Multiselective Pyramidal Decompositions of Images: How to exploit wavelets with adaptive angular selectivity

Laurent Jacques*,[1] and Jean-Pierre Antoine[2]

Abstract

Many techniques have been devised these last ten years to add an appropriate directionality concept in decompositions of images from the (affine) transformations of a small set of atomic functions (e.g. directional wavelets, steerable filters, curvelets, wave atoms). Generally features which are best represented are straight lines (as those defining contours of objects), smooth curves (e.g. curvelets processing) or oriented textures (e.g. wave atoms). However, real images present also a set of details less oriented and more isotropic (like corners, spots, texture components, ...). This paper aims at developing one possible 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 (linear) frame of directional wavelets with variable angular selectivity. Inside such a decomposition, wavelets inherit some particularities of the (biorthogonal) circular multiresolution framework. This simple link qualifies our method of multiselectivity analysis. 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.

Keywords: image decomposition, multiselectivity, directional wavelet, frames, multiresolution analysis

* Laurent.Jacques@uclouvain.be, corresponding author

[1] Communications and Remote Sensing Laboratory, Université catholique de Louvain (UCL), B-1348 Louvain-la-Neuve, Belgium

[2] Institut de Physique Théorique (FYMA), Université catholique de Louvain, Chemin du Cyclotron 2. B-1348 Louvain-la-Neube, Belgium

Electronic version of an article published as IJWMIP, Vol. 5(5), pp. 785-814 (2007), doi:10.1142/S0219691307002051

CONTENTS

I. 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,[1], [2] that of Simoncelli *et al.* on oriented steerable filters,[3], [4] that of Candès and Donoho on the curvelet representation,[5], [6] and that of Demanet and Ying on wave atoms.[7]

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[18]), the advantages of the composite multiselective frame over a non-adaptive fixed selectivity method.

II. THE FRAMEWORK

Let us describe an *image* by a continuous function f in $\mathbb{R}^2$, i.e. a *light intensity* function in every point $\vec{x} \in \mathbb{R}^2$ of the image. If $f \in L^2(\mathbb{R}^2) = \{g(\vec{x}) : \|g\| = \int_{\mathbb{R}^2} |g(\vec{x})|^2 \mathrm{d}^2\vec{x} < \infty\}$ and given an admissible wavelet $\psi \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$, the continuous wavelet transform[8] of f is

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

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

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

where * is the complex conjugation and $\psi_{\vec{b},a,\theta}$ is an $L^1(\mathbb{R}^2)$-normalized copy of ψ, 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_θ is the usual 2×2 rotation matrix, and the hat denotes the standard Fourier transform on $L^2(\mathbb{R}^2)$, that is, $\hat{f}(\vec{k}) = \int_{\mathbb{R}^2} f(\vec{x})\, e^{-i\vec{k}\vec{x}}\, \mathrm{d}^2\vec{x}$, with inverse $f(\vec{x}) = (2\pi)^{-2} \int_{\mathbb{R}^2} \hat{f}(\vec{k})\, e^{i\vec{k}\vec{x}}\, \mathrm{d}^2\vec{k}$. We will forget in the sequel the index of $\mathcal{W}_f$ and we will write simply $\mathcal{W}$.

From now on, we take for ψ a *directional* wavelet,[2], [8] that is, a wavelet whose frequency support is essentially contained in a convex cone with apex at the origin. In addition, we assume that ψ is separable in polar coordinates:

$$\hat{\psi}(\vec{k}) = \rho(k)\, \varphi(\kappa), \tag{2.4}$$

where $\vec{k} \equiv (k, \kappa),\ k = |\vec{k}|,\ \kappa = \arg \vec{k}$, ρ is a function in $L^2(\mathbb{R}, k\mathrm{d}k)$ and φ is a positive function in $L^2(S_1, \mathrm{d}\kappa)$. Notice that if φ has a compact support in S_1 with a width strictly inferior to π, the directional wavelet ψ is said *conical* since $\operatorname{supp} \hat{\psi}$ is then exactly contained in a convex cone..

A directional wavelet is characterized by its *angular selectivity* (or *Angular Resolving Power*[8]), 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 frequency representation of our separable wavelet ψ is given in Figure 1.

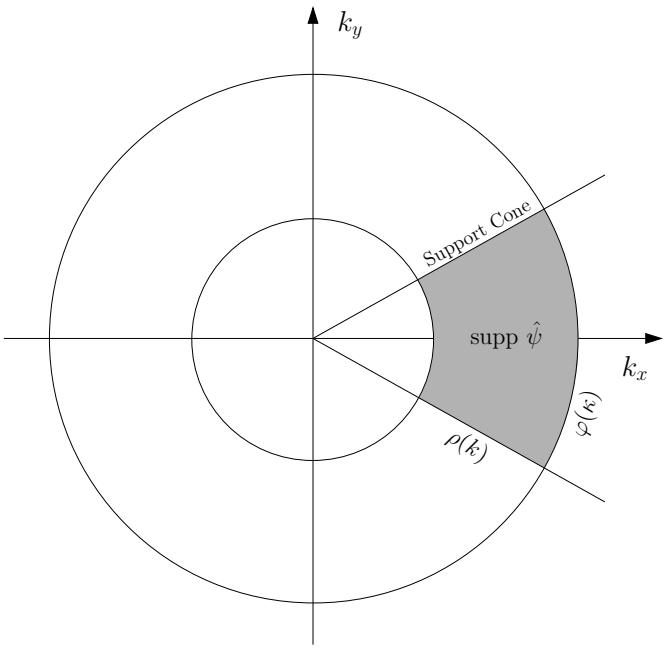


Fig. 1. Frequency representation of ψ

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 equation 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 2π-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, the 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 φ_θ localized around $\theta \in S_1$.

It is now very tempting to add a dilation in this scheme. If φ corresponds to the periodization of a function $\nu : \mathbb{R} \to \mathbb{R}$, we may for instance "dilate" φ by dilating ν,[9] that is,

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

Let us replace now φ_θ 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 ϵ, we have a wavelet

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

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

III. 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,[10] 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 $[0, 2\pi)$).

