Complex Wavelet Bases, Steerability, and the Marr-Like Pyramid

Dimitri Van De Ville *Member*, Michael Unser *Fellow*

***Abstract*— Our aim in this work is to tighten the link between wavelets, some classical image-processing operators, and David Marr's theory of early vision. The cornerstone of our approach is a new complex wavelet basis that behaves like a smoothed version of the Gradient-Laplace operator. Starting from first principles, we show that a single-generator wavelet can be defined analytically and that it yields a semi-orthogonal complex basis of $L_2(\mathbb{R}^2)$, irrespective of the dilation matrix used. We also provide an efficient FFT-based filterbank implementation.**

We then propose a slightly redundant version of the transform that is nearly translation-invariant and that is optimized for better steerability (Gaussian-like smoothing kernel). We call it the Marr-like wavelet pyramid because it essentially replicates the processing steps in Marr's theory of early vision. We use it to derive a primal wavelet sketch which is a compact description of the image by a multiscale, subsampled edge map. Finally, we provide an efficient iterative algorithm for the reconstruction of an image from its primal wavelet sketch.

***Index Terms*— wavelet design, steerable filters, feature extraction, primal sketch**

MULTISCALE transforms are powerful tools for signal and image processing, computer vision, and for modeling biological vision. A prominent example is the 1-D wavelet transform, which acts as a multiscale version of an Nth-order derivative operator, where N is the number of vanishing moments of the wavelet [1]. Its extension to multiple dimensions and to 2-D, in particular, is typically achieved by forming tensor-product basis functions. However, such separable wavelets are not well matched to the singularities occuring in images such as lines and edges which can be arbitrarily oriented and even curved. Consequently, there has been a considerable research effort in developing alternative multiscale transforms that are better tuned to the geometry of natural images. Notable examples of these "geometrical x-lets" include biologically-inspired 2-D Gabor transforms [2], wedgelets [3], ridgelets [4], [5], curvelets [6], [7], contourlets [8], bandelets [9], [10], directional wavelet frames [11], and directionlets [12]. A second category of methods takes advantage of the spectral separation into positive and negative frequency bands that can be achieved via a 1-D complex wavelet transform whose real and imaginary parts are in quadrature [13], [14], or, equivalently, via the Hilbert transform [15]. Proper combinations of positive and negative frequency bands in multiple dimensions then allow one to separate various orientations (e.g., six for the dual-tree wavelet transform [13], [16] or even more, using its M-band extension [17]). The last important class of multiscale transforms are 2-D directional filterbanks [18] and, most notably, steerable pyramids [19]–[21], which combine the structure of the Laplacian pyramid [22] with steerable filters [23], [24]. The derivation of the filters is not based on wavelets, but rather obtained through a numerical optimization process. Special constraints are imposed to ensure that the frequency response of the filters is essentially polar-separable and that the decomposition is simple to invert numerically (approximate tight-frame property).

The authors are with the Biomedical Imaging Group (BIG), EPFL, Lausanne, Switzerland. This work was supported by the Swiss National Science Foundation under Grant 200020-109415 and by the Center for Biomedical Imaging of the Geneva-Lausanne Universities and the EPFL, as well as the foundations Leenaards and Louis-Jeantet.

In this paper, we present an alternative approach based on an explicit analytical and spline-based formulation of complex wavelet bases of $L_2(\mathbb{R}^2)$. Special care is given to design basis functions that best match the properties of the visual system, in accordance with Marr's theory of early vision [25]. To motivate our construction, we specify a number of properties that are highly desirable and that are fulfilled, sometimes implicitly, by a number of classical image-processing algorithms:

- *Invariance*. We aim at invariance with respect to elementary geometric operations such as translation, scaling, and rotation. Traditional wavelet transforms only satisfy these properties to some extent, since a trade-off needs to be found between the conciseness of the representation—its (non-)redundancy—and the degree of scale and translation invariance. Rotation invariance is probably the most challenging property for wavelets, especially if one insists on having basis functions (non-redundant representation).
- *Feature detection*. Image differentials are valuable clues for feature detection. Most classical edge-detection algorithms are gradient- or Laplacian-based, and, sometimes, a combination of both. A wavelet transform that behaves like a multiscale version of these fundamental operators could therefore be of great use for image analysis.
- *Steerability* (a.k.a. rotation covariance). This property is satisfied when all rotated versions of a basis function remain in the span of the wavelet at that particular scale [23]. It is one of the keys for efficient directional analysis and, more importantly, for designing image-processing algorithms that are truly rotation-invariant.
- *Localization*. A wavelet can always be expressed as a multiple-order derivative of some smoothing function. The smoothing function should be isotropic, for good steerability, and achieve a good trade-off between space and frequency localization. Clearly, the Gaussian would be an ideal choice, but it is incompatible with the wavelets being orthogonal between different scales. Accordingly, one should consider biorthogonal or even mildly-redundant designs to alleviate this limitation.
- *Simplicity*. Multidimensional wavelet transforms typically

lead to multi-channel data structures because the information at a particular scale is encoded using several basis functions. From a practical point of view, it would be advantageous if there was a unique analysis wavelet so that the coefficients at a given scale could be stored into an image array, similar to what is done in the quincunx case. This would reinforce the (joint) interpretation of the wavelet coefficients and facilitate to the development of in-band processing algorithms.

Design Method for Complex Wavelet Basis

Unlike most traditional wavelet designs, our starting point is not filterbanks, but rather the selection of an appropriate differential operator. We shall see that imposing three fundamental properties (translation and scale invariance, rotation covariance) narrows down the choice to a family of complex operators which involve some fractional iterate of the Laplacian as well as the complex-gradient operator $(j\partial/\partial x_1 + \partial/\partial x_2)$. Clearly, this operator gives access to derivatives in any direction by suitable multiplication with a complex number of unit modulus.

We then adopt a spline-like construction where the approximation space is spanned by the integer-shifts of the Green function of the operator. Since these Green's functions are not in L_2, we need to search for a localized basis function. For a pure Laplacian operator (associated to the rotation-invariance property), we can select the polyharmonic B-splines that are described in [26]–[28]. Incorporating the complex gradient leads to complex polyharmonic B-splines, which have first been explored in [29].

The next step is to specify wavelet functions that span the orthogonal complement between approximation spaces at two subsequent scales. Remarkably, the inter-scale orthogonality constraint automatically enforces the operator-like behavior of the wavelet, a mechanism that is well understood for the 1-D case [30] and for (real-valued) polyharmonic B-spline wavelets [28]. While the present construction heavily relies on the theoretical results of Forster *et al.* [29], it goes an important step further by providing a solution to the wavelet-design problem for arbitrary subsampling lattices. Remarkably, our new operator-like wavelet is independent upon the choice of the dilation matrix and it always generates a semi-orthogonal basis of $L_2(\mathbb{R}^2)$ through simple shifts and dilations.

If the input signal is real, then the fact of expanding it on complex-valued wavelets results into a redundancy factor of 2.

Marr-Like Wavelet Pyramid and Primal Sketch

The conceptual motivation for the present construction was to provide a mathematical interpretation of Marr's theory for the early stage of vision [31]. Based on similarities with the primate's visual system, Marr proposed to analyze the zero-crossings of the image filtered with a Laplacian-of-Gaussian (LoG). He then determined the orientation of the zero-crossing segments and used these to define a "raw primal sketch", which, he argued, could serve as input to higher-level visual processing. We can transpose those ideas to the present setup by combining both operations (LoG and orientation of zero-crossings) into a composite complex gradient-Laplace operator, for which our wavelets provide a multiscale version. The corresponding wavelet-domain local maxima give the intensity (by their magnitude) and the orientation (by their phase) of an edge at a particular scale. We coin the term "Marr-like wavelet pyramid" for this decomposition, and we introduce the "wavelet primal sketch" as a compact multiscale version of the raw primal sketch (see Fig. 15 (a)). Specifically, we propose a scheme to obtain a contour-based description of images and complement it with an algorithm that reconstructs a high-quality approximation of the original image.

Organization of the Paper

We present our design method in Sect. I, which starts with the identification of a class of rotation-covariant differential operator; this yields a complex wavelet basis of $L_2(\mathbb{R}^2)$. The main technical contributions there are the specification of the operator-like wavelets and the proof that they generate Riesz bases for arbitrary subsampling configurations. Next, we introduce the Marr-like wavelet pyramid and the wavelet primal sketch representation in Sect. II. In Sect. III, we illustrate essential features of the Marr-like wavelet pyramid such as the angular selectivity and the importance of phase and magnitude. We also propose an algorithm to efficiently reconstruct an image from its wavelet primal sketch. Finally, in Sect. IV, we conclude the paper with a discussion and comparison of the Marr-like wavelet pyramid against other state-of-the-art multiscale decompositions.

I. Operator-Like Wavelets

A. Operator Design Principle

Our design starts with the specification of a suitable differential operator L that satisfies some desirable invariance properties. This operator admits a Green function $\rho(\mathbf{x}) = \rho(x_1, x_2)$ such that $\mathrm{L}\{\rho\} = \delta$. The integer shifts of ρ can be used to define a sequence of embedded approximation spaces

$$\mathcal{V}_i = \mathrm{span}\left\{\rho(2^i\mathbf{x} - \mathbf{k})\right\}_{\mathbf{k}\in\mathbb{Z}^2} \cap L_2(\mathbb{R}^2). \tag{1}$$

For most differentiation operators, the Green's functions ρ are not in L_2, which makes the representation (1) not practical for implementation purposes. Yet, it is quite useful conceptually because it focuses on the essentials (see [32] for a discussion of the 1-D case which brings out an interesting connection with fractals). We defer to subsection I-C the specification of a valid scaling function of $\mathcal{V}_0$.

Let us now consider the function $\psi = \mathrm{L}^\star\{\phi\}$, where $\mathrm{L}^\star$ is the adjoint operator of L and ϕ is a smoothing function. The orthogonality condition between ψ at position $\mathbf{x}_0$ and the function space $\mathcal{V}_0$ can then be expressed as

$$\begin{aligned}\langle\psi(\cdot - \mathbf{x}_0), \rho(\cdot - \mathbf{k})\rangle &= \langle \mathrm{L}^\star\{\phi\}(\cdot - \mathbf{x}_0), \rho(\cdot - \mathbf{k})\rangle \\ &= \langle \phi, \mathrm{L}\{\rho(\cdot - \mathbf{k} + \mathbf{x}_0)\}\rangle \\ &= \langle \phi, \delta(\cdot - \mathbf{k} + \mathbf{x}_0)\rangle \\ &= \phi(\mathbf{k} - \mathbf{x}_0) = 0, \end{aligned} \tag{2}$$

and can be enforced by a judicious choice of ϕ and $\mathbf{x}_0$. Using this fundamental property together with the fact that ψ is in the finer approximation space $\mathcal{V}_1$, we want to specify the wavelet subspace

$$\mathcal{W}_0 = \text{span}\{\psi(\mathbf{x} - \mathbf{k}/2)\}_{\mathbf{k}\in\mathbb{Z}^2\setminus 2\mathbb{Z}^2}, \tag{3}$$

which is spanned by a *single* wavelet function at all positions $(\mathbf{k}/2)\setminus\mathbb{Z}^2$. We are going to show that this is possible; in other words, that ϕ can be selected such that ψ is a Riesz-basis generator of $\mathcal{W}_0$, while maintaining the orthogonality condition with respect to $\mathcal{V}_0$. The wavelet spaces can then be embedded as

$$\mathcal{W}_i = \text{span}\{2^i\psi(2^i\mathbf{x} - \mathbf{k}/2)\}_{\mathbf{k}\in\mathbb{Z}^2\setminus 2\mathbb{Z}^2}, \tag{4}$$

where the wavelet function $\psi = \mathrm{L}^*\{\phi\}$ behaves as a multiscale version of L.

This strategy was used implicitly in one of the earliest constructions of 1-D polynomial spline wavelets [33]. In that case, the operator L was chosen to be d^N/dx^N, for which the causal Green's function is the one-sided monomial $(x)_+^{N-1}/(N-1)!$. The smoothing function ϕ corresponded to the spline interpolant of order $2N$ dilated by a factor $1/2$. It can be readily verified that (2) holds in that case for $x_0 = 1/2$.

The above derivation provides us with an important insight on how the approximation space and wavelet space are tightly linked together through the operator and its Green's function. In the sequel, we clarify the different ingredients—operator, scaling function, and smoothing function—that lead to the Marr-like wavelet pyramid, the wavelet primal sketch, and its corresponding reconstruction algorithm.

B. A Suitable Operator for Image Processing

In classical (separable) wavelet theory, the underlying operator L is necessarily an N-th order derivative, where N corresponds to the number of vanishing moments. Here, we are interested in selecting a true 2-D operator that has invariance properties that are better suited for image processing. Specifically, we have identified the following three desirable properties:

Definition 1 (Translation Invariance): The operator L is said to be translation-invariant when it commutes with the shift operator: $\mathrm{L}\{s(\cdot - \mathbf{x}_0)\}(\mathbf{x}) = \mathrm{L}\{s(\cdot)\}(\mathbf{x} - \mathbf{x}_0)$.

Definition 2 (Scale Invariance): The operator L is said to be scale-invariant when it commutes (up to a constant) with the dilation operator; that is, when $\mathrm{L}\{s(T\cdot)\}(\mathbf{x}) = c_T\mathrm{L}\{s(\cdot)\}(T\mathbf{x})$ where c_T is an appropriate real-valued constant.

Definition 3 (Rotation Covariance): The operator L is said to be rotation-covariant when it commutes with the rotation operator up to a rotation-encoding constant with unit magnitude; specifically, when $\mathrm{L}\{s(\mathbf{R}_\theta\cdot)\}(\mathbf{x}) = c_\theta\mathrm{L}\{s(\cdot)\}(\mathbf{R}_\theta\mathbf{x})$ where $c_\theta \in \mathbb{C}$ with $|c_\theta| = 1$.

Since we assume L to be a convolution operator, which requires linearity and translation invariance, we can characterize it by its Fourier transform $\hat{L}(\boldsymbol{\omega}) = \hat{L}(\omega_1, \omega_2)$ in the distributional sense. We have that $\mathrm{L}\{s\}(\mathbf{x}) \leftrightarrow \hat{L}(\boldsymbol{\omega})\hat{s}(\boldsymbol{\omega})$, where $\hat{s}(\boldsymbol{\omega}) = \int_{\mathbb{R}^2} s(\mathbf{x})e^{-j\boldsymbol{\omega}^T\mathbf{x}}\mathrm{d}\mathbf{x}$ is the Fourier transform of $s(\mathbf{x})$. Here, we want to further restrict ourselves to the class of scale-invariant and rotation-covariant operators. This imposes strong constraints on the form of $\hat{L}(\boldsymbol{\omega})$ as specified by Lemma 1, and, ultimately, leads to the identification of the class of complex operators in Theorem 1, which are a combination of the Laplace operator and complex gradient or Wirtinger-type operators [34]. The proofs are given in Appendices A and B, respectively.

