Available online at www.sciencedirect.com
SCIENCE DIRECT®
Signal Processing: *Image Communication* 20 (2005) 569–581
SIGNAL PROCESSING: IMAGE COMMUNICATION
www.elsevier.com/locate/image

Implications of invariance and uncertainty for local structure analysis filter sets

Hans Knutsson*, Mats Andersson

Medical Informatics, Department of Biomedical Engineering, Center for Medical Image Science and Visualization, Linköping University, University Hospital, SE-581 15 Linköping, Sweden

Received 14 March 2005

Abstract

The paper discusses which properties of filter sets used in local structure estimation that are the most important. Answers are provided via the introduction of a number of fundamental invariances. Mathematical formulations corresponding to the required invariances leads up to the introduction of a new class of filter sets termed *loglets*. Loglets are polar separable and have excellent uncertainty properties. The directional part uses a spherical harmonics basis. Using loglets it is shown how the concepts of quadrature and phase can be defined in n-dimensions. It is also shown how a reliable measure of the certainty of the estimate can be obtained by finding the deviation from the signal model manifold.

Local structure analysis algorithms are quite complex and involve a lot more than the filters used. This makes comparisons difficult to interpret from a filter point of view. To reduce the number 'free' parameters and target the filter design aspects a number of simple 2D experiments have been carried out. The evaluation supports the claim that loglets are preferable to other designs. In particular it is demonstrated that the loglet approach outperforms a Gaussian derivative approach in resolution and robustness.

Keywords: Generalized phase; Invariance; Uncertainty; Quadrature filters; Local structure tensor; Spherical harmonics; Orientation estimation; Velocity estimation

1. Introduction

The first steps towards analysis of images were taken 4 decades ago. From the very start detecting edges and lines in images was considered a fundamental operation [31,17]. Since these early days new and more advanced schemes for analysis

*Corresponding author. Tel.: +46 13227294; fax: +46 13101902.
E-mail address: knutte@imt.liu.se (H. Knutsson).
URL: http://www.imt.liu.se/mi.

0923-5965/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.image.2005.03.011

of local image structure has been suggested in a seemingly never ending stream Local image orientation, scale, frequency, phase, motion and locality of estimates are prominent examples of features that have been considered central in the analysis [12,13,24,27,6,5,28,36,19,29,20,9,32,21,2,26,34,35,22,15,8,4,11,33].

Apart from sheer curiosity, the main force driving the research has been the need to analyze data produced by increasingly capable imaging devices. Presently produced data are also often intrinsically more complex. Both the outer and the inner dimensionality can be higher, e.g. volume sequence data and tensor field data respectively.

Regardless of this development the first stages in the analysis remain the same. In most cases the processing starts by performing local linear combinations of image values, e.g. convolution operators. Perhaps somewhat surprising after 30 years of research the design of these filters is still debated. In fact the object of this paper is to contribute to this discussion in a way that hopefully will help in bringing it to an end by providing what we believe to be a valid line of reasoning.

1.1. Image signal models

The mathematical framework used to characterize image signals is the foundation for development and evaluation of all image processing methods. Methods in image processing can in most cases be classified as belonging to the deterministic or the statistical world. This is unfortunate since none of the approaches alone is well suited for modeling real life images. Edges, lines, corners etc are naturally deterministically modeled whereas textures and noise belong naturally to the statistical world. Real life images clearly has components from both worlds. Objects moving in front of textured backgrounds and borders between textured areas provide obvious examples. A fusion of appropriate parts from both worlds will potentially provide a much more powerful image analysis model. This is the spirit in which the remainder of this paper should be read.

2. What, exactly, is orientation and motion

There is a strong correspondence between the problems of estimating velocity and estimating signal orientation. If the signal is band-limited so as to not contain frequencies above the Nyquist limit the problems are in fact identical. For the case of constant illumination this identity is manifested in the Fourier domain by that all non-zero values can be found on a plane through the origin. The normal to the plane, $\hat{\boldsymbol{m}}$, is directly related to the velocity through:

$$\boldsymbol{v} = \frac{\boldsymbol{P}_x \hat{\boldsymbol{m}}}{\boldsymbol{P}_t \hat{\boldsymbol{m}}}, \tag{1}$$

where $\boldsymbol{P}_x$ projects $\hat{\boldsymbol{m}}$ onto the spatial frequency plane, $\boldsymbol{P}_t$ projects $\hat{\boldsymbol{m}}$ onto the temporal frequency axis.

2.1. Invariances and images of reality

In a Newtonian world the true motion and orientation of a rigid object is a well defined entity that is obviously independent of the visual appearance of the object itself. When orientation and/or velocity is estimated using images it is, however, equally obvious that the properties of, for example, the imaging device, the light sources and the object surface directly influence the transfer of pertinent information, see e.g. [16]. For this reason, a fundamental part of any estimation method is the incorporation of appropriate invariances. The implications of a number of important invariances are discussed below.

2.2. Event identity and estimate locality

For a single highly localized feature it is, all else being equal, desirable to maximize spatial locality of the feature estimate. Retaining feature identity also requires that the estimate is smoothly varying and centered on the feature.[1] These two requirements counteract each other in a fundamental way and there exist many possibilities to define

[1] It would be reasonable to require a unimodal response but we will not do so as it may be perceived as giving the quadrature approach to much of an advantage.

a measure of goodness. Reasonable definitions will, however, produce similar results. We have decided to use the traditional uncertainty product as a quality measure, [12], as this product is relatively shape tolerant and yet severely punishes large deviations from the desired behavior.

2.3. Sample shift invariance

In most cases the signal and the sampling process are not synchronized and it is natural to require that estimates are insensitive to the precise space-time position of the sample grid. For a properly band-limited signal a shifting of the signal by $\boldsymbol{\Delta}_{\boldsymbol{x}}$ can be obtained by multiplying the Fourier transform $F(\boldsymbol{u})$ by $e^{-i\boldsymbol{u}^T\boldsymbol{\Delta}_x}$. For test purposes a signal with spectrum $\cos(u/2)$ is used. This is the most spatially concentrated band-limited signal there is, [12]. The upper left part of Fig. 5 shows the corresponding spatial signal computed at eight different positions over a period of one pixel (one for each row). This signal will efficiently reveal any sampling shift dependencies in the signal processing.