A. 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,

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

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 case, there exists, for each space V_l ($l \in \mathbb{Z}$), a complementary subspace W_l, with scaling and translation properties similar to those of V_l, such that $V_{l+1} = V_l \oplus W_l$. 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

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

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

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

We turn now to the biorthogonal case. We take a second multiresolution analysis, with subspaces $\{\widetilde{V}_l\}$ and $\{\widetilde{W}_l\}$, and generating functions $\widetilde{\phi}$ and $\widetilde{\xi}$, respectively, satisfying the biorthogonality relations

$$\widetilde{V}_l \perp W_l \quad \text{and} \quad \widetilde{W}_l \perp V_l. \tag{3.4}$$

In other words, $\widetilde{\phi}$ and $\widetilde{\xi}$ define Riesz bases $\{\widetilde{\phi}_{l,n}(t) = 2^{l/2}\widetilde{\phi}(2^l t - n),\, n \in \mathbb{Z}\}$ and $\{\widetilde{\xi}_{l,n}(t) = 2^{l/2}\widetilde{\xi}(2^l t - n),\, n \in \mathbb{Z}\}$ of $\widetilde{V}_l$ and $\widetilde{W}_l$, respectively, and verify the biorthogonality relations

$$\begin{array}{rclcrcl} \langle \xi_{l,n'} | \, \widetilde{\phi}_{l,n} \rangle & = & 0, & & \langle \phi_{l,n} | \, \widetilde{\xi}_{l,n'} \rangle & = & 0, \\ \langle \phi_{l,n'} | \, \widetilde{\phi}_{l,n} \rangle & = & \delta_{n,n'}, & & \langle \xi_{l',n'} | \, \widetilde{\xi}_{l,n} \rangle & = & \delta_{l,l'}\, \delta_{n,n'}, \end{array} \tag{3.5}$$

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

$$g(\omega) = e^{-i\omega}\, \widetilde{h}^*(\omega + \pi), \quad \widetilde{g}(\omega) = e^{-i\omega}\, h^*(\omega + \pi). \tag{3.6}$$

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

Given any $f \in L^2(\mathbb{R})$ and any $L \in \mathbb{Z}$, the reconstruction formulas read

$$f = \sum_{n \in \mathbb{Z}} c_L[n]\, \widetilde{\phi}_{L,n} + \sum_{l=L}^{+\infty} \sum_{n \in \mathbb{Z}} d_l[n]\, \widetilde{\xi}_{l,n} \tag{3.7}$$

$$= \sum_{l \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} d_l[n]\, \widetilde{\xi}_{l,n}, \tag{3.8}$$

with $c_l[n] = \langle \phi_{l,n} | \, f \rangle$ and $d_l[n] = \langle \xi_{l,n} | \, f \rangle$ for $l, n \in \mathbb{Z}$.

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

$$\begin{aligned} c_l[n] &= I\, h * c_{l+1}\,[2n], \\ d_l[n] &= I\, g * c_{l+1}\,[2n], \\ c_{l+1}[n] &= \widetilde{h} * U c_l\,[n] + \widetilde{g} * U d_l\,[n], \end{aligned} \tag{3.9}$$

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

$$Uq[n] = \begin{cases} q[n/2] & \text{if } n \text{ is even,} \\ 0 & \text{if } n \text{ is odd,} \end{cases} \tag{3.10}$$

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

$$u * v[n] \;=\; \sum_{m \in \mathbb{Z}} u[n-m]\, v[m]. \tag{3.11}$$

B. Circular MRA

Consider a multiresolution analysis of $L^2(\mathbb{R})$, such as described in the previous section, with generating functions[1] $\phi^{\mathbb{R}}(t)$ and $\xi^{\mathbb{R}}(t)$ well localized in space, in the sense that $\exists\, \epsilon > 0$ such that $|\phi^{\mathbb{R}}(t)|, |\xi^{\mathbb{R}}(t)| \leqslant C(1 + |t|)^{-1-\epsilon}$. The

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

periodized version of $\phi^{\mathbb{R}}_{l,n}$ is

$$\phi_{l,n}(t) = \sum_{m\in\mathbb{Z}} \phi^{\mathbb{R}}_{l,n}(t+m) \tag{3.12}$$

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 $\widetilde{\xi}_{l,n}$.

Hence, from the subspaces $V^{\mathbb{R}}_l, W^{\mathbb{R}}_l \in L^2(\mathbb{R})$, 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 $\widetilde{V}_l, \widetilde{W}_l$ induce the periodic subspaces $\widetilde{V}_l, \widetilde{W}_l$. It is easy to see that these new spaces are finite dimensional, since

$$\phi_{l,n} = \phi_{l,n+2^l r}, \quad \xi_{l,n} = \xi_{l,n+2^l r}, \tag{3.13}$$

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,[10] that is, $\sum_{m\in\mathbb{Z}} \phi^{\mathbb{R}}(t-m) = 1$ for all $t\in\mathbb{R}$. 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 the orthogonality (3.5) between V_l and $\widetilde{W}_l$ (resp. between $\widetilde{V}_l$ and W_l) is inherited from the one between V_l and $\widetilde{W}_l$ (resp. between $\widetilde{V}_l$ and W_l),[10] that is,

$$\langle \xi_{l,n} |\, \widetilde{\phi}_{l,n'} \rangle_{C_1} = 0, \quad \text{etc.}, \quad l,l' \in \mathbb{N},\ 0 \leqslant n < 2^l,\ 0 \leqslant n' < 2^{l'}. \tag{3.14}$$

In conclusion, since the dilations are performed before the periodization of the nonperiodic functions, ϕ, ξ, $\widetilde{\phi}$ 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

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

$$\xi_{l-1,0}(t) = \sqrt{2} \sum_{n=0}^{2^l-1} g_l[n]\, \phi_{l,n}(t), \tag{3.16}$$

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

The reconstruction formula (3.7) now becomes

$$f = c_0[0] + \sum_{l\in\mathbb{N}} \sum_{n=0}^{2^l-1} d_l[n]\, \widetilde{\xi}_{l,n}, \tag{3.17}$$

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 (3.9) become

$$\begin{aligned} c_{l-1}[n] &= I\, h_l \star c_l\,[2n], \\ d_{l-1}[n] &= I\, g_l \star c_l\,[2n], \\ c_{l+1}[n] &= \widetilde{h}_{l+1} \star U c_l\,[n] + \widetilde{g}_{l+1} \star U d_l\,[n], \end{aligned} \tag{3.18}$$

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

$$u \star v[n] \;=\; \sum_{m=0}^{P-1} u[n - m \bmod P]\, v[m], \tag{3.19}$$

for two sequences u and v of length P.

IV. 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.[17] As we 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 linear directional frame.

A. Directional Frames

We begin by recalling the definition of a linear directional frame. According to the theory of dyadic (or half-continuous) frames,[8], [11] the continuous wavelet transform given in (2.1) can be discretized in its parameters a and θ, 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 $\theta_n = \frac{2\pi}{K} n$ ($0 \leqslant n < K$, $K \in \mathbb{N}^*$), assume that the *frame property* is satisfied, that is, there are two positive constants m, M such that

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

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

$$f(\vec{x}) = \sum_{j\in\mathbb{Z}} \sum_{n=0}^{K-1} \; \mathcal{W}_{j,n} * \widetilde{\psi}_{a_j,\theta_n}(\vec{x}), \quad \text{a.e. on } \mathbb{R}^2, \tag{4.2}$$