Lemma 1: A scale-invariant and rotation-covariant convolution operator L is necessarily γth-order scale invariant and μth-order rotation-covariant. Its frequency response is such that $\hat{L}(T\mathbf{R}_\theta\boldsymbol{\omega}) = T^\gamma e^{j\mu\theta}\hat{L}(\boldsymbol{\omega})$ for any $\theta \in \mathbb{R}$ and $T > 0$, where $\gamma \in \mathbb{R}$ and $\mu \in \mathbb{Z}$.

Theorem 1 (Wirtinger-Laplace operator): The convolution operator L is γth-order scale invariant and Nth-order rotation covariant if and only if its Fourier transform can be written (up to some complex multiplicative factor) as

$$\hat{L}(\boldsymbol{\omega}) = \hat{L}_{\gamma,N}(\boldsymbol{\omega}) = ||\boldsymbol{\omega}||^{\gamma-N}(\omega_1 - j\omega_2)^N, \tag{5}$$

where N is an arbitrary positive integer and $\gamma \geq N$ is real-valued.

In the spatial domain, the Wirtinger-Laplace operator from (5) corresponds to

$$\mathrm{L}_{\gamma,N} = (-\Delta)^{\frac{\gamma-N}{2}}\left(-j\frac{\partial}{\partial x_1} - \frac{\partial}{\partial x_2}\right)^N, \tag{6}$$

for which the Green's function ρ is known to be

$$\rho_{\gamma,N}(x_1, x_2) = \begin{cases} c_1\,||\mathbf{x}||^{\gamma-N-2}(x_1 + ix_2)^N, \\ \quad \text{for } \gamma \notin 2\mathbb{N}, \\ c_2\,||\mathbf{x}||^{\gamma-N-2}(x_1 + ix_2)^N(\ln\pi\,||\mathbf{x}|| + c_3), \\ \quad \text{for } \gamma \in 2\mathbb{N}, \end{cases}$$

where c_1, c_2, c_3 are some appropriate constants [29]. The novel contribution here is to show that this class of operators is complete with respect to the desired invariance properties. Notice that, for $\gamma = 3, N = 1$, we obtain the complex gradient-Laplace operator.

C. Complex Polyharmonic B-Splines

Definition 1 is not useful computationally because the Green's function ρ is not in L_2. Therefore, we need to find a localized basis function $\varphi(\mathbf{x}) = \sum_{\mathbf{k}} v[\mathbf{k}]\rho(\mathbf{x} - \mathbf{k})$ that is a Riesz-basis generator of $\mathcal{V}_0$, the space spanned by the integer shifts of ρ. Our solution is somewhat different from the one initially proposed in [29].

Definition 4 (Complex Polyharmonic B-Spline): The polyharmonic B-spline associated with the complex gradient-Laplace operator $\mathrm{L}_{\gamma,N}$, where the order γ is real-valued with $\gamma > 1$ and $N \in \mathbb{N}$, is defined via its Fourier transform

$$\hat{\beta}_{\gamma,N}(\boldsymbol{\omega}) = \frac{V_{\gamma,N}(e^{j\boldsymbol{\omega}})}{||\boldsymbol{\omega}||^{\gamma-N}(\omega_1 - j\omega_2)^N} = \frac{V_{\gamma,N}(e^{j\boldsymbol{\omega}})}{\hat{L}(\boldsymbol{\omega})}. \tag{7}$$

The numerator is the localization filter with polar representation $V_{\gamma,N}(e^{j\boldsymbol{\omega}}) = V_\gamma(e^{j\boldsymbol{\omega}}) \cdot e^{j\theta(\boldsymbol{\omega})}$, where the modulus and

phase are given by

$$V_\gamma(e^{j\boldsymbol{\omega}}) = \left(\frac{8}{3}\left(\sin^2\left(\frac{\omega_1}{2}\right)+\sin^2\left(\frac{\omega_2}{2}\right)\right) +\frac{2}{3}\left(\sin^2\left(\frac{\omega_1+\omega_2}{2}\right)+\sin^2\left(\frac{\omega_1-\omega_2}{2}\right)\right)\right)^{\gamma/2},$$

$$\theta(\boldsymbol{\omega}) = \angle(j[\omega_1]-j[\omega_2])^N,$$

and where $[\omega]$ stands for the unique ω in $[-\pi,\pi[$ such that $\omega-[\omega]=2n\pi$ for some integer n.

The amplitude term $V_\gamma(e^{j\boldsymbol{\omega}})$ makes the Fourier transform well-behaved at $\boldsymbol{\omega}=0$, where it "cancels" the zero of the denominator exactly at the order γ. The pure phase term $e^{j\theta(\boldsymbol{\omega})}$, on the other hand, is a 2π-periodic all-pass filter that is specified to provide a zero-phase behavior in the whole Nyquist band $\boldsymbol{\omega}\in[-\pi,\pi[^2$.

The space-domain counterpart of (7) is the so-called Green's function localization formula, which yields the equivalent B-spline definition

$$\beta_{\gamma,N}(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^2} v_{\gamma,N}[\mathbf{k}]\rho_{\gamma,N}(\mathbf{x}-\mathbf{k}). \tag{8}$$

This construction is such that $\beta_{\gamma,0}$ coincides with the isotropic polyharmonic B-splines that were earlier introduced by us and that are associated with the Laplacian operator alone [28]. Moreover, our complex extension has the desirable feature that $\left|\hat{\beta}_{\gamma,N}\right| = \hat{\beta}_{\gamma,0}$, so that our new B-splines share the same autocorrelation filter as the polyharmonic ones, with

$$A_\gamma(e^{j\boldsymbol{\omega}}) = \sum_{\mathbf{k}\in\mathbb{Z}^2}\left|\hat{\beta}_\gamma(\boldsymbol{\omega}+2\pi\mathbf{k})\right|^2. \tag{9}$$

Another important function, also independent of N, is the unique interpolant within the polyharmonic-spline space of augmented order spanned by $\rho(\cdot)*\rho(-\cdot)$ or equivalently by $\beta_{2\gamma,0}$. It is such that $\phi_{2\gamma}(\mathbf{k})=\delta_{\mathbf{k}}$ and its Fourier transform is given by

$$\hat{\phi}_{2\gamma}(\boldsymbol{\omega}) = \frac{\left|\hat{\beta}_\gamma(\boldsymbol{\omega})\right|^2}{A_\gamma(e^{j\boldsymbol{\omega}})}. \tag{10}$$

Based on our earlier results [29], we can deduce that the complex polyharmonic B-splines satisfy all the requirements for generating a valid multiscale analysis of $L_2(\mathbb{R}^2)$ (see [1]):

- Their integer shifts form a Riesz basis of the signal space $\mathcal{V}_0=\operatorname{span}_{\mathbf{k}\in\mathbb{Z}^2}\{\beta_{\gamma,N}(\mathbf{x}-\mathbf{k})\}$;
- They satisfy the partition-of-unity property $\sum_{\mathbf{k}}\beta_{\gamma,N}(\mathbf{x}-\mathbf{k})=1$;
- They satisfy a scaling relation for any admissible subsampling matrix $\mathbf{D}$ that consists of a rotation combined with a dilation and $|\det\mathbf{D}|\geq 2$. The scaling relation is given in the frequency domain by

$$\hat{\beta}_{\gamma,N}(\mathbf{D}^T\boldsymbol{\omega}) = \frac{1}{|\det\mathbf{D}|}H(e^{j\boldsymbol{\omega}})\hat{\beta}_{\gamma,N}(\boldsymbol{\omega}), \tag{11}$$

  where $H(e^{j\boldsymbol{\omega}})$ is the scaling filter.

These complex polyharmonic B-splines can therefore be used to construct wavelets that span the orthogonal complements $\mathcal{W}_i=\mathcal{V}_i\ominus_\perp\mathcal{V}_{i-1}$ between the sequence of embedded spaces $\mathcal{V}_i=\operatorname{span}\left\{\beta_{\gamma,N}(\mathbf{D}^i\mathbf{x}-\mathbf{n})\right\}_{\mathbf{n}\in\mathbb{Z}^2}$.

D. Operator-Like Wavelets

Real-valued polyharmonic wavelets (i.e., $N=0$) have been proposed both for the quincunx subsampling matrix [28], which corresponds to a two-channel design in 2-D, and for the dyadic subsampling matrix [35], [36]. So far, complex polyharmonic spline wavelets have only been specified explicitly for the simpler quincunx case [29], which corresponds to a classical two-channel design with a single wavelet generator.

In this work, we propose a more general construction that yields operator-like wavelets for any admissible subsampling matrix $\mathbf{D}$. Another remarkable feature of our approach is that the wavelet spaces are generated using a single wavelet instead of $|\det(\mathbf{D})|-1$ distinct ones as in classical designs (e.g., separable constructions). It is also important to mention that our wavelets are uniquely defined in the sense that they do not depend on the specific choice of the generator of the approximation space, but only on the approximation space itself. We define $M=|\det(\mathbf{D})|-1$ as the canonical number of wavelets.

Definition 5 (Operator-Like Wavelets): The operator-like wavelets that span the detail space $\mathcal{W}_{-1}$ are defined as

$$\psi^{(n)}_{\gamma,N}(\mathbf{D}^{-1}\mathbf{x}) = (-\Delta)^{\frac{\gamma-N}{2}}\left(j\frac{\partial}{\partial x_1}-\frac{\partial}{\partial x_2}\right)^N\{\phi_{2\gamma}\}(\mathbf{x}+\mathbf{e}_n), \quad n=1,\ldots,M, \tag{12}$$

where $\hat{\phi}_{2\gamma}$ is specified by (10) and $\mathbf{e}_n$ is in the cosets $\mathcal{C}\setminus\mathbf{e}_0$ of the subsampling matrix $\mathbf{D}$, with $\mathbf{e}_0=(0,0)$.

Our wavelet space is generated by a single function $\psi^{(0)}_{\gamma,N}=\psi_{\gamma,N}$ that is spatially shifted on all coset positions $\mathbf{e}_n$. Therefore, we compactly denote the wavelet space $\mathcal{W}_i$ as

$$\mathcal{W}_i = \operatorname{span}\{|\det(\mathbf{D})|^{i/2}\ \psi_{\gamma,N}(\mathbf{D}^i\mathbf{x}-\mathbf{D}^{-1}\mathbf{k})\}_{\mathbf{k}\in\mathbb{Z}^2\setminus\mathbf{D}\mathbb{Z}^2}, \tag{13}$$

where the factor $|\det(\mathbf{D})|^{i/2}$ ensures a proper L_2-normalization.

Theorem 2 (Riesz-Basis Property): For $\gamma>1$, the operator-like wavelet $\psi_{\gamma,N}$ generates a semi-orthogonal basis of $L_2(\mathbb{R}^2)$ for any admissible subsampling matrix $\mathbf{D}$. Their construction is such that

1) The space $\mathcal{W}_i$ is contained in the finer approximation space $\mathcal{V}_{i+1}$;
2) The space $\mathcal{W}_i$ is orthogonal to the approximation space $\mathcal{V}_i$ at the same scale;
3) The basis $\psi_{\gamma,N}$ generates a Riesz basis of $\mathcal{W}_0$, and, by rescaling, of all the spaces $\mathcal{W}_i$.

Proof: To establish the first property, we rewrite the Fourier counterpart of (12) as

$$\hat{\psi}^{(n)}_{\gamma,N}(\mathbf{D}^{\mathrm{T}}\boldsymbol{\omega}) = \|\boldsymbol{\omega}\|^{\frac{\gamma-N}{2}}(\omega_1+j\omega_2)^N\, e^{j\boldsymbol{\omega}^T\mathbf{e}_n}\,\frac{\hat{\beta}^\star_{\gamma,N}(\boldsymbol{\omega})\hat{\beta}_{\gamma,N}(\boldsymbol{\omega})}{A_\gamma(e^{j\boldsymbol{\omega}})} = \underbrace{\frac{V^\star_{\gamma,N}(e^{j\boldsymbol{\omega}})}{A_\gamma(e^{j\boldsymbol{\omega}})}e^{j\boldsymbol{\omega}^T\mathbf{e}_n}}_{W^{(n)}(e^{j\boldsymbol{\omega}})}\hat{\beta}_{\gamma,N}(\boldsymbol{\omega}), \tag{14}$$

which shows that $\psi^{(n)}_{\gamma,N}(\mathbf{D}^{-1}\mathbf{x})\in\mathcal{W}_{-1}$ can be expressed as a linear combination of integer shifts of $\beta_{\gamma,N}(\mathbf{x}-\mathbf{k})$. This also

identifies the wavelet filter $W^{(n)}(e^{j\boldsymbol{\omega}})$ that needs to be used with the scaling function $\beta_{\gamma,N}$.

To prove the second property, we rewrite the inner product between a wavelet and the scaling function at scale $i=0$ as

$$\begin{aligned}&\left\langle \psi^{(n)}_{\gamma,N}(\cdot), \beta_{\gamma,N}(\cdot+\mathbf{n})\right\rangle\\ &= \left\langle \mathrm{L}^{\star}_{\gamma,N}\left\{\phi_{2\gamma}\right\}(\mathbf{D}\cdot+\mathbf{e}_n-\mathbf{D}\mathbf{n}), \beta_{\gamma,N}(\cdot)\right\rangle\\ &= \left\langle \phi_{2\gamma}(\mathbf{D}\cdot+\mathbf{e}_n-\mathbf{D}\mathbf{n})), \mathrm{L}_{\gamma,N}\left\{\beta_{\gamma,N}\right\}(\cdot)\right\rangle. \quad (15)\end{aligned}$$

Applying the operator $\mathrm{L}_{\gamma,N}$ to $\beta_{\gamma,N}$ cancels the denominator of (7) so that the only remaining term is the filter $V_{\gamma,N}(e^{j\boldsymbol{\omega}})$. The corresponding space-domain expression is

$$\mathrm{L}_{\gamma,N}\left\{\beta_{\gamma,N}\right\}(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^2} v_{\gamma,N}[\mathbf{k}]\delta(\mathbf{x}-\mathbf{k}), \quad (16)$$

where $v_{\gamma,N}$ is the inverse Fourier transform of $V_{\gamma,N}(e^{j\boldsymbol{\omega}})$. Therefore, the inner product further simplifies to

$$\sum_{\mathbf{k}\in\mathbb{Z}^2} v_{\gamma,N}[\mathbf{k}]\phi_{2\gamma}(\mathbf{D}\mathbf{k}+\mathbf{e}_n-\mathbf{D}\mathbf{n}) = 0, \quad (17)$$

which is zero since the interpolating function $\phi_{2\gamma}$ is sampled at non-zero integer positions only.

