# Multiselective Pyramidal Decomposition of Images: Wavelets with Adaptive Angular Selectivity

Laurent Jacques[*]

Jean-Pierre Antoine[†]

(Updated December 13th, 2006)

### Abstract

Many techniques have been devised these last ten years to add an appropriate directionality concept in decompositions of images from the specific transformations of a small set of atomic functions. Let us mention for instance works on directional wavelets, steerable filters, dual-tree wavelet transform, curvelets, wave atoms, ridgelet packets, . . . In general, features that are best represented are straight lines or smooth curves as those defining contours of objects (e.g. in curvelets processing) or oriented textures (e.g. wave atoms, ridgelet packets). However, real images present also a set of details less oriented and more isotropic, like corners, spots, texture components, . . . This paper develops an adaptive representation for all these image elements, ranging from highly directional ones to fully isotropic ones. This new tool can indeed be tuned relatively to these image features by decomposing them into a Littlewood-Paley frame of directional wavelets with variable angular selectivity. Within such a decomposition, 2D wavelets inherit some particularities of the biorthogonal circular multiresolution framework in their angular behavior. Our method can therefore be seen as an angular multiselectivity analysis of images. Two applications of the proposed method are given at the end of the paper, namely, in the fields of image denoising and $N$-term nonlinear approximation.

## 1  Introduction

Real world images contain very different features, ranging from very oriented ones, like straight edges, to more isotropic objects, like corners and spots. Between these two extreme behaviors, we find, for instance, curves with variable curvature radius and texture points. During the last ten years, many techniques have been designed for obtaining good representations of images with oriented features. We may quote, e.g., the work of Antoine, Murenzi, and Vandergheynst on directional wavelets [3, 4], that of Simoncelli *et al.* on oriented steerable filters [5, 6], that of Candès, Donoho and Flesia on the curvelet representation and the ridgelet packets [8, 9, 11], the wave atoms of Demanet and Ying [10], and the dual tree wavelet transform of Kingsbury, Selesnik and Baraniuk [24, 25].

---

[*]Laurent.Jacques@uclouvain.be (corresponding author), Laboratoire de Télécommunications et Télédétection (TELE), Université catholique de Louvain, Place du Levant 2. B-1348 Louvain-la-Neuve, Belgium.

[†]Antoine@fyma.ucl.ac.be, Institut de Physique Théorique (FYMA), Université catholique de Louvain, Chemin du Cyclotron 2. B-1348 Louvain-la-Neuve, Belgium

In this paper, we propose a method based on directional wavelet decomposition for representing the various objects mentioned above *adaptively* on each point of the image. In order to achieve this result, we will define particular directional wavelets based on a circular multiresolution analysis, whose main property is to combine with each other to form directional frames of different angular selectivity (actually, half-continuous frames, that is, discretized only in scales and angles). A criterion to select the "Best Directional Frame" is then defined.

Finally, we will present in two applications, namely image denoising and $\tau\%$-term nonlinear approximation (a simple variant of the $N$-term nonlinear approximation [23]), the advantages of the composite multiselective frame over a non-adaptive fixed selectivity method. Beyond this final comparison, this paper does not claim, however, to replace state of the art techniques tuned to these specific applications.

## 2   The Framework

We describe an *image* by a real function $f$ on $\mathbb{R}^2$, representing the *light intensity* at every point $\vec{x} \in \mathbb{R}^2$ of the image. If $f \in L^2(\mathbb{R}^2) = \{g(\vec{x}) : \|g\|^2 = \int_{\mathbb{R}^2} |g(\vec{x})|^2 \mathrm{d}^2\vec{x} < \infty\}$ and given an admissible wavelet $\psi \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$, the continuous wavelet transform [13] of $f$ is

$$\mathcal{W}_f(\vec{b}, a, \theta) = \langle \psi_{\vec{b},a,\theta} \,|\, f \rangle \tag{2.1}$$

$$= \frac{1}{a^2} \int_{\mathbb{R}^2} \psi^*(a^{-1}r_\theta^{-1}(\vec{x} - \vec{b})) \, f(\vec{x}) \, \mathrm{d}^2\vec{x} \tag{2.2}$$

$$= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \hat{\psi}^*(ar_\theta^{-1}\vec{k}) \, \hat{f}(\vec{k}) \, e^{i\vec{k}\cdot\vec{b}} \, \mathrm{d}^2\vec{k}, \tag{2.3}$$

where $^*$ is the complex conjugation and $\psi_{\vec{b},a,\theta}$ is an $L^1(\mathbb{R}^2)$-normalized copy of $\psi$, translated by $\vec{b} \in \mathbb{R}^2$, dilated by $a \in \mathbb{R}_+$ and rotated by $\theta \in S_1 \simeq [0, 2\pi)$. In these equations, $r_\theta$ is the usual $2 \times 2$ rotation matrix, and the hat denotes the standard Fourier transform on $L^2(\mathbb{R}^2)$, that is, $\hat{f}(\vec{k}) = \int_{\mathbb{R}^2} f(\vec{x}) \, e^{-i\vec{k}\cdot\vec{x}} \, \mathrm{d}^2\vec{x}$, with inverse $f(\vec{x}) = (2\pi)^{-2} \int_{\mathbb{R}^2} \hat{f}(\vec{k}) \, e^{i\vec{k}\cdot\vec{x}} \, \mathrm{d}^2\vec{k}$. In the sequel, we will omit the index of $\mathcal{W}_f$ and write simply $\mathcal{W}$.

From now on, we take for $\psi$ a *directional* wavelet [2, 3, 4, 13], that is, a wavelet whose frequency support is essentially contained in a convex cone with apex at the origin. In addition, we assume that $\psi$ is separable in polar coordinates:
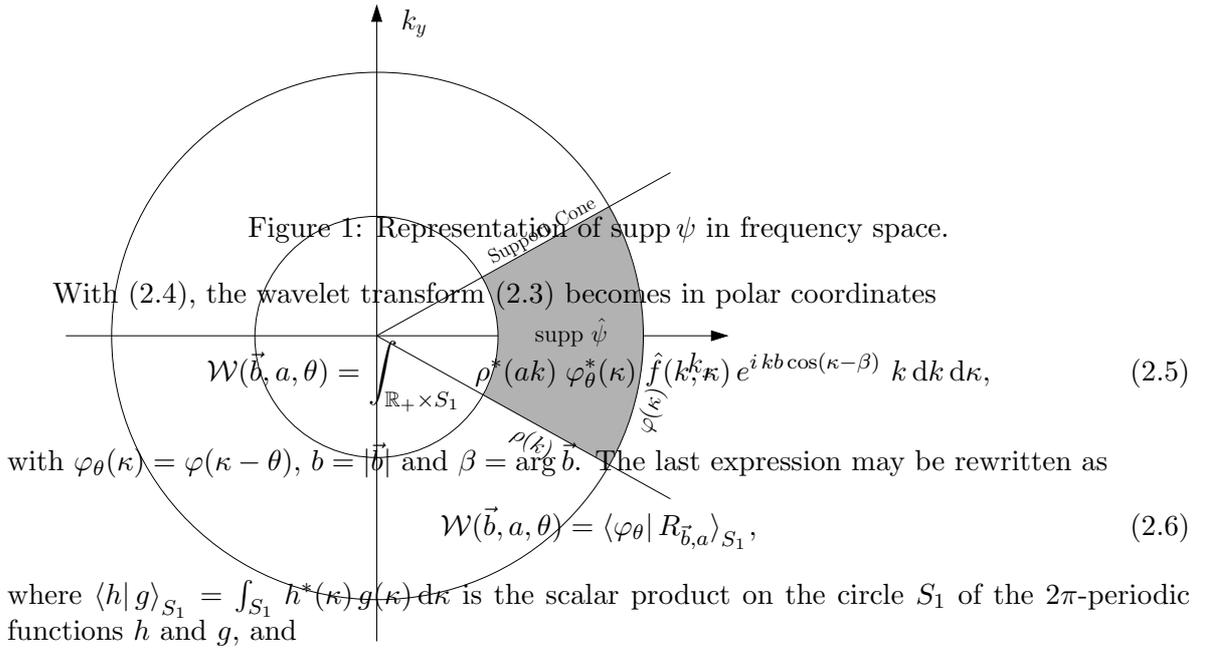
$$\hat{\psi}(\vec{k}) = \rho(k)\, \varphi(\kappa), \tag{2.4}$$

where $\vec{k} \equiv (k, \kappa)$, $k = |\vec{k}|$, $\kappa = \arg \vec{k}$, $\rho$ is a function in $L^2(\mathbb{R}, k\mathrm{d}k)$ and $\varphi$ is a positive function in $L^2(S_1, \mathrm{d}\kappa)$. If the width of the support of $\varphi$ in $S_1$ is strictly smaller than $\pi$, the directional wavelet $\psi$ is called *conical*, since supp $\hat{\psi}$ is then contained in a convex cone (Ref. [13], Sec.3.3.4).

A directional (conical) wavelet is characterized by its *angular selectivity* (or *Angular Resolving Power* [2, 13]), that is, its ability to distinguish features with close orientations. This quantity is by definition inversely proportional to the aperture of the support cone of $\hat{\psi}$ : the sharper the cone, the higher the angular selectivity. The support in frequency space of a separable wavelet $\psi$ of the type (2.4) is presented in Figure 1.

Recall that the *anisotropy* of a wavelet does not imply its directionality. An anisotropic wavelet is just noninvariant under rotation around its center. A *width* and a *length* can thus be defined in the spatial domain or in the frequency domain with a ratio different from one.

However, this property alone does not make it able to distinguish different orientations in images as it is suggested by the counterexample of the anisotropic Mexican Hat [1, 2, 13]. Notice that curvelets are directional filters with an anisotropy ratio *width/length* depending (parabolically) of the frequency scale [8] (see Sec.10).



Figure 1: Representation of supp $\psi$ in frequency space.

With (2.4), the wavelet transform (2.3) becomes in polar coordinates

$$\mathcal{W}(\vec{b}, a, \theta) = \int_{\mathbb{R}_+ \times S_1} \rho^*(ak) \, \varphi_\theta^*(\kappa) \, \hat{f}(k, \kappa) \, e^{i\,kb\cos(\kappa-\beta)} \, k \, \mathrm{d}k \, \mathrm{d}\kappa, \tag{2.5}$$

with $\varphi_\theta(\kappa) = \varphi(\kappa - \theta)$, $b = |\vec{b}|$ and $\beta = \arg \vec{b}$. The last expression may be rewritten as

$$\mathcal{W}(\vec{b}, a, \theta) = \langle \varphi_\theta | \, R_{\vec{b},a} \rangle_{S_1}, \tag{2.6}$$

where $\langle h | \, g \rangle_{S_1} = \int_{S_1} h^*(\kappa) g(\kappa) \mathrm{d}\kappa$ is the scalar product on the circle $S_1$ of the $2\pi$-periodic functions $h$ and $g$, and

$$R_{\vec{b},a}(\kappa) = \int_{\mathbb{R}_+} \rho^*(ak) \, \hat{f}(k, \kappa) \, e^{i\,kb\cos(\kappa-\beta)} \, k\mathrm{d}k. \tag{2.7}$$

In other words, relation (2.6) means that the wavelet coefficients of the image $f$ can be interpreted as the projection of $R_{\vec{b},a}$ on a kind of scaling function $\varphi_\theta$ localized around $\theta \in S_1$.

It is now very tempting to add a dilation in this scheme. If $\varphi$ corresponds to the periodization of a function $\nu : \mathbb{R} \to \mathbb{R}$, we may for instance "dilate" $\varphi$ by dilating $\nu$ [14], that is, we consider the following functions:

$$\varphi_\epsilon(\kappa) \;=\; \frac{1}{\epsilon} \sum_{n \in \mathbb{Z}} \nu\big(\frac{\kappa - n}{\epsilon}\big), \quad \epsilon \in \mathbb{R}_+. \tag{2.8}$$

3

Let us replace now $\varphi_\theta$ in (2.6) by the family of functions $\varphi_{\epsilon,\theta}(\kappa) = \varphi_\epsilon(\kappa - \theta)$, leading to the new coefficients

$$\mathcal{W}(\vec{b}, a, \epsilon, \theta) = \langle \varphi_{\epsilon,\theta} |\, R_{\vec{b},a} \rangle_{S_1}.$$

Of course, in this five-dimensional parameter space, the redundancy of the transformation is highly increased, but, for each $\epsilon$, we have a wavelet

$$\hat{\psi}_\epsilon(\vec{k}) = \rho(k)\, \varphi_\epsilon(\kappa), \tag{2.9}$$

of angular selectivity controlled by $\epsilon$. We may expect that isotropic objects in images will be better represented by coefficients $\mathcal{W}(\vec{b}, a, \epsilon, \theta)$ with large $\epsilon$ (small angular selectivity), whereas very directional ones will correspond to a small $\epsilon$ (and high angular selectivity). In the next section, we will show how a circular multiresolution framework [15] can clarify these hints.

# 3 Biorthogonal Multiresolution on the Circle

We start with the usual biorthogonal multiresolution analysis (MRA) on *the line*, that is on $L^2(\mathbb{R})$, then, following Daubechies' procedure (Ref. [15], Sec.9.3), we extend the multiresolution framework to the circle $C_1 \simeq [0, 1)$ (throughout this paper, we reserve the notation $S_1$ to the circle identified with the interval $[0, 2\pi)$).