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

A particular case ('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

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

B. Angular Multiselectivity

As we have seen in Section II, when the wavelet ψ 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 ϕ and a wavelet ξ, we apply the discretization of Section IV-A and project $R_{\vec{b},j} := R_{\vec{b},a_j}$ onto the functions $\varphi_{l,n}(\kappa) = \phi_{l,n}\big(\frac{\kappa}{2\pi}\big)$ and $\eta_{l,n}(\kappa) = \xi_{l,n}\big(\frac{\kappa}{2\pi}\big)$. This gives rise to the new coefficients

$$\mathcal{W}^{\mathrm{a}}_{j,l,n}(\vec{b}) = \langle \varphi_{l,n} |\, R_{\vec{b},j} \rangle_{S_1}, \tag{4.6}$$

$$\mathcal{W}^{\mathrm{d}}_{j,l,n}(\vec{b}) = \langle \eta_{l,n} |\, R_{\vec{b},j} \rangle_{S_1}, \tag{4.7}$$

$$\mathcal{W}^{\mathrm{i}}_{j}(\vec{b}) = \langle \varphi_{0,0} |\, R_{\vec{b},j} \rangle_{S_1} = \langle\, 1 \,|\, R_{\vec{b},j} \rangle_{S_1}, \tag{4.8}$$

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

$$\hat{\psi}^{\mathrm{a}}_{l,n}(\vec{k}) = \rho(k)\, \varphi_{l,n}(\kappa), \tag{4.9}$$

$$\hat{\psi}^{\mathrm{d}}_{l,n}(\vec{k}) = \rho(k)\, \eta_{l,n}(\kappa), \tag{4.10}$$

$$\hat{\psi}^{\,\mathrm{i}}(\vec{k}) = \hat{\psi}^{\mathrm{a}}_{0,0}(\vec{k}) = \rho(k), \tag{4.11}$$

where 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 reads

$$\psi^{\mathrm{a}}_{\vec{b},j,l,n}(\vec{x}) = a_j^{-2}\, \psi^{\mathrm{a}}_{l,n}\big(\frac{\vec{x} - \vec{b}}{a_j}\big), \tag{4.12}$$

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

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

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

Let us end this section by noticing that, for each selectivity level $l \in \mathbb{N}$, the family $\{\psi^{\mathrm{a}}_{j,l,n}\}$ may easily generate a linear frame.

Proposition 1: Let $a_j = a_0 2^{-j}$ be a dyadic scale discretization. If $\sum_{j\in\mathbb{Z}} \rho(a_j k) = 1$ a.e. for $k \in \mathbb{R}_+$, then, for any $l \in \mathbb{N}$, the family $\{\psi^{\mathrm{a}}_{j,l,n} : j \in \mathbb{Z},\, 0 \leqslant n < 2^l\}$ is a linear 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 all $t \in \mathbb{R}$. Indeed, with $u = \frac{\kappa}{2\pi}$,

$$\sum_{j\in\mathbb{Z}} \sum_{n=0}^{2^l-1} \hat{\psi}^{\mathrm{a}}_{j,l,n}(\vec{k}) = \sum_{j\in\mathbb{Z}} \sum_{n=0}^{2^l-1} \rho(a_j k)\, \varphi_{l,0}(\kappa - n\tfrac{2\pi}{2^l}) = \sum_{n=0}^{2^l-1} \phi_{l,0}(u - \tfrac{n}{2^l}) \tag{4.13}$$

$$= 2^{\frac{l}{2}} \sum_{m\in\mathbb{Z}} \sum_{n=0}^{2^l-1} \phi^{\mathbb{R}}(2^l u + 2^l m - n) \tag{4.14}$$

$$= 2^{\frac{l}{2}}, \tag{4.15}$$

and (4.3) is satisfied.

C. Recursion Formulas

For any $l \in \mathbb{N}^*$, φ and η verify simple extensions of the scaling rules (3.15) and (3.16) :

$$\varphi_{l-1,0}(\kappa) = \sum_{n=0}^{2^l-1} h_l[n]\, \varphi_{l,n}(\kappa), \tag{4.16}$$

$$\eta_{l-1,0}(\kappa) = \sum_{n=0}^{2^l-1} g_l[n]\, \varphi_{l,n}(\kappa). \tag{4.17}$$

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

$$\varphi_{l+1,n}(\kappa) = \sum_{n'=0}^{2^l-1} \widetilde{h}_{l+1}[n-2n']\, \varphi_{l,n}(\kappa) + \sum_{n'=0}^{2^l-1} \widetilde{g}_{l+1}[n-2n']\, \eta_{l,n}(\kappa) \tag{4.18}$$

If we project (4.16), (4.17) and (4.18) onto $R_{\vec{b},j}$, we obtain the following relations for the decomposition:

$$\mathcal{W}^{\mathrm{a}}_{j,l-1,n}(\vec{b}) = \sum_{n'=0}^{2^l-1} h^*_l[n'-2n]\, \mathcal{W}^{\mathrm{a}}_{j,l,n'}(\vec{b}) = \big(\bar{h}_l \star \mathcal{W}^{\mathrm{a}}_{j,l,\cdot}(\vec{b})\big)_{2n}, \tag{4.19}$$

$$\mathcal{W}^{\mathrm{d}}_{j,l-1,n}(\vec{b}) = \sum_{n'=0}^{2^l-1} g^*_l[n'-2n]\, \mathcal{W}^{\mathrm{a}}_{j,l,n'}(\vec{b}) = \big(\bar{g}_l \star \mathcal{W}^{\mathrm{a}}_{j,l,\cdot}(\vec{b})\big)_{2n}, \tag{4.20}$$

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

As for the reconstruction, we get

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

$$= \big(\widetilde{h}^*_{l+1} \star U\mathcal{W}^{\mathrm{a}}_{j,l,\cdot}(\vec{b})\big)_n + \big(\widetilde{g}^*_{l+1} \star U\mathcal{W}^{\mathrm{d}}_{j,l,\cdot}(\vec{b})\big)_n. \tag{4.21}$$

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 (3.10) and (3.18). The notation $(\,\cdot\,)_n$ means simply that we select the n^{th}

angular element of this convolution. These relations are summarized in Figs. 2 and 3 where the operator D is the downsampling operator which turns a sequence u of length $2P$ into a sequence $(Du)[n] = u[2n]$ of length P.

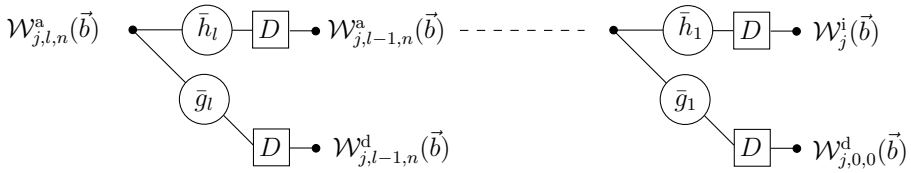


Fig. 2. Decomposition of the wavelet coefficients according to their angular selectivity

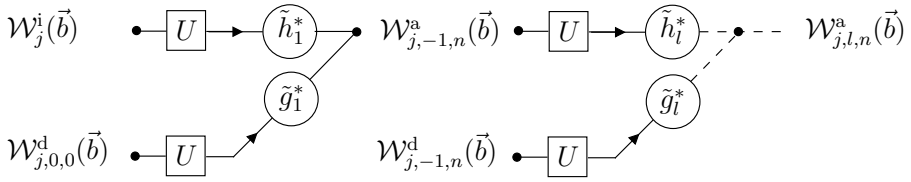


Fig. 3. Reconstruction of the wavelet coefficients according to their angular selectivity

Notice that (4.19), (4.20) and (4.21) are the exact counterpart of the usual recursion relations for wavelet coefficients.[10] They rely on the fact that, according to (4.16), (4.17), several wavelets of level l merge into a single wavelet at level $l-1$, i.e., with half the angular selectivity.

