PYRAMIDAL ALGORITHMS FOR LITTLEWOOD-PALEY DECOMPOSITIONS

M.A. Muschietti*, B. Torrésani **

Abstract: It is well known that with any usual multiresolution analysis of $L^2(\mathbb{R})$ is associated a pyramidal algorithm for the computation of the corresponding wavelet coefficients. It is shown that an approximate pyramidal algorithm may be associated with more general Littlewood-Paley decompositions. Accuracy estimates are provided for such approximate algorithms. Finally, some explicit examples are studied.

May 1993
CPT-93/P.2910

keywords: Wavelets, Quadrature Mirror Filters, Littlewood-Paley decompositions.

*: Departamento de Matemática, UNLP, Argentina
**: Groupement De Recherches Ondelettes, CNRS.
PYRAMIDAL ALGORITHMS FOR LITTLEWOOD-PALEY DECOMPOSITIONS

M.A. Muschietti*, B. Torrésani†

May 1993

Abstract

It is well known that with any usual multiresolution analysis of \( L^2(\mathbb{R}) \) is associated a pyramid algorithm for the computation of the corresponding wavelet coefficients. It is shown that an approximate pyramidal algorithm may be associated with more general Littlewood-Paley decompositions. Accuracy estimates are provided for such approximate algorithms. Finally, some explicit examples are studied.

1 INTRODUCTION

Wavelet analysis has emerged in the past ten years as a completely generic methodology for solving problems in many different areas such as mathematical analysis and operator theory, numerical analysis, signal and image processing, computer vision, computer music, turbulence, astrophysics for instance. Among the advantages of wavelet decompositions, their relative simplicity and the existence of associated fast algorithms are ones of the most important [1] [6].

The main drawback of continuous wavelet decompositions is that there is a priori no associated fast algorithm for the computation of the corresponding wavelet transform. To get a finer analysis, one is interested in keeping redundancy on the wavelet transform, to get a low-pass filter. These filters are in general not discrete, but may in some situations be well approximated by discrete filters, the localization of which can be directly related to the regularity of the scaling function.

The paper is organized as follows. Section 2 is devoted to a description the version of continuous wavelet decompositions that we will use in the sequel. In section 3 we present the algorithmic aspects we are interested in, and in particular our main result (Theorem 4). We present some examples in section 4 and section 5 is devoted to conclusion.

Throughout this paper we shall use the following notations. We shall denote by \( \| \cdot \|_p \) the \( L^p(\mathbb{R}) \) and \( L^p(\mathbb{R}) \) norms. The \( L^p(\mathbb{R}) \) norm is normalized as:

\[ \| f \|_p = \left( \int_{\mathbb{R}} |f(x)|^p dx \right)^{\frac{1}{p}} \]

Our conventions for Hermitian product and Fourier transform in \( L^2(\mathbb{R}) \) are the following ones:

\[ (f,g) = \int_{\mathbb{R}} f(x)g^*(x) dx \]

where the star denotes complex conjugation, and

\[ \hat{f}(\omega) = \int_{\mathbb{R}} f(x)e^{-i\omega x} dx \]

2 CONTINUOUS WAVELET DECOMPOSITIONS

Let us start from standard notions of continuous wavelet analysis. We will focus on the analysis of \( L^2(\mathbb{R}) \), and sometimes describe in a few words the corresponding results in the \( H^2(\mathbb{R}) \) context (we will denote here by \( \hat{H}^2(\mathbb{R}) = \{ f \in L^2(\mathbb{R}), \hat{f}(\xi) = 0 \forall 0 \leq \xi \leq 1 \} \), the complex Hardy space).

We shall be interested here in two "decomposition-reconstruction" schemes, corresponding to different "reconstruction wavelets".

2.1 The bilinear scheme

Generically, an infinitesimal wavelet (or mother wavelet) is a function \( \psi \in L^2(\mathbb{R}) \) such that the following admissibility condition holds:

\[ a_0 = \int_{\mathbb{R}} |\psi(u)|^2 du = \int_{\mathbb{R}} |\hat{\psi}(\omega)|^2 du = 1 \]  (1)

(in such a case, \( \psi \) is in general taken to be a real-valued function). If \( \psi \) is say differentiable, equation 1 basically means that \( \psi(0) = 0 \), otherwise stated:

\[ \int_{\mathbb{R}} \psi(x) dx = 0 \]  (2)

Such a mother wavelet provides the following analysis of \( L^2(\mathbb{R}) \): for any \( (b,a) \in \mathbb{R} \times \mathbb{R}_+^* \), one introduces the wavelet:

\[ \psi_{b,a}(x) = \frac{1}{a} \psi \left( \frac{x-b}{a} \right) \]  (3)

and one has the following representation theorem, the proof of which is well known and can be found for instance in [3] or [2].

