GET: The Connection Between Monogenic Scale-Space and Gaussian Derivatives

Michael Felsberg[1]* and Ullrich Köthe[2]

[1] Linköping University, Computer Vision Laboratory,
SE-58183 Linköping, Sweden,
mfe@isy.liu.se, http://www.isy.liu.se/cvl/
[2] University of Hamburg, Cognitive Systems Group,
D-22527 Hamburg, Germany,
koethe@informatik.uni-hamburg.se,
http://kogs-www.informatik.uni-hamburg.de/

Abstract. In this paper we propose a new operator which combines advantages of monogenic scale-space and Gaussian scale-space, of the monogenic signal and the structure tensor. The gradient energy tensor (GET) defined in this paper is based on Gaussian derivatives up to third order using different scales. These filters are commonly available, separable, and have an optimal uncertainty. The response of this new operator can be used like the monogenic signal to estimate the local amplitude, the local phase, and the local orientation of an image, but it also allows to measure the coherence of image regions as in the case of the structure tensor. Both theoretically and in experiments the new approach compares favourably with existing methods.

1 Introduction

In this paper we derive a connection between features of the monogenic scale-space [1] of an image and its Gaussian scale-space [2], respectively the derivatives of the latter. Thus, it becomes possible to compute monogenic features from Gaussian derivatives. The advantages of the proposed method are:

- Many people have implementations of Gaussian derivatives available so that they can use monogenic features without implementing new basis filters.
- The Gaussian derivatives are separable and decay faster than the Poisson filter and its Riesz transform resulting in more efficient computational schemes.
- The additional feature (coherence) of the derivative-based method directly indicates the validity of the monogenic phase model which is based on the assumption of locally 1D signals.

A key assumption of this paper is of course that the local phase is useful for the processing and analysis of images. Therefore, we give a short motivation of phase-based image processing in the subsequent section. Although most of the discussions focus on images, the reflections about phase based signal processing generalize to signals of arbitrary dimension.

* This work has been supported by EC Grant IST-2002-002013 MATRIS and by EC Grant IST-2003-004176 COSPAL.

1.1 Phase-Based Image Processing

First of all, there is some evidence that the human visual system makes use of local phase to analyze the image contents [3]. Since the human visual system performs remarkably well in analyzing images, it is reasonable to design technical systems accordingly. However, there are also purely technical observations which motivate the use phase. In [4] the authors present several experiments which show that the Fourier phase contains the major part of the signal information. The same applies to the local phase. For the definition of local phase, we assume an image (patch) model according to

$$I(\mathbf{x}) = A(\mathbf{x})\cos(\varphi(\mathbf{x})) + \bar{I} \tag{1}$$

where $\mathbf{x} = (x, y)^T$ indicates the spatial coordinate vector, $I(\mathbf{x})$ the image (patch), $\bar{I}$ the average intensity (DC level), $A(\mathbf{x})$ the local amplitude (non-negative), and $\varphi(\mathbf{x})$ the local phase. The average intensity is irrelevant for the analysis of the image contents and is largely compensated already during the image acquisition in the human visual system. What remains is to analyze the relation of local amplitude and local phase. Although the decomposition in (1) seems to be ambiguous, this is not the case due to the non-negativity of the amplitude. Due to the latter, zero crossings in $I(\mathbf{x}) - \bar{I}$ must be covered by zeros of $\cos(\varphi(\mathbf{x}))$ and zero crossings are in direct correspondence to the full phase [5]. Therefore, the local phase becomes a uniquely defined feature.

If the image is decomposed into its amplitude and phase information, it becomes evident that the local amplitude is basically a measure for the confidence of the extracted phase, i.e., in technical terms it represents the signal-to-noise ratio (SNR), cf. Fig. 1, center. The local phase represents most of the image structure, cf. Fig. 1, left. In the areas where the amplitude is close to zero, thus meaning 'no confidence', the local phase contains mainly noise. In the regions of non-zero confidence, the cosine of the local phase results in a visual impression which comes very close to the original, bandpass-filtered image, cf. Fig. 1, right.

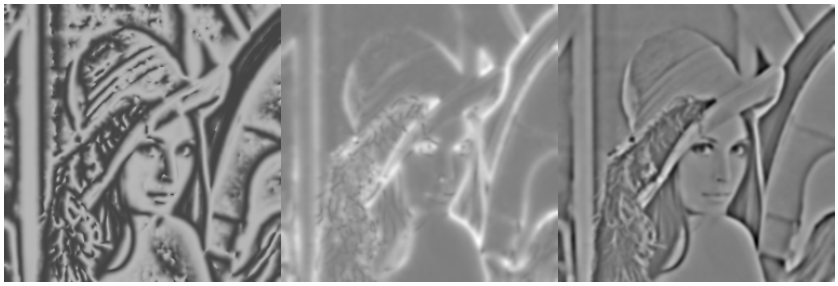

Fig. 1. Decomposing a bandpass image into its local phase and its local amplitude. From left to right: $\cos(\varphi(\mathbf{x}))$, $A(\mathbf{x})$, and $I(\mathbf{x})$, where the intensities where adjusted to obtain similar intensity ranges. Grey means zero, white means positive values, and black means negative values. $I(\mathbf{x})$ is obtained from a bandpass-filters with center frequency $\pi/6$.