For the third property, we first introduce the elements of the Gram matrix as

$$r^{(n_1,n_2)}[\mathbf{k}] = \left\langle \psi^{(n_1)}_{\gamma,N}(\cdot), \psi^{(n_2)}_{\gamma,N}(\cdot+\mathbf{k})\right\rangle, \quad (18)$$

which has the discrete Fourier transform $R^{(n_1,n_2)}(e^{j\boldsymbol{\omega}})$. We need to prove that the determinant of the corresponding $M\times M$ Fourier matrix $\mathbf{R}(e^{j\boldsymbol{\omega}}) = [R^{(n_1,n_2)}(e^{j\boldsymbol{\omega}})]_{n_1,n_2}$ is bounded by two positive constants $0 < A_1 \le A_2 < \infty$. The entries of this matrix are expressed as (using $\boldsymbol{\pi} = (\pi,\pi)$)

$$\begin{aligned}&R^{(n_1,n_2)}(e^{j\boldsymbol{\omega}})\\ &= \sum_{\mathbf{k}\in\mathbb{Z}^2} \hat{\psi}^{(n_1)\star}_{\gamma,N}(\boldsymbol{\omega}+2\pi\mathbf{k})\hat{\psi}^{(n_2)}_{\gamma,N}(\boldsymbol{\omega}+2\pi\mathbf{k})\\ &= \sum_{n=0}^{M} e^{j(\boldsymbol{\omega}^T\mathbf{D}^{-1}+\pi\mathbf{e}_n^T)(\mathbf{e}_{n_2}-\mathbf{e}_{n_1})} \frac{\left|V_{\gamma,N}(e^{j(\mathbf{D}^{-T}\boldsymbol{\omega}+\mathbf{e}_n^T\boldsymbol{\pi})})\right|^2}{A_\gamma(e^{j(\mathbf{D}^{-T}\boldsymbol{\omega}+\mathbf{e}_n^T\boldsymbol{\pi})})}\\ &= \sum_{n=0}^{M} e^{j(\boldsymbol{\omega}^T\mathbf{D}^{-1}+\pi\mathbf{e}_n^T)(\mathbf{e}_{n_2}-\mathbf{e}_{n_1})}\\ &\qquad \left(\sum_{\mathbf{k}} \frac{1}{||\mathbf{D}^{-T}\boldsymbol{\omega}+\mathbf{e}_n^T\boldsymbol{\pi}+2\pi\mathbf{k}||^{2\gamma}}\right)^{-1}\\ &= \sum_{n=0}^{M} e^{j(\boldsymbol{\omega}^T\mathbf{D}^{-1}+\pi\mathbf{e}_n^T)(\mathbf{e}_{n_2}-\mathbf{e}_{n_1})}\\ &\qquad \frac{||\mathbf{D}^{-T}\boldsymbol{\omega}+\mathbf{e}_n^T\boldsymbol{\pi}||^{2\gamma}}{1+\sum_{\mathbf{k}\neq\mathbf{0}} \frac{||\mathbf{D}^{-T}\boldsymbol{\omega}+\mathbf{e}_n^T\boldsymbol{\pi}||^{2\gamma}}{||\mathbf{D}^{-T}\boldsymbol{\omega}+\mathbf{e}_n^T\boldsymbol{\pi}+2\pi\mathbf{k}||^{2\gamma}}}. \quad (19)\end{aligned}$$

As $R^{(n_1,n_2)}(e^{j\boldsymbol{\omega}})$ is 2π-period, we can restrict our analysis to $[-\pi,\pi[^2$. It is easy to see that (19) is upper-bounded because the denominator and the numerator are trivially lower- and upper-bounded, respectively. Consequently, the determinant of $\mathbf{R}(e^{j\boldsymbol{\omega}})$ is upper-bounded too. For the lower bound, we can use the fact that $\det(\mathbf{R}(e^{j\boldsymbol{\omega}})) > 0$ if the M "fiber" vectors [37]

$$\mathbf{v}^{(n)} = \left(\hat{\psi}^{(n)}_{\gamma,N}(\boldsymbol{\omega}+2\pi\mathbf{k})\right)_{\mathbf{k}\in\mathbb{Z}^2}, \quad n=1,\ldots,M \quad (20)$$

are linearly independent for any fixed $\boldsymbol{\omega}$. This can be verified by checking that

$$\sum_{n=1}^{M} \lambda_n \mathbf{v}^{(n)} = \mathbf{0} \quad (21)$$

requires $\lambda_n = 0$. Component-wise, we can rewrite (21) as

$$\sum_{n=1}^{M} \lambda_n [\mathbf{v}^{(n)}]_{\mathbf{k}} = \hat{\psi}^{(0)}(\boldsymbol{\omega}+2\pi\mathbf{k}) \sum_{n=1}^{M} \lambda_n e^{j(\boldsymbol{\omega}+2\pi\mathbf{k})^T\mathbf{D}^{-1}\mathbf{e}_n} = 0,$$

which really reduces to $M+1$ conditions to be fulfilled. We get the same equation for all $\mathbf{k} = \mathbf{D}^T\mathbf{k}'$, $\mathbf{k}' \in \mathbb{Z}^2$, and $M+1$ different ones can be identified for $\mathbf{k} = \mathbf{e}_l + \mathbf{D}^T\mathbf{k}'$, $l = 0,\ldots,M$, so that

$$\begin{aligned}&\hat{\psi}^{(0)}(\boldsymbol{\omega}+2\pi(\mathbf{e}_l+\mathbf{D}^T\mathbf{k})) \sum_{n=1}^{M} \lambda_n e^{j(\boldsymbol{\omega}+2\pi\mathbf{e}_l+2\pi\mathbf{D}^T\mathbf{k})^T\mathbf{D}^{-1}\mathbf{e}_n}\\ &= \underbrace{\hat{\psi}^{(0)}(\boldsymbol{\omega}+2\pi(\mathbf{e}_l+\mathbf{D}^T\mathbf{k}))}_{c_l} \sum_{n=1}^{M} \lambda_n \underbrace{e^{j\boldsymbol{\omega}^T\mathbf{D}^{-1}\mathbf{e}_n}}_{c'_n} e^{j2\pi\mathbf{e}_l^T\mathbf{D}^{-1}\mathbf{e}_n}\\ &= 0, \quad (22)\end{aligned}$$

where $|c'_n| = 1$ and $|c_l| > 0$, except for $\boldsymbol{\omega} = \mathbf{0}, l = 0$. It is convenient to rewrite the system of linear equations as

$$\underbrace{\begin{bmatrix} c_0 & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & c_M\end{bmatrix}}_{(M+1)\times(M+1)} \underbrace{\begin{bmatrix} 1 & \cdots & 1\\ e^{j2\pi\mathbf{e}_1^T\mathbf{D}^{-1}\mathbf{e}_1} & \cdots & e^{j2\pi\mathbf{e}_1^T\mathbf{D}^{-1}\mathbf{e}_M}\\ \vdots & \ddots & \vdots\\ e^{j2\pi\mathbf{e}_M^T\mathbf{D}^{-1}\mathbf{e}_1} & \cdots & e^{j2\pi\mathbf{e}_M^T\mathbf{D}^{-1}\mathbf{e}_M}\end{bmatrix}}_{\mathbf{E}:(M+1)\times M}$$

$$\underbrace{\begin{bmatrix} c'_1 & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & c'_M\end{bmatrix}}_{M\times M} \begin{bmatrix}\lambda_1\\ \vdots\\ \lambda_M\end{bmatrix} = \mathbf{0}.$$

We identify the aliasing offsets with respect to $\mathbf{D}$ as $\mathbf{w}_l = 2\pi\mathbf{D}^{-T}\mathbf{e}_l$. In the case $\boldsymbol{\omega} = \mathbf{0}$, we have that $c_0 = 0$, which removes the first equation, leaving us with the system

$$\underbrace{\begin{bmatrix} c_1 & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & c_M\end{bmatrix}}_{M\times M} \underbrace{\begin{bmatrix} e^{j\mathbf{w}_1^T\mathbf{e}_1} & \cdots & e^{j\mathbf{w}_1^T\mathbf{e}_M}\\ \vdots & \ddots & \vdots\\ e^{j\mathbf{w}_M^T\mathbf{e}_1} & \cdots & e^{j\mathbf{w}_M^T\mathbf{e}_M}\end{bmatrix}}_{\mathbf{E}:M\times M}$$

$$\underbrace{\begin{bmatrix} c'_1 & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & c'_M\end{bmatrix}}_{M\times M} \begin{bmatrix}\lambda_1\\ \vdots\\ \lambda_M\end{bmatrix} = \mathbf{0}.$$

The null solution is the only one if $\mathbf{E}$ is full-rank, which is true in general as the matrix $\mathbf{E}$ is a submatrix of the $(M+1)\times(M+1)$ generalized discrete Fourier-transform matrix with respect to $\mathbf{D}$ [38]. Specifically, in the quincunx case, we have that

$$\mathbf{D} = \begin{bmatrix} 1 & 1\\ -1 & 1\end{bmatrix}, \quad (23)$$

leading to $\mathbf{E} = [-1]$, which is clearly full rank. In the dyadic case, we have that

$$\mathbf{D} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \tag{24}$$

leading to

$$\mathbf{E} = \begin{bmatrix} -1 & 1 & -1 \\ 1 & -1 & -1 \\ -1 & -1 & 1 \end{bmatrix} \tag{25}$$

which has full rank again as well.

For the case $\boldsymbol{\omega} \neq \mathbf{0}$, we need to add the first equation again ((22) for $l = 0$), which makes the system overdetermined. Due to the previous result, $\mathbf{E}$ has still rank M and thus the only solution is still $\lambda_n = 0$, $n = 1, \ldots, M$. ■

We use the term semi-orthogonality to indicate that the wavelet is orthogonal to its dilations, but not to its translations. The fact that we have a Riesz basis does also imply the existence of an equivalent orthogonal basis but it is unlikely to have a single generator, except, of course, for the quincunx case where $M = 1$. It is also important to note that the operator-like wavelet generator at scale $i = 0$ is essentially independent upon $\mathbf{D}$ (up to a geometrical transformation) and that it is not tied to any particular choice of scaling function. In fact, it is completely determined by the operator alone, which is a clear indication of its fundamental nature. The same can also be said for the higher-order interpolant $\phi_{2\gamma}$ appearing in the formula which is uniquely tied to the function space, and hence, to the operator.

Let us now review a few more properties.

Property 1 (Relationship Between Real and Imaginary Part): The operator-like wavelets have the following $\pi/2$-rotation property between their real and imaginary parts:

$$\mathrm{Re}\{\psi_{\gamma,N}(x_1, x_2)\} = -\mathrm{Im}\{\psi_{\gamma,N}(x_2, x_1)\}. \tag{26}$$

Property 2 (Multiscale Operator Signal Analysis): Analyzing a function f with operator-like wavelets results into complex coefficients for which we have that

$$\left\langle f, \psi_{\gamma,N}(\mathbf{D}^i \cdot - \mathbf{k}) \right\rangle = (-\Delta)^{\frac{\gamma - N}{2}} \left(-j\frac{\partial}{\partial x_1} - \frac{\partial}{\partial x_2} \right)^N \left\{ f(\cdot) * \phi_{2\gamma}(\mathbf{D}^i \cdot) \right\} (\cdot - \mathbf{k}),$$

where $\phi_{2\gamma}$ is a max-flat lowpass function such that

$$\hat{\phi}_{2\gamma}(\boldsymbol{\omega}) = 1 + O(\|\boldsymbol{\omega}\|^{2\gamma}). \tag{27}$$

Moreover, as $\gamma \to \infty$, we have that $\hat{\phi}_{2\gamma}(\boldsymbol{\omega}) \to \mathrm{rect}(\boldsymbol{\omega})$ (see [28]).

For the choice $\gamma = 3, N = 1$, we obtain the behavior of the "pure" complex gradient-Laplace operator as

$$\begin{bmatrix} \mathrm{Re}\left\langle f, \psi_{3,1}(\mathbf{D}^i \cdot - \mathbf{k}) \right\rangle \\ \mathrm{Im}\left\langle f, \psi_{3,1}(\mathbf{D}^i \cdot - \mathbf{k}) \right\rangle \end{bmatrix} = \begin{bmatrix} \partial/\partial x_2 \\ \partial/\partial x_1 \end{bmatrix} \Delta \left\{ f(\cdot) * \phi_{2\gamma}(\mathbf{D}^i \cdot) \right\} (\cdot - \mathbf{k}).$$

In Fig. 1, we show the real and imaginary parts of the operator wavelet in (a) and (b), respectively. In (c), we display the linear combination $\cos(\pi/4)\mathrm{Re}\{\psi_{\gamma,N}\} + \sin(\pi/4)\mathrm{Im}\{\psi_{\gamma,N}\}$, which is the wavelet steered at $\pi/4$. However, the steerability is not perfect since the smoothing function $\phi_{2\gamma}$ is interpolating and thus not isotropic. We show in Section III how this can be fixed.

Property 3 (Analysis-Synthesis): The operator-like wavelet signal analysis of a function $f \in L_2(\mathbb{R}^2)$ can be inverted using the dual synthesis wavelets; specifically,

$$f(\mathbf{x}) = \sum_{i \in \mathbb{Z}} \sum_{\mathbf{k} \in \mathbb{Z}^2} \sum_{n=1}^{M} d_i^{(n)}[\mathbf{k}] \, |\det(\mathbf{D})|^{i/2} \; \tilde{\psi}_{\gamma,N}^{(n)}(\mathbf{D}^i \mathbf{x} - \mathbf{k}) \tag{28}$$

with

$$d_i^{(n)}[\mathbf{k}] = \left\langle f(\cdot), |\det(\mathbf{D})|^{i/2} \; \psi_{\gamma,N}^{(n)}(\mathbf{D}^i \cdot - \mathbf{k}) \right\rangle, \tag{29}$$

where the $\tilde{\psi}_{\gamma,N}^{(n)}$, $n = 1, \ldots, M$, are the unique functions in $\mathcal{W}_i$ such that

$$\left\langle \tilde{\psi}_{\gamma,N}^{(n_1)}(\mathbf{D}^{i_1} \cdot - \mathbf{k}_1), \psi_{\gamma,N}^{(n_2)}(\mathbf{D}^{i_2} \cdot - \mathbf{k}_2) \right\rangle = \delta_{n_1 - n_2} \cdot \delta_{i_1 - i_2} \cdot \delta_{\mathbf{k}_1 - \mathbf{k}_2}.$$

E. Implementation

We show in Fig. 2 the filterbank implementation of the complex-wavelet decomposition. The derived filters

$$H(\mathbf{z}) = |\det(\mathbf{D})|^{1-\gamma} \frac{V_{\gamma,N}(\mathbf{z}^{\mathbf{D}})}{V_{\gamma,N}(\mathbf{z})}, W^{(n)}(\mathbf{z}) = \frac{V_{\gamma,N}^*(\mathbf{z})}{A_\gamma(\mathbf{z})} \mathbf{z}^{\mathbf{e}_n} \tag{30}$$

are placed on the analysis side, while their duals are used for the reconstruction; these can be found by imposing the perfect-reconstruction condition and solving a linear system of equations. Since the filters are known explicitly in the Fourier domain, the transform is evaluated efficiently using the FFT (implicit periodic boundary conditions).