## 3.1 MRA on the line

As it is well-known, a multiresolution analysis of $L^2(\mathbb{R})$ is an increasing sequence of closed subspaces, interpreted as approximation spaces,[1]

$$\ldots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \ldots, \tag{3.1}$$

with $\bigcap_{l \in \mathbb{Z}} V_l = \{0\}$ and $\bigcup_{l \in \mathbb{Z}} V_l$ dense in $L^2(\mathbb{R})$, and such that

(1) $f(x) \in V_l \Leftrightarrow f(2x) \in V_{l+1}$

(2) There exists a function $\phi \in V_0$, called a *scaling function*, with nonvanishing integral, such that the family $\{\phi(x - k), k \in \mathbb{Z}\}$ is a Riesz basis of $V_0$.

It follows that, for each $l \in \mathbb{Z}$, the family of functions $\{\phi_{l,n}(t) = 2^{l/2}\phi(2^l t - n) : n \in \mathbb{Z}\}$ is a Riesz basis of $V_l$.

In the orthogonal scheme, one considers, for each $l \in \mathbb{Z}$, the orthogonal complement $W_l$ of $V_l$ in $V_{l+1}$, so that $V_{l+1} = V_l \oplus W_l$. The (wavelet) subspaces $W_l$ inherit the scaling and translation properties of $V_l$. Then, the central result is that $W_l$ has a Riesz basis $\{\xi_{l,n}(t) = 2^{l/2}\xi(2^l t - n), n \in \mathbb{Z}\}$, and thus $\{\xi_{l,n}(t) = 2^{l/2}\xi(2^l t - n), l, n \in \mathbb{Z}\}$ is a Riesz basis of $L^2(\mathbb{R})$. The functions $\xi_{l,n}$ are the wavelets.

The inclusions $\phi \in V_0 \subset V_1$ and $\xi \in W_0 \subset V_1$ imply the refinement (or scaling) equations

$$\phi(t) = \sqrt{2} \sum_{n \in \mathbb{Z}} h[n]\, \phi(2t - n), \tag{3.2}$$

$$\xi(t) = \sqrt{2} \sum_{n \in \mathbb{Z}} g[n]\, \phi(2t - n), \tag{3.3}$$

---

[1] Our convention is different from that of Ref. [15]

4

for certain sequences $h[n]$ and $g[n]$ belonging to $l^2(\mathbb{Z})$. Denote by $h(\omega) = \frac{1}{\sqrt{2}} \sum_{n \in \mathbb{Z}} h[n] e^{-in\omega}$, $g(\omega) = \frac{1}{\sqrt{2}} \sum_{n \in \mathbb{Z}} g[n] e^{-in\omega}$ the corresponding filters.

We turn now to the biorthogonal scheme. Starting from the scale (3.1), one considers, for each $l \in \mathbb{Z}$, the basis$\{\widetilde{\phi}_{l,n}(t) = 2^{l/2}\widetilde{\phi}(2^l t - n),\ n \in \mathbb{Z}\}$, dual to $\{\phi_{l,n}, n \in \mathbb{Z}\}$, that is, the vectors defined by the relation $\langle \phi_{l,n'} | \widetilde{\phi}_{l,n} \rangle = \delta_{nn'}$ (as it is well-known, these relations do *not* determine the new basis vectors uniquely[15, 23]). Let $\widetilde{V}_l$ denote the closed subspace generated by $\{\widetilde{\phi}_{l,n},\ n \in \mathbb{Z}\}$. The outcome is a second multiresolution scale $\{\widetilde{V}_l\}$, with exactly the same properties as $\{V_l\}$. Next, for each $l \in \mathbb{Z}$, one defines a subspace $W_l$ by the two conditions

$$W_l \subset V_{l+1} \text{ and } W_l \perp \widetilde{V}_l,$$

and similarly $\widetilde{W}_l \subset \widetilde{V}_{l+1}$ and $\widetilde{W}_l \perp V_j$. In this way one obtains two sequences of subspaces $\{W_l\}$ and $\{\widetilde{W}_l\}$, with Riesz bases $\{\xi_{l,n}(t) = 2^{l/2}\xi(2^l t - n),\ n \in \mathbb{Z}\}$ and $\{\widetilde{\xi}_{l,n}(t) = 2^{l/2}\widetilde{\xi}(2^l t - n),\ n \in \mathbb{Z}\}$, respectively. The four bases satisfy the biorthogonality relations

$$
\begin{array}{rclcrcl}
\langle \xi_{l,n'} | \widetilde{\phi}_{l,n} \rangle & = & 0, & & \langle \phi_{l,n} | \widetilde{\xi}_{l,n'} \rangle & = & 0, \\
\langle \phi_{l,n'} | \widetilde{\phi}_{l,n} \rangle & = & \delta_{n,n'}, & & \langle \xi_{l',n'} | \widetilde{\xi}_{l,n} \rangle & = & \delta_{l,l'}\,\delta_{n,n'},
\end{array}
\tag{3.4}
$$

for all $l, l', n, n' \in \mathbb{Z}$. Since the corresponding filters $\{\widetilde{h}, g\}$ and $\{h, \widetilde{g}\}$ are linked by the relations

$$g(\omega) = e^{-i\omega}\,\widetilde{h}^*(\omega + \pi), \quad \widetilde{g}(\omega) = e^{-i\omega}\,h^*(\omega + \pi), \tag{3.5}$$

the whole biorthogonal multiresolution analysis is determined by the filters $h$ and $\widetilde{h}$.

Given any $f \in L^2(\mathbb{R})$, define the coefficients $c_l[n] = \langle \phi_{l,n} | f \rangle$ and $d_l[n] = \langle \xi_{l,n} | f \rangle$, for $l, n \in \mathbb{Z}$. Then, for any $L \in \mathbb{Z}$, the reconstruction formulas read

$$f = \sum_{n \in \mathbb{Z}} c_L[n]\,\widetilde{\phi}_{L,n} + \sum_{l=L}^{+\infty} \sum_{n \in \mathbb{Z}} d_l[n]\,\widetilde{\xi}_{l,n} \tag{3.6}$$

$$= \sum_{l \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} d_l[n]\,\widetilde{\xi}_{l,n}. \tag{3.7}$$

Finally, the refinement equations entail that the coefficients $c_l$ and $d_l$ satisfy recursion formulas which allow their fast computation, namely,

$$
\begin{aligned}
c_l[n] &= I\,h * c_{l+1}\,[2n], \\
d_l[n] &= I\,g * c_{l+1}\,[2n], \\
c_{l+1}[n] &= \widetilde{h} * U c_l\,[n] + \widetilde{g} * U d_l\,[n],
\end{aligned}
\tag{3.8}
$$

where, for any sequence $q \in l^2(\mathbb{Z})$, $I\,q[n] = q[-n]$ and

$$U q[n] = \begin{cases} q[n/2] & \text{if } n \text{ is even,} \\ 0 & \text{if } n \text{ is odd,} \end{cases} \tag{3.9}$$

that is, the operation $q \mapsto Uq$ represents oversampling by a factor of 2. We also use the standard convolution $*$ of two sequences defined by

$$u * v[n] = \sum_{m \in \mathbb{Z}} u[n - m]\,v[m]. \tag{3.10}$$

## 3.2 Circular MRA

Consider a multiresolution analysis of $L^2(\mathbb{R})$, such as described in the previous section, with generating functions[2] $\phi^\mathbb{R}(t)$ and $\xi^\mathbb{R}(t)$ well localized in space, in the sense that there are constants $\epsilon > 0$ and $C > 0$ such that $|\phi^\mathbb{R}(t)|, |\xi^\mathbb{R}(t)| \leqslant C(1 + |t|)^{-1-\epsilon}$. The periodized version of $\phi^\mathbb{R}_{l,n}$ (Ref. [15], Sec.9.3) is

$$\phi_{l,n}(t) = \sum_{m \in \mathbb{Z}} \phi^\mathbb{R}_{l,n}(t + m) \qquad (3.11)$$

and similarly for $\xi^\mathbb{R}_{l,n}$, and also for the dual functions $\widetilde{\phi}^\mathbb{R}$ and $\widetilde{\xi}^\mathbb{R}$, leading to $\xi_{l,n}$, $\widetilde{\phi}_{l,n}$ and $\widetilde{\xi}_{l,n}$. By the periodicity, we may clearly restrict all the functions to $C_1 \sim [0, 1]$.

Hence, from the subspaces $V_l^\mathbb{R}, W_l^\mathbb{R} \in L^2(\mathbb{R})$, generated by $\{\phi^\mathbb{R}_{l,n}, n \in \mathbb{Z}\}$ and $\{\xi^\mathbb{R}_{l,n}, n \in \mathbb{Z}\}$, respectively, we deduce the subspaces $V_l, W_l \in L^2(C_1)$, generated by $\{\phi_{l,n}, n \in \mathbb{Z}\}$ and $\{\xi_{l,n}, n \in \mathbb{Z}\}$, respectively. In the same way, the dual subspaces $\widetilde{V}_l^\mathbb{R}, \widetilde{W}_l^\mathbb{R}$ induce the periodic subspaces $\widetilde{V}_l, \widetilde{W}_l$. It is easy to see that these new spaces are finite dimensional, since

$$\phi_{l,n} = \phi_{l,n+2^l r}, \quad \xi_{l,n} = \xi_{l,n+2^l r}, \qquad (3.12)$$

for all $l \geqslant 0$, $n \in \mathbb{Z}$ and $r \in \mathbb{Z}$. Therefore, $\dim V_l = \dim W_l = 2^l$ and the dual spaces have the same dimension.

In addition, from the periodicity of the circle, the scale parameter $l$ is restricted to positive values. Indeed, if $\sum_{m \in \mathbb{Z}} h[2m] = \sum_{m \in \mathbb{Z}} h[2m + 1] = \frac{1}{\sqrt{2}}$, then $\phi^\mathbb{R}$ realizes a partition of the line [15], that is, $\sum_{m \in \mathbb{Z}} \phi^\mathbb{R}(t-m) = 1$ for every $t \in \mathbb{R}$ (pointwise). Under the same assumption, we have also $\sum_{m \in \mathbb{Z}} \xi^\mathbb{R}(t - \frac{m}{2}) = 0$.

Therefore, $\phi_{0,n} = 1$ and $\xi_{-1,n} = 0$ for all $n \in \mathbb{Z}$. The spaces $V_0$ and $W_{-1}$ are thus respectively the set of the constant functions on $[0, 1)$ and the null function set $\{0\}$. In consequence, from now on, we will restrict $l$ to positive integers.

Notice that orthogonality relations analogous to (3.4) hold between the subspaces $V_l$ and $\widetilde{W}_l$ (resp. between $\widetilde{V}_l$ and $W_l$) of $L^2(C_1)$, since they are inherited from the ones between $V_l^\mathbb{R}$ and $\widetilde{W}_l^\mathbb{R}$ (resp. between $\widetilde{V}_l^\mathbb{R}$ and $W_l^\mathbb{R}$) [15], that is,

$$\langle \xi_{l,n} | \widetilde{\phi}_{l,n'} \rangle_{C_1} = 0, \quad \langle \phi_{l,n} | \widetilde{\xi}_{l,n'} \rangle_{C_1} = 0, \quad l, l' \in \mathbb{N}, \ 0 \leqslant n < 2^l, \ 0 \leqslant n' < 2^{l'}, \qquad (3.13)$$

and similarly for the other two relations in (3.4).

In conclusion, since the dilations are performed before the periodization of the nonperiodic functions, the periodic functions $\phi, \xi, \widetilde{\phi}$ and $\widetilde{\xi}$ generate a biorthogonal multiresolution analysis with subspaces $V_l \subset V_{l+1}$, $W_l \subset W_{l+1}$, $\widetilde{V}_l \subset \widetilde{V}_{l+1}$ and $\widetilde{W}_l \subset \widetilde{W}_{l+1}$.

Notice that, for $l \in \mathbb{N}^*$, the scaling rules (3.2) and (3.3) become

$$\phi_{l-1,0}(t) = \sqrt{2} \sum_{n=0}^{2^l-1} h_l[n] \, \phi_{l,n}(t), \qquad (3.14)$$

$$\xi_{l-1,0}(t) = \sqrt{2} \sum_{n=0}^{2^l-1} g_l[n] \, \phi_{l,n}(t), \qquad (3.15)$$

where $q_l[n] = \sum_{m \in \mathbb{Z}} q[n + 2^l m]$ denotes the $2^l$-periodization of the sequence $q[n]$. The dual periodic refinement equations are obtained in the same way.

---

[2]In the sequel , we will add the label $(\cdot)^\mathbb{R}$ to every function or space related to a MRA on $L^2(\mathbb{R})$.

The reconstruction formula (3.6) now becomes

$$f = c_0[0] + \sum_{l \in \mathbb{N}} \sum_{n=0}^{2^l - 1} d_l[n] \, \widetilde{\xi}_{l,n}, \quad f \in L^2(C_1), \tag{3.16}$$

with $c_l[n] = \langle \phi_{l,n} | f \rangle_{C_1}$ and $d_l[n] = \langle \xi_{l,n} | f \rangle_{C_1}$. Remark that $c_0[0] = \langle f \rangle_{C_1} = \int_{C_1} f$, since $\phi_{0,0} = 1$.