D. 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.[12] In particular, we use the results of Cohen-Daubechies-Feauveau[13] (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 I). This ensures that the resulting $\psi^{\mathrm{a}}_{l,n}$ (given below) is a conical wavelet with quadratic regularity on the edges of its conical frequency support.

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

$$\hat{\psi}^{\mathrm{a}}_{L,n}(\vec{k}) = \phi^{\mathbb{R}}(\log_2 k)\, \varphi_{L,n}(\kappa), \tag{4.22}$$

$$\hat{\psi}^{\mathrm{d}}_{L,n}(\vec{k}) = \phi^{\mathbb{R}}(\log_2 k)\, \eta_{L,n}(\kappa), \tag{4.23}$$

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

Remark that, choosing $\rho(k) = \phi^{\mathbb{R}}(\log_2 k)$, as in (4.22), 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 $\sum_{m\in\mathbb{Z}} \phi^{\mathbb{R}}(t-m) = 1$. From Proposition 1, the wavelets $\{\psi^{\mathrm{a}}_{j,l,n}\}$ constructed from (4.22) thus constitute a linear frame for each $l \in \mathbb{N}$.

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 I

CDF DIRECT AND DUAL FILTERS OF 3 AND 7 VANISHING MOMENTS.

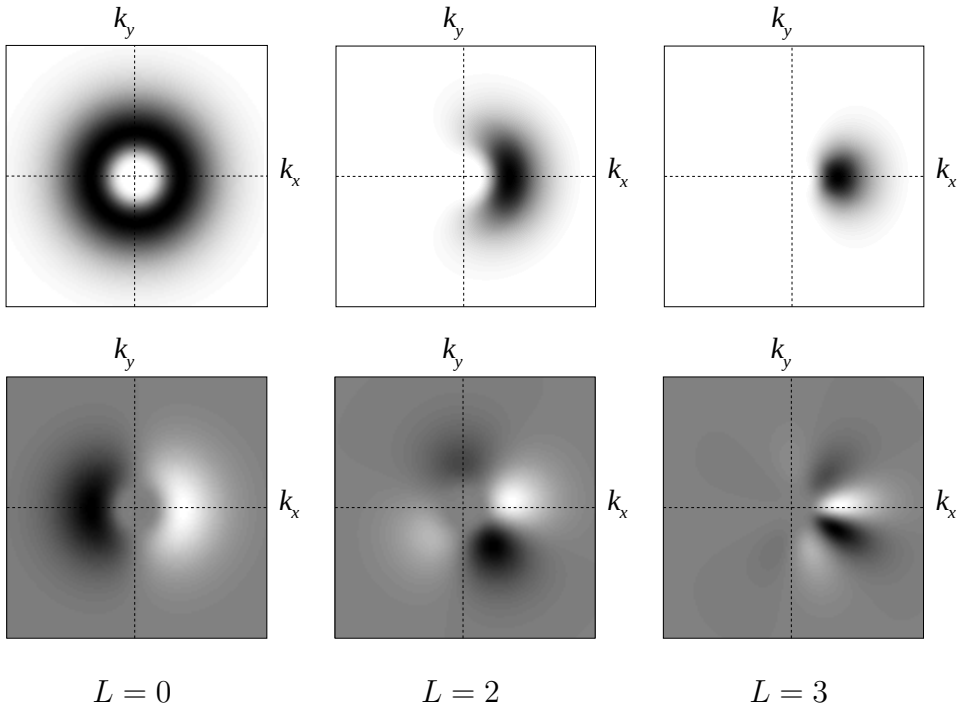


Fig. 4. Wavelets $\hat{\psi}^{\mathrm{a}}_{L,n}(\vec{k})$ (top row) and $\hat{\psi}^{\mathrm{d}}_{L,n}(\vec{k})$ (bottom) for several value of L and for $n = 0$.

The resulting function $\hat{\psi}^{\mathrm{a,d}}_{L,n}$ are presented in the frequency domain on Figure 4 for several value of L and for $n = 0$. Notice that the aperture of the supporting cones is actually decreasing with the increasing of L. As explained before, the aperture $2\alpha_L$ of the cone supporting $\hat{\psi}^{\mathrm{a}}_{L,n}$ and $\hat{\psi}^{\mathrm{d}}_{L,n}$ is proportional to 2^{-L}, which means that the angular selectivity of these wavelets growths with L. This behaviour is illustrated more clearly in Figure 5 for the CDF$-(3, 7)$ framework. Since $\operatorname{supp}\phi^{\mathbb{R}} = [-3/2, 3/2]$, α_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 $\psi^{\mathrm{a}}_{L,n}$ are conical[8] (i.e. $\alpha_L < \pi/2$) for $L \geqslant 3$.

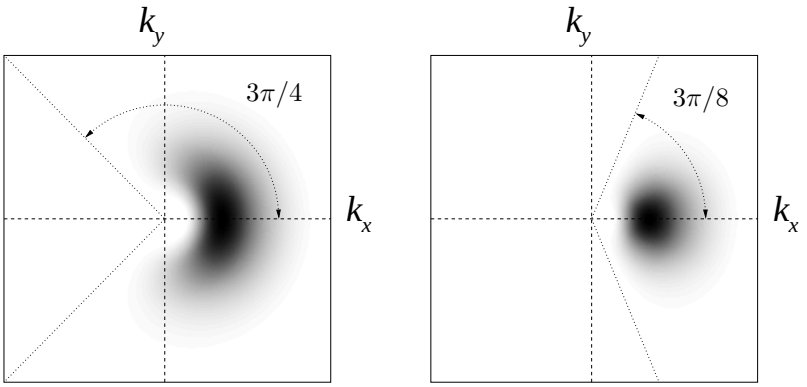


Fig. 5. $\hat{\psi}^{\mathrm{a}}_{2,0}(\vec{k})$ with $\alpha_2 = 3\pi/4$ (left). $\hat{\psi}^{\mathrm{a}}_{3,0}(\vec{k})$ with $\alpha_3 = 3\pi/8$; this is a conical wavelet (right).

V. BEST FRAME SELECTION

Starting from a selectivity level $L \in \mathbb{N}$, we have seen in Section IV-D that we can generate inductively $L+1$ frames with $\psi^{\mathrm{a}}_{l,n}$ ($l \in [0, L]$), characterized by an angular selectivity 2^l.

In particular, for each level $l \in [0, L]$, Proposition 1 shows that the wavelets $\psi^{\mathrm{a}}_{j,l,n}$ generate a frame of constant $A = 2^{l/2}$ which, according to (4.4), may be used to reconstruct the original image. In addition, as expressed in the following proposition, we can mix the different frames inside the same reconstruction formula.

