16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

# THE MARR WAVELET PYRAMID AND MULTISCALE DIRECTIONAL IMAGE ANALYSIS

*Dimitri Van De Ville, Daniel Sage, Katarina Balać, Michael Unser*

École Polytechnique Fédérale de Lausanne, Biomedical Imaging Group
Station 17, CH1015, Lausanne (VD), Switzerland
phone: + (41) 21 6935142, http://bigwww.epfl.ch

## ABSTRACT

The "Marr wavelet pyramid" is a wavelet decomposition that implements a multiscale version of the complex gradient-Laplace operator. It is closely linked to a multiresolution analysis of $L_2(\mathbb{R}^2)$ and it has a fast filterbank implementation. We show how the Marr wavelets, which are essentially steerable, can be used to extract a multiscale version of the structure tensor. This yields a multiscale characterization of an image in terms of various features such as local gradient energy, orientation, and coherency.

We provide an implementation of the proposed system as a Java plug-in for ImageJ, and we illustrate its applicability to directional image analysis which is useful in domains such as biological imaging and material science.

## 1. INTRODUCTION

The structure tensor is a powerful tool in image processing [1] with a wide range of applications; e.g., feature detection (to retrieve edges, corners, and junctions), diffusion filtering [2], and texture analysis. The underlying principle is to determine the orientation that best matches the local gradient information. It can be applied both to scalar images, by considering a local spatial neighborhood, or to multivalued images, by merging the information from the various channels [3]. The tensor captures the predominant orientation of a local image neighborhood. It also yields other useful features: orientation, isotropy versus anisotropy (coherency), and "cornerness". Bigun et al. [4] also proposed a generalized structure tensor for curved patterns.

The structure tensor has been extended within the scale-space framework [5], and also combined with the wavelet transform [6]. In the latter case, two fully redundant 1-D wavelet transforms are applied separately along rows and columns, respectively. The wavelets are defined as in [7]; i.e., they are the horizontal and vertical derivatives of a 2-D smoothing function, respectively.

Recently, we have proposed a new complex wavelet basis design that is intrinsically 2-D [8, 9]. The wavelet behaves like a multiscale version of the complex gradient-Laplace operator; i.e., the real and imaginary parts of the wavelet coefficients provide the gradient information of a bandpass filtered version of the image. In Sect. 2, we describe our main computational tool, which is a slightly redundant extension of the scheme named the "Marr wavelet pyramid". It has better shift-invariance and rotation-covariance for a total redundancy factor of only $8/3$.

In Sect. 3, we make explicit the link between the Marr wavelet pyramid and the structure tensor. This provides a compact multiscale tensor representation and its associated feature measures. An advantage of our approach is that the Marr wavelet pyramid can be reconstructed, after processing using structure tensor information. Finally, in Sect. 4, we demonstrate the applicability of our approach to some concrete image analysis tasks.

## 2. MARR WAVELET PYRAMID

### 2.1 Complex polyharmonic B-splines

Complex polyharmonic B-splines $\beta_{\gamma,N}$ are 2-D basis functions that are associated with the family of complex gradient-Laplace operators

$$\mathrm{L}_{\gamma,N} = (-\Delta)^{\frac{\gamma-N}{2}} \left(-j\frac{\partial}{\partial x_1} - \frac{\partial}{\partial x_2}\right)^N, \tag{1}$$

where $\gamma \in \mathbb{R}^+$ and $N \in \mathbb{N}$. Let us consider the scaled approximation spaces

$$\mathscr{V}_i = \mathrm{span}\left\{2^i \beta_{\gamma,N}(2^i\mathbf{x}-\mathbf{k})\right\}_{\mathbf{k}\in\mathbb{Z}^2}. \tag{2}$$

Applying the operator $\mathrm{L}_{\gamma,N}$ to a signal of $\mathscr{V}_i$,

$$s_i(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^2} c_i[\mathbf{k}]\ 2^i \beta_{\gamma,N}(2^i\mathbf{x}-\mathbf{k}), \tag{3}$$

results into a train of weighted Diracs, which indicate the knots of the polyharmonic spline.