Theorem 1 (Calderon) Let \( \psi \) be a mother wavelet. Then any \( f \in L^2(\mathbb{R}) \) decomposes as follows:

\[ f = \int_{\mathbb{R} \times \mathbb{R}_+^*} T_f(b,a) \psi_{b,a} db da \]

\[ T_f(b,a) = \hat{f}(\psi_{b,a}) \]  (4)

strongly in \( L^2(\mathbb{R}) \).
The corollary also holds in the $H^1(R)$ context. Let us now set

$$D_j(x) = \int_{2^{-j} \mathbb{R}} d_2(x) \, \frac{du}{u}.$$  

Then $D_j \in L^2(R)$ represent the details of $f(x)$ visible at scale $2^{-j}$ and not at scale $2^j$.

Introducing the function $\Phi(x)$, such that:

$$\Phi(x) = f(x) \left( \frac{2^j}{a} \right)^ \frac{1}{2}$$

one then has:

$$\tilde{\Phi}_j(x) = \tilde{\Phi}(x) \left( \frac{2^j}{a} \right)^ \frac{1}{2}.$$  

We will refer to the $\Phi(x)$ function as the (global or integrated) wavelet. Notice that equation 14 does not completely define the wavelet $\Phi$. Once again, one can restrict to wavelets with positive-valued Fourier transform, but this is not necessary. By construction, the integrated wavelets lead to a partition of unity in the Fourier space as follows:

$$\sum_{j=\infty}^{a} |\tilde{\Phi}(2^j \xi)|^2 + \sum_{j=-a}^{\infty} |\tilde{\Phi}(2^j \xi)|^2 = 1$$

for all $\xi \in \mathbb{R}$. We also have:

$$|\tilde{\Phi}(\xi)|^2 = |\tilde{\Phi}(0)|^2 - |\tilde{\Phi}(\xi)|^2.$$

Defining the dilates and translates of the $\Phi$ and $\Phi^2$ as

$$\Phi^2(x) = 2^{-j} \Phi(2^{-j}(x - b))$$

one then has:

$$\tilde{\Phi}_j(x) = \tilde{\Phi}(x) \left( \frac{2^j}{a} \right)^ \frac{1}{2}.$$  

we then have:

**Theorem 2** Let $\Psi$ an infinitesimal wavelet, and let $\Phi$ and $\Psi$ be associated integrated wavelet and scaling functions as in equations 14 and 6. Then any $f \in L^2(R)$ can be decomposed as:

$$f = \sum_{j=\infty}^{a} \int_{\mathbb{R}} \langle f, \tilde{\Phi}_j \rangle \Phi_j \, db + \sum_{j=-a}^{\infty} \int_{\mathbb{R}} \langle f, \tilde{\Psi}_j \rangle \Psi_j \, db$$

strongly in $L^2(R)$.

We shall use the following notations

$$T_j f(b) = \langle f, \tilde{\Phi}_j \rangle$$

$$S_j f(b) = \langle f, \tilde{\Psi}_j \rangle$$

2.2 The linear scheme

It is well known that the reconstructing and the analyzing wavelets can be decoupled i.e. one can use different infinitesimal wavelets for the computation of the coefficients and for the reconstruction of the analyzed function from the coefficients. In such a case, the admissibility condition 1 has to be modified accordingly.
A particular example of such a decoupling, which has been known for a long time, consists in taking formally a Dirac distribution for the reconstructing wavelet. Assuming instead of equation 1 that
\[ k_\nu = \int_0^\infty \psi(u) \frac{du}{u} = \int_0^\infty (-\psi(-u)) \frac{du}{u} = 1 , \]
on one has the following decomposition of any \( f \in L^2(\mathbb{R}) \):
\[ f(x) = \int_{\mathbb{R}^d} \langle f, \varphi(x,\omega) \rangle \frac{d\omega}{a} \]
This is the so-called Morlet reconstruction formula of \( f \) from its wavelet coefficients. Such a linear analysis (linear in the \( \psi \) function) generates a continuous multiresolution analysis as follows: introduce the linear scaling function \( \varphi \in L^2(\mathbb{R}) \), defined by:
\[ \varphi(t) = \int_{\mathbb{R}} \psi(u \omega) \frac{du}{u} \]
\( \varphi \) is also such that \( \varphi(u \omega) = -u \varphi(u \omega) \) for all \( \omega \in \mathbb{R} \). Associate to \( \varphi \) the following functions:
\[ \varphi_n(x) = \varphi(x-n) \]
Finally introduce:
\[ \delta_n(x) = \varphi_n(x,a) = (f, \varphi_n(x,\omega)) \]
and
\[ \sigma_n(x) = (f, \varphi_n(x,\omega)) \]
One then has the linear analogue of Theorem 1 and the corresponding corollary:

**Theorem 3** Let \( \psi \in L^2(\mathbb{R}) \) be a mother wavelet, such that equation 21 holds, and let \( \varphi \) be the associated linear scaling function. Then any \( f \in L^2(\mathbb{R}) \) can be decomposed as:
\[ f = \lim_{n \to \infty} (f, \varphi_n(x,\omega)) \frac{d\omega}{a} \]
\[ = \varphi_0 + \sum_{n \in \mathbb{Z}} \delta_n(x) \]
\[ = \varphi_0 + \sum_{n \in \mathbb{Z}} \theta_n(x) \]
where \( \delta_n(x) = (f, \varphi_n(x,\omega)) \) and \( \theta_n(x) = (f, \varphi_n(x,\omega)) \).