1.2 The Monogenic Scale-Space: A Brief Survey

The monognic scale-space is a framework to estimate the local phase, the local orientation, and the local amplitude from an image at different scales [1]. The starting point is to compute the Poisson scale-space $p(\mathbf{x}, s)$ of the image. The corresponding figure flow is obtained as the Riesz transform $\mathbf{q} = (q_1,\, q_2)^T$

$$\mathbf{q}(\mathbf{x}, s) = \int_{\mathrm{R}^2} \frac{\mathbf{x}'}{2\pi|\mathbf{x}'|^3} p(\mathbf{x}-\mathbf{x}', s)\, d\mathbf{x}' = \int_{\mathrm{R}^2} \frac{\mathbf{x}'}{2\pi(|\mathbf{x}'|^2+s^2)^{3/2}} p(\mathbf{x}-\mathbf{x}', 0)\, d\mathbf{x}' \quad (2)$$

of the image at each scale s. Together, the blurred image and its Riesz transform form a monogenic signal [6] at the respective scale.

The monogenic signal contains of three components at each position, i.e., for a fixed scale s_0 it is a function $\mathrm{R}^2 \to \mathrm{R}^3 :\ \mathbf{x} \mapsto (q_1(\mathbf{x}, s_0),\, q_2(\mathbf{x}, s_0),\, p(\mathbf{x}, s_0))^T$. For convenience, we sometimes omit the arguments $\mathbf{x}$ and s in the following. The 3D co-domain is then transformed into polar coordinates, cf. Fig. 2, left, resulting in a triplet $(A, \varphi, \theta) \in \mathrm{R}^+ \times [0, 2\pi) \times [0, \pi)$ where $A = \sqrt{q_1^2 + q_2^2 + p^2}$ is an estimate for the local amplitude, $\varphi = \arg(p + i\,\mathrm{sign}\,(q_1)|q|)$ for the local phase, and $\theta = \tan^{-1}(q_2/q_1)$ for the local orientation of the image if the image varies locally only in the orientation θ (intrinsic dimensionality of one [7]). Since the direction of an image does not follow from its local orientation [8], an ambiguity concerning the sign of the phase is obtained. In order to obtain a continuous representation of orientation and phase, they can be combined to form a 2D phase vector $\mathbf{r} = \varphi(-\sin\theta,\, \cos\theta)^T = \mathbf{q}^\perp/|\mathbf{q}| \tan^{-1}(|\mathbf{q}|/p)$, cf. Fig. 2, right.

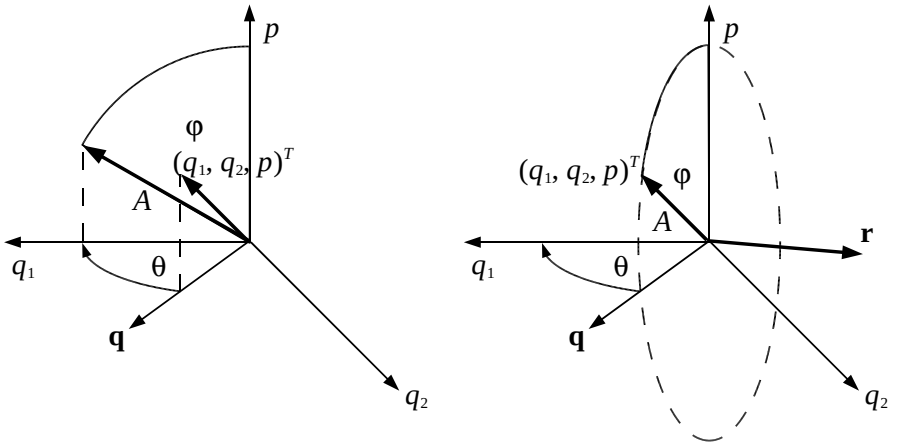

Fig. 2. Phase models used in context of the monogenic signal. Left: the 3D vector is derotated by the local orientation θ, such that it lies in the (q_1, p)-plane. The amplitude and phase are then extracted like in the 1D case as vector length and argument. Right: the 3D vector together with the p-axis define a plane in which the rotation takes place. The normal of this plane multiplied by the (directed) rotation angle φ results in the rotation vector $\mathbf{r}$.

Further features can be derived from the local features of the monogenic signal respectively the monogenic scale-space, e.g., local frequency and phase congruency as the spatial and scale derivatives of the local phase. The consideration of these features is however out of the scope of this paper.

In order to estimate the local features, implementations of the monogenic signal and the monogenic scale-space are required. This can either be done by local operators, which combine a radial bandpass filter with its Riesz transform [9,10], or by a global eigentransform solution [11]. The problem is, however, that the involved Poisson kernel decays quite slowly, resulting in either large truncation errors of the filter masks or non-locality of the output.