Proposition 2: If ρ respect the same condition than in Prop.1, for all 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 by

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

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

$$\sum_{n=0}^{2^{l}-1} 2^{-\frac{l}{2}}\, \mathcal{W}^{\mathrm{a}}_{j,l,n}(\vec{x}) \;=\; \sum_{n=0}^{2^{l'}-1} 2^{-\frac{l'}{2}}\, \mathcal{W}^{\mathrm{a}}_{j,l',n}(\vec{x}) \;=\; \mathcal{W}^{\mathrm{i}}_j(\vec{x}), \tag{5.2}$$

for any $l, l' \in \mathbb{N}$.

The precious property unveiled by this last proposition provides us a new degree of freedom to adaptively describe images. 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.

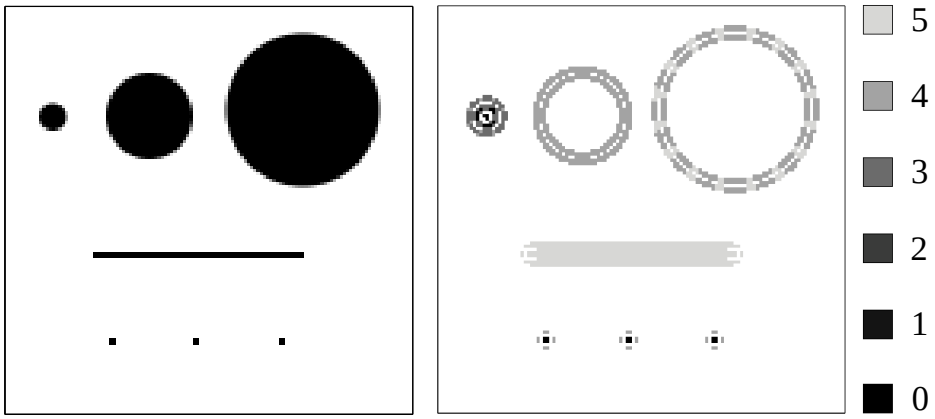
Therefore, we decide simply to choose the frame which offers the best match between the image and the wavelets, that is,

$$\ell(\vec{b}, j) \quad = \quad \operatorname*{arg\,max}_{l\in[0,L]} \max_{n\in[0,2^l)} \frac{|\langle \psi^{\mathrm{a}}_{\vec{b},j,l,n} | f \rangle|}{\|\psi^{\mathrm{a}}_{\vec{b},j,l,n}\|}. \tag{5.3}$$

The reconstruction procedure will be then defined by

$$f(\vec{x}) = \sum_{j\in\mathbb{Z}} \sum_{n=0}^{2^{\ell}-1} 2^{-\ell/2}\, \mathcal{W}^{\mathrm{a}}_{j,\ell,n}(\vec{x}), \quad \ell = \ell(\vec{x}, j), \tag{5.4}$$

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



(a) (b)

Fig. 6. Value of $\ell(\vec{x}, 4)$ on a toy example for $L = 5$.

On Figure 6(b), the computation of ℓ is presented for $j = J - 1 = 4$ (see next section for details on the scale discretization) and $L = 5$ on a toy image containing a set of simple geometric objects (Fig. 6(a)). Values of ℓ 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, ℓ follows closely and locally the directional aspect of the analysed objects. Indeed, the selectivity level ℓ 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 last object, the straight line, produces the maximum level as expected.

VI. 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 $\vec{p} = (p, q) \in \mathbb{Z}^2$. By the Shannon theorem, the Fourier transform of f_{d} is the 2π-periodization of $\hat{f}$, both in k_x and k_y directions. In other words, f is completely determined from the values of f_{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 $B_\pi = [-\pi, \pi) \times [-\pi, \pi)$. In this particular context, the linear frame condition (4.3) must hold only at points $\vec{k} \in B_\pi$.

However, since $a_j = a_0 2^{-j}$, $\rho(a_j k)$ is moving to high frequencies when j increases, and there is a maximal value $J > 0$ for j. Hence, in our choice of wavelet, since the quadratic spline ϕ is centered on the origin and has a support equal to $[-\frac{3}{2}, \frac{3}{2}]$, $\rho(k) = \phi(\log_2 k)$ will be centered on $k = 1$ 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 $\frac{\pi}{2}$ inside the support $[\frac{\pi}{8}, \pi]$, ensuring that $\psi^{\mathrm{a}}_{\vec{b},j,l,n} \in \mathcal{B}_\pi$ for $j < J$. Let us now gather together wavelets with $j \geqslant J$ in B_π by defining

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

$$= \chi_{B_\pi}(\vec{k})\, \varphi_{l,n}(\kappa) \sum_{j=J}^{\infty} \phi(j + \log_2 a_0 k), \tag{6.2}$$

where *high* stands for *high* frequency components and $\chi_{B_\pi}(\vec{k})$ is the characteristic function of B_π equals 1 if $\vec{k} \in B_\pi$ and 0 elsewhere.

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 which are larger than the original image f, that is

$$\hat{\psi}^{\,\mathrm{low}}(\vec{k}) = \sum_{j=-\infty}^{-1} \phi(\log_2 a_j k). \tag{6.3}$$

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

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

$$\{\psi^{\mathrm{high}}_{\vec{p},L,n},\ 2^{L/2}\psi^{\mathrm{low}}_{\vec{p}},\ \psi^{\mathrm{a}}_{\vec{p},j,L,n}\ :\ \vec{p} \in \mathbb{Z}^2,\ j \in [0, J-1],\ n \in [0, 2^L)\} \tag{6.4}$$

is a linear frame of constant $A = 2^{L/2}$ in B_π.

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

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

In a multiselective context, this becomes

$$f(\vec{p}) = \mathcal{W}^{\,\mathrm{low}}(\vec{p}) + \sum_{j=0}^{J-1} \sum_{n=0}^{2^\ell-1} 2^{-\frac{\ell}{2}}\, \mathcal{W}^{\mathrm{a}}_{j,\ell,n}(\vec{p}) + \sum_{n=0}^{2^\ell-1} 2^{-\ell/2}\, \mathcal{W}^{\mathrm{high}}_{\ell,n}(\vec{p}), \tag{6.6}$$

with $\ell = \ell(\vec{p}, j)$ defined in (5.3).

VII. 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 pretend to obtain the best denoising algorithm. Our aim is rather to compare, in a linear frame or Littlewood-Paley decomposition context, a fixed selectivity denoising, using directional wavelet 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[14] before the image reconstruction.

This procedure will be then extended to the composite frame associated with the multiselective scheme.