### 3 ASSOCIATED APPROXIMATE FILTERS

#### 3.1 Pyramid algorithms

Let us recall the usual algorithmic structure associated with a multiresolution analysis. Let \( \varphi \) and \( \psi \) be respectively a scaling function and a wavelet associated with the multiresolution analysis, and set, \( f \in L^2(\mathbb{R}) \):
\[ T_j(f)(n) = (f, \varphi_n) \]
\[ S_j(f)(n) = (f, \varphi_{n-1}) \]
Notice that for any value of the scale parameter \( a = 2^j \), we sample the corresponding wavelet and scaling function transform at unit sampling frequency.

Then if \( \varphi \) and \( \psi \) are related by
\[ \psi(t) = m_0(t) \varphi(t) \]
\[ \varphi(t) = m_1(t) \varphi(t) \]
where \( m_0 \) and \( m_1 \) are the \( 2 \pi \)-periodic low-pass and high-pass filters
\[ m_0(t) = \sum_{k \in \mathbb{Z}} \hat{h}_k e^{i \omega t} \]
\[ m_1(t) = \sum_{k \in \mathbb{Z}} \hat{g}_k e^{i \omega t} \]
then the coefficients may be computed using the following pyramidal algorithm
\[ T_j(f)(n) = \sum_{k \in \mathbb{Z}} \hat{h}_k S_{j-1}(f)(n-k 2^j) \]
\[ S_j(f)(n) = \sum_{k \in \mathbb{Z}} \hat{g}_k T_{j-1}(f)(n-k 2^j) \]
The algorithm is called pyramidal since scaled copies of the same filters are used throughout the calculation, and the coefficients are obtained by successive convolutions with such filters. It is easy to see that the total number of multiplications necessary to process \( N \) samples of any \( S_jf \) is proportional to \( N \log(N) \). It is schematically described in figure 1 (in the particular situation where the \( m_0(t) \) filter has only 3 non vanishing coefficients).

#### 3.2 Approximate filters

We address now the problem of discretization of the previous wavelet decompositions. Up to now, we have only obtained decompositions that are discrete with respect to the scale, and continuous with respect to the position. The problem is that no discrete filters are "a-priori" available.
We shall work with both the linear and the bilinear analysis-reconstruction schemes at the beginning, and specify our choice later on. We shall assume from now on that a pair of functions
\[ m_1(e) = \frac{\hat{\phi}(e)}{\| \hat{\phi} \|_2} \]  
\[ m_2(e) = \frac{\hat{\psi}(e)}{\| \hat{\psi} \|_2} \]
can be defined almost everywhere in \( \mathbb{R} \). This is clearly the case in the bilinear scheme, where \( [\hat{\phi}] \) is monotoneous for both \( \xi \geq 0 \) and \( \xi \leq 0 \). The problem is that in general, such a \( m_2 \)-function is not 2\( \pi \)-periodic, and thus cannot be used in a pyramid algorithm.

Nevertheless, a modification is possible. Indeed, in \( \mathbb{R} \), it is concentrated around the origin, in say the interval \([{-\pi}, \pi]\), then one may expect that \( \phi \) "does not see too much" the non-2\( \pi \)-periodicity of \( m_2(e) \), and that \( m_1(e) \hat{\phi}(e) \) can be well approximated by \( m_2(e) \hat{\psi}(e) \), for some 2\( \pi \)-periodic function \( \hat{\psi}(e) \).

Let us introduce here for convenience the following subspace of \( L^2(\mathbb{R}) \):
\[ \mathcal{V}_0 = \{ f \in L^2(\mathbb{R}) : f = \sum \alpha_n \phi(x-n), \ (\alpha_n) \in \ell^2(\mathbb{Z}) \} \]

We will assume that the collection \( \{ \phi(x-k), k \in \mathbb{Z} \} \) is a Riesz basis of \( \mathcal{V}_0 \), or, equivalently, there exist two finite and nonzero constants \( A \) and \( B \) such that
\[ A \leq \sum_n |\hat{\phi}(x+2n\pi)|^p \leq B \text{ almost everywhere} \]

Then it follows from general results that there exists a function \( \phi \) in \( L^2(\mathbb{R}) \) such that the sequence \( \{ \phi(x-k), k \in \mathbb{Z} \} \) is the biorthogonal basis of \( \{ \hat{\phi}(x-k), k \in \mathbb{Z} \} \). \( \chi \) is given by its Fourier transform
\[ \hat{\chi}(e) = \sum_{n} |\hat{\phi}(e+2n\pi)|^2 \]

We will denote by \( \mathcal{V}_0 \) the biorthogonal space of \( \mathcal{V}_0 \), i.e. the closed linear span of \( \{ \chi(x-k), k \in \mathbb{Z} \} \).