In our implementation, we use a prefiltering step to obtain the initial approximation coefficient $c_0[\mathbf{k}]$ at scale 0. We assume that the sampled input signal is bandlimited, and we represent it by its cardinal series

$$s(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{Z}^2} s(\mathbf{k}) \mathrm{sinc}(\mathbf{x} - \mathbf{k}). \tag{31}$$

The orthogonal projection onto $\mathcal{V}_0$ is then obtained as

$$c_0[\mathbf{k}] = \left\langle s(\cdot), \tilde{\beta}_{\gamma,N}(\cdot - \mathbf{k}) \right\rangle, \tag{32}$$

where $\tilde{\beta}_{\gamma,N}$ is the unique dual function in $\mathcal{V}_0$. In the Fourier domain, this translates into

$$C_0(e^{j\boldsymbol{\omega}}) = S(e^{j\boldsymbol{\omega}}) \frac{\hat{\beta}_{\gamma,N}([\boldsymbol{\omega}])}{A_\gamma(e^{j\boldsymbol{\omega}})} \tag{33}$$

where $[\boldsymbol{\omega}]$ is the component-wise extension of $[\omega]$, and where $C_0(e^{j\boldsymbol{\omega}})$ and $S(e^{j\boldsymbol{\omega}})$ are the discrete Fourier transforms of c_0 and s, respectively.

II. The Marr-Like Pyramid

In the previous section, we exhibited semi-orthogonal wavelets that we derived from the complex gradient-Laplace operator. We now want to relate these to a wavelet-pyramid decomposition inspired by Marr's theory of early vision [25], [31].

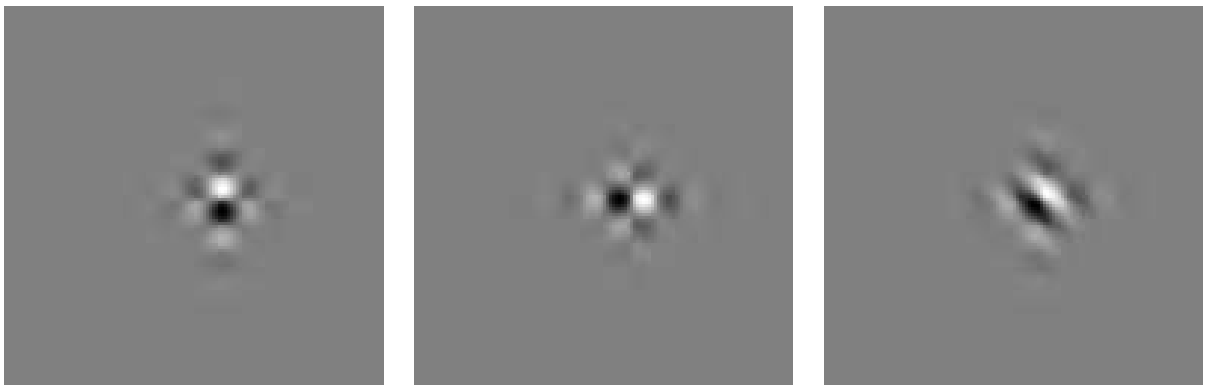


Fig. 1. Semi-orthogonal operator-like wavelets ($\gamma = 3, N = 1$). (a) Real part. (b) Imaginary part. (c) Steered at $\pi/4$.

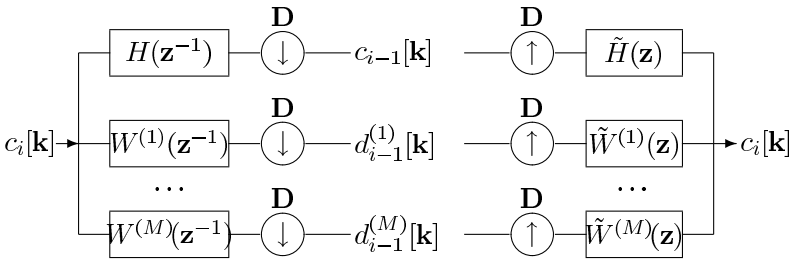


Fig. 2. Filterbank implementation of the complex wavelet decomposition.

A. Marr's Theory of Vision and the Raw Primal Sketch

In his seminal work on vision [25], David Marr proposed an influential model for the primate's visual system. The first (early) stage of the vision process leads to the "raw primal sketch" [31]. It consists of three essential steps. First, the image is smoothed by a Gaussian which is a low-pass filter with optimal joint spatial-spectral localization. Next, the Laplacian operator is applied and edges are detected as zero-crossings. Last, zero-crossing segments are identified and their orientation determined. The characterization of the image by zero-crossing segments is called the "raw primal sketch", which is then processed by higher-level vision mechanisms.

The Laplacian-of-Gaussian (LoG) operation is performed at several scales simultaneously [39]. Mathematically, the theory of linear scale space elegantly deals with the continuous stack of images blurred by a Gaussian kernel with increasing width [40], [41]. Applying the Laplacian operator to the scale-space representation and detecting the zero-crossings leads to the "scale-space primal sketch" [42].

An outstanding theoretical question is whether or not the multiscale primal sketch carries all the information of the original image. Therefore, one is interested in reconstructing an image from its primal sketch representation. To be practical, the representation also needs to be stable, so that a small perturbation of the primal sketch does only slightly modify the image.

Several researchers have investigated the inversion problem from zero-crossings of band-limited signals (or, equivalently, using an ideal band-pass wavelet function) in the 1-D case [43]. These results rely on the band-limitedness of the signal to compute its analytic extension [44]–[46], which is a nonstable characterization. The properties of zero-crossings of functions convolved with a LoG have also been studied, more specifically its multiscale version [47], [48]. The completeness of the representation is guaranteed for polynomial signals, but not its stability [49]. To alleviate this problem, additional information of the function is retained, such as the gradient values at the zero-crossings [50]. Yet, one should note that the scale-space primal sketch has a continuous scale, which makes this approach unpractical. As an alternative, Mallat proposed a stabilized zero-crossing representation using the redundant wavelet transform at a sequence of dyadic scales $(2^i)_{i\in\mathbb{Z}}$ [49]. In addition to the zero-crossing positions, he records the integral between two zero-crossings. The proposed reconstruction algorithm, which is based on iterative projections, is able to recover a close approximation of the original signal. Unfortunately, the proposed representation is not guaranteed to be complete [51], [52]. In [51], Mallat and Zhong propose another representation that stores the positions and the values of the modulus maxima of the wavelet coefficients. In this case, the wavelet is chosen to be a derivative of a B-spline. The scheme extends for images using a separable wavelet transform. In that case, local modulus maxima are extracted by a Canny edge-detector-like procedure [53]. The reconstruction algorithm is similar to the 1-D case and applied row-by-row and column-by-column. The modulus maxima representation uses dyadic scales, but is fully redundant inside each scale. Moreover, two separate (redundant) wavelet transforms are required to obtain horizontal and vertical derivatives.

What we are after here is a decomposition/reconstruction procedure that has the flavor of the one proposed by Mallat and Zhong, but that uses a much lesser degree of redundancy in the construction of the edge map. In other words, we seek a pyramidal representation as opposed to a fully redundant decomposition.

B. The Marr-Like Wavelet Pyramid

The operator-like wavelet basis constitutes our starting point for specifying a wavelet decomposition that mimicks the first stage of vision processing proposed by Marr. The innovation of this paper is (1) a pyramid structure with mild redundancy for better translation invariance; (2) the adaptation of the smoothing function to make it more Gaussian-like.

1) Pyramid Structure and Dyadic Subsampling Scheme: The wavelet spaces $\mathcal{W}_i$ are spanned by appropriate shifts and dilations of the prototype wavelet. They are compactly written as

$$\mathcal{W}_i = \text{span}\{\sqrt{|\det \mathbf{D}^i|}\ \psi_{\gamma,N}(\mathbf{D}^i\mathbf{x}-\mathbf{D}^{-1}\mathbf{k})\}_{\mathbf{k}\in\mathbb{Z}^2\setminus\mathbf{D}\mathbb{Z}^2}. \quad (34)$$

In Fig. 3 (a), we show an equivalent polyphase representation of the filterbank structure corresponding to a wavelet analysis into $\mathcal{W}_i$. It suggests that one can filter once and keep all but one out of $|\det \mathbf{D}|$ coefficients. It is then very tempting to consider a true pyramid structure where the "missing" wavelet would be included as well, as shown in Fig. 3 (b). The corresponding "augmented" wavelet space is

$$\mathcal{W}_i^+ = \text{span}\{\sqrt{|\det \mathbf{D}^i|}\ \psi_{\gamma,N}(\mathbf{D}^i\mathbf{x} - \mathbf{D}^{-1}\mathbf{k})\}_{\mathbf{k}\in\mathbb{Z}^2}. \quad (35)$$

Clearly, we now have that $\mathcal{W}_i \subset \mathcal{W}_i^+$, but we loose the semi-orthogonality property as the additional wavelet is not orthogonal to the scaling function of $\mathcal{V}_i$. Yet, since every wavelet is in the span of the scaling function at the finer resolution, it still holds that $\mathcal{W}_i^+ \subset \mathcal{V}_{i+1}$.

In this slightly redundant scheme, the reconstruction of the pyramid decomposition is performed using the subband regression algorithm, which we illustrate in Fig. 3 (c). Then, the pyramid data is reprojected onto the (non-redundant) wavelet basis such as to minimize a quadratic consistency error criterion. More details and explicit filter expressions can be found in [54].

From now on, we focus on the dyadic subsampling scheme that corresponds to the matrix $\mathbf{D} = 2\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. As illustrated in Fig. 3 (d), the three wavelets and the additional one can be rearranged in a single subband on a Cartesian grid of high (non-subsampled) resolution. The pyramid structure brings along two important advantages: (1) only the approximation subband is subsampled and thus the translation invariance of the wavelets is improved, (2) in-band processing can be done using conventional image-processing algorithms designed for the Cartesian grid.

2) Smoothing Function: An important feature of the Marr framework is smoothing with a Gaussian filter. It is therefore tempting to replace the smoothing function $\phi_{2\gamma}$ in (12) by another one that is closer to a Gaussian. While the possibility of doing so in $\mathcal{W}_i$ is somewhat restricted (i.e., the zeros of the smoothing function must be guaranteed according to (2)), it is much easier to achieve within our enlarged wavelet space $\mathcal{W}_i^+$. Our proposal is to consider appropriate linear combinations of functions within $\mathcal{W}_i^+$, which corresponds to a filter $Q(e^{j\boldsymbol{\omega}/2})$ at scale 0. Here, we pick the Marr smoothing function to be the polyharmonic B-spline

$$\begin{aligned}\hat{\phi}_{2\gamma}^{\text{Marr}}(\boldsymbol{\omega}/2) &= \hat{\phi}_{2\gamma}(\boldsymbol{\omega}/2)A_\gamma(e^{j\boldsymbol{\omega}/2}) \\ &= \left|\hat{\beta}_\gamma(\boldsymbol{\omega}/2)\right|^2 = \hat{\beta}_{2\gamma}(\boldsymbol{\omega}/2). \end{aligned} \quad (36)$$

This smoothing function closely resembles a Gaussian due to the properties of the polyharmonic B-spline. In fact, it quickly converges to a Gaussian with standard deviation $\sqrt{\gamma/12}$ as γ increases [28]. In Fig. 4, we show the smoothing functions of both the operator-like wavelet and the new proposed one for $\gamma = 3$.

(a) (b)

Fig. 4. Smoothing functions of order $\gamma = 3$ for (a) the quasi-steerable wavelets and (b) the Marr-like wavelets. Note that (a) is an interpolator while (b) is an isotropic polyharmonic B-spline of order 6.

3) Pure Complex Gradient-Laplace Behavior: For our purpose, the natural choice of the parameters of the operator $\mathrm{L}_{\gamma,N}$ is $\gamma = 3$ and $N = 1$, which corresponds to the pure complex gradient-Laplacian behavior. The Laplacian part has, combined with the smoothing function, a LoG behavior. By adding the gradient, we can easily map the zero-crossings of the Laplacian onto the local extrema of the complex gradient-Laplace. Additionally, the phase and magnitude of the complex wavelet coefficient contain important information on the edges.

Definition 6 (Marr-Like Wavelet Pyramid): The Marr-like wavelet pyramid is specified by a sequence of analysis wavelets $\psi(2^i\mathbf{x} - \mathbf{k}/2)$ where

$$\psi(\mathbf{x}) = \Delta^\alpha \left(j\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}\right)\{\beta_{2\gamma}(2\cdot)\}(\mathbf{x}), \quad (37)$$

with $\alpha = (\gamma-1)/2$. In the sequel, we select $\alpha = 1$ and $\gamma = 3$.

Note that the Marr-like wavelets span our augmented wavelet spaces

$$\begin{aligned}\mathcal{W}_i^+ &= \text{span}\{\psi(2^i\mathbf{k} - \mathbf{k}/2)\}_{\mathbf{k}\in\mathbb{Z}^2} \\ &= \text{span}\{\psi_{(\gamma-1)/2,1}(2^i\mathbf{x} - \mathbf{k}/2)\}_{\mathbf{k}\in\mathbb{Z}^2},\end{aligned}$$

which follows directly from (36).