Finally, the recursion rules (3.8) become

$$\begin{aligned}
c_{l-1}[n] &= I \, h_l \star c_l \, [2n], \\
d_{l-1}[n] &= I \, g_l \star c_l \, [2n], \\
c_{l+1}[n] &= \widetilde{h}_{l+1} \star U c_l \, [n] + \widetilde{g}_{l+1} \star U d_l \, [n],
\end{aligned} \tag{3.17}$$

where $\star$ denotes the circular convolution defined as,

$$u \star v[n] \;=\; \sum_{m=0}^{P-1} u[n - m \bmod P] \, v[m], \tag{3.18}$$

for any two sequences $u$ and $v$ of length $P$.

# 4 Angular Multiselectivity Analysis

The circular multiresolution framework developed in the previous section can be used to define new 2-D wavelets with multiple, but controlled, angular selectivities [22]. As we will see below, their main properties are

- they combine with each other in a pyramidal scheme to form less selective directional wavelets until one obtains a totally isotropic one;

- they define for each *selectivity level* a Littlewood-Paley[3] (LP) directional frame.

## 4.1 Directional frames

We begin by recalling the definition of a directional frame. According to the theory of dyadic (or half-continuous) frames (Ref. [13], Sec. 2.4, Ref. [16], Chap. VI), the continuous wavelet transform given in (2.1) can be discretized in its parameters $a$ and $\theta$, while preserving a perfect reconstruction formula.

Indeed, let $\psi \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$ be a polar separable wavelet of the form (2.4). Given a dyadic sampling of the scales $a_j = a_0 2^{-j}$ ($j \in \mathbb{Z}$, $a_0 \in \mathbb{R}_+$) and a regular sampling of the angles $\theta_n = n \frac{2\pi}{K}$ ($0 \leqslant n < K$, $K \in \mathbb{N}^*$), assume that the *frame property* is satisfied, that is, there are two positive constants $\mathsf{m}, \mathsf{M}$ such that

$$\mathsf{m} \;\leqslant\; \sum_{j \in \mathbb{Z}} \sum_{n=0}^{K-1} |\rho(a_j k)|^2 \, |\varphi(\kappa - \theta_n)|^2 \;\leqslant\; \mathsf{M}. \tag{4.1}$$

---

[3]In this paper, we prefer to qualify these frames of *Littlewood-Paley frames* since they induce a Littlewood-Paley reconstruction formula (see (4.4) below). Other authors [16] prefer to call them *linear* frame.

Then, there exists a (nonunique) dual wavelet $\widetilde{\psi}$ which yields the following reconstruction formula for any $f \in L^2(\mathbb{R}^2)$ :

$$f(\vec{x}) = \sum_{j \in \mathbb{Z}} \sum_{n=0}^{K-1} \mathcal{W}_{j,n} * \widetilde{\psi}_{a_j, \theta_n}(\vec{x}), \tag{4.2}$$

where $*$ stands here for the standard convolution between two functions of $L^2(\mathbb{R}^2)$.

A particular case Littlewood-Paley (LP) or 'linear' frame arises when

$$\sum_{j \in \mathbb{Z}} \sum_{n=0}^{K-1} \rho(a_j k)\, \varphi(\kappa - \theta_n) = A, \quad A \in \mathbb{R}_+^* \ . \tag{4.3}$$

Then, $\widetilde{\psi}$ is simply a Dirac distribution $\delta^{(2)}(\vec{x})$ and the reconstruction formula (4.2) reduces to a Littlewood-Paley decomposition

$$f(\vec{x}) = \frac{1}{A} \sum_{j \in \mathbb{Z}} \sum_{n=0}^{K-1} \mathcal{W}_{j,n}(\vec{x}). \tag{4.4}$$

## 4.2 Angular multiselectivity

As we have seen in Section 2, when the wavelet $\psi$ has the form (2.4), then the resulting 2-D continuous wavelet transform $\mathcal{W}(\vec{b}, a, \theta)$ of an image $f \in L^2(\mathbb{R}^2)$ is given by the scalar product

$$\mathcal{W}(\vec{b}, a, \theta) = \langle \varphi_\theta | R_{\vec{b}, a} \rangle_{S_1}. \tag{4.5}$$

Given a biorthogonal multiresolution analysis of the circle, with a scaling function $\phi$ and a wavelet $\xi$, we apply the discretization described in Section 4.1 and project $R_{\vec{b}, j} := R_{\vec{b}, a_j}$ onto the functions $\varphi_{l,n}(\kappa) = \phi_{l,n}\left(\frac{\kappa}{2\pi}\right)$ and $\eta_{l,n}(\kappa) = \xi_{l,n}\left(\frac{\kappa}{2\pi}\right)$. Thus we get the new coefficients

$$\begin{aligned}
\mathcal{W}_{j,l,n}^{\mathrm{a}}(\vec{b}) &= \langle \varphi_{l,n} | R_{\vec{b}, j} \rangle_{S_1}, \\
\mathcal{W}_{j,l,n}^{\mathrm{d}}(\vec{b}) &= \langle \eta_{l,n} | R_{\vec{b}, j} \rangle_{S_1}, \\
\mathcal{W}_j^{\mathrm{i}}(\vec{b}) &= \langle \varphi_{0,0} | R_{\vec{b}, j} \rangle_{S_1} = \langle 1 | R_{\vec{b}, j} \rangle_{S_1},
\end{aligned} \tag{4.6}$$

for $l \in \mathbb{N}$ and $0 \leqslant n < 2^l$. These amount, respectively, to the projection of the image $f$ on translated and dilated copies of the functions $\psi_{l,n}^{\mathrm{a}}$, $\psi_{l,n}^{\mathrm{d}}$ and $\psi^{\mathrm{i}}$, respectively, defined in frequency space by

$$\hat{\psi}_{l,n}^{\mathrm{a}}(\vec{k}) = \rho(k)\, \varphi_{l,n}(\kappa), \tag{4.7}$$

$$\hat{\psi}_{l,n}^{\mathrm{d}}(\vec{k}) = \rho(k)\, \eta_{l,n}(\kappa), \tag{4.8}$$

$$\hat{\psi}^{\mathrm{i}}(\vec{k}) = \hat{\psi}_{0,0}^{\mathrm{a}}(\vec{k}) = \rho(k). \tag{4.9}$$

Here the exponent $a$ stands for angular *approximation*, $d$ for angular *details*, and $i$ for *isotropic*. The full parametrization of the $L^1(\mathbb{R}^2)$-normalized wavelets thus reads

$$\psi_{\vec{b}, j, l, n}^{\mathrm{a}}(\vec{x}) = a_j^{-2}\, \psi_{l,n}^{\mathrm{a}}\left(\frac{\vec{x} - \vec{b}}{a_j}\right), \tag{4.10}$$

and similarly for $\psi_{\vec{b},j,l,n}^{\mathrm{d}}(\vec{x})$ and $\psi_{\vec{b},j}^{\mathrm{i}}$.

Besides the scale and translation parameters $a_j$ and $\vec{b}$, the rotation of the 2-D wavelets is obviously given by the parameter $n$, $0 \leqslant n < 2^l$, which precisely translates a function on $S^1$ by an angle $2\pi n 2^{-l}$. Since the aperture of the frequency space cones containing these wavelets is proportional to $2^{-l}$, the angular selectivity of $\psi_{l,n}^{\mathrm{a}}$ and $\psi_{l,n}^{\mathrm{d}}$ is proportional to $2^l$ (see Section 2). Keeping this in mind, we will call the parameter $l$ the *(angular) selectivity level*.

We have thus generated a new family of wavelets, $\{\psi_{l,n}^{\mathrm{a}}, \psi_{l,n}^{\mathrm{d}}, \psi^i\}$ ranging from very directional ones to a totally isotropic one, depending on the value of $l$.

We conclude this section by noticing that, for each selectivity level $l \in \mathbb{N}$, the family $\{\psi_{j,l,n}^{\mathrm{a}}\}$ generates a LP frame. More precisely, the following proposition holds.

**Proposition 4.1** *Let $a_j = a_0 2^{-j}$, $j \in \mathbb{Z}$, be a dyadic scale discretization and suppose that $\sum_{j\in\mathbb{Z}} \rho(a_j k) = 1$ a.e. for $k \in \mathbb{R}_+$. Then, for any $l \in \mathbb{N}$, the family $\{\psi_{j,l,n}^{\mathrm{a}} : j \in \mathbb{Z}, 0 \leqslant n < 2^l \}$ is a LP frame of $L^2(\mathbb{R}^2)$, i.e., it obeys (4.3) with $A = 2^{l/2}$.*

This is a simple consequence of the fact that $\phi^{\mathbb{R}}$ realizes a partition of the line, i.e., $\sum_{m\in\mathbb{Z}} \phi^{\mathbb{R}}(t + m) = 1$ for every $t \in \mathbb{R}$. Indeed, with $u = \frac{\kappa}{2\pi}$, where $\kappa = \arg \vec{k}$,

$$
\sum_{j\in\mathbb{Z}} \sum_{n=0}^{2^l-1} \hat{\psi}_{j,l,n}^{\mathrm{a}}(\vec{k}) = \sum_{j\in\mathbb{Z}} \sum_{n=0}^{2^l-1} \rho(a_j k)\, \varphi_{l,0}(\kappa - n\tfrac{2\pi}{2^l}) = \sum_{n=0}^{2^l-1} \phi_{l,0}(u - \tfrac{n}{2^l})
$$

$$
= 2^{l/2} \sum_{m\in\mathbb{Z}} \sum_{n=0}^{2^l-1} \phi^{\mathbb{R}}(2^l u + 2^l m - n)
$$

$$
= 2^{l/2},
$$

and (4.3) is satisfied.

## 4.3    Recursion formulas

For any $l \in \mathbb{N}^*$, the functions $\varphi$ and $\eta$ verify simple extensions of the scaling rules (3.14) and (3.15):

$$
\varphi_{l-1,0}(\kappa) = \sum_{n=0}^{2^l-1} h_l[n]\, \varphi_{l,n}(\kappa), \tag{4.11}
$$

$$
\eta_{l-1,0}(\kappa) = \sum_{n=0}^{2^l-1} g_l[n]\, \varphi_{l,n}(\kappa). \tag{4.12}
$$

We have also, for every $l \in \mathbb{N}$,

$$
\varphi_{l+1,n}(\kappa) = \sum_{n'=0}^{2^l-1} \widetilde{h}_{l+1}[n - 2n']\, \varphi_{l,n}(\kappa) + \sum_{n'=0}^{2^l-1} \widetilde{g}_{l+1}[n - 2n']\, \eta_{l,n}(\kappa). \tag{4.13}
$$

If we project (4.11), (4.12) and (4.13) onto $R_{\vec{b},j}$, we obtain the following relations for the decomposition:

$$\mathcal{W}^{\mathrm{a}}_{j,l-1,n}(\vec{b}) = \sum_{n'=0}^{2^l-1} h_l^*[n'-2n]\,\mathcal{W}^{\mathrm{a}}_{j,l,n'}(\vec{b}) = \big(\bar{h}_l \star \mathcal{W}^{\mathrm{a}}_{j,l,\cdot}(\vec{b})\big)_{2n}, \qquad (4.14)$$

$$\mathcal{W}^{\mathrm{d}}_{j,l-1,n}(\vec{b}) = \sum_{n'=0}^{2^l-1} g_l^*[n'-2n]\,\mathcal{W}^{\mathrm{a}}_{j,l,n'}(\vec{b}) = \big(\bar{g}_l \star \mathcal{W}^{\mathrm{a}}_{j,l,\cdot}(\vec{b})\big)_{2n}, \qquad (4.15)$$

where we write $\bar{u} = I\,u^*$ for any sequence $u$.

As for the reconstruction, we get

$$\mathcal{W}^{\mathrm{a}}_{j,l+1,n}(\vec{b}) = \sum_{n'=0}^{2^l-1} \widetilde{h}^*_{l+1}[n-2n']\,\mathcal{W}^{\mathrm{a}}_{j,l,n'}(\vec{b}) \;+\; \sum_{n'=0}^{2^l-1} \widetilde{g}^*_{l+1}[n-2n']\,\mathcal{W}^{\mathrm{d}}_{j,l,n'}(\vec{b})$$

$$= \big(\widetilde{h}^*_{l+1} \star U\mathcal{W}^{\mathrm{a}}_{j,l,\cdot}(\vec{b})\big)_n + \big(\widetilde{g}^*_{l+1} \star U\mathcal{W}^{\mathrm{d}}_{j,l,\cdot}(\vec{b})\big)_n. \qquad (4.16)$$

Note that the oversampling operation $U$ as well as the circular convolution $\star$ are performed on the angular parameter of the wavelet coefficients, as in (3.9) and (3.17). The notation $(\,\cdot\,)_n$ means simply that we select the $n^{\mathrm{th}}$ angular element of this convolution. These relations are summarized in Figures 2 and 3, where the operator $D$ is the downsampling operator that turns a sequence $u$ of length $2P$ into a sequence $(Du)[n] = u[2n]$ of length $P$. Notice that the last term in the upper row of Figure 2 and the first term in the upper row of Figure 3 carry a label '$i$' instead of '$a$', since these terms correspond to the fully isotropic wavelet.