Consider now the discretization of the functions \( T_1(f) \) and \( S_1(f) \), which we shall denote by \( T_1^e(n) \) and \( S_1^e(n) \) respectively,
\[ S_1^e(n) = S_1(n) \forall n \in \mathbb{Z} \]
\[ T_1^e(n) = T_1(n) \forall n \in \mathbb{Z} \]

Let \( m_1(e) \in L^2([-\pi, \pi]) \) and \( m_2(e) \in L^2([-\pi, \pi]) \) be two (2\( \pi \)-periodic) candidates for approximate filters, and denote by \( \{ \phi_n(k), n \in \mathbb{Z} \} \) and \( \{ \psi_n(k), n \in \mathbb{Z} \} \) their respective Fourier coefficients. We will then set:
\[ T_1^e(n) = \sum_k \phi_1^e(n-k) \]
\[ S_1^e(n) = \sum_k \psi_1^e(n-k) \]
and
\[ T_2^e(n) = \sum_k \phi_2^e(n-k) \]
\[ S_2^e(n) = \sum_k \psi_2^e(n-k) \]

Our purpose is to compare such "algorithmic expressions" with the exact expressions \( S_1^e(f) \) and \( T_1^e(f) \), and to find "best approximants" for the \( m_1 \) and \( m_2 \) filters.

The first remark is the following.

Proposition 1
1. \( \text{Ker}(S_1^e) = \{0\} \).
2. \( S_1^e \cdot U_0 = T_1^e \cdot U_0 = 0 \) for any \( j = 1, \ldots \).

Proof: The first part is a direct consequence of the definition of \( U_0 \), and implies the second part by definition 43.

Our main result is the following.

Theorem 4
1. There exists a unique pair of 2\( \pi \)-periodic filters \( m_1(e) = \phi_1(e) \) minimizing \( \|\phi_1 - m_1\|_2 \), given by:
\[ m_1(e) = \frac{\sum_{k} \phi_1^e(n-k) \hat{\phi}(e+2n\pi)}{\sum_{k} \| \hat{\phi}(e+2n\pi) \|^2} \]
\[ m_1(e) = \frac{\sum_{k} \phi_1^e(n-k) \hat{\phi}(e+2n\pi)}{\sum_{k} \| \hat{\phi}(e+2n\pi) \|^2} \]

2. For the above choice of filters, and setting
\[ C_i = \max_{x \in [0,1]} |\phi_i(x)|, \ i = 0, 1 \]

the following inequalities hold:
\[ \|S_1^e f - S_1 f\|_2 \leq \frac{\|m_1 - m_2\|_2}{\|\phi_1\|_2} \left( 1 - \frac{C_1 \sqrt{2}}{\|\phi_1\|_2} \right) \|f\|_2 \]
\[ \|T_1^e f - T_1 f\|_2 \leq \frac{\|m_1 - m_2\|_2}{\|\phi_1\|_2} \left( 1 - \frac{C_1 \sqrt{2}}{\|\phi_1\|_2} \right) \|f\|_2 \]

3. For any \( f \in \mathcal{V}_0 \cup \mathcal{U}_0 \),
\[ S_1 f = S_1^e f \]
\[ T_1 f = T_1^e f \]

Before giving the proof of the theorem, let us give the following immediate corollary:

*The assumption is motivated by the fact that we will sample the wavelet transform and the scaling function transform at unit sampling frequency

*A natural candidate for \( m_2(e) \) is the periodicization \( \sum k \phi_2^e(n-k) \) of \( m_2(e) \), but as we shall see, there are many other choices. 
Corollary 2 Assume that the infinitesimal wavelet \( \psi(x) \) is associated with a linear "analysis-reconstruction" scheme. Then the associated approximate QMFs \( \psi_0(x) \) given by equations (43 and 44) satisfy

\[ m_0(x) + m_1(x) = 1 \]  

so that we have the "reconstruction algorithm"

\[ S_0(f)(n) = \sum_{k \geq 0} T_k^*(n) \]  

This reconstruction formula is the discrete counterpart of Morlet's reconstruction formula 22. It was also obtained in a slightly different context by Beylkin and Saito [5].