In Fig. 5, we show the real and imaginary parts of the Marr-like wavelet in (a) and (b), respectively. The wavelet steered at $\pi/4$ is shown in (c). Note the improved steerability with respect to the complex operator-like wavelets. The total

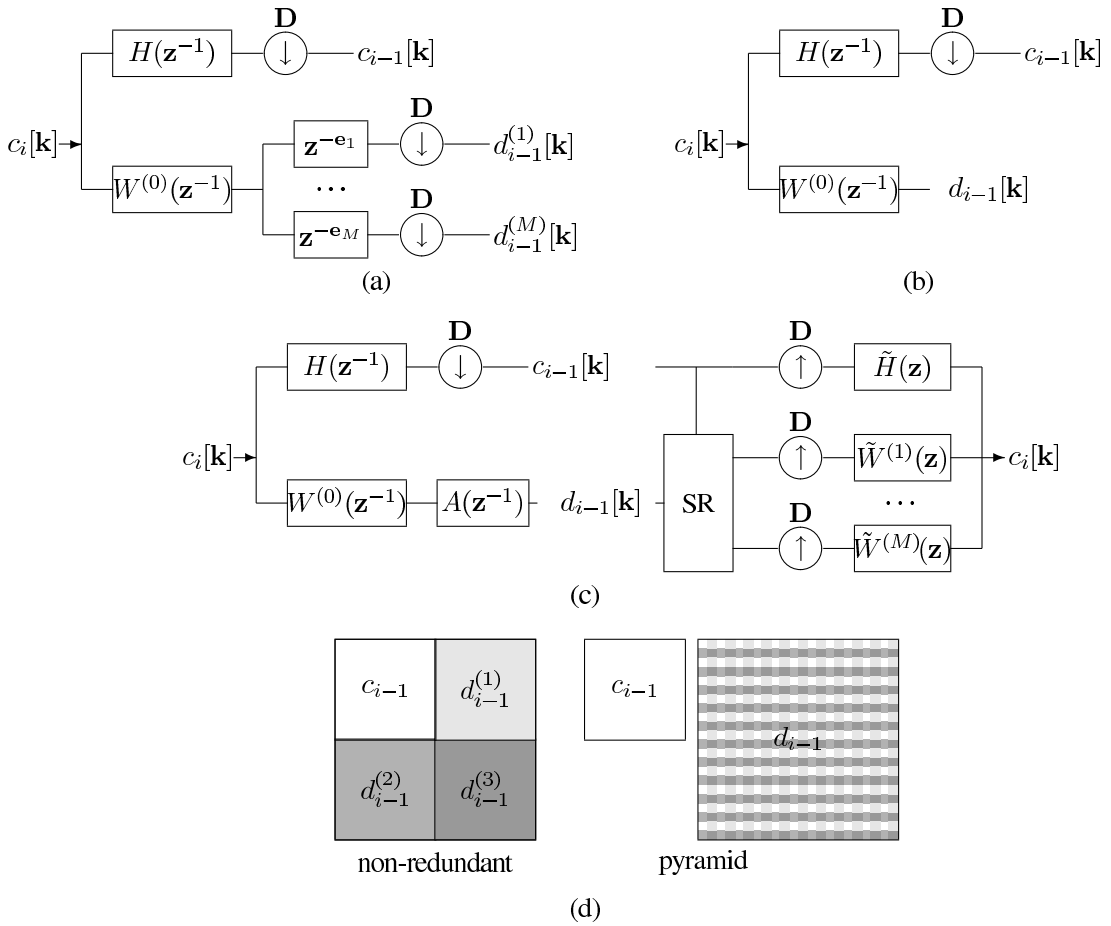


Fig. 3. (a) Filterbank at the analysis side of the non-redundant wavelet decomposition where the same filter $W^{(0)} = V_{\tilde{\gamma},N}^{*}/A_{\gamma}$ can be applied to the data and shifted afterwards. (b) The pyramid extension of (a) can be obtained by adding the non-shifted wavelet subband $W^{(0)}$ and regrouping all coefficients in a single subband. (c) Marr-like wavelet pyramid with subband regression (module "SR") reconstruction. (d) Pyramid organization for the dyadic subsampling scheme.

redundancy of the Marr-like wavelet pyramid is $8/3$, which corresponds to a factor of $4/3$ from the pyramid structure, and a factor of 2 due to the complex nature of the wavelet coefficients.

The coefficients for the Marr-like wavelet pyramid with J decomposition levels at scales $i = -1, \ldots, -J$ are given by

$$d_i[\mathbf{k}] = \left\langle f(\cdot), 2^i \psi(2^i \cdot -\mathbf{k}/2) \right\rangle. \tag{38}$$

These are complemented with the coarser-scale approximation coefficients

$$c_{-J}[\mathbf{k}] = \left\langle f(\cdot), 2^{-J} \beta_{3,1}(2^{-J} \cdot -\mathbf{k}) \right\rangle. \tag{39}$$

Any function $f \in L^2(\mathbb{R}^2)$ is thereby uniquely specified by its Marr-like wavelet pyramid $(d_i[\mathbf{k}], c_{-J}[\mathbf{k}])_{i=-J,\ldots,+\infty;\mathbf{k}\in\mathbb{Z}^2}$, which constitutes a frame represented by the operator W. The reconstruction (operator W^{-1}) is performed according to the algorithm described in Fig. 3(d) which reprojects the data onto the complex wavelet basis.

In Fig. 6, we visualize the Marr-like wavelet pyramid (b) for a synthetic test image "disc" (a). The left side of the pyramid corresponds to the real part, while the right side to the imaginary one. The scale gets coarser from top to bottom. The small thumbnail image at the bottom row represents the approximation coefficients from the last decomposition level. In (c), we show a zoom using a vector plot of the lower-left corner of the subband at scale $i = -2$. Finally, we show the Marr-like wavelet pyramid of the well-known "Einstein" image in Fig. 7.

C. Wavelet Primal Sketch

We now have all the tools at hand to extract the wavelet primal sketch, as illustrated by the flowchart in Fig. 8. First, the image is decomposed into its Marr-like wavelet pyramid. The resulting images are interpreted as multiscale gradients and fed into a Canny-like edge detector (Fig. 9). Specifically, the gradient phase and magnitude are computed and the coefficients that are not local maxima are suppressed. The remaining magnitude values are subjected to a hysteresis threshold within each subband [53]. This yields a multiscale edge map of the image. Note that we retain the positions of the edges, as well as the phase and magnitude of the detected wavelet coefficients.

Our wavelet primal sketch is compact when compared to other state-of-the-art primal sketch representations. Indeed, the complete Marr-like wavelet pyramid has a small redundancy (factor of $8/3$), and there is only a small portion of coefficients

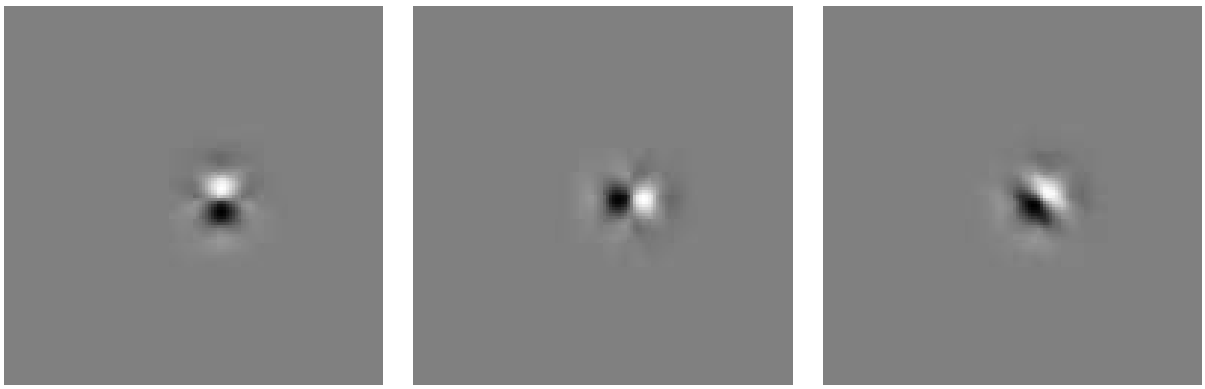


Fig. 5. Marr-like wavelets from Eq. (37). (a) Real part. (b) Imaginary part. (c) Steered at $\pi/4$.

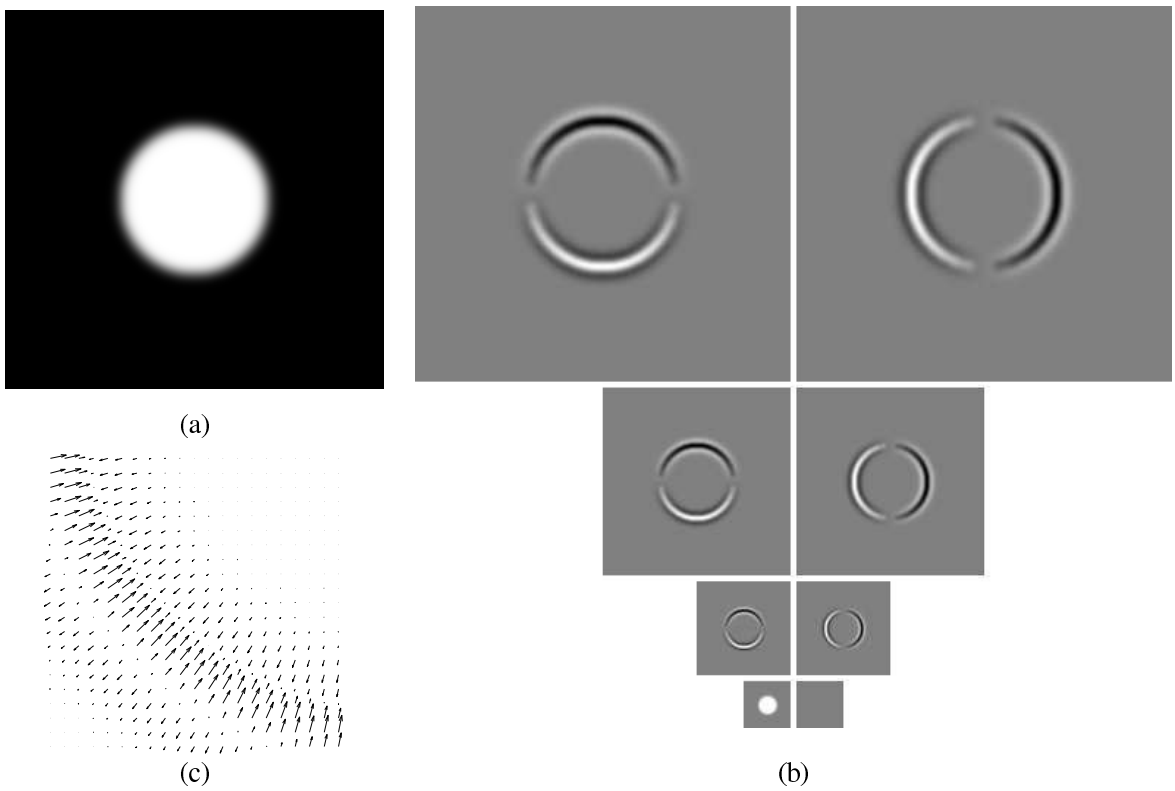


Fig. 6. (a) Disc test image (size 256×256; central disc of radius 50; Gaussian smoothing with $\sigma = 4$ pixels). (b) Marr-like pyramid ($J = 3$ decomposition levels) of the disc image. (c) Vector plot of the lower-left corner of the wavelet subband at scale $i = -2$.

that is retained after hysteresis thresholding. These "primal-sketch" wavelet coefficients carry essential information that can be used to reconstruct the image, as we show in the next section.

Mathematically, we denote $(d_i|_{\Omega_i})$ the wavelet primal sketch of $f \in L_2(\mathbb{R}^2)$ with Marr-like wavelet pyramid coefficients $(d_i[\mathbf{k}])_{i,\mathbf{k}}$, where Ω_i is the index domain at scale i indicating the edge positions. The masks corresponding to the wavelet primal sketches of the images "disc" and "Einstein" are shown in Fig. 14 (b) and Fig. 15 (a), respectively.

III. Experimental Results

We now present a series of experiments to illustrate the properties and possible applications of the complex Marr-like wavelets. In particular, we describe a reconstruction algorithm that is able of re-synthesizing a close approximation of the original image from this reduced primal wavelet sketch.

A. Angular Selectivity

Due to the rotation covariance of the complex gradient-Laplace operator, the wavelet coefficients of the Marr-like pyramid have angular selectivity. The "zoneplate" test image, shown in Fig. 10 (a), forms an ideal test case. The Marr-like

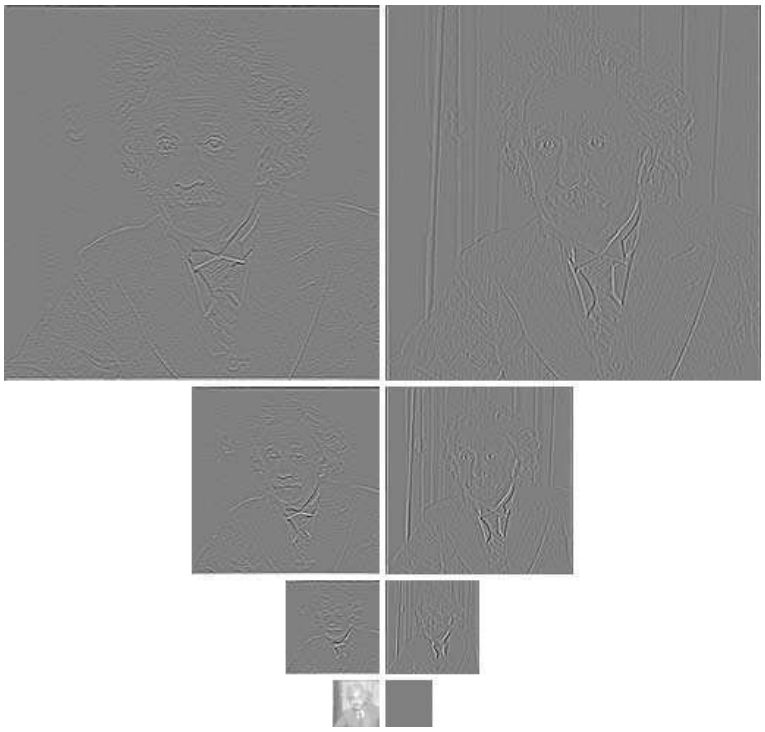


Fig. 7. Marr-like pyramid ($J = 3$ decomposition levels) of the "Einstein" image.

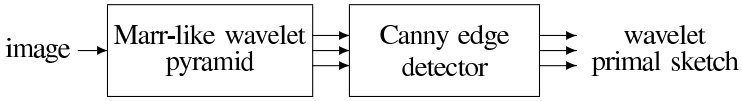


Fig. 8. Flowchart of how to extract the "wavelet primal sketch" from the Marr-like wavelet pyramid. The Canny edge-detection procedure is applied to every subband. Real and imaginary parts of the wavelet coefficients are interpreted as vertical and horizontal derivatives, respectively.

