory*, vol. 41, pp. 613–627, May 1995.

[16] M. Lang, H. Guo, J. E. Odegard, C. S. Burrus, and R. O. Wells, Jr., "Noise reduction using an undecimated discrete wavelet transform," *IEEE Signal Processing Letters*, vol. 3, pp. 10–12, Jan. 1996. Also Tech. report CML TR95-06.