Proof of the theorem: Using the inequality \( \| f \|_1 \leq \| f \|_2 \sqrt{2} \), we shall directly work in the Fourier space. First of all, we clearly have

\[ \| \hat{f} - \hat{f}_a \|_2^2 \leq \sum_{k \geq 0} k^2 \| \hat{m}(k) - m_k(x + 2k) \| \| \hat{\psi} (x + 2k) \|_2 \leq \sum_{k \geq 0} k^2 \| \hat{m}(k) \|_2 \| \hat{\psi} (x + 2k) \|_2 \]

\[ \leq 2\pi \| f \|_2 \]  

or, otherwise stated

\[ \int_{\mathbb{R}} \hat{m}(k) \| \hat{\psi}(x + 2k) \|_2^2 \leq 2\pi \| f \|_2 \]  

The unique solution is precisely that given in equation 45. The estimation of \( \| \hat{f} - \hat{f}_a \|_2 \) is completely similar, and leads to

\[ \int_{\mathbb{R}} \hat{m}(k) \| \hat{\psi}(x + 2k) \|_2^2 \leq 2\pi \| f \|_2 \]  

or, otherwise stated

\[ \int_{\mathbb{R}} \hat{m}(k) \| \hat{\psi}(x + 2k) \|_2^2 \leq 2\pi \| f \|_2 \]  

Assume now that \( f \in \mathcal{L}^2 \). This means that in the Fourier space, \( f \) is of the form

\[ \hat{f}(k) = F(k) \]  

Then \( \| f \|_1 = \| f \|_2 \) and \( \| f \|_2 \) yields

\[ \| \hat{f} - \hat{f}_a \|_2 = \| \hat{m}(k) \|_2 \| \hat{\psi}(x + 2k) \|_2 \int_{\mathbb{R}} F(k) \]  

by definition of the \( \chi \) function (see equation 40 above). This concludes the proof of the theorem.

Remark: asymptotic behaviour:

It is interesting to see what is the asymptotic behaviour of the estimates when \( j \to \infty \). Consider for instance the estimate of \( \| \hat{f}_j - \hat{f}_a \|_2 \), the coefficient of \( \| f \|_2 \) goes as:

\[ \| \hat{f}_j - \hat{f}_a \|_2 \leq \| f \|_2 \]  

(61)

The second term is estimated as follows

\[ \| \hat{f}_j - \hat{f}_a \|_2 \leq \int_{\mathbb{R}} \| m_0(2^{-j-1}x) \| \| \hat{\psi}(x + 2k) \|_2 \| f \|_2 \]  

\[ \leq \int_{\mathbb{R}} \| m_0(2^{-j-1}x) \| \| \hat{\psi}(x + 2k) \|_2 \]  

(62)

The same kind of estimates yields the error estimate for the \( \mathcal{T}_f^* \) coefficients. This achieves the proof of the two first items of the theorem.

Let us turn to the third part of the theorem. Let us then assume that \( f \in \mathcal{L}^2 \). This means that in the Fourier space, \( f \) is of the form

\[ \hat{f}(k) = F(k) \]  

(63)

The same computation as before yields

\[ \| \hat{f}_j - \hat{f}_a \|_2 = \| \hat{m}(k) \|_2 \| \hat{\psi}(x + 2k) \|_2 \int_{\mathbb{R}} F(k) \]  

(64)

and \( \{g_k \} \) approximate filters can be directly related to the regularity properties of the scaling function, as follows.

3.3 Decay of approximate filter coefficients

The localization properties of the \( \{h_k \} \) and thus \( \{g_k \} \) approximate filters can be directly related to the regularity properties of the scaling function, as follows.
Theorem 5 \textit{Let } \phi \in L^1(\mathbb{R}) \textit{ be a } p \textit{-times differentiable scaling function, and let } n_0(t) = \sum_{k} h_k(t) \textit{ be the low-pass filter defined according to equation 45. Then if for any } m = 0, 1, \ldots, p \textit{ \begin{equation} \left| \frac{d^m \tilde{\phi}(t)}{dt^m} \right| \leq \frac{K_m}{(1 + |t|)^{h_m}} \end{equation} \textit{ for some positive constants } K_m, \textit{ then } n(t) = n_0(t) \end{equation}} \textit{as high-pass filter. One has in such a case to be careful with the zeros of the } n_0(t) \textit{ filter. As an alternative, the same kind of analysis as before can be performed in the bilinear "analysis-reconstruction" scheme. The previous arguments have to be applied to the details and approximations instead of the wavelet coefficients themselves.} \begin{equation} \tilde{\phi}_t(t) = \tilde{\phi}(2t) \end{equation}