Figure 2: Decomposition of the wavelet coefficients according to their angular selectivity.

Figure 3: Reconstruction of the wavelet coefficients according to their angular selectivity.

As a matter of fact, (4.14), (4.15) and (4.16) are the exact counterpart of the usual recursion relations for wavelet coefficients [15]. They rely on the fact that, according to (4.11), (4.12), several wavelets of level $l$ merge into a single wavelet at level $l-1$, i.e., with half the angular selectivity.

## 4.4  Choice of the wavelet

We must now choose a wavelet for our multiselectivity analysis. We select for $\phi^{\mathbb{R}}$ and $\xi^{\mathbb{R}}$ the *B-spline* scaling function and wavelet, respectively [17]. In particular, we use the results of Cohen-Daubechies-Feauveau [18] (CDF) on the compactly supported, spline biorthogonal wavelet bases with several vanishing moments. We propose to use the filters $h$ and $\widetilde{h}$, with 3 and 7 vanishing moments, respectively (see Table 1). This ensures that the resulting $\psi_{l,n}^{\mathrm{a}}$ (given below) is a conical wavelet with quadratic regularity on the edges of its conical frequency support.

| $n$ | $h[n]$ | $\widetilde{h}[n]$ |
|:---:|:---:|:---:|
| $0,1$ | 0.53033008588991 | 0.95164212189718 |
| $-1,2$ | 0.17677669529664 | $-0.02649924094535$ |
| $-2,3$ | | $-0.30115912592284$ |
| $-3,4$ | | 0.03133297870736 |
| $-4,5$ | | 0.07466398507402 |
| $-5,6$ | | $-0.01683176542131$ |
| $-6,7$ | | $-0.00906325830378$ |
| $-7,8$ | | 0.00302108610126 |

Table 1: CDF direct and dual filters of 3 and 7 vanishing moments.

Given an initial selectivity level $L \in \mathbb{N}$, we define $\psi_{L,n}^{\mathrm{a}}$ and $\psi_{L,n}^{\mathrm{d}}$ in frequency as

$$\hat{\psi}_{L,n}^{\mathrm{a}}(\vec{k}) = \phi^{\mathbb{R}}(\log_2 k)\, \varphi_{L,n}(\kappa), \tag{4.17}$$

$$\hat{\psi}_{L,n}^{\mathrm{d}}(\vec{k}) = \phi^{\mathbb{R}}(\log_2 k)\, \eta_{L,n}(\kappa), \tag{4.18}$$

with $0 \leqslant n < 2^L$, $\varphi_{L,n}(\kappa) = \phi_{L,n}(\frac{\kappa}{2\pi})$, $\eta_{L,n}(\kappa) = \xi_{L,n}(\frac{\kappa}{2\pi})$, and $\phi$ and $\xi$ the periodization of $\phi^{\mathbb{R}}$ and $\xi^{\mathbb{R}}$, respectively. This yields $K = 2^L$ differently oriented wavelets.

Remark that, choosing $\rho(k) = \phi^{\mathbb{R}}(\log_2 k)$, as in (4.17), implies that $\sum_{j \in \mathbb{Z}} \rho(a_j k) = \sum_{j \in \mathbb{Z}} \phi^{\mathbb{R}}(\log_2 a_0 k - j) = 1$, since $\sum_{m \in \mathbb{Z}} \phi^{\mathbb{R}}(t - m) = 1$. Thus, by Proposition 4.1, the wavelets $\{\psi_{j,l,n}^{\mathrm{a}}\}$ constructed from (4.17) constitute a LP frame for each $l \in \mathbb{N}$.

The resulting functions $\hat{\psi}_{L,n}^{\mathrm{a,d}}$ are presented in the frequency domain on Figure 4 for several values of $L$ and for $n = 0$. Notice that the aperture of the supporting cones is decreasing for increasing $L$. As explained before, the aperture $2\alpha_L$ of the cone supporting $\hat{\psi}_{L,n}^{\mathrm{a}}$ and $\hat{\psi}_{L,n}^{\mathrm{d}}$ is proportional to $2^{-L}$, which means that the angular selectivity of these wavelets growths with $L$. This behavior is illustrated more clearly in Figure 5 for the CDF$-(3,7)$ framework. Since $\operatorname{supp}\phi^{\mathbb{R}} = [-3/2, 3/2]$, $\alpha_L$ is not defined for $0 \leqslant L < 2$ (i.e., $\operatorname{supp}\varphi = S_1$), and $\alpha_L = 3\pi\, 2^{-L}$ for $L \geqslant 2$, so that the wavelets $\psi_{L,n}^{\mathrm{a}}$ are conical (i.e., $\alpha_L < \pi/2$) for $L \geqslant 3$ (Ref. [13], Sec.3.3.4).

## 5  Best Frame Selection

Starting from a selectivity level $L \in \mathbb{N}$, we have seen in Section 4.4 that we can generate inductively $L + 1$ frames with $\psi_{l,n}^{\mathrm{a}}$ ($0 \leqslant l \leqslant L$), characterized by an angular selectivity $2^l$.
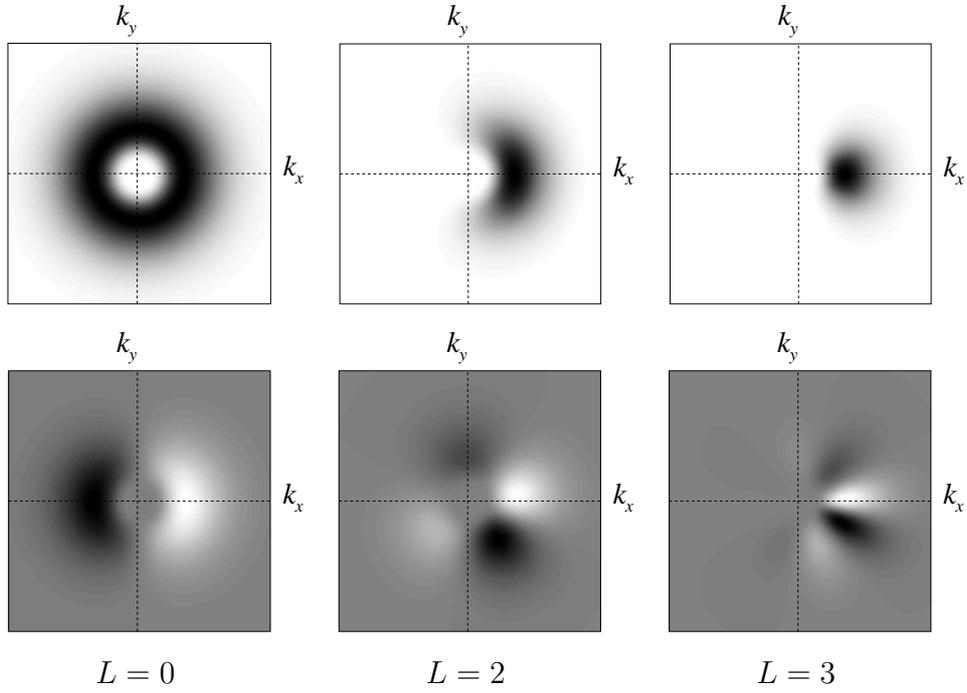
Figure 4: Wavelets $\hat{\psi}_{L,n}^{\mathrm{a}}(\vec{k})$ (top row) and $\hat{\psi}_{L,n}^{\mathrm{d}}(\vec{k})$ (bottom) for several values of $L$ and for $n = 0$.
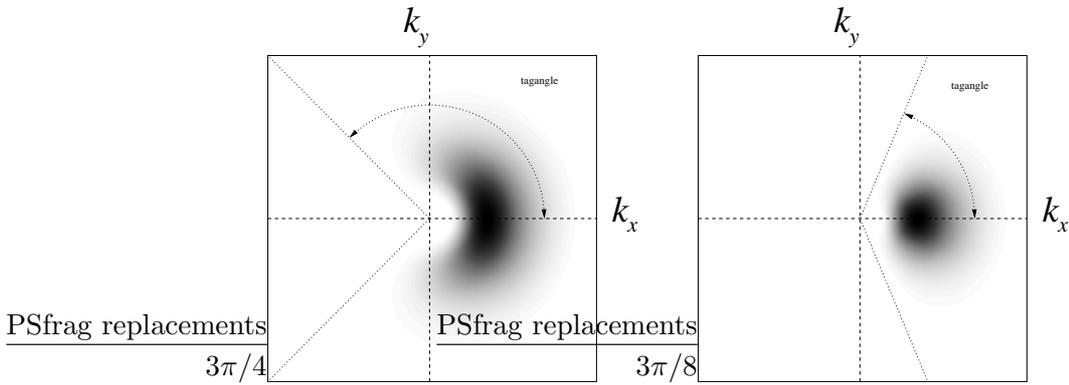


Figure 5: $\hat{\psi}_{2,0}^{\mathrm{a}}(\vec{k})$ with $\alpha_2 = 3\pi/4$ (left) and $\hat{\psi}_{3,0}^{\mathrm{a}}(\vec{k})$ with $\alpha_3 = 3\pi/8$ (right) ; the latter is a conical wavelet.

In particular, for each level $l \in [0, L]$, Proposition 4.1 shows that the wavelets $\psi_{j,l,n}^{\mathrm{a}}$ generate a LP frame of constant $A = 2^{l/2}$ which, according to (4.4), may be used to reconstruct the original image. But there is more. Indeed, as shown in the following proposition, we can mix *different* frames inside the *same* reconstruction formula.

**Proposition 5.1** *If $\rho$ satisfies the same condition as in Proposition 4.1, then, given any function $\tilde{l} : (\vec{x}, j) \in \mathbb{R}^2 \times \mathbb{Z} \mapsto \tilde{l}(\vec{x}, j) \in \mathbb{N}$, a function $f \in L^2(\mathbb{R}^2)$ can be decomposed as*

$$f(\vec{x}) \;=\; \sum_{j \in \mathbb{Z}} \sum_{n=0}^{2^{\tilde{l}}-1} 2^{-\tilde{l}/2} \, \mathcal{W}_{j,\tilde{l},n}^{\mathrm{a}}(\vec{x}), \quad \tilde{l} \equiv \tilde{l}(\vec{x}, j). \tag{5.1}$$

This is a simple consequence of the identities $\sum_{j \in \mathbb{Z}} \mathcal{W}_j^{\mathrm{i}}(\vec{x}) = f$ and

$$\sum_{n=0}^{2^l-1} 2^{-l/2} \, \mathcal{W}_{j,l,n}^{\mathrm{a}}(\vec{x}) = \sum_{n=0}^{2^{l'}-1} 2^{-l'/2} \, \mathcal{W}_{j,l',n}^{\mathrm{a}}(\vec{x}) \;=\; \mathcal{W}_j^{\mathrm{i}}(\vec{x}), \text{ for any } l, l' \in \mathbb{N}.$$

The precious property uncovered by this last proposition provides us with a new degree of freedom to describe images adaptively. Indeed, at each point $\vec{b} \in \mathbb{R}^2$ and each scale $j \in \mathbb{Z}$, we may search the "best frame", that is, the selectivity level $l(\vec{b}, j)$ characterizing best the content of $f$. To that effect, we simply choose the frame that offers the best match between the image and the wavelets, that is, at $(\vec{b}, j)$, we choose the selectivity level

$$\ell(\vec{b}, j) = \arg\max_{l \in [0, L]} \; \max_{n \in [0, 2^l)} \; \frac{|\langle \psi_{\vec{b}, j, l, n}^{\mathrm{a}} \,|\, f \rangle|}{\|\psi_{\vec{b}, j, l, n}^{\mathrm{a}}\|}. \tag{5.2}$$

Obviously, the previous proposition garantees that other selection rules may be chosen (e.g. a sparsity criterion based on the $\ell^1$ minimization of normalized coefficients, see Sec. 11). In this paper, however, we prefer this best matching procedure to study when the similarity between the analyzed image and the wavelets generated by our method is interesting for the final decomposition.
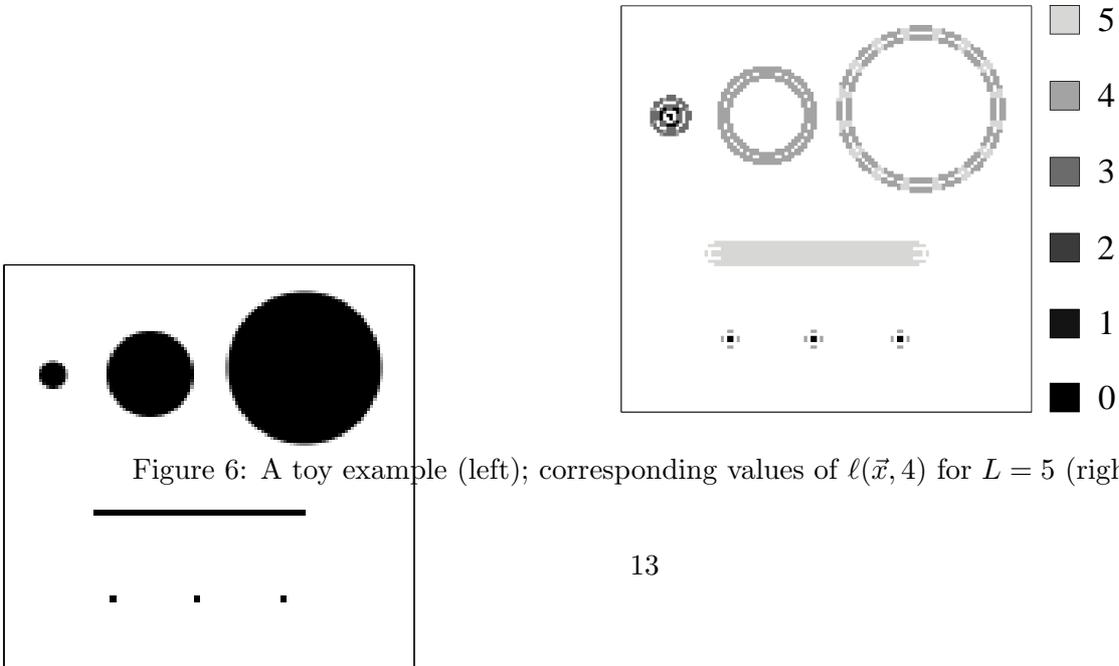


Figure 6: A toy example (left); corresponding values of $\ell(\vec{x}, 4)$ for $L = 5$ (right).

