Continuous wavelet transform with arbitrary scales and $O(N)$ complexity

Arrate Muñoz *, Raphaël Ertlé, Michael Unser

Department of Microtechnology, Biomedical Imaging Group/DMT, Swiss Federal Institute of Technology Lausanne, CH-1015 Lausanne EPEL, Switzerland

Abstract

The continuous wavelet transform (CWT) is a common signal-processing tool for the analysis of nonstationary signals. We propose here a new B-spline-based method that allows the CWT computation at any scale. A nice property of the algorithm is that the computational cost is independent of the scale value. Its complexity is of the same order as that of the fastest published methods, without being restricted to dyadic or integer scales. The method reduces to the filtering of an auxiliary (pre-integrated) signal with an expanded mask that acts as a kind of modified ‘à trous’ filter. The algorithm is well-suited for a parallel implementation. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Continuous wavelet transform; Arbitrary scale; B-spline

1. Introduction

The continuous wavelet transform (CWT) of a signal $f$ with the wavelet $\psi$ is defined as

$$W_\psi f(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x) \psi \left( \frac{b-x}{a} \right) \, dx.$$  

(1)

It can be interpreted as the correlation of the input signal with a time-reversed version of $\psi$ rescaled by a factor of $a$. For a 1-D input signal, the result is a 2-D description of the signal with respect to time $b$ and scale $a$. The scale $a$ is inversely proportional to the central frequency of the rescaled wavelet $\psi_0(x) = \psi(x/a)$ which is typically a bandpass function; $b$ represents the time location at which we analyze the signal. The larger the scale $a$, the wider the analyzing function $\psi_0$, and hence smaller the corresponding analyzed frequency. The output value is maximized when the frequency of the signal matches that of the corresponding dilated wavelet. The main advantage over the Fourier transform (FT) analysis is that the frequency description is localized in time. The advantage over the short-time Fourier transform (STFT) is that the window size varies; low frequencies are analyzed over wide time windows, and high frequencies over narrow time windows, which is more effective than to use a fixed-size analysis. Typical applications of the CWT are the detection and characterization of singularities [3,14], pattern recognition [6], image processing [4,15], fractal analysis [2,12,23], noise reduction [11] and the analysis of biomedical signals [7,10,25].

The main contribution of this paper is the development of a fast algorithm for the computation of the CWT at any real scale $a$ and integer time localization $b$. Mallat’s fast wavelet algorithm [12] uses the multiresolution properties of the wavelet to compute the

* Corresponding author. Tel.: +41-21-693-5142; fax: +41-21-69-3701.

E-mail addresses: arrate.munoz@epfl.ch (A. Muñoz), raphael.ertle@compaq.com (R. Ertlé), michael.unser@epfl.ch (M. Unser).
CWT at dyadic scales \( a = 2^k \) and time shifts \( b = 2^k k \), \( k \in \mathbb{Z} \) [17]; it achieves an overall \( O(N) \) complexity. Other techniques compute the wavelet transform at dyadic scales and integer time points with a ‘\( \text{à trous} \)’ approach. Their complexity per scale is \( O(N) \), the same as Mallat’s algorithm, but with a larger leading constant [3,5,9,16].

Despite their speed, these methods may not be precise enough for some applications, since a dyadic scale progression cannot be finer than an octave sub-band decomposition. To achieve a better scale resolution, other approaches have been proposed, either based on \( M \)-band decomposition inside an octave [17,24] or on a generalization of the two-scale relation to general integer \( N \)-scale relations [8,22]. However, none of these algorithms can handle arbitrary scales.

Our purpose here is to develop a novel and fast algorithm that works for any real value of \( a \). It takes advantage of a B-spline decomposition of the input signal and of the mother wavelet. The method exploits the fact that B-splines are compactly supported and that the convolution of two B-splines can be expressed analytically [13].

2. Operators and definitions

First, we introduce some operators and definitions that will be helpful to solve our problem.

We express a B-spline of degree \( n \) as

\[
b^n(x) = A^{n+1} + \frac{x^n}{n!} + \delta \left( x + \frac{n+1}{2} \right),
\]

where \( x^n = \max(x, 0)^n \) is a one-sided power function; \( A^{n+1} \) denotes the \((n+1)\)-fold iteration of the finite difference operator \( A = \delta(x) - \delta(x-1) \). This latter operator also corresponds to a discrete convolution (digital filter) whose \( z \)-transform is \( A(z) = 1 - z^{-1} \).

