J Math Imaging Vis (2017) 57:138–163
DOI 10.1007/s10851-016-0671-7

A Sinusoidal Image Model Derived from the Circular Harmonic Vector

Ross Marchant[1] · **Paul Jackway**[2]

Received: 22 October 2015 / Accepted: 20 June 2016 / Published online: 11 July 2016
© Springer Science+Business Media New York 2016

Abstract We introduce a sinusoidal image model consisting of an oriented sinusoid plus a residual component. The model parameters are derived from the circular harmonic vector, a representation of local image structure consisting of the responses to the higher-order Riesz transforms of an isotropic wavelet. The vector is split into sinusoidal and residual components. The sinusoidal component gives a phase-based description of the dominant local linear symmetry, with improved orientation estimation compared to previous sinusoidal models. The residual component describes the remaining parts of the local structure, from which a complex-valued representation of intrinsic dimension is derived. The usefulness of the model is demonstrated for corner and junction detection and parameter-driven image reconstruction.

Keywords Riesz transform · Steerable wavelets · Feature detection · Sinusoidal model

1 Introduction

Often the first step of image analysis is to obtain a description of the local image structure, from which detection or parametrisation of low-level features, such as line, edges, corners and junctions, can be made. Methods vary according the structural information obtained, and their robustness to image changes such as noise, illumination, contrast and rotation.

Phase-based representations are useful as they describe lines and edges in the one signal model [32,33,38,55]. The signal model is that of a sinusoid, where amplitude describes feature strength separately to phase, which describes feature type, such as line or edge. To obtain these parameters a set of symmetric (even) and anti-symmetric (odd) filters are required. This can consist of a pair of filters related by the Hilbert transform, such as in [14], or two or more filters related by the Riesz transform [10,50,53,57]. Depending on the method, an estimate of the orientation of the main axis of linear symmetry is also obtained.

One such method is the monogenic signal [10], which estimates orientation using the 1st-order Riesz transform (RT). However, the orientation estimate is poor near the centre of symmetric structures, such as lines, in the presence of noise. In this paper we propose a model of local image structure consisting of an oriented sinusoid plus a residual component, which is derived from higher-order RT responses. Using higher-order RTs improves the estimation of amplitude, phase and orientation, while the residual component is useful in that it describes the remainder of the local image structure that is not well-modelled by a sinusoid.

The structure of the paper is as follows: In Sect. 2 we review current sinusoidal image models and introduce the RT, then in Sect. 3 the proposed model is described. In Sect. 4 we introduce the circular harmonic vector as a primary descriptor of local image structure. It consists of the responses to the higher-order RTs of an isotropic wavelet. We show how to estimate the parameters of a general image model by splitting the vector into model and residual parts.

✉ Ross Marchant
ross.marchant@my.jcu.edu.au

Paul Jackway
paul@Neginfinity.net

[1] James Cook University, Townsville, Australia

[2] Neginfinity Pty. Ltd., Brisbane, Australia

Table 1 Mathematical notation used in this paper

Notation	Description
$f(\mathbf{z})$	2D image
$f_S(\mathbf{z})$	2D sinusoidal image
A	Sinusoid amplitude
ϕ	Sinusoid phase
θ	Sinusoid orientation
$\mathcal{R}^n$	n-th order complex RT
$\mathbf{f}$	Circular harmonic (CH) vector of wavelet coefficients at a point in an image
$\mathbf{f}_S$	CH vector at a point in a sinusoidal image
$\boldsymbol{\epsilon}$	Residual CH vector
$\mathbf{f}_e$	Even sinusoidal wavelet CH vector
$\mathbf{f}_o$	Odd sinusoidal wavelet CH vector
$\mathbf{W}$	CH vector weighting matrix
$\mathbf{S}_\theta$	CH vector rotation matrix
N	Maximum RT order in the CH vector
$\mathbf{d}$	Complex intrinsic dimension measure
γ	Intrinsic dimension angle

In Sect. 5 this method is used to solve for the sinusoidal model specifically, by creating wavelets that match a purely sinusoidal image. The effect of increasing the number of RT orders and their weighting is also discussed. In Sect. 6 both the sinusoidal and the residual components are used to create a complex-valued intrinsic dimension representation, from which a robust corner and junction detector is obtained. Finally, in Sect. 7 the usefulness of the approach is demonstrated for model-driven image reconstruction and wavelet-based denoising.

A number of different functions and variables are used throughout the paper, the most important of which are listed in Table 1 as a reference for the reader.

