Characterization of color images with multiscale monogenic maxima

Raphaël Soulard and Philippe Carré

Abstract—Can we build a feature-based analysis that fully characterizes images? The literature answers with edge-based reconstruction methods inspired by Marr's paradigm but limited to the greyscale case. This paper studies the color case.

A new sparse representation is carried out with the monogenic concept and the Mallat-Zhong wavelet maxima method.

Our monogenic maxima provide efficient contour shape and color characterization, as a sparse set of local features including amplitude, phase, orientation and ellipse parameters. This rich description takes the wavelet maxima representation further towards the wide topic of keypoint analysis.

We propose a reconstruction process that retrieves the image from its monogenic maxima. While known works all rely on constrained optimization, implying an iterative use of the filterbank, we propose to interpolate the data in the feature domain by exploiting the visual knowledge from the feature-set. This direct retrieval is accurate enough so that no iteration is required.

The main question is finally answered with comparative experiments. It is shown that a reasonably small amount of features is sufficiently informative for visually appealing image retrieval. The features appear numerically stable to rotation, and can be intuitively simplified to perform image regularization.

Index Terms—feature extraction, image color analysis, image reconstruction, monogenic wavelets, wavelet maxima

I. INTRODUCTION

The main information in images is often contained in irregularities and transient phenomena, observed as contours and textures. In the computer vision literature, the contours are generally analyzed by using the *gradient* operator, sometimes used in a multiscale fashion [3], and/or enhanced with the structure tensor formalism [13]. This kind of analysis based on first order derivatives is able to detect edges while characterizing their strength and orientation, as well as their regularity in the multiscale case. As a low-level feature extraction step, it is generally aimed at being exploited in a more general edge-based pattern recognition application.

Since David Marr conjectured that the edge information is sufficient to fully retrieve images [23], several reconstruction methods from low-level features have been proposed [14], [18], [22]. The pioneering "Mallat-Zhong" wavelet maxima representation [22] opened the way to sparse low-level image analysis. Inspired by the Canny edge detection, this tool is able to extract the edges and estimate their regularity in a multiscale fashion. Thanks in part to the wavelet theory, a close approximation of the input image can be retrieved from the edge-related information (and a low frequency component).

Fifteen years later, the well-known limits of *separable* schemes motivated to define the *steerable* "Marr-like pyramid" [31], recently optimized in [5], [25], [29], where the rotation-covariant *Riesz transform* replaces the gradient. These non-separable wavelet frames greatly enhance the invariance of the analysis and the quality of the reconstruction.

Retrieving approximate reconstruction from features confirms the Marr conjecture, and also gives important insights about the information carried by the features. While the literature is strictly limited to greyscale images, we here address the color case: Is it possible to reconstruct a *color* image based on its contours, and if so, how should the *color* contours be defined? This has not yet been answered to our knowledge.

In our previous work [27], we have proposed a color monogenic wavelet transform that leads to a new generation of low-level features, with multivalued coefficients being directly interpretable in terms of local shape and color properties. This can be viewed as an enriched multiscale gradient-like analysis, where not only the local orientation is estimated, but also the phase and color properties. The present paper extends the wavelet maxima representation to a *color monogenic* version, allowing the use of multi-channel images and with improved contour analysis. This richer description takes the wavelet maxima representation further towards the wide topic of *keypoint analysis* by taking monogenic maxima as "feature sets". On the other hand, the finer analysis of the contours can help improving the reconstruction. While the above cited papers all rely on iterative constrained optimization, we propose to exploit the intuition about features to interpolate the wavelet coefficients around the maxima, so that no iteration is required. This choice is also aimed at improving our understanding of the color monogenic features - still recent concepts. In contrast with many pattern recognition approaches for which the analysis method is driven by statistical tools and recognition performance, we here focus on the geometric aspect of the low-level measures.

The (approximative) reconstruction property confirms the "color Marr conjecture" being addressed in the present paper. It also makes the maxima-based analysis a *sparse representation*, which has several advantages. First, the quality and relative completeness of the analysis can be assessed by observing the associated reconstruction. A good similarity between the original image and the retrieved one confirms that the contour information is efficiently extracted at the analysis stage. Second, visual elements can be clearly identified by isolating some features and observing the corresponding reconstructed image. One can control the feature selection/quantization to enhance sparsity. Finally, these tools are not only useful for

pattern *analysis* but also open the way to *processing* tasks such as denoising.

Because the monogenic theory takes wavelets to a higher level by extending coefficients to actual feature vectors, directly interpretable as shape and color information, we expect better sparsity from the wavelet maxima representation. This paper proposes the first color extension of the wavelet maxima based image representation.

Section II reviews the Riesz-based multiscale analysis of contours used in the literature, as well as its color extension from [27]. Two original contributions are added: The color extension is formulated for any number of color channels, and an original wavelet design more adapted to the purpose is proposed. In section III, we present the color monogenic version of the wavelet maxima representation. The multiscale analysis of color contours is described in details. Section IV describes the proposed reconstruction method from the color contour information, and examples of reconstruction, image rotation and denoising are finally presented in section V.

Notations :
D-vectors: $\boldsymbol{x} = (x_1, \ldots, x_D) \in \mathbb{R}^D$
Euclidean norm: $\|\boldsymbol{x}\| = \sqrt{\sum x_i^2}$
Complex algebra: $z = \Re\{z\} + \mathbf{j}\Im\{z\} = |z|e^{\mathbf{j}\arg(z)}$
Fourier transform: $\mathcal{F}\{f\}(\omega) = \int f(t)e^{-\mathbf{j}\omega t}dt$

II. MULTISCALE CONTOUR ANALYSIS

This section presents the multiscale *edge* analysis used in the literature, its evolution to a more general *contour* analysis, and its color extension from our previous work.

A. Riesz-based edge analysis

Given any D-dimensional signal $s(\boldsymbol{x})$, edge detection is generally based on the gradient defined as:

$$(\nabla s)(\boldsymbol{x}) = \left[\frac{\partial s(\boldsymbol{x})}{\partial x_1} \ \cdots \ \frac{\partial s(\boldsymbol{x})}{\partial x_D}\right]^{\mathsf{T}} \overset{\mathcal{F}}{\longleftrightarrow} -\mathbf{j}\boldsymbol{\omega}\,\hat{s}(\boldsymbol{\omega}) \quad (1)$$

The orignal *multiscale gradient* is carried out by prior low-pass filtering with varying cutoff frequency, resulting in a sequence of band-pass operations (since ∇ is naturally high-pass). But a new generation of gradient-like tools is rather based on the Riesz transform:

$$(\mathcal{R}s)(\boldsymbol{x}) = p.v. \int \frac{s(\boldsymbol{x}-\boldsymbol{\tau})\boldsymbol{\tau}}{2\pi\|\boldsymbol{\tau}\|^3}d\boldsymbol{\tau} \overset{\mathcal{F}}{\longleftrightarrow} -\frac{\mathbf{j}\boldsymbol{\omega}}{\|\boldsymbol{\omega}\|}\hat{s}(\boldsymbol{\omega}) \quad (2)$$

The Riesz transform has a unitary frequency gain, making it a pure phase-shifting filter that can be seen as an "all-pass gradient filter". The *multiscale Riesz-transform* is naturally performed jointly with an isotropic filterbank. It has been shown that $\mathcal{R}$ maps any isotropic wavelet frame to another frame [12], [29], with D-valued coefficients (gradient-like x- and y- components for D = 2). We will use the Riesz wavelet frames from [29], that are perfectly isotropic, and tunable by a proper choice of the radial frequency response.
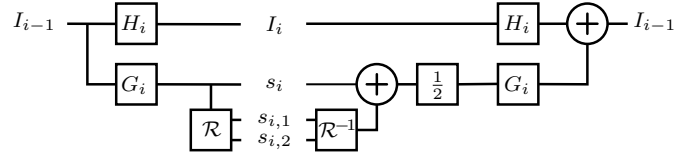

Fig. 1. Undecimated monogenic filterbank flowchart for D = 2.

B. Numerical scheme

We propose to use the undecimated filterbank design, like in [22], because it allows for robust spatial localization of the maxima as well as interscale connection. Note that the further maxima selection will dramatically reduce the redundancy, so that the final amount of data will not be affected by such a choice, compared to the pyramid-like dyadic subsampling scheme from [29]. The tight frame condition is thus released, and the resulting filterbank is depicted in the flowchart of the Fig. 1. This flowchart shows the i-th iteration of both the isotropic filterbank and its corresponding Riesz transform. The synthesis part on the right hand side shows a combined reconstruction of the two filterbanks, according to the monogenic setting explained below. The scale goes from finest $i = 1$ to coarsest $i = L$. The input image $I(\boldsymbol{x}) = I_0(\boldsymbol{x})$ is encoded through its low-frequency residual $I_L(\boldsymbol{x})$ and L isotropic subbands $s_i(\boldsymbol{x})$, as well as L scales of the *multiscale Riesz-transform* as:

$$(\mathcal{R}s_i)(\boldsymbol{x}) = (s_{i,1}(\boldsymbol{x}) \ \ldots \ s_{i,\mathrm{D}}(\boldsymbol{x}))^{\mathsf{T}} \quad (3)$$

Because the essential action of the Riesz transform is a pure phase-shifting operation, the low-pass and high-pass filters H_i and G_i are required to be perfectly neutral with respect to the signal's phase. Their frequency response must then be radial, positive and real-valued. In this context, the perfect reconstruction of the filterbank is ensured by:

$$\hat{H}_i(\boldsymbol{\omega}) = \sqrt{1 - \hat{G}_i(\boldsymbol{\omega})^2} \qquad \hat{G}_i(\boldsymbol{\omega}) = \sqrt{1 - \hat{H}_i(\boldsymbol{\omega})^2} \quad (4)$$

The filters verify the dyadic rule $\hat{H}_i(\boldsymbol{\omega}) = \hat{H}_1(2^{i-1}\boldsymbol{\omega})$, and we propose the following high-pass filter:

$$\hat{G}_1(\boldsymbol{\omega}) = 1 - e^{-\|\boldsymbol{\omega}\|^2/2} \quad (5)$$

This definition has the advantage of being simple and generating a compact wavelet. The filterbank is normalized with constant factors, and computed in the FFT domain [29].

Whether it is based on the gradient (Mallat-Zhong) or the Riesz transform (Marr-like pyramid), the existing edge-based maxima representations are only optimal for *edge*-like contours, while the *line*-like contours are not properly encoded. This *phase-variance* is due to the sole use of $\mathcal{R}s_i$. We propose to overcome this limitation with the monogenic framework, able to detect the local amplitude and phase of D-dimensional signals, by a joint use of s_i and $\mathcal{R}s_i$.

C. Local amplitude, phase and frequency

The local amplitude and phase concepts are well known in 1d within the framework of the *analytic signal*, based on the

Hilbert transform. The *amplitude* draws an 'envelope' over the signal's variations, and the *phase* describes the kind of variation occurring where the amplitude is high. The local *frequency* is finally obtained by differentiating the phase.

Their extension to D-dimensional signals have long been handled with a set of quadrature filters and limited to 2d signals [1], [2], [9], [10]. A definitive method has more recently become popular as the *monogenic analysis* [7], [16], [24], which is isotropic (rotation covariant) and D-dimensional.