A. Fixed Selectivity Method

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

$$f_\sigma(\vec{x}) = f(\vec{x}) + \sigma n(\vec{x}), \tag{7.1}$$

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

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

A common procedure to determine $f_{\rm e}$ (see, for instance, Refs. [15], [4], [14]) 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 latter. Since the pure image f is rather concentred in a limited number of coefficients, and the noise spread uniformly on all coefficients, this thresholding has for effect 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_σ (7.1), compute the coefficients

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

  for $n \in [0, 2^L)$, $\vec{p} \in \mathbb{Z}^2$ and $j \in [0, J-1]$.
- Softly threshold the wavelet coefficients according the following rules:

  $$\widetilde{\mathcal{W}}_{j,L,n}^{\rm a}(\vec{p}) = T[\mu\,\sigma_{j,L,n}] \cdot \mathcal{W}_{j,L,n}^{\rm a}(\vec{p}), \tag{7.2}$$

  $$\widetilde{\mathcal{W}}_{L,n}^{\rm high}(\vec{p}) = T[\mu\,\sigma_{L,n}^{\rm high}] \cdot \mathcal{W}_{L,n}^{\rm high}(\vec{p}), \tag{7.3}$$

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

  $$T[t]\cdot u = \begin{cases} (|u|-t)\,\operatorname{sign} u\,, & \text{if } |u| > t, \\ 0\,, & \text{otherwise.} \end{cases} \tag{7.4}$$

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

The particular parameters $\sigma_{L,n}^{\rm high}$ 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 the noise, i.e., $f = \sigma n$. 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 simple linear transformation of f, as a consequence of the Wiener-Khintchine theorem (see, for instance, Ref. [16]), wavelet coefficients of σn have also a zero-mean Gaussian distribution with

$$\sigma_{j,L,n} \;=\; \sigma\|\psi_{j,L,n}^{\rm a}\| \;=\; a_j^{-1}\,\sigma\,\|\psi_{L,0}^{\rm a}\|, \tag{7.5}$$

$$\sigma_{L,n}^{\rm high} \;=\; \sigma\|\psi_{L,n}^{\rm high}\|. \tag{7.6}$$

In other words, thresholding wavelet coefficients which are below (a multiple of) these values amounts to keep values with high probabilities to be due to f and not to noise. In the sequel, $\sigma_{L,n}^{\rm high}$, which is not independent of n from the restriction of $\hat{\psi}_{L,n}^{\rm high}$ to B_π (6.1), has been approximated by $\sigma_{L,n}^{\rm high} \simeq a_J^{-1}\,\sigma\,\|\psi_{L,0}^{\rm a}\| = 2\,\sigma_{J-1,L,0}$.

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

The last parameter μ, controlling the thresholding strength relatively to these standard deviations, has been empirically set to 2 to obtain interesting PSNR between f and f_{e}.

B. 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 IV-D), $\sigma_{j,l,n}$ is constant as soon as the support of $\varphi_{l,n}(\kappa)$ is strictly included to $[0, 2\pi)$. This is a simple consequence of the L^2 normalization of $\phi^{\mathbb{R}}$ generating φ 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[5], [6] 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 more 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 of Figure 7. The percentage of points $\vec{b}$ with value $\ell(\vec{b}, j) = l$ for $0 \leqslant l \leqslant 4$ and $0 \leqslant j < J = 5$ has been determined for two images : the `Lena` picture (Fig. 8(a)) and a purely noisy image, i.e. our previous $n(\vec{x})$. 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 spreaded 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 (Fig. 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 favour 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 the highest selectivity level $L \in \mathbb{N}$.
- Given the noisy image f_σ of (7.1), compute the coefficients

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

  for $l \in [0, L]$, $n \in [0, 2^L)$, $\vec{p} \in \mathbb{Z}^2$ and $j \in [0, J-1]$.
- Determine $\ell(\vec{b}, j)$ from (5.3).
- Softly threshold the wavelet coefficients according the following rules:

$$\widetilde{\mathcal{W}}_{j,\ell,n}^{\mathrm{a}}(\vec{p}) = T[\mu\, \gamma_{j,\ell}\, \sigma_{j,l,n}] \cdot \mathcal{W}_{j,\ell,n}^{\mathrm{a}}(\vec{p}), \tag{7.7}$$

$$\widetilde{\mathcal{W}}_{\ell,n}^{\mathrm{high}}(\vec{p}) = T[\mu\, \gamma_{-1,\ell}\, \sigma_{l,n}^{\mathrm{high}}] \cdot \mathcal{W}_{\ell,n}^{\mathrm{high}}(\vec{p}), \tag{7.8}$$

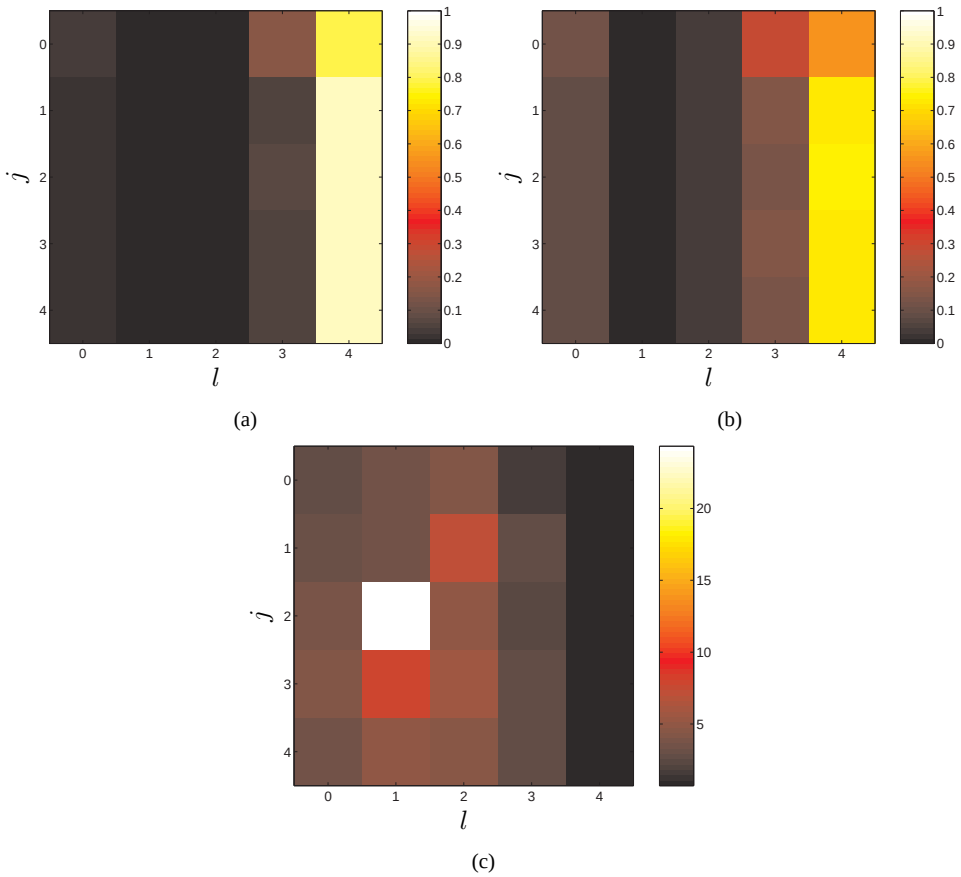


Fig. 7. (a) Percentage of points $\vec{b}$ in the Lena picture (Fig. 9(a)) with value $\ell(\vec{b}, j) = l$ for $0 \leqslant l \leqslant 4$ and $0 \leqslant j < J = 5$. (b) Same measure but for purely noisy image n. (c) Ratio of noise and Lena images percentages.

with $\gamma_{j,l}$ given by

$$\gamma_{j,l} = \lambda^{2^j \frac{L-l}{L}}, \tag{7.9}$$

where $\lambda > 1$ is a parameter which tunes the thresholding operation on low selectivity levels.

