COLOR MONOGENIC WAVELETS FOR IMAGE ANALYSIS

Raphaël Soulard and Philippe Carré

Xlim-SIC laboratory, University of Poitiers, France

raphael.soulard@univ-poitiers.fr carre@sic.univ-poitiers.fr

ABSTRACT

We define a color monogenic wavelet transform. This is based on the recent grayscale monogenic wavelet transform and a non-marginal extension to color signals. To our knowledge, wavelet based color image processing schemes have always been made by using a grayscale tool separately on color channels. This may have some unexpected effect on colors because those marginal schemes are not necessarily justified. Here we propose a definition that considers a color (vector) image right at the beginning of the mathematical definition and so brings an actual color wavelet transform - which has not been done so far to our knowledge. This so provides a promising multiresolution color geometric analysis of images.

Index Terms— Color Wavelets, Analytic, Monogenic, Wavelet transforms, Image analysis

1. INTRODUCTION

Wavelets have been widely used for handling images for more than 20 years. It seems that the human visual system sees images through different channels related to particular frequency bands and directions; and wavelets provide such decompositions. Since 2001, the *analytic signal* and its 2D generalizations have brought a great improvement to wavelets [1, 2, 3] by a natural embedding of an AM/FM analysis in the subband coding framework. This yields an efficient representation of geometric structures in grayscale images thanks to a *local phase* carrying geometric information complementary to an *amplitude envelope* having good invariance properties. So it codes the signal in a more coherent way than standard wavelets. The last and seemingly most appropriate proposition [3] of *analytic wavelets* for image analysis is based on the *monogenic signal* [4].

In parallel a *color monogenic signal* was proposed [5] as a mathematical extension of the monogenic signal; paving the way to non-marginal color tools especially by using geometric algebra and above all by considering a color signal right at the foundation of the mathematical construction.

We define here a *color monogenic wavelet transform* that extends the monogenic wavelets of [3] to color. These *analytic wavelets* are defined for color 2D signals (images) and avoid the classical pitfall of marginal processing (grayscale tool used separately on color channels) by relying on a sound mathematical definition. We may so expect to handle coherent information of multiresolution color geometric structure; which would make easier any wavelet based color image processing. To our knowledge color wavelets have not been proposed so far.

We first give a technical study of analytic signal/wavelets with the intent to popularize them since they rely on non-trivial concepts of geometric algebra, complex/harmonic analysis, as well as non-separable wavelet frames. Then we describe our *color monogenic wavelet transform*.

This work is part of the French ANR project CAIMAN.

Notations :

2-vector coordinates : $\boldsymbol{x} = (x, y)$, $\boldsymbol{\omega} = (\omega_1, \omega_2) \in \mathbb{R}^2$; $\boldsymbol{k} \in \mathbb{Z}^2$
Euclidean norm : $\|\boldsymbol{x}\| = \sqrt{x^2 + y^2}$
Complex imaginary number : $\boldsymbol{j} \in \mathbb{C}$
Argument of a complex number : $\arg$
Convolution symbol : $*$
Fourier transform : $\mathcal{F}$

2. ANALYTIC SIGNAL AND 2D GENERALIZATION

An *analytic signal* s_A is a multi-component signal associated to a real signal s to analyze. The definition is well known in the 1D case where $s_A(t) = s(t) + \boldsymbol{j}\ (h * s)(t)$ is the complex signal made of s and its Hilbert transform (with $h(t) = \frac{1}{\pi t}$).

The polar form of the 1D analytic signal provides an AM/FM representation of s with $|s_A|$ being the *amplitude envelope* and $\varphi = \arg(s_A)$ the *instantaneous phase*. This classical tool can be found in many signal processing books and is used in communications for example.

Interestingly we can also interpret the phase in terms of signal shape *i.e.* there is a direct link between the angle φ and the *local structure* of s. Such a link between a 2D phase and local geometric structures of images would be very attractive in image processing. That is why there were several attempts to generalize it for 2D signals; and among them the *monogenic signal* [4] seems the most advanced since it is rotation invariant.

The Monogenic Signal

Without going beyond strictly needed details we here review the key points of the fundamental construction of the monogenic signal; which will be necessary to understand the color extension.