The 1d amplitude and phase concepts can be transferred by locally considering the signal as directional. To this end, the *directional Riesz-transform* is defined as:

$$(\mathcal{R}_{\boldsymbol{u}_i} s_i)(\boldsymbol{x}) = (\mathcal{R}s_i(\boldsymbol{x}))^{\mathsf{T}} \boldsymbol{u}_i(\boldsymbol{x}) \tag{6}$$

where $\boldsymbol{u}_i(\boldsymbol{x})$ is a unit D-vector that locally *steers* the analysis (Note that this definition is identical to the directional derivative based on the gradient).

To properly analyze the amplitude envelope and phase of contours, the steering must be locally set to the signal's main orientation. A natural choice is the pointwise Riesz direction $\boldsymbol{u}_i = (\mathcal{R}s_i)/\|\mathcal{R}s_i\|$, analogous to gradient direction, which reduces $\mathcal{R}_{\boldsymbol{u}_i} s_i$ to the Riesz norm $\|\mathcal{R}s_i\|$. Although the original works on monogenic analysis do this, it is sometimes better to use the more robust local orientation information obtained from the *structure tensor* (presented below), that may slightly differ from the pointwise Riesz direction (see also [27], [30]).

The directional Riesz transform steered towards the signal's main orientation is the D-dimensional counterpart of the Hilbert transform. It performs $\pi/2$-phase-shifting, so that the multiscale amplitude and phase features can be defined by:

$$A_i(\boldsymbol{x}) = |\, s_i(\boldsymbol{x}) + \mathbf{j}(\mathcal{R}_{\boldsymbol{u}_i} s_i)(\boldsymbol{x})\,| \qquad \text{(amplitude)} \quad (7)$$

$$\varphi_i(\boldsymbol{x}) = \arg\{\, s_i(\boldsymbol{x}) + \mathbf{j}(\mathcal{R}_{\boldsymbol{u}_i} s_i)(\boldsymbol{x})\,\} \qquad \text{(phase)} \quad (8)$$

We obtain the following modulation model of the signal:

$$s_i(\boldsymbol{x}) = A_i(\boldsymbol{x}) \cos(\varphi_i(\boldsymbol{x})) \tag{9}$$

where A_i is slowly varying with respect to φ_i, and forms an efficient contour detector. The advantage of the monogenic amplitude A_i over gradient-like norms is its *phase invariance i.e.* it detects *lines* and *edges* as well, without producing the well-known "double response" around lines. As a new feature, the phase φ_i provides a line/edge classification of contours, complementary to the orientation information from $\boldsymbol{u}_i$. Analogously to the classical gradient-based wavelet maxima representation, the local maxima of A_i indicate the locations of contours. At those points, the phase φ_i and orientation θ_i describe the shape in a richer way than the usual gradient direction. We will also use the *local frequency* information, computed as the $\boldsymbol{u}_i$-directional derivative of φ_i:

$$\nu_i(\boldsymbol{x}) = (\nabla \varphi_i(\boldsymbol{x}))^{\mathsf{T}} \boldsymbol{u}_i(\boldsymbol{x}) \qquad \text{(frequency)} \quad (10)$$

The frequency was pointed out as a "texture feature" [1], and also contains fine hints about line width and edge smoothness, in combination with the evolution of the amplitude through scales. At the reconstruction stage, ν_i will be a crucial piece of information to linearly interpolate the phase.

An efficient local description of the whole image can then be provided by recording the values of $(A_i, \varphi_i, \theta_i, \nu_i)$ at the locations of the amplitude maxima. This extended "feature vector" with phase-invariance should improve the existing method based on the sole gradient norm and direction. Note that the original translation invariance property still holds.

The monogenic framework does not handle non-directional structures. In a future work, evolutions of the monogenic model [8], [32] taking into account more complex shapes could be used to further enrich the local description of the signal.

We propose now to extend the feature set with color-related features from our previous work [27]. Here, we rewrite the elliptical monogenic analysis for any number of color channels C, so as to open the way to multi-spectral monogenic representation, which is an interesting potential area of research.

D. Multi-channel local orientation

The first step towards the color monogenic framework is the well established color analysis of *edges* popularized by Di Zenzo [6], relying on the structure tensor. The structure tensor formalism combines the gradient with local smoothing to perform robust orientation analysis, and is well defined for multidimensional and multivalued images [13], [28]. Replacing the gradient by the Riesz transform may have an impact on the frequency response of the analysis, but actually allows the same optimal orientation estimation [15], [30]. Given a D-dimensional C-channel image, we consider its i-th scale:

$$\boldsymbol{s}_i(\boldsymbol{x}) = \begin{bmatrix} s_i^1(x_1, \ldots, x_{\mathsf{D}}) \\ \ldots \\ s_i^{\mathsf{C}}(x_1, \ldots, x_{\mathsf{D}}) \end{bmatrix} \tag{11}$$

obtained by applying channelwise the isotropic part of the filterbank (Section II-A). The Riesz-based structure tensor is defined by the following symmetric D × D matrix:

$$\mathrm{T}_i(\boldsymbol{x}) = f * \sum_{c=1}^{\mathsf{C}} (\mathcal{R}s_i^c)(\mathcal{R}s_i^c)^{\mathsf{T}} \tag{12}$$

where $f(\boldsymbol{x})$ is an isotropic smoothing filter.

It can be shown that the local variation of $\boldsymbol{s}_i(\boldsymbol{x})$ - in the sense of Euclidean distance - is maximum in the spatial direction of $\boldsymbol{v}_i$, where $\boldsymbol{v}_i$ is the eigenvector tied to the highest eigenvalue of T_i. This makes the structure tensor a powerful edge detector with accurate orientation analysis.

In the case of images (D = 2), and whatever the number of color channels, the main eigenvector is:

$$\boldsymbol{v}_i = [\, \cos(\theta_i(\boldsymbol{x})) \quad \sin(\theta_i(\boldsymbol{x}))\,]^{\mathsf{T}} \tag{13}$$

with the following closed-form expression for the orientation:

$$\theta_i(\boldsymbol{x}) = \frac{1}{2} \arg\{\mathrm{T}_{i_{1,1}} - \mathrm{T}_{i_{2,2}} + 2\mathbf{j}\,\mathrm{T}_{i_{1,2}}\} \quad \in \left[-\frac{\pi}{2}; \frac{\pi}{2}\right] \tag{14}$$

This value can replace the Riesz orientation for more robust monogenic analysis. This modification alters the perfect reconstruction of the monogenic filterbank, but can be made negligible by choosing a small smoothing filter. In this work, f is a Gaussian filter with variance $\sigma^2 = 0.5^2$. This is sufficiently small for the estimated θ to be very close to the Riesz orientation, and sufficiently large for a stable computation.

The analysis can be enriched with the *coherence* measure:

$$c_i(\boldsymbol{x}) = \frac{\sqrt{\left(T_{i_{1,1}} - T_{i_{2,2}}\right)^2 + 4T_{i_{1,2}}^2}}{T_{i_{1,1}} + T_{i_{2,2}}} \qquad \in [0;1] \qquad (15)$$

The coherence is as close to 1 as the local structure is directional and the orientation measure relevant. In our work, this feature will be used to discriminate isotropic shapes. A discussion about 3d coherence measure can be found in [11], and the topic is open for $D > 3$.

E. Color monogenic analysis

We review here the elliptical monogenic features proposed in [27], and inspired by [19].

The signal's local orientation is defined for any dimension D and any number of color channels C thanks to the structure tensor (see II-D). We assume that we can consider this same spatial orientation for all the channels of $\boldsymbol{s}_i$ so we set

$$\boldsymbol{u}_i = \boldsymbol{v}_i \qquad (16)$$

in eq. (6) to form a vector-valued directional Riesz transform:

$$(\boldsymbol{\mathcal{R}}_{\boldsymbol{u}_i}\boldsymbol{s}_i)(\boldsymbol{x}) = \left[\, (\mathcal{R}_{\boldsymbol{u}_i}s_i^1)(\boldsymbol{x}) \quad \ldots \quad (\mathcal{R}_{\boldsymbol{u}_i}s_i^{\mathsf{C}})(\boldsymbol{x}) \,\right]^{\mathsf{T}} \quad (17)$$

This operation performs channel-wise phase-shifting of $\boldsymbol{s}_i(\boldsymbol{x})$ in the direction $\boldsymbol{u}_i(\boldsymbol{x})$. The direction of analysis is locally the same for all color channels, accordingly to the *directional* model upon which both the structure tensor and the monogenic frameworks are based. This gives access to scalar amplitudes A_i^c and phases φ_i^c obtained from eqs. (7)-(8) as:

$$\boldsymbol{s}_i(\boldsymbol{x}) + \mathbf{j}(\boldsymbol{\mathcal{R}}_{\boldsymbol{u}_i}\boldsymbol{s}_i)(\boldsymbol{x}) = \begin{bmatrix} A_i^1(\boldsymbol{x})e^{\mathbf{j}\varphi_i^1(\boldsymbol{x})} \\ \ldots \\ A_i^{\mathsf{C}}(\boldsymbol{x})e^{\mathbf{j}\varphi_i^{\mathsf{C}}(\boldsymbol{x})} \end{bmatrix} \qquad (18)$$

Then a non-marginal approach is adopted so that these 2C parameters are converted into more intuitive features: unique *amplitude* and *phase* completed by color parameters.

As pointed out in [7], the monogenic amplitude is expected to be locally smooth w.r.t. phase variations, while the phase itself is expected to be locally linear. The frequency, as the derivative of a linear function, should be locally constant. We need to add the following hypothesis: the local frequency is expected to be *the same* at all color channels. This is true for most real-life contours for which the color-channels are highly correlated. Whether it is an edge or a line, the channels often have similar behaviors - up to a multiplicative factor. By assuming that every channel's phase evolves with the same frequency $\nu_i(\boldsymbol{x})$, the signal locally boils down to a 1d vector sinusoid drawing a cyclic elliptical path in $\mathbb{R}^{\mathsf{C}}$, with frequency $\nu_i(\boldsymbol{x})$. This curve is the generalization of the sinusoid from which new amplitude and phase concepts can be defined.