2.4. Invariance to the 'signal cross section'

Estimates of object velocity should be invariant to the image of the object itself. A generalization of this statement is that for a spatio-temporal signal that is constant on lines oriented by $\hat{\boldsymbol{m}}$, i.e. $s(\boldsymbol{x}) = f([\boldsymbol{I} - \hat{\boldsymbol{m}}\hat{\boldsymbol{m}}^T]\boldsymbol{x})$ the estimate of $\hat{\boldsymbol{m}}$, should be invariant to the 'signal cross section' $g(\cdot)$. This seemingly harmless and simple requirement has far reaching consequences. It is equivalent to the statement that the estimate of the normal to the non-zero plane in the Fourier domain should be invariant to what the signal looks like *in* the plane. To make the argument as clear as possible we will discuss the 2-dimensional case where the non-zero plane reduces to a line.

For the estimate to be invariant to the signal on the line it is required that the ratio between the filters involved are the same everywhere on the line. This must hold for all velocities, i.e. line orientations, implying that the Fourier domain ratios between the filters involved can depend only on orientation. More precisely the filter set can be expressed as: $F_k(\boldsymbol{u}) = D_k(\hat{\boldsymbol{u}})G(\boldsymbol{u})$. It is simple to see that using filters that are *polar separable* in the Fourier domain will solve this problem, i.e. $F_k(\boldsymbol{u}) = D_k(\hat{\boldsymbol{u}})G(\rho)$. The generalization of this argument to higher dimensional signals is straightforward.

Perhaps the simplest example of a polar separable filter set is the Gaussian derivative filters used in traditional optical flow analysis. In the Fourier domain these filters are given by

$$F_k(\boldsymbol{u}) = u_k G(\rho) = \rho G(\rho)\cos(\varphi_k), \tag{2}$$

where ρ is the radius in the frequency domain, φ_k the angle between $\boldsymbol{u}$ and the u_k axis and $G(\rho)$ is a Gaussian.

In a 3D spatio-temporal signal the orientational part of these filters are spherical harmonics[2] of *order* 1 as $\cos(\varphi_k)$ can be expressed in this basis, see also Section 4.3, [18]. The use of only order 1 functions limits the performance of the velocity estimation. Furthermore the implied radial function, $(R(\rho) = \rho G(\rho))$, have a number of undesirable consequences that can be avoided by using the proposed loglets filter set. (See Sections 2.5 and 3.2).

2.5. Invariance to illumination

Many image processing tasks require that the analysis is insensitive to spatio-temporally varying lighting conditions. In particular, it is natural to make the analysis invariant to the following two properties when object motion estimation is the task:

- Slowly varying mean level.
- Slowly varying signal amplitude.

Note that, in the present context, this is in sharp contrast to the optical flow approach which sets out to estimate the local spatio-temporal structure

[2]Spherical harmonics allow rotation invariant formulations, [18]. A general n:th order polynomial can be expressed by spherical harmonics of order 0 to n on the unit sphere. Spherical harmonics exist for spaces of any dimension. In 2D spherical harmonics correspond to the common $\cos(n\varphi)$ and $\sin(n\varphi)$ functions. The number of spherical harmonics of order n are $(2n+1)$ in 3D and $(n+1)^2$ in 4D.

of the image sequence itself and not the object motion. Standard optical flow estimates are invariant to global mean level and global signal amplitude but not to variations in these. The analysis can be made invariant to certain classes of such variations by the use of suitable filters. We will in the following assume that changes in illumination can be modeled as an addition of a low-order spatio-temporal polynomial to the image of the moving object.[3] Invariance to illumination can then be obtained by using filters that are orthogonal to the subspace spanned by the polynomial basis.

3. Loglets—vector valued filters

We here introduce a new set of filters termed *loglets* which are suitable for the purposes discussed above and in addition have a number of other useful properties which are discussed in Sections 3.2 and 3.3. Loglets have distinct similarities to the filter-banks introduced by Knutsson [19], and are polar separable in the Fourier domain as opposed to e.g. Gabor filter-bank approaches [13,1].

The definition of loglets is given by the product of a scalar radial part and a vector valued directional part.

$$\boldsymbol{Q}_{sk}(\boldsymbol{u}) = R_s(\rho)\boldsymbol{D}_k(\hat{\boldsymbol{u}}), \quad \rho = \|\mathbf{u}\|. \tag{3}$$

The subscript s indicates the scale and k indicates the orientation of the individual loglet.

3.1. Multi-dimensional directional part

The directional part can be defined for multi-dimensional signals and consist of a set of vector functions, $\boldsymbol{D}_k(\hat{\boldsymbol{u}})$ given by

$$\boldsymbol{D}_k(\hat{\boldsymbol{u}}) = \begin{pmatrix} 1 \\ \hat{\boldsymbol{u}} \end{pmatrix} (\hat{\boldsymbol{u}}^{\mathrm{T}}\hat{\boldsymbol{n}}_k)^{2a}, \tag{4}$$

where $\hat{\boldsymbol{u}}$ is a unit vector in the frequency domain, $\hat{\boldsymbol{n}}_k$ is a filter directing unit vector, $a \geqslant 0$ is an integer setting the directional selectivity of the loglet. (Half integers may be used but will imply an alternate phase definition.)

The last part of Eq. (4) define the direction and selectivity of the angular envelope of the filter. The first part define angular modulations corresponding to spherical harmonics of order zero (the '1') and one ($\hat{\boldsymbol{u}}$), see Fig. 2. This approach enables a straightforward generalization of the phase concept, see Section 3.3.

The realization of the envelope part will require spherical harmonics of even order from $[0,\ldots,2a]$ (odd order $[1,\ldots,2a]$ if a is half-integer). Consequently the required set of directional filters span a spherical harmonics space of order $2a+1$. Thus the individual filters can simply be constructed as a weighted sum of spherical harmonics basis functions. In this way filtering results for a large number of orientations can be obtained in a highly efficient way. Also note that the odd part of $\boldsymbol{D}_k(\hat{\boldsymbol{u}})$ corresponds to the Hilbert transform in the 1-dimensional case and the Riesz transform for higher dimensions [30].

3.2. Radial part

The radial function set, $R_s(\rho)$, is given by

$$R_s(\rho) = f_{s+1}(\rho) - f_s(\rho),$$

$$f_s(\rho) = \tfrac{1}{2}\mathrm{erf}[\alpha \log(\beta^{s-\frac{1}{2}}\rho/\rho_0)], \tag{5}$$

where s is an integer defining the scale of the filter, $\beta > 1$ sets the relative ratio of adjacent scales, α determines the filter shape and overlap.

