Loglets: Generalized Quadrature and Phase for Local Spatio-Temporal Structure Estimation

Hans Knutsson and Mats Andersson

Dept. of Biomedical Engineering, Linköping University, SE-581 85 Linköping, Sweden
{knutte,matsa}@imt.liu.se, www.imt.liu.se/mi

Abstract. The question of which properties of a local structure estimator are important is discussed. 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*. Using loglets it is shown how the concepts of quadrature and phase can be defined in n-dimensions. A number of experiments support 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 to variations in object illumination. It is also shown how a measure of the certainty of the estimate can be obtained using the consistency of the generalized phase with respect to orientation.

1 Introduction

The first steps towards analysis of images were taken more than 30 years ago. From the very start detecting edges and lines in images was considered a fundamental operation. Since these early days new and more advanced schemes for analysis of local image structure has been suggested in a seemingly never ending stream. Local image orientation, scale, frequency and phase are prominent examples of features that have been considered central in the analysis.

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 thirty 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 and end.

2 Estimation of 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

J. Bigun and T. Gustavsson (Eds.): SCIA 2003, LNCS 2749, pp. 741–748, 2003.

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, $\hat{\boldsymbol{n}}$, is directly related to the velocity through:

$$\boldsymbol{v} \;=\; \frac{\boldsymbol{P}_{\boldsymbol{x}}\,\hat{\boldsymbol{n}}}{\boldsymbol{P}_t\,\hat{\boldsymbol{n}}} \tag{1}$$

Where: $\boldsymbol{P}_{\boldsymbol{x}}$ projects $\hat{\boldsymbol{n}}$ onto the spatial frequency plane
$\boldsymbol{P}_t$ projects $\hat{\boldsymbol{n}}$ onto the temporal frequency axis

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. [1]. 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.

Fundamental estimate behaviour - 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 a measure of goodness. Reasonable definitions will, however, produce similar results. We have decided to use the traditional uncertainty product as a quality measure, [2], as this product is relatively shape tolerant and yet severely punishes large deviations from the desired behavior.

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 $e^{-i\boldsymbol{u}^T\boldsymbol{\Delta}_{\boldsymbol{x}}}$. The upper left part of fig.3 shows the spatial signal computed at eight different positions over a period of one pixel. The magnitude of the test signal in the FD is shown in the adjacent plot. A $\cos(u/2)$ frequency function is chosen as it is the most spatially concentrated band-limited signal there is, [2]. This signal will efficiently reveal any sampling shift dependencies in the signal processing.

Invariance to the 'signal section' - Estimates of object velocity should be invariant to the image of the object itself. A generalization of this statement is that for a fully oriented signal, i.e. $s(\boldsymbol{x}) \;=\; g(\;[\boldsymbol{I} - \hat{\boldsymbol{n}}\hat{\boldsymbol{n}}^T]\;\boldsymbol{x}\;)$ estimates of the orientation, $\hat{\boldsymbol{n}}$, should be invariant to the 'signal section' $g(\cdot)$. This seemingly harmless and simple requirement has far reaching consequences. It is equivalent to the statement that the estimation of the normal of the non-zero plane in the

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

Fourier domain should be invariant to what the signal looks like *in* the plane. To make the following argument as clear as possible we will discuss the case with only one spatial dimension 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 line orientations implying that the Fourier domain ratios between the filters involved can depend only on orientation. Perhaps the simplest example of such a set of filters is the Gaussian derivative filters used in traditional optical flow analysis which in the Fourier domain is given by:

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

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

We will, however, show that a better set of filters can be found.

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.

The analysis can be made invariant to certain classes of such variations by the use of suitable filters.

Loglets - 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[2]. Invariance to illumination can then be obtained by using filters that are orthogonal to the subspace spanned by the polynomial basis. We here introduce a new set of filters termed *loglets* which naturally lend themselves to this purpose and in addition have a number of other useful properties. Loglets have distinct similarities to the filter-banks introduced by Knutsson, [3], and are polar separable in the Fourier domain as opposed to e.g. Gabor filter-bank approaches [4, 5].

