# Rapid Computation of the Continuous Wavelet Transform by Oblique Projections

Michael J. Vrhel, *Member, IEEE*, Chulhee Lee, and Michael Unser, *Senior Member, IEEE*

***Abstract*— We introduce a fast simple method for computing the real continuous wavelet transform (CWT). The approach has the following attractive features: It achieves $O(N)$ complexity per scale, the filter coefficients can be analytically obtained by a simple integration, and the algorithm is faster than a least squares approach with negligible loss in accuracy. Our method is to use $P$ wavelets per octave and to approximate them with their oblique projection onto a space defined by a compactly supported scaling function. The wavelet templates are expanded to larger sizes (octaves) using the two-scale relation and zero-padded filtering. Error bounds are presented to justify the use of an oblique projection over an orthogonal one. All the filters are FIR with the exception of one filter, which is implemented using a fast recursive algorithm.**

## I. Introduction

THE CONTINUOUS wavelet transform (CWT) is an often used tool in the analysis of nonstationary and fractal signals (e.g., EEG) [10], [18]. The typically long length of these signals and the large number of scales computed in their analysis provides a motivation for considering fast algorithms to compute the CWT. Here, we define the real CWT of the signal $s(t)$ as the inner product

$$W_\psi s(\alpha,\tau) = \frac{1}{\sqrt{\alpha}}\left\langle s(t), \psi\left(\frac{\tau - t}{\alpha}\right)\right\rangle = \frac{1}{\sqrt{\alpha}}\int_{-\infty}^{+\infty} s(t)\psi\left(\frac{\tau - t}{\alpha}\right) dt \qquad (1)$$

where $\alpha$ and $\tau$ are, respectively, the continuously varying scaling and shifting parameters, and the real function $\psi(t)$ is the mother wavelet.[1]

In practice, the variables $\alpha$ and $\tau$ are sampled over the plane of values. Fast algorithms exist for computing the wavelet transform at the dyadic scales $\alpha = 2^i$ when the wavelet is associated with a multiresolution [1], [4], [9], [13]. In this paper, we are interested in a finer sampling of the scale axis.

Previous methods for computing at nondyadic scales include an algorithm for computing the CWT at the integer sample values $\alpha = i, \tau = k$ with $O(N)$ number of operations per scale, where $N$ is the length of the input signal [18]. Another approach that computes the CWT along arbitrary scales, again with $O(N)$ complexity per scale, is discussed in [15]. This last method is restricted to Gabor-like wavelets (i.e., modulated Gaussians). A method that achieved $O(N)$ operations per scale, with no restrictions on the shape of the wavelet, and with arbitrarily fine exponential sampling along the scale axis was introduced in [19]. That approach approximated wavelets at several scales by their orthogonal projection onto a space defined by a compact scaling function. Most other algorithms to date are implemented using an FFT-based approach and require $O(N\log(N))$ computations per scale [5], [7], [11].

Here, we introduce a more general framework that includes the method [19] as a particular case while providing advantages from the computational point of view. In particular, we will show that this new system can result in a much simpler specification of the wavelet filters. A surprising result is that this increase in speed, flexibility, and simplicity is achieved with a negligible loss in accuracy.

In the algorithm, we approximate a set of $P$ wavelets by their oblique projection into a space defined by one compact scaling function orthogonal to a second compact analysis function. While this approach yields an approximation to the actual CWT, we have full control over the approximation error, and we can achieve any desired level of accuracy.

The algorithm consists of an FIR filter bank ($P$ FIR filters for computing $P$ scales or voices per octave) and a fast recursive IIR filter. The system is shown in Fig. 1, where the additional FIR filter $[h(k)]_{\uparrow 2^i}$ performs the two scale stretching of the approximating scaling function. This two-scale relationship enables us to use the same FIR filters in computing the CWT for $P$ scales within each octave (with appropriate zero padding between the filter coefficients). In the simplest form, the filter coefficients for the FIR filter bank are analytically determined by integration of the wavelet

$$q_{\alpha_i}(k) = \int_{k-1/2}^{k+1/2} \psi(t/\alpha_i)\, dt, \qquad i = 0, \cdots, P-1$$

where the $\alpha_i$'s are the corresponding scale parameters.

If, in addition, we perform the approximation using cubic splines, then the refinement filter $h(k)$ is the symmetric binomial kernel

$$H(z) = \tfrac{1}{8}(z^{-2} + 4z^{-1} + 6 + 4z + z^2)$$

and the IIR filter

$$Q_{12}(z) = \frac{384}{(z^2 + 76z + 230 + 76z^{-1} + z^{-2})}$$

is implemented using an efficient recursive routine that is described in Appendix A.

Manuscript received February 1, 1996; revised September 20, 1996. The associate editor coordinating the review of this paper and approving it for publication was Dr. Farokh Marvasti.

The authors are with the Biomedical Engineering and Instrumentation Program, National Center for Research Resources, National Institutes of Health, Bethesda, MD 20892 USA.

Publisher Item Identifier S 1053-587X(97)02581-6.

[1] To simplify the notation throughout the paper, we use a definition of the wavelet transform that is a time-reversed version of the conventional one.

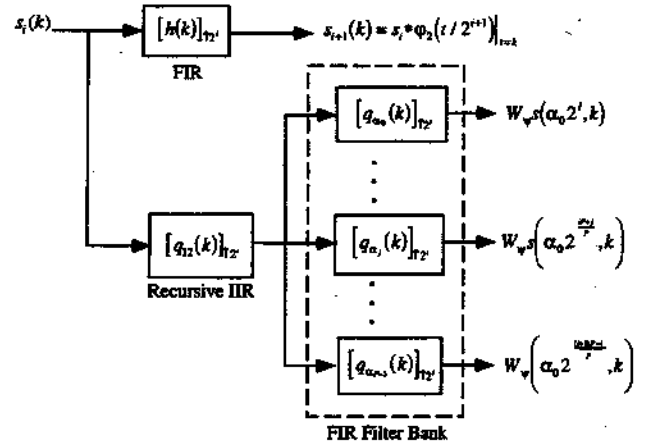


Fig. 1. System diagram for the computation of the CWT in the $i$th octave. All filters are expanded by the factor $2^i$ using zero padding.

The organization of the paper is as follows: In Section II, we provide all relevant definitions and introduce the oblique projection operator, after which the equations describing the system in Fig. 1 quickly follow. In Section III, we study the approximation error and its properties. In Section IV, we consider the admissibility of the projected wavelet and its vanishing moments. In Section V, we present the implementation as a series of design steps and describe a particular implementation, providing explicit formulas for determining the FIR filter coefficients. Finally, in Section VI, we present some results and compare the performance of the algorithm with an FFT-based method.

## II. General Principle of the Approach

Our goal is to efficiently compute the CWT at $P$ scales per octave. This will be achieved by constructing a set of $P$ auxiliary wavelets $\{\psi_i(t) \cong \alpha_i^{-1/2}\psi(t/\alpha_i)\}_{i=0,\cdots,P-1}$, which are the oblique projections of the wavelets $\{\alpha_i^{-1/2}\psi(t/\alpha_i)\}_{i=0,\cdots P-1}$ into a space defined by the scaling function $\varphi_2(t)$ orthogonal to the space defined by the analysis function $\varphi_1(t)$. In the remainder of the paper, both functions will be assumed to be compactly supported. Here, we will list the properties that these functions must satisfy and introduce the oblique projection operation. The accuracy of the approximation and the admissibility of the projected wavelet will be addressed in Sections III and IV, respectively.