This function set was designed in the spirit of the wavelet approach and has features particularly interesting in this context: (1) $\frac{\partial^n}{\partial\rho^n} R_s(\rho) = 0$; $\forall n$ or in wavelet vocabulary $R_s(\rho)$ has an infinite number of vanishing moments, [25]. (2)The limit $\beta \to 1$ yields a lognormal filter [19]. (3) The sum of $R_s(\rho)$ over all scales is constant for all non-zero frequencies. For a finite number of consecutive scales the sum is given by: $R_{\mathrm{sum}}(\rho) = f_{s_{\max}}(\rho) - f_{s_{\min}}(\rho)$. i.e. the sum will tend to be flat inside the 'pass-band' range. This is the main advantage over a lognormal filter set.

The latter makes the continuous *loglets* orthogonal to all polynomials indicating the suitability

[3] This simplifies the following discussion but for this assumption to be reasonable the logarithm of the image signal should be used.

for the present purpose. In practice discrete loglets will have to be used naturally limiting the number of vanishing moments. Discrete loglets can be optimized using multiple subspace criteria as outlined in [23]. Fig. 1 shows $R_s(\rho)$ for a sequence of scales.

3.3. Generalized quadrature and phase

Applying the loglets to a d-dimensional signal will produce a $(d+1)$-dimensional vector response signal, $\boldsymbol{q}$, for each loglet. The amplitude, q, of the response is simply defined as the norm of the response and the generalized phase, $\boldsymbol{\theta}$, as the normalized response, i.e.

$$q = \|\boldsymbol{q}\| \quad \text{and} \quad \boldsymbol{\theta} = \hat{\boldsymbol{q}} \quad \text{giving } \boldsymbol{q} = q\hat{\boldsymbol{\theta}}. \tag{6}$$

This definition of phase is identical to the definition due to Knutsson presented in [15,14]. For 1-dimensional signals it reduces to the classical definition of amplitude and phase: $\boldsymbol{f} = f e^{i \arg(f)} = f\hat{\boldsymbol{\theta}}$ [3]. More recent related work can be found in, e.g. [10]. The construction of classical quadrature filters, [19], is based on the analytic signal and requires a pre-defined filter direction. The generalized quadrature approach elegantly removes this requirement.

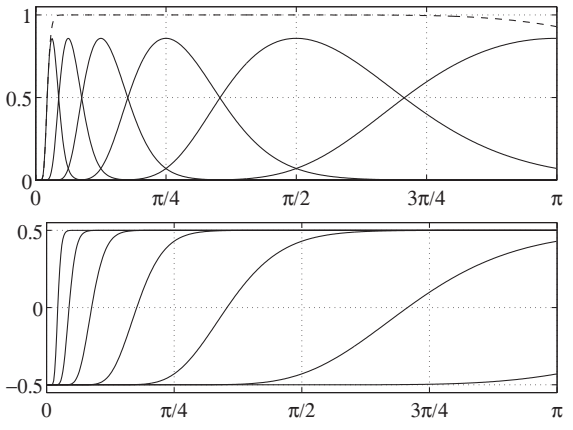

Fig. 1. Top: $R(\rho)$ for six loglets separated by one octave and the sum, $\rho_0 = \pi$, $\beta = 2$, $\alpha = 3$ and $s = [0, 1, \ldots, 5]$. Bottom: the radial parts are computed as differences of the erf-functions shown, see Eq. (5).

The generalized phase concept also naturally lends itself to a further extension. An order n phase can be defined by including spherical harmonics of order $0-n$ in the vector part of Eq. (4). This will increase the descriptive power but will also decrease the spatial resolution and significantly increase the number of filters involved.

4. The local structure tensor

Representation of orientation has had a long standing central position in the development of the image processing framework of today. In 1978 Granlund suggested a representation for 2-dimensional orientation, [13]. The requirements for representing 3-dimensional orientation were discussed and outlined by Knutsson 1985 [20]. In 1987 this work lead to the formulation of the local structure tensor approach, [21,22], that is now common practice.

In order to take the quadrature/gradient comparison one step further we will use the Gaussian derivative outer product matrix suggested by Bigün and Granlund [2].[4] Although the two approaches are fundamentally different they end up using the same orientation representation which allows for a direct comparison.

In the following, for reasons that will soon become apparent, a slightly modified form using the l_2 norm of the filter responses will be used:

$$\boldsymbol{T} = \sum_{k=1}^{K} \|\boldsymbol{q}_k\|^2 \boldsymbol{M}_k, \tag{7}$$

where $\boldsymbol{q}_k$ are the responses from single scale loglets, $\boldsymbol{M}_k$ are the filter orientation tensors, [22,14].

4.1. Direct tensor estimation from spherical harmonics basis

As mentioned in Section 3.1 the spherical harmonics basis allows the number of orientations, K, used in Eq. (7) to be increased with only

[4] In this work orientation representation is not the issue, the approach is based on a least squares problem formulation and no mention of tensors is made.

a minor additional computational cost as the result for any orientation can be computed from the responses of the basis filters. An interesting observation can be made if the number of orientations is increased to the limit, $K \to \infty$. The equivalent of Eq. (7) for this case is given by the integral:

$$\boldsymbol{T} = \int_{\hat{\boldsymbol{n}}} \|\boldsymbol{q}(\hat{\boldsymbol{n}})\|^2 \boldsymbol{M}(\hat{\boldsymbol{n}})\, \mathrm{d}\hat{\boldsymbol{n}}, \tag{8}$$

where $\boldsymbol{q}(\hat{\boldsymbol{n}})$ is the response from a loglet in the $\hat{\boldsymbol{n}}$ direction, $\boldsymbol{M}(\hat{\boldsymbol{n}})$ the dual tensor corresponding to $\hat{\boldsymbol{n}}$.

Without explicitly going through the calculations it can be realized that computing $\boldsymbol{T}$ from Eq. (8) will result in weighted sets of sums of products of the basis filter outputs, one sum for each tensor component, t_n, i.e.

$$t_n = \sum_{j=1}^{J} \sum_{l=1}^{j} a_{njl} H_j(\hat{\boldsymbol{u}}) H_l(\hat{\boldsymbol{u}}), \tag{9}$$

where J is the total number of spherical harmonics functions used, a_{njl} the weighting coefficient for filter product (j, l) and tensor component n, $H_j(\hat{\boldsymbol{u}})$ the j:th spherical harmonics function.

Note that the phase invariant property of the quadrature filter based tensor, of Eq. (7) is preserved in Eq. (9) for $J \geqslant 2\alpha + 1$ although the individual spherical harmonic filters are not phase invariant. Further, Eq. (8) can be seen as a projection onto $\boldsymbol{M}(\hat{\boldsymbol{n}})$. As $\boldsymbol{M}(\hat{\boldsymbol{n}})$ only has zeroth and second-order components only filter products containing order 0 and/or order 2 components will contribute to the sum in Eq. (9). In fact most of the coefficients, a_{njl}, will be equal to zero.