The definition of the 1D case given above can be interpreted in terms of *signal processing* : the Hilbert transform makes a "pure $\frac{\pi}{2}$-dephasing". But such a dephasing is not straightforward to define in 2D (same issue with many 1D signal tools) so let us look at the equivalent *complex analysis* definition of the 1D analytic signal. It says that s_A is the *holomorphic extension* of s restricted to the real line. But complex algebra is impeding for generalizations to higher dimensions. To bypass this limitation we can see a *holomorphic function* like a 2D *harmonic field* that is an equivalent *harmonic analysis* concept involving the 2D Laplace equation $\Delta f = 0$. It so can be generalized within the framework of 3D harmonic fields by using the 3D Laplace operator $\Delta_3 = \left(\frac{\delta}{\delta x} + \frac{\delta}{\delta y} + \frac{\delta}{\delta z}\right)$. The whole generalization relies on this natural choice and remaining points are analogous to the 1D case (see [4] for more details). Note that in

Felsberg's thesis this construction is expressed in terms of *geometric algebra* but here we avoided it for simplicity's sake. Finally the 2D monogenic signal s_A associated to s is the 3-vector valued signal :

$$s_A(\boldsymbol{x}) = \begin{bmatrix} s(\boldsymbol{x}) \\ s_{r1}(\boldsymbol{x}) = \frac{x}{2\pi\|\boldsymbol{x}\|^3} * s(\boldsymbol{x}) \\ s_{r2}(\boldsymbol{x}) = \frac{y}{2\pi\|\boldsymbol{x}\|^3} * s(\boldsymbol{x}) \end{bmatrix} \quad (1)$$

Where s_{r1} and s_{r2} are analogous to the imaginary part of the complex 1D analytic signal. Interestingly, this construction reveals the two components of a Riesz transform (that is convolution of a function by the two kernels $\frac{x}{2\pi\|\boldsymbol{x}\|^3}$ and $\frac{y}{2\pi\|\boldsymbol{x}\|^3}$) in the same way that the 1D case exhibits a Hilbert transform. Note that we get back to a *signal processing* interpretation since the Riesz transform can also be viewed like a pure 2D dephasing. In the end, by focusing on the *complex analysis* definition of the analytic signal we end up with a convincing generalization of the Hilbert transform.

Now recall that the motivation to build 2D analytic signals arises from the strong link existing between the phase and the geometric structure. To define the 2D phase related to the Riesz transform the actual monogenic signal must be expressed in spherical coordinates that yield the following amplitude envelope and 2-angle phase :

$$\begin{array}{ll|l} \text{Amplitude :} & A = \sqrt{s^2 + s_{r1}^2 + s_{r2}^2} & s = A\cos\varphi \\ \text{Orientation :} & \theta = \arg\left(s_{r1} + \boldsymbol{j}\ s_{r2}\right) & s_{r1} = A\sin\varphi\cos\theta \\ \text{1D Phase :} & \varphi = \arccos\left(\frac{s}{A}\right) & s_{r2} = A\sin\varphi\sin\theta \end{array} \quad (2)$$

Felsberg shows a direct link between the angles θ and φ and the geometric local structure of s. The signal is so expressed like an "A-strong" 1D structure with orientation θ. φ is analogous to the 1D local phase and indicates if the structure is rather a line or an edge. A direct drawback is that intrinsically 2D structures are not handled. Yet this tool found many applications in image analysis from contour detection to motion estimation (see [3] and references therein p. 1).

From a *signal processing* viewpoint the AM/FM representation provided by an analytic signal is accordingly well suited for narrow-band signals. That is why it seems natural to embed it in a wavelet transform that performs subband decomposition. We now present the monogenic wavelet analysis proposed in [3].

3. MONOGENIC WAVELETS

So far there is one proposition of computable monogenic wavelets in the literature [3]. It provides 3-vector valued monogenic subbands consisting of a rotation-covariant *magnitude* and this new 2D *phase*. This representation - specially defined for 2D signals - is a great theoretic improvement of the complex and quaternion wavelets [1, 2]; as well as the monogenic signal itself is an improvement of its complex and quaternion counterparts.

The proposition of [3] consists of one real-valued "primary" wavelet transform in parallel with an associated complex-valued wavelet transform. Both transforms are linked each other by the Riesz transform so they carry out a multiresolution monogenic analysis. We end up with 3-vector coefficients forming subbands that are monogenic.