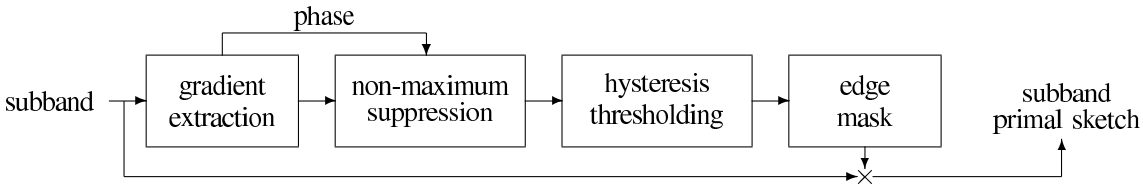


Fig. 9. Flowchart of how the "Canny edge detector" is applied to a subband of the Marr-like pyramid decomposition. First, the gradient is extracted by considering real and imaginary parts of the wavelet coefficients. Then, non-maximum suppression is applied along the gradient direction, followed by hysteresis thresholding. Finally, a complex coefficient is restored for the coefficients that are detected.

wavelet pyramid is shown in Fig. 10 (c). Note the different vertical/horizontal sensitivity between real and imaginary parts, and the varying frequency selectivity according to the scale. In Fig. 10 (b), we show the reconstruction corresponding to the coefficients whose phase is within a wedge of $\pi/4 \pm \pi/8$; the other coefficients are set to zero. Note that we have a complete directional control of the transform thanks to the steerability of the basis functions; in other words, we could as well have extracted an angular wedge in any other direction.

B. Significance of Phase and Modulus

It is a well-known property of the Fourier transform that the main perceptual information of the image is carried by the phase of the Fourier coefficients rather than their magnitude. As a consequence, interchanging phase and magnitude between two images reveals the image from which the phase was selected. An example is shown in Fig. 11 (b), using the "cameraman" (magnitude) and "Einstein" (phase) images. It is informative to perform the equivalent experiment with our complex-valued wavelet coefficients. In Fig. 11 (a), we interchanged the phases and magnitudes of the Marr-like wavelet pyramid of the same images. As with the Fourier transform, the reconstruction mainly reveals the image from which the phase was selected. However, since the wavelet basis functions are more localized, the magnitude retains some "diffuse" spatial content from; i.e., the halo from the "cameraman" can be recognized as well.

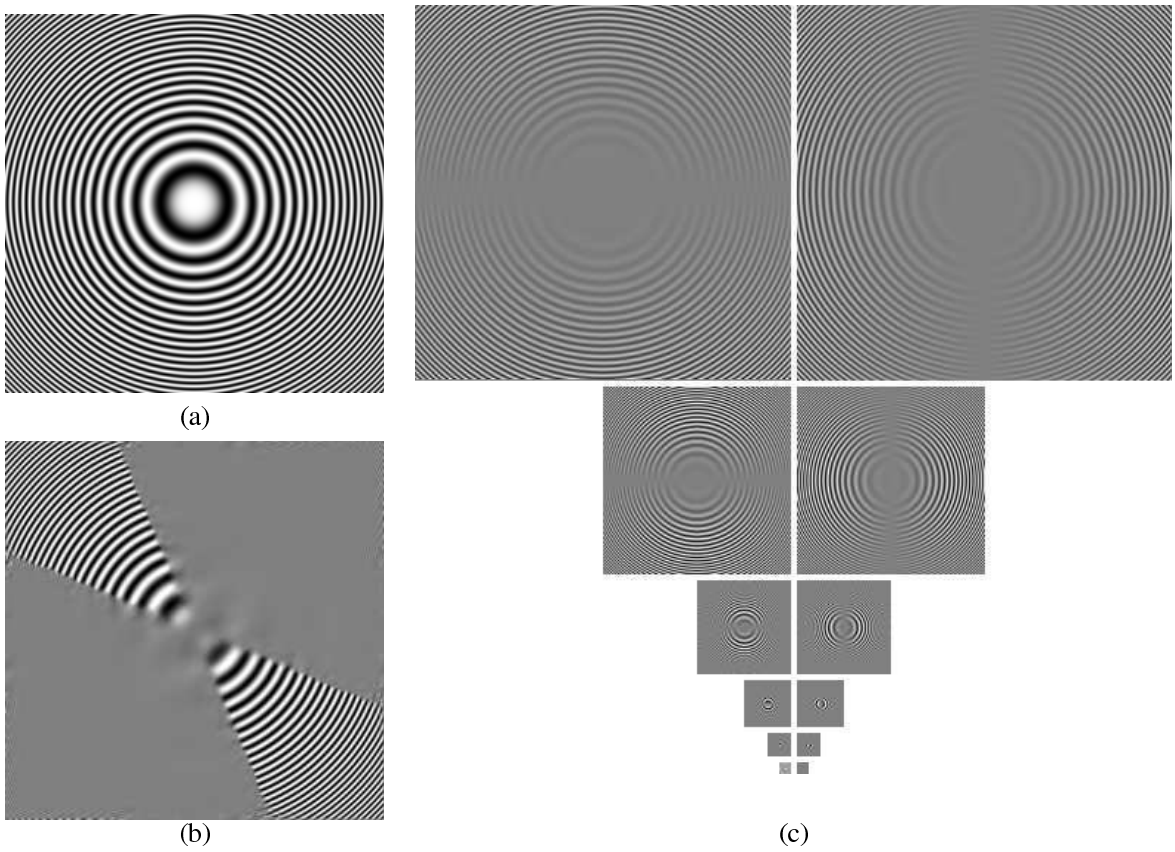


Fig. 10. (a) Test image "zoneplate" (256×256). (b) Marr-lik wavelet pyramid ($J = 5$). (c) Phase selectivity: only coefficients with phase within a wedge $\pi/4 \pm \pi/8$ are retained for the reconstruction. The phase of the coefficients is interpreted modulo π.

Fig. 11. Effect of interchanging phase and magnitude between two images. The phase is recovered from the "Einstein" image, and the magnitude from the "cameraman" image. (a) Marr-like wavelet pyramid, full decomposition ($J = 8$), (b) Discrete Fourier transform.

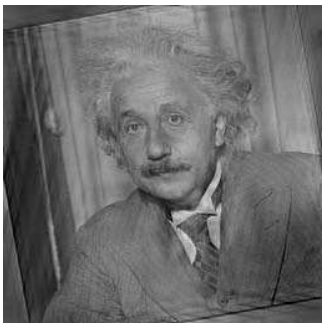


Fig. 12. Reconstruction of the "Einstein" image after rotating each subband in the Marr-like wavelet pyramid by $\pi/18$. The wavelet coefficients are also multiplied by a phase factor $e^{j\pi/18}$.

C. Translation Invariance and Rotation Covariance

Due to the rotation covariance of the operator L, we know that

$$\mathrm{L}\{s(\mathbf{R}_\theta \cdot)\}(\mathbf{x}) = e^{j\theta}\mathrm{L}\{s(\cdot)\}(\mathbf{R}_\theta \mathbf{x}). \tag{40}$$

Since the Marr-like wavelet pyramid is a multiscale version of the operator, it should approximately maintain this property. We demonstrate this feature by geometrically transforming an image by individually rotating each subband. Ideally, we would like to set the rotated wavelet coefficients to

$$d_i'[\mathbf{k}] = e^{j\theta}\big\langle f(\cdot), 2^i \psi(2^i \cdot - \mathbf{R}_\theta \mathbf{k})\big\rangle. \tag{41}$$

In practice, we approximate this step by a bicubic interpolation step of the original subband (Matlab implementation `imrotate`). As shown in Fig. 12 for the "Einstein" image, the reconstruction of the rotated pyramid closely resembles the original (but rotated) image. This would not work at all if we were rotating the coefficients of a conventional separable wavelet transform, because the wavelets cannot be steered.

D. Image Reconstruction from Wavelet Primal Sketch

We now present a practical algorithm for reconstructing an image that is consistent with the primal wavelet sketch. First, we define the space $\mathcal{M}$ of all Marr-like pyramid transforms of real-valued functions in $L^2(\mathbb{R}^2)$. Any such Marr-like pyramid transform is invariant under the projector

$$P_{\mathcal{M}} = W \odot \mathrm{Re} \odot W^{-1}, \tag{42}$$

which consists of three operations: the inverse wavelet transform W^{-1} is an orthogonal projector onto the image with respect to the norm that is minimized by the subband regression algorithm; the pointwise transform Re is an orthogonal projector onto the space of real-valued images; and the wavelet transform W provides the Marr-like wavelet pyramid of the new image.

Second, we are interested in those Marr-like pyramids that have the same values on the domain Ω_i of the wavelet primal sketch. In order to restrict the range of solutions and to impose some unicity, we search for a solution that is consistent with the wavelet primal sketch (i.e., the desired values $d_i[\mathbf{k}]$ on Ω_i) and that minimizes some ad-hoc wavelet-domain energy. Specifically, we would like the coefficients to be small and essentially non-oscillating. Therefore, we propose to minimize the regularizing norm

$$|(d_i[\mathbf{k}])_{i,\mathbf{k}}|_R^2 = \sum_{i\in\mathbb{Z}} \sum_{\mathbf{k}\in\mathbb{Z}^2} |d_i[\mathbf{k}]|^2 + \lambda_i \left(|d_i[\mathbf{k}] - d_i[\mathbf{k}-\mathbf{e}_1]|^2 + |d_i[\mathbf{k}] - d_i[\mathbf{k}-\mathbf{e}_2]|^2 \right), \tag{43}$$

where $\mathbf{e}_1$ and $\mathbf{e}_2$ are the unit horizontal and vertical vectors, respectively. Let us define $\mathcal{R}$ as the space of pyramids with a consistent primal sketch; i.e., $(d_i)_{i,\mathbf{k}}$ such that $(d_i|_{\Omega_i})$ match. A critical component of our algorithm is the projector $P_{\mathcal{R}}$ onto $\mathcal{R}$, which is orthogonal with respect to the norm (43), and which is made explicit in Appendix C.

We are now able to precisely formalize the objective of the reconstruction algorithm: We are searching for an image corresponding to an admissible Marr-like wavelet pyramid such that

$$(d'_i[\mathbf{k}])_{i,\mathbf{k}} \in \mathcal{M}, \text{ subject to } \min_{(d'_i|_{\Omega_i})=(d_i|_{\Omega_i})} |(d'_i[\mathbf{k}])_{i,\mathbf{k}}|_R^2, \tag{44}$$

where $(d_i|_{\Omega_i})$ is the given wavelet primal sketch. This objective can be reached by alternating between the projectors $P_{\mathcal{M}}$ and $P_{\mathcal{R}}$, a procedure that is proven to converge to the orthogonal projection on $\Lambda = \mathcal{M} \cap \mathcal{R}$ [55]. If we start from the zero element of $\mathcal{M}$, we are eventually going to converge to the element of Λ whose norm (43) is minimum. We give the flowchart of the algorithm in Fig. 13. Our procedure has similarities with the approach of Mallat and Zhong [51] who consider a more redundant representation. However, in our case the method is adapted for the Marr-like wavelet pyramid, which is a compact representation and has an efficient $P_{\mathcal{M}}$ projection (simple inverse transform). In other words, we directly take into account the discrete nature of the subbands. In practice, the algorithm is applied for a limited number of decomposition levels (J), while retaining the unaltered coarsest-approximation subband.

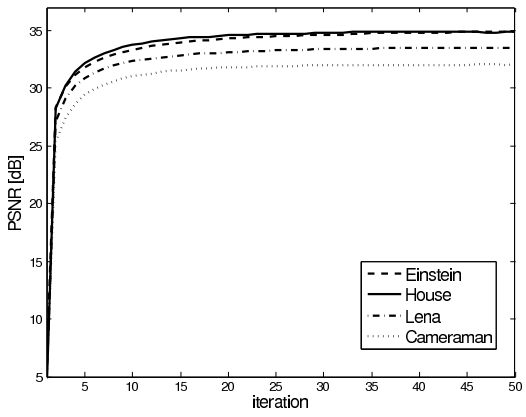


Fig. 16. PSNR as a function of the iteration index for the recovery of various test images from their wavelet primal sketch ($J = 3$ decomposition levels). The maximum number of coefficients in each subband of the wavelet primal sketch is limited at 15%.

The projector $P_{\mathcal{R}}$ can be further refined to suppress spurious edges. Let us denote Ω'_i as the domain at scale i of the wavelet primal sketch of the current estimate. We extend the boundary conditions by imposing on $\Omega'_i \backslash \Omega_i$ that the magnitude should be reduced at the low threshold value of the Canny edge detector. Note that this associated space with the refined projector changes at each iteration, meaning that convergence is no longer guaranteed. Yet, we have observed experimentally that this strategy converges; it has the advantage of somewhat reducing ringing artifacts.

In Fig. 14, we show the results of the reconstruction algorithm for the disc image. The mask of the wavelet primal sketch for $J = 4$ is shown in (b), while a partial vector representation of the subband at $i = -2$ is given in (c). Since the image is noiseless, all local maxima were retained by the Canny edge detector (about 2% of the coefficient at each scale). The reconstruction reaches more than 40dB PSNR after 5 iterations, and 50dB at full convergence (50 iterations).

In Fig. 15, we show the results for the test image "Einstein". The high threshold value of Canny's hysteresis threshold is chosen in an empirical way as a factor $3(1 + 2^i)$ of median-absolute-deviation (MAD) at each scale i. The low threshold value is then adapted such that the number of wavelet coefficients for the primal sketch at scale i remains limited at 15%. The evolution through the iterations is shown in (b). In Fig. 16, we plot the PSNR performances for a couple of popular test images. Convergence at 32–35dB can be observed after about 10 iterations only.

IV. Discussion & Conclusion

We summarize in Table I the essential features of the complex wavelet basis, the Marr-like wavelet pyramid, and comparable state-of-the-art transforms. The steerable pyramid [20] is an orientation-sensitive extension of the well-known Laplacian pyramid [22]. The filters are obtained by a constrained numerical optimization and can be implemented either in the spatial or the Fourier domain (using the FFT). The complex dual-tree wavelet transform [16] relies on the 1-D Hilbert transform to separate negative from positive frequency

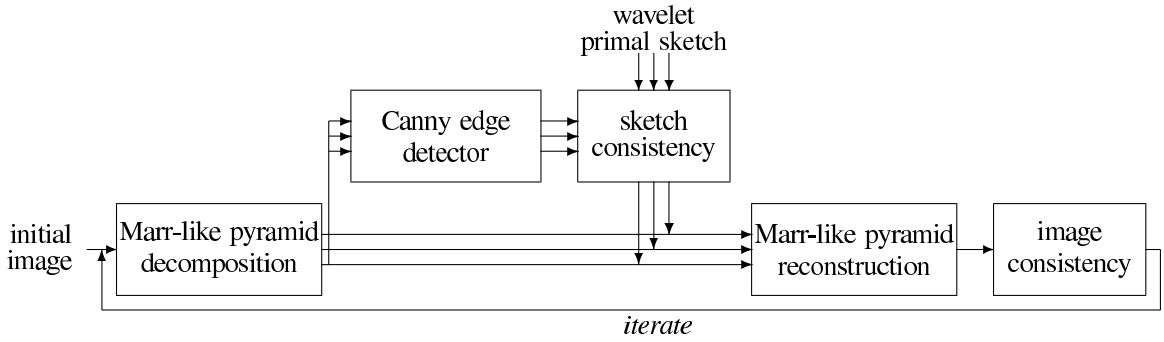


Fig. 13. Flowchart of the algorithm to recover the image from the "wavelet primal sketch". The consistency of the sketch of the intermediate image is verified: Wavelet coefficients of the known sketch are imposed with respect to a regularizing norm. The consistency in the image domain removes the imaginary part of the reconstructed image.