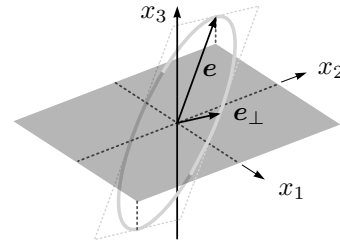The amplitude is intuitively re-defined as the ellipse's size:

$$A_i(\boldsymbol{x}) = \sqrt{\sum_{c=1}^{\mathsf{C}} (A_i^c(\boldsymbol{x}))^2} \qquad \text{(Amplitude)} \qquad (19)$$

The local maxima of A_i will be studied in the next section, as a phase-invariant color generalization of Canny-like edge detection. At maxima locations, all other local features will be recorded. Let us now extend the *phase* concept. The classical definition can be understood as the relative coordinate of a sinusoid's peak. The "peak" of an elliptical oscillation is represented by its *apogee*, which 1d coordinate can be found analytically, providing the definition of the phase:

$$\varphi_i(\boldsymbol{x}) = \frac{1}{2}\arg\left\{\sum_{c=1}^{\mathsf{C}} (A_i^c(\boldsymbol{x}))^2\, e^{2\mathbf{j}\,\varphi_i^c(\boldsymbol{x})}\right\} \qquad \text{(Color phase)} \qquad (20)$$

This formula can be roughly viewed as a weighted average of the channel-wise phases. We will see that φ_i still performs the same efficient classification of lines and edges. A complete representation of $\boldsymbol{s}$ by its features requires additional parameters, able to fully characterize the ellipse's semi-major and semi-minor axes $\boldsymbol{e}$ and $\boldsymbol{e}_\perp$ defined by:

$$\boldsymbol{e} + \mathbf{j}\boldsymbol{e}_\perp = \frac{1}{A_i e^{\mathbf{j}\varphi_i}} \begin{bmatrix} A_i^1 e^{\mathbf{j}\varphi_i^1} \\ \ldots \\ A_i^{\mathsf{C}} e^{\mathbf{j}\varphi_i^{\mathsf{C}}} \end{bmatrix} \qquad (21)$$

so that the multivalued signal is retrieved as the real part of:

$$\boldsymbol{s}(\boldsymbol{x}) + \mathbf{j}(\boldsymbol{\mathcal{R}}_{\boldsymbol{u}}\boldsymbol{s})(\boldsymbol{x}) = A(\boldsymbol{x})\, e^{\mathbf{j}\varphi(\boldsymbol{x})}\,(\boldsymbol{e}(\boldsymbol{x}) + \mathbf{j}\boldsymbol{e}_\perp(\boldsymbol{x})) \quad (22)$$

This is the proposed color monogenic signal. The $2\mathsf{C} - 2$ parameters of $\boldsymbol{e} + \mathbf{j}\boldsymbol{e}_\perp$ are expected to convey comprehensive color information. The first one is the *linearity* feature:

$$\lambda = \|\boldsymbol{e}\|^2 - \|\boldsymbol{e}_\perp\|^2 = 1 - 2\|\boldsymbol{e}_\perp\|^2 \quad \in [0;1] \qquad (23)$$

This parameter defines the semi-major and semi-minor axes lengths $\|\boldsymbol{e}\|$ and $\|\boldsymbol{e}_\perp\|$. In most cases, λ is close to 1, meaning that the semi-minor axis $\boldsymbol{e}_\perp$ is negligible, which reduces the ellipse to a *linear* path between two main colors, pointed by $\boldsymbol{e}$ and $-\boldsymbol{e}$. The vector $\boldsymbol{e}$ carries the information of main *color axis*, and can be encoded by $\mathsf{C}-1$ angles. For 3 color channels ($\mathsf{C} = 3$), we use the two angle encoding proposed in [27]:

$$\boldsymbol{e} = \sqrt{\frac{1+\lambda}{2}} \begin{bmatrix} \cos\alpha_1 \cos\alpha_2 \\ \sin\alpha_1 \cos\alpha_2 \\ \sin\alpha_2 \end{bmatrix} \qquad (24)$$

In some cases like around textured areas, $\boldsymbol{e}_\perp$ may be significant ($\lambda < 1$). It can be defined by its spherical coordinates within the subspace orthogonal to $\boldsymbol{e}$, which requires $\mathsf{C} - 2$ angles. For $\mathsf{C} = 3$, we use the angle α_3 proposed in [27]:

$$\boldsymbol{e}_\perp = \sqrt{\frac{1-\lambda}{2}} R_{\alpha_1} R_{\alpha_2} \begin{bmatrix} 0 \\ \cos\alpha_3 \\ \sin\alpha_3 \end{bmatrix} \qquad (25)$$

with the rotation matrices:

$$R_{\alpha_1} = \begin{bmatrix} \cos\alpha_1 & -\sin\alpha_1 & 0 \\ \sin\alpha_1 & \cos\alpha_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad R_{\alpha_2} = \begin{bmatrix} \cos\alpha_2 & 0 & -\sin\alpha_2 \\ 0 & 1 & 0 \\ \sin\alpha_2 & 0 & \cos\alpha_2 \end{bmatrix} \qquad (26)$$

Of course different axes could have been chosen for the spherical coordinates. As well, different angle wrappings are possible and several angle settings are equivalent (see Appendix). We here retain the setting from [27] where $(\varphi, \alpha_3) \in [-\pi; \pi]$ while $(\theta, \alpha_1, \alpha_2) \in [-\pi/2; \pi/2]$.

This elliptical model combined with the Riesz-based structure tensor extends the monogenic analysis to any spatial dimension D and any number of color channels C. It applies on 2d as well as 3d images, taking values in several color spaces, including multispectral images. In all cases, unique amplitude A, phase φ and linearity λ are extracted, the color axis e is encoded by $C - 1$ angles and the orthogonal axis by $C - 2$ angles. Note that the color space should be Euclidean so that the structure tensor and the elliptical path are well defined. In this paper, we use the RGB space.

Now that the local features are defined, let us use the color monogenic filterbank for maxima-based multiscale contour analysis.

III. CHARACTERIZATION OF CONTOURS WITH MONOGENIC MAXIMA

This section presents the proposed maxima-based representation of color images, using the multiscale monogenic analysis introduced above.

A. Geometric and color information

Let us here study in details how the color monogenic features describe the color contours. We propose to observe the response to a simple synthetic input image containing a yellow rectangle on a red background with a blue line inside. By focusing on one row that crosses the image, we expect to detect from left to right: one *edge* from red to yellow, one blue *line* on yellow background, and one other *edge* from yellow to red. We plot on the Fig. 2 the monogenic analysis on several scales, along this row. On top of the figure is the trimmed image, containing the two edges and the line, and from top to bottom are plotted three groups of curves, corresponding to scales 3 to 5. Within each group, we plotted the amplitude A, the phase φ, the two angles of the color axis (α_1, α_2) and finally the orientation θ.

According to the above mentionned phase-invariance of the monogenic contour detector, we verify that the amplitude A shows well centered peaks above edges and lines as well, which is crucial for the *maxima* representation. Now we can observe the behavior of the other features around the peaks.

Let us begin with the phase. This circular data plotted in $[-\pi; \pi]$ shows "false" discontinuities around $\pm\pi$, but has actually a continuous behavior (wrapping effect). For all displayed scales, the phase φ takes a stable value of $\pi/2$ for the first edge, π for the line and $-\pi/2$ for the second edge. This is exactly what is expected from the monogenic model in terms of contour classification. We experimentally noticed that the evolution of φ in the neighbourhood looks linear, and seems to become "exponential-like" in a larger neighbourhood, with apparent assymptotic behavior towards $\pm\pi/2$. The local frequency ν is of interest to characterize the slope of φ. The frequency tends to evolve by octaves
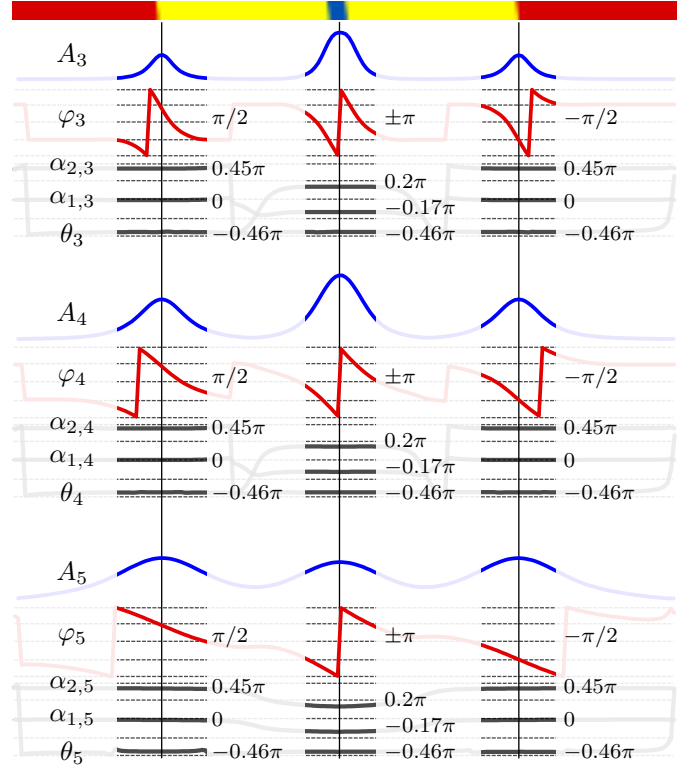

Fig. 2. Linewise observation of monogenic features on 3 scales.

through the dyadic scales, suggesting a Fourier-like harmonic decomposition of the contours. It is clear that the recording of φ and ν at the sole location of the amplitude's maximum is sufficient to characterize φ in the neighbourhood.

Now let us observe the color axis. The values of α_1 and α_2 are very stable and constant in a large neighbourhood. In addition, they take the same values for the two edges, according to the fact that these edges involve the two same colors. More precisely, the encoded color axis is parallel to the axis that joins the red and the yellow in the RGB space. In the case of the blue line, the values of α_1 and α_2 define an axis parallel to the axis joining the blue and the yellow. This confirms the assumption that the *color axis* is locally constant and can be trivially characterized by its value at maxima.

The orientation is as well very stable, equaling -0.46π ($\approx -83°$) according to the orientation of the objects in the test image (Also displayed in the Fig. 3 with lower resolution). The behavior of the features is also stable through scales.

Not plotted here, we have $\lambda = 1$ everywhere, at all scales. Our contours involve a simple linear color path between two main colors, which requires only the color axis (α_1, α_2) to be fully encoded. Thus the angle α_3 is irrelevant (not plotted).

The fact that the monogenic feature values at maxima location are sufficient to describe their neighbouring values proves that they convey rich information. Let us now exploit this property through a maxima based image representation.

B. Bidimensional amplitude maxima detection

The Canny-edge detection, as well as the Mallat-Zhong maxima method and the Marr-like pyramid based edge de-
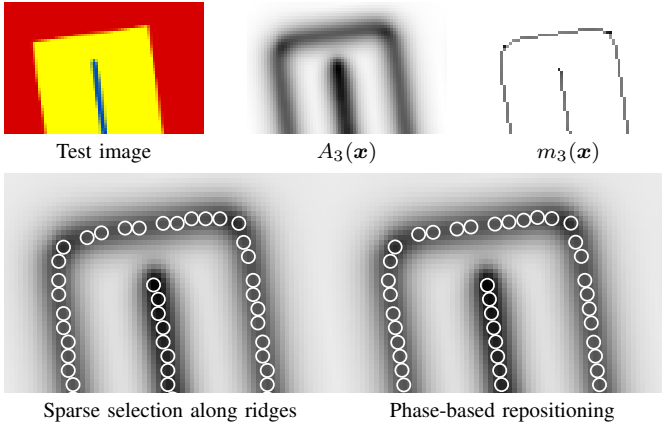
Fig. 3. Amplitude maxima detection and selection at third scale.

tection, all use an oriented 1d maximum detection, using the prior knowledge of the signal's orientation. Given an amplitude sample $A_i(\boldsymbol{x}_0)$, the corresponding local orientation $\theta_i(\boldsymbol{x}_0)$ is quantized to compare with the two neighbour samples $A_i(\boldsymbol{x}_-)$ and $A_i(\boldsymbol{x}_+)$ in the direction orthogonal to the contour. Here we consider the following local differences:

$$d_+(\boldsymbol{x}_0) = A_i(\boldsymbol{x}_0) - A_i(\boldsymbol{x}_+) \tag{27}$$
$$d_-(\boldsymbol{x}_0) = A_i(\boldsymbol{x}_0) - A_i(\boldsymbol{x}_-) \tag{28}$$

The location $\boldsymbol{x}_0$ is recorded if $d_+ > 0$ and $d_- > 0$, indicating that a strict 1d maximum is found.