### A. Scaling Functions and Oblique Projections

An $L$th-order scaling function $\varphi(t) \in L_2$ is a function that satisfies the following three conditions:

i) $0 < A \le \sum_{k\in Z} |\hat{\varphi}(\omega + 2\pi k)|^2 \le B < +\infty$

ii) $\hat{\varphi}(0) = 1$ and $\hat{\varphi}^{(m)}(2\pi k) = 0$, $k \in Z, k \neq 0$ for $m = 0, \cdots, L-1$

iii) $\varphi\left(\frac{t}{2}\right) = \sum_{k\in Z} h(k)\varphi(t-k)$

where $\hat{\varphi}(\omega)$ is the Fourier transform of $\varphi(t)$, and $\hat{\varphi}^{(m)}(2\pi k)$ is the $m$th derivative of $\hat{\varphi}(\omega)$ evaluated at $2\pi k$. Property i) ensures that the subspace

$$V(\varphi) = \left\{ h(t) = \sum_{k\in Z} c(k)\varphi(t-k) \quad c \in l_2 \right\}$$

is a well-defined (closed) subspace of $L_2$ [2]. Property ii) implies that $\varphi(t)$ reproduces all polynomials of degree $L-1$, which will be of importance with regard to the approximation errors (cf. Section III). Property iii) is referred to as the two-scale relation, and it allows us to dilate all the wavelet filters by a power of two. An analysis function is defined as a function that satisfies only properties i) and ii).

Criteria for the selection of a scaling or analysis function should be based on its approximating power, smoothness, and shortness of its support, to name a few. These properties will be discussed in detail in Sections III and IV.

For a pair of analysis functions $\varphi_1$ and $\varphi_2$, the oblique projection of the wavelet $\psi_{\alpha_1}(t) = \psi(t/\alpha_i)$ into $V(\varphi_2)$ orthogonal to $V(\varphi_1)$ can be expressed in terms of the basis generated by $\varphi_2$ and a set of coefficients $p_{\alpha_i}(k)$, where

$$\psi_i(t) = \sum_{k\in Z} p_{\alpha_i}(k)\varphi_2(t-k). \tag{2}$$

The original wavelet is measured in terms of the analysis function $\varphi_1$ providing the values

$$q_{\alpha_i}(k) = \langle \psi_{\alpha_i}(t), \varphi_1(t-k) \rangle. \tag{3}$$

The coefficients $p_{\alpha_i}(k)$ can be obtained from the values $q_{\alpha_i}(k)$ while enforcing a projection constraint by the use of the digital filter $q_{12}$, which is the convolution inverse of the cross-correlation sequence $\langle \varphi_1(t-k), \varphi_2(t) \rangle$, i.e.

$$Q_{12}(z) = \frac{1}{\sum_{k\in Z} \langle \varphi_1(t-k), \varphi_2(t) \rangle z^{-k}}$$

(cf. [16]). Note that if $\varphi_1$ and $\varphi_2$ both have compact support, then $Q_{12}(z)$ is an all-pole filter. Thus, if the filter is stable, the corresponding sequence $q_{12}(k)$ has exponential decay.

Using the filter $q_{12}(k)$ to obtain the coefficients $p_{\alpha_i}(k)$, we have

$$p_{\alpha_i}(k) = (q_{\alpha_i} * q_{12})(k). \tag{4}$$

Mathematical details of the oblique projection operator are contained in [16]. Note that if $\psi$ and $\varphi_1$ both have compact support, then the FIR filters in the filter bank of Fig. 1 $\{q_{\alpha_i}\}_{i=0,\cdots P-1}$ also have finite support. In addition, we have an orthogonal projection if $\varphi_1 \in V(\varphi_2)$; in particular, $\varphi_1 = \varphi_2$ is the solution considered in [19].

### B. CWT Approximation

In order to achieve $O(N)$ complexity per scale, we will replace the computation of the convolution in (1) by its approximation

$$\tilde{W}_\psi s(\alpha_i, \tau) = (s * \psi_i)(\tau). \tag{5}$$

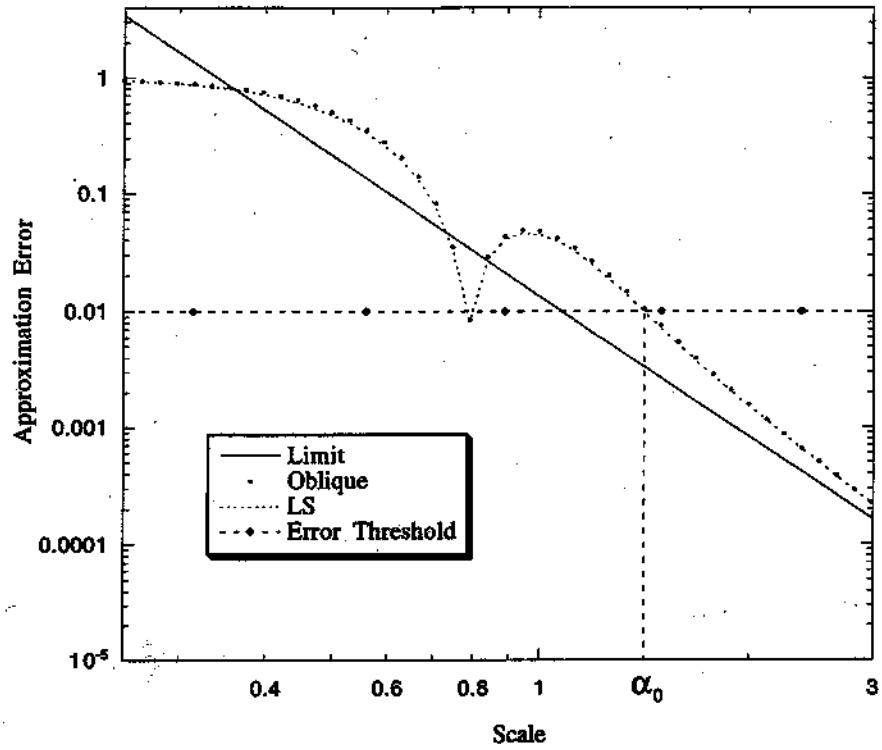


Fig. 2. Approximation errors for the second derivative Gaussian wavelet.

Substituting (2) into the above, we have the following equations:

$$s_0(t) = (s * \varphi_2)(t) \tag{6}$$

$$\tilde{W}_\psi s(\alpha_i, \tau) = \sum_{k \in Z} p_{\alpha_i}(k) s_0(\tau - k). \tag{7}$$

In principle, we need a filter $p_{\alpha_i}(k)$ for every scale we wish to compute. However, we will make use of the two-scale relationship of $\varphi_2$ given as property iii) (hence, $\varphi_2$ is a scaling function), which allows us to use the same set of filters for each octave.