2 Related Signal Models

In 1D, the analytic signal consists of a signal and its Hilbert transform. It is represented as a complex exponential,

$$f_A(x) = f(x) + \mathrm{i}\mathcal{H}f(x) \tag{1}$$

$$= A(x)e^{\mathrm{i}\phi(x)}, \tag{2}$$

from which the original signal is modelled as a sinusoid, $f(x) = A(x)\cos(\phi(x))$, where $A(x)$ is the amplitude and $\phi(x)$ is the phase. Typically the response is localised by convolution with an isotropic band-pass filter centred at the frequency (scale) of interest. The combination of filter and its Hilbert transform is referred to as a quadrature filter.

The isotropic part of the quadrature filter is an even (symmetric) function and thus has a high response at locally symmetric parts of the signal, such as roof edges and peaks. The Hilbert transformed part is an odd (anti-symmetric) function and thus has a high response at locally anti-symmetric parts, such as step edges. It has been shown that the sinusoid amplitude is high at locations of each of these features, and that the phase can be used to discriminate their type [32,33,38,55].

In 2D, lines and edges are analogous to peaks and steps in 1D, therefore a sinusoidal model of these features would also be useful. One approach is to use a 2D quadrature filter, consisting of an even (or odd) directional filter and its Hilbert transform along the directional axis. However, there is an additional complexity in that the axis of filter symmetry should match that of the local signal structure. Rotated copies of quadrature filter at discrete orientations can be used as an approximation. Alternatively, the filter can be constructed from a steerable basis from which any directional response can be synthesised [14,39,46]. However, for these filters the Hilbert transformed part is an approximation [14].

2.1 Monogenic Signal

Other approaches have focussed on extending the 1D analytic signal to 2D images. However, early attempts applied the Hilbert transform along discrete orientations, and therefore varied with rotations [5,15,49]. The first rotationally invariant extensions [10,22] used the Riesz transform (RT) instead of the Hilbert transform to generate the required quadrature filters.

Of these, the monogenic signal models local image structure as a single 2D oriented sinusoid [10]. Assuming a local coordinate system where the point of interest is located at $\mathbf{z} = \mathbf{0}$, the monogenic signal model is

$$f_S(\mathbf{z}) = A\cos(\langle \mathbf{z}, \mathbf{o}\rangle + \phi) \tag{3}$$

with parameters: amplitude, A, phase, ϕ, and orientation, θ. The parameters are obtained using the 0th and 1st-order 2D complex RT, $\mathcal{R}^n$, according to

$$A = \left|\mathcal{R}^0 f + \mathrm{i}|\mathcal{R}^1 f|\right|, \tag{4}$$

$$\phi = \arg(\mathcal{R}^0 f + \mathrm{i}|\mathcal{R}^1 f|), \tag{5}$$

$$\theta = \arg(\mathcal{R}^1 f), \tag{6}$$

where $A \in \mathbb{R}^+$, $\phi \in [0, \pi)$ and $\theta \in [0, 2\pi)$ [10].

Figure 1 shows the monogenic signal model of the *House* image, localised using a Cauchy filter [44] with centre wavelength 8 pixels and bandwidth factor $a = 2$ [4]. As is observed for the analytic signal, the amplitude is high at the

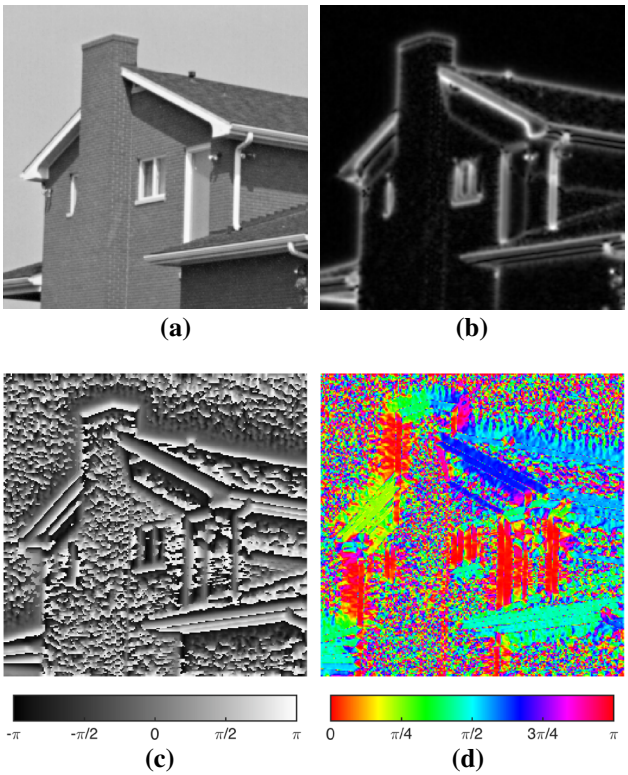


Fig. 1 *House* image (**a**) and its sinusoidal model amplitude (**b**), phase (**c**) and orientation (**d**) given by the monogenic signal.