In this paper, we propose to go one step further by not only detecting maxima, but also measuring their *sharpness*, so as to refine the selection by discarding the "weakest" ones. This can be used to adjust the sparsity of the representation. The sharpness measure of a retained maximum at location $\boldsymbol{x}_0$ is defined by the estimated *slope* of the underlying ridge:

$$m_i(\boldsymbol{x}_0) = \frac{d_+(\boldsymbol{x}_0) + d_-(\boldsymbol{x}_0)}{A_i(\boldsymbol{x}_0)[\cos\theta_i(\boldsymbol{x}_0) \, \sin\theta_i(\boldsymbol{x}_0)](\boldsymbol{x}_+ - \boldsymbol{x}_-)} \tag{29}$$

The term on the numerator is the average absolute difference between amplitude samples. The weighting term A_i normalizes the measure so that small but sharp maxima are well considered. The remaining terms in the denominator correct the bias introduced by the quantization according to a directional ridge model. Note that in practice, very small maxima are discarded with an experimental amplitude threshold of 0.04. An example of sharpness measure is shown on the Fig. 3, where we can see that all the contours are well detected.

1) Enhance the sparsity: In the literature [5], [22], [25], [31], all detected maxima are kept in order to favor closed contours, which facilitates the reconstruction procedure. On undecimated coarser scales, the overly smooth amplitude generates a very redundant succession of maxima. In the present work, we propose to handle a certain resolution adapted to the scale, by selecting the most important maximum in a given neighbourhood of radius σ. The radius should be increasing with the scale, and according to the low-pass filter spatial support. Our wavelet design turns out to correspond to a near-Gaussian low-pass filter, which RMS width experimentally provides $\sigma = 0.7 \times 2^{i-1}$. The process is to sort all maxima by

decreasing sharpness m_i, and to progressively insert them by checking that they do not overlap a previously inserted one. The Fig. 3 (lower row, first plot) shows the obtained regular positioning of the most important maxima, with average distance superior to $\sigma = 2.8$ pixels ($i = 3$). We will see on the next section that an efficient reconstruction is still possible in spite of the least amount of maxima, with equal visual quality.

Note that it is possible to use the phase information to estimate subpixel maxima positions, as illustrated on the Fig. 3 (last image). We have proposed to estimate the positions from a linear phase model. This aspect has no influence on the reconstruction, this is why the method will not be detailed here due to a lack of space. However the phase-based repositioning method will be important for future application (for example color image indexing).

Let us now study the whole feature extraction aimed at constructing local descriptors.

C. Local feature extraction

Once a maximum location has been found, the monogenic feature set can be formed by just picking up the corresponding samples in the wavelet transform. In cases where the location has been refined with subpixel accuracy, the nearest sample still contains a stable information for orientation θ and color features $(\lambda, \alpha_1, \alpha_2, \alpha_3)$, because these features are locally constant. However, the remaining features - amplitude, phase and coherence - need a deeper consideration, and the local frequency has to be computed.

1) Amplitude: If the maximum lies between samples, the nearest amplitude sample should correspond to a mildly attenuated version of the *actual* value. The attenuation can be experimentally computed by using the interpolation function that will be defined in the next section, fed with the relative coordinate of the nearest sample. The final amplitude feature is then obtained by boosting the nearest sample accordingly.

2) Phase: The subpixel shifted cases indicate by definition that the phase value should be a multiple of $\pi/2$. Then we just have to set the feature vector with this multiple

3) Coherence: This feature equals $c = 1$ on straight edges and lines. Note the particular case of lines, for which the Riesz-part is *null* on maxima (as a gradient-like analysis) and the information is fetched from the neighbourhood by the tensor. This is why we still have $c = 1$ and stable θ value on lines. However, this no longer applies on isotropic shapes, well detected by the monogenic framework with a central amplitude maximum and a phase of 0 or $\pm\pi$, but considered irrelevant by the structure tensor, with $c < 1$. The analysis is yet coherent *around* the center of the shape, which borders are seen like *edges*, making the coherence go back to $c = 1$ in a close neighbourhood. This one is as small as the tensor's smoothing kernel is compact (We have chosen a very compact Gaussian kernel with $\sigma = 0.5$). This results in unstable behaviors for the coherence on isotropic shapes. In this work, isotropic shapes are detected by simply thresholding c. The coherence value in the feature vector will be simply picked up from the nearest sample. Note that c is the sole feature that will not be interpolated (Section IV).

$A \in \mathbb{R}^+$	Amplitude	Energy of contour
$\varphi \in [-\pi; \pi]$	Phase	Type of contour (line or edge)
$\theta \in [-\frac{\pi}{2}; \frac{\pi}{2}]$	Orientation	Orientation of contour ($c \approx 1$)
$\lambda \in [0; 1]$	Linearity	Color complexity
$\alpha_1 \in [-\frac{\pi}{2}; \frac{\pi}{2}]$	Ellipse position 1	Color axis
$\alpha_2 \in [-\frac{\pi}{2}; \frac{\pi}{2}]$	Ellipse position 2	Color axis
$\alpha_3 \in [-\pi; \pi]$	Ellipse position 3	Second axis direction ($\lambda < 1$)
$c \in [0; 1]$	Coherence	Directionality of contour
$\nu \in [0; \pi]$ (rad/px)	Frequency	Contour width/Texture feature

(a) Elliptical monogenic feature set.

(b) Maxima on scales 2 (green), 3 (blue) and 4 (brown).

Fig. 4. Elliptical monogenic maxima representation (C = 3, D = 2).

4) Frequency: Like in [27], we compute the local frequency by central finite differences. This requires the neighbouring phase samples to be unwrapped around $\varphi(\boldsymbol{x}_0)$, as detailed in the Appendix. Then the local orientation θ has to be quantized for proper sample selection in the same way as for amplitude sharpness (see section III-B). The frequency is computed as:

$$\nu_i(\boldsymbol{x}_0) = \frac{\varphi(\boldsymbol{x}_+) - \varphi(\boldsymbol{x}_-)}{[\cos\theta_i(\boldsymbol{x}_0) \ \sin\theta_i(\boldsymbol{x}_0)](\boldsymbol{x}_+ - \boldsymbol{x}_-)} \quad (30)$$

where the denominator includes bias correction due to the quantization of θ and based on a linear model.

The complete proposed feature set is listed in table 4a, and an example of analysis is shown on the Fig. 4b, with a real color image. The circles represent the very small number of retained coefficients for the characterization of the contours at each scale. This kind of local description is promising for working towards keypoint descriptors like the well known SIFT method [20], that could be enriched with the phase concept and our color extension. We are currently working on this topic. In the present paper, these features will be used to drive the reconstruction around the maxima (Section IV).

D. Maxima evolution through scales

As already studied with the separable multiscale gradient [22], the structure of objects can be appreciated from the cross-scale evolution of their maxima. This information is related to the regularity of the shapes, which can be formalized by Lipschitz exponents. In the case of monogenic (phase-invariant) analysis, we often observe a clear "maxima jump" between two consecutive specific scales. This jump generally goes from a peak- or line- model with phase 0 or π and centered in the middle of the object, towards a set of edge-maxima on its contours. The Fig. 5 shows different maxima jumps for three objects of different sizes. The amplitude responses show that a same object can be viewed either as an edge or as a large peak, according to the scale of interest. The smaller disk is detected as a contour from scales 1 to 3, and as a peak at all coarser scales. The medium disk shows the jump between the scales 4 and 5, and the bigger disk shows it between 5 and 6 (6th scale is not plotted). This illustrates the significance of using the monogenic framework in a multiscale fashion. We believe that the maxima representation can be efficiently used by connecting those being related to a same
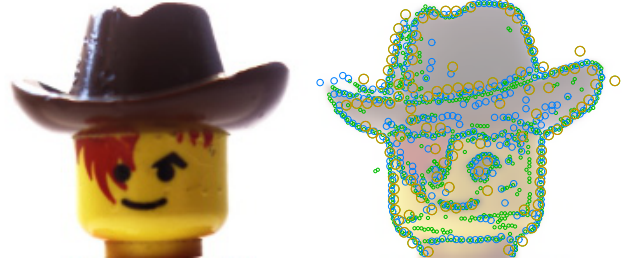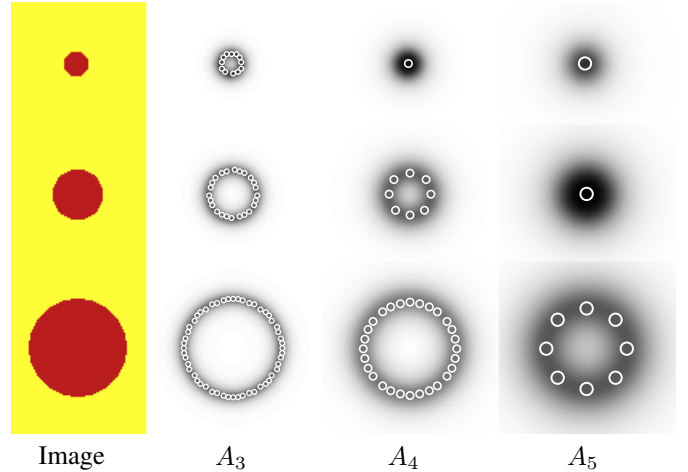

Fig. 5. Maxima jumps through scales.

object at different scales. We expect to reveal a high cross-scale correlation for the features, as already seen in the section III-A, excepted for the amplitude which evolution will be dependent on the regularity. This correlation could be used for enhancing the sparsity, and designing specific color denoising algorithms for example.

As for classical wavelet maxima, the evolution properties of extrema are not guaranteed when the kernel used for generating the scale-space is not a Gaussian or its derivatives. However, the particular scaling function and wavelet we are using is a good approximation to a Gaussian and its first-order derivative. It is reasonable to assume that the scale-space generated with such a wavelet possesses the causality property. Indeed, we have not observed any violations of causality for a large class of signals with which we have experimented. Our wavelet maxima can be viewed as samplings of wavelet maxima traces in the scale-space: some wavelet maxima at different dyadic scales arise from the same edge in the signal; they belong to the same wavelet maxima trace ; the waveforms of the signal have some hierarchy that is shown by the arches in the wavelet maxima traces.

Our objective is to retrieve the structural information that is contained in the monogenic wavelet maxima traces from their dyadic samplings. By using the structural information from the feature set, we extend the Mallat-Zhong method to a representation that will be more suitable for applications of

color image processing and analysis.