In practice, the input signal will actually be the sample values $s[k] = s(t)|_{t=k}$. Computing the continuous convolution in (6) requires an interpolation of these sample values. Alternatively, one can use a Riemann approximation for the convolution integral resulting in the simple initialization

$$s_0[k] \cong (s * b_2)(k)$$

where $b_2(k) = \varphi_2(t)|_{t=k}$.

Incorporating property iii) into (4), (6), and (7), sampling, and performing simple algebraic manipulations, we obtain the following algorithm:

$$s_{i+1}(k) = (s_i * [h]_{\uparrow 2^i})(k) \tag{8}$$

$$\tilde{s}(k) = (s_i * [q_{12}]_{\uparrow 2^i})(k) \tag{9}$$

$$\tilde{W}_\psi s(2^i \alpha_j, k) = (\tilde{s}_i * [q_{\alpha_j}]_{\uparrow 2^i})(k), \qquad j = 0, \cdots, P-1 \tag{10}$$

where $[h]_{\uparrow 2^i}$ is $h(k)$ with $2^i - 1$ zeros between each sample (i.e., expanded by a factor of $2^i$). Equation (8) performs the two-scale filtering defined by property iii) for $\varphi_2$; equation (9) is the correction filtering that ensures that we have an oblique projection of our wavelet; equation (10) is the FIR filtering that constructs the approximated wavelets in terms of the scaling function and computes the wavelet coefficients. These equations describe the algorithm as it is shown in Fig. 1. Note that only $\varphi_2$ must have property iii)—the two-scale relationship.

## III. The Approximation Error

In the previous section, we replaced the convolution of the CWT by its approximation (5). Here, we consider the behavior of the approximation error. In particular, we show how we can tune our algorithm to maintain the error below an acceptable threshold. Reduction of the error is achieved by selecting the appropriate functions $\varphi_1$ and $\varphi_2$, and adjusting the fine-scale resolution $\alpha_0$.

The approximation power of a scaling function depends on its ability to reproduce polynomials up to a specific degree $n$. This degree plus one $(L = n + 1)$ provides the order of accuracy of the approximation function. The properties of the approximation error for the orthogonal projection case are given by the Strang–Fix conditions [14] . We will first describe the relationship between the oblique and the orthogonal approximation errors and then translate this relationship to the Strang–Fix conditions.

The oblique projection into $V(\varphi_2)$ orthogonal to $V(\varphi_1)$ is related to the orthogonal projection into $V(\varphi_2)$ by the following: (cf. [16, Th. 3])

$$\|\psi_\alpha - P_2\psi_\alpha\| \le \|\psi_\alpha - P_{2\perp 1}\psi_\alpha\| \le \frac{1}{\cos(\theta)}\|\psi_\alpha - P_2\psi_\alpha\| \tag{11}$$

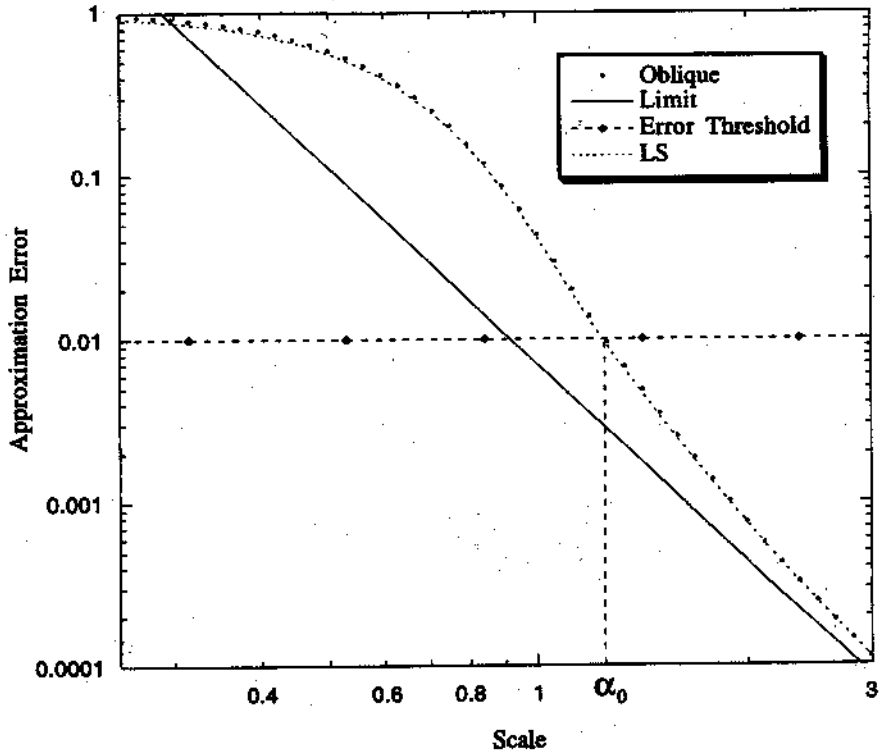


Fig. 3. Approximation errors for the first derivative Gaussian wavelet.

where

$P_2$ orthogonal projection operator into $V(\varphi_2)$,

$P_{2\perp 1}$ oblique projection operator into $V(\varphi_2)$ orthogonal to $V(\varphi_1)$,

$\theta$ largest angle between $V(\varphi_1)$ and $V(\varphi_2)$,

and

$$\cos(\theta) = \operatorname*{ess\,inf}_{\omega\in(0,\pi]} \frac{\sum_{k\in Z}\hat{\varphi}_2(\omega+2\pi k)\cdot\sum_{k\in Z}\hat{\varphi}_1^*(\omega+2\pi k)}{\sqrt{\sum_{k\in Z}|\hat{\varphi}_1(\omega+2\pi k)|^2}\cdot\sqrt{\sum_{k\in Z}|\hat{\varphi}_2(\omega+2\pi k)|^2}}.$$

As the angle between the subspaces decreases, the bound (11) becomes tighter, and the oblique projection error approaches the orthogonal error. In particular, if $\varphi_1 \in V(\varphi_2)$, then $\cos(\theta) = 1$, and $V(\varphi_1) = V(\varphi_2)$.

As already mentioned, the behavior of the least squares approximation error as a function of the scale is described by the Strang–Fix theory (cf. [19]). For the $L_2$th-order scaling function $\varphi_2$ [cf. condition ii)], the theory provides the relationship

$$\|\psi_\alpha - P_2\psi_\alpha\| \le C_{\varphi_2}\frac{\|\psi^{(L_2)}\|}{\alpha^{L_2}}$$

where the constant $C_{\varphi_2}$ is a function of $\varphi_2$, and $\|\psi^{(L_2)}\|$ is the norm of the $L_2$th derivative of $\psi$. In other words, the error typically decays like $O(\alpha^{-L_2})$. This bound translates to a bound on the oblique projection in the following fashion using relationship (11):

$$\|\psi_\alpha - P_{2\perp 1}\psi_\alpha\| \le \frac{C_{\varphi_2}}{\cos(\theta)}\frac{\|\psi^{(L_2)}\|}{\alpha^{L_2}}. \tag{12}$$

Hence, we can expect the same kind of decay as in the least squares case.