Likewise, \( A^{n+1}(z) = (1 - z^{-1})^{n+1} \).

We have the equivalence

\[
\frac{1}{a} b^a \left( \frac{x}{a} \right) = A^{n+1} + \frac{1}{a^{n+1}} \frac{x^n}{n!} \delta \left( x + \frac{n+1}{2} \right),
\]

where \( A_a \) is the rescaled finite-difference operator,

\[
A_a^{n+1} = \sum_{k=0}^{n+1} \binom{n+1}{k} (-1)^k \delta(x - ak).
\]

Using equivalence (3), the rescaled B-spline is rewritten as

\[
\frac{1}{a} b^a \left( \frac{x}{a} \right) = A_a^{n+1} + \frac{1}{a^{n+1}} D^{-(n+1)} \delta \left( x + a \frac{n+1}{2} \right).
\]

From definition (2) of the B-spline and from equivalence (3), we find that the integral of a B-spline of a given degree is a spline with the degree increased by 1. It is given by the following expression (see [13] for the proof):
for the proof)
\[ D^{-(n_1+1)} B_{n_2}(x) = A^{-(n_1+1)} \ast \beta^{(n_1+n_2+1)} \left( x - \frac{n_1+1}{2} \right), \quad (7) \]
where \( A^{-1} \) is the inverse finite-differences operator defined as \( A^{-1}(x) = \sum_{n \geq 0} \delta(x-n) \) whose z-transform is \( A^{-1}(z) = (1 - z^{-1})^{-1} \). It can be defined as the running sum filter, \((A^{-1} s)_k = \sum_{n \leq k} s_n\). Note that \( y_k = (A^{-1} \ast s)_k \) can be implemented very efficiently using the recursive equation \( y_k = y_{k-1} + s_k \).

Next, we consider the mixed convolution of a continuous signal \( v(x) \) with a discrete sequence \( p_k \) when the intersample distance is equal to \( a \).

\[ g(x) = \left( \sum_{k \in \mathbb{Z}} p_k \delta(x - ak) \right) \ast v(x) = \sum_{k \in \mathbb{Z}} p_k v(x - ak). \quad (8) \]

The graphical interpretation of this formula is given in Fig. 4. The mixed convolution is a weighted sum of shifted replicates of the signal \( v(x) \) separated by a distance \( a \).

3. Fast continuous wavelet transform algorithm

3.1. Spline wavelets

Among all existing wavelet bases, B-spline wavelets have the advantage of possessing an explicit formula [1]; most wavelets are defined only implicitly by means of a refinement filter. For example, the well-known Haar wavelet is a weighted sum of two B-splines of degree 0. Other wavelets, such as the first derivative or the second derivative of a Gaussian (Mexican hat wavelet), can be closely approximated by linear combination of B-splines of sufficiently high degrees \( n \geq 2 \) [19].

The description of wavelets on to B-spline basis allows for an efficient computation of the convolution products of the CWT that takes advantage of the convolution properties of B-splines. Thus, we choose to express our mother wavelet \( \psi \) on a B-spline basis of order \( n_1 \). The wavelet at scale \( a \) is represented by its B-spline expansion
\[ \psi \left( \frac{x}{a} \right) = \sum_{k=-K}^K d_k \beta^{n_1} \left( \frac{x}{a} - k \right), \quad (9) \]
where \( d_k \) are the B-spline coefficients (see [21] for additional information on how to choose their value).

3.2. CWT computation

3.2.1. Spline wavelet transform

Using the expression for the rescaled B-spline (6), the CWT (1) for the wavelet (9) becomes

\[ W_{\psi} f(a, b) = \frac{1}{\sqrt{a}} \left( f(\cdot) \ast \sum_{k=-K}^K d_k \beta^{n_1} \left( \frac{\cdot}{a} - k \right) \right)(b) \]
\[ = \left( \frac{1}{a^{n_1+1/2}} \left( \sum_{k=-K}^K d_k \delta(\cdot - ak) \ast \beta^{n_1+1}(\cdot) \right) \ast D^{-(n_1+1)} f \left( \cdot + a \left( \frac{n_1+1}{2} \right) \right) \right)(b). \]

We see that the CWT can be calculated from the application of a discrete convolution operator to the \((n_1+1)\)th integral of a shifted version of the analyzed function \( f \).