3.1. Primary transform

The primary transform is real-valued and relies on a dyadic pyramid decomposition tied to a wavelet frame. Only one 2D wavelet is needed and the dyadic downsampling is done only at the low frequency branch; leading to a redundancy of 4:3. The scaling function φ_γ and mother wavelet ψ are defined in the Fourier domain :

$$\varphi_\gamma \overset{\mathcal{F}}{\longleftrightarrow} \frac{\left(4(\sin^2\frac{\omega_1}{2} + \sin^2\frac{\omega_2}{2}) - \frac{8}{3}\sin^2\frac{\omega_1}{2}\sin^2\frac{\omega_2}{2}\right)^{\frac{\gamma}{2}}}{\|\boldsymbol{\omega}\|^\gamma} \quad (3)$$

$$\psi(\boldsymbol{x}) = (-\Delta)^{\frac{\gamma}{2}}\ \varphi_{2\gamma}(2\boldsymbol{x}) \quad (4)$$

Note that φ_γ is a cardinal polyharmonic spline of order γ and spans the space of those splines with its integer shifts. It also generates - as a scaling function - a valid multiresolution analysis. In this paper we set $\gamma = 1.00001$ that gave satisfying experimental results.

This particular construction is made by an extension of a wavelet basis (non-redundant) related to a critically-sampled filterbank. This extension to a wavelet frame (redundant) adds some degrees of freedom used by the authors to tune the involved functions. In addition a specific *subband regression* algorithm is used at the synthesis side. The construction is fully described in [6].

3.2. The monogenic transform

The second "Riesz part" transform is a complex-valued extension of the primary one. We define the associated complex-valued wavelet by including the Riesz components :

$$\psi' = -\left(\frac{x}{2\pi\|\boldsymbol{x}\|^3} * \psi(\boldsymbol{x})\right) + \boldsymbol{j}\left(\frac{y}{2\pi\|\boldsymbol{x}\|^3} * \psi(\boldsymbol{x})\right) \quad (5)$$

It can be shown that it generates a valid wavelet basis and that it can be extended to the pyramid described above. The joint consideration of both transforms form monogenic subbands from which can be extracted the amplitude and phase for an overall redundancy of 4:1.

So far no applications of the monogenic wavelets have been proposed. In [3] a demonstration of AM/FM analysis is done with fine orientation estimation and gives very good results in terms of coherency and accuracy. Accordingly this tool may be rather used for analysis tasks than processing.

Motivated by the powerful analysis provided by the monogenic wavelet transform we propose now to extend it for color images.

4. COLOR MONOGENIC WAVELETS

We define here our proposition that combines a fundamental generalization of the monogenic signal to color with the monogenic wavelets described above. The challenge is to avoid the classical *marginal* definition that would be applying a *grayscale* monogenic transform on each of the three color channels of a color image. We believe that the monogenic signal has a favorable theoretical framework for a color extension and this is why we propose to start from this particular wavelet transform rather than from a more classical one.

The color generalization of the monogenic signal is expressed within the *geometric algebra* framework. This algebra is very general and embeds the complex and quaternion as subalgebras. Its elements are "multivectors" naturally linked with various geometric entities. The use of this fundamental tool is gaining popularity in the literature because it allows rewriting sophisticated concepts with simpler algebraic expressions and so paves the way to innovative ideas and generalizations in many fields.

For simplicity's sake and since anyway we would not have enough space to present the fundamentals of geometric algebra we here express the construction in classical terms; as we already did section 2. Yet we may sometimes point out some necessary specific mechanisms but we refer the reader to [4, 5] for further details.

4.1. The Color Monogenic Signal

Starting from Felsberg's approach that is originally expressed in the geometric algebra of $\mathbb{R}^3$; the extension proposed in [5] is written in the geometric algebra of $\mathbb{R}^5$ for 3-vector valued 2D signals of the form (s_R, s_G, s_B). By simply increasing the dimensions we can embed each color channel along a different axis and the original equation from Felsberg involving a 3D Laplace operator can be generalized in 5D with $\Delta_5 = \left(\frac{\delta}{\delta x_1} + \frac{\delta}{\delta x_2} + \frac{\delta}{\delta x_3} + \frac{\delta}{\delta x_4} + \frac{\delta}{\delta x_5}\right)$.

