Elliptical monogenic representation of color images and local frequency analysis

Raphaël Soulard, Philippe Carré

To cite this version:
ELLIPSOIDAL MONOGENIC REPRESENTATION OF COLOR IMAGES AND LOCAL FREQUENCY ANALYSIS

Raphaël Soulard and Philippe Carré

XLIM-SIC Department UMR CNRS 7252, University of Poitiers, France.

ABSTRACT

We define a new color extension for the monogenic representation of images by using an ellipsoidal tri-valued oscillation model jointly with the vector structure tensor formalism. The proposed method provides a rich local colorimetric and geometric analysis, in particular a color phase concept, which can be computed by a numerically stable algorithm. This representation is finally used to estimate the local frequency of color images.

Index Terms— wavelet transform, color, analytic signal, monogenic signal, monogenic wavelets, local frequency

1. THE MONOGENIC REPRESENTATION

As a 2D extension of the analytic signal, the monogenic signal [1] provides a 2D AM/FM representation of gray-scale images. Thanks to its rotation invariance property and its underlying signal shape analysis; the monogenic representation has been used for numerous image applications, including analysis of textures [2], contours [1], and motion or stereo disparity [3, 4]; segmentation [5, 6] etc.

Given a 2D real signal \( s(x) \), the associated monogenic signal \( s_M(x) \) is the complex-valued Riesz transform of \( s \):

\[
\begin{align*}
\{R s(x)\} & = p.v. \int \frac{t_1 + j t_2}{2 \pi |\tau|} s(x - \tau) \ d\tau = t_1 \omega_2 - j\omega_1 \frac{\omega_2 - j\omega_1}{\|\omega\|} \hat{s}(\omega) \\
\end{align*}
\]

(2)

Parameters of \( a \) are directly obtained from separated Fourier analyses to form the complex valued 3-vector

\[
\begin{align*}
A & = \sqrt{s^2 + |R s|^2} \\
\theta & = \arg\{R s\} \in [-\pi; \pi[ \\
\varphi & = \arg\{s + j|R s|\} \in [0; \pi] \\
\end{align*}
\]

(3)

This representation is equivalent to the pointwise estimation of the underlying plane wave that locally resembles the signal. The signal is modeled by an “A-strong” plane wave with orientation \( \theta \), and phase-shifted according to \( \varphi \) - which explicitly indicates if the structure is rather a line or an edge.

The goal of this paper is to study the generalization of this tool for color images. Analyzing color data is essential for a lot of applications. However, due to a lack of definition in multidimensional signal processing, color images are most often handled by a suboptimal use of scalar tools, either by only considering luminance, or by working separately on every color channel (“marginal” scheme [9]). Note that a color monogenic signal is defined in [10] by generalizing the Cauchy-Riemann equations within Geometric algebra.

However, we showed in [11] that this definition does not provide satisfactory physical interpretation of the obtained features. To go one step further, this paper handles the issue from the physical point of view.

We first define the 1D color elliptical model in section 2 in order to well understand what a “color oscillation” is. Then we extend it to our 2D color monogenic representation using the structure tensor formalism in section 3; and finally exploit it to estimate local frequency of color images in section 4.

2. THE COLOR 1D OSCILLATION

This section aims at redefining the classical amplitude and phase concepts for color signals. Our starting point is the sinusoidal signal model, building block of the Fourier analysis. It is well known that a Fourier coefficient is made of an amplitude and a phase value, directly giving the two parameters of a sinusoidal oscillation. We must then define what a color oscillation is.

Regarding the numerous generalized Fourier transforms in the mathematical literature e.g. [12, 13], it seems reasonable to us to consider independent Fourier analyzers at each color channel in the first place. Therefore, our proposition is to define the color oscillation by:

\[
\begin{align*}
A(t) & = \begin{bmatrix} a_R(t) \\ a_G(t) \\ a_B(t) \end{bmatrix} = \begin{bmatrix} A_R \cos(\omega t + \varphi^R) \\ A_G \cos(\omega t + \varphi^G) \\ A_B \cos(\omega t + \varphi^B) \end{bmatrix} \\
\end{align*}
\]

(4)

Parameters of \( a \) are directly obtained from separated Fourier analyses to form the complex valued 3-vector

\[
\begin{align*}
\hat{a} & = \begin{bmatrix} A_R e^{j\varphi^R} \\ A_G e^{j\varphi^G} \\ A_B e^{j\varphi^B} \end{bmatrix}^T \\
\end{align*}
\]

(5)

The problem is that those parameters do not allow any intuitive interpretation of how the oscillation is modulated by them. This is why a non-marginal construction must be carried out by some proper conversion of the 3 amplitudes and 3 phases into clearly identified geometric and colorimetric features. In particular, we may prefer to have an unique amplitude, and a physical color phase concept.