Even and odd filters based on, e.g., Gaussian derivatives, are preferable concerning locality, but these filters do not allow to estimate the local phase or phase invariant features in a linear framework, since their respective amplitude responses differ. To combine the locality of Gaussian derivatives with phase invariant feature extraction and phase estimation is the main topic of the present paper. The key idea is to use a quadratic operator in order to avoid using the Riesz transform. This idea is based on the concept of the 1D energy operator, which is briefly introduced in the subsequent section.

1.3 The 1D Energy Operator

This brief review of the 1D energy operator is based on [12]. The purpose of the energy operator is to compute directly the squared local amplitude of a signal without using the Hilbert transform, since the Hilbert transform based methods suffer from the same phenomena as the implementations of the monogenic scale-space. The energy operator is defined for continuous 1D signals $s(t)$ as

$$\Psi_c[s(t)] = [\dot{s}(t)]^2 - s(t)\ddot{s}(t) \ . \tag{3}$$

It is obviously not positive semi-definite, but it tracks the energy of simple harmonic oscillators. Moreover, for constants A, r, and ω_0 and for any s_1, s_2

$$\Psi_c[Ar^t \cos(\omega_0 t + \varphi_0)] = A^2 r^{2t} \omega_0^2 \tag{4}$$

$$\Psi_c[s_1(t)s_2(t)] = s_1(t)^2 \Psi_c[s_2(t)] + s_2(t)^2 \Psi_c[s_1(t)] \ . \tag{5}$$

If we instead just consider $[\dot{s}(t)]^2$, likewise the orientation tensor in higher dimensions, we obtain

$$[\frac{d}{dt} A\cos(\omega_0 t + \varphi_0)]^2 = \frac{1}{2} A^2 \omega_0^2 (1 - \cos(2\omega_0 t + 2\varphi_0)) \ ,$$

which is obviously not phase invariant and might even suffer from aliasing if ω_0 is larger than half the Nyquist frequency. Apparently, the second part $s(t)\ddot{s}(t)$ of the energy operator exactly compensates the spurious modulation components at $2\omega_0$. A possible 2D generalization of the energy operator is the energy tensor [13], which we introduce in the subsequent section.

1.4 The 2D Energy Tensor

For continuous, 2D bandpass signals $b(\mathbf{x})$, the 2D energy tensor is defined as [13]

$$\Psi_c[b(\mathbf{x})] = [\nabla b(\mathbf{x})][\nabla b(\mathbf{x})]^T - b(\mathbf{x})[\mathbf{H}b(\mathbf{x})] \ , \tag{6}$$

where $\nabla = (\partial_x, \partial_y)^T$ indicates the gradient and $\mathbf{H} = \nabla\nabla^T$ indicates the Hessian. Likewise in the 1D case, this operator is not positive semi-definite in general, but for a simple harmonic oscillator it results in a energy-frequency-weighted orientation tensor. Moreover, we obtain

$$\Psi_c[Ar^{x+y}\cos(\mathbf{u}_0^T\mathbf{x} + \varphi_0)] = A^2 r^{2x+2y}\mathbf{u}_0\mathbf{u}_0^T \tag{7}$$

$$\Psi_c[s_1(\mathbf{x})s_2(\mathbf{x})] = s_1(\mathbf{x})^2\Psi_c[s_2(\mathbf{x})] + s_2(\mathbf{x})^2\Psi_c[s_1(\mathbf{x})] \ . \tag{8}$$

If we just consider the first part of (6), i.e., the structure / orientation tensor according to [14,15] (but without spatial averaging), we obtain

$$[\nabla A\cos(\mathbf{u}_0^T\mathbf{x} + \varphi_0)][\nabla A\cos(\mathbf{u}_0^T\mathbf{x} + \varphi_0)]^T = \frac{1}{2}A^2\mathbf{u}_0\mathbf{u}_0^T(1 - \cos(2\mathbf{u}_0^T\mathbf{x} + 2\varphi_0)) \ ,$$

which is likewise in the 1D case not phase invariant and might show aliasing artifacts.

The energy tensor is a second order symmetric tensor like the structure tensor. The latter is included in the energy tensor, but it is combined with a product of even filters, which provides phase invariance for simple harmonic oscillators and products thereof. The energy tensor can hence be classified as a phase invariant, orientation equivariant second order tensor [16]. Same as the 2D structure tensor, the energy operator can be converted into a complex double angle orientation descriptor [17]:

$$o(\mathbf{x}) = \Psi_c[b(\mathbf{x})]_{11} - \Psi_c[b(\mathbf{x})]_{22} + i2\Psi_c[b(\mathbf{x})]_{12} \tag{9}$$

which is equivalent to the 2D energy operator defined in [18]. As one can easily show, $|o(\mathbf{x})| = |\lambda_1(\mathbf{x}) - \lambda_2(\mathbf{x})|$, where $\lambda_1(\mathbf{x}), \lambda_2(\mathbf{x})$ are the eigenvalues of the energy tensor. Since the trace of the tensor is given by the sum of eigenvalues, we obtain $2\lambda_{1,2} = \mathrm{tr}(\Psi_c[b(\mathbf{x})]) \pm |o(\mathbf{x})|$, which can be subject to the same analysis in terms of coherence as suggested in [19,8] or for the Harris detector [20].