In practice, the oblique error is much closer to the orthogonal error than what is indicated by the above worst-case bound. In fact, it can be shown that the oblique and orthogonal errors have the same limiting behavior as $\alpha \to \infty$. Specifically, for $\alpha$ sufficiently large, the approximation error behaves as

$$\|\psi_\alpha - P_{2\perp 1}\psi_\alpha\| = C_2\frac{\|\psi^{(L_2)}\|}{\alpha^{L_2}} + O\left(\frac{1}{\alpha^{L_2+1}}\right) \tag{13}$$

where $L_2$ is again the order of approximation of $\varphi_2$, and

$$C_2 = \frac{1}{L_2!}\left(\sum_{k\neq 0}|\hat{\varphi}_2^{(L_2)}(2\pi k)|^2\right)^{1/2}.$$

This is exactly the same relation as the one given for the least squares case in [19] (including the value of the constants).

Fig. 2 displays the approximation error curve for the cubic LS approximation of a second derivative Gaussian wavelet and an oblique approximation using the zeroth- and third-degree $B$-splines for $\varphi_1$ and $\varphi_2$, respectively. Fig. 3 displays a similar plot for the first derivative Gaussian wavelet. Remarkably, the difference between the LS and oblique errors is very small. It is also clear that both curves have the same asymptote, which is given by (13) (for cubic splines, the value of the constant is $C_2 = 9.092\,41 \times 10^{-4}$.) In both examples, the loss of performance is negligible; in fact, the discrepancy is much less than the factor $1/\cos(\theta)$ (here, $\cos(\theta) = 0.892$) predicted by the theory (worst-case scenario). Thus, we can conclude that for all practical purposes, the orthogonal and oblique projections are equivalent in terms of their approximation power.

## IV. Wavelet Properties

Here, we consider the admissibility of the projected wavelet as well as the number of vanishing moments it exhibits.

### A. Admissibility

A wavelet $\psi$ is admissible if it satisfies the following conditions [3, p. 24]:

$$\exists \beta > 0 \text{ such that } \int (1+|t|)^{\beta} |\psi(t)|\, dt < \infty \qquad (14)$$

and

$$\int \psi(t)\, dt = 0. \qquad (15)$$

Here, we will consider only the bounding condition (14) since the projected wavelet is shown to satisfy the zero mean property in Section IV-B.

*Proposition 4.1:* If the analysis function $\varphi_1$, the scaling function $\varphi_2$, and the wavelet $\psi$ all have compact support, then the oblique projection of $\psi_{\alpha_i} = \psi(t/\alpha_i)$ into $V(\varphi_2)$ orthogonal to $V(\varphi_1)$ satisfies $\int (1+|t|)^{\beta} |\psi_i(t)|\, dt < \infty$ for some $\beta > 0$, where $\psi_i$ is the oblique projection of $\psi_{\alpha_i}$.

*Proof:* Substituting the projected wavelet given by (2) into (14), we want to guarantee

$$\int (1+|t|)^{\beta} \left| \sum_{k \in Z} p_{\alpha_i}(k) \varphi_2(t-k) \right| dt < \infty \quad \text{for some } \beta > 0.$$

Since $\varphi_2$ has a compact support, this condition will obviously be satisfied (for all $\beta > 0$) if $p_{\alpha_i}$ has exponential decay; that is, if

$$|p_{\alpha_i}(k)| \le C e^{-\gamma |k|} \quad \text{for some } \gamma > 0.$$

Clearly, the sequence $q_{\alpha_i}(k) = \langle \psi_{\alpha_i}(t), \varphi_1(t-k) \rangle$ has a finite support. In addition, $q_{12}(k)$ is the convolution inverse of $\langle \varphi_1(x-k), \varphi_2(x) \rangle$ and decays exponentially since $Q_{12}(z)$ is an all-pole filter. From (4), we have $p_{\alpha_i} = (q_{12} * q_{\alpha_i})(k)$. Therefore, $p_{\alpha_i}$ satisfies the exponential decay requirement. □

### B. Vanishing Moments

The ability of the analyzing wavelet to detect a singularity typically depends on the interaction between its number of vanishing moments and the Hölder exponent associated with the singularity [8]. For this reason, wavelets are often designed with a specific number of vanishing moments. The number of vanishing moments of the approximated wavelet will depend on the number in the actual wavelet and on the approximation power of $\varphi_1$. This is stated in the following proposition.

*Proposition 4.2:* If $\varphi_1$ is an $L_1$th-order analysis function, $\psi$ has $M$ vanishing moments (i.e., $\hat{\psi}^{(m)}(\omega)|_{\omega=0} = 0, m = 0, \cdots, M-1$), and $\varphi_1, \varphi_2$, and $\psi$ are all compactly supported, then the oblique projection of $\psi$ into $V(\varphi_2)$ orthogonal to $V(\varphi_1)$ has $\min(M, L_1)$ vanishing moments.

*Proof:* From (3), we have in the Fourier domain:

$$\hat{Q}(\omega) = \sum_{k \in Z} \hat{\psi}(\omega + 2\pi k) \hat{\varphi}_1(\omega + 2\pi k).$$

Differentiating $\hat{Q}(\omega)$ $i$ times and evaluating at $\omega = 0$ produces

$$\hat{Q}^{(i)}(0) = \sum_{k \in Z} \sum_{j=0}^{i} \binom{i}{i-j} \hat{\psi}^{(j)}(2\pi k) \hat{\varphi}_1^{(i-j)}(2\pi k).$$

From property ii) in Section II-A and the initial assumptions, we have

$$\hat{\varphi}_1^{(i-j)}(2\pi k) = 0, \qquad i - j = 0, \cdots, L_1 - 1 \quad \text{and} \quad k \neq 0$$
$$\hat{\psi}^{(i)}(0) = 0, \qquad i = 0, \cdots, M-1$$

which implies that

$$\hat{Q}^{(i)}(0) = 0, \qquad i = 0, \cdots, \min(M-1, L_1 - 1).$$

From (4), we have

$$\hat{P}(\omega) = \hat{Q}(\omega) \hat{Q}_{12}(\omega).$$

Differentiating the above $l$ times and evaluating at $\omega = 0$ leads to

$$\hat{P}^{(l)}(0) = \sum_{i=0}^{l} \binom{l}{l-i} \hat{Q}^{(i)}(0) \hat{Q}_{12}^{(l-i)}(0).$$

From the proof of Proposition 4.1, we know that $q_{12}(k)$ decays exponentially, which ensures that $\hat{Q}_{12}^{(i)}(0) \propto \Sigma_{k \in Z}\, k^i q_{12}(k)$ is finite. Therefore

$$\hat{P}^{(l)}(0) = 0, \qquad l = 0, \cdots, \min(M-1, L_1 - 1).$$

Now, differentiating $\hat{\psi}_i(\omega)$ $m$ times and evaluating at $\omega = 0$ leads to

$$\hat{\psi}_i^{(m)}(0) = \sum_{l=0}^{m} \binom{m}{m-l} \hat{P}^{(l)}(0) \hat{\varphi}_2^{(m-l)}(0).$$

Since $\varphi_2$ is compactly supported, $\hat{\varphi}_2^{(m-l)}(0) \propto \int x^{m-l} \varphi_2(x)\, dx$ is finite, which implies that

$$\hat{\psi}_i^{(m)}(0) = (-j)^m \int x^m \psi_i(x)\, dx = 0,$$
$$m = 0, \cdots, \min(M-1, L_1 - 1). \qquad \square$$

An implication of the above proposition is that in applications for which a large number of vanishing moments are required, it is necessary to use a high-order function for $\varphi_1$. In addition, since the minimum number of vanishing moments for admissibility is one, we have the following corollary.