Radial part - The radial function set, $R_s(\rho)$, is given by:

$$R_s(\rho) \;=\; \mathrm{erf}[\alpha \log(\frac{\beta^{s+\frac{1}{2}}}{\rho_0}\;\rho)] - \mathrm{erf}[\alpha \log(\frac{\beta^{s-\frac{1}{2}}}{\rho_0}\;\rho)] \tag{3}$$

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

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

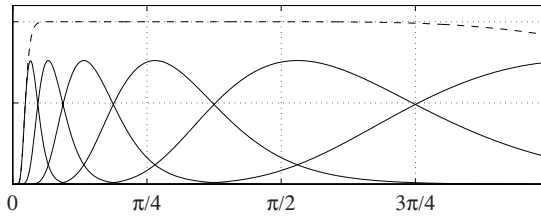

Fig. 1. Six loglets separated by one octave and the sum.

features particularly interesting in this context: **1.** The limit $\beta \rightarrow 1$ yields a lognormal filter [3]. **2.** The sum of $R_s(\rho)$ over all scales is a constant in the Fourier domain (as opposed to a lognormal filter set). **3.** $\frac{\partial^n}{\partial \rho^n} R_s(\rho) = 0;\ \forall n$ or in wavelet vocabulary $R_s(\rho)$ has an infinite number of vanishing moments, [6]. The latter makes the continuous *loglets* orthogonal to polynomials indicating that the suitability for the present purpose. In practise 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 [7]. Figure 1 shows $R_s(\rho)$ for a sequence of scales and the sum.

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:

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

Where: $\hat{\boldsymbol{u}}$ is a unit vector in the frequency domain
$\hat{\boldsymbol{n}}_k$ is a filter directing unit vector
$a \geq 0$ is an integer setting the directional selectivity of the loglet

The vector valued loglet is given by the product of the radial and the directional parts.

$$\boldsymbol{L}_{sk}(\boldsymbol{u}) \ = \ R_s(\rho)\ \boldsymbol{D}_k(\hat{\boldsymbol{u}}) \tag{5}$$

The directional filters span a spherical harmonics space of order $2a + 1$ and 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 the Hilbert transform in the 1-dimensional case and the Riesz transform for higher dimensions, [8].

Generalized Quadrature and Phase - Applying the loglets to a d-dimensional signal will produce an $(d{+}1)$-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.

$$q \ = \ \|\boldsymbol{q}\| \qquad \text{and} \qquad \boldsymbol{\theta} \ = \ \hat{\boldsymbol{q}} \tag{6}$$

This definition of phase is identical to the definition due to Knutsson presented in [9] and [10]. For 1-dimensional signals it reduces to the classical amplitude and phase of the analytic signal [11]. More recent related work can be found in, e.g. [12]. The construction of classical quadrature filters, [3], is based on

the analytic signal and requires a pre-defined filter direction. The generalized quadrature approach elegantly removes this requirement.

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 to n in the vector part of eqn. 4. This significantly increases the number of filter involved but will also provide a proportionally higher descriptive power.

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, [4]. The requirements for representing 3-dimensional orientation were discussed and outlined by Knutsson 1985 [13]. In 1987 this work lead to the formulation of the local structure tensor approach, [14], that is now common practise.

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, [15][3]. Although the two approaches are fundamentally different they end up using the same orientation representation which allows for a direct comparison. To make the comparison as fair as possible a slightly modified form involving squared quadrature filter magnitudes is used:

$$\boldsymbol{T} = \sum_k \|\boldsymbol{q}_k\|^2 \; \boldsymbol{M}_k \tag{7}$$

Where: $\boldsymbol{q}_k$ are the responses from single scale loglets
$\boldsymbol{M}_k$ are the filter orientation tensors, [16, 10]

It should be noted, however, that the difference in the tests below due to this modification is, in fact, completely insignificant.

Phase consistency based certainty - Certainty measures have played an important role in many image processing applications. A new measure of certainty in orientation can be obtained using generalized phase.

$$c_{\boldsymbol{\theta}} \;=\; \frac{\|\sum_k \boldsymbol{q}_k\|}{\sum_k \|\boldsymbol{q}_k\|} \tag{8}$$

Ideally the phase should be the same for all $\hat{\boldsymbol{n}}_{\mathbf{k}}$ making $c_{\boldsymbol{\theta}} = 1$. A varying phase will produce a smaller value. This certainty measure captures information that differs from the traditional certainty based on the eigenvalues of $\boldsymbol{T}$: $c_\lambda = \frac{\lambda_1 - \lambda_2}{\lambda_1}$, see section 3 and fig.6.

3 Results

The first experiment is based on the test signal described in section 2 which is shown in the upper left part of fig.3. The local energy is computed using both gradient and quadrature filters. The quadrature filter is designed to mimic the gradient filter for positive frequencies, see fig2. The dotted LP-filter is used to average the square of the gradient response. Figure 3 shows the result, the σ

[3] 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.