To this end, we must view \( a(t) \) like a point-trajectory (a “color path”) within the 3D color space. Figures 1a and 1b show an example of color sinusoid, and its representation in the color space as a point-trajectory. We also plot the corresponding 2D wave in the image world (Fig. 1d) which will be treated later.

What follows is based on the work in [14] that analyzes trajectories of particles in the field of seismology, which is perfectly analogous to our point. We here use it while adding the “color” concept.

It can be verified [14] that \( a(t) \) draws an ellipse within the color space - as illustrated on figures 1a and 1b with arbitrary values of amplitudes and phases. The amplitude concept can intuitively be
extended in the color case to the ellipse’s size. The phase concept can also be handled, by defining it as the time-relative position of the oscillication’s maximum. In the general case, the ellipse’s apogees can be viewed like the sinusoid’s maxima. So the phase \( \varphi \) can be defined by the time-relative position of the ellipse’s apogee, as illustrated in figure 1c. The remaining ellipse parameters are expected to convey the angle \( \alpha \), and is based on the following rotation:

\[
\begin{bmatrix}
\hat{a}_1 \\
\hat{a}_2 \\
0
\end{bmatrix} = \begin{bmatrix}
\cos \alpha_1 & -\cos \alpha_2 & 0 \\
\sin \alpha_1 & \cos \alpha_1 & \sin \alpha_2 \\
0 & \sin \alpha_2 & \cos \alpha_2
\end{bmatrix} \begin{bmatrix}
A^R e^{j\varphi^R} \\
A^G e^{j\varphi^G} \\
A^B e^{j\varphi^B}
\end{bmatrix}
\]

(11)

The phase \( \varphi \) is such that \( \hat{a}(-\varphi) \) coincides with the ellipse’s apogee: \( ||\hat{a}(\varphi)|| = r_+ \).

We refer the reader to [14] for a detailed explanation, as well as for the reverse calculus to retrieve the 3 independent Fourier coefficients.

Eventually, we have converted 3 independent Fourier analyzes into intuitive amplitude and phase data completed by 4 colorimetric features \( \lambda, \alpha_1, \alpha_2 \) and \( \alpha_3 \). Let us now put this into the 2D monogenic framework.

### 3. Elliptical Monogenic Representation

According to the monogenic formalism, the signal to analyze must be combined with its Riesz transform (“RT”). In order to better handle the link from 1D to 2D, we propose to rewrite the monogenic formalism in terms of directional RT, defined for any scalar 2D signal \( s(x) \) by:

\[
\mathcal{R}_s(x) = \cos(\theta) \Re \{\mathcal{R}_s(x)\} + \sin(\theta) \Im \{\mathcal{R}_s(x)\}
\]

(12)

\[
|\mathcal{R}_s(x)| \cos(\theta - \arg(\mathcal{R}_s(x)))
\]

(13)

Here the RT is used like a pair of steerable filters, of which a linear combination can provide any rotation of them. The amplitude and phase turn out to be rewritable as:

\[
A(x) e^{j\varphi(x)} = s(x) + j\mathcal{R}_s(x)
\]

(14)

where \( \theta(x) = \arg(\mathcal{R}_s(x)) \) corresponds to the analysis of the local main direction. We already know how to deal with amplitudes and phases in the color case thanks to the elliptical model, so the last key is the color orientation analysis. The problem is that, at local position \( x_1 \), three channels may have three different main orientations \( \arg(\mathcal{R}_s^s(x_1)), \arg(\mathcal{R}_s^G(x_1)) \) and \( \arg(\mathcal{R}_s^B(x_1)) \).

This issue happens to be well known in differential geometry, through the structure tensor concept, in particular with Di Zenzo’s color gradient [15, 16]. This method locally finds the orientation along which the vector signal has the maximum variation in terms of Euclidean distances. We have proposed in [17] a Riesz counterpart of it, so that we locally extract a single main orientation maximizing the signal’s color variation.

Based on the fact that the RT is equal to a gradient operator combined with an isotropic low-pass filter [18, 19], a simple “replacement” by the RT provides the same quality of analysis, while bringing the connexion with the monogenic framework.