*Corollary 4.3:* If $\Sigma_{k \in Z}\, \varphi_1(x+k) = 1$ and $\psi, \varphi_1, \varphi_2$ are all compactly supported, then the projection of an admissible wavelet into $V(\varphi_2)$ orthogonal to $V(\varphi_1)$ is admissible as well.

*Proof:* Using Poisson's summation formula, we have that

$$\sum_{k \in Z} \varphi_1(x+k) = \sum_{k \in Z} \hat{\varphi}_1(2\pi k) e^{j 2\pi k}.$$

Thus, the conditions

$$\sum_{k \in Z} \varphi_1(x+k) = 1 \quad \text{and} \quad \hat{\varphi}_1(2\pi k) = \delta_k$$

are equivalent, and $\varphi_1$ has at least a first order of approximation ($L_1 \ge 1$). The proof now follows from the application of Propositions 4.1 and 4.2. □

Note that the shortest and simplest analysis function that satisfies this property is the box function: a $B$-spline of degree zero.

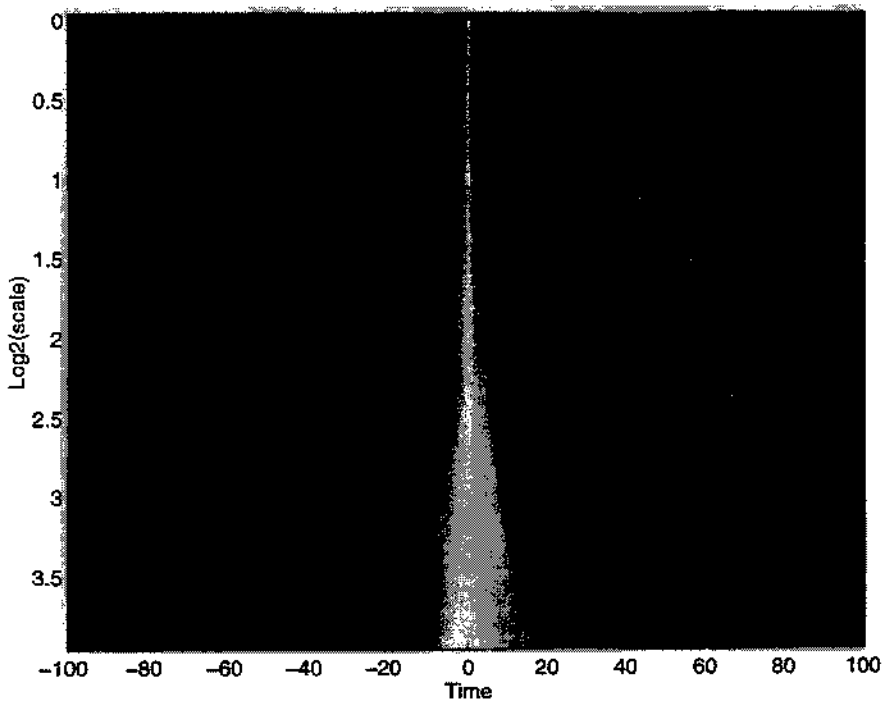


Fig. 4. Impulse response of the system for the second derivative Gaussian wavelet.

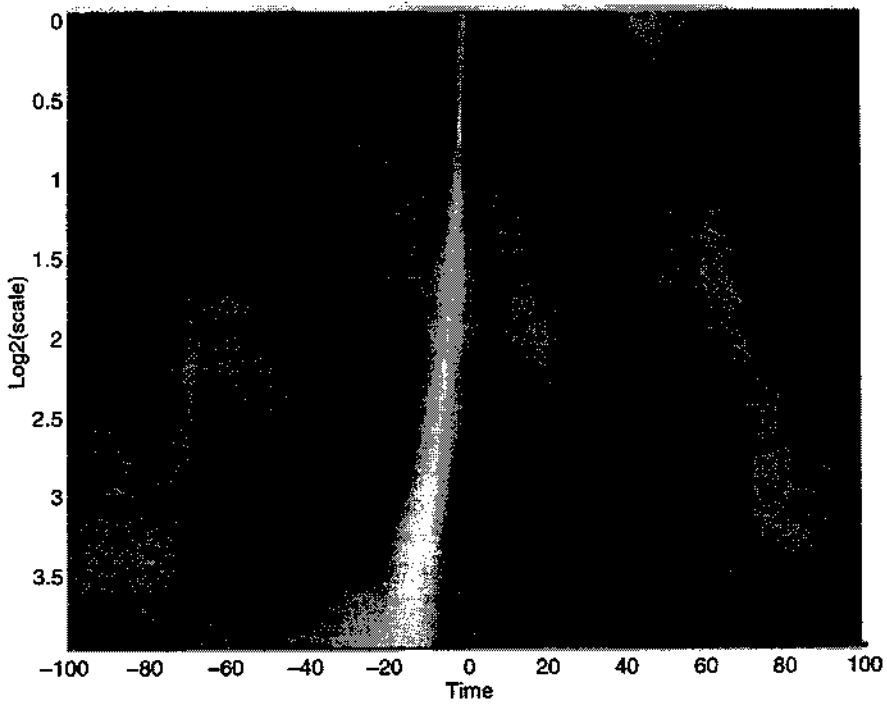


Fig. 5. Impulse response of the system for the first derivative Gaussian wavelet.

## V. Implementation

In this section, we will summarize the implementation of the algorithm as a series of steps, each of which are illustrated with an example. The performance of our design is discussed in Section VI.

*Step 1:* Select the compactly supported wavelet function for the application at hand. For illustrative purposes, we consider the following two wavelet shapes: a truncated version of the first derivative of a Gaussian

$$\psi_{\text{deriv}}(t) = \begin{cases} -K_o \dfrac{t}{\alpha_o} e^{-(t/\alpha_o)^2/2}; & \left|\dfrac{t}{\alpha_o}\right| \leq 5 \\ 0; & \text{otherwise} \end{cases}$$

and a truncated version of the second derivative of a Gaussian

$$\psi_{mex}(t) = \begin{cases} K_o \left(1 - \left(\dfrac{t}{\alpha_o}\right)^2\right) e^{-(t/\alpha_o)^2} - K; & \left|\dfrac{t}{\alpha_o}\right| \leq 5 \\ 0; & \text{otherwise} \end{cases}$$