The connection of the maxima from two consecutive scales is established if they are spatially close to each other, for which a scale-dependent maximum distance has to be chosen. In most cases, two maxima from a same object have the same location, up to one pixel. We experimentally noticed that a maxima jump occurs from scale i to $i + 1$ for a disk which radius is between 2^{i-1} and 2^i. This means that the distance between the coarser maximum and its potential i-th scale *children* maxima may reach 2^i pixels in the worst case. Note that the decision to connect two maxima could also be influenced by the similarity between the other features, which is an open topic discussed in the last experimental section.

Before studying the color image reconstruction from the monogenic maxima, we have to handle a numerical issue related to the first scale.

E. The special case of first scale

The monogenic first scale is particular because the discretization of the Riesz transform is ill-defined around the Nyquist frequency (The response $\mathbf{j}\boldsymbol{\omega}/\|\boldsymbol{\omega}\|$ conflicts with spectrum periodicity). It produces spatial artifacts - see A_1 in Fig. 6 - that may lead to false maxima. The artifacts are even stronger
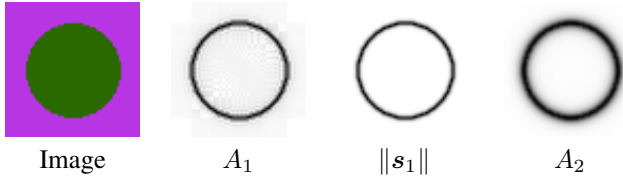

Fig. 6. Monogenic amplitude at 1st scale.

with the "Simoncelli" and "$h_{U_{1_D}}$" wavelets from [25], [26] (Although this is overcome in [26] by a high-pass residual). At subsequent scales, the near-zero Fourier response in the Nyquist area regularizes the discontinuity, making the issue negligible (see how A_2 is smooth).

In the present work, we propose not to use the Riesz transform at first scale. This can be handled theoretically by considering that $\boldsymbol{\mathcal{R}}_{\boldsymbol{u}_1}\boldsymbol{s}_1$ vanishes at all sampling points - or equivalently $\varphi_1 \in \{0, \pi\}$. The amplitude thus boils down to:

$$A_1(\boldsymbol{x}) \leftarrow \|\boldsymbol{s}_1(\boldsymbol{x})\| \tag{31}$$

From the contour detection viewpoint, the sole use of the isotropic part $\boldsymbol{s}_1$ (Behaving like a Laplacian) is known to restrict the analysis to *lines* and *peaks*, while producing a double response on both sides of *edges*. However at this scale, the critical sampling makes this theoretical double response an apparent two-sample thick single response - see $\|\boldsymbol{s}_1\|$ on Fig. 6. It turns out that $\|\boldsymbol{s}_1\|$ is an efficient artifact-free approximation of the expected monogenic amplitude.

Now the other features must be extracted. Based on the strong dependencies observed in the section III-A, between two neighbour scales, we propose to extrapolate the first scale features from the second scale ones. The contribution of the first scale coefficients is then reduced to *locations* and *amplitudes* of the maxima, while the phase and all other features are inherited from the second scale. This simplification

restricts the first scale data to contour harmonics, and tends to discard the possible high-frequency textural content. This can be viewed as regularization, in a denoising context.

Because most high-frequency edge harmonics are centered between two samples, subpixel accuracy is required. Thus we propose to interpolate A_1 at all half-coordinates with bicubic splines, so as to reveal explicit edge peaks, and capture their subpixel positions. This is classical with the Mallat-Zhong approach, but in our case the phase value has then to be refined. Note that the spatial density of the detected maxima is still controlled by the cleaning process explained above, so that the interpolation does not increase the actual amount of maxima.

IV. RECONSTRUCTION FROM FEATURES

After having studied the contour model, we try in this section to verify the color Marr conjecture. The goal is to retrieve the visual content of color images from its contours (and a low frequency component). This result is aimed at confirming that our monogenic features carry significant and complete information, thus validating the proposed analysis. We first explain the synthesis part of the monogenic wavelet transform, and then propose a method to retrieve all the wavelet coefficients from the sole list of monogenic maxima.

A. Monogenic wavelet synthesis

Here is reviewed the color monogenic filterbank synthesis proposed in [27] as the "pure-directional method". This part of the algorithm naturally comes *after* the interpolation step presented below.

The 3-channel 2d monogenic filterbank provides at each scale 9-valued coefficients which Cartesian coordinates are $s_i^1, s_{i,1}^1, s_{i,2}^1, s_i^2, s_{i,1}^2, s_{i,2}^2, s_i^3, s_{i,1}^3, s_{i,2}^3$. We have converted them into the feature vector $(A, \varphi, \theta, \lambda, \alpha_1, \alpha_2, \alpha_3)$ in section II-E. To feed the synthesis part of the filterbank, we need to convert these features back into the Cartesian terms. The isotropic part s_i^1, s_i^2, s_i^3 is trivially retrieved from the real part of eq. (22), which uses all the features except θ, while the directional Riesz transform comes from the imaginary part as $\boldsymbol{\mathcal{R}}_{\boldsymbol{u}_i}\boldsymbol{s}_i$. The x- and y- components of the Riesz transform are finally obtained by using θ as:

$$\begin{bmatrix} s_{i,1}^1 & s_{i,1}^2 & s_{i,1}^3 \end{bmatrix}^{\mathsf{T}} + \mathbf{j} \begin{bmatrix} s_{i,2}^1 & s_{i,2}^2 & s_{i,2}^3 \end{bmatrix}^{\mathsf{T}} \leftarrow (\boldsymbol{\mathcal{R}}_{\boldsymbol{u}_i}\boldsymbol{s}_i)e^{\mathbf{j}\theta} \tag{32}$$

This assignment turns $\boldsymbol{\mathcal{R}}_{\boldsymbol{u}}\boldsymbol{s}$ back to the original Riesz coefficients with the condition that θ coincides with the pointwise Riesz direction *i.e.* $\boldsymbol{u}_i = (\mathcal{R}s_i^c)/\|\mathcal{R}s_i^c\|$ on every channel c separately. This is not the case for two reasons. First, we have chosen the more robust structure tensor based local orientation measure $\boldsymbol{u}_i = \boldsymbol{v}_i$, which is slightly different from the Riesz direction (Sec. II-C) (Even if the choice of a small smoothing kernel limits this difference). Second, we have chosen to estimate a unique local direction θ for the whole color contour, which implies a loss of information in cases where the color channels show different directions (Sec. II-E). Once the Cartesian terms are obtained, the classical monogenic

synthesis, depicted in the Fig. 1 for the greyscale case, is based on the inverse Riesz transform

$$\mathcal{R}^{-1}[\, s_{i,1}\,,\, \ldots\,,\, s_{i,\mathrm{D}}\,]^{\mathsf{T}} = -\mathrm{tr}\,([\, \mathcal{R}s_{i,1}\,,\, \ldots\,,\, \mathcal{R}s_{i,\mathrm{D}}\,]) \tag{33}$$

verifying $\mathcal{R}^{-1}\mathcal{R} = 1$, and the summation of all the outputs (see [27] for a detailed study of the monogenic reconstruction). The great property of this "pure-directional" reconstruction method is that the precise shape of every synthesis wavelet is parameterized by its associated monogenic "coefficient" that is in fact a full "feature set" $(A, \varphi, \theta, \lambda, \alpha_1, \alpha_2, \alpha_3)$. This allows to retrieve isotropic peaks as well as directional edges, while handling their geometric and colorimetric parameters in a unified wavelet theory. As a counterpart, the filterbank inversion is not exact theoretically. However, we have experimentally noticed no visible artifact with usual color images [27]. Note that it only holds when wavelet coefficients are all kept unchanged, which is not the case in a maxima based representation. The last step of the algorithm is then to retrieve the missing coefficients by interpolating around the maxima.

B. A feature-domain approach

The existing papers on wavelet maxima based reconstruction [5], [22], [25], [31] always rely on constrained optimization implying the iteration of the filterbank. In addition, the interpolation is classically performed on all the Cartesian coordinates separately. For example, the Mallat-Zhong method [22] or the Marr-like pyramid [31], interpolate the x- and y-components of the multiscale gradient/Riesz coefficients.

We have already distinguished our work by proposing a sparse selection of the maxima, while the existing literature just keeps all maxima samples. Let us here again take an original approach for the reconstruction. Because we have driven our study by the interpretation of the monogenic features, we will perform the interpolation in the *feature domain*. The coefficients will be retrieved in the neighbourhoods of the maxima by taking into account the local color and shape information conveyed by the features, as well as the knowledge about them - local constancy of angles, smoothness of amplitude, linearity of the phase... From the scientific viewpoint, we believe that the understanding of their nature can be improved by such a study. Moreover, this knowledge can be used in an applicative context to regularize or simplify the data contained in the extracted maxima *i.e.* enhance the sparsity. Finally, we expect this direct interpolation to be accurate enough so that no iterative smoothing will be required, which would provide a computational advantage over existing works. Promising experiments will be presented in the last part of the paper.

The feature-domain interpolation around maxima first requires a 2d segmentation step.

C. Segmentation around maxima

While the original Mallat-Zhong method trivially interpolates *between* two maxima on all rows and columns, the more recent works just do not interpolate, and let the iteration of the filterbank "fill in the blanks" progressively. Here the row-column method is not adapted to the isotropy of the

Original A_4 Segmented areas   Reconstructed A_4
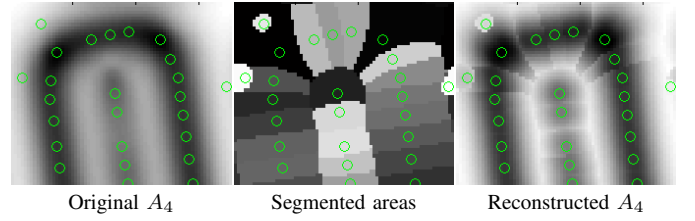
Fig. 7. Segmentation example at 4-th scale.

tool. Instead, interpolating "between" maxima in 2d requires an adapted 2d segmentation process defining the areas of influence around each maximum[1].

In other words, all spatial positions $\boldsymbol{k}$ need to be labelled with the index of the locally most influential maximum. The first idea is to retain the closest one in terms of Euclidean distance. However, high amplitude maxima must intuitively take up *larger* areas. It is then better to weight the distance of a maximum by its amplitude value, so that the sizes of the areas are proportional to maxima's height.

The Fig. 7 shows an example of segmentation with weighted distances. We can see that the segmentation allows to draw an amplitude envelope that resembles the original (The interpolation method is explained below). The junctions between the segmented areas may produce discontinuities, but they will be smoothed when finally calling the synthesis filterbank.

D. Interpolation

Once the area of influence around every maximum has been defined, these neighbourhoods have to be filled with interpolated values of each monogenic feature $(A_0, \varphi_0, \theta_0, \ldots)$. While this interpolation is trivial for most of them - assumed locally constant - the amplitude and the phase require a deeper study - begun in section III-A. The apparent behavior of the analysis on typical lines and edges leads us to propose experimental local models for the amplitude and the phase.