Then the system can be simplified by splitting it into three systems with a 3D Laplace equation, reducing to applying Felsberg's condition to each color channel. At this stage appears the importance of *geometric algebra* since an algebraic simplification between vectors leads to a 5-vector color monogenic signal that is non-marginal. Instead of naively applying the Riesz transform to each color channel, this fundamental generalization carries out the following color monogenic signal : $s_A = (s_R, s_G, s_B, s_{r1}, s_{r2})$ where s_{r1} and s_{r2} are the Riesz transform applied to $s_R+s_G+s_B$.Note that this simple sum of color channels is obtained by calculation. This is coherent with the idea that s_{r1} and s_{r2} carry structural information - independent of color.

Now the color extension of Felsberg's monogenic signal is defined let us construct the color extension of the monogenic wavelets.

4.2. The Color Monogenic Wavelet Transform

We can now define a wavelet transform whose subbands are color monogenic signals. The goal is to obtain vector coefficients of the form $(c_R, c_G, c_B, c_{r1}, c_{r2})$ such that $c_{r1} = \frac{x}{2\pi\|\boldsymbol{x}\|^3} * (c_R + c_G + c_B)$ and $c_{r2} = \frac{y}{2\pi\|\boldsymbol{x}\|^3} * (c_R + c_G + c_B)$.

It turns out that we can very simply use the transforms presented above by applying the *primary* one on each color channel and the *Riesz part* on the sum of the three. The five related color wavelets illustrated Fig. 1 and forming one color monogenic wavelet ψ_A are :

$$\psi_R = \begin{pmatrix} \psi \\ 0 \\ 0 \end{pmatrix} \quad \psi_G = \begin{pmatrix} 0 \\ \psi \\ 0 \end{pmatrix} \quad \psi_B = \begin{pmatrix} 0 \\ 0 \\ \psi \end{pmatrix} \tag{6}$$

$$\psi_{r1} = \begin{pmatrix} \frac{x}{2\pi\|\boldsymbol{x}\|^3} * \psi \\ \frac{x}{2\pi\|\boldsymbol{x}\|^3} * \psi \\ \frac{x}{2\pi\|\boldsymbol{x}\|^3} * \psi \end{pmatrix} \quad \psi_{r2} = \begin{pmatrix} \frac{y}{2\pi\|\boldsymbol{x}\|^3} * \psi \\ \frac{y}{2\pi\|\boldsymbol{x}\|^3} * \psi \\ \frac{y}{2\pi\|\boldsymbol{x}\|^3} * \psi \end{pmatrix} \tag{7}$$

$$\psi_A = (\psi_R\ ,\ \psi_G\ ,\ \psi_B\ ,\ \psi_{r1}\ ,\ \psi_{r2}) \tag{8}$$

We then get 5-vector coefficients verifying our conditions and so forming a color monogenic wavelet transform. The associated decomposition is described by the diagram Fig. 2. This provides a multiresolution color monogenic analysis made of a 5-vector valued pyramid transform.The construction can be viewed like a vector extension of the scalar wavelet ψ. The 5 decompositions of two images are shown Fig. 3 from left to right. Each one consists of 4 juxtaposed image-like subbands resulting from 3-level decomposition.

Let us look at the first 3 graymaps. These are the 3 primary transforms c_R, c_G and c_B where white (resp. black) pixels are high positive (resp. negative) values. Note that our transform is non-separable and so provides at each scale only *one* subband related to *all* orientations. We are not subjected to the arbitrarily separated horizontal, vertical and diagonal analyses of usual wavelets. This advantage is even greater in color. Whereas marginal separable transforms show 3 arbitrary orientations within each color channel - which is not easily interpretable - the color monogenic wavelet transform provides a more compact *energy* representation of the color image content regardless of the local orientation. The color information is well separated through c_R, c_G and c_B : see that blue contours of first image are present only in c_B. And in each of the 3 decompositions it is clear that every orientation is equally represented all along the round contours. That is different from separable transform that privileges particular directions. The multiresolution framework makes the horizontal blue low frequency structure of second image be coded mainly in the third scale of c_B.