If the signal is not a (product of) simple harmonic oscillations, the operator (6) does not result in a positive response in general. However, if the local signal region adheres to the model (1) with slowly varying amplitude and frequency, the response is positive. This is the case if we prefilter the signal with a bandpass, which avoids low frequencies (DC component and changes of local amplitude) and high frequencies. Removing high frequencies can be considered as a regularization that allows the computation of derivatives for discrete data. The prefiltering is necessary in most practical situations, since natural images $I(\mathbf{x})$ are typically no bandpass signals $b(\mathbf{x})$.

2 The GET Operator

As pointed out above, the signal needs to be bandpass filtered in order to obtain small frequency ranges, and hence, positive responses. For the high frequency regularization, we prefer Gaussian functions due to their high localization in both domains. However, Gaussian filters are not DC-free, which is a central requirement in context of the energy tensor. If we consider a difference of Gaussian filters as in [13], we implicitly lift the level of spatial differentiation by two. According to the equation of linear diffusion [2], the scale derivative of a Gaussian filter is equivalent to the Laplacian of the Gaussian, i.e., a combination of second order derivatives. Hence, applying the Hessian to the Laplacian of the Gaussian means to consider fourth order derivatives instead of second order derivatives. Due to angular aliasing however, one cannot compute fourth order derivatives on a local support [10]. Therefore, we propose an operator below which makes use of Gaussian derivatives up to order three, but avoids the zeroth order Gaussian, i.e., the DC-component is removed.

2.1 The Gradient Energy Tensor

The idea to define the gradient energy tensor (GET) follows from the previous considerations. We introduce the tensor in three steps. First, we plug the gradient of the image into (6) and use tensor notation instead of matrix notation:

$$\begin{aligned}\mathrm{GET}\,\{I(\mathbf{x})\} &= \Psi_c[\nabla I(\mathbf{x})] \\ &= [\nabla \otimes \nabla I(\mathbf{x})] \otimes [\nabla \otimes \nabla I(\mathbf{x})] \\ &\quad -\frac{1}{2}([\nabla I(\mathbf{x})] \otimes [\nabla \otimes \nabla \otimes \nabla I(\mathbf{x})] + [\nabla \otimes \nabla \otimes \nabla I(\mathbf{x})] \otimes [\nabla I(\mathbf{x})])\end{aligned} \tag{10}$$

where we symmetrized the tensor by replacing the second term with the corresponding anticommutator term. The obtained operator has 16 coefficients, where 6 can be omitted due to symmetry and one further coefficient is a linear combination of two others. Hence, 9 independent coefficients are left.

In a second step, we contract the tensor. This becomes possible, since there is no practical gain from the coefficients that are omitted in the contraction:

$$\begin{aligned}\mathrm{GET}\,\{I(\mathbf{x})\} &= [\nabla \otimes \nabla I(\mathbf{x})] \cdot [\nabla \otimes \nabla I(\mathbf{x})] \\ &\quad -\frac{1}{2}([\nabla I(\mathbf{x})] \otimes [\nabla \cdot \nabla \otimes \nabla I(\mathbf{x})] + [\nabla \otimes \nabla \cdot \nabla I(\mathbf{x})] \otimes [\nabla I(\mathbf{x})]) \\ &= [\mathbf{H}I(\mathbf{x})][\mathbf{H}I(\mathbf{x})] - \frac{[\nabla I(\mathbf{x})][\nabla \Delta I(\mathbf{x})]^T + [\nabla \Delta I(\mathbf{x})][\nabla I(\mathbf{x})]^T}{2}\end{aligned} \tag{11}$$

In this formula $\Delta = \nabla^T \nabla$ denotes the Laplacian. Due to the non-linearity of the operator, it is difficult to show which degrees of freedom are lost in the contraction, but we can consider certain different cases. Assuming a simple harmonic oscillation $I(\mathbf{x}) = \cos(u\,x + v\,y + \phi)$, we obtain for the full tensor

$$\mathrm{GET}\left\{I(\mathbf{x})\right\} = \begin{bmatrix} \begin{bmatrix} u^4 & u^3\,v \\ u^3\,v & u^2\,v^2 \end{bmatrix} & \begin{bmatrix} u^3\,v & u^2\,v^2 \\ u^2\,v^2 & u\,v^3 \end{bmatrix} \\ \begin{bmatrix} u^3\,v & u^2\,v^2 \\ u^2\,v^2 & u\,v^3 \end{bmatrix} & \begin{bmatrix} u^2\,v^2 & u\,v^3 \\ u\,v^3 & v^4 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} u^2 \begin{bmatrix} u^2 & u\,v \\ u\,v & v^2 \end{bmatrix} & u\,v \begin{bmatrix} u^2 & u\,v \\ u\,v & v^2 \end{bmatrix} \\ u\,v \begin{bmatrix} u^2 & u\,v \\ u\,v & v^2 \end{bmatrix} & v^2 \begin{bmatrix} u^2 & u\,v \\ u\,v & v^2 \end{bmatrix} \end{bmatrix}$$

and for the contracted tensor

$$\mathrm{GET}\left\{I(\mathbf{x})\right\} = \begin{bmatrix} u^2 & \left(u^2+v^2\right) & u\,v & \left(u^2+v^2\right) \\ u\,v & \left(u^2+v^2\right) & v^2 & \left(u^2+v^2\right) \end{bmatrix} \quad . \tag{12}$$