Replacing \( \beta^{n_1+1} \) by its definition (5), the CWT is expressed as a mixed convolution
\[ W_{\psi} f(a, b) = \left( \left( \sum_{k=-K}^{n_1+1} p_k \delta(\cdot - al) \right) \ast v(\cdot) \right)(b), \quad (10) \]
where
\[ p_k = \frac{1}{a^{n_1+1/2}} (d \ast q)_k \]
\[ = \frac{1}{a^{n_1+1/2}} \sum_{l=0}^{n_1+1} d(k-l)q(l), \quad (11) \]
\( q_l \) being the finite differences coefficients in (5) and
\[ v(x) = D^{-(n_1+1)} f \left( x + a \left( \frac{n_1+1}{2} \right) \right). \quad (12) \]
3.2.2. Spline input signal

We assume that the continuous input signal \( f(x) \) is a spline that interpolates the discrete input samples \( f(k) \). Thus, we have \( f(x) = \sum_{k \in \mathbb{Z}} c_k \beta^{n_2}(x - k) \) with \( c = f \ast (\beta^{n_2})^{-1} \), where \((\beta^{n_2})^{-1}\) is the inverse filter of the B-spline interpolation filter \( \beta^{n_2}(x) \big|_{x=k} \), as shown elsewhere [20]. Then, we rewrite \( v(x) \) as

\[
v(x) = c \ast D^{-(n_1+1)} \beta^{n_2} \left( x + a \left( \frac{n_1 + 1}{2} \right) \right). \tag{13}
\]

Using the integral formula for a B-spline (7), we get

\[
v(x) = g \ast \beta^{n_1+n_2+1}(x + \tau) \tag{14}
\]

where \( \tau = (a - 1)(n_1 + 1)/2 \) is a shift factor and \( g \) is a discrete sequence obtained by iterative summation of the interpolation coefficients

\[
g = A^{-(n_1+1)} \ast c. \tag{15}
\]

We introduce this relation in the CWT expression (10) which yields,

\[
W_g f(a, b) = \sum_{k=-K}^{K} \sum_{i=0}^{n_1+n_2+1} g_{i_0} w_{a,b}(k, i), \tag{16}
\]

where \( i_0 = \left\lfloor b - ak + \tau - (n_1 + n_2 + 2)/2 \right\rfloor \) and where

\[
w_{a,b}(k, i) = \frac{1}{a^{n_1+1/2}} p_k \beta^{n_1+n_2+1}(b - ak - i - i_0 + \tau). \tag{17}
\]

We have used here the compact-support property of B-splines to reduce the number of terms of the sum over \( i \). In this way, the computation of the CWT reduces to the inner product with the coefficients \( g_i \) of a series of precalculated weights \( w_{a,b}(k, i) \) (which we can store in a look-up table).

In practice, we are typically only interested in the values of \( b \) that correspond to the time locations of the original samples, that is, for integer \( b \). Then, we can use the fact that \( w_{a,b}(k, i) = w_{a,0}(k, i - b) \) if \( b \in \mathbb{Z} \). The algorithm (16) is then equivalent to a discrete convolution. This reduces considerably the number of weights to be precalculated, since only the values \( w_{a,0} = w_a \) are required.

Note the interesting spatial structure of the filter \( w_a \): ‘clusters’ of weights separated by a distance \( a \) of each other. Parameters: \( n_1 = 3, n_2 = 3, K = 1, a = 16 \) and \( b = 128 \).