4.2. Direct tensor estimation in 2D

It is helpful to have a more detailed look at the 2D case. Here it is convenient to express the spherical harmonics basis filters of order $0-3$ as complex functions in the Fourier domain:

$$\boldsymbol{H}_j = R(\rho) \mathrm{e}^{\mathrm{i}j\varphi}; \quad -3 \leqslant j \leqslant 3, \tag{10}$$

where ρ, φ are polar coordinates in the Fourier domain.

Writing $\boldsymbol{T}$ as

$$\boldsymbol{T} = \begin{pmatrix} t_{11} & t_{12} \\ t_{12} & t_{22} \end{pmatrix} \tag{11}$$

and denoting the complex output from filter $\boldsymbol{H}_j$ by $\boldsymbol{h}_j$ it turns out that $\boldsymbol{T}$ is given by

$$\begin{aligned} t_{11} &= \frac{1}{24}(\Re[6\boldsymbol{h}_0\boldsymbol{h}_2 - 3\boldsymbol{h}_1^2 + 3\boldsymbol{h}_3\boldsymbol{h}_{-1}] \\ &\quad + 4|\boldsymbol{h}_0|^2 + 5|\boldsymbol{h}_1|^2 + 2|\boldsymbol{h}_2|^2 + |\boldsymbol{h}_3|^2), \\ t_{22} &= \frac{1}{24}(-\Re[6\boldsymbol{h}_0\boldsymbol{h}_2 - 3\boldsymbol{h}_1^2 + 3\boldsymbol{h}_3\boldsymbol{h}_{-1}] \\ &\quad + 4|\boldsymbol{h}_0|^2 + 5|\boldsymbol{h}_1|^2 + 2|\boldsymbol{h}_2|^2 + |\boldsymbol{h}_3|^2), \\ t_{12} &= \frac{1}{24}(\Im[6\boldsymbol{h}_0\boldsymbol{h}_2 - 3\boldsymbol{h}_1^2 + 3\boldsymbol{h}_3\boldsymbol{h}_{-1}]). \end{aligned} \tag{12}$$

The computation is straightforward but cumbersome and is for that reason left out in this presentation. Note that the sum of the indices of the filters making up the output products involved is either 2 (for the terms in the square brackets) or 0 (for the rest). Also note that the filters with negative indices are given by $\boldsymbol{H}_{-j} = \boldsymbol{H}_j^*$ implying that these responses need not be separately computed.

4.3. Derivatives, spherical harmonics and quadrature

In the traditional gradient approach the local structure is estimated by the Gaussian derivative outer product matrix. In loglet terms this implies that the radial part of the filter functions is set to $\rho G(\rho)$ and only the first-order spherical harmonics filters, $\mathbf{h}_1$ in Eq. (12), is used. In fact derivative operators of any order are polar separable in the Fourier domain having order n spherical harmonics as the directional part and ρ^n as the radial part. This observation provides links to a large number of traditional image processing approaches and in particular to the work of Danielsson et al. e.g. [6,7].

The concept of quadrature is however in direct conflict with the use of derivative operators. To attain a quadrature response all filters making up

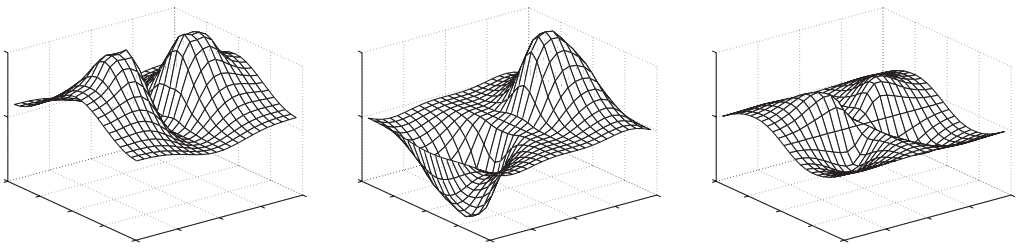

Fig. 2. A 2D loglet i the x-direction in the Fourier domain. From left: orientation selective envelope and the envelope modulated by $\cos(\varphi)$ and $\sin(\varphi)$.

the response must have the same radial function and the combined sensitivity must be the same for even and odd signals. As odd and even order derivatives necessarily have different radial behavior this cannot be accomplished by combining outputs from derivative operators.[5] That Eq. (12) will produce quadrature responses for all tensor components is shown by the fact that the sum of the magnitude of the coefficients are equal for the odd and the even filter products.

5. Local certainty estimation

Certainty estimates play an important role in many image processing applications. Certainty can be interpreted as a measure of how well the present signal fits the adopted signal model. In the following we will refer to the part of the signal that is not explained by the signal model as the residue. A measure of residue magnitude is only possible to attain when the measured properties of the signal space contain more information than is required by the signal model. A classical example of relative residue estimation is the certainty estimate based on the eigenvalues of $\boldsymbol{T}$: $c_\lambda = \frac{\lambda_1 - \lambda_2}{\lambda_1}$ where $\mathbf{T} = \lambda_1 \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1^{\mathrm{T}} + \lambda_2 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_2^{\mathrm{T}}$, see Fig. 3. Here $\lambda_1 \gg \lambda_2 \approx 0$ corresponds to a neighborhood with a well defined orientation without noise while and an isotropic neighborhood is characterized by $\lambda_1 \approx \lambda_2$. Another example of how the residue can manifest itself is given by the behavior of loglet phase.

[5] Using BP-pyramids these problems will be reduced which is not surprising as doing so will, in fact, yield resulting filters that are more 'loglet like'.

5.1. Phase consistency based certainty

A new measure of certainty estimate can be obtained using a generalized definition of phase.

$$c_{\boldsymbol{\theta}} = \frac{\sum_k q_k \boldsymbol{\theta}_k}{\sum_k q_k} = \frac{\|\sum_k \boldsymbol{q}_k\|}{\sum_k \|\boldsymbol{q}_k\|}. \tag{13}$$

For simple signals, i.e. signals that can be expressed as $s(\boldsymbol{x}) = g(\hat{\boldsymbol{n}}^{\mathrm{T}}\boldsymbol{x})$, [14], the loglet phase will be the same for all $\boldsymbol{q}_k$ making the certainty estimate $c_{\boldsymbol{\theta}}$ equal to unity. Local deviations from a simple signal will tend to cause a varying phase and, as a consequence, produce a smaller value.