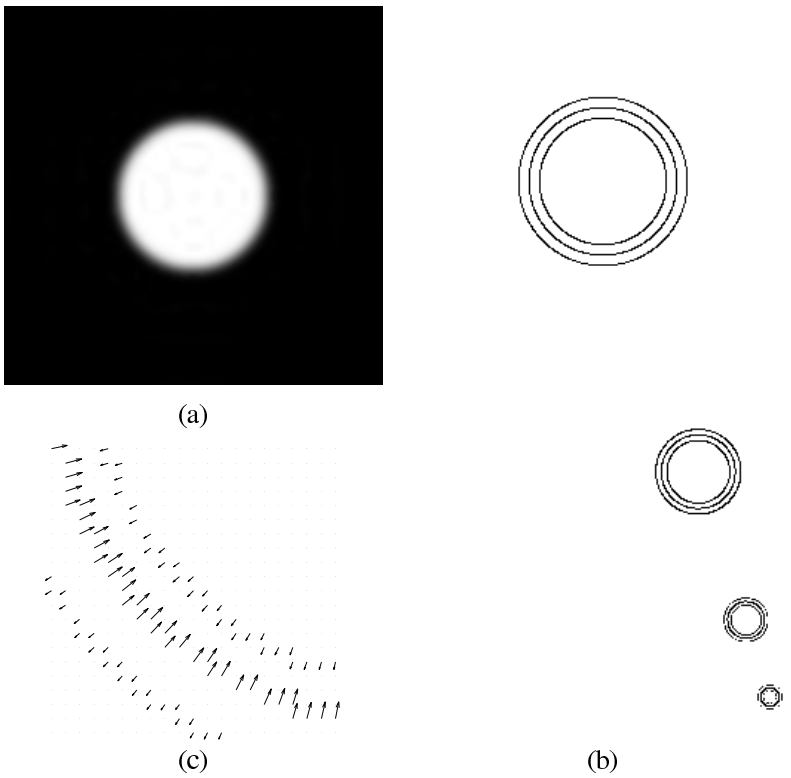


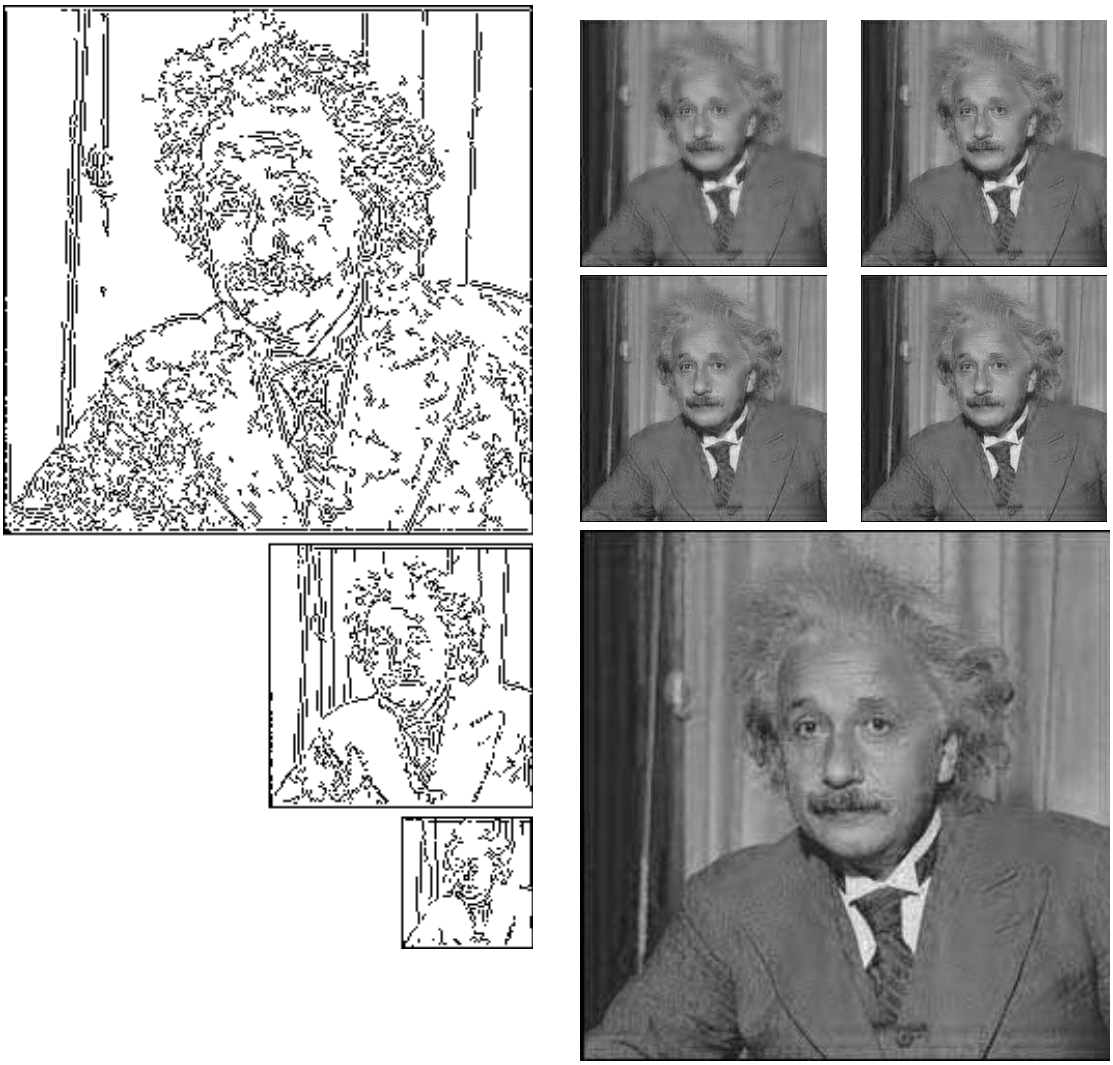
Fig. 14. (a) Disc test image recovered from the wavelet primal sketch after 50 iterations, PSNR=50dB. (b) Mask indicating the wavelet coefficients selected for the wavelet primal sketch ($J = 4$). (c) Vector plot of the lower-left corner of the wavelet subband at scale $i = -2$.

TABLE I

OVERVIEW OF THE ESSENTIAL FEATURES OF THE NEW WAVELET DECOMPOSITIONS AND SOME POPULAR COUNTERPARTS.

	separable orthogonal	Laplacian pyramid	steerable pyramid	complex dual-tree	operator-like complex basis	Marr-like wavelet pyramid
translation invariance	$-$	$+$	$+$	$+$	$-$	$+$
steerability	$-$	$-$	$++^{(1)}$	$+^{(2)}$	$+^{(2)}$	$+++^{(3)}$
number of orientations	$-$	$-$	$2K$	6	2	2
vanishing moments	yes	2	no	yes, 1-D	$\lceil\gamma\rceil$	$\lceil\gamma\rceil$
implementation	filterbank/FFT	filterbank	filterbank/FFT	filterbank	FFT	FFT
construction type	basis	frame	tight frame	frame	complex basis	complex frame
redundancy factor	1	$4/3$	$4K/3+1$	4	2	$8/3$
localization	$+^{(4)}$	$++^{(5)}$	$+^{(4)}$	$+^{(4)}$	$+^{(4)}$	$++^{(5)}$
analytical formulas	yes (splines only)	no	no	no	yes	yes

$^{(1)}$sinc-like smoothing function; $^{(2)}$non-isotropic smoothing function; $^{(3)}$Gaussian-like smoothing function; $^{(4)}$essentially localized; $^{(5)}$Gaussian-like localization



(a) (b)

Fig. 15. The "Einstein" image reconstruction from its wavelet primal sketch. (a) The edge map as detected by the Canny procedure (limited at 15% for each scale). (b) Iterative reconstruction. First 4 iterations left to right, top to bottom. The larger image is obtained after 50 iterations, corresponding to PSNR=33.7dB.

components. This aim is pursued by designing two 1-D wavelet transforms with a half-sample shift difference between them. Making suitable combinations in 2-D yields 6 distinct orientations.

The steerable pyramid is most similar to our construction when the number of orientations is selected as $K = 1$. In that case, the redundancy is $8/3 + 1 = 11/3$, which is slightly higher than in our case due to the special treatment at the first decomposition level. The constraints imposed on the filterbank design lead to a tight "sinc-like" frame. A remarkable feature is that the construction generalizes to a larger number of channels with improved angular selectivity, which can be very useful for applications in computer vision [20], [21]. We should note, however, that the price to pay for higher-order steerability is a comparable increase in redundancy, which suggests a fundamental incompatibility with the specification of basis functions with that type of property.

The dual-tree and Marr-like wavelet transforms are both complex, but they differ in the kind of information that is encoded in the phase. In the former, the phase of a wavelet coefficient gives an indication of relative displacement within the direction of analysis (similar to the correlation with a complex sinusoid), while in the latter it corresponds to the orientation of the underlying pattern. This is consistent with the interpretation of the modulus as the maximal correlation along the dimension encoded by the phase.

While all methods have their strength and niche of applications, the complex wavelet basis and Marr-like wavelet pyramid bring together some interesting analytical properties. First, these wavelets are, by design, multiscale versions of the gradient-Laplace operator which plays such an essential role in image processing (e.g. contour detection). The corresponding wavelet transforms are non-separable and they share all the good theoretical properties of earlier spline transforms:

- They have an order of approximation γ, which means that the approximation error of a scale-truncated expansion

decays like the γth power of that scale.
- They have $\lceil\gamma\rceil$ moments, which implies that the wavelet coefficients are essentially zero in smooth image areas where the image is well represented by its lower-order Taylor series.
- The basis functions have a maximal degree of smoothness for the given order (γ derivatives in the L_2-sense).
- The wavelets have explicit analytical formulas in both space and frequency domains.

The last two properties are direct consequences of the functional construction approach (specification of the Green's function of the underlying operator); they are specific to splines. The final crucial ingredient, which is much less standard with regard to classical wavelet theory, is that the decomposition works for any 2-D subsampling matrix with a straightforward characterization of the corresponding wavelet basis of $L_2(\mathbb{R}^2)$. More precisely, the subsampling matrix determines a single (universal) wavelet generator that is appropriately dilated and shifted to the coset positions. As far as we know, there is no counterpart for any of these continuous-domain properties in the case of the other directional wavelet decompositions because the corresponding formulations are strictly discrete.

An important feature of the present construction is that the complex wavelet basis is associated with a frame (consisting of Marr-like wavelets) with a redundancy factor that is competitive in comparison to the other decompositions. Thanks to the true pyramid structure, one gains in terms of translation-invariance and ease of in-band processing. In particular, this opens the door to the application of standard image-processing operations in the wavelet domain (such as Canny's edge detector). Our design also takes advantage of the redundancy to improve upon the steerability of the wavelets so that the resulting smoothing kernel closely resembles a Gaussian. These modifications considerably improve the feature-extraction capabilities of the transform, as illustrated by our examples. The Marr-like wavelet analysis is coupled with an effective reconstruction algorithm which consists of re-projecting the pyramid data onto the wavelet basis. In this way, we end up with a hybrid basis-frame approach that combines the best of both worlds. It results in improved properties for signal analysis (good translation invariance, rotation covariance, and steerability) and computationally-efficient reconstruction.

We coined the term "Marr-like wavelet pyramid" for the new decomposition, due to its resemblance with David Marr's theory [25] for early-stage processing in the primate's visual system [56]. In particular, the complex gradient-Laplace operator can be seen as a combination of Laplacian-like processing in the ganglion cells of the retina and directional selectivity in the primary visual cortex. Multiscale is also available, such as indicated by the presence of ganglion cells with small (P-type) and large (M-type) receptive fields. The fact that there are many more P- than M-type cells also suggests a compact organization to represent in coarser way the information coming from large basis functions. The localization of the wavelet by the Gaussian(-like) smoothing function is also similar to Marr's framework.

Finally, we have proposed a way to obtain a compact primal wavelet sketch in the form of a multiscale edge map, which is much less redundant than Mallat's wavelet-modulus-maxima representation [51]. This compact representation opens up interesting perspectives, not only for image analysis, but also for image processing because of the existence of a reverse synthesis procedure. Specifically, we have proposed an iterative scheme that reconstructs a high-quality approximation of the original image from its primal wavelet sketch. The algorithm could probably be improved even further by introducing more sophisticated regularization constraints.

APPENDIX A
PROOF OF LEMMA 1

Let us consider a scale-invariant and rotation-covariant convolution operator L. Note that the effect of the dilation and rotation operators can be combined into a single multiplication with a complex-valued function $f(T,\theta) = c_T c_\theta$, where c_T is real-valued and c_θ a pure phase term. As $\hat{L}(\boldsymbol{\omega})$ is a distribution, it acts as a linear functional on the test functions φ_k in Schwartz' class $\mathcal{S}$ and satisfies the continuity condition $\left\langle \hat{L}, \varphi_k \right\rangle \to \left\langle \hat{L}, \varphi \right\rangle$ when $\varphi_k \to \varphi$ as $k \to \infty$. This implies continuity at $(T,\theta) = (1, 2\pi n)$ of the function $f(T,\theta)$, for any $n \in \mathbb{Z}$:

- By making a change of variables $\boldsymbol{\omega} \to T\mathbf{R}_\theta\boldsymbol{\omega}$, we have that
$$\left\langle \hat{L}(T\mathbf{R}_\theta\boldsymbol{\omega}), \varphi(\boldsymbol{\omega}) \right\rangle = \left\langle \hat{L}(\boldsymbol{\omega}), T^{-1}\varphi(T^{-1}\mathbf{R}_\theta\boldsymbol{\omega}) \right\rangle. \quad (45)$$
Using scale invariance and rotation covariance, this proves that
$$f(T,\theta)\left\langle \hat{L}(\boldsymbol{\omega}), \varphi(\boldsymbol{\omega}) \right\rangle = \left\langle \hat{L}(\boldsymbol{\omega}), T^{-1}\varphi(T^{-1}\mathbf{R}_\theta\boldsymbol{\omega}) \right\rangle. \quad (46)$$
- The limit of $T^{-1}\varphi(T^{-1}\mathbf{R}_\theta\boldsymbol{\omega})$ as $(T,\theta) \to (1, 2\pi n)$ is obviously $\varphi(\boldsymbol{\omega})$. Therefore, the right-hand side of (46) tends to $\left\langle \hat{L}(\boldsymbol{\omega}), \varphi(\boldsymbol{\omega}) \right\rangle$. This proves that the left-hand side is convergent as well when $(T,\theta) \to (1, 2\pi n)$, and, finally, that $f(T,\theta) \to 1$.