3.3. Fast implementation

Let us now describe the fast algorithm based on the expansion (16). In the initialization step, the B-spline...
Fig. 6. Schematic representation of the fast wavelet transform: $(b^{p_2})^{-1}$, computation of the interpolation coefficients $c$; $A^{-1(n_1+1)}$, calculation of the $(n_1+1)$-fold integral of $c$; $w_a(k,i)$, look-up table calculation where $k \in [-K,K+n_1+1]$, $i \in [0,n_1+n_2+1]$ and $a_i \in [a_1,a_N]$; $N$ is the number of scales; $w_{an}$, filtering with the mask calculated for each scale; $W_{cd}f(a_i,b)$, wavelet transform of $f$ for scale $a_i$ at position $b$.

expansion coefficients $c$ of the sampled signal $f$ are calculated, and the running-sum operator $A^{-1}$ is applied $(n_1+1)$ times; it is computed recursively by iterating (2) (see Section 3.4 for the implementation details). The intermediate result $g$ does not depend on the scale $a$. For a given scale $a$, we compute the weights $w_a$ and store them in a 2-D look-up table of dimensions $(2K+n_1+2)(n_1+n_2+2)$. These values are then convolved with the precomputed sequence $g$. The values $w_a$ and the inter-cluster distance for the filtering depend on $a$, but the computational complexity is constant and does not depend on $a$. Moreover, the values $w_a$ do not depend on the signal $f$. Thus, the computational complexity per point only depends on the values of $K$, $n_1$, and $n_2$. Note that the independence between scales allows for a straightforward parallel implementation (Fig. 6).

3.4. Boundary conditions

To minimize boundary artifacts, we extend our signal $\{s_k\}_{k=0,...,N-1}$ using symmetric mirror boundary conditions defined as $s_{-k} = s_k$, and $s_{N-1-k} = s_{N-1-k}$, for $k = 0, 1, \ldots, N-1$. The boundary conditions defined above are repeated on further extensions to yield a $(2N-2)$ periodic signal. To implement the running sums (inverse finite difference operator) in a consistent manner, we apply it to zero-mean signals, which is perfectly acceptable since the DC-component is filtered out by the wavelet (vanishing moments) anyway. Thus, assuming that $s$ is zero mean, the inverse finite differences operator is given by

$$\left(A^{-1} \ast s\right)_k = \sum_{l=0}^{k \mod(2N-2)} s_l. \quad (18)$$

For zero-mean signals, the application of the above operator reverses the symmetry of the boundary conditions. That is to say, if the input signal has symmetric boundary conditions, the output will have anti-symmetric boundary conditions defined as $s_k = -s_{-k-1}$, and $s_{N-1-k} = -s_{N-2-k}$, for $k = 0, 1, \ldots, N-1$. An antisymmetric signal is always zero mean. Thus, in our implementation, we alternate between symmetric and anti-symmetric input boundary extensions depending on the parity of the iteration number.

Finally, the boundary conditions—either symmetric or antisymmetric—of the $(n_1+1)$-fold integrated interpolation coefficients will be handled accordingly by the rest of the algorithm (in our case, convolution with the sequence of precalculated weights).

4. Results

Here, we discuss the implementation of our fast CWT algorithm and compare its execution time with a FFT-based implementation. As an example of application, we show the analysis of a biomedical signal.

4.1. Comparison with FFT-based computation

The FFT has an overall $O(N \log N)$ complexity and is therefore asymptotically slower than our method which has an $O(N)$ complexity. This can be observed from the experimental comparison of the computation times shown in Fig. 7. We see that, for long input signals, our method is indeed faster. The interpolation degree for the input signal was zero. The CWT was computed over four octaves with 12 scales per octave. The wavelet was the second derivative of the quintic (see Fig. 8(b)), quartic and cubic B-spline, respectively [22]. The FFT-based method used a radix-2 algorithm when the signal length was a power of 2 and a mixed-radix method for other signal lengths (MATLAB’s FFT algorithm). The region where the algorithm is faster than the FFT-radix 2 method with zero padding is colored in gray. The time required to compute the wavelet in the time domain before its FFT computation was neglected. A parallel implementation for the scale-dependent part of each algorithm would speed up the computations.
4.2. Analysis of a biomedical signal

We have applied our method to the analysis of bowel movements. A magnetically active capsule was swallowed and its gastrointestinal transit was monitored. The measures consisted of its three spatial coordinates and the angles that describe its orientation [18].

Fig. 10(a) corresponds to the $x$-coordinate of the capsule. The sampling time was 70 ms. We have analyzed it using both the real (Fig. 10(b)) and the complex CWT (Fig. 10(c)) for cubic spline interpolation of the input signal. The $y$-axis corresponds to a normalized scale in seconds; it is given by $a_0 = a/f_0$, where $f_0$ is the central frequency of the wavelet. In the real case, the wavelet was the first derivative of a quartic spline expanded by a factor of 2 as shown in Fig. 8(a). A visual inspection highlights the band of period 3 s, which corresponds to the breathing of the patient. We chose to calculate the complex CWT by an extension of the method proposed by Unser [19] for integer scales. The analysis wavelet is the Morlet-like wavelet $\beta^3(x)e^{-4\pi i}$ (Fig. 9).