The fundamental link between the continuous-domain operator $\mathrm{L}_{\gamma,N}$ and the complex polyharmonic B-spline $\beta_{\gamma,N}$ is also reflected in its Fourier domain definition:

$$\hat{\beta}_{\gamma,N} = \frac{V_{\gamma,N}(e^{j\boldsymbol{\omega}})}{\hat{L}_{\gamma,N}(\boldsymbol{\omega})}, \tag{4}$$

where the numerator $V_{\gamma,N}(e^{j\boldsymbol{\omega}})$ is the localization filter, and the denominator $\hat{L}_{\gamma,N}(\boldsymbol{\omega}) = ||\boldsymbol{\omega}||^{\gamma-N}(\omega_1 - j\omega_2)^N$ is the Fourier transform of (1) in the distributional sense.

The localization filter should regularize the singularity of the Fourier transform at the origin. To that aim, we consider the polar representation $V_{\gamma,N}(e^{j\boldsymbol{\omega}}) = V_\gamma(e^{j\boldsymbol{\omega}}) \cdot e^{j\theta_N(\boldsymbol{\omega})}$ and we specify the modulus and phase as

$$\begin{aligned} V_\gamma(e^{j\boldsymbol{\omega}}) &= \left(\frac{8}{3}\left(\sin^2\left(\frac{\omega_1}{2}\right)+\sin^2\left(\frac{\omega_2}{2}\right)\right)\right. \\ &\quad \left.+\frac{2}{3}(\sin^2\left(\frac{\omega_1+\omega_2}{2}\right)+\sin^2\left(\frac{\omega_1-\omega_2}{2}\right))\right)^{\gamma/2}, \\ \theta_N(\boldsymbol{\omega}) &= \angle(j[\omega_1]-j[\omega_2])^N, \end{aligned}$$

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

Notice that the choice of our localization filter is different from [8], where the phase at the origin was not compensated (i.e., $\theta_N(\boldsymbol{\omega}) = 0$). Our construction is also such that $\beta_{\gamma,0} = \beta_\gamma$ coincides with the isotropic polyharmonic B-splines [10] that were earlier introduced as extensions of the elementary polyharmonic B-splines [11].

The complex polyharmonic B-splines satisfy all the requirements for generating a dyadic multiscale analysis of $L_2(\mathbb{R}^2)$ [8]: their integer shifts form a Riesz basis; they satisfy the partition of unity property; and they satisfy a scaling relation. The latter can be expressed in the Fourier domain as

$$\hat{\beta}_{\gamma,N}(2\boldsymbol{\omega}) = \frac{1}{2} H(e^{j\boldsymbol{\omega}}) \hat{\beta}_{\gamma,N}(\boldsymbol{\omega}), \tag{5}$$

where $H(e^{j\boldsymbol{\omega}})$ is the scaling filter. So the approximation spaces are be embedded as

$$\ldots \mathscr{V}_{i-1} \subset \mathscr{V}_i \subset \mathscr{V}_{i+1} \subset \ldots L_2(\mathbb{R}^2). \tag{6}$$

We denote approximation coefficients at scale $i$ as $c_i[\mathbf{k}]$.

### 2.2 Complex polyharmonic wavelet basis

Based on the complex polyharmonic B-splines, we have shown that it is possible to construct a semi-orthogonal wavelet basis[1]:

$$\begin{aligned} \mathscr{W}_{i-1} &= \mathscr{V}_i \ominus_\perp \mathscr{W}_i \\ &= \operatorname{span}\left\{2^i \psi_{\gamma,N}(2^i \mathbf{x} - \mathbf{k}/2)\right\}_{\mathbf{k}\in\mathbb{Z}^2\setminus 2\mathbb{Z}^2}. \end{aligned}$$

The fundamental observation is that computing the wavelet transform of a signal is equivalent to applying a multiscale version of the operator (1).