Hence, no information is lost by the contraction under the assumed signal model. If we extend the model to two different frequencies in the same direction, the tensor coefficients are multiplied by a spurious modulation factor.[3] However, this modulation is the same for all coefficients, and therefore, the full tensor does not provide additional information. By repeating this procedure for more frequencies in the same direction, the result will always be the same, and hence, of locally 1D signals there is no gain from the full tensor.

If we assume the signal to contain two arbitrary *perpendicular* oscillations, i.e., $I(\mathbf{x}) = \cos(u\,x + v\,y + \phi) + \cos(l\,v\,x - l\,u\,y + \psi)$, the contracted tensor shows no modulations:

$$\mathrm{GET}\left\{I(\mathbf{x})\right\} = \begin{bmatrix} u^4 + \left(1+l^4\right) u^2\,v^2 + l^4\,v^4 & -\left(\left(-1+l^4\right) u\,v\,\left(u^2+v^2\right)\right) \\ -\left(\left(-1+l^4\right) u\,v\,\left(u^2+v^2\right)\right) & \left(u^2+v^2\right)\left(l^4\,u^2+v^2\right) \end{bmatrix} , \tag{13}$$

but the full tensor does. Performing the contraction, the modulations exactly compensate each other, which is a very strong argument in favor of the contracted tensor. As it can be seen from this example, the 2×2 tensor obtained from (11) allows to estimate two perpendicular oscillations with different frequencies at the same time, i.e., it covers the same model as the structure multivector in [10].

Due to the non-linear behavior of the tensor it is impossible to calculate the response for a general signal. However, one can analyze it in terms of null-spaces and it turns out that the contraction does not change the null-space of the operator. This follows from the equality

$$\mathrm{GET}\left\{I(\mathbf{x})\right\} = \Psi_c[\partial_x I(\mathbf{x})] + \Psi_c[\partial_y I(\mathbf{x})] \quad . \tag{14}$$

Deeper investigations of the null-spaces and the number of independent components will be subject to a future publication.

2.2 Regularization and Gaussian Derivatives

The results from the previous section are obtained for idealized, continuous signals. In practice, however, we have to deal with non-ideal, noisy, and discrete signals. The most common thing to do is therefore to regularize the derivative operators from (11) with Gaussian kernels. A Gaussian regularization is the optimal choice if nothing is known about the signal and its noise characteristic.

[3] The spatial modulation is undesired. The response should have constant amplitude.

Therefore, we replace the derivatives in (11) with Gaussian derivatives of order one to three.

The scales for the regularization are chosen such that the variance increases linearly with the order of the derivative, cf. (13) in [21]. In [22] we discuss the choice of scales and different regularizations more in detail.

2.3 Extraction of Monogenic Features

The monogenic signal provides three features: local amplitude, local phase, and local orientation [6]. In case of signals with intrinsic dimensionality one, i.e., $I(\mathbf{x}) = s(\mathbf{n}^T\mathbf{x})$ ($s : \mathrm{R} \to \mathrm{R}$, $|\mathbf{n}| = 1$), the GET is of rank one:

$$\begin{aligned}\mathrm{GET}\,\{I(\mathbf{x})\} &= [\mathbf{n}\mathbf{n}^T\ddot{s}(\mathbf{n}^T\mathbf{x})][\mathbf{n}\mathbf{n}^T\ddot{s}(\mathbf{n}^T\mathbf{x})] \\ &\quad - \frac{[\mathbf{n}\dot{s}(\mathbf{n}^T\mathbf{x})][\mathbf{n}\dddot{s}(\mathbf{n}^T\mathbf{x})]^T + [\mathbf{n}\dddot{s}(\mathbf{n}^T\mathbf{x})][\mathbf{n}\dot{s}(\mathbf{n}^T\mathbf{x})]^T}{2} \\ &= \mathbf{n}\mathbf{n}^T[\ddot{s}(\mathbf{n}^T\mathbf{x})^2 - \dddot{s}(\mathbf{n}^T\mathbf{x})\dot{s}(\mathbf{n}^T\mathbf{x})] \ .\end{aligned}$$

The first eigenvector of this expression is $\pm\mathbf{n}$, i.e., the local orientation of the signal. The first eigenvalue (or its trace, aka the second eigenvalue is zero) of the GET is more difficult to analyze, except for the single-frequency case, where we obtain according to (12) $|\mathbf{u}|^4A^2$ for an oscillation with amplitude A.