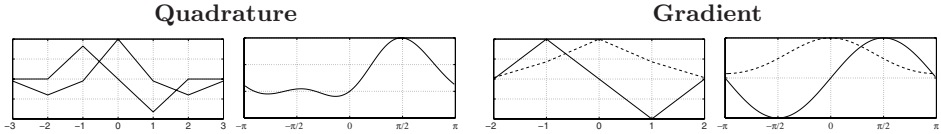

Fig. 2. Compact support quadrature and gradient filters in the spatial (left) and the Fourier (right) domain. The dotted LP-filter is used to average the square of the gradient response.

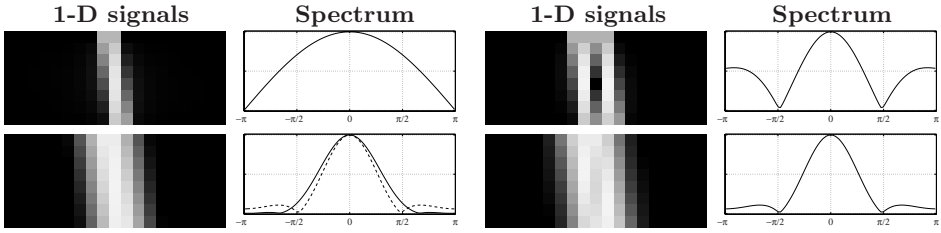

Fig. 3. Signals and corresponding spectra. Left/top: Original signal. Left/bottom: Squared quadrature output. Right/top: Squared gradient. Right/bottom Averaged squared gradient.

of the low-pass filter is chosen such that the variance of the spectrum is equal for the quadrature and gradient filters, see table1. Note that this σ is a bit too small to completely remove the two peaks in the gradient image. From fig.3 and the uncertainty relation in table1 it is obvious that the localization of the quadrature filters are superior compared to the conventional 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 in the top right of fig.3 may suggest that the performance of the gradient method can be improved by oversampling. The results for the large filters in table 1 shows that this is not the case.

The next experiment comprise a 2D spatio-temporal test signal where a small dark object is moving in front of a moving background, fig.4 left. Also, for test purposes, a slight vertical shading has been added. The velocity (or orientation) is estimated by both quadrature filters and gradient filters. The gradient filters in fig.2 where used but the quadrature filters are now loglets. The orientation errors where estimated from the gradient outer product matrix and the quadrature tensors as:

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

	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

Table 1. Variances and uncertainty products for the quadrature and gradient responses in fig.3. Theoretical uncertainty min value is 0.5. The table comprise two different scales. The left part correspond to the filters in fig.2 and the right part to more LP oriented filter sets with a larger spatial support.

Fig. 4. Left: two dimensional spatio-temporal test image. Note the slight vertical shading. Orientation error on spatiotemporal test image for gradient- (middle) and quadrature- (right) filters.

	no shading	with shading
gradient	13.5°	30.5°
quadrature	11.7°	12.5°

Table 2. Orientation error.

where $\hat{\boldsymbol{e}}_{\mathbf{1}}$ is the 'largest' eigenvector and $\hat{\boldsymbol{n}}$ defines the true velocity (orientation) of the image. The results are displayed in fig.4. It is apparent that the slight shading present in the signal misleads the estimation of the velocity using the gradient approach. If the shading is removed the performances are more similar but the quadrature approach still outperforms the gradient approach most likely due to better localization properties, see table2. The last experiment refer to

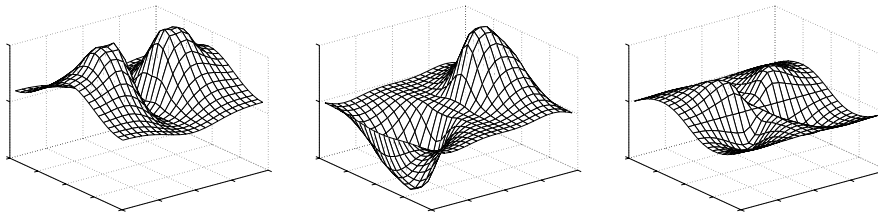

Fig. 5. Super quadrature filter in the FD. From left: orientation selective envelope and the envelope modulated by $\cos(\varphi)$ and $\sin(\varphi)$.

the part *Generalized Quadrature and Phase* in section 2. Here $\boldsymbol{q}$ is defined by the three filters in fig.5 where the radial part is a loglet, see eq.3. This order one phase representation comprise a $\cos^2(\varphi)$ angular window (left in fig.5) which is modulated by $\cos(\varphi)$ and $\sin(\varphi)$, (middle and right). By using an orientation selective window in three or more directions a phase invariant structure tensor can be estimated (eq.7) as well as the certainty in orientation c_θ, eq.8. The left part in fig.6 shows a consistency estimation 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_θ. Note that certainty estimates c_λ and c_θ clearly have different properties. Combining these estimates provides provide a more robust certainty estimate.