Notice that the same wavelet is spatially shifted on the (non-zero) coset positions; i.e., the grid $\mathbb{Z}^2\setminus 2\mathbb{Z}^2$ at scale $-1$. Consequently, the corresponding filterbank, shown in Fig. 1, can be implemented using a single wavelet filter $W(\mathbf{z})$ that links the generating wavelet function to the scaling function:

$$\hat{\psi}_{\gamma,N}(2\boldsymbol{\omega}) = \frac{1}{4} W(e^{j\boldsymbol{\omega}}) \hat{\beta}_{\gamma,N}(\boldsymbol{\omega}). \tag{7}$$

The filters at the synthesis side are derived from the perfect reconstruction condition or from considering the dual scaling function and wavelet (in the same spaces). Notice that all filters are non-separable.

### 2.3 Marr wavelet pyramid

The Marr wavelet pyramid is a slightly redundant version of the complex polyharmonic wavelet basis. It includes the "missing" shift from the previous construction and corresponds to the enlarged analysis spaces

$$\mathscr{W}_{i-1}^{+} = \operatorname{span}\left\{2^i \psi_{\gamma,N}(2^i \mathbf{x} - \mathbf{k}/2)\right\}_{\mathbf{k}\in\mathbb{Z}^2}. \tag{8}$$

This leads to a pyramid-like decomposition structure which has an overall redundancy factor of $8/3$. We have also chosen the parameters $\gamma = 3$, $N = 1$, which correspond to the pure complex gradient-Laplace operator. A further refinement is to use the available redundancy to improve its steerability. Specifically, we select the modified analysis wavelet

$$\psi(\mathbf{x}/2) = \underbrace{\Delta\left(-j\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}\right)}_{\mathrm{L}_{3,1}^*} \left\{\beta_{6,0}\right\}(\mathbf{x}), \tag{9}$$

where $\beta_{6,0}$ is the isotropic polyharmonic B-spline, which converges to a Gaussian as $\gamma$ increases. Practically, this amounts to the filterbank for this pyramid-style decomposition is shown in Fig. 2. The shaping filter $A_3(\mathbf{z})$ corresponds to the B-splines autocorrelation filter, which is given in the Fourier domain by

$$A_\gamma(e^{j\boldsymbol{\omega}}) = \sum_{\mathbf{k}} \left|\hat{\beta}_\gamma(\boldsymbol{\omega} + 2\pi\mathbf{k})\right|^2. \tag{10}$$

The synthesis side is performed by projecting the redundant representation onto the basis again according to the subband regression principle [12].

The wavelet coefficients at scale $i$ are denoted as $d_i[\mathbf{k}]$ and correspond to

$$d_i[\mathbf{k}] = \left\langle f, 2^i \psi(2^i \mathbf{x} - \mathbf{k}/2)\right\rangle. \tag{11}$$

We coin the term "Marr wavelet pyramid" for the proposed decomposition as it closely mimicks the basic operations of David Marr's framework for early vision [13]. We refer to [9] for more details.

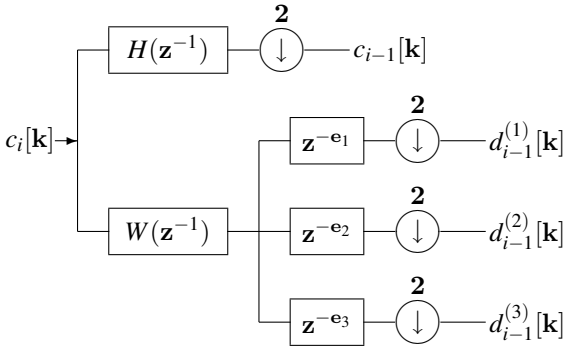


Figure 1: Filterbank implementation of the (non-redundant) complex polyharmonic wavelet basis (analysis side). The vectors $\mathbf{e}_k$, $k = 1,2,3$, represent the non-zero coset positions of the dyadic subsampling matrix; i.e., $[0\ 1]^T$, $[1\ 0]^T$, $[1\ 1]^T$.