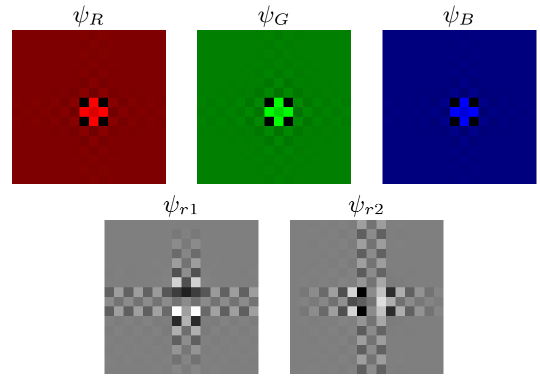

Fig. 1. Space representation of the 5 color wavelets.

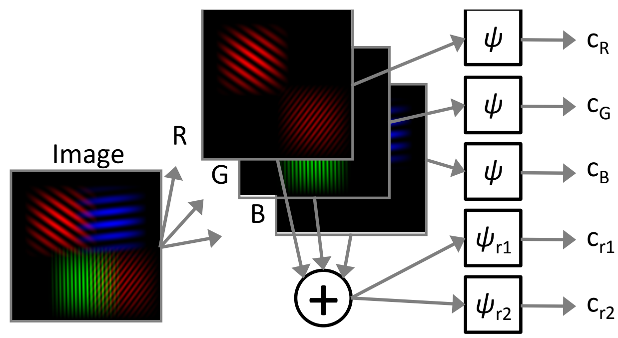

Fig. 2. Color MWT scheme. Each color channel is analyzed with primary wavelet transform ψ; and the sum "$R+G+B$" is analyzed with the "Riesz part" wavelet transform (ψ_{r1} and ψ_{r2} blocs).

But the directional analysis is not lost thanks to the Riesz part that completes this representation. Now look at the "2-in-1" last decomposition forming the Riesz part. It is displayed in one color map where the geometric energy $\sqrt{c_{r1}^2 + c_{r2}^2}$ is encoded into the intensity (with respect to the well known HSV color space) and the orientation $\arg\left(c_{r1} + \boldsymbol{j} c_{r2}\right)(\pi)$ is encoded in the hue (*e.g.* red is for $\{0, \pi\}$ and cyan is for $\pm\frac{\pi}{2}$). This way of displaying the Riesz part well reveals the provided geometric analysis of the image.

The Riesz part makes a precise analysis that is *local* both in space and scale. If there is a local color geometric structure in the image at a certain scale the Riesz part exhibits a high intensity in the corresponding position and subband. This is completed with an orientation analysis (hue) of the underlying structure. For instance a horizontal (resp. vertical) structure in the image will be coded by a cyan (resp. red) intense point in the corresponding subband. The orientation analysis is strikingly coherent and accurate. See for example that color structures with constant orientation (second image)