The reconstruction procedure will be then defined by

$$f(\vec{x}) = \sum_{j \in \mathbb{Z}} \sum_{n=0}^{2^\ell - 1} 2^{-\ell/2} \, \mathcal{W}_{j,\ell,n}^{\mathrm{a}}(\vec{x}), \quad \ell = \ell(\vec{x}, j), \tag{5.3}$$

since, for fixed $j$, the inner sum on $n$ equals $\mathcal{W}_j^{\mathrm{i}}(\vec{x})$, exactly as in (4.4).

In Figure 6, the computation of $\ell$ is presented for $j = J - 1 = 4$ (see the next section for details on the scale discretization) and $L = 5$ on a toy image containing a set of simple geometric objects. Values of $\ell$ are displayed only in areas where $|W_4^{\mathrm{i}}(\vec{x})| \geqslant \frac{1}{10} |\max_{\vec{x}} W_4^{\mathrm{i}}(\vec{x})|$.

As we can see, $\ell$ follows closely and locally the directional aspect of the analyzed objects. Indeed, the selectivity level $\ell$ is increasing with the curvature radius of the three disks on their edges. The three singularities (spots) on the bottom are however linked with the 0 level, i.e., the isotropic level. The straight line produces the maximum level, as expected.

# 6 Discretization

In practice, images are discretized on a regular grid of pixels, i.e., an image $f$ is defined on samples $f_{\mathrm{d}}[\vec{p}] = f(\vec{p})$, where $\vec{p} = (p, q) \in \mathbb{Z}^2$ (and in fact, of course, with only a finite number of samples, see below). By the Shannon theorem, the Fourier transform of $f_{\mathrm{d}}$ is the $2\pi$-periodization of $\hat{f}$, both in $k_x$ and $k_y$ directions. In other words, $f$ is completely determined from the values of $f_{\mathrm{d}}$ if $f$ is band-limited, that is, if $f \in \mathcal{B}_\pi = \{g \in L^2(\mathbb{R}^2) : \hat{g}(\vec{k}) = 0 \text{ if } \vec{k} \notin B_\pi\}$ with if $f \in \mathcal{B}_\pi = \{g \in L^2(\mathbb{R}^2) : \operatorname{supp} \hat{g} \subset B_\pi := [-\pi, \pi) \times [-\pi, \pi)\}$. In this particular context, the LP frame condition (4.3) must hold only at points $\vec{k} \in B_\pi$.

However, since $a_j = a_0 2^{-j}$, the function $\rho(a_j k)$ spreads to higher frequencies when $j$ increases. Thus there is a maximal value allowed for $j$, we call it $J - 1$ (with $J > 0$). Hence, with our choice of wavelet, since the quadratic spline $\phi$ is centered at the origin and has the support equal to $[-3/2, 3/2]$, $\rho(k) = \phi(\log_2 k)$ will be centered at $k = 1$ with $\operatorname{supp} \rho = [2^{-3/2}, 2^{3/2}]$. So, with $a_0 = \pi^{-1} 2^{J+1/2}$, $\rho_j(k) = \rho(a_j k)$ is centered for $j = J - 1$ on $\pi/2$ inside the support $[\pi/8, \pi]$, ensuring that $\psi_{b,j,l,n}^{\mathrm{a}} \in \mathcal{B}_\pi$ for $j < J$. Let us now gather together wavelets with $j \geqslant J$ in $B_\pi$ by defining

$$\begin{aligned}
\hat{\psi}_{l,n}^{\mathrm{high}}(\vec{k}) &= \chi_{B_\pi}(\vec{k}) \, \varphi_{l,n}(\kappa) \sum_{j=J}^{\infty} \phi(\log_2 a_j k) \\
&= \chi_{B_\pi}(\vec{k}) \, \varphi_{l,n}(\kappa) \sum_{j=J}^{\infty} \phi(j + \log_2 a_0 k), \tag{6.1}
\end{aligned}$$

where *high* stands for *high* frequency components and $\chi_{B_\pi}(\vec{k})$ is the characteristic function of $B_\pi$.

Since the discretized image is also limited in space, say of size $N \times N$ pixels, we define also the isotropic *low* frequency function to gather wavelets that are of a larger size than that of the original image $f$, that is

$$\hat{\psi}^{\mathrm{low}}(\vec{k}) = \sum_{j=-\infty}^{-1} \phi(\log_2 a_j k). \tag{6.2}$$

This two-dimensional *scaling function* is in fact fully determined by $J$ : because of the support properties of $\phi$, $\hat{\psi}^{\text{low}}$ is contained inside a disk of radius $\pi \, 2^{-J}$. By inspection of $\psi^{\text{low}}$ and $\psi_{j,l,n}$ in the spatial domain, one sees that imposing roughly $J$ smaller than $\log_2 N/8$ guarantees that these functions are essentially smaller than the image and sufficiently discretized in frequency.

Finally, for any $L \in \mathbb{N}$, the family

$$\{\psi_{\vec{p},L,n}^{\text{high}}, \ 2^{L/2}\psi_{\vec{p}}^{\text{low}}, \ \psi_{\vec{p},j,L,n}^{\text{a}} \ : \ \vec{p} \in \mathbb{Z}^2, \ 0 \leqslant j \leqslant J-1, \ 0 \leqslant n < 2^L\} \tag{6.3}$$

is a LP frame in $\mathcal{B}_\pi$, with constant $A = 2^{L/2}$.

With $\mathcal{W}^{\text{low}} = (\psi^{\text{low}} * f)$, $\mathcal{W}_{j,L,n}^{\text{a}} = (\psi_{j,L,n}^{\text{a}} * f)$ and $\mathcal{W}_{l,n}^{\text{high}} = (\psi_{l,n}^{\text{high}} * f)$, the reconstruction formula reads

$$f(\vec{p}) = \mathcal{W}^{\text{low}}(\vec{p}) + \sum_{j=0}^{J-1}\sum_{n=0}^{2^L-1} 2^{-L/2}\,\mathcal{W}_{j,L,n}^{\text{a}}(\vec{p}) + \sum_{n=0}^{2^L-1} 2^{-L/2}\,\mathcal{W}_{L,n}^{\text{high}}(\vec{p}). \tag{6.4}$$

In a multiselective context, this becomes

$$f(\vec{p}) = \mathcal{W}^{\text{low}}(\vec{p}) + \sum_{j=0}^{J-1}\sum_{n=0}^{2^\ell-1} 2^{-\ell/2}\,\mathcal{W}_{j,\ell,n}^{\text{a}}(\vec{p}) + \sum_{n=0}^{2^\ell-1} 2^{-\ell/2}\,\mathcal{W}_{\ell,n}^{\text{high}}(\vec{p}), \tag{6.5}$$

with $\ell = \ell(\vec{p},j)$ defined in (5.2).

# 7    Numerical Complexity

Let us discuss in this section the numerical complexity of the coefficients computation in the multiselective scheme.

The coefficients $W_{j,l,n}^{\text{a}}(\vec{b})$ can be evaluated by two methods. On the one hand, fast convolutions of the image $f$ with the filters $\psi_{j,l,n}^{\text{a}}$ can be perfomed simply using the FFT algorithm. Indeed, from the convolution theorem,

$$W_{j,L,n}^{\text{a}} \ = \ \mathcal{F}^{-1}\big[\,\mathcal{F}(f)\,\hat{\psi}_{j,L,n}^{\text{a}*}\,\big],$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ are the forward and inverse FFT on a $N{\times}N$ grid, and with similar formulas for $\mathcal{W}^{\text{low}}$ and $\mathcal{W}_{l,n}^{\text{high}}$. Hence, for a $N{\times}N$ pixel size image, for a fixed $j$ and given a maximal selectivity level $L$,

$$\sum_{l=0}^{L} 2^l\,O(N^2\log_2^2 N) \ = \ O\big((2^{L+1}-1)(N^2\log_2^2 N)\big)$$

operations are required to compute all the coefficients $W_{j,l,n}^{\text{a}}(\vec{b})$ for $l \in [0,L]$ and $n \in [0, 2^l-1]$.

On the other hand, we can use the pyramidal scheme presented in Section 4.3 (Figures 2 and 3). In the initial step, the $2^L$ functions $W_{j,L,n}^{\text{a}}(\vec{b})$ ($n \in [0, 2^L-1]$) are computed with the FFT and stored in memory. Next, the other selectivity levels are obtained with the help of the sequence $h_l$ in the linear combination (4.14).

Consequently, for fixed $j$ and $l$, and assuming $h$ has length $K$, the computation of the $2^l$ functions $W_{j,l,n}^{\text{a}}$ from the $2^{l+1}$ functions $W_{j,l+1,n}^{\text{a}}$ requires $O\big(2^l\min(K,2^l)N^2\big)$ operations,

since the periodization of $h$ in the $2^l$-length sequence $h_l$ implies that $\operatorname{supp} h_l = \min(2^l, K)$. Therefore,

$$O\big(2^L N^2 \log_2^2 N + \sum_{l=0}^{L-1} 2^l \min(2^l, K) N^2\big) \leqslant O\big(2^L N^2 \log_2^2 N + K(2^L - 1)N^2)\big) \qquad (7.1)$$

operations are thus necessary to evaluate the $2^{L+1} - 1$ images $W_{j,l,n}^{\mathrm{a}}(\vec{b})$ (with $l \in [0, L]$ and $n \in [0, 2^l - 1]$). If $K \ll \log_2^2 N$, which seems realistic for our purposes (e.g. $\log_2^2 N = 64$ for $N = 256$ and $K = 4$ for CDF filters), this second method is roughly twice faster than the first one, even if the asymptotic complexity is the same in $N$.

Notice finally that, even if no scaling rule can be exploited in the radial part of the wavelets (no simple relation between $j$ and $j-1$), further complexity improvement can be realized using the bandwidth of the true filters, as it is done for instance in the steerable filters radial design [5]. Indeed, for all $\vec{p}$, $L$ and $n$, functions $\psi_{\vec{p},L,n}^{\mathrm{high}}$ and $\psi_{\vec{p}}^{\mathrm{low}}$ and $\psi_{\vec{p},j,L,n}^{\mathrm{a}}$ belong to the spaces $\mathcal{B}_\pi$, $\mathcal{B}_{\pi/2^J}$ and $\mathcal{B}_{\pi/2^{J-1-j}}$, respectively, with $j \in [0, J-1]$. The different inverse FFTs used above can thus be computed on a subset of the original frequency grid for $W^{\mathrm{low}}$ and $W_{j,l,n}^{\mathrm{a}}$. The memory allocation in the recording of wavelet coefficients benefits also of this effect equivalent to a spatial downsampling of the coefficients. However, these technical improvements will not be pursued in the rest of this paper. We prefer to restrict our explanations on the main ideas concerning the multiselective scheme rather than on the possible numerical optimizations of the final decomposition.

# 8 Image Denoising

As a first application, we propose in this section to observe how the multiselective scheme is able to clean noisy images. We do not claim to obtain the best denoising algorithm. Our aim is rather to compare, in a Littlewood-Paley decomposition context (LP frame), a fixed selectivity denoising, using directional wavelets with the same selectivity level, with an adaptive multiselective denoising.

We will start by describing the whole process for the fixed selectivity method using a soft-thresholding of wavelet coefficients [19] before the image reconstruction. Therelated procedure will be then extended to the composite frame associated with the multiselective scheme.

## 8.1 Fixed selectivity method

Let $f \in \mathcal{B}_\pi$ be an image corrupted by a Gaussian white noise of vanishing mean and variance $\sigma^2$, that is,

$$f_\sigma(\vec{x}) = f(\vec{x}) + \sigma \mathcal{N}(\vec{x}), \qquad (8.1)$$

where $f_\sigma$ is the noisy image and $\mathcal{N} \underset{i.i.d.}{\sim} N(0, 1)$.