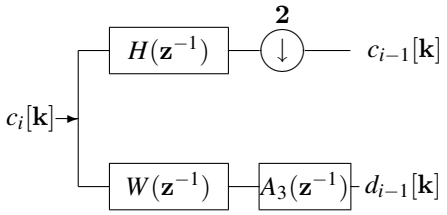


Figure 2: Filterbank implementation of the Marr wavelet pyramid (analysis side).

[1] The quincunx case is treated in [8], while the wavelet design for more general subsampling matrices is dealt with in [9].

## 3. MULTISCALE STRUCTURE TENSOR

### 3.1 From Marr wavelet pyramid to multiscale tensor

To establish the link between the Marr wavelet pyramid and the structure tensor, let us first reconsider the wavelet coefficients of (11), using (9), as

$$\begin{bmatrix} \mathrm{Im}(d_i[\mathbf{k}]) \\ \mathrm{Re}(d_i[\mathbf{k}]) \end{bmatrix} = 2^i \underbrace{\begin{bmatrix} \partial/\partial x_1 \\ \partial/\partial x_2 \end{bmatrix}}_{\nabla} \underbrace{\Delta\left\{\left(f(\cdot) * \beta_6(2^i\cdot)\right)(2^{-i-1}\mathbf{k})\right\}}_{g_i(\mathbf{k})}.$$

Notice that $g_i(\mathbf{k})$ corresponds to a bandpass version of $f(\mathbf{x})$, sampled at rate $2^{-i-1}$. The equivalent standard deviation of the smoothing kernel $\beta_6(2^i\cdot)$ is $\sqrt{2^{-i-1}}$.

The multiscale structure tensor can now be introduced as a $2\times 2$ symmetric, nonnegative definite matrix

$$\mathbf{J}_i(\mathbf{k}) = \sum_{\mathbf{n}\in\Omega} w[\mathbf{n}]\left(\nabla g_i(\mathbf{k}+\mathbf{n})\nabla g_i(\mathbf{k}+\mathbf{n})^T\right), \qquad (12)$$

where the weights $w[\mathbf{n}]$, $\mathbf{n}\in\Omega$, define a non-negative symmetric observation window (typically Gaussian). Notice that each element of $\mathbf{J}_i[\mathbf{k}]$ is a weighted inner product $\langle\cdot,\cdot\rangle_w$ between derivatives of $g_i$. The associated norm of the derivative of $g_i$ along any direction $\mathbf{u} = [\cos\theta \;\; \sin\theta]^T$ can then be obtained as

$$||D_{\mathbf{u}} g_i||_w^2 = \left\langle \mathbf{u}^T\nabla g_i, \nabla g_i^T\mathbf{u}\right\rangle_w = \mathbf{u}^T\mathbf{J}_i\mathbf{u}. \qquad (13)$$

Therefore, maximizing (13) under the constraint $||\mathbf{u}|| = 1$ yields the eigenvector equation. The direction of maximal change is indicated by the eigenvector $\mathbf{v}^+$ which is associated with the largest eigenvalue $\lambda^+$. The smallest eigenvalue $\lambda^-$ is zero, while the eigenvector $\mathbf{v}^-$ is perpendicular to $\mathbf{v}^+$.

While the wavelet coefficients $d_i[\mathbf{k}]$ reflects the local behavior (due to the smoothing kernel), the tensor representation needs multiple observations (through the window $w[\mathbf{n}]$) to become full rank. In the case of identical wavelet coefficients $d_i[\mathbf{n}]$, $\mathbf{n}\in\Omega$, or of a window $w[\mathbf{n}] = \delta_{\mathbf{n}}$, the tensor's rank will be one. The original coefficient can then be retrieved, up to its sign, as