1) Interpolating the amplitude: We saw that the amplitude is bell-shaped and directional over both lines and edges. As a first step, we propose the following bell-shaped radial model:

$$g(r) = \left(\frac{2}{3} e^{\frac{-r^2}{2\sigma^2}} + \frac{1}{3} e^{\frac{-r^2}{4(2\sigma^2)}} \right) \tag{34}$$

The width σ is fixed for every scale, and the anisotropy will be controlled by coordinate changes. This model has a higher tail than a simple Gaussian, and experimentally fits better the amplitude's behavior (See Fig. 8, upper part) for both lines and edges. Now this model must be given a 2d shape.

First the coherence is used to discriminate between isotropic cases ($c_0 < 1$) and directional ones ($c_0 \approx 1$). The former show near 0 or π phase, and we use isotropic coordinate $r = \|\boldsymbol{x}\|$ to set $A(\boldsymbol{x}) = A_0 g(\|\boldsymbol{x}\|)$ with $\sigma = 0.42 \times 2^{i-1}$. The latter are handled by exploiting the anisotropy to lengthen the amplitude kernel along the contour. Moreover, the local frequency ν_0 informs about contour's *thickness*. We can modulate its width with respect to an expected frequency due to the scale. So as to

---

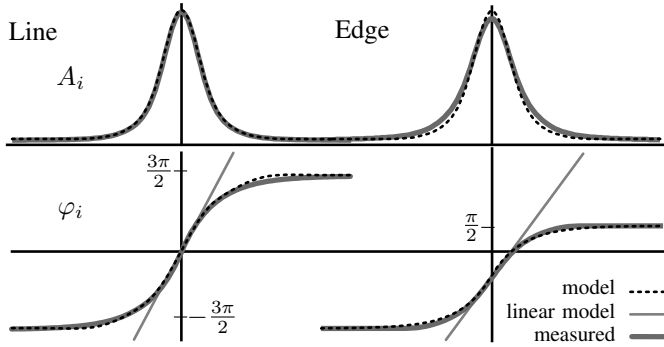[1]This could be related to the concept of *cone of influence* [21]

Fig. 8. Modeling of amplitude and phase for ideal lines and edges.

independently dilate the interpolation along and away from the contour, we consider the orthogonal and parallel coordinates:

$$x_\perp = [\cos\theta_0 \sin\theta_0]\boldsymbol{x} \qquad (35)$$

$$x_\parallel = [-\sin\theta_0 \cos\theta_0]\boldsymbol{x} \qquad (36)$$

The modulation of the contour width is controlled by the following dilation rule depending on the ratio between the measured and expected local frequencies:

$$x_\perp \leftarrow x_\perp \frac{\nu_0}{\nu_{\text{ref}}} \quad \nu_{\text{ref}} = \begin{cases} 0.95\pi/2^{i-1} & \text{for lines} \\ 0.69\pi/2^{i-1} & \text{for edges} \end{cases} \qquad (37)$$

Now the envelope can be lengthened in the direction parallel to the contour so as to properly join between maxima chaining along an amplitude ridge. The proposed lengthening rule is:

$$x_\parallel \leftarrow \begin{cases} |x_\parallel| - l & \text{if } |x_\parallel| > l \\ 0 & \text{otherwise} \end{cases} \quad l = c^{16} \times 0.35 \times 2^{i-1} \quad (38)$$

The envelope is then maintained at its maximum value in the interval $[-l; l]$, and decays outside. We set l as half the *cleaning radius* (Section III-B1), weighted by the coherence. This value allows proper junction between maxima. The coherence weighting term c^{16} reduces the lengthening for less directional cases. Interpolated amplitude can finally be computed as:

$$A(\boldsymbol{x}) = A_0 \, g\left(\|(x_\parallel, x_\perp)\|\right) \qquad (39)$$

with $\sigma = 0.6 \times 2^{i-1}$. The example on the Fig. 7 shows well retrieved amplitude ridges, even between maxima, and faithful thickness thanks to the use of ν.

2) Interpolating the phase: The most obvious interpolation method for the phase data is based on the linear model:

$$\varphi(\boldsymbol{x}) = \varphi_0 + \nu_0 x_\perp \qquad (40)$$

However, as seen in section III-A, the observed behavior of the phase actually fits the linear model in the closer neighbourhood, but shows an exponential-like behavior further away, and tends to some $\pm\pi/2$ constant. The asymptote $\varphi_0 + K = \pm\pi/2$ (modulo 2π) experimentally verifies:

$$K = \begin{cases} \text{sgn}(\nu_0)\left(\frac{\pi}{2}(2\lfloor q/2 \rfloor + 3) - \varphi_0\right) & \text{if } x_\perp \geq 0 \\ \text{sgn}(\nu_0)\left(\frac{\pi}{2}(2\lceil q/2 \rceil - 3) - \varphi_0\right) & \text{if } x_\perp < 0 \end{cases} \qquad (41)$$

with $q = \text{round}(\varphi_0/(\pi/2))$. We thus propose the following exponential phase model, defined to have a slope at the origin

close to ν and to reach the asymptote K on both sides:

$$\varphi(\boldsymbol{x}) = \varphi_0 + K\left(1 - e^{|x_\perp| \log_2\left(1 - \frac{|\nu_0|}{K}\right)}\right) \qquad (42)$$

The lower part of the Fig. 8 shows how both phase models fit the measured phase values. The exponential model will be used in the subsequent experiments.

3) Interpolating the other features: All other features are considered constant over the whole neighbourhood of the maximum. This includes orientation θ, and ellipse parameters $(\lambda, \alpha_1, \alpha_2, \alpha_3)$. We have observed no experimental issue due to this very simple aspect of the method. This is a main advantage of representing an image by its monogenic maxima: most features are very consistent.

4) Using the synthesis filterbank: When all the monogenic coefficient have been interpolated, the last step is to call the synthesis filterbank. The redundancy factor of the set of color monogenic wavelet coefficients is 9 times the number of scales, which is quite important. The advantage of this redundancy is that the synthesis filterbank intrinsically performs some regularization, by projecting our artificial set of coefficients onto the set of feasible wavelet transforms. This property is used in the literature to iteratively approach the best fit between the known maxima and the possible wavelet transforms. In the present work, it is used only once, because we focused on a direct high quality interpolation.

V. Experimental validation of the representation

After having defined the feature extraction and a full reconstruction method, this last section aims at evaluating the quality of the proposed keypoint-based representation, in order to confirm the "color Marr conjecture", and to characterize the information carried by the features. We focus on partial reconstructions, which quality is expected to gradually increase with the number of maxima. Comparison is made with channel-wise ("marginal") use of the classical Mallat-Zhong method, and an illustrative preliminary denoising experiment is shown.

A. Examples of reconstruction

For this type of work the subjective quality assessment is most important. Nevertheless, some insight can be gained from the use of objective algorithmic quality measures. While many methods are known for greyscale images (PSNR, SSIM...) we found only one usable source code for *color* quality assessment: The DSSIM measure [17]. This tool is an advanced multiscale color extension of the classical SSIM based on the L*a*b color space. It equals 0 when the image is identical to the reference, and increases as the quality degrades.

Partial reconstructions from our monogenic maxima are shown on Fig. 9-(b,c,d). The main observation is that a visually very good reconstruction is obtained with about 12000 maxima (Fig. 9-(b)), while the original image contains about 200000 color pixels. Note that the low-frequency band is very coarse, ensuring that the contour information is actually carried by the maxima. Color and shape of all contours are properly retrieved. The dim shadow-like artifacts around the hat and shoulders are due to coarse scale interpolation bias, that do not really
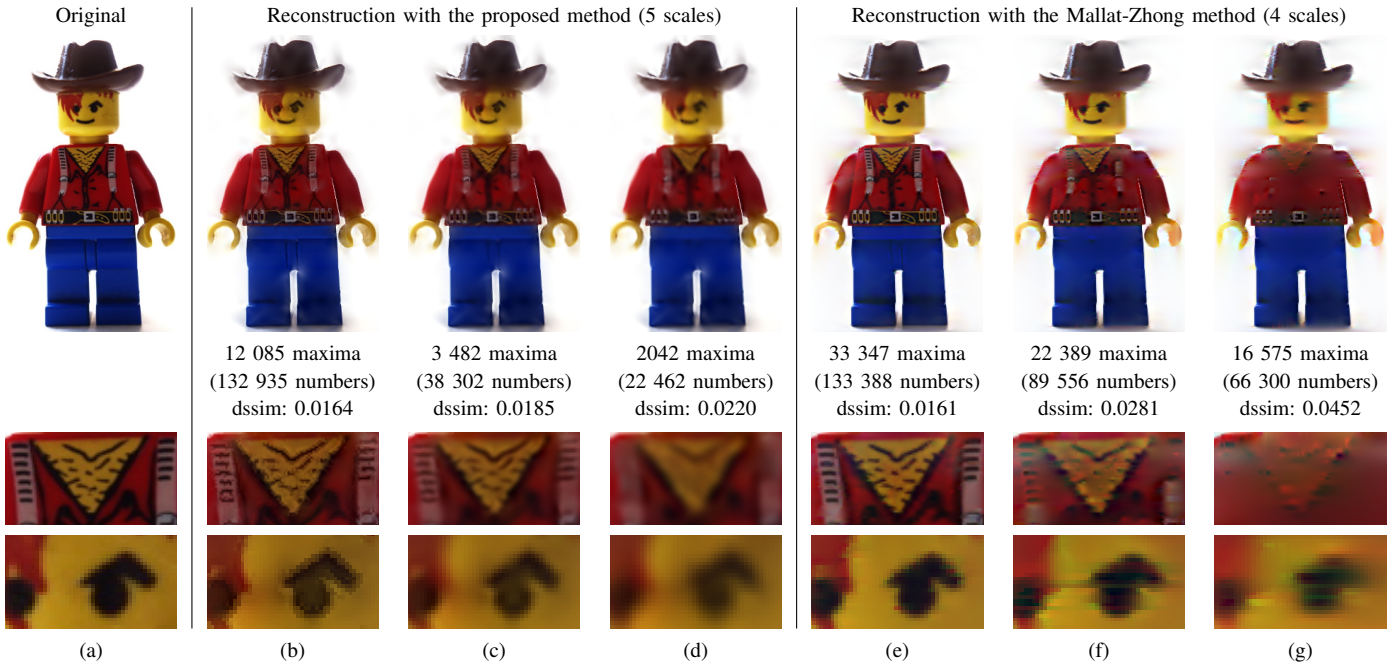
Original	Reconstruction with the proposed method (5 scales)			Reconstruction with the Mallat-Zhong method (4 scales)		
	12 085 maxima (132 935 numbers) dssim: 0.0164	3 482 maxima (38 302 numbers) dssim: 0.0185	2042 maxima (22 462 numbers) dssim: 0.0220	33 347 maxima (133 388 numbers) dssim: 0.0161	22 389 maxima (89 556 numbers) dssim: 0.0281	16 575 maxima (66 300 numbers) dssim: 0.0452
(a)	(b)	(c)	(d)	(e)	(f)	(g)

Fig. 9. Reconstructions based on 5-scale monogenic maxima and 4-scale channel-wise Mallat-Zhong method. Lower rows: zoomed areas.
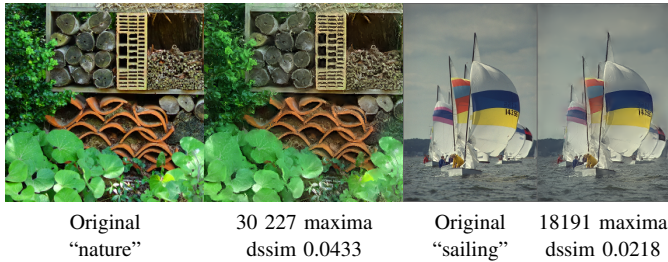
Fig. 10. Reconstructions with the proposed method on 5 scales.

compromise the overall quality. At this stage we can say that it *is* possible to reconstruct a color image based on a point-wise description of its contours. The Fig. 10 shows some other examples of reconstruction, validating the hypothesis of the paper. Now let us verify the consistency of the representation.