We want to estimate $f$ from $f_\sigma$. In our method, we will always compare $f$ and its *estimator* $f_{\mathrm{e}}$ with the *Peak Signal to Noise Ratio* (PSNR) determined by $\mathrm{PSNR} = 20 \log_{10} 256/\sigma_{\mathrm{e}}$, assuming the quantification of $f$ has 256 gray levels. Here $\sigma_{\mathrm{e}}^2$ is the estimated noise variance $E[(f - f_{\mathrm{e}})^2]$ [4]. The higher the PSNR, the better our estimator $f_{\mathrm{e}}$.

---

[4]We assume that $E[(f - f_{\mathrm{e}})] = 0$, that is, $E[f_{\mathrm{e}}]$ correctly estimates the mean value of $f$ with no bias.

A common procedure to determine $f_e$ (see, for instance, Refs. [6, 19, 20]) is to decompose the image in a basis of functions, to threshold the computed coefficients at some specific levels, and finally to reconstruct an estimated image from the remaining coefficients. Since the pure image $f$ is rather concentrated in a limited number of coefficients, and the noise is spread uniformly on all coefficients, the effect of this thresholding is to separate the noise artifacts from the real signal features in each "band" of the transformation.

We propose here to apply this framework with the following algorithm:

- Fix $J$ and the selectivity level $L \in \mathbb{N}$.

- Given the noisy image $f_\sigma$ in (8.1), compute the coefficients

$$\mathcal{W}^{\mathrm{high}}_{L,n}(\vec{p}), \mathcal{W}^{\mathrm{low}}(\vec{p}) \text{ and } \mathcal{W}^{\mathrm{a}}_{j,L,n}(\vec{p})$$

for $0 \leqslant n < 2^L$, $\vec{p} \in \mathbb{Z}^2$ and $0 \leqslant j \leqslant J - 1$.

- Softly threshold the wavelet coefficients according to the following rules:

$$\widetilde{\mathcal{W}}^{\mathrm{a}}_{j,L,n}(\vec{p}) = T[\mu\,\sigma_{j,L,n}]\mathcal{W}^{\mathrm{a}}_{j,L,n}(\vec{p}), \tag{8.2}$$

$$\widetilde{\mathcal{W}}^{\mathrm{high}}_{L,n}(\vec{p}) \tag{8.3}$$

where $T[t]$ is the *soft thresholding* operator of threshold $t \geqslant 0$ defined by

$$T[t]u = \begin{cases} (|u| - t)\,\mathrm{sign}\,u\,, & \text{if } |u| > t, \\ 0\,, & \text{otherwise,} \end{cases} \tag{8.4}$$

and $\mu$ is a parameter to be adjusted later.

- Reconstruct $f_e$ with $\widetilde{\mathcal{W}}^{\mathrm{high}}_{L,n}, \mathcal{W}^{\mathrm{low}}$ and $\widetilde{\mathcal{W}}^{\mathrm{a}}_{j,L,n}$ according to (6.4).

The particular parameters $\sigma^{\mathrm{high}}_{L,n}$ and $\sigma_{j,L,n}$ appearing above stand for the standard deviations of, respectively, the high-frequency coefficients and the wavelet coefficients, when the input image consists only of noise, i.e., $f = \sigma \mathcal{N}$. Since the noise has no preferred direction, it is clear that $\sigma_{j,L,n}$ does not depend on the angular index $n$. In fact, since wavelet coefficients are obtained by simple linear transformations of $f$, as a consequence of the Wiener-Khintchine theorem (see, for instance, Ref. [21]), wavelet coefficients of $\sigma \mathcal{N}$ have also a zero-mean Gaussian distribution with

$$\sigma_{j,L,n} = \sigma \|\psi^{\mathrm{a}}_{j,L,n}\| = a_j^{-1}\,\sigma\,\|\psi^{\mathrm{a}}_{L,0}\|, \tag{8.5}$$

$$\sigma^{\mathrm{high}}_{L,n} = \sigma \|\psi^{\mathrm{high}}_{L,n}\|. \tag{8.6}$$

In other words, thresholding wavelet coefficients that are smaller than these values amounts to keeping values that have a high probability to be due to $f$ and not to noise. In the sequel, $\sigma^{\mathrm{high}}_{L,n}$, which is not independent of $n$ because of the restriction (6.1) of $\hat{\psi}^{\mathrm{high}}_{L,n}$ to $B_\pi$, has been approximated by $a_J^{-1}\,\sigma\,\|\psi^{\mathrm{a}}_{L,0}\| = 2\,\sigma_{J-1,L,0}$.

Finally, the parameter $\mu$, which controls the thresholding strength relatively to these standard deviations, has been empirically set to 2 to obtain interesting PSNR between $f$ and $f_e$.

## 8.2 Multiselective method

For the multiselective method, the thresholding process has to take into account the selectivity level. It is easy to prove that, for our choice of wavelet (see Section 4.4), the variance $\sigma_{j,l,n}$ is constant as soon as the support of $\varphi_{l,n}(\kappa)$ is strictly included in $[0, 2\pi)$. This is a simple consequence of the $L^2$-normalization of the function $\phi^{\mathbb{R}}$ which generates $\varphi$ by periodization.

However, even if this behavior is independent of the scale $j$, we may conjecture that the pure image is less and less directional when $j$ decreases. We follow in fact the work of E. Candès and D. Donoho [8, 9] on the multiscale geometric study of $C^2$-edges in images. At small scales, these are well described by very elongated atoms, while at large scales, more isotropic functions are better adapted. In addition, in comparison to the noise, at a fixed $0 \leqslant j < J$, points $\vec{b}$ with high $\ell(\vec{b}, j)$ must be more numerous since many substructures inside real images define curved and straight edges.

This is confirmed by the results presented in Figure 7. For each $j \in [0, 4]$ ($J = 5$), the distribution of points $\vec{b}$ with selectivity level $l \in [0, 4]$, that is

$$\varpi(l, j) = N^{-2} \#\{\vec{b} : \ell(\vec{b}, j) = l\},$$

has been determined for two images of size $N^2 = 256^2$ : the `Lena` picture (Figure 8(a)) and a purely noisy image, i.e., our previous $\mathcal{N}(\vec{x})$. The observation of Figures 7(a) and 7(b) shows us that the `Lena` picture has globally a higher percentage of points with high selectivity level for any $j$ than the noise image. In addition, for small $j$, both images display percentages more spread on smaller $l$.

The noise image has also many more points associated to $l = 0$ for all $j$. The ratio of the percentages of the noise image and of the `Lena` picture (Figure 7(c)) confirms this effect for small values of $l$ too, with particularly high ratios in high frequencies (large $j$). To conclude this analysis, noise seems to favor small selectivity levels comparing to real images, and this trend is stronger in high frequencies.

Therefore, in our previous thresholding procedure, we propose to add a new thresholding factor taking into account our statement:

- Fix $J$ and $L \in \mathbb{N}$ and compute the coefficients of the noisy image, as before.

- Determine $\ell \equiv \ell(\vec{b}, j)$ from (5.2).

- Softly threshold the wavelet coefficients according the following rules:

$$\widetilde{\mathcal{W}}_{j,\ell,n}^{\mathrm{a}}(\vec{p}) = T[\mu \, \gamma_{j,\ell} \, \sigma_{j,l,n}] \mathcal{W}_{j,\ell,n}^{\mathrm{a}}(\vec{p}), \tag{8.7}$$

$$\widetilde{\mathcal{W}}_{\ell,n}^{\mathrm{high}}(\vec{p}) = T[\mu \, \gamma_{-1,\ell} \, \sigma_{l,n}^{\mathrm{high}}] \mathcal{W}_{\ell,n}^{\mathrm{high}}(\vec{p}), \tag{8.8}$$

  with $\gamma_{j,l}$,

$$\gamma_{j,l} = \lambda^{2^j \frac{L-l}{L}}, \tag{8.9}$$

  where $\lambda > 1$ is a parameter which tunes the thresholding operation on low selectivity levels. In agreement with the results presented in Figure 7, the additional thresholding factor $\gamma_{j,l}$ increases with $2^j \propto a_j^{-1}$ and decreases when the selectivity level $\ell$ is high.

- Reconstruct $f_{\mathrm{e}}$ with $\widetilde{\mathcal{W}}_{\ell,n}^{\mathrm{high}}, \mathcal{W}^{\mathrm{low}}$ and $\widetilde{\mathcal{W}}_{j,\ell,n}^{\mathrm{a}}$ according to (6.5).

(a) Distribution of selectivity for Lena

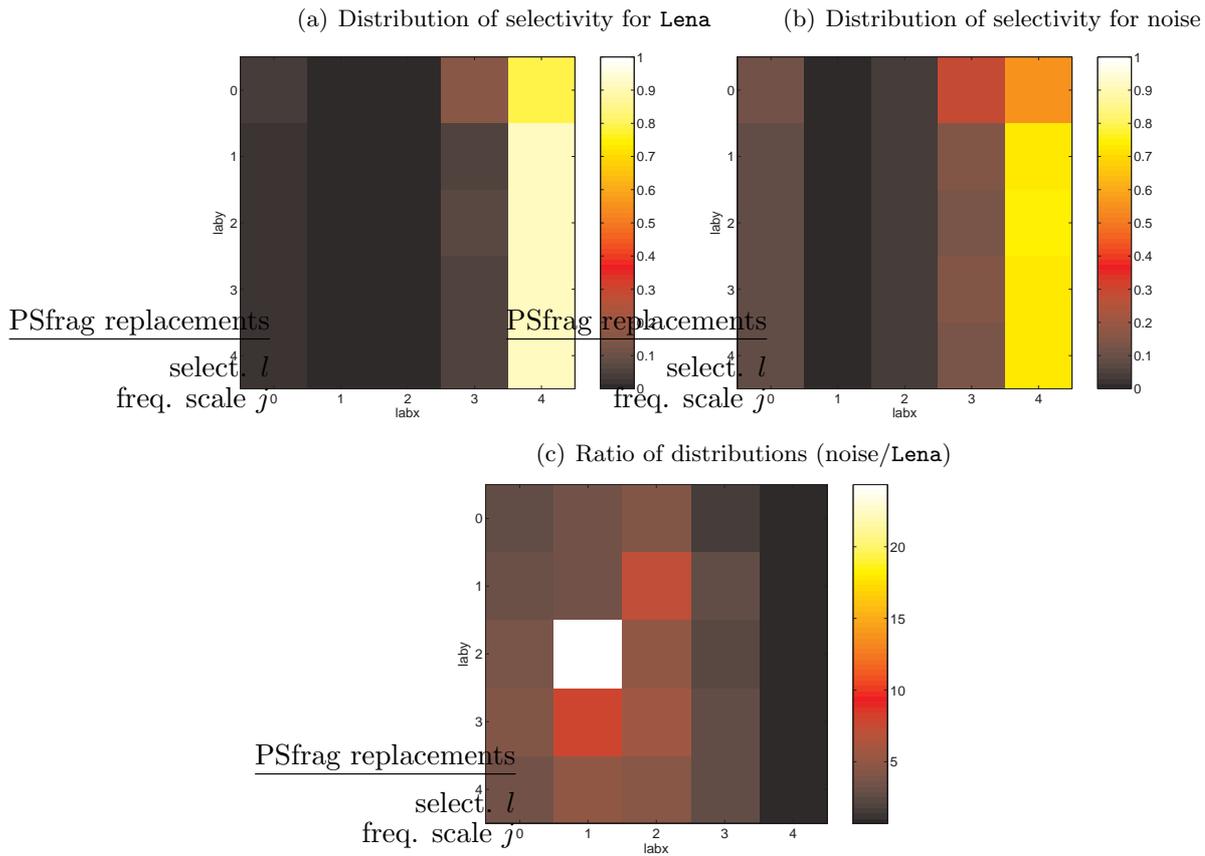(b) Distribution of selectivity for noise

(c) Ratio of distributions (noise/Lena)

Figure 7: (a) Distribution of selectivity levels $\varpi(l,j)$ for the Lena picture (Figure 9(a)) with $l \in [0,4]$ and $j \in [0,4]$. (b) Same measure but for purely noisy image $\mathcal{N}$. (c) Ratio of noise and Lena images percentages.

Notice that the computation of $\ell(\vec{p}, j)$ is performed on the noisy coefficients, but it is fully equivalent to evaluate it on the thresholded coefficients. Indeed, since $\ell(\vec{p}, j)$ corresponds to the selectivity level for which one orientation maximizes all the ratios of (5.2) for all $l$ and $n$, two situations may arise if wavelet coefficients are thresholded. First, after this thresholding, all the ratios are zero and the value of $\ell$ has no effect on the reconstruction since it simply disappears. Second, at least one coefficient is different from zero in (5.2), and since thresholding preserves the order of ratios above the threshold, $\ell$ is unchanged compared to its computation in the non-thresholded situation.