- Reconstruct f_{e} with $\widetilde{\mathcal{W}}^{\mathrm{high}}_{\ell,n}, \mathcal{W}^{\mathrm{low}}$ and $\widetilde{\mathcal{W}}^{\mathrm{a}}_{j,\ell,n}$ according to (6.6).

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 (5.3) 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 ℓ has no effect on the reconstruction since it simply disappears. Second, at least one coefficient is different of zero in (5.3), and since thresholding preserves the order of ratios above the threshold, ℓ is unchanged compared to its computation in the non-thresholded situation.

C. Results

We have tested our denoising method on two 256×256 images with 256 gray levels: the familiar Lena picture (Fig. 8(a)) and the cameraman image (Fig. 10(a)). In both cases, we have added an artificial Gaussian noise of

(a) (b) (c)

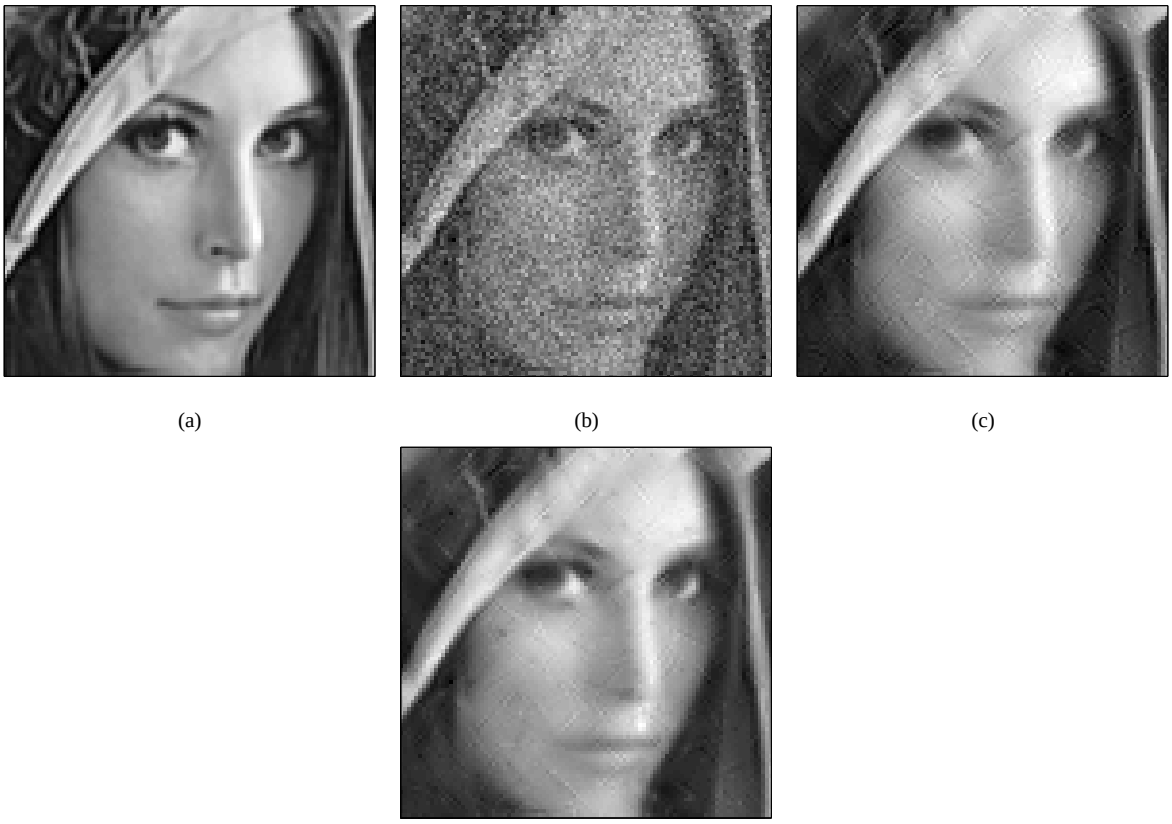
(d)

Fig. 8. Denoising of the `Lena` picture. (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).

standard deviation $\sigma = 256/10$, 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 less reconstruction artifacts.

Cameraman results : For $\lambda = 1.04$, the multiselective PSNR (26.46dB, Fig.10(d)) is again better than that of the fixed selectivity (26.32dB, Fig.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.



(a) (b) (c)

(d)

Fig. 9. Zooms on images of Figure 8. (a) Zoom on original image; (b) Noisy image; (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).