Next, we verify that $f(T,\theta)$ has to satisfy the chain rule

$$\begin{aligned} f(T_1T_2, \theta_1+\theta_2)\hat{L}(\boldsymbol{\omega}) &= \hat{L}(T_1T_2\mathbf{R}_{\theta_1+\theta_2}\boldsymbol{\omega}) \\ &= \hat{L}(T_1T_2\mathbf{R}_{\theta_1}\mathbf{R}_{\theta_2}\boldsymbol{\omega}) \\ &= f(T_1,\theta_1)\hat{L}(T_2\mathbf{R}_{\theta_2}\boldsymbol{\omega}) \\ &= f(T_1,\theta_1)f(T_2,\theta_2)\hat{L}(\boldsymbol{\omega}). \end{aligned}$$

Using complex analysis, we show that functions that satisfy the chain rule $f(T_1T_2, \theta_1+\theta_2) = f(T_1,\theta_1)f(T_2,\theta_2)$ and are continuous at $(T,\theta) = (1, 2\pi n)$ with $f(1, 2\pi n) = 1$, are necessarily of the form $f(T,\theta) = T^\gamma e^{j\mu\theta}$, where $\gamma \in \mathbb{R}$ and $\mu \in \mathbb{Z}$. In this way, we identify $c_T = T^\gamma$ and $c_\theta = e^{j\mu\theta}$.

APPENDIX B
PROOF OF PROPOSITION 1

From Lemma 1, the frequency response of L satisfies $\hat{L}(T\mathbf{R}_\theta\boldsymbol{\omega}) = T^\gamma e^{j\mu\theta}\hat{L}(\boldsymbol{\omega})$. We make use of complex analysis and identify $\hat{L}(\omega_1,\omega_2)$ with $\hat{L}(z,\overline{z})$ using complex-valued variable $z = \omega_1 + j\omega_2$. We further introduce the complex constant $a = Te^{-j\theta}$ such that we can rephrase the condition

as $\hat{L}(az, \overline{az}) = a^{\frac{\gamma-\mu}{2}} \overline{a}^{\frac{\gamma+\mu}{2}} \hat{L}(z, \overline{z})$. We can now make use of Gelfand's theory for homogeneous distributions of degree $(\frac{\gamma-\mu}{2}, \frac{\gamma+\mu}{2})$ [57, A-6, p. 501]. We find that $\hat{L}(z, \overline{z})$ is unique and proportional, up to some complex constant, to $z^{\frac{\gamma-\mu}{2}} \overline{z}^{\frac{\gamma+\mu}{2}}$, which can be reinterpreted as

$$||\boldsymbol{\omega}||^{\gamma} e^{j\angle(\omega_1 - j\omega_2)^{\mu}} = ||\boldsymbol{\omega}||^{\gamma-\mu} (\omega_1 - j\omega_2)^{\mu}. \qquad (47)$$

APPENDIX C
PROJECTION OPERATOR $P_{\mathcal{R}}$

We consider the projector $P_{\mathcal{R}}$ that transforms a pyramid $(d'_i[\mathbf{k}])_{i,\mathbf{k}} \in \mathcal{M}$ into the closest sequence $(d''_i[\mathbf{k}])_{i,\mathbf{k}} \in \mathcal{R}$ with respect to the regularizing norm (43). We define the update term $\epsilon_i[\mathbf{k}] = d'_i[\mathbf{k}] - d''_i[\mathbf{k}]$. The coefficients $\epsilon_i[\mathbf{k}]$ have to be chosen such that: (1) the norm $|(\epsilon_i[\mathbf{k}])_{i,\mathbf{k}}|^2_R$ is minimum; (2) the boundary conditions imposed by the wavelet primal sketch of f, denoted as $(d_i|_{\Omega_i})$, are satisfied with $\epsilon_i|_{\Omega_i} = d_i|_{\Omega_i} - d'_i|_{\Omega_i}$, for every scale i.

The minimization of $|(\epsilon_i[\mathbf{k}])_{i,\mathbf{k}}|^2_R$ is solved at every scale i for $\mathbf{k} \in \mathbb{Z}^2 \setminus \Omega_i$. We can easily derive from (43) that the constraint

$$\epsilon_i[\mathbf{k}] - \left(\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} * \epsilon_i \right)[\mathbf{k}] = 0 \qquad (48)$$

should hold. The solution is found by a simple numerical implementation that alternates between applying the update formula

$$\epsilon_i^{\text{next}} = \frac{1}{5} \left(\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} * \epsilon_i^{\text{prev}} \right) \qquad (49)$$

and imposing the boundary conditions.

ACKNOWLEDGEMENTS

The authors would like to thank I. Khalidov for the fruitful discussions, K. Balać for the Java implementation, and Dr. P. Thévenaz for the thorough proofreading.

REFERENCES

[1] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego (CA), 1998.
[2] J. G. Daugman, "Complete discrete 2-D gabor transform by neural networks for image analysis and compression," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 36, no. 7, pp. 1169–1179, July 1988.
[3] D. L. Donoho, "Wedgelets: Nearly minimax estimation of edges," *Annals of Statistics*, vol. 27, no. 3, pp. 859–897, June 1999.
[4] E. J. Candès and D. L. Donoho, "Ridgelets: A key to higher-dimensional intermittency?," *Phil. Trans. R. Soc. Lond. A.*, pp. 2495–2509, 1999.
[5] M. N. Do and M. Vetterli, "The finite ridgelet transform for image representation," *IEEE Transactions on Image Processing*, vol. 12, no. 1, pp. 16–28, Jan. 2003.
[6] E. Candès, "The curvelet transform for image denoising," in *Proceedings of the International Conference on Image Processing*, 7-10 Oct. 2001, vol. 1, p. 7.
[7] J. L. Starck, E. J. Candès, and D. L. Donoho, "The curvelet transform for image denoising," *IEEE Transactions on Image Processing*, vol. 11, no. 6, pp. 670–684, June 2002.
[8] M. N. Do and M. Vetterli, *Beyond Wavelets*, chapter Contourlets, Academic Press, 2003.
[9] E. LePennec and S. Mallat, "Sparse geometric image representation with bandelets," *IEEE Transactions on Image Processing*, vol. 14, no. 4, pp. 423–438, Apr. 2005.
[10] G. Peyré and S. Mallat, "Surface compression with geometric bandelets," *ACM Transactions on Graphics*, vol. 14, no. 3, 2005.
[11] P. Vandergheynst and J.-F. Gobbers, "Directional dyadic wavelet transforms: Design and algorithms," *IEEE Transactions on Image Processing*, vol. 11, no. 4, pp. 363–372, Apr. 2002.
[12] V. Velisavljevic, B. Beferull-Lozano, M. Vetterli, and P. Dragotti, "Directionlets: Anisotropic multidirectional representation with separable filtering," *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 1916–1933, 2006.
[13] N. Kingsbury, "Image processing with complex wavelets," *Philos. Trans. R. Soc. London A., Math. Phys. Sci.*, vol. 357, no. 1760, pp. 2543–2560, Sept. 1999.
[14] N. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Applied and Computational Harmonic Analysis*, vol. 10, no. 3, pp. 234–253, May 2001.
[15] F. C. A. Fernandes, R. L. C. van Spaendonck, and C. S. Burrus, "A new framework for complex wavelet transforms," *IEEE Transactions on Signal Processing*, vol. 51, no. 7, pp. 1825–1837, July 2003.
[16] I. W. Selesnick, R. G. Baraniuk, and N. G. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 123–151, Nov. 2005.
[17] C. Chaux, L. Duval, and J.-C. Pesquet, "Image analysis using a dual-tree m-band wavelet transform," *IEEE Transactions on Image Processing*, vol. 15, no. 8, pp. 2397–2412, Aug. 2006.
[18] R. H. Bamberger and M. J. T. Smith, "A filter bank for the directional decomposition of images: Theory and design," *IEEE Transactions on Signal Processing*, vol. 40, no. 4, pp. 882–893, Apr. 1992.
[19] E. P. Simoncelli, W. T. Freeman, and E. H. Adelson, "Shiftable multiscale transforms," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 587–607, Mar. 1992.
[20] E. Simoncelli and W. Freeman, "The steerable pyramid: a flexible architecture for multi-scale derivative computation," in *International Conference on Image Processing*, 23-26 Oct. 1995, vol. 3, pp. 444–447.
[21] A. A. Bharath and J. Ng, "A steerable complex wavelet construction and its application to image denoising," *IEEE Transactions on Image Processing*, vol. 14, no. 7, pp. 948–959, July 2005.
[22] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Transactions on Communications*, vol. COM-31, no. 4, pp. 532–540, Apr. 1983.
[23] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 9, pp. 891–906, Sept. 1991.
[24] G. Granlund and H. Knutsson, *Signal Processing for Computer Vision*, Kluwer, Norwell, MA, 1994.
[25] D. Marr, *Vision*, Freeman Publishers, San Francisco, 1982.
[26] C. Rabut, "Elementary m-harmonic cardinal B-splines," *Numerical Algorithms*, vol. 2, pp. 39–62, 1992.
[27] C. Rabut, "High level m-harmonic cardinal B-splines," *Numerical Algorithms*, vol. 2, pp. 63–84, 1992.
[28] D. Van De Ville, T. Blu, and M. Unser, "Isotropic polyharmonic B-splines: Scaling functions and wavelets," *IEEE Transactions on Image Processing*, vol. 14, no. 11, pp. 1798–1813, Nov. 2005.
[29] B. Forster, T. Blu, D. Van De Ville, and M. Unser, "Shift-invariant spaces from rotation-covariant functions," *Applied and Computational Harmonic Analysis*, in press.
[30] M. Unser and T. Blu, "Wavelet theory demystified," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 470–483, Feb. 2003.
[31] D. Marr and E. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society of London. Series B, Biological Sciences*, vol. 207, no. 1167, pp. 187–217, Feb. 1980.
[32] T. Blu and M. Unser, "Wavelets, fractals, and radial basis functions," *IEEE Transactions on Signal Processing*, vol. 50, no. 3, pp. 543–553, Mar. 2002.
[33] C. K. Chui and J.-Z. Wang, "A cardinal spline approach to wavelets," *Proceedings of the American Mathematical Society*, vol. 113, no. 3, pp. 785–793, Nov. 1991.
[34] R. Remmert, *Theory of Complex Functions*, Springer, 1991.
[35] C. Micchelli, C. Rabut, and F. I. Utreras, "Using the refinement equation for the construction of pre-wavelets III: Elliptic splines," *Numerical Algorithms 1*, , no. 4, pp. 331–352, 1991.
[36] B. Bacchelli, M. Bozzini, C. Rabut, and M.-L. Varas, "Decomposition and reconstruction of multidimensional signals using polyharmonic pre-wavelets," *Applied and Computational Harmonic Analysis*, vol. 18, no. 3, pp. 282–299, May 2005.
[37] C. Deboor, R. A. Devore, and A. Ron, "The structure of finitely generated shift-invariant spaces in $L_2(R^2)$," *Journal of Functional Analysis*, vol. 119, no. 1, pp. 37–78, jan 1994.

[38] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice-Hall, Englewood Cliffs, NJ, 1993.
[39] S. Hochstein and H. Spitzer, "Zero-crossing detectors in primary visual cortex?," *Biol. Cybern.*, vol. 51, pp. 195–199, 1984.
[40] A. P. Witkin, "Scale-space filtering," in *Proc. International Joint Conference on Artificial Intelligence*, Karlsruhe, Germany, 1983, pp. 1019–1023.
[41] J. J. Koenderink, "The structure of images," *Biol. Cybern.*, vol. 50, pp. 363–370, 1984.
[42] T. Lindeberg and J.-O. Eklundh, "Scale-space primal sketch: Construction and experiments," *Image and Vision Computing*, vol. 10, no. 1, pp. 3–18, 1992.
[43] B. Logan, "Information in the zero-crossings of band pass signals," *Bell Syst. Tech. J.*, vol. 56, pp. 510, 1977.
[44] S. Curtis, S. Shitz, and V. Oppenheim, "Reconstruction of non-periodic two-dimensional signals from zero-crossings," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 35, pp. 890–893, 1987.
[45] Y. Zeevi and D. Rotem, "Image reconstruction from zero-crossings," *IEEE Acoust. Speech Signal Processing*, vol. 34, pp. 1269–1277, 1986.
[46] J. Sanz and T. Huang, "Theorem and experiments on image reconstruction from zero-crossings," Res. Rep. RJ5460, IBM, jan 1987.
[47] R. Hummel and R. Moniot, "Reconstruction from zero-crossings in scale-space," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 37, no. 12, pp. 2111–2130, 1989.
[48] A. Yuille and T. Poggio, "Scaling theorems for zero-crossings," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, no. 1, pp. 15–25, Jan. 1986.
[49] S. Mallat, "Zero-crossings of a wavelet transform," *IEEE Transactions on Information Theory*, vol. 37, no. 4, pp. 1019–1033, July 1991.
[50] R. Hummel and R. Moniot, "A network approach to reconstruction from zero-crossings," in *Proc. IEEE Workshop Computer Vision*, Dec. 1987, pp. 8–13.
[51] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 7, pp. 710–732, July 1992.
[52] Z. Berman, "The uniqueness question of discrete wavelet maxima representation," Tech. Rep. TR 91-48, University of Maryland, College Park, Apr. 1991.
[53] J. F. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.
[54] M. Unser and D. Van De Ville, "The pairing of a wavelet basis with a mildly redundant analysis via subband regression," *IEEE Transactions on Image Processing*, in press.
[55] D. Youla and H. Webb, "Image restoration by the method of convex projections," *IEEE Trans. Med. Imaging*, vol. MI-1, pp. 81–101, Oct. 1982.
[56] E. R. Kandel, J. H. Schwartz, and T. M. Jessell, Eds., *Principles of Neural Science*, McGraw-Hill, 2000.
[57] I. M. Guelfand, M. I. Graev, and N. J. Vilenkin, *Les distributions*, vol. 5, Dunod, Paris, 1970.