Much more interesting is the extraction of the local phase, which is obtained in two steps. First, we consider the two addends of the GET separately. The first one represents the symmetric (even) parts of the signal, whereas the second one represents the antisymmetric (odd) parts of the signal. However, both parts are quadratic expressions, such that we have to consider their square-roots:

$$q_{\mathrm{even}} = \pm\sqrt{\mathrm{trace}(\mathbf{T}_{\mathrm{even}})} \qquad \text{and} \qquad q_{\mathrm{odd}} = \pm\sqrt{\mathrm{trace}(\mathbf{T}_{\mathrm{odd}})} \qquad \text{where}$$

$$\mathbf{T}_{\mathrm{even}} = [\mathbf{H}I(\mathbf{x})][\mathbf{H}I(\mathbf{x})] \qquad \text{and} \tag{15}$$

$$\mathbf{T}_{\mathrm{odd}} = -\frac{[\nabla I(\mathbf{x})][\nabla\Delta I(\mathbf{x})]^T + [\nabla\Delta I(\mathbf{x})][\nabla I(\mathbf{x})]^T}{2} \ . \tag{16}$$

In a second step, the correct signs for both parts are selected, such that $\arg(q_{\mathrm{even}} + iq_{\mathrm{odd}})$ gives the local phase of the signal. Comparing the signs in different quadrants of a harmonic oscillation results in the following procedure. Let $\mathbf{T} = \mathbf{T}_{\mathrm{even}} + \mathbf{T}_{\mathrm{odd}}$ denote the GET response, $o = \mathbf{T}_{11} - \mathbf{T}_{22} + i2\mathbf{T}_{12}$ its complex double angle orientation representation [17], and $\mathbf{o} = (\mathrm{real}(\sqrt{o}), \mathrm{imag}(\sqrt{o}))^T$ the orientation vector. We then define the two signs as

$$s_{\mathrm{even}} = -\mathrm{sign}\,(\mathbf{o}^T[\mathbf{H}I(\mathbf{x})]\mathbf{o}) \qquad \text{and} \qquad s_{\mathrm{odd}} = -\mathrm{sign}\,(\mathbf{o}^T\nabla I(\mathbf{x})) \tag{17}$$

such that

$$\varphi = \arg(q_{\mathrm{even}} + iq_{\mathrm{odd}}) = \arg(s_{\mathrm{even}}\sqrt{\mathrm{trace}(\mathbf{T}_{\mathrm{even}})} + is_{\mathrm{odd}}\sqrt{\mathrm{trace}(\mathbf{T}_{\mathrm{odd}})}) \tag{18}$$

is consistent with the definition of the monogenic phase. This can easily be verified by inserting $\cos(ux + vy)$ into the previous four expressions, resulting in $\varphi = ux + vy$ if $(u, v)^T$ lies in the upper half-plane and $\varphi = -ux - vy$ otherwise.

If the underlying signal is non-simple, i.e., it has intrinsic dimensionality two, the analysis becomes more difficult. Following the strategy of the structure multivector in [10], the first eigenvector is extracted from $\mathbf{T}$. Then, the even tensor and the odd tensor are projected onto the first eigenvector and onto the orthogonal vector (aka the second eigenvector). This gives two even components and two odd components, which are then combined with appropriate signs to extract two phases for the two perpendicular orientations.

Note also that in the latter case not a single amplitude is obtained, but two eigenvalues, which correspond to the local amplitudes of the two perpendicular components. These eigenvalues can then be used for coherence analysis or corner detection likewise the eigenvalues of the structure tensor.

3 Comparisons

In this section we compare the results of the GET operator with those of the DCT-based implementation [11], the spherical quadrature filters [10], and the structure tensor (ST) by outer products of gradients (see e.g. [19]). The latter approach is not suited for phase-estimation per se, but one can easily extend it for this purpose in the following way. Assuming that the outer product of gradients of a cosine oscillation results in a trace which is $A\sin^2(\varphi)$ and assume further that local averaging can be replaced with integration over entire periods, the trace of the (averaged) tensor becomes $\frac{A}{2}$. Hence, the sine and the cosine are obtained up to a sign-ambiguity by

$$q_{\text{odd}} = \pm\sqrt{t(\mathbf{x})} \qquad \text{and} \qquad q_{\text{even}} = \pm\sqrt{-t(\mathbf{x}) + 2\sum t(\mathbf{x})}\ ,$$

where $t(\mathbf{x}) = \text{trace}([\nabla I(\mathbf{x})][\nabla I(\mathbf{x})]^T)$. For the subsequent comparison only the second sign needs to be recovered. If we have locally vanishing DC components and single frequency, it is obtained by the sign of $I(\mathbf{x})$, otherwise we use the sign of $-\Delta I$. In order to remove some outliers, the signs are median-filtered.