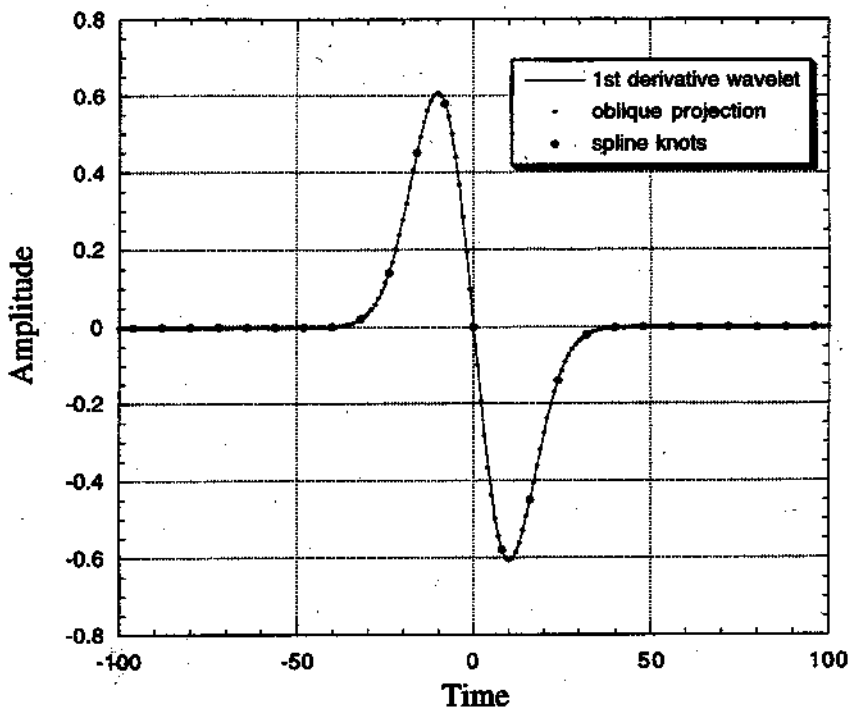


Fig. 6. Comparison of the first derivative Gaussian wavelet and its oblique approximation.

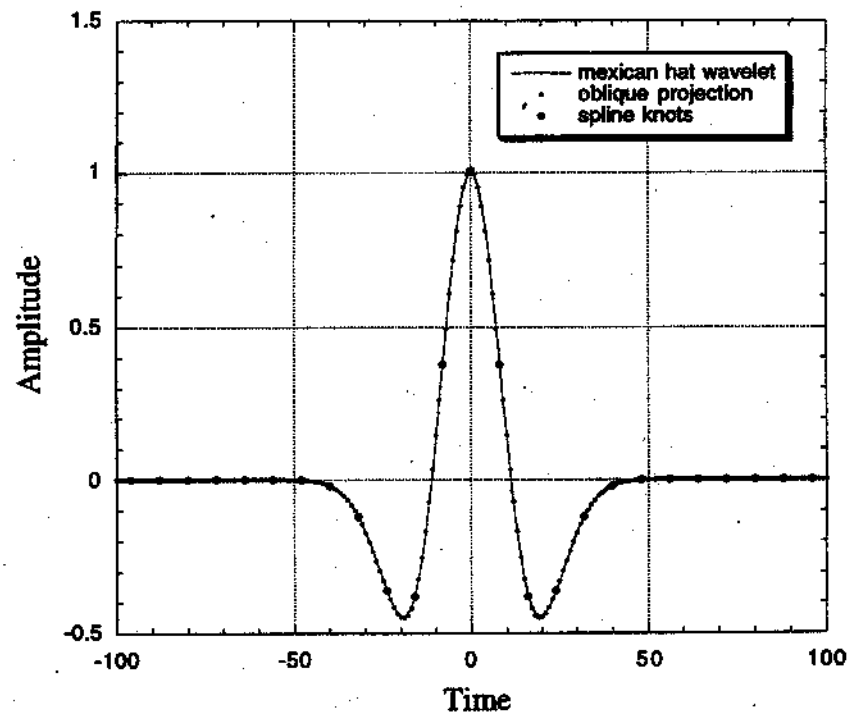


Fig. 7. Comparison of the second derivative Gaussian wavelet and its oblique approximation.

where $K_o$ is a constant ensuring that $\|\psi\| = 1$ and $K_1$ is an offset that guarantees the admissibility of $\psi_{mex}$.

*Step 2:* Select the scaling function $\varphi_2$ and the analysis function $\varphi_1$. Here, it is advantageous to use $B$-splines because of their attractive features:

i) They are the shortest functions for a given order of approximation (this will result in shorter filters).
ii) They are symmetrical, which makes them ideal for approximating symmetrical or anti-symmetrical wavelets.
iii) They have a simple analytical form, which makes them easy to manipulate.
iv) They are smooth well-behaved functions.

To keep things as simple as possible but still maintain a high degree of accuracy, we will use the zero degree $B$-spline (first-order approximation) on the analysis side $\varphi_1 = \beta^0$ and the cubic $B$-spline on the synthesis side $\varphi_2 = \beta^3$. The advantages of this particular selection are as follows:

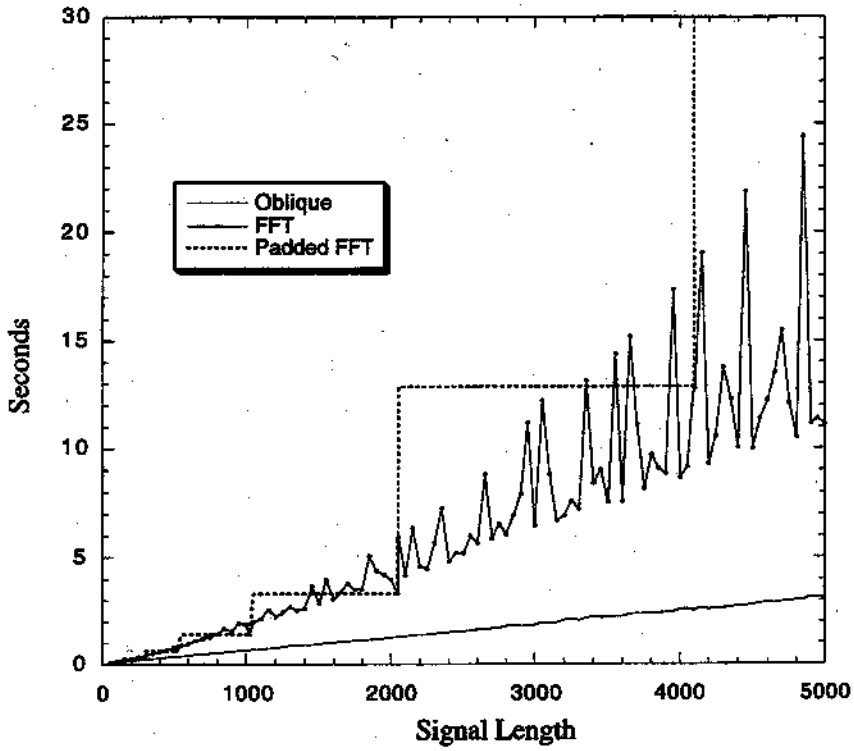


Fig. 8. Speed comparison of FFT based method and the oblique projection method.

i) The FIR filters $\{q_{\alpha_i}(k)\}_{i=o,\cdots M-1}$ are determined analytically by simple integration of the wavelet $\psi(t/\alpha_i)$

$$q_{\alpha_i}(k) = \langle \psi(t/\alpha_i), \beta^0(t-k) \rangle = \int_{k-1/2}^{k+1/2} \psi(t/\alpha_i)\, dt. \tag{16}$$

ii) The FIR filters are shorter, and the IIR filter is of a lower order as compared with the orthogonal projection into $V(\varphi_2)$ considered in [19]. This also implies that the oblique algorithm will be faster than the orthogonal (or least squares) one.

iii) The approximation error is very close to the minimum, which is achieved by the orthogonal projection (cf. Section III).