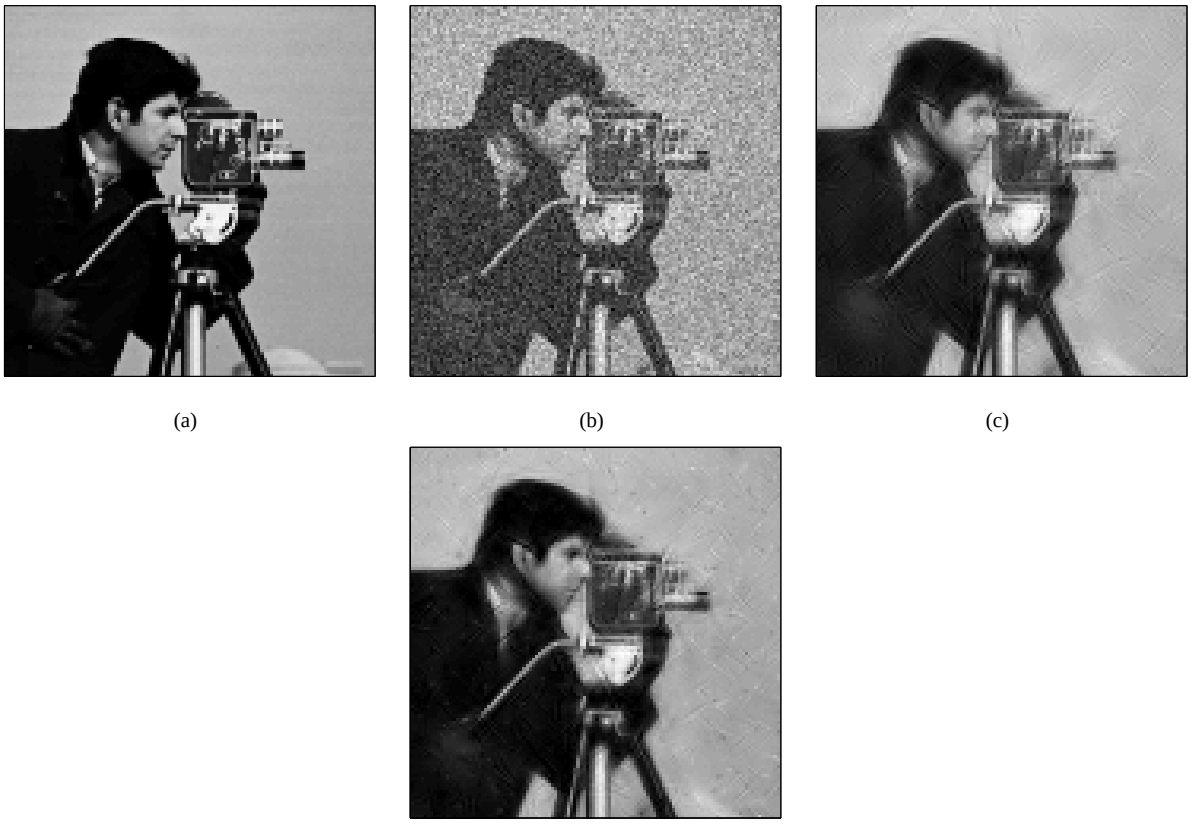
VIII. NONLINEAR APPROXIMATIONS

For our 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 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.

A. Definitions

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

$$f_N \;=\; \sum_{k=1}^{N} \langle \psi_{\zeta_k} | \, f \rangle \; \tilde{\psi}_{\zeta_k}, \tag{8.1}$$



(a) (b) (c)

(d)

Fig. 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).

where $\widetilde{\mathcal{F}} = \{\tilde{\psi}_\zeta \in L^2(\mathbb{R}^2)\}$ is the dual frame of $\mathcal{F}$. Parameters ζ_k are a reordering of the indices ζ 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_ζ is called the *magnitude* of the coefficient $\langle \psi_\zeta | f \rangle$.

Unlike the case of orthogonal bases,[18] it is not guaranteed for frames that f_N is the best N-term nonlinear approximation. We will assume however that the error $\Theta_f[N] = \|f - f_N\|$ is globally decreasing with N.

Our main objective is now to use nonlinear approximations to compare the fixed selectivity and the multiselective methods.[3] However, these two frames do not share the same number of elements. In consequence, 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.

Fig. 11. Sunflower field picture, original image

B. Results

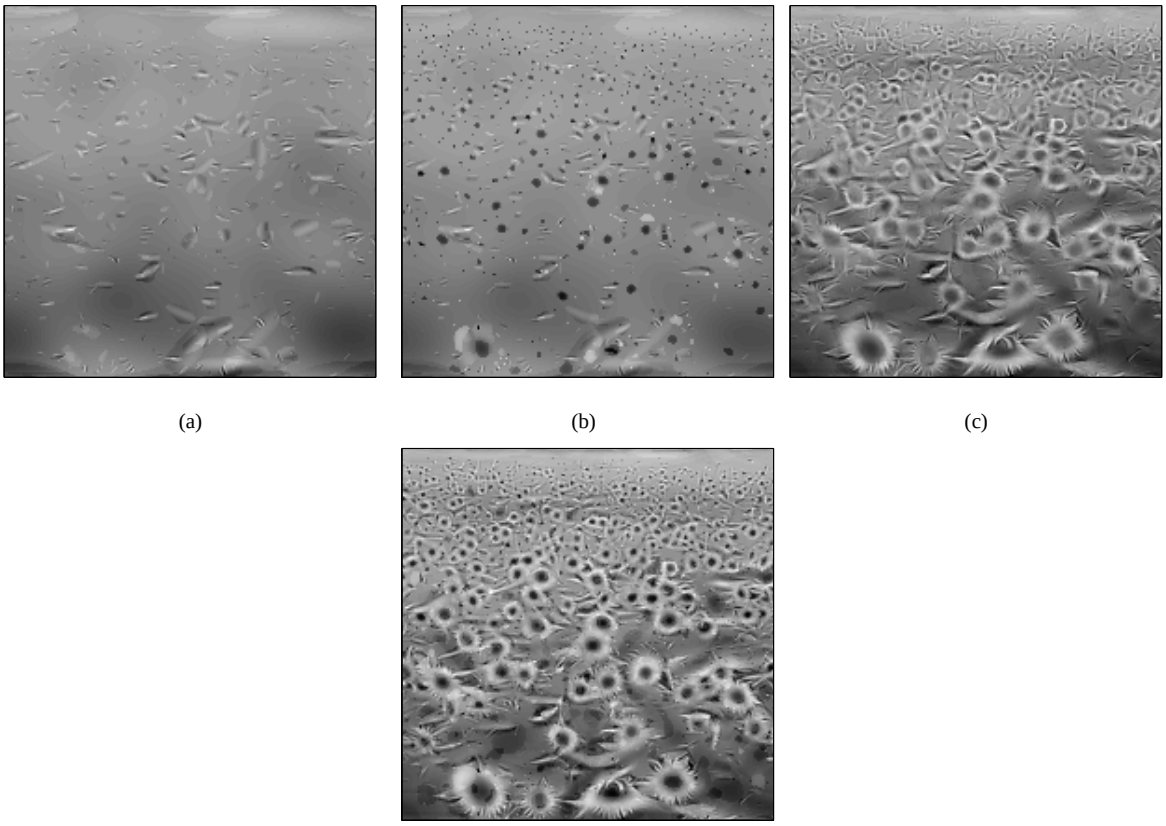
To evaluate the fixed selectivity and the multiselective methods, we analyse the image of a sunflower field (Fig. 11). This picture presents directional objects, like the sticks and the leaves of the plants, as well more isotropic features like the dark center of the flowers. In addition, due to the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

Figures 12(a) and 12(b) show nonlinear approximations obtained for 1% of the total number of terms in the fixed and adaptative 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 dB and 14.27 dB. We can observe that, without losing the main directional objects, the adaptative 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 dB (Fig. 12(c)), while the adaptative one yields a quality of 18.22 dB (Fig. 12(d)). The explanation to this phenomenon comes from the number of coefficients needed to render an object. To take an example, 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, which are then used to describe other features.

IX. CONCLUSION

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

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



(a) (b) (c)

(d)

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

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.

REFERENCES

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

[2] J-P. Antoine, R. Murenzi, and P. Vandergheynst, Directional wavelets revisited: Cauchy wavelets and symmetry detection in patterns, *Appl. Comput. Harmon. Anal.* **6** (199) 314–345.

[3] 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.

[4] 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.

[5] E. Candes 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.

[6] E. Candes, L. Demanet, D. Donoho and L. Ying. Fast Discrete Curvelet Transforms. SIAM Multiscale Model. Simul., to appear.

[7] L. Demanet, L. Ying. Wave atoms and sparsity of oscillatory patterns. to appear, June 2006

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

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

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

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

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

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

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

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

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

[17] Laurent Jacques. *Ondelettes, Repères et Couronne Solaire*. thèse, Université catholique de Louvain, Louvain-la-Neuve, Belgique, 2004.

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