3.1 Experiment: Extraction of Phase and Orientation

In this experiment, we applied all three methods to a synthetic pattern, cf. Fig. 3, top left, and a real image, cf. Fig. 3, bottom row. We added Gaussian noise to the pattern (SNR 3.0dB) and selected a mask for the feature comparison. The scales of all methods were chosen such that the local amplitude estimates were comparably similar, although the higher spatial-frequency localization of the GET and the ST leads to a narrower ridge then for the other two methods, cf. Fig. 3, second row. Instead of showing the phase estimates, we reconstructed the signal from the phase estimates (cf. Fig. 3, third row) and computed their SNRs. Furthermore, we computed the orientation error according to [23]. The results are summarized in Tab. 1, which shows that if we are interested in simultaneously estimating orientation and phase, the GET gives the best results.

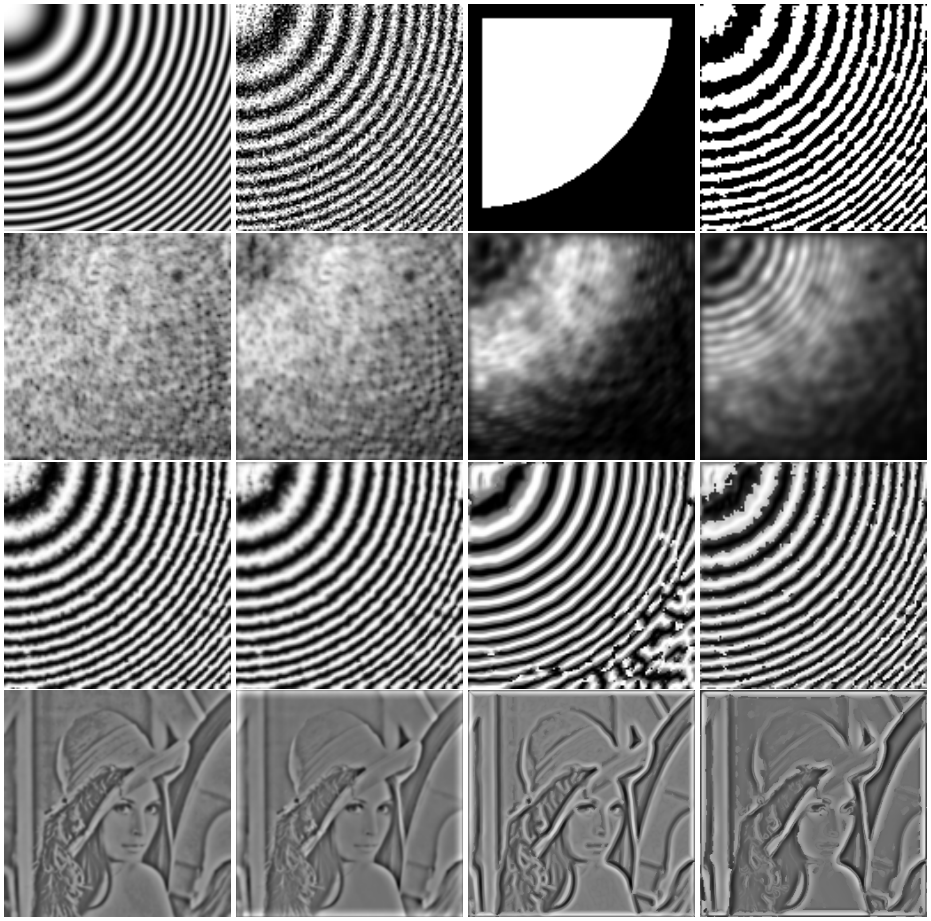

Fig. 3. Top row from left to right: test-pattern, test-pattern with noise (SNR 3.0dB), mask for error evaluation, and sign extraction for ST. Second row: amplitudes of DCT, SQF, GET, and ST (from left to right). Third row: respective reconstructions. Bottom row: respective amplitude-weighted reconstructions of a real image.

Table 1. SNR of reconstruction and orientation error from estimates, cf. Fig. 3.

method	DCT	SQF	GET	ST
SNR	12.8dB	13.7dB	13.5dB	11.1dB
$\Delta\theta$	19.6°	18.5°	7.0°	3.6°

3.2 Complexity Analysis

The computational complexity might also be an important aspect when it comes to the selection of suitable methods. Since the extraction of phase and orientation has to be done in all cases, we only compare the complexities up to that point.

The complexity of the DCT-based method is given by ten 2D FFTs [11], since we have to compute three scales. Hence, we obtain $30N^2 \log_2 N$ floating point operations (FLOPs) for an image of size $N \times N$ if N is a power of two. For our test image we have $N = 128$, such that we applied about $3.4 \cdot 10^6$ FLOPs.

For the SQF filter set, the complexity depends on the filter size. In our experiment, we used three 23×23 filters. The filters are not separable, but we can exploit a four-fold symmetry for the even filter and an eight-fold symmetry for the odd filter pair. Hence, we end up with $1850N^2$ FLOPs, which is about $3 \cdot 10^7$ FLOPs in our test.

For the GET operator, the complexity also depends on the filter size. In our experiment, we used 2σ-truncation (for the largest scale) resulting in seven 17×17 filters. These filters are separable and each of the 1D filters can exploit a two-fold symmetry. Hence, we get $357N^2$ FLOPs, i.e., $6 \cdot 10^6$ FLOPs in our special case. For the structure tensor, the computational effort is about the same if we take into account the calculation of the sign (bandpass filter and median filter).