Fig. 3 shows an example of certainty estimation. Here $\boldsymbol{q}$ is defined by the loglet shown in Fig. 2. The generalized order one phase representation comprise a $\cos^2(\varphi)$ angular window (left in Fig. 2) which is modulated by $\cos(\varphi)$ and $\sin(\varphi)$, (middle and right). Using 12 orientations the phase invariant local structure tensor, Eq. (7), and the phase consistency based certainty, $c_{\boldsymbol{\theta}}$ in Eq. (13), was estimated. The left part in Fig. 3 shows the test image. The middle image shows the traditional certainty based on the eigenvalues of $\boldsymbol{T}$: $c_\lambda = (\lambda_1 - \lambda_2)/\lambda_1$. The right image shows the phase consistency based certainty $c_{\boldsymbol{\theta}}$. Note that certainty estimates c_λ and $c_{\boldsymbol{\theta}}$ clearly have different properties picking up different parts of the local 'non-simplicity'.

5.2. Certainty and the signal representation manifold

A more general approach to certainty estimation is to view the representation space as a manifold embedded in the signal space defined by the set of basis filters used. Estimating certainty

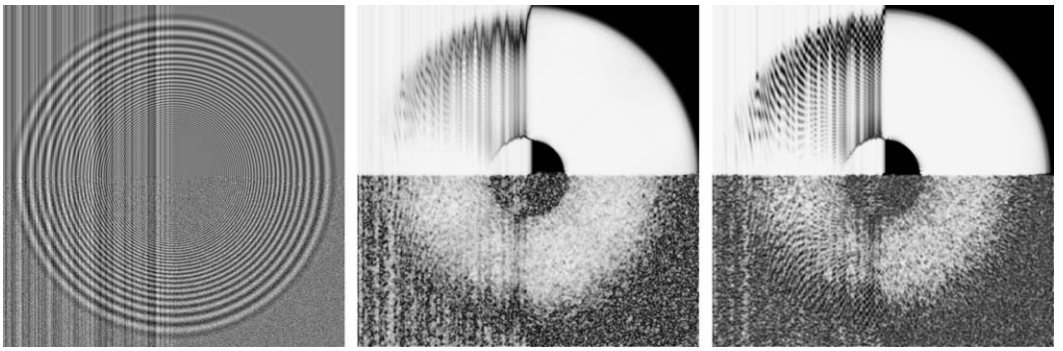

Fig. 3. Certainty estimation example. Left: test image comprising three overlayed structures, a circular modulated sine wave pattern, a random vertical stripe pattern in the left part and white noise in the lower half (may be less visible due to the print resolution), making up four differently structured image quadrants. Middle: orientation consistency $c_\lambda = (\lambda_1 - \lambda_2)/\lambda_1$. Right: phase consistency based certainty c_θ.

then becomes a task of measuring the distance to the representation manifold, i.e. the residue. The main advantage of the representation manifold view is that the intrinsic dimensionality of the residue can be found. Suppose there are J filters and L different independent ways to change an ideal signal in a way that can be seen by the filters. The intrinsic dimensionality of the ideal signal manifold, and of any consistent representation, will then be L. The intrinsic dimensionality of the residue will be $J - L$. In the 2D case the signal representation manifold is described by the following four constraints:

$$|\boldsymbol{h}_0| - |\boldsymbol{h}_2| = 0,$$

$$|\boldsymbol{h}_1| - |\boldsymbol{h}_3| = 0,$$

$$|\boldsymbol{h}_1^2 + \boldsymbol{h}_0\boldsymbol{h}_2| = |\boldsymbol{h}_1^2| + |\boldsymbol{h}_0\boldsymbol{h}_2|,$$

$$|\boldsymbol{h}_1^2 + \boldsymbol{h}_3\boldsymbol{h}_1^*| = |\boldsymbol{h}_1^2| + |\boldsymbol{h}_3\boldsymbol{h}_1^*|. \tag{14}$$

The intrinsic dimensionality of simple signals viewed through loglets is $n + 1$. The joint intrinsic dimensionality of an ideal (rank 1) local structure tensor and the local phase is also $n + 1$. The dimensionality of the loglet space (using $\alpha = 1$ in Eq. (4)) is 7 in 2D and 16 in 3D. This gives residue intrinsic dimensionalities of 4 in 2D and 12 in 3D. The higher this number the better the estimates of the residue will be from a statistical point of view.

Using only first-order spherical harmonics, e.g. first-order derivatives, implies that n filter will be used. This will exactly suffice to represent orientation and magnitude. Neither signal phase nor deviations from the ideal signal case can be estimated locally. In other words, all signal are seen as ideal. It should be noted, however, that averaging of local estimates remedies this drawback to some extent.

5.3. Certainty weighted tensor averaging

Using loglets local certainty estimates can be used to improve the orientation estimates. Estimating certainty is preferably done based on the discussion in Section 5.2. The estimated residue can be thought of as local noise. Traditional estimation theory states that if several estimates of the same signal are available they should be weighted together using weights inversely proportional to the corresponding noise variances. This is the background for suggesting local certainty, c, to be computed in practice as

$$c = \frac{\gamma(E_{\text{tot}} - E_{\text{res}})}{E_{\text{res}} + \gamma E_{\text{tot}}}, \tag{15}$$

where E_{tot} is the total signal energy picked up by the filter set used, E_{res} is the energy of the residue, γ is a factor related to how much the actual filters used deviate from the ideal filter shape.

This certainty estimate will range from 0 to 1. The sensitivity of the estimate can be adjusted through γ.

6. Experiments and locality evaluation

Commonly used local structure analysis algorithms are quite complex and involve a lot more than the filters used. The effects of the filters cannot be separated from the rest of the analysis and this makes comparisons difficult to interpret from a filter point of view. In addition different approaches can be expected to be preferable in different situation. To reduce the number 'free' parameters and target the filter design aspects a number of simple 2D experiments have been carried out. The experiments in this section were designed to demonstrate the high locality of the loglet approach as well as the performance in a few simple situations. The intention is obviously not to prove that our approach is always better but simply to indicate the nature of the qualities gained.