Sorting the maxima by descending amplitude and keeping the highest ones recovers simplified reconstructions, see Fig. 9-(c,d). See how the details progressively vanish while main elements remain, with no particular artifact.

Note that in spite of our interest for compactness, this work is not intended to perform any *compression* algorithm. Such methods require non-redundant transforms, optimized quantization, entropy coding etc. which is out of the scope of this paper.

B. Partial reconstruction: Comparison with existing method

A related work can be found in [5], proposing a monogenic extension of the image reconstruction from edges. However, the approach is very different from the present paper. Like in [22], [25], [31], it is based on constrained optimization and limited to greyscale images. The reconstruction acts on the Cartesian coordinates, while we have focused on the physical interpretation of the features in terms of local descriptors. We

believe that our approach is necessary for a color extension that surpasses the *marginal* method (Channel-wise analysis).

Yet, in order to provide some intuition about the relation with existing methods, we choose to compare with the state-of-the-art reference Mallat-Zhong maxima representation [22]. This greyscale tool can be used for color images in two ways. Either the image is restricted to its luminance, then no color can be retrieved, and some isoluminant contours may disappear; or the tool is applied channel-wise (*marginally*). The latter method produces 3 sets of gradient-like maxima, one by color channel. Their amplitude is defined as the gradient norm, so that we can sort them as we did above. Note that keeping n maxima implies $4n$ coefficients with the Mallat-Zhong method (two gradient values plus two spatial coordinates), and $11n$ with our method (9 features plus two coordinates). We must compare regarding the equivalent number of coefficients instead of the actual number of maxima. The number of scales is considered so that the low-frequency component contains an equivalent amount of information in both methods. We use 5 monogenic scales and only 4 Mallat-Zhong scales, which in both cases implies a very coarse residual, like out of a strong Gaussian smoothing. The reconstructions with the Mallat-Zhong method are displayed on the Fig. 9-(e,f,g). By using the same amount of data as for reconstruction (b), we obtain a globally comparable quality (e). But comparing the zoomed areas on (b) and (e), the Mallat-Zhong method raises the famous *false color* issue (see the green and red pixels around the eye, (e) zoomed area) while our method is totally stable regarding color. The shadow-like artifacts also look stronger. The advantage of the proposed approach becomes obvious when decreasing the number of maxima. Comparing (c) *vs.* (f) and (d) *vs.* (g), the *marginal* Mallat-Zhong method shows more and more false colors as well as horizontal and vertical strong artifacts due to its separable design, whereas our method
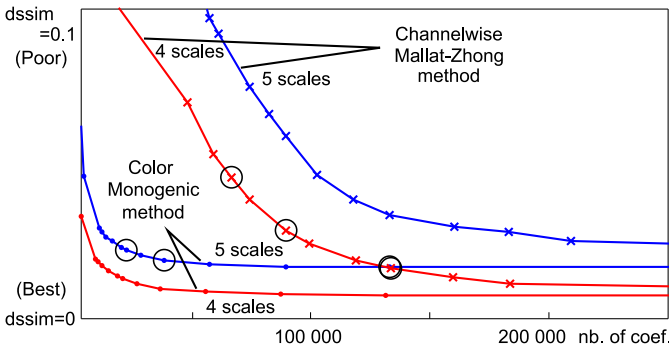
Fig. 11. Trade-off between reconstruction quality (y-axis) and maxima data size (x-axis). The LEGO image from Fig 9 (a) is used. The circled points indicate the examples shown on Fig. 9.

Fig. 12. Image rotation by coordinate change of the maxima on 5 scales. Upper row: Original and reconstructed. Lower row: zoomed regions.

remains stable. Furthermore, the Mallat-Zhong method turns out to require much more maxima than the proposed method, which is more compact thanks to a better encoding of the color information and the cleaning procedure (Sec. III-B1).

Let us now make a more objective comparison by using the DSSIM quality measure. The Fig. 11 shows the progression of the quality according to the amount of maxima (number of coefficients). Both 4-scale and 5-scale settings are shown to avoid ambiguities about the contribution of the LF band. These curves clearly show the superiority of our method in terms of sparsity *i.e.* better content can be retrieved with far less coefficients. The reference Mallat-Zhong method is yet able to reach, with a very large number of maxima, a better overall quality; while the proposed approach quickly tends to a limited quality that will not be improved by adding more and more data (even when disabling the cleaning procedure).

Our color monogenic maxima representation reveals itself a sparse way to describe color images based on a set of keypoint-like feature sets. The sparsity could further be improved by exploiting the redundancy within the maxima themselves, especially their expected significant cross-scale correlation.

C. Color Monogenic maxima and Rotation

Let us here illustrate the robustness of features against discrete rotation, by using the monogenic maxima representation to rotate a discrete color image. The visual content is described by local descriptors (maxima) with explicit $x - y$ coordinates and θ orientation, so the simple modification of these values should rotate the whole image. More precisely, rotating the image by an angle α around a central point (x_c, y_c) implies to modify its monogenic maxima as:

$$\theta \leftarrow \theta + \alpha \tag{43}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} \leftarrow \begin{bmatrix} x_c \\ y_c \end{bmatrix} + \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x - x_c \\ y - y_c \end{bmatrix} \tag{44}$$

The values are not required to be integral thanks to the subpixel accuracy of our maxima representation. Note that if $\theta + \alpha$ is pushed outside $[-\pi/2; \pi/2]$, φ and α_3 must be shifted accordingly (see Appendix). In parallel, the low-frequency component can be trivially rotated with classical cubic spline interpolation. The Fig. 12 shows an example with $\alpha = \pi/8$, with a 5-scale monogenic decomposition. The image on the
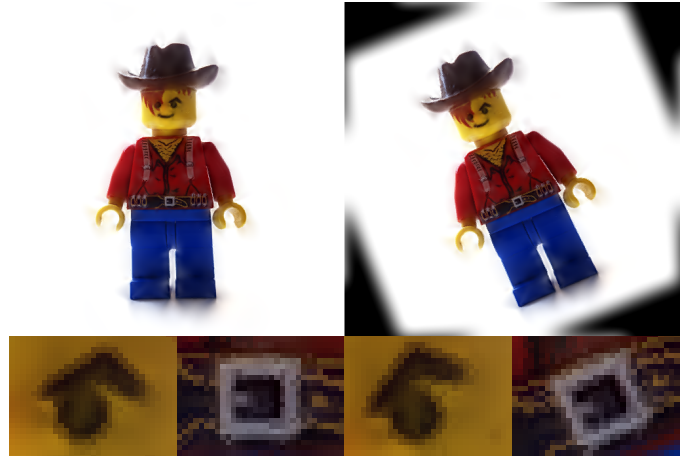
left shows the reconstruction from original maxima, while this on the right shows the reconstruction after having modified (x, y) and θ in every maximum. This maxima-based rotation is visually coherent. Note that the reconstruction artifacts present in the regular reconstruction (dark shades around the hat...) are equally reproduced in the rotated version, which shows the stability of the representation. Here the rotation covariance of the monogenic analysis is a crucial ingredient, and this experiment confirms the stability of our representation with respect to rotation changes.

D. Color denoising

To conclude the paper, let us open a lead towards using monogenic maxima for image denoising. Not competing with the already existing powerful state of the art methods [4], this section aims at illustrating how our representation allows for selecting important information as well as regularizing it. By a lack of space, we only point out a few details of the algorithm.

Wavelet based denoising generally simply discards small wavelet coefficients, considered noisy. With the wavelet maxima representations and the study of cross-scale dependencies, denoising methods were pushed further by suggesting to keep small coefficients being yet related to important contours. In this work, we propose to go one step further by not only discarding the noisy maxima, but also regularizing the remaining ones. This regularization can be driven by the strong assumptions we have about the color monogenic features.

The first step is to discard low amplitude maxima from the L-th scale. This basic thresholding is considered sufficient because the coarsest scale is known to be the least noisy one. Then we span all the maxima from coarser scale $L - 1$ to finest scale 1, and for each one the algorithm is twofold.

First, we decide if the maximum is noisy and has to be discarded. We keep it if it has a parent maximum in the coarser scale, and if their two color axes (α_1, α_2) are similar.

Second, when the maximum has been kept, its features are regularized with respect to its parent. If it is positioned on the same point as its parent, phase and orientation are replaced by
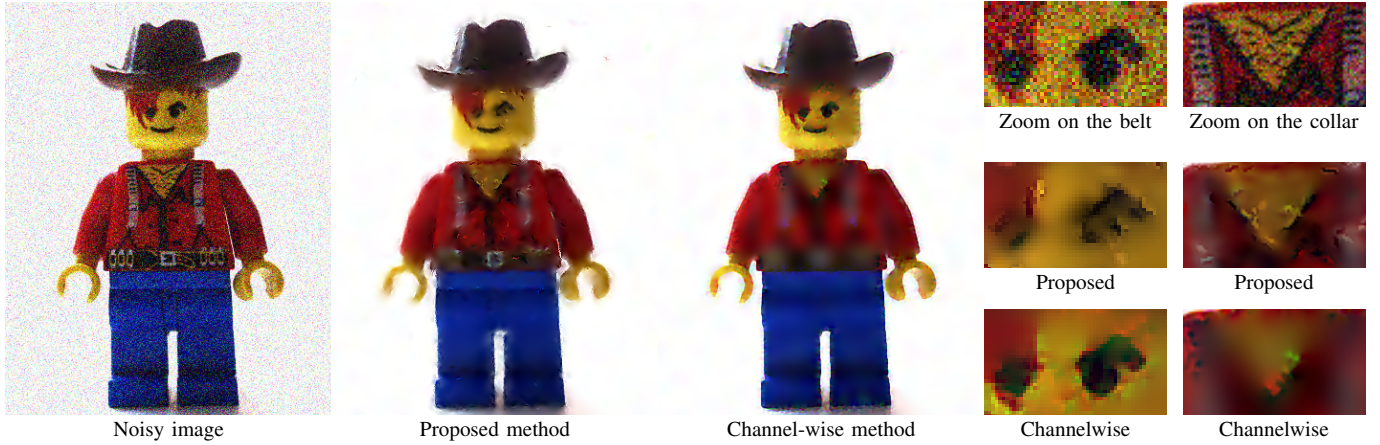
Fig. 13. Monogenic maxima-based denoising compared to channel-wise Mallat-Zhong based method. Right hand side: zoomed areas.

their coarser values. If the parent is relatively shifted ("jump"), the phase is rounded to the closest multiple of $\pi/2$ so as to enhance the corresponding line or edge. In both cases, the 4 ellipse parameters are updated with the values from the parent.