References

1. B. K. P. Horn. *Robot vision.* The MIT Press, 1986.
2. D. Gabor. Theory of communication. *J. Inst. Elec. Eng.*, 93(26):429–457, 1946.

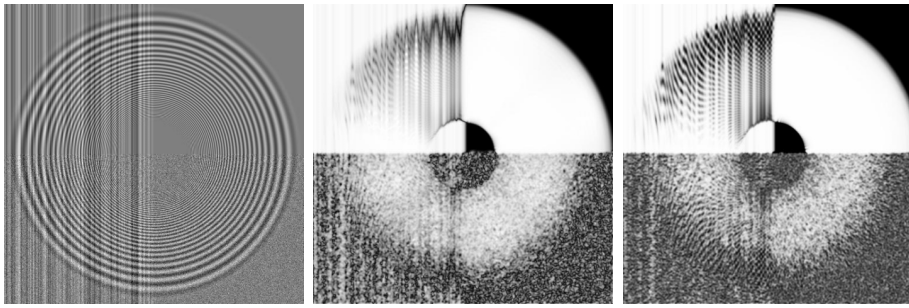

Fig. 6. Generalized quadrature. Left: test image. Middle: orientation consistency $c_\lambda = (\lambda_1 - \lambda_2)/\lambda_1$. Right: phase consistency based certainty c_θ.

3. H. Knutsson. *Filtering and Reconstruction in Image Processing.* PhD thesis, Linköping University, Sweden, 1982. Diss. No. 88.
4. G. H. Granlund. In search of a general picture processing operator. *Computer Graphics and Image Processing*, 8(2):155–178, 1978.
5. 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*, pages C–184–187. IEEE Computer Society, 1994.
6. Roy B. Leipnik. On lognormal random variables: I - the characteristic function. *J. Austral. Math. Soc.*, pages 327–347, 1991. Ser. B 32.
7. H. Knutsson, M. Andersson, and J. Wiklund. Advanced filter design. In *Proceedings of the Scandinavian Conference on Image analysis*, Greenland, June 1999. SCIA.
8. M. Riesz. Sur les fonctions conjuge'es. *Math. Zeit.*, 27:218–244, 1927.
9. L. Haglund, H. Knutsson, and G. H. Granlund. On phase representation of image information. In *The 6th Scandinavian Conference on Image Analysis*, pages 1082–1089, Oulu, Finland, June 1989.
10. G. H. Granlund and H. Knutsson. *Signal Processing for Computer Vision.* Kluwer Academic Publishers, 1995. ISBN 0-7923-9530-1.
11. R. Bracewell. *The Fourier Transform and its Applications.* McGraw-Hill, 2nd edition, 1986.
12. M. Felsberg and G. Sommer. Image features based on a new approach to 2d rotation invariant quadrature filters. In A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, editors, *Computer Vision - ECCV 2002*, volume 2350 of *Lecture Notes in Computer Science*, pages 369–383. Springer, 2002.
13. 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*, pages 175–182, Miami Beach, Florida, November 1985. IEEE. Report LiTH–ISY–I–0843, Linköping University, Sweden, 1986.
14. H. Knutsson. A tensor representation of 3-D structures. In *5th IEEE-ASSP and EURASIP Workshop on Multidimensional Signal Processing*, Noordwijkerhout, The Netherlands, September 1987. Poster presentation.
15. J. Bigün and G. H. Granlund. Optimal orientation detection of linear symmetry. In *Proceedings of the IEEE First International Conference on Computer Vision*, pages 433–438, London, Great Britain, June 1987.
16. H. Knutsson. Representing local structure using tensors. In *The 6th Scandinavian Conference on Image Analysis*, pages 244–251, Oulu, Finland, June 1989. Report LiTH–ISY–I–1019, Computer Vision Laboratory, Linköping University, Sweden, 1989.