$$d_i[\mathbf{k}] = \pm\sqrt{\lambda^+}[j\;1]\mathbf{v}^+. \qquad (14)$$

The essential information of the structure tensor (that is, the eigenvalues $\lambda^+, \lambda^-$ and the direction $\theta$ of minimal inertia) can be obtained directly from the Marr wavelet coefficients, which exploits the link with complex moments similar to the procedure in [4]; i.e., we have

$$\begin{aligned} M_i^{(1)}(\mathbf{k}) &= 2^{-2i}\langle d_i[\mathbf{k}+\cdot], d_i[\mathbf{k}+\cdot]\rangle_w = (\lambda^+ - \lambda^-)e^{j2\theta}, \\ M_i^{(2)}(\mathbf{k}) &= 2^{-2i}\langle |d_i[\mathbf{k}+\cdot]|, |d_i[\mathbf{k}+\cdot]|\rangle_w = \lambda^+ + \lambda^-, \end{aligned}$$

without the need to construct $\mathbf{J}_i(\mathbf{k})$.

### 3.2 Features from the structure tensor

There are various measures associated with the structure tensor that have been proposed to characterize local image content. Thanks to our multiscale framework, these can be obtained at various scales and positions. From $M_i^{(1)}(\mathbf{k})$, $M_i^{(2)}(\mathbf{k})$, the most popular structure tensor features can be derived as:

- *Energy.* The local gradient energy is given by the trace

$$\text{energy} = \lambda^+ + \lambda^- = M_i^{(2)}(\mathbf{k}). \qquad (15)$$

- *Orientation.* The direction of minimal inertia is indicated by:

$$\theta = \frac{\arg M_i^{(1)}(\mathbf{k})}{2}. \qquad (16)$$

- *Coherency.* The coherency measure reveals the degree of anisotropy of the local structure:

$$\text{coh} = \frac{\lambda^+ - \lambda^-}{\lambda_+ + \lambda^-} = \frac{\left|M_i^{(1)}(\mathbf{k})\right|}{M_i^{(2)}(\mathbf{k})}, \qquad (17)$$

  which ranges from 0 to 1. Large coherency indicates one dominant orientation, small coherency indicates isotropy.

Many other feature measures exist, such as for corners and junctions; e.g., the Harris corner detector [14].

## 4. RESULTS

### 4.1 Examples

In Fig. 4, we first show the Marr wavelet pyramid decomposition for the test image of Fig. 3. The orientation of the complex wavelet coefficient $d_i[\mathbf{k}]$, wrapped within the interval $[0,\pi]$, is shown in Fig. 5 (a). It also corresponds to the direction of the eigenvector $\mathbf{v}^+$ for the case $w[\mathbf{n}] = \delta_{\mathbf{n}}$.

By contrast, the structure tensor is computed using a Gaussian averaging window, which is more reliable for feature detection and texture analysis. Here, we used a Gaussian window with standard deviation $\sigma = 1.5$. The tensor orientation, shown in Fig. 5 (b), gives a much more robust impression of the local direction than in Fig. 5 (a). Finally, in Fig. 5 (c), we have combined the three features in a single image using the HSB channels: hue for orientation, saturation for squared coherency, and brightness for energy. Interestingly, the zoneplate components, which have most of there energy in the second subband (scale $i = -2$), are indeed clearly visible at that scale.

### 4.2 ImageJ plug-in

We have implemented the proposed approach in Java as a plug-in for ImageJ[2], which is freely available to the image processing community. In Fig. 6, we show a screenshot of the application dialog window and the result when analyzing a biological test image. The various parameters of the method and the color encoding for visualization can be configured in a flexible way.

## 5. CONCLUSION

We have proposed a multiscale version of the structure tensor that is based on the Marr wavelet pyramid, a pyramid-type wavelet decomposition with limited redundancy (i.e., a factor $8/3$). This tool can be useful for efficient directional image analysis and tensor-based processing. Further on, using the Marr wavelet pyramid allows to reconstruct the image, after processing based on structure tensor information.