An example is given on the Fig. 13. The image is corrupted by an additive white Gaussian noise with $\sigma = 0.15$, and the denoising algorithm is run on 4 scales. It can be seen that the noisy regions have been highly smoothed while preserving the contours, with efficient color retrieval. There is no false color in spite of a highly non-linear method. This is due to the construction of a true multi-channel extension, instead of the separate channel-wise use of a greyscale tool.

For comparison, we propose a similar experiment based on the channel-wise use of the Mallat-Zhong representation. Since the color features are unavailable, cross-scale linking is based on comparing the orientation feature, and regularization is done on the gradient norm. The Fig. 13 (Third image and right-hand side lower row) shows the expected annoying false colors at many locations. See the orange-green artifacts around the eyes and the collar (zoomed areas). Noisy colors are also present around most contours, contrary to the proposed version which clearly benefits from the regularization of the color axis (α_1, α_2). The marginal method fails again to handle the color information, while the proposed method provides more adapted complementary shape and color features, as well as more advanced maxima connnections.

VI. CONCLUSION

In this article, we expanded the wavelet maxima representation developed by Mallat and Zhong [22]. The key concept is the color monogenic wavelet maxima representation, which includes innovative algorithms for the maxima selection and the interpolation of features. We showed that multiscale color contours can be detected and characterized from the maxima of our color monogenic wavelet transform.

The reconstruction algorithm that is proposed directly in the feature domain recovers a close approximation of the original image. We conducted a number of reconstruction experiments by using this new technique with color images. The new reconstruction method based on monogenic features appears excellent in preserving contours and color information. We

have illustrated that the properties of a color image can be modified by processing its color monogenic features, which confirms the here proposed "color Marr conjecture": It is possible to reconstruct a color image based on its contour information, thanks to the monogenic maxima features.

Future work includes studying the cross-scale evolution of features, which could bring substantial improvements in terms of sparsity and for any applicative work. We are also working on color image indexing/retrieval application.

APPENDIX
UNWRAPPING OF ANGLES

The elliptical monogenic representation contains several *circular* features, implying the handling of angle wrapping, which is explained here for the case of 3-valued 2d signals.

A. Orientation

The first wrapping effect comes from the spatial orientation θ. Its "double-angle" computation from the structure tensor makes no difference between opposite directions $\boldsymbol{u}$ and $-\boldsymbol{u}$, and wraps θ in $[-\pi/2; \pi/2]$. Unstable "π-shifts" may occur in areas where $\theta \approx \pm\pi/2$. When comparisons have to be made in such a neighbourhood, the "opposite" θ samples must be π-shifted back towards the central reference sample, which unwraps them to $[-\pi; \pi]$. This artificial shift turns the whole analysis in the opposite direction, so that the eq. (22) becomes its conjugate value (since $\mathcal{R}_{-\boldsymbol{u}} = -\mathcal{R}_{\boldsymbol{u}}$). The ellipse parameters must then be updated according to:

$$(\boldsymbol{s} + \mathbf{j}\mathcal{R}_{\boldsymbol{u}}\boldsymbol{s}) \leftarrow (\boldsymbol{s} + \mathbf{j}\mathcal{R}_{-\boldsymbol{u}}\boldsymbol{s}) \quad \Leftrightarrow \quad \begin{aligned} \theta &\leftarrow \theta + \pi \\ \varphi &\leftarrow -\varphi \\ \alpha_3 &\leftarrow \alpha_3 + \pi \end{aligned} \quad (45)$$

This corresponds to inverting the phase, and tilting the ellipse by $180°$ around its major axis (so that $\boldsymbol{e}_\perp \leftarrow -\boldsymbol{e}_\perp$).

B. Ellipse angles

Once the orientation is unwrapped and the corresponding ellipse parameters updated accordingly, we have:

$$
\begin{array}{ccccc}
(& \varphi & \alpha_1 & \alpha_2 & \alpha_3 &) \\
\equiv (& \varphi & \alpha_1+\pi & -\alpha_2+\pi & \alpha_3+\pi &) \\
\equiv (& \varphi+\pi & \alpha_1+\pi & -\alpha_2 & -\alpha_3 &) \\
\equiv (& \varphi+\pi & \alpha_1 & \alpha_2+\pi & \pi-\alpha_3 &)
\end{array} \quad (46)
$$

for which $s+\mathbf{j}\mathcal{R}_u s$ remains identical with respect to eq. (22). In [27], the color axis is arbitrarily chosen to lie in the half space where α_1 and α_2 are in $[-\pi/2;\pi/2]$. When computing the features from the Cartesian coefficients, the above equivalences are used to update φ and α_3.

For the computation of the local frequency ν, we need the phase φ to be properly unwrapped, while the other angles - assumed locally constant - are "free of π-shifts". We propose the following procedure:

- Get the reference angle values from the central samples;
- If $\theta_{\text{ref}} \approx \pm\pi/2$, detect π-shifts in the neighbourhood and update $(\theta, \varphi, \alpha_3)$;
- Do the same for α_1 and α_2 by choosing the proper line from eq. (46), and update φ and α_3.

Note that $\alpha_3 \in [-\pi;\pi]$ is not wrapped. The last step is to wrap back the neighbouring phase samples φ in $[\varphi_{\text{ref}} - \pi; \varphi_{\text{ref}} + \pi]$ by classical 2π-shifts, so that they lie around the central value. The phase samples are now comparable one another and can be differentiated for local frequency computation.

REFERENCES

[1] A. C. Bovik. *Handbook of Image and Video Processing*. Academic Press, May 2000.

[2] T. Bülow. Hypercomplex spectral signal representation for the processing and analysis of images. *Thesis*, Aug. 1999.

[3] J. F. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8:679–698, 1986.

[4] P. Chatterjee and P. Milanfar. Is denoising dead? *IEEE Trans. Image Process.*, 19(4):895–911, 2010.

[5] N. Chenouard and M. Unser. 3D steerable wavelets in practice. *IEEE Transactions on Image Processing*, 21(11):4522–4533, November 2012.

[6] S. Di Zenzo. A note on the gradient of a multi-image. *Computer Vision, Graphics, and Image Processing*, 33(1):116–125, 1986.

[7] M. Felsberg. Low-level image processing with the structure multivector. *Thesis*, 2002.

[8] O. Fleischmann, L. Wietzke, and G. Sommer. Image analysis by conformal embedding. *J. of Math. Im. and Vis.*, 40(3):305–325, 2011.

[9] G. H. Granlund and H. Knutsson. *Signal processing for computer vision*. Kluwer, 1995.

[10] S. L. Hahn. Multidimensional complex signals with single-orthant spectra. *Proc. IEEE*, 80(8):1287–1300, 1992.

[11] H. Haußecker and B. Jähne. A tensor approach for local structure analysis in multi-dimensional images. In *Proc. 3D Image Analysis and Synthesis*, pages 18–19, Nov. 1996.

[12] S. Held, M. Storath, P. Massopust, and B. Forster. Steerable wavelet frames based on the riesz transform. *IEEE Trans. Image Process.*, 19(3):653–667, 2010.

[13] B. Jähne. *Digital Image Processing 6th Edition*. Springer, Berlin, 2005.

[14] B. J. Janssen and R. Duits. Linear image reconstruction by sobolev norms on the bounded domain. *Int'l J. of Comp. Vis.*, 84(2), 2009.

[15] U. Köthe and M. Felsberg. Riesz-transforms versus derivatives: On the relationship between the boundary tensor and the energy tensor. In J. Weickert R. Kimmel, N. Sochen, editor, *Proc. Scale-Space, LNCS 3459*, pages 179–191. Springer, 2005.

[16] K. G. Larkin, D. Bone, and M. A. Oldfield. Natural demodulation of two-dimensional fringe patterns: I. general background to the spiral phase quadrature transform. *J. Opt. Soc. Am.*, 18 (8):1862–1870, 2001.

[17] K. Lesiński. Source code for color multiscale SSIM image quality measure. https://kornel.ski/dssim, 2017.

[18] M. Lillholm, M. Nielsen, and L.D. Griffin. Feature-based image analysis. *Int'l J. of Comp. Vis.*, 52(73), 2003.

[19] J. M. Lilly. Modulated oscillations in three dimensions. *IEEE Trans. Signal Process.*, 59(12):5930–5943, 2011.

[20] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *Int'l J. of Comp. Vis.*, 60(2):91–110, 2004.

[21] S. Mallat. *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. Academic Press, 3 edition, Dec. 2008.

[22] S. Mallat and S. Zhong. Characterization of signals from multiscale edges. *IEEE Trans. Pattern Anal. Mach. Intell.*, 14(7):710–732, 1992.

[23] D. Marr. *Vision*. W. H. Freeman and Company, New York, 1982.

[24] M. N. Nabighian. Toward a three-dimensional automatic interpretation of potential field data via generalized Hilbert transforms: Fundamental relations. *Geophysics*, 49(6):780–786, 1984.

[25] P. Pad, V. Uhlmann, and M. Unser. Maximally localized radial profiles for tight steerable wavelet frames (preprint). *IEEE Trans. Image Process.*, 2016.

[26] J. Portilla and E. P. Simoncelli. A parametric texture model based on joint statistics of complex wavelet coefficients. *Int'l J. of Comp. Vis.*, 40(49), 2000.

[27] R. Soulard and P. Carré. Elliptical monogenic wavelets for the analysis and processing of color images. *IEEE Trans. Signal Process.*, 66(6), 2015.

[28] D. Tschumperlé and R. Deriche. Vector-valued image regularization with PDEs: A common framework for different applications. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(4):506–517, 2005.

[29] M. Unser and N. Chenouard. A unifying parametric framework for 2D steerable wavelet transforms. *SIAM J. Imaging Sci.*, 6(1), 2013.

[30] M. Unser, D. Sage, and D. Van De Ville. Multiresolution monogenic signal analysis using the riesz-laplace wavelet transform. *IEEE Trans. Image Process.*, 18(11):2402–2418, 2009.

[31] D. Van De Ville and M. Unser. Complex wavelet bases, steerability, and the marr-like pyramid. *IEEE Trans. Image Process.*, 17(11), 2008.

[32] L. Wietzke and G. Sommer. The signal multi-vector. *Journal of Mathematical Imaging and Vision*, 37:132150, June 2010.

Dr Raphaël Soulard received the PhD degree in 2012 from the University of Poitiers, France. He is currently a postdoctoral research fellow at the XLIM laboratory, France. He received the national prize "Blanc-Lapierre" granted by the SEE (Society of the electricity, the electronics, and the information technologies) for his master's thesis in 2009. His research interests include wavelets and filterbanks, analytic/monogenic representation methods, extension of signal and image processing tools for multi-dimensional/multivalued signals like color images.

Prof. Philippe Carré received the Engineer degree in Computer Engineering from Compiegne University of Technology, France, in 1995, and the PhD degree in Signal Processing and Computer Science from the University of Poitiers, France, in 2000 for his work on the definition of discrete atomic Time-Frequency representations. He is presently Professor at the XLIM laboratory, France. His interests include signal and image processing, wavelets, Time-Frequency and multiscale decompositions theory and their applications to the fields of denoising, segmentation and watermarking.