Using complex analysis (Fig. 10(c)), we discovered three relevant frequency bands: first, the breathing with a period close to 3 s; then, two more bands with periods around 12 and 20–25 s due to the contractions of the stomach.

5. Discussion: integer scale method

In a previous paper [22], Unser et al. describe a fast algorithm for the CWT computation at integer scales using B-splines as basis functions. Their method can be shown to be equivalent to ours when $a$ is an integer. This follows from the identity

\[
A^{-(n+1)}(z)A_a^{n+1}(z) = \left(\frac{1-z^a}{1-z}\right)^{n+1} = \left(U_a^0(z)\right)^{n+1}.
\]
Next we summarize the two ways of computing the CWT at integer scales. We start with the method of Unser et al. for which we write
\[ W_s(a, b) = f \ast (b_{n_2}^{n_1})^{-1} \ast b_{n_1+1}^{n_1+1} \ast (u_0^0)^{n_1+1} \ast [d]_{\uparrow a}, \]

where the upsampling operator \([ \cdot ]_{\uparrow a}\) is defined as
\[ [d]_{\uparrow a}(k) = \begin{cases} d(\frac{k}{a}) & \text{if } a \text{ divides } k, \\ 0 & \text{elsewhere} \end{cases} \]

and the \(z\)-transform of the sequence \(u_0^0\) is given by
\[ U_a^0(z) = z^{\frac{a-1}{2}} \sum_{k=0}^{a-1} z^{-k}. \]

We then summarize our method by
\[ W_\phi s(a, b) = f \ast (b_{n_2}^{n_1})^{-1} \ast f_{\text{scale-dependent}} \ast (u_0^0)^{n_1+1} \ast [d]_{\uparrow a}, \]

In the first approach, \(u_0^0\) can be computed recursively by a moving-sum method that requires only one addition and one subtraction per point. For the first method (not taking into account the initialization) we need \(2(n_1 + 1)\) additions for the computation of \((u_0^0)^{n_1+1}\) and \((2K + 1)\) multiplications and \(2K\) additions per point for the convolution with the upsampled version of the sequence \(d\), \([d]_{\uparrow a}\). Then, for one point, one needs \((2(n_1 + 1) + 2K)\) additions and \((2K + 1)\) multiplications in total.

In the second method (not taking into account the initialization step either) we need \((2K + n_1 + 2)(n_1 + n_2 + 2)\) multiplications and \((2K + n_1 + 1)(n_1 + n_2 + 1)\) additions per point, corresponding to the filtering of \(g\) with the weights \(w_a\).

Thus, the fast implementation of Unser et al. outperforms ours for integer scales and is therefore preferable in this particular setting. Our method, on
the other hand, is more versatile and can be used for arbitrary scale progressions.

6. Conclusions

We have presented a novel B-spline-based CWT algorithm that is able to compute the CWT at any real scale, making it possible to use arbitrary scale progressions. The computational complexity per computed coefficient is O(1), as is the case with the most efficient wavelet algorithms for dyadic or integer scales. The overall operation count only depends on the wavelet shape and on the degrees of the B-spline basis on which the wavelet and the input signal are described, but is independent of the value of the scale. Thanks to the good approximation properties of B-splines, virtually any wavelet can be used (either via their B-spline interpolation or projection). Moreover, the algorithm lends itself well to a parallel implementation as it is not iterative across scales.

The price to pay for the generality of this algorithm is that the leading constant in the O(N) complexity can be large (typically 56 for a cubic spline Mexican-hat wavelet and cubic interpolation of the input signal). For smaller sizes, it may be more efficient to use a simpler FFT-based method. Note, however, that the specialized version of the algorithm for integer scales beats the FFT in all cases.

A demonstration of our method is available on the web at http://bigwww.epfl.ch/demo/cwtspline.

Acknowledgements

This work was funded by the Swiss National Science Foundation, Grant #2100-053540. We also thank the reviewers for their insightful comments.

References