[2] http://bigwww.epfl.ch/demo/orientation/

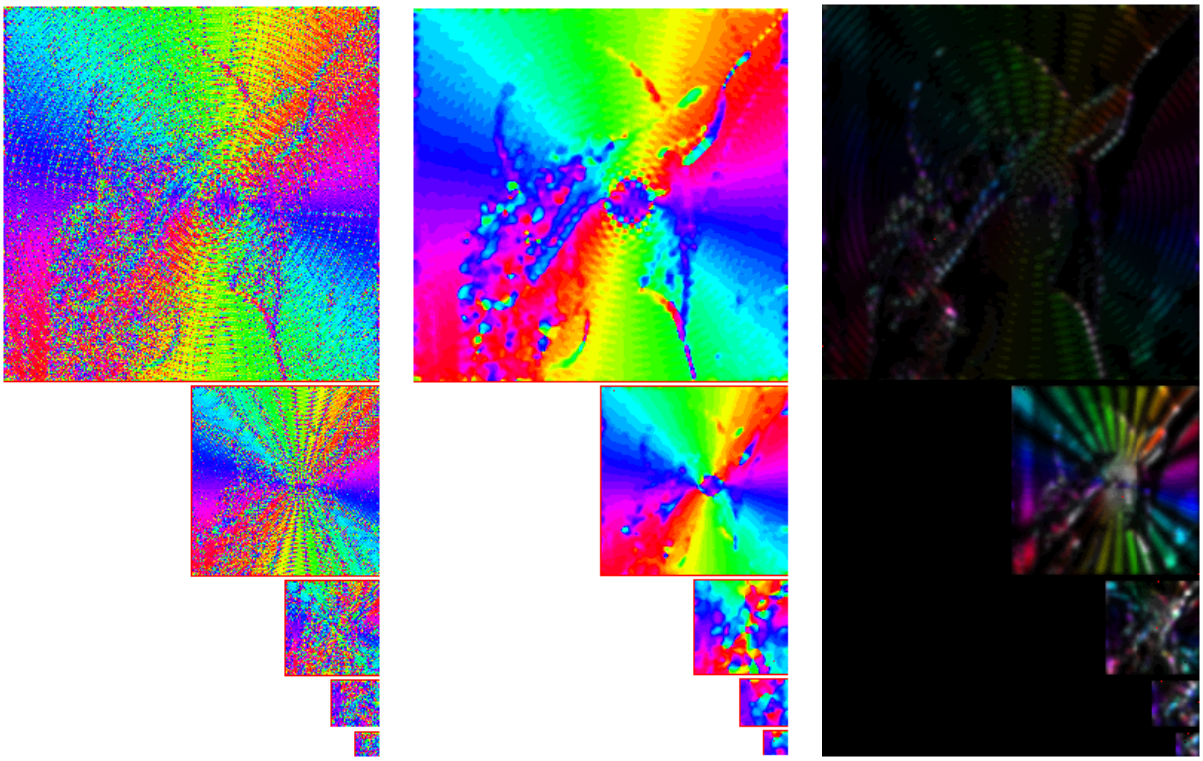

Figure 5: Features from the multiscale structure tensor derived from the Marr wavelet pyramid in Fig. 4 (the orientation is encoded in the hue channel). (a) Orientation without neighborhood ($N = 1$). (b) Orientation with Gaussian observation window ($\sigma = 1.5$). (c) Combined features from the multiscale structure: orientation in hue, squared coherency in saturation, and local energy in brightness.