The Riesz based structure tensor is the 2 by 2 matrix defined by:

\[
T = h * \sum_{c \in \{R,G,B\}} \begin{bmatrix}
\Re \{\mathcal{R}_s^c\}^2 & \Re \{\mathcal{R}_s^c\} \Im \{\mathcal{R}_s^c\} \\
\Re \{\mathcal{R}_s^c\} \Im \{\mathcal{R}_s^c\} & \Im \{\mathcal{R}_s^c\}^2
\end{bmatrix}
\]

(15)

where \( h \) is a smoothing filter applied on the four elements of \( T \). The local orientation is obtained from the eigenvector tied to the largest eigenvalue of \( T \):

\[
\theta_\perp = \arg(T_{11} - T_{22} + 2T_{12})/2
\]

(16)

We here benefit from the proper analysis of color discontinuities thanks to a differential geometry approach.

Now that the local color orientation is given, we can define the color monogenic analysis. We propose to locally steer the color...
Riesz transform towards the same main direction $\theta_+$ for each color channel, so that the vector amplitude-phase estimation is done coherently.

The 6 amplitude-phase parameters of eq. 5 can be defined by

$$\hat{A} = \begin{bmatrix} s_i^R(x) + jR_{s_i}^R s_i^R(x) \\ s_i^G(x) + jR_{s_i}^G s_i^G(x) \\ s_i^B(x) + jR_{s_i}^B s_i^B(x) \end{bmatrix} = \begin{bmatrix} A^R e^{j\phi^R} \\ A^G e^{j\phi^G} \\ A^B e^{j\phi^B} \end{bmatrix}$$

and converted into ellipse parameters with formulas given in section 2, to form the elliptical monogenic signal:

$$s_M = [A(x) \ \varphi(x) \ \lambda(x) \ \alpha_1(x) \ \alpha_2(x) \ \alpha_3(x)]$$

The previous definition that we gave in [17] had the same calculus of orientation but the phase-shifted signal was limited to the Euclidean norm of the whole vector RT, instead of the more complete Euclidean color directional RT proposed here. As a result, the extraction of the phase data fully benefits from the efficient ellipse parameter estimation, that properly takes into account all the rotations involved in the multidimensional space.

The vector modeling of the color 2D signal results in the conversion of the whole data into meaningful features; and makes our tool a truly non-marginal method.

The elliptical monogenic analysis is illustrated on the figure 2, from a synthetic test image featuring explicit color contours. The first part of the calculus is put on the upper row. The input image $s$ is isotropically filtered to provide $s_i$. Since the result is oscillating around (0, 0, 0), the graphic has been centered-normalized around the mid-gray [128, 128, 128]. This illustration technique is also used for the real and imaginary parts of the Riesz transform, showed next to it. The main local orientation $\theta_+$ obtained from the Riesz structure tensor is the fifth illustration. In order to respect the circularity of the data, it is displayed as the hue component in the HSV color space, while the saturation is controlled by the amplitude $A$, so as to whiten irrelevant values. Here the smoothing kernel $h$ is a Gaussian filter with $\sigma = 1$. The measured orientation is clearly in accord with our perception of contours, and equally efficient on lines and edges. This is why the color directional Riesz transform $R_{\theta_+}$ also performs well by phase-shifting the local oscillations from $s_i$ towards the proper direction (see how the yellow line is clearly turned into a curved blue/yellow edge, with respect to $x$- and $y$- components of the RT).

Let us now observe the elliptical features on the second row.

The amplitude reveals edges as well as lines, by showing a maximum on their center, contrary to what would give a simple gradient-based contour analysis. This is due to the embedding of the type of contour at the core of the signal model.

As a complement, the phase $\varphi$ provides the contour type information. It is here wrapped in one quadrant of the trigonometric circle, according to the symmetry of the related information [20]. We actually find $0 \pm \pi$ (red) and $\pm \pi/2$ (blue) at the centers of line- and edge-like structures respectively. This experimental result is clearly improved with respect to our previous work, in which the phase generally suffers from numerical unstability - due to the circular nature of the data being less well handled by the previous signal model.

The linearity of local color ellipse $\lambda$ is near 1 almost everywhere, excepted around unsignificant low-amplitude areas. This means that the chosen test image contains ‘simple’ oscillations involving two main colors. For example, the yellow line in test image is encoded by a variation between blue and yellow, drawing a very thin ellipse.