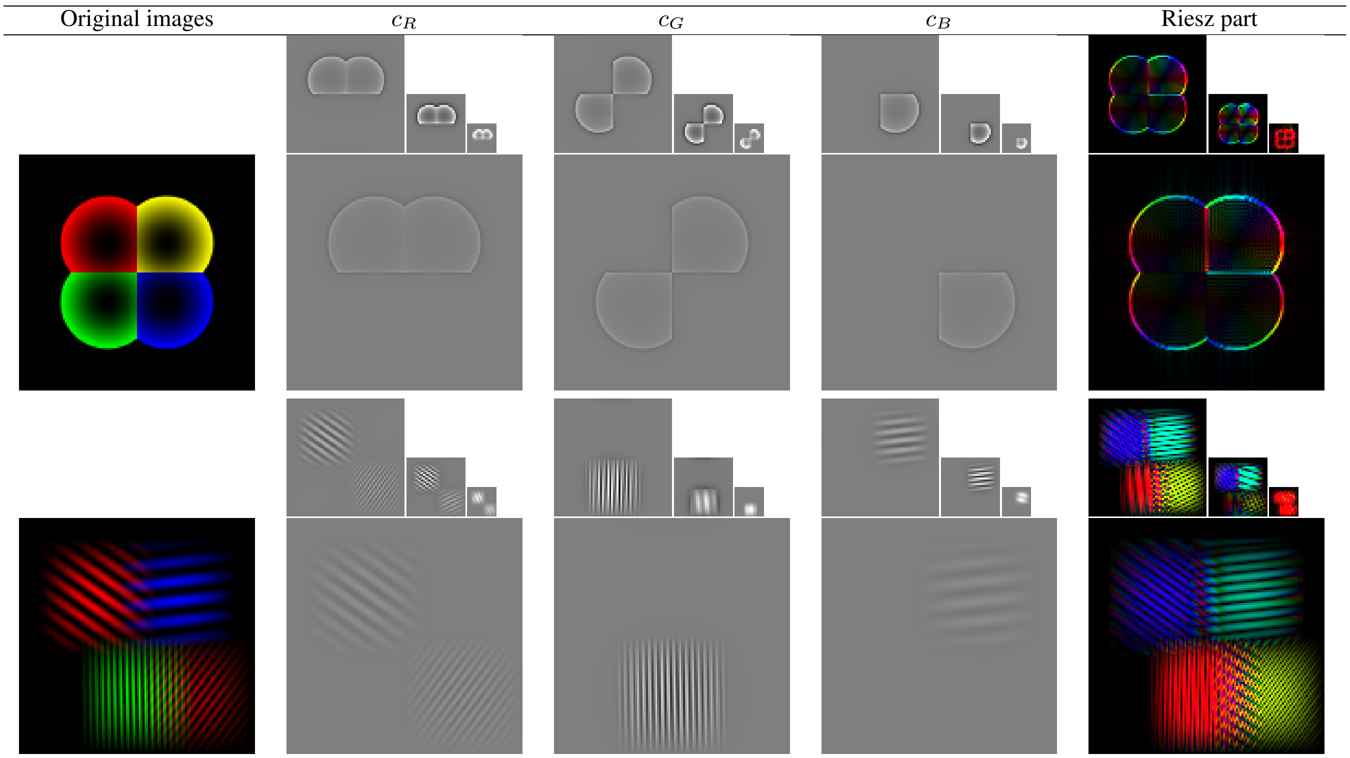

Fig. 3. Color MWT of images. The two components of the *Riesz part* are displayed in the same graphic with the magnitude of $c_{r1} + \boldsymbol{j} c_{r2}$ encoded in the intensity and argument (local orientation) encoded in the hue.

exhibit a *constant* hue in the Riesz part over the whole structure.

Note that low intensity corresponds to "no structure" *i.e.* where the image has no geometric information. It is coherent not to display the orientation (low intensity makes the hue invisible) for these coefficients since this data has no sense in those cases.

In short the color and geometric information of the image are well separated from each other and the orientation analysis is very accurate. In addition the invariance properties of the primary and Riesz wavelet transforms are kept in the color extension for a slight overall redundancy of $20:9 \approx 2.2$.

5. CONCLUSION

We define a color extension of the recent monogenic wavelet transform proposed in [3]. This extension is non-marginal since it takes care of considering a vector signal at the very beginning of the fundamental construction and leads to a definition basically different from the marginal approach. This has not been proposed before to our knowledge. The use of non-separable wavelets joint with the monogenic framework allows for a good orientation analysis well separated from the color information. This color transform can be a great color image analysis tool thanks to this good separation of information through various data. We are currently working on color denoising by using this transform.

Although it is not marginal the color generalization has a marginal style since it reduces to apply the Riesz transform on the intensity of the image. So the geometric analysis is done without considering the color information and it would be much more attractive to have a complete representation of the color monogenic signal into magnitude and phase(s) with color/geometric interpretation.

Another limit is the restriction to analysis applications. The redundancy of the grayscale transform makes it not well suited *a priori* to image compression.

6. REFERENCES

[1] I. W. Selesnick, R. G. Baraniuk, and N. G. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Processing Magazine [123] November*, 2005.

[2] W.L. Chan, H.H. Choi, and R.G. Baraniuk, "Coherent multiscale image processing using dual-tree quaternion wavelets," *IEEE Transactions on Image Processing*, vol. 17, no. 7, pp. 1069–1082, July 2008.

[3] M. Unser, D. Sage, and D. van de Ville, "Multiresolution monogenic signal analysis using the riesz-laplace wavelet transform," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2402–2418, November 2009.

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

[5] G. Demarcq, L. Mascarilla, and P. Courtellemont, "The color monogenic signal: a new framework for color image processing.," *IEEE Proceedings of the International Conference on Image Processing (ICIP)*, 2009.

[6] 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*, vol. 17, no. 11, pp. 1–13, November 2008.