Proof: \textit{Assume that } \tilde{\phi}(t) \textit{ is } p \textit{-times differentiable. Then after } p \textit{ differentiations, equation 45 yields} \begin{equation} \frac{d^m \tilde{\phi}_t(t)}{dt^m} = \sum_k G(t + 2k) \frac{d^m \tilde{\phi}(t)}{dt^m} \end{equation} \textit{where } G(t) \textit{ is a finite linear combination of terms of the form} \begin{equation} \frac{d^m \tilde{\phi}(t)}{dt^m} \frac{d^m \tilde{\phi}(t)}{dt^m} \end{equation} \textit{and their complex conjugates. Then the estimate 56 gives} \begin{equation} \frac{d^m \tilde{\phi}_t(t)}{dt^m} \in L^\infty([-2t, 0)] \end{equation} \textit{and} \begin{equation} \frac{d^m \tilde{\phi}_t(t)}{dt^m} \in C(\mathbb{R}) \end{equation} \textit{Moreover,} \begin{equation} \frac{d^m \tilde{\phi}_t(t)}{dt^m} \textit{ and } \frac{d^m \tilde{\phi}_t(t)}{dt^m} \textit{ and similarly} \end{equation} \textit{It is worth noticing that in such a case, the bilinear scheme is well suited for this pair of filters, and ensures the validity of the usual QMF relation:} \begin{equation} \left| \tilde{\phi}_t(t)^2 + \tilde{\phi}_t(t)^2 \right| = 1 \end{equation} \textit{Moreover, it is easy to derive the "bilinear counterpart" of theorem 5, relating the length of the approximate filters to the regularity of the scaling function.} \begin{equation} n_0(t) = n_0(t) \textit{ as high-pass filter, and} \end{equation} \textit{as high-pass filter. One has in such a case to be careful with the zeros of the } n_0(t) \textit{ filter. As an alternative, the same kind of analysis as before can be performed in the bilinear "analysis-reconstruction" scheme. The previous arguments have to be applied to the details and approximations instead of the wavelet coefficients themselves.} \begin{equation} \tilde{\phi}_t(t) = \tilde{\phi}(2t) \end{equation}

Using again approximate filters to evaluate the coefficients, one is then naturally led to the quantity: \begin{equation} \tilde{\phi}_t^2(t) = \left| n_0(t)^2 + n_1(t)^2 \right| dt \end{equation}

The minimization of such a quantity naturally leads to \begin{equation} \left| \tilde{\phi}_t(t)^2 + \tilde{\phi}_t(t)^2 \right| = 1 \end{equation} \textit{Moreover, it is easy to derive the "bilinear counterpart" of theorem 5, relating the length of the approximate filters to the regularity of the scaling function.} \begin{equation} n_0(t) = n_0(t) \textit{ as high-pass filter, and} \end{equation} \textit{as high-pass filter. One has in such a case to be careful with the zeros of the } n_0(t) \textit{ filter. As an alternative, the same kind of analysis as before can be performed in the bilinear "analysis-reconstruction" scheme. The previous arguments have to be applied to the details and approximations instead of the wavelet coefficients themselves.} \begin{equation} \tilde{\phi}_t(t) = \tilde{\phi}(2t) \end{equation}

Using again approximate filters to evaluate the coefficients, one is then naturally led to the quantity: \begin{equation} \tilde{\phi}_t^2(t) = \left| n_0(t)^2 + n_1(t)^2 \right| dt \end{equation}

The minimization of such a quantity naturally leads to \begin{equation} \left| \tilde{\phi}_t(t)^2 + \tilde{\phi}_t(t)^2 \right| = 1 \end{equation} \textit{Moreover, it is easy to derive the "bilinear counterpart" of theorem 5, relating the length of the approximate filters to the regularity of the scaling function.}
3. Assume that we are in the case of a scaling function with exponential decay in the Fourier space (i.e. \( \hat{\Phi}(e) \leq C e^{-ae} \) for some positive \( a \)). Then it is not very difficult to show that (in the case of a unit sampling frequency) the approximation of the filters obtained by sampling the inverse Fourier transform of \( \Phi \) leads to an error on the scaling function coefficients of the order \( e^{-a} \). In the same way, defining the approximate \( 2\pi \)-periodic high-pass filter by the QMF-relation leads to the same kind of error estimate for the wavelet coefficients.

4. Obviously, it follows from the expressions of the approximate filters (both in the linear and bilinear schemes) that if \( \Phi \) and \( \Psi \) are associated with an usual multiresolution analysis, with \( 2\pi \)-periodic filters, one recovers \( m_0 = m_0 \) and \( m_1 = m_1 \).

5. It was shown in [2] how to use Calderon's formula to get descriptions of the Fourier space different from the Littlewood-Paley one, by replacing the powers of 2 by an arbitrary monotonic sequence of scale parameters. It sounds reasonable to think of corresponding approximate algorithms similar to the one described above, at least for rational scale parameters. However this has not been done at the present time.

4 EXAMPLES

There are many examples of continuous wavelets for which an efficient algorithm is needed. We describe here some very simple examples (the filters coefficients have been computed using the Mathematica Package).

4.1 The LOG and DOG wavelets

The LOG wavelets are widely used in the context of computer vision. LOG stands for "Laplacian Of Gaussians". As stressed in [3], in the linear "decomposition-reconstruction" scheme, if

\[
\psi(x) = \frac{1}{\sqrt{2}}(1 - x^2)e^{-x^2/2}
\]

the associated scaling function and integrated wavelet are given by

\[
\theta(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}
\]

and

\[
\theta(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2 - 2xe^{-x^2/2}}
\]

The integrated wavelet is then a DOG wavelet ("Difference Of Gaussians"), and there is no problem to derive the details coefficients \( T_f \) from the approximations \( S_f \). But one clearly needs an efficient algorithm to compute the approximations. In general, the scaling function and wavelet have to be scaled properly in order that the corresponding transforms can be accurately sampled at unit sampling frequency. We shall then consider more general scaling functions

\[
\psi(x) = e^{-\alpha x^2}
\]

with the corresponding integrated wavelets. We give as examples the plots of the approximate low-pass filters \( \theta_\alpha(x) \) (the high-pass filter \( \theta_\alpha(x) \) is easy to deduce), and the coefficients of \( \theta_\alpha(x) \) and \( \phi_\alpha(x) \) for \( \alpha = 4 \) and \( \alpha = 6 \) respectively. It is worth noticing that in both cases (and in fact for any positive \( \alpha \)), the \( \{a_n\} \) and \( \{g_k\} \) sequences are rapidly decreasing, as a consequence of theorem 5.