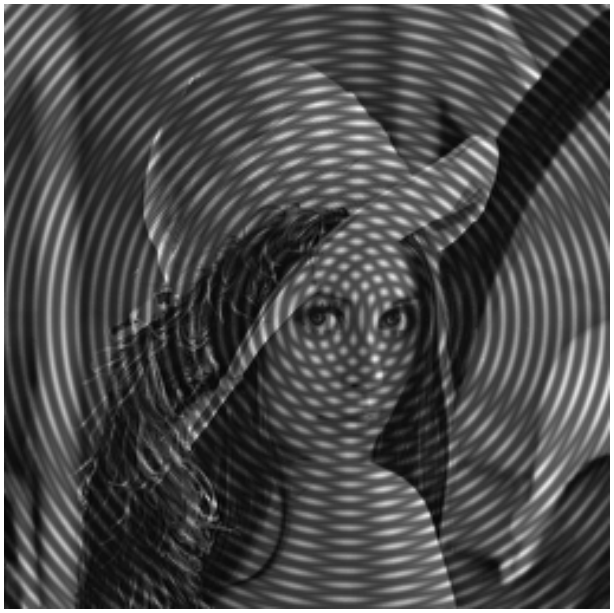

Figure 3: "Psychedelic Lena": combining "Lena" with concentric zoneplate images (of constant frequency) positioned on both eyes.

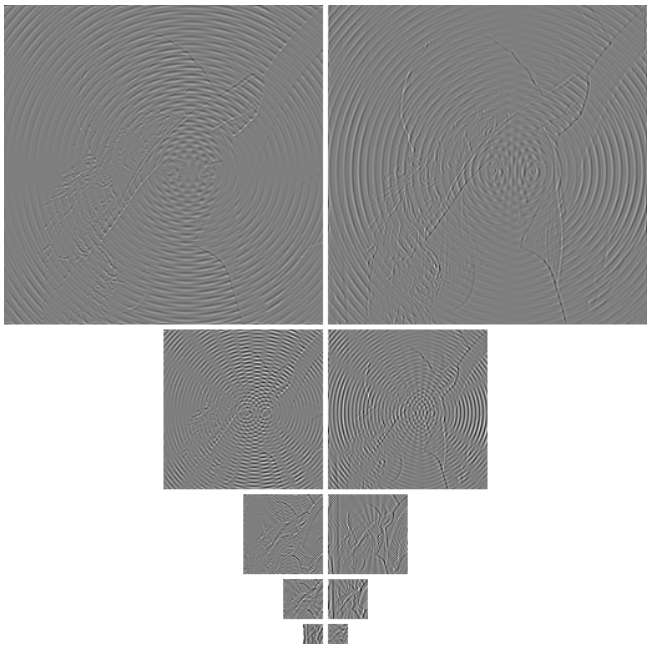

Figure 4: Marr wavelet pyramid for the test image of Fig. 3.