The first experiment is based on the test signal described in Section 2 which is shown in the upper left part of Fig. 5. The local energy is computed using both gradient and quadrature filters. The quadrature filters were designed to mimic the gradient filter for positive frequencies, see Fig. 4. The dotted LP-filter is used to average the square of the gradient response. Fig. 5 shows the results. The spatial standard deviation, σ, of the low-pass filter is chosen such that the variance of the spectrum is equal for the quadrature and gradient filters, see Table 1. Note that this σ is a bit too

Table 1
Variances and uncertainty products for the quadrature and gradient responses in Fig. 5

	Small filters			Large filters		
	σ_x	σ_u	$\sigma_x\sigma_u$	σ_x	σ_u	$\sigma_x\sigma_u$
Loglets	0.894	0.566	0.506	1.672	0.300	0.502
Gradient	1.106	0.567	0.627	2.224	0.300	0.667

Theoretical uncertainty min value is 0.5. The table refer to the filters in Fig. 4.

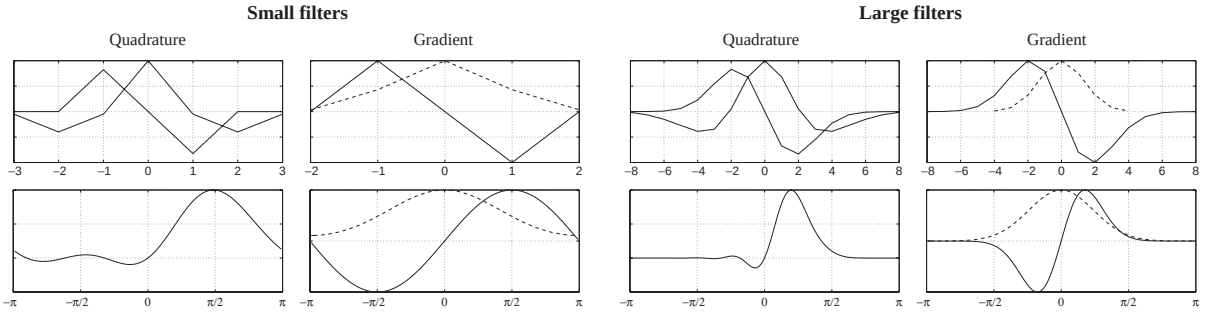

Fig. 4. Compact support quadrature and gradient filters in the spatial domain (top row) and the Fourier domain (bottom row). The dotted LP-filter is used to average the square of the gradient response.

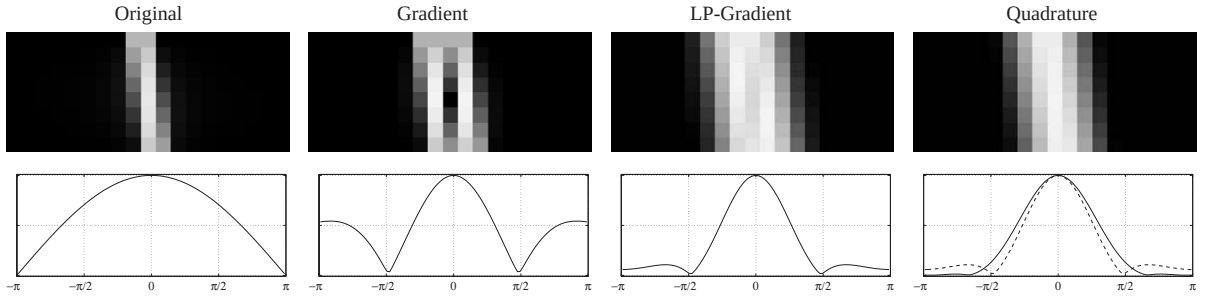

Fig. 5. Test signal and spatial results for eight different positions (top row) and corresponding spectra (bottom row). The LP-gradient spectra is repeated (dotted) in the quadrature plot for comparison.

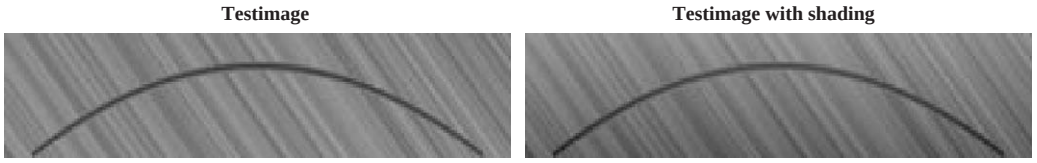

Fig. 6. Left: 2-dimensional spatio-temporal test image of an accelerating small dark object in front of a constant velocity background. Right: the same image with a slight vertical shading added.

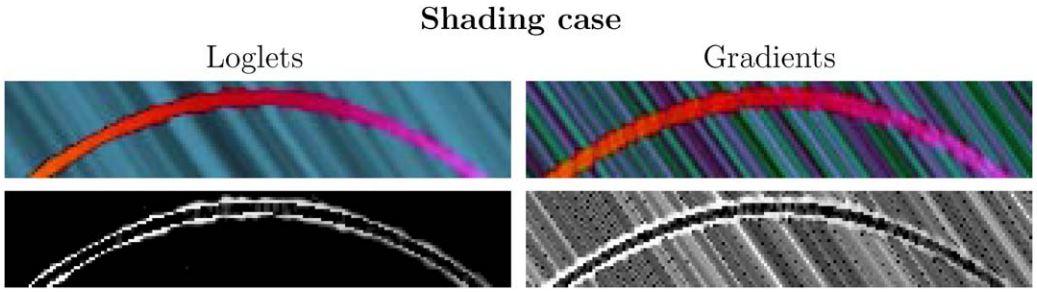

Fig. 7. Result and orientation errors for the shaded spatio-temporal test image. The very small shading markedly increases the error of the gradient based estimate but the loglet estimate seems to be unaffected.

small to completely remove the two peaks in the gradient image.

From Fig. 5 and the uncertainty products in Table 1 it is obvious that the localization of the quadrature filters are superior compared to the gradient filters. The experiment was carried out in two scales and in both experiments the width in the spatial domain (and the uncertainty product) is more than 25% smaller for the quadrature filters. The apparent aliasing for the small gradient filter in Fig. 5 may suggest that the performance of the gradient method can be improved by oversampling. The results for the large filters in Table 1 show that this is not the case.

6.1. Orientation/velocity estimation

The second experiment comprises a 2D spatio-temporal test signal where a small dark object is moving in front of a moving background, Fig. 6 left. Also, for test purposes, a second test image was constructed by adding a slight vertical shading (barely visible) to this image, Fig. 6 right.