4.1.1 \( \alpha = 4 \)

\[
\begin{align*}
\text{Figure 2: Approximate low-pass filter for the DOG wavelet with } & \alpha = 4 \\
a_n \text{ coefficients:} & \\
& \{0.3256232760276189, 0.2334898320421699, 0.085798244731082, -0.01625749135447542, -0.00154899267327945, -0.0000922406284933434, 6.3324e-10, 5.04664e-6, -3.0372e-6, 1.84053e-6, -1.1595e-7, 6.76772e-7, -4.10463e-7, 2.48954e-7, -1.50997e-7, 9.15844e-8\}
\end{align*}
\]

\[
\begin{align*}
\text{Figure 3: Approximate low-pass filter for the DOG wavelet with } & \alpha = 6 \\
g_k \text{ coefficients:} & \\
& \{0.6743672599723812, -0.2334898320421699, -0.085798244731082, -0.01625749135447542, -0.00154899267327945, -0.0000922406284933434, 6.3294e-10, -5.04664e-6, 3.0372e-7, -1.84053e-6, 1.1595e-7, -6.76772e-7, 4.10463e-7, -2.48954e-7, 1.50997e-7, -9.15844e-8\}
\end{align*}
\]

4.1.2 \( \alpha = 6 \)

\[
\begin{align*}
\text{Figure 2: Approximate low-pass filter for the DOG wavelet with } & \alpha = 6 \\
a_n \text{ coefficients:} & \\
& \{0.3256232760276189, 0.2334898320421699, -0.085798244731082, -0.01625749135447542, -0.00154899267327945, -0.0000922406284933434, 6.3294e-10, -5.04664e-6, 3.0372e-7, -1.84053e-6, 1.1595e-7, -6.76772e-7, 4.10463e-7, -2.48954e-7, 1.50997e-7, -9.15844e-8\}
\end{align*}
\]

4.1.2 \( \alpha = 6 \)
The corresponding $h_i$ coefficients read

\[
\{0.3972771041574456, 0.2433136948393599, 0.05323383400865474, 0.004788125238023543, -0.00002934485917262438, 0.00007751303939803943, -0.00003576989742401233, 0.00001689424764443, -7.97394 \times 10^{-8}, 3.7664510^{-6}, -1.8751810^{-7}, 8.4039810^{-8} \}
\]

and the $g_i$ coefficients are

\[
\{0.6027228958425545, -0.2433136948393598, -0.05323383400865478, -0.004788125238023527, 0.00002934485917263695, \\
-0.00007751303939803943, 0.00003576989742401233, 0.00001689424764447, -7.97394 \times 10^{-8}, 3.7664510^{-6}, -1.8751810^{-7}, 8.4039810^{-8} \}
\]

4.1.3 Precision of the algorithm:

As we have seen, the estimate of the accuracy of the approximate algorithm is governed by the functional $p(m_0, r_0)$. We present here numerical estimation of this quantity for the DOG wavelets, for various values of the $\alpha$ parameter. For instance, $p(m_0, r_0) = 1.03632 \times 10^{-17}$ for $\alpha = 1$, $p(m_0, r_0) = 1.63909 \times 10^{-6}$ for $\alpha = 3$, and $p(m_0, r_0) = 0.0012859$ for $\alpha = 6$. Figure 4 represents the logarithm of $p(m_0, r_0)$ as a function of $\alpha$.

![Figure 4: Logarithm of $p(m_0, r_0)$ as a function of $\alpha$](image)

4.2 "Exponential-type" wavelets

These wavelets are real-valued wavelets, characterized by their exponential decay in the Fourier space. Let

\[
\bar{\psi}_n(t) = \frac{1}{(n - 1)!} e^{-\beta |t|^\alpha}
\]

$n = 1, \ldots \infty$ then controls the number of vanishing moments, and $\bar{\psi}_n(t)$ has exponential decay for all $n$. A direct computation yields the corresponding scaling function

\[
\psi_n(t) = e^{-\beta |t|^\alpha} \sum_{k=0}^{n-1} \frac{1}{k!} \left( \frac{\beta |t|^\alpha}{k!} \right)^k
\]

Notice that $\bar{\psi}_n(t) \sim_{n \to \infty} 1 + O(|t|^\alpha)$. Then $\bar{\psi}_n \in C^{n-1}([0,\infty)$ and $\bar{\psi}_n (|d^n\bar{\psi}_n|)$ has exponential decay $\forall m = 0, \ldots, n - 1$, which implies that $h_n = O(|t|^\alpha)$. The integrated wavelets are easy to deduce, and the associated low-pass filter is represented in figure 5, in the case $n = 1$ (with $\alpha = .3$).