## 8.3  Results
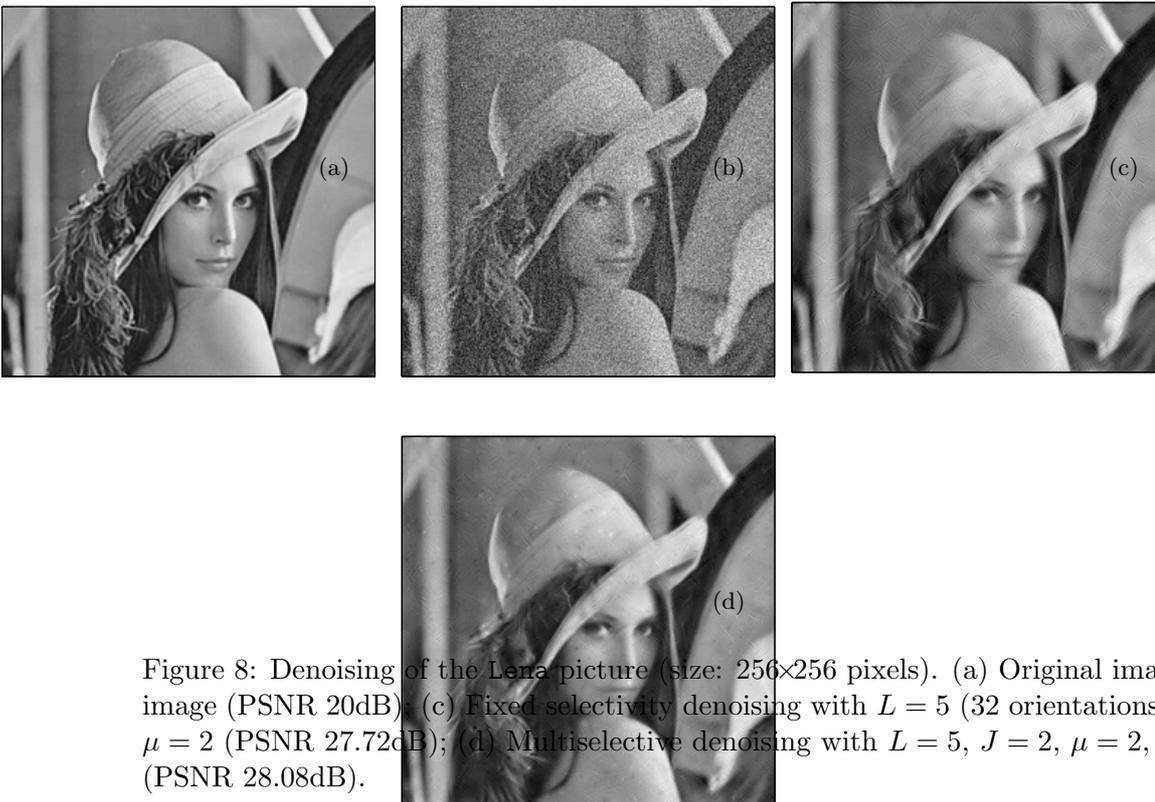


Figure 8: Denoising of the Lena picture (size: 256×256 pixels). (a) Original image; (b) Noisy image (PSNR 20dB); (c) Fixed selectivity denoising with $L = 5$ (32 orientations), $J = 3$, and $\mu = 2$ (PSNR 27.72dB); (d) Multiselective denoising with $L = 5$, $J = 2$, $\mu = 2$, and $\lambda = 1.05$ (PSNR 28.08dB).

We have tested our denoising method on two $256 \times 256$ images with 256 gray levels: the familiar Lena picture (Figure 8(a)) and the cameraman image (Figure 10(a)). In both cases, we have added an artificial Gaussian noise of standard deviation $\sigma = 25.6$, giving PSNRs of 20dB relatively to the original images. For the two denoisings, we have chosen the parameters $L = 5$, $J = 3$ and $\mu = 2$.

**Lena results :** For $\lambda = 1.05$, the multiselective scheme gives a slightly better PSNR (28.08dB, Figures 8(d) and 9(d)) than that of the fixed selectivity method (27.72dB, Figures 8(c) and 9(c)). However, we may remark, for instance, that more isotropic features, such as Lena's right nostril or the tip of her nose, are better preserved in the multiselective procedure. The smooth areas, like the right cheek or the forehead, have fewer reconstruction artifacts.



Figure 9: Zoom on the four images of Figure 8, with the same captions.

**Cameraman results :** For $\lambda = 1.04$, the multiselective PSNR (26.46dB, Figure 10(d)) is again better than that of the fixed selectivity (26.32dB, Figure 10(c)). Isotropic features like the cameraman's right eye and right ear, or the camera fixings, are also better defined. Artifacts decrease in the black area of the cameraman's coat.

## 9 Nonlinear Approximations

As a second application, we focus now on *nonlinear approximations* of images. In short, this technique consists in decomposing an image and rebuilding it only from a certain number of
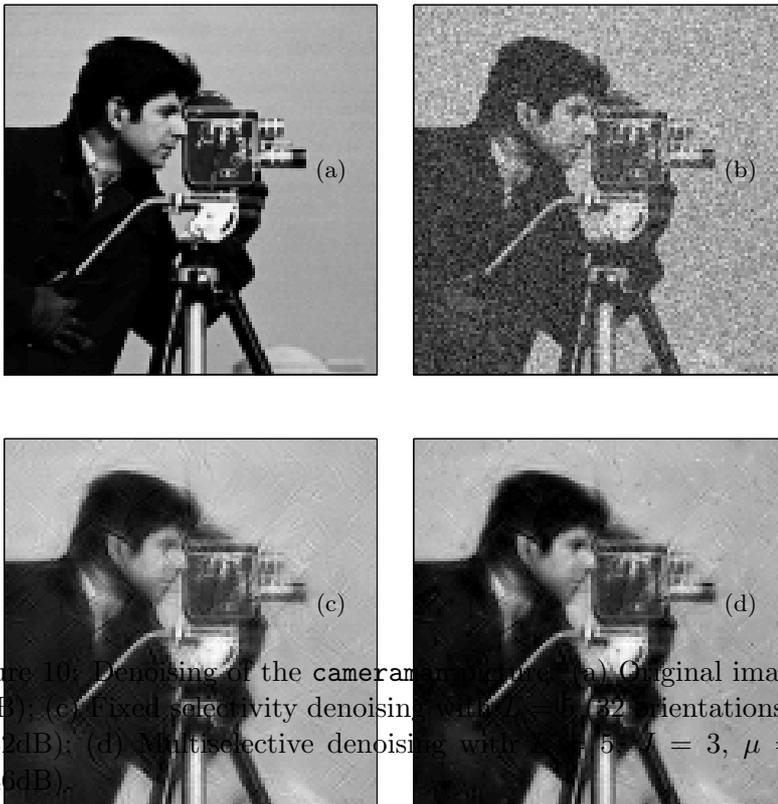
Figure 10: Denoising of the `cameraman` picture. (a) Original image; (b) Noisy image (PSNR 20dB); (c) Fixed selectivity denoising with $L = 5$ (32 orientations), $J = 3$, and $\mu = 2$ (PSNR 26.32dB); (d) Multiselective denoising with $L = 5$, $J = 3$, $\mu = 2$, and $\lambda = 1.04$ (PSNR 26.46dB).

its "highest" coefficients. After briefly reviewing the general definitions of this method, we will show how the multiselective scheme obtains better approximated images than the fixed selective method by saving up coefficients on less directional image features.

## 9.1 Definitions

Given a frame $\mathcal{F} = \{\psi_\zeta \in L^2(\mathbb{R}^2)\}$ where $\zeta$ stands collectively for all the wavelet parameters, we define the $N$-term nonlinear approximation of a function $f \in L^2(\mathbb{R}^2)$ by

$$f_N = \sum_{k=1}^{N} \langle \psi_{\zeta_k} | f \rangle \, \tilde{\psi}_{\zeta_k}, \tag{9.1}$$

where $\widetilde{\mathcal{F}} = \{\tilde{\psi}_\zeta \in L^2(\mathbb{R}^2)\}$ is the dual frame of $\mathcal{F}$. The parameters $\{\zeta_k\}$ result from a reordering of the indices $\zeta$ such that

$$m_{\zeta_k} := \|\psi_{\zeta_k}\|^{-1} |\langle \psi_{\zeta_k} | f \rangle| \geqslant m_{\zeta_{k+1}}, \quad \forall k \in \mathbb{N}.$$

The value $m_\zeta$ is called the *magnitude* of the coefficient $\langle \psi_\zeta | f \rangle$.

Unlike the case of orthogonal bases [23], for frames it is not guaranteed that $f_N$ is the best $N$-term nonlinear approximation. We will assume, however, that the error $\Theta_f(N) = \|f - f_N\|_2$ is globally decreasing with $N$. Indeed, if $\mathcal{F}$ has frame bounds $\mathsf{m}$ and $\mathsf{M}$ ($0 < \mathsf{m} \leqslant \mathsf{M}$), that is,

$$\mathsf{m}\|g\|^2 \leqslant \sum_{k=1}^{\infty} |\langle \psi_{\zeta_k} | g \rangle|^2 \leqslant \mathsf{M}\|g\|^2, \quad \forall g \in L^2(\mathbb{R}^2),$$

by taking $g = f - f_N$ and using the biorthogonality of $\mathcal{F}$ and $\widetilde{\mathcal{F}}$, it is easy to see that

$$\mathsf{M}^{-1/2} \, h_f(N) \leqslant \Theta_f(N) \leqslant \mathsf{m}^{-1/2} \, h_f(N), \tag{9.2}$$

with $h_f(N) = (\sum_{k=N+1}^{\infty} |\langle \psi_{\zeta_k} | f \rangle|^2)^{1/2}$. If $\|\psi_\zeta\|$ is constant for all $\zeta$, then $h_f$ is a strictly decreasing function and $\Theta_f$ is bounded by two decreasing functions. In addition, the control provided by inequality (9.2) is better if the frame bounds are close.

Our main objective is now to use nonlinear approximations for comparing the fixed selectivity and the multiselective methods.[5] However, the two frames do not have the same number of elements. Therefore, we define the $\tau\%$-term nonlinear approximation (with $\tau \in [0, 100]$) as the approximation obtained with $N = \lfloor \frac{\tau}{100} M \rfloor$ of the "best" terms, with $M$ representing the total number of elements in the frame. We will work also scale by scale in the $\tau\%$-term counting in order to highlight the directional effects of the two procedures.

## 9.2 Results

To evaluate the fixed selectivity and the multiselective methods, we analyse the image of a sunflower field (Figure 11). This picture presents directional objects, like the sticks and the leaves of the plants, as well more isotropic features like the dark center of the flowers. In addition, due to the angle of view of the camera, these elements appear at various scales, depending on their distance to the objective.

---

[5]Assuming we can "count" the positions, in view of the discretization occuring for bandlimited functions $f \in \mathcal{B}_\pi$

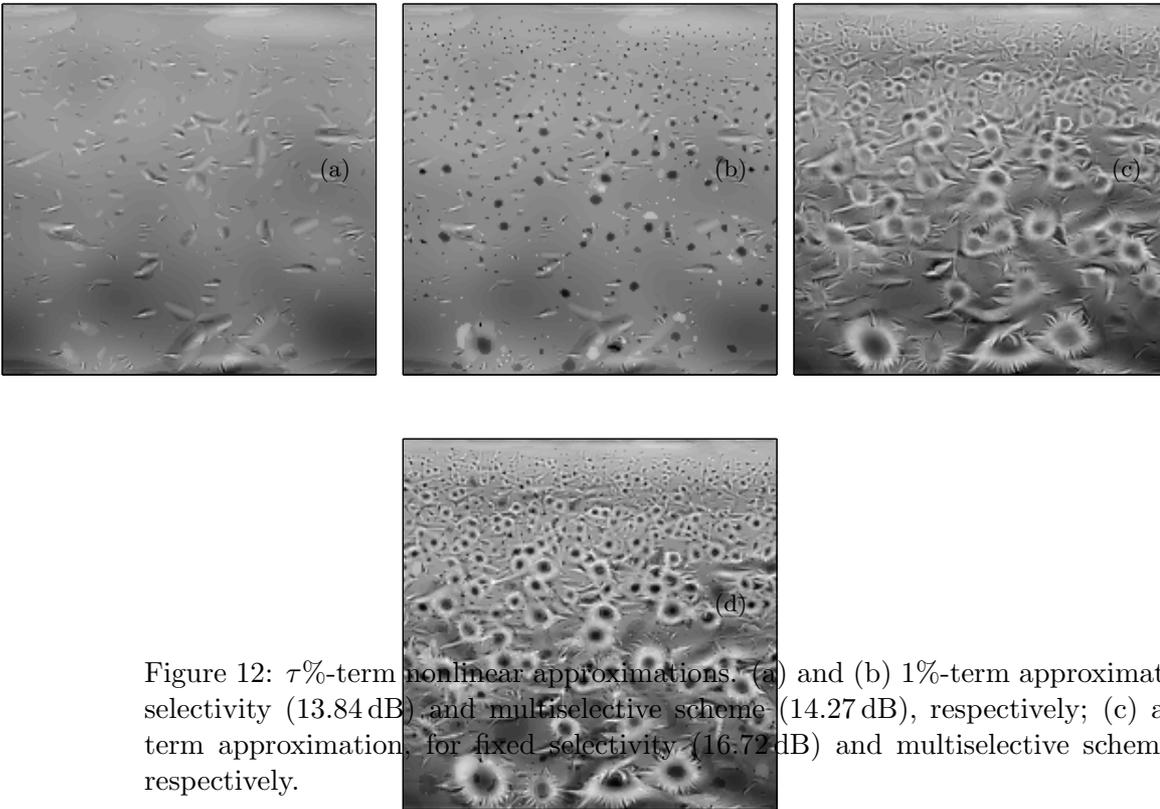Figure 11: Sunflower field picture, original image (size: 256×256 pixels).