In this case, we gain in simplicity and speed with little loss in accuracy.

Other properties to consider are the approximation power of the function and the number of vanishing moments desired in the projected wavelet. The latter is, of course, closely tied to the first and second steps, and the former is related to the third step.

*Step 3:* Select an acceptable error level and, hence, a fine-scale resolution. This step should actually be performed in conjunction with the selection of the order of the scaling function since its approximation power determines the exact error characteristics. For a particular error threshold, one can compute a plot such as that shown in Figs. 2 and 3 to determine a suitable order of approximation and the fine-scale resolution $\alpha_0$. Here, we selected an error threshold of 0.01, which is shown as a dashed line in Figs. 2 and 3. This places our fine-scale resolution at 1.26 for the first derivative wavelet and 1.41 for the second derivative wavelet.

*Step 4:* Select the number of scales per octave. This will set the number of FIR filters. We used $P = 12$ scales per octave, where $\alpha_i = \alpha_o 2^{i/12}$ $i = 0, \cdots, 11$, which corresponds to the musical notes $(A, A^{\#}, B, C, C^{\#}, \cdots)$.

With these design steps completed, (16) provides the means for determining the filter coefficients for the FIR filter bank. The remaining FIR filter is the refinement filter $h(k)$. For a $B$-spline of degree $n$ ($n$ odd; $\varphi_2(t) = \beta^n(t)$), the filter $h(k)$ is given by the binomial coefficients

$$u_2^n(k) = \begin{cases} \dfrac{1}{2^n}\binom{n+1}{k+(n+1)/2}; & |k| \le (n+1)/2 \\ 0; & \text{otherwise.} \end{cases}$$

In the most general $B$-spline implementation with $\varphi_1 = \beta^{n_1}$ and $\varphi_2 = \beta^{n_2}$, the IIR filter is given by

$$Q_{12}(z) = \frac{1}{B^{n_1+n_2+1}(z)} \tag{17}$$

where $B^{n_1+n_2+1}(z)$ is the $z$-transform of the sampled $B$-spline kernel of degree $n_1 + n_2 + 1$ (cf. [17]). This filter is implemented in a fast recursive fashion as described in Appendix A.

## VI. Results

We implemented the algorithm in MATLAB with specific routines (zero-padded filtering) coded in C to speed up the computations. The impulse response of the system over four octaves is shown for the wavelets $\psi_{mex}$ and $\psi_{\text{deriv}}$ in Figs. 4 and 5, respectively. The fine scale within each octave contains the same worst-case error. The fine-scale slice in the fourth

TABLE I
ALGORITHM IIR FILTER PARAMETERS

| Spline degree $n_1+n_2+1$ | IIR filter $q_{12}$ parameters | |
|---|---|---|
| | $d_0$ | $\lvert z_i\rvert<1$ |
| 1 | 1 | |
| 2 | 8 | $z_1=-3+2\sqrt{2}$ |
| 3 | 6 | $z_1=-2+\sqrt{3}$ |
| 4 | 384 | $z_1=-0.361341$<br>$z_2=-0.0137254$ |
| 5 | 120 | $z_1=-0.430575$<br>$z_2=-0.0430963$ |
| 6 | 46080 | $z_1=-0.488295$<br>$z_2=-0.0816793$<br>$z_3=-0.00141415$ |
| 7 | 5040 | $z_1=-0.53528$<br>$z_2=-0.122555$<br>$z_3=-0.00914869$ |

octave for both scalograms is compared with the actual function in Figs. 6 and 7. Note that the oblique projection is almost indistinguishable from the actual function.

We compared the performance of our algorithm with an FFT-based method that used a radix-2 algorithm when the signal length $N$ was a power of 2 and a mixed radix method for other signal lengths (MATLAB's FFT algorithm). The input was an electroencephalograph signal (EEG). Fast algorithms for analyzing such signals are of interest for applications that require real-time detection of brain seizures [12].

The length of the signal was varied, and the time required to compute four octaves is shown as a function of the signal length in Fig. 8. The dashed line represents the FFT algorithm when the signal was padded to a power of two. The line with the circles is the FFT algorithm for various signal lengths. The solid line, which is the oblique projection algorithm, clearly demonstrates the $O(N)$ characteristics of the method. In this comparison, the present algorithm appears to be advantageous even for relatively small signal lengths, e.g., $N<128$. The speed improvement over the least squares in [19] is roughly 12%. The more important advantage, though, is that the determination of the wavelet filters is much more straightforward. For the oblique algorithm, the majority of the computation was spent in the FIR filter bank. Additional speedup could be achieved by performing each FIR filter in parallel.

## VII. CONCLUSION

We have introduced a method for the rapid computation of the CWT. The algorithm has the following properties:

- It has a complexity of $O(1)$ per computed wavelet coefficient. This constitutes an $O(\log(N))$ improvement over conventional FFT-based methods.
- The FIR filter coefficients are obtained analytically by simple integration of the wavelet, which is the main practical advantage over a comparable least squares design.
- The approximation error is simple to control by adjusting the fine scale resolution or increasing the order of the scaling function.
- The method is flexible enough to approximate a variety of wavelet shapes and achieves an arbitrarily fine sampling of the scale axis.

Finally, the present theoretical formulation is general enough to contain most previous methods as particular cases.

## APPENDIX A
## THE IIR FILTER

Here, we discuss the recursive implementation of the IIR filter

$$[q_{12}(k)]_{\uparrow m}=[(b^{n_1+n_2+1})^{-1}(k)]_{\uparrow 2^i}.$$

Since $q_{12}$ is a symmetrical all-pole filter, the $z$-transform of its upsampled version $[q_{12}(k)]_{\uparrow m}$ can be written in the following standard form:

$$A_{n_o}(z^m)=\frac{d_0}{[z^{n_o m}+z^{-n_o m}]+\left(\sum_{k=1}^{n_o-1}c_k[z^{km}+z^{-km}]\right)+c_0}$$

where $d_0$ and $\{c_k, k=0,\cdots n_o-1\}$ are constant coefficients, and

$$n_o=\left\lfloor\frac{(n_1+n_2+1)}{2}\right\rfloor.$$

The filter is expressed as a cascade of simple first-order causal/anticausal components

$$A_{n_1}(z^m)=d_o\prod_{i=1}^{n_o}A(z^m;z_i)$$

where $A(z^m;z_i)$ is defined as

$$A(z^m;z_i)=\frac{1}{(1-z_iz^{-m})}\,\frac{-z_i}{(1-z_iz^m)}.$$

Values of $d_0$ and $\{z_i, i=1,\cdots n_o\}$ for different orders are given in Table I.