![Figure 5: Approximate low-pass filter for "exponential-type" wavelet with 1 vanishing moment](image)

The case $n = 1$ is not very interesting numerically because $\psi$ is not differentiable at $t = 0$ and the $m_0$ filter has slow decay. We shall then show the case $n = 5$, $\alpha = .3$ for which the low-pass filter is shown in figure 6.

![Figure 6: Approximate low-pass filter for "exponential-type" wavelet with 5 vanishing moments](image)

and the 16 top low-pass and high-pass filter coefficients are given by:

$h_n$ coefficients:

\[
\{0.3689009, 0.2119350, 0.1180216, 0.047464751, 0.012065044, -0.001261182, -0.004691885, -0.004500855, -0.0034217616, -0.0023426834, -0.0015318216, -0.00096772754, -0.00060530854, -0.00037287531, -0.00023103465, -0.00014240093 \}
\]

$g_n$ coefficients:

\[
\{0.7391091, -0.2119350, -0.1180216, -0.047464751, -0.012065044, 0.001261182, 0.004691885, 0.004500855, 0.0034217616, 0.0023426834, 0.0015318216, 0.00096772754, 0.00060530854, 0.00037287531, 0.00023103465, 0.00014240093 \}
\]

It is worth noticing that all such coefficients are easy to obtain numerically.
4.3 The Cauchy wavelets

The same filters as before can be used to work with the wavelets that were used by T. Paul in a quantum mechanical context (they are canonically associated with the radial Schrödinger equation for the Hydrogen atom for instance. They are of the form

$$\varphi_n(\xi) = \begin{cases} \frac{(-1)^{n-1}}{n!} \xi^{n-1} & \text{for positive values of } \xi \\ 0 & \text{otherwise} \end{cases} \quad (93)$$

$n = 1, \ldots$ controls the number of vanishing moments, and $\varphi_n(\xi)$ has exponential decay for all $\xi$. A direct computation yields the corresponding scaling function

$$\psi_n(\xi) = \begin{cases} (-1)^{n-1} \sum_{p=0}^{n-1} \frac{\partial^p}{\partial \xi^p} \xi^n & \text{for positive values of } \xi \\ 0 & \text{otherwise} \end{cases} \quad (94)$$

The first case is particularly interesting since it is canonically related with the Cauchy kernel. Indeed, the scaling function coefficients of a function $f(\xi) \in H^2(\mathbb{R})$ form an analytic function of $\xi = b + ia$, that is the analytic continuation of $f(\xi)$ to the upper half-plane. The corresponding wavelet transform is then (up to a factor $a$) the derivative of $f(\xi)$ with respect to its imaginary part:

$$\gamma_{(b,a)} f(\xi) = -a \frac{d}{d\xi} f(b+ia).$$

The previous filters can then be used to get a fast approximate algorithm for wavelet transform with such wavelets. Of course, adapted filters can also be obtained by using directly the formula yielding the $\delta_0$ filters. However, $\delta_0$ being discontinuous at the origin for any $n$, such filters are not suitable for numerical use since they have slow decay.

5 CONCLUSIONS

We have described in this paper a method that associates a pair of ($2\pi$-periodic) filters with a Littlewood-Paley (or dyadic wavelet) decomposition, yielding a pyramidal algorithm for the computation of a corresponding approximate transform.

We have in particular shown that in the case where the Littlewood-Paley decomposition comes from a "linear scheme" of infinitesimal wavelet analysis (as described in [2]), such filter fulfill a kind of "linear" QMF relation, leading to simple reconstruction formulæ from the approximate coefficients. Our main result was an estimate of the accuracy of the approximate algorithm. The problem of finding approximate filters was transformed into a minimization problem having a unique solution. Of course when there already exist a pair of $2\pi$-periodic filters naturally associated with the wavelet, this solution coincides with it.

As in the case of usual multiresolution analysis, the localization (i.e. decay properties) of the approximate filters is directly related to the regularity of the scaling function.

We also have discussed some simple examples, in particular those of the LOG and DOG wavelets in the linear scheme, familiar to computer vision specialists, and wavelets of exponential type. If the corresponding scaling functions are sufficiently well localized in the Fourier space, good error estimates are obtained.

Let us notice that $n$-dimensional generalizations of our method with the tensor-product construction of filters are straightforward.

Acknowledgements

This research was partly supported by a CNRS/CNRS/CT Grant, and the Physics department of Université de Provence, Aix-Marseille. We also wish to thank A. Grossmann and Ph. Tchamitchian, for stimulating discussions.

References