The velocity (or orientation) is estimated by both quadrature filters and gradient filters. The gradient filters in Fig. 4 where used but the quadrature filters are now loglets. Both for the gradient outer product matrix and the quadrature tensors the orientation errors where estimated as

$$\Delta\phi = \sin^{-1}\sqrt{\frac{1}{2N}\sum_{l=1}^{N}\|\hat{\boldsymbol{n}}\hat{\boldsymbol{n}}^{\mathrm{T}} - \hat{\boldsymbol{e}}_1\hat{\boldsymbol{e}}_1^{\mathrm{T}}\|^2}, \tag{16}$$

where $\hat{\boldsymbol{e}}_1$ is the eigenvector belonging to the largest eigenvalue, $\hat{\boldsymbol{n}}$ defines the true velocity (orientation) of the image, N is the number of instances summed.

The results are displayed in Figs. 7 and 8. In Fig. 7 it is apparent that the slight shading present in the signal misleads the estimation of the velocity using the gradient approach. Without shading the performances are more comparable but the quadrature approach still outperforms the gradient approach. This is mainly due to better localization properties as can be studied in Fig. 8.

In Fig. 9 plots of estimation errors are shown for a range of tensor LP-filter sizes.

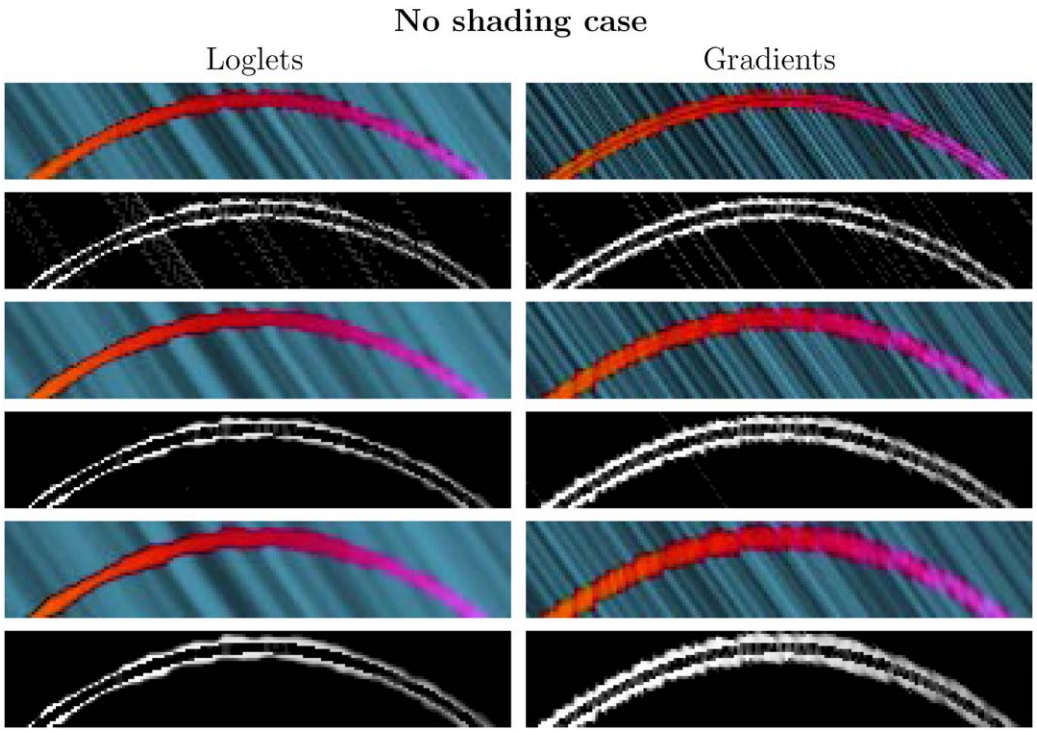

Fig. 8. Results and errors for the spatio-temporal test image without shading. Results for three different sizes of tensor LP-filters are shown, the spatial standard deviations are: top 0.4, middle 0.8, bottom 1.2. All of the loglet results are still better than the gradient results.

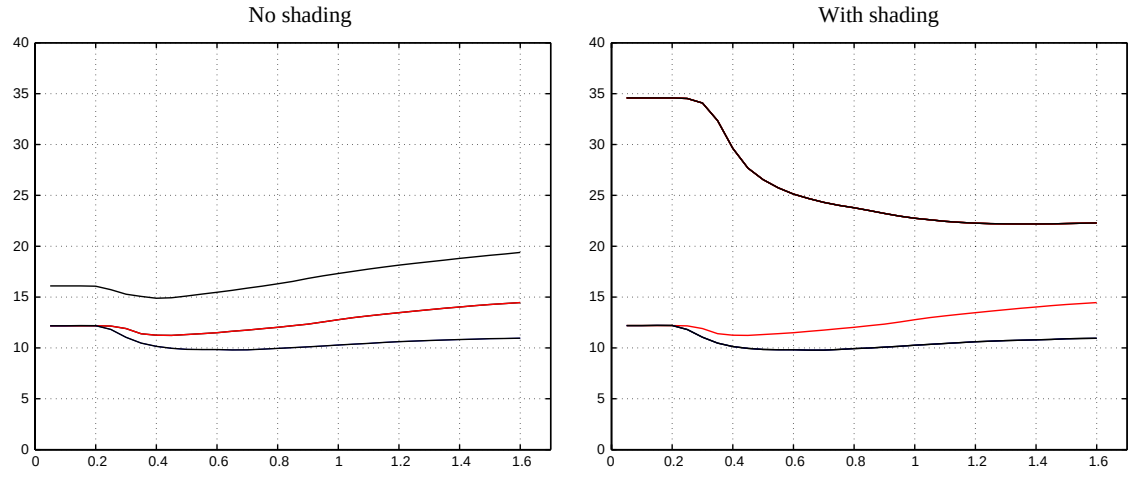

Fig. 9. Plots showing estimation errors (in degrees) for a range of tensor LP-filter sizes. The top curves are the gradient based errors, the middle curves are the loglet errors and the bottom curves are loglet errors using certainty weighted tensor averaging.

The top curves are the gradient based errors, the middle curves are the loglet errors and the bottom curves show the improvement that can be attained by using certainty weighted averaging of the tensors using c from Eq. (15) ($\gamma = 0.01$).

Acknowledgements

The authors wishes to thank the medical image science team in Linköping, in particular Ola Friman for suggesting to try and solve the integral in Eq. (8). This work is part of research carried out within The Center for Medical Informatics Image Science and Visualization (CMIV) and the European Network of Excellence *SIMILAR*. The work was supported by The center for non invasive medical measurements NiMed, Amersham Health Norway, The Swedish Agency for Innovation Systems (VINNOVA) and The Swedish Research Council (VR).