One problem with these complexity estimates are the missing complexities for memory accesses, which become more and more important nowadays. As an side-effect of this, the SQF filters are 1.5 times faster than the DCT based method and the GET operator is 2 times faster than the SQF filters.

4 Conclusion

In this paper we have described an alternative way of extracting the image features of the monogenic signal, i.e., local amplitude, local phase, and local orientation, by using a quadratic form. The proposed method of the gradient energy tensor is the contraction of a fourth order tensor built from image derivatives of order one to three. The new tensor is compatible to the structure tensor concerning eigensystem analysis, but it is phase-invariant without spatial averaging. Using Gaussian regularization of the derivatives leads to a connection of monogenic scale-space and Gaussian scale-space via the quadratic form.

We provided formulas to extract the local phase from the two different parts of the GET, and compared the extracted features phase and orientation to those of previous approaches. Considering both estimates at once, the GET provides the best estimates and it is also among the fastest operators. For non-simple signals, it even provides the two additional features of second eigenvalue and second phase, which makes it comparable to the much slower structure multivector.

References

1. Felsberg, M., Sommer, G.: The monogenic scale-space: A unifying approach to phase-based image processing in scale-space. JMIV **21** (2004) 5–26
2. Koenderink, J.J.: The structure of images. Biol. Cybernetics **50** (1984) 363–370
3. Mechler, F., Reich, D.S., Victor, J.D.: Detection and discrimination of relative spatial phase by V1 neurons. Journal of Neuroscience **22** (2002) 6129–6157
4. Oppenheim, A., Lim, J.: The importance of phase in signals. Proc. of the IEEE **69** (1981) 529–541
5. Hurt, N.E.: Phase Retrieval and Zero Crossings. Kluwer Academic, Dordrecht (1989)
6. Felsberg, M., Sommer, G.: The monogenic signal. IEEE Transactions on Signal Processing **49** (2001) 3136–3144
7. Krieger, G., Zetzsche, C.: Nonlinear image operators for the evaluation of local intrinsic dimensionality. IEEE Trans. on Image Processing **5** (1996) 1026–1041
8. Granlund, G.H., Knutsson, H.: Signal Processing for Computer Vision. Kluwer Academic Publishers, Dordrecht (1995)
9. Felsberg, M.: On the design of two-dimensional polar separable filters. In: 12th European Signal Processing Conference, Vienna, Austria. (2004)
10. Felsberg, M.: Low-Level Image Processing with the Structure Multivector. PhD thesis, Christian-Albrechts-University of Kiel (2002) TR no. 0203.
11. Felsberg, M., Duits, R., Florack, L.: The monogenic scale space on a rectangular domain and its features. IJCV (2004) accepted publ. in future issue.
12. Maragos, P., Bovik, A.C., Quartieri, J.F.: A multi-dimensional energy operator for image processing. In: SPIE Conference on Visual Communications and Image Processing, Boston, MA (1992) 177–186
13. Felsberg, M., Granlund, G.: POI detection using channel clustering and the 2D energy tensor. In: 26. DAGM Symposium Mustererkennung, Tübingen. (2004)
14. Förstner, W., Gülch, E.: A fast operator for detection and precise location of distinct points, corners and centres of circular features. In: ISPRS Intercommission Workshop, Interlaken. (1987) 149–155
15. Bigün, J., Granlund, G.H.: Optimal orientation detection of linear symmetry. In: Proc. IEEE First International Conference on Computer Vision (1987) 433–438
16. Nordberg, K.: Signal Representation and Processing using Operator Groups. PhD thesis, Linköping University (1995) Dissertation No 366, ISBN 91-7871-476-1.
17. Bigün, J., Granlund, G.H., Wiklund, J.: Multidimensional orientation estimation with applications to texture analysis and optical flow. PAMI **13** (1991) 775–790
18. Larkin, K.G., Oldfield, M.A., Bone, D.J.: Demodulation and phase estimation of two-dimensional patterns. Australian patent AU 200110005 A1 (2001)
19. Jähne, B.: Digitale Bildverarbeitung. Springer, Berlin (1997)
20. Harris, C.G., Stephens, M.: A combined corner and edge detector. In: 4th Alvey Vision Conference. (1988) 147–151
21. Florack, L., Duits, R.: Regularity classes for locally orderless images. In Griffin, L.D., Lillholm, M., eds.: Scale-Space Theory in Computer Vision. Volume 2695 of LNCS., Springer, Heidelberg (2003) 255–265
22. Köthe, U., Felsberg, M.: Riesz-transforms vs. derivatives: On the relationship between the boundary tensor and the energy tensor. In: Proc. Scale Space Conference (this volume). (2005)
23. Knutsson, H., Andersson, M.: Robust N-dimensional orientation estimation using quadrature filters and tensor whitening. In: Proceedings of IEEE International Conference on Acoustics, Speech, & Signal Processing, Adelaide, Australia (1994)