Figure 12: $\tau$%-term nonlinear approximations. (a) and (b) 1%-term approximation, for fixed selectivity (13.84 dB) and multiselective scheme (14.27 dB), respectively; (c) and (d) 10%-term approximation, for fixed selectivity (16.72 dB) and multiselective scheme (18.22 dB), respectively.

Figures 12(a) and 12(b) show nonlinear approximations obtained for 1% of the total number of terms in the fixed and adaptive methods, respectively. In each case, we use $L = 4$ (16 orientations) with $J = 5$ number of scales. The corresponding qualities of the approximations, expressed in PSNR, are equal to $13.84\,\mathrm{dB}$ and $14.27\,\mathrm{dB}$. We can observe that, without losing the main directional objects, the adaptive method displays most of the dark centers of the flowers, while they are completely absent in the fixed selectivity method. This effect can be tested at higher percentages. For instance, for 10%-term approximations, the fixed selectivity gives a PSNR of $16.72\,\mathrm{dB}$ (Figure 12(c)), while the adaptive one yields a quality of $18.22\,\mathrm{dB}$ (Figure 12(d)). This phenomenon may be explained by the the number of coefficients needed to render an object. For instance, if a feature at point $\vec{x}$ and scale $a_j$ corresponds to a selectivity level $\ell(\vec{x}, j) = L - 1$, the multiselective scheme saves up $2^L - 2^{L-1} = 2^{L-1}$ coefficients, compared to a fixed selectivity decomposition of level $L$. These may then be used to describe other features, so that the result is better.

## 10  Related Work

As mentioned in the introduction, many techniques have been designed these last twenty years to improve the (sparse) representation of image features like $C^2$ curves, textures, ... The main idea is to go beyond the limitations of 2-D orthogonal separable bases (as in the usual discrete wavelet transform) which are unable to produce a good expansion of piecewise-smooth 2-D signals [7]. In this section we compare our approach to some other interesting techniques.

Our multiselective scheme can first be linked to the curvelet transform approach [8, 9] by considering the *anisotropy ratio* of both methods. For the curvelet transform, the filters obey a *parabolic* scaling rule that sparsifies the corresponding representation of any $C^2$ curve. Roughly speaking, at a scale $2^{-j}$, a curvelet has a *width* $\propto 2^{-j}$ and a *length* $\propto 2^{-j/2}$. Similarly, in the frequency domain, a curvelet has a support mainly contained in a circular sector of radial *width* $\propto 2^j$ and of angular *length* $\propto 2^{j/2}$ (see Ref. [9] for details). The parabolic name comes then from the fact that, whatever the space,

$$width \;\approx\; length^2.$$

In our case, using the same measurements, Sections 4.2 and 4.4 shown us that the multiselective wavelets $\psi_{j,l,n}^{\mathrm{a,d}}$ are contained in the frequency plane in a circular sector of radial width $\propto 2^j$ and, in rough approximation, of angular length $\propto 2^j 2^{-l}$ when $l$ is high enough to make the wavelets conical. In the end, for our scheme

$$width \approx length\, 2^l,$$

which reflects the flexibility of our wavelet parametrization. The selectivity selection (5.2) of $\ell(\vec{b}, j)$ allowed by Proposition 5.1 makes the choice of $l = \ell(\vec{b}, j)$ adaptive with respect of the local image geometry. Other features than $C^2$ curves can thus be conveniently represented.

Beyond this simple measurement comparison, we may remark also that the curvelet transform and the multiselective analysis do not share the same kind of reconstruction formula. Indeed, curvelets realize a tight frame of $L^2(\mathbb{R}^2)$ and so, the synthesis filters correspond to the analysis filters. In our case, Proposition 5.1 relies on the LP frame property of the multiselective analysis. In that case, the reconstruction is performed by summation of the wavelet

coefficients as in (5.3), whereas position parameters $\vec{b}$ are identified with points $\vec{x}$ in the image. In other words, the synthesis filter is here a simple Dirac distribution instead of the usual convolution used in tight frame reconstruction.

Second, Kingsbury, Seleskik and Baraniuk [24, 25] have introduced a very nice framework to achieve a directional decomposition of images, namely, the Dual-Tree complex wavelet Transform (or DTT). In this work, two classical discrete wavelet trees are combined, with filters forming (approximate) Hilbert pairs. This last property guarantees that *(i)* the final decomposition is more shift-invariant than the usual discrete wavelet transform (DWT); *(ii)* a selectivity of the filters can be defined outside the vertical, horizontal and (almost) diagonal directions of the DWT; and *(iii)* the redundancy factor is only 4 (and $2^d$ is a $d$-dimensional space).

This formalism is extended by Chaux et al. [26] into a $M$-band extension. Roughly speaking, this improvement leads to a refinement in $M \times M$ "octaves" of the frequency plane partitioning associated to the previous dual-tree WT. It is interesting to notice, in the denoising application of Ref. [26], that different choices of the band number $M$ lead to different denoising qualities, with best results obtained for high $M$.

At frequency scale $j$, the support of each directional filter of the $M$-bands DTT is made of two antipodal boxes of size $(2^j/M) \times (2^j/M)$ in the frequency plane. In other words, compared to the multiselective framework, $M$ is linked to $2^L$ with $L$ the maximal selectivity level. However, no adaptive gathering of these frequency boxes is currently proposed in the $M$-band DTT to access, for instance, the $M'$-band DTT decompositions for $M' < M$. This point precludes any further comparison with our the multiselective framework.

Finally, let us mention the work of Flesia et al. on the Ridgelet packets. This approach develops in a first step an orthonormal basis of $L^2(\mathbb{R}^2)$ by defining new *ridgelets* filters in polar frequency plane. This construction is very similar to that we adopt in (4.17) and (4.18). In short, the frequency angular part of these ridgelets corresponds to elements of periodized Lemarié scaling functions and periodized Meyer wavelets, while the radial part is the Fourier transform of Meyer wavelets in $L^2(\mathbb{R})$. Next the authors create a polar partioning of the frequency domain that best matches the image features. This is achieved by generating *ridgelet packets* in frequency by choosing separately their radial and angular parts, respectively, in a wavelet packets dictionary and in a wavelet packets dictionary or a cosine packets dictionary. A best basis strategy is then applied on the packet construction and the optimal partitioning corresponds to that which minimizes an additive measure of "entropy", or sparsity [12].

Although ridgelet packets form an adaptive orthonormal basis of $L^2(\mathbb{R}^2)$, this scheme has a serious limitation, in that it lacks a translation parameter in the filter generation. Consequently, the resulting decomposion is not truly local. Its application is then more restricted to image having global directional properties. Notice that, in order to solve this issue, the authors suggest to partition the spatial domain in boxes where the ridgelet packet procedure could be developed, following the definition of first generation curvelets [8].

Our multiselectivity decomposition of images does not provide an orthonormal basis or a tight frame of $L^2(\mathbb{R}^2)$. Its redundancy is very high but, in the end, angular adaptivity and locality can cohabit.

# 11   Conclusion and Perspectives

We have presented a new (Littlewood-Paley) decomposition of real images, based on the concept of angular multiselectivity.

The idea is that the angular selectivity of the wavelet should be adapted to the degree of isotropy of the analyzed point. Highly directional wavelets are needed for reproducing correctly sharply oriented features, but may constitute a hindrance at points around which the image is roughly isotropic. Thus, as always in wavelet analysis, the emphasis is on the *local* character of the procedure: The analysis tool must be adapted in a dynamical way to the local features of the image, and orientation is the relevant characteristic in the present context.

Finally, as a test of the new concept, we have shown in two applications, namely image denoising and $\tau\%$-term nonlinear approximation, that the multiselective scheme presents a clear improvement compared to a nonadaptive fixed selectivity method.

To conclude this paper, we remark that further research should be performed on the selectivity selection (5.2). Rather than selecting the directional frame that best matches the image features, which is equivalent to seek the $l$ that maximizes the $\ell^\infty$ norm of the angular coefficients (with respect to $n \in [0, 2^l)$), we could rather search for the $l$ that minimizes the $\ell^1$ norm of the same coefficients. This sparsity criterion can lead to interesting results, for instance, for the $\tau$-term approximation of Section 9. In addition, we expect that the efficiency of the multiselective method will increase with the size of the image. The more pixels we have, the more degrees of freedom we have in choosing selectivity levels.

## Acknowledgments

## References

[1] J-P. Antoine, P. Carrette, R. Murenzi and B. Piette, Image analysis with two-dimensional continuous wavelet transform, *Signal Processing* **31** (1993), 241–272.

[2] J-P. Antoine and R. Murenzi, Two-dimensional directional wavelets and the scale-angle representation, *Signal Processing* **52** (1996), 259–281.

[3] J-P. Antoine, R. Murenzi, and P. Vandergheynst, Two-dimensional wavelet analysis in image processing, *Int. J. Imag. Syst. Tech.* **7** (1996) 152–165.

[4] J-P. Antoine, R. Murenzi, and P. Vandergheynst, Directional wavelets revisited: Cauchy wavelets and symmetry detection in patterns, *Appl. Comput. Harmon. Anal.* **6** (1999) 314–345.

[5] A. Karasaridis and E. Simoncelli, A filter design technique for steerable pyramid image transforms, in *Int'l Conf. Acoustics Speech and Signal Processing*, (Atlanta GA), May 1996.

[6] E. P. Simoncelli, Bayesian denoising of visual images in the wavelet domain, in *Bayesian Inference in Wavelet Based Models*, P. Müller and B. Vidakovic, eds., 291–308, Springer-Verlag, New York, 1999.

[7] M. N. Do, P. L. Dragotti, R. Shukla, and M. Vetterli, On the compression of two-dimensional piecewise smooth functions, in *IEEE Int. Conf. on Image Proc. – ICIP '01*, Thessaloniki, Greece, Oct. 2001.

[8] E. Candès and D. Donoho, Curvelets: A surprisingly effective nonadaptive representation for objects with edges, in *Curve and Surface Fitting*, Nashville, TN, 1999. eds. L. L. Schumaker et al., Vanderbilt University.

[9] E. Candès, L. Demanet, D. Donoho, and L. Ying, Fast discrete curvelet transforms, *SIAM Multiscale Model. Simul.*, **5** (2006) 861–899.

[10] L. Demanet and L. Ying, Wave atoms and sparsity of oscillatory patterns, preprint, June 2006, submitted.

[11] A. G. Flesia, H. Hel-Or, A. Averbuch, E. J. Candès, R. R. Coifman and D. L. Donoho. Digital implementation of ridgelet packets, *Beyond Wavelets*, J. Stoeckler and G. V. Welland eds., Academic Press, 2002.

[12] R. R. Coifman and M. V. Wickerhauser. Entropy Based Algorithms for Best Basis Selection, *IEEE Transactions on Information Theory*, **32**:712–718, March 1992.

[13] J-P. Antoine, R. Murenzi, P. Vandergheynst, and S.T. Ali, *Two-dimensional Wavelets and Their Relatives*, Cambridge University Press, Cambridge (UK), 2004.

[14] M. Holschneider, Wavelet analysis on the circle, *J. Math. Phys.* **31**, 39–44, 1990.

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

[16] B. Torrésani, *Analyse continue par ondelettes*, InterÉditions/CNRS Éditions, Paris, 1995.

[17] M. Unser, A. Aldroubi, and M. Eden, B-Spline signal processing: Part I—Theory, *IEEE Trans. Signal Process.* **41**:821–833. (1993)

[18] A. Cohen, I. Daubechies, and J. Feauveau, Biorthogonal bases of compactly supported wavelets, *Commun. Pure Appl. Math.* **45** (1992) 485–560.

[19] D. L. Donoho, De-noising by soft-thresholding, *IEEE Trans. Information Theory* **41** (1995) 613–627.

[20] J. Starck, E. Candès, and D. Donoho, The curvelet transform for image denoising, *IEEE Trans. Image Process.* **11** (2002) 670–684.

[21] S. Haykin, *Communication Systems*, 4th ed, New York, NY: Wiley, 2000.

[22] L. Jacques. *Ondelettes, Repères et Couronne Solaire.* PhD thesis, Université catholique de Louvain, Louvain-la-Neuve, Belgium, 2004.

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

[24] N.G. Kingsbury, Complex wavelets for shift invariant analysis and filtering of signals, *Appl. Comput. Harmon. Anal.*, **3**, pp. 234–253, 2001.

[25] I. W. Selesnick, R. G. Baraniuk, and N. Kingsbury. The dual-tree complex wavelet transform - A coherent framework for multiscale signal and image processing, *IEEE Signal Processing Magazine*, **22**(6):123-151, November 2005.

[26] Caroline Chaux, Laurent Duval and Jean-Christophe Pesquet. Image Analysis Using a Dual-Tree M-Band Wavelet Transform, *IEEE Transactions on Image Processing*, **15**(8):2397-2412, Aug. 2006.