At this stage of the work, the interpretation of the expected colorimetric information from ellipse’s position angles $\alpha_1, \alpha_2, \alpha_3$ is an open issue. Yet these angles are necessary to handle well the multidimensionality of color oscillations i.e. they are the counterpart of the good quality of the amplitude-phase analysis. This method eventually provides numerically efficient phase data, which we can now exploit to achieve local frequency analysis from a color image.

4. LOCAL COLOR FREQUENCY

The local frequency $\nu$ is basically obtained by differentiating the local phase $\varphi$. In 2D, we have to differentiate along the main direction of the signal, toward which the phase makes sense. This direction is explicitly given by color monogenic feature $\theta_+$.

We apply the central finite difference on $\varphi(x)$. In order to steer it properly, our continuous orientation data $\theta_+$ must be quantized.
so as to select the neighbor “left” $\varphi_l$ and “right” $\varphi_r$ samples to be compared to the local value $\varphi(x)$. We propose the following quantization formula

$$\theta_q = (\pi/4) \text{round}(\theta_+/\pi/4)$$ \hspace{1cm} (19)

giving 4 possible directions: horizontal (0), vertical ($\pi/2$ and $-\pi/2$ are merged) and the two diagonals ($\pm \pi/4$).

For example, if $\theta_+$ lies in $[-\pi/2,\pi/2]$; then the wavefront is approximately parallel to the $y$-axis, and the phase evolves mainly along the $x$-axis. So $\varphi(x_1, x_2)$ has to be differentiated with $\varphi_l = \varphi(x_1 - 1, x_2)$ and $\varphi_r = \varphi(x_1 + 1, x_2)$.

The circularity of angle data can be handled with complex exponentials, and the local frequency defined by:

$$\nu = \arg \left( \left( \exp(\varphi_r) / \exp(\varphi_l) \right)^{1/2} \right) / \cos(\theta_+ - \theta_q)$$ \hspace{1cm} (20)

The left side term compensates the bias due to orientation quantization, according to a local linear model of $\varphi$. Note that for diagonal cases ($\theta_+ = \pm \pi/4$), $\nu$ must also be divided by $\sqrt{2}$. There is not enough space here to give all details, this is why the source code has been made publicly available [21].

The performance of this method is illustrated on figure 3. This first test is done on a synthetic image, featuring explicit local frequencies for various examples of colorimetric and geometric parameters. We expect our method to measure constant values of frequencies for various examples of colorimetric and geometric parameters, which experimentally confirms the soundness of the method.

We finally test it on a natural image - fourth graphic of fig. 3. This image features a perceptually constant unique frequency, oscillating horizontally. The difficulty is that colorimetric properties are various - blue, green and yellow. We also expect measurements not to be as regular as in the synthetic case, in particular due to richer frequency content, and randomness of real data. This is why a prior frequency band selection has been done with an isotropic filter, and the estimated $\nu(x)$ has been regularized by a $11 \times 11$ median filtering (see [21] for more details).

The figure 4 below shows the estimated frequency as a 3D mesh, together with a histogram of the measured $\nu(x)$. Apart from a mild random variation due to the original image itself, we can see that $\nu$ still lies within a small interval around $\nu = 0.6$ rad/pixel. This value corresponds to a period of around 10 pixels, which is well in accord with the visible pattern in this image, actually of this kind of periodicity. To our knowledge, no existing work deals with local frequency of color images. This experiment shows that our elliptical monogenic representation actually provides sound data, according to an oscillatory model in the spirit of the classical Fourier analysis, with an intuitive extension of the amplitude and phase concepts.

We are currently working on integrating this tool into a reversible filterbank algorithm, so as to carry out new color wavelet transforms and apply them to image enhancement and compression.

5. CONCLUSION

We have defined a new color monogenic representation providing amplitude, phase, orientation and colorimetric local features. The building blocks are the elliptical model, the directional Riesz transform and the structure tensor. The construction is driven by the physical meaning of the data, ending up with an easily usable local phase, thanks to a truly non-marginal definition. This phase is numerically differentiated to carry out the local frequency: a new concept for color images. The algorithm proves to be efficient on both synthetic and natural images. We have seen that the different pieces of information are well separated and independent from each other. This could be the basis for a new generation of SIFT-like detectors. Our current work investigates the wavelet counterpart of this framework.

R. Soulard is partially supported by Crescen2o.

6. REFERENCES