References

[1] J. Bigun, Speed, frequency, and orientation tuned 3-D gabor filter banks and their design, in: Proceedings of International Conference on Pattern Recognition, ICPR, Jerusalem, IEEE Computer Society, Silver Spring, MD, 1994, pp. C-184–187.

[2] J. Bigün, G.H. Granlund, Optimal orientation detection of linear symmetry, in: Proceedings of the IEEE First International Conference on Computer Vision, London, Great Britain, 1987, pp. 433–438.

[3] R. Bracewell, The Fourier Transform and its Applications, second ed., McGraw-Hill, New York, 1986.

[4] A.D. Calway, H. Knutsson, R. Wilson, Multiresolution estimation of 2-d disparity using a frequency domain approach, in: Proceedings of the British Machine Vision Conference, Leed, UK, 1992.

[5] T.A.C. Claasen, W.F.G. Mecklenbrauker, The Wigner distribution—a tool for time-frequency signal analysis, part ii: Discrete-time signals, Philips J. Res. 35 (4/5) (1980) 276–300.

[6] P.E. Danielsson, Rotation invariant operators with directional response, in: Proceedings of the Fifth International Conference on Pattern Recognition, Miami Beach, Florida, 1980.

[7] P. Danielsson, O. Seger, Rotation invariance in gradient and higher derivative detectors, Comput. Vision Graphics Image Process. 49(2).

[8] I. Daubechies, The wavelet transform, time-frequency localization and signal analysis, IEEE Trans. Inform. Theory 36 (5) (1990) 961–1005.

[9] J.G. Daugman, Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters, J. Opt. Soc. Amer. 2 (1985) 1160.

[10] M. Felsberg, G. Sommer, Image features based on a new approach to 2d rotation invariant quadrature filters, in: A. Heyden, G. Sparr, M. Nielsen, P. Johansen (Eds.), Computer Vision—ECCV 2002, Lecture Notes in Computer Science, vol. 2350, Springer, Berlin, 2002, pp. 369–383.

[11] D.J. Fleet, A.D. Jepson, Stability of phase information, IEEE Trans. Pattern Anal. Mach. Intell. 15 (12) (1993) 1253–1268.

[12] D. Gabor, Theory of communication, J. Inst. Elec. Engrg. 93 (26) (1946) 429–457.

[13] G.H. Granlund, In search of a general picture processing operator, Comput. Graphics Image Process. 8 (2) (1978) 155–178.

[14] G.H. Granlund, H. Knutsson, Signal Processing for Computer Vision, Kluwer Academic Publishers, Dordrecht, 1995 ISBN 0-7923-9530-1.

[15] L. Haglund, H. Knutsson, G.H. Granlund, On phase representation of image information, in: The Sixth Scandinavian Conference on Image Analysis, Oulu, Finland, 1989, pp. 1082–1089.

[16] B.K.P. Horn, Robot Vision, The MIT Press, Cambridge, MA, 1986.

[17] A. Hueckel, An operator which locates edges in digital pictures, J. ACM 20 (1971) 113–125.

[18] M.N. Jones, Spherical Harmonics and Tensors for Classical Field Theory, Research Studies Press, 1985.

[19] H. Knutsson, Filtering and reconstruction in image processing, Ph.D. Thesis, Linköping University, Sweden, Diss. No. 88, 1982.

[20] H. Knutsson, Producing a continuous and distance preserving 5-D vector representation of 3-D orientation, in: IEEE Computer Society Workshop on Computer Architecture for Pattern Analysis and Image Database Management—CAPAIDM, IEEE, Miami Beach, Florida, 1985, pp. 175–182, Report LiTH-ISY-I-0843, Linköping University, Sweden, 1986.

[21] H. Knutsson, A tensor representation of 3-D structures, in: Fifth IEEE-ASSP and EURASIP Workshop on Multidimensional Signal Processing, Noordwijkerhout, The Netherlands, 1987, poster presentation.

[22] H. Knutsson, Representing local structure using tensors, in: The Sixth Scandinavian Conference on Image Analysis, Oulu, Finland, 1989, pp. 244–251, Report LiTH-ISY-I-1019, Computer Vision Laboratory, Linkoping University, Sweden, 1989.

[23] H. Knutsson, M. Andersson, J. Wiklund, Advanced filter design, in: Proceedings of the Scandinavian Conference on Image analysis, SCIA, Greenland, 1999.

[24] H. Knutsson, G.H. Granlund, Fourier domain design of line and edge detectors, in: Proceedings of the Fifth International Conference on Pattern Recognition, Miami, Florida, 1980.

[25] R.B. Leipnik, On lognormal random variables: I—the characteristic function, J. Austral. Math. Soc. Ser. B 32 (1991) 327–347.

[26] R. Lenz, Rotation-invariant operators and scale space filtering, Pattern Recognition Lett. 6 (1987) 151–154.

[27] D. Marr, E. Hildreth, Theory of edge detection, in: Proceedings of the Royal Society of London, vol. 207, 1980, pp. 187–217.

[28] A.V. Oppenheim, J.S. Lim, The importance of phase in signals, Proc. IEEE 69 (1981) 529.
[29] D.A. Pollen, S.F. Ronner, Visual cortical neurons as localized spatial frequency filters, IEEE Trans. Systems Man Cybernet. 13 (5) (1983) 907–915.
[30] M. Riesz, Sur les fonctions conjuge'es, Math. Zeit. 27 (1927) 218–244.
[31] L.G. Roberts, Machine perception of three-dimensional solid, in: J.T. Tippell (Ed.), Optical and Electro-Optical Information Processing, MIT Press, Cambridge, MA, 1965, pp. 159–197.
[32] V. Torre, T.A. Poggio, On edge detection, IEEE Trans. Pattern Anal. Mach. Intell. 8 (2) (1986) 147–163.
[33] C.-F. Westin, H. Knutsson, Estimation of Motion Vector Fields using Tensor Field Filtering, in: Proceedings of the IEEE International Conference on Image Processing, IEEE, Austin, Texas, 1994, pp. 237–242.
[34] R. Wilson, A. Calway, A general multiresolution signal descriptor and its application to image analysis, in: Signal Processing, EURASIP, 1988, pp. 663–666.
[35] R. Wilson, H. Knutsson, Uncertainty and inference in the visual system, IEEE Trans. Systems Man Cybernet. 18 (2) (1988) 305–312.
[36] S.W. Zucker, R.M. Hummel, A three-dimensional edge operator, IEEE Trans. Pattern Anal. Mach. Intell. 3 (3) (1981) 324–331.