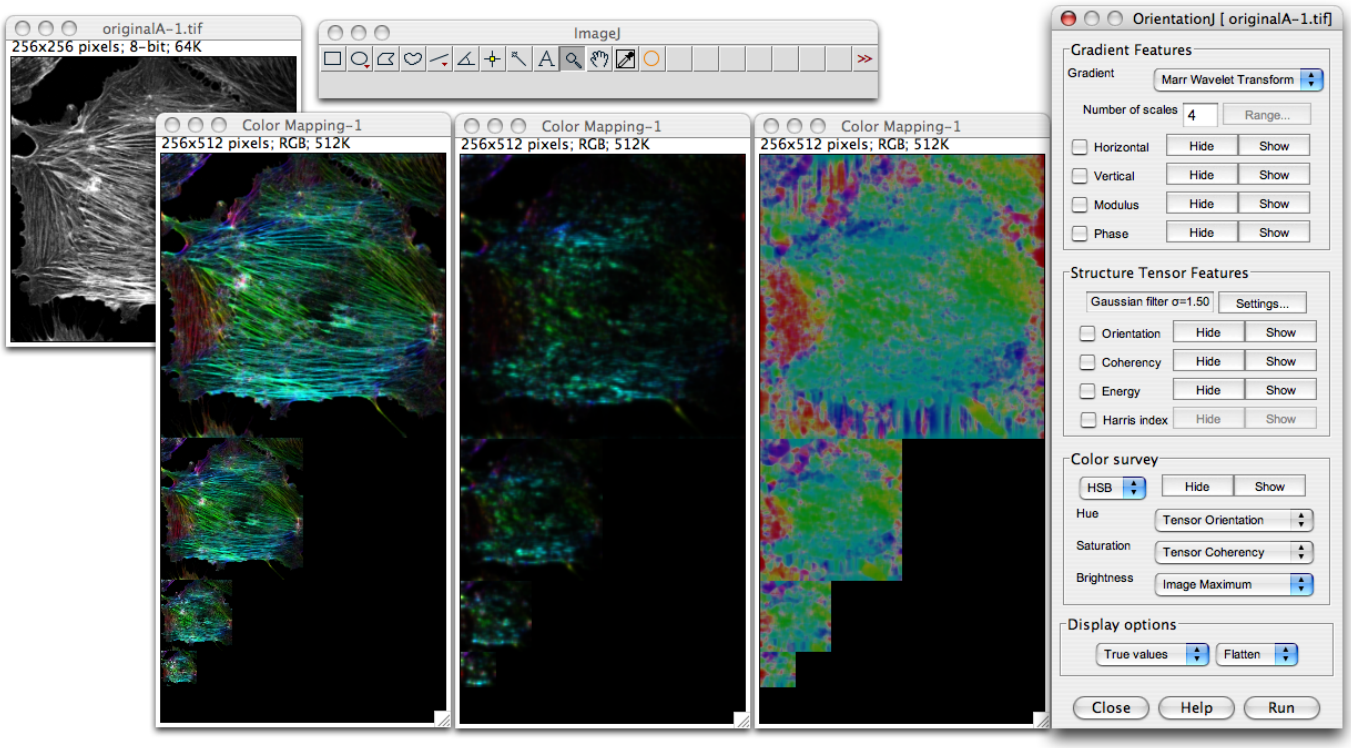


Figure 6: Screenshot of the ImageJ plug-in at work on a biological test image. The Gaussian observation window was set for $\sigma = 1.5$. The various color mapping correspond to (from left to right): (1) hue: orientation, saturation: coherency, brightness: original image; (2) hue: orientation, saturation: coherency, brightness: energy; (3) hue: orientation, saturation: coherency, brightness: maximum. The image is courtesy of Caroline Aemisegger (Zentrum für Mikroskopie und Bildanalyse, Zürich); it shows actin fibres of the cytoskeleton by immunofluorescence.

## 6. ACKNOWLEDGEMENTS

The authors would like to thank François Aguet for providing the original test image (Fig. 3). This work was supported by the Swiss National Science Foundation (KB, MU) under Grant 200020-109415 and by the Center for Biomedical Imaging (DVDV) of the Geneva-Lausanne Universities and the EPFL, as well as the foundations Leenaards and Louis-Jeantet.

## REFERENCES

[1] B. Jähne, *Digital Image Processing*, Springer, 6th edition, 1997.

[2] J. Weickert, "Coherence-enhancing diffusion filtering," *International Journal of Computer Vision*, vol. 31, pp. 111–127, 1999.

[3] S. Di Zenzo, "A note on the gradient of a multi-image," *Comput. Vis. Graph. Image Process.*, vol. 33, pp. 116–125, 1986.

[4] J. Bigun, T. Bigun, and K. Nilsson, "Recognition by symmetry derivatives and the generalized structure tensor," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 12, pp. 1590–1605, Dec. 2004.

[5] T. Lindeberg, *Scale-Space Theory in Computer Vision*, Kluwer Academic Publishers, Dordrecht, 1994.

[6] P. Scheunders, "A multivalued image wavelet representation based on multiscale fundamental forms," *IEEE Transactions on Image Processing*, vol. 10, no. 5, pp. 568–575, May 2002.

[7] 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.

[8] 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.

[9] D. Van De Ville and M. Unser, "Complex wavelet bases, steerability, and the Marr pyramid," submitted.

[10] 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.

[11] C. Rabut, "Elementary *m*-harmonic cardinal B-splines," *Numerical Algorithms*, vol. 2, pp. 39–62, 1992.

[12] 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*, submitted.

[13] 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.

[14] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. of the 4th ALVEY Vision Conference*, 1988, pp. 147–151.