The above yields the following recursive filter equations:

$$y^+(k)=x(k)+z_iy^+(k-m),\quad (k=m,\cdots,N-1)$$
$$y(k)=z_i(y(k+m)-y^+(k)),\quad (k=N-1-m,\cdots,0)$$

for the input $\{x(k)\}, k=0,\cdots,N-1$. In order to calculate $y^+(k)$ recursively, we need to know $y^+(k)$ for $k=0,\cdots,m-1$. These initial values are computed using

$$y^+(k)=\sum_{j=0}^{k_0}x(k-mj)z_i^j;\quad (k=0,\cdots,m-1)$$

where $k_o=\log(\epsilon)/\log(|z_i|)$ and $\epsilon$ is a prespecified level of precision. Likewise, we determine the initial values for

$y(k), k = N-m, \cdots, N-1$, using the sum decomposition

$$y(k) = \frac{-z_i}{(1-z_i^2)}(y^+(k) - y^-(k) - 1)$$

and a computation of $y^-(k)$ similar to the above. In computing the initial values, the signal is extended by using mirror boundary conditions.

## REFERENCES


[1] P. Abry and A. Aldroubi, "Designing multiresolution analysis-type wavelets and their fast algorithms," *J. Fourier Anal. Applications*, vol. 2, no. 2, pp. 135–159, 1995.

[2] A. Aldroubi and M. Unser, "Sampling procedures in function spaces and asymptotic equivalence with Shannon's sampling theory," *Numer. Funct. Anal. Optimization*, vol. 15, no. 1 and 2, pp. 1–21, Feb. 1994,.

[3] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.

[4] P. Dutilleux, "An implementation of the algorithme à trous to compute the wavelet transform," in *Proc. Wavelets: Time-Frequency Methods Phase Space*, 1989, pp. 298–304.

[5] D. L. Jones and R. G. Baraniuk, "Efficient approximation of continuous wavelet transforms," *Electron. Lett.*, vol. 27, no. 9, pp. 748–750, 1991.

[6] S. Maes, "A fast quasi-continuous wavelet transform algorithm," in *Proc. Workshop Time, Frequency, Wavelets Multiresolution Theory*, INSA-Lyon, 1994.

[7] ——, "The wavelet transform in signal processing, with application to the extraction of the speech modulation model features," Ph.D. dissertation, Universite Catholique De Louvain, 1994.

[8] S. G. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Trans. Inform. Theory*, vol. 38, pp. 617–643, Mar. 1992,.

[9] S. G. Mallat, "A theory of multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 674–693, July 1989.

[10] J. F. Muzy, E. Barcy, and A. Arneodo, "Wavelets and multifractal formalism for singular signals: Application to turbulence data," *Phys. Rev. Lett.*, vol. 67, no. 25, pp. 3515–3518, 1991.

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

[12] S. J. Schiff, J. Heller, S. L. Weinstein, and J. Milton, "Wavelet transforms and surrogate data for electroencephalographic spike and seizure detection," *Opt. Eng.*, vol. 33, no. 7, pp. 2162–2169, 1994.

[13] M. J. Shensa, "The discrete wavelet transform : wedding the à trous and Mallat algorithms," *IEEE Trans. Signal Processing*, vol. 40, pp. 2464–2482, Oct. 1992.

[14] G. Strang and G. Fix, "A Fourier analysis of the finite element variational method," in *Proc. Conf. Constructive Aspect Functional Anal., Edizioni Cremonese*, Rome, 1971, pp. 796–830.

[15] M. Unser, "Fast Gabor-like windowed Fourier and continuous wavelet transforms," *IEEE Signal Processing Lett.*, vol. 1, pp. 76–79, May 1994.

[16] M. Unser and A. Aldroubi, "A general sampling theory for non-ideal acquisition devices," *IEEE Trans. Signal Processing*, vol. 42, pp. 2915–2925, Nov. 1994.

[17] M. Unser, A. Aldroubi, and M. Eden, "$B$-spline signal processing. Part I: Theory," *IEEE Trans. Signal Processing*, vol. 41, pp. 821–833, Feb. 1993.

[18] 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.

[19] M. J. Vrhel, C. Lee, and M. Unser, "Fast continuous wavelet transform: A least squares formulation," *Signal Processing*, to appear.


**Michael J. Vrhel** (M'94) was born in St. Joseph, MI, in 1964. He received the B.S. degree in electrical engineering from Michigan Technological University, Houghton, in 1987. He received the M.S. and Ph.D. degrees in electrical engineering from North Carolina State University, Raleigh, in 1989 and 1993, respectively.

He was the recipient of a Kodak Fellowship from Eastman Kodak Company, Rochester, NY, from 1989 to 1993. During 1992, he researched problems in color reproduction at the Eastman Kodak Company. From 1993 to 1996, he was a National Research Council Associate at the National Institutes of Health, Biomedical Engineering and Instrumentation Program. Currently, he is a Senior Staff Fellow at the National Institutes of Health, Bethesda, MD. His research interests include color reproduction, signal restoration/reconstruction, fast algorithms, and wavelets.

Dr. Vrhel is an associate member of Sigma Xi.

**Chulhee Lee** received the B.S. and M.S. degrees in electronics engineering from Seoul National University, Korea, in 1984 and 1986, respectively, and the Ph.D. degree in electrical engineering from Purdue University, West Lafayette, IN, in 1992.

From 1986 to 1987, he was a researcher in te Acoustic Laboratory at the Technical University of Denmark (DTH). From 1993 to 1996, he worked with the National Insitutes of Health, Bethesda, MD. In 1996, he joined the Faculty of the Department of Electronic Engineering, Yonsei University, Seoul. His research interests include image/signal processing, pattern recognition, and neural networks.

Dr. Lee is a member of Tau Beta Pi, Eta Kappa Nu, and KSEA.

**Michael Unser** (M'88–SM'94) was born in Zug, Switzerland, on April 9, 1958. He received the M.S. (summa cum laude) and Ph.D. degrees in electrical engineering in 1981 and 1984, respectively, from the Swiss Federal Institute of Technology, Lausanne, Switzerland.

He has been with the Biomedical Engineering and Instrumentation Program, National Institutes of Health, Bethesda, MD, since 1985, where he is presently a Visiting Scientist and Head of the Image Processing Group. His research interests include the application of image processing and pattern recognition techniques to various biomedical problems, multiresolution algorithms, wavelet transforms, and the use of splines in signal processing. He is the author of more than 60 published papers in these areas.

Dr. Unser has served as an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING, is currently an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS, and is a member of the Multidimensional Signal Processing Committee of the IEEE Signal Processing Society. He is also on the editorial boards of *Signal Processing* and *Pattern Recognition*. He co-organized the 1994 IEEE-EMBS Workshop on Wavelets in Medicine and Biology and serves as Chair for SPIE's Conference on Wavelet Applications in Signal and Image Processing, which has been held annually since 1993. He received the Dommer Prize for Excellence from the Swiss Federal Institute of Technology in 1981, the research prize of the Brown-Boweri Corporation for his thesis in 1984, and the IEEE Signal Processing Society's 1995 Best Papaer Award (in the IMDSP technical area) for a Transactions paper co-written with A. Aldroubi and M. Eden on B-spline signal processing.