locations of strong image features [10]. The phase value is a measure of the local symmetry independent of amplitude; phase values near 0 or ϕ indicate an even structure, such as a line, and phase values near $\pm\pi/2$ indicate an odd structure, such as an edge. The orientation parameter describes the main axis of symmetry of the local structure [10]. The monogenic signal has been applied to diverse problems such as optical flow measurement [8], image registration [29] and segmentation [1].

However, care must be taken when interpreting the model. Firstly, the 0th-order RT responds to both even and isometric structures, meaning that blobs, which would not be considered sinusoid-like, also give a large response. Secondly, the 0th-order operator is isotropic, therefore orientation is only calculated from the 1st-order RT, which only responds to odd structures. Thus the orientation estimate is poor near the centre of even features in the presence of noise [7]. This is noticeable in Fig. 1d as discontinuities along some features.

2.2 Higher-Order Signals

The poor orientation estimate around even structures is a drawback of using the monogenic signal in practical applications. Two solutions that have been proposed are to average the phase vector near even structures [7], or to include higher-order RT responses using an expanded signal model [7,11,57,58].

The n-th order complex RT is given by [53,54,57]

$$\mathcal{R}^n f(\mathbf{z}) \quad \overset{\mathcal{F}}{\longleftrightarrow} \quad e^{\mathrm{i}n\phi}\hat{f}(\omega,\phi), \tag{7}$$

where ω and ϕ are the radial and angular co-ordinates of the frequency spectrum, respectively. The 0th-order RT, $\mathcal{R}^0$, is the identity operator. The complex embedding allows for easy rotation of the impulse response by multiplying by a complex exponential [51] as

$$\mathcal{R}^n\{\delta\}(\mathbf{R}_\theta \mathbf{z}) = e^{-\mathrm{i}n\theta}\mathcal{R}\{\delta\}(\mathbf{z}), \tag{8}$$

where $\mathbf{R}_\theta$ is a matrix that rotates the image axes by θ. Like the Hilbert transform, the RT has an infinite impulse response and requires the entire image to compute. To construct a more localised operator, we can combine the RT with an isotropic band-pass filter with enough vanishing moments [51], resulting in a spherical quadrature filter (SQF) [7,10].

Similar to the Hilbert transform in 1D, the RT of a sinusoidal signal is equivalent to a phase shift. Consider an image given by the addition of K 2D sinusoidal images,

$$f(\mathbf{z}) = \sum_{k=1}^{K} A_k \cos(w\langle \mathbf{z}, \mathbf{o}_k\rangle + \phi_k), \tag{9}$$

where $\mathbf{o} = [\cos\theta, \sin\theta]$ is the orientation vector with $\theta \in [0,\pi)$, $A \in \mathbb{R}^+$ is the amplitude, and $\phi \in [0, 2\pi)$ is the phase of the sinusoid. It has been shown by means of the Radon transform [10,57] that the n-th order RT of this signal at the point of interest, $\mathbf{z} = \mathbf{0}$, is

$$\mathcal{R}^n f = \begin{cases} \sum_k A_k e^{\mathrm{i}n\theta_k} \cos(\phi_k), & n \text{ is even}, \\ \sum_k A_k e^{\mathrm{i}n\theta_k} \,\mathrm{i}\sin(\phi_k), & n \text{ is odd}. \end{cases} \tag{10}$$

These higher-order RT responses give more estimates of sinusoid parameters, albeit with the orientation range reduced to $2\pi/n$ for RT order n. Therefore, to estimate orientation from even structures, RTs orders from 2 or above must be used.

Sinusoidal image models derived from these higher-order responses include the structure multi-vector [7,11], 2D analytic signal [58] and the signal multi-vector [57]. The structure multi-vector model consists of two sinusoids at right angles,

$$f(\mathbf{z}) = \sum_{k=1}^{2} A_k \cos(\langle \mathbf{z}, \mathbf{o}_k\rangle + \phi_k), \tag{11}$$

where $\theta_2 = \theta_1 + \pi/2$, and is calculated using the 0th to 3rd-order RTs.

The 2D analytic signal model consists of two sinusoids with the same phase,

$$f(\mathbf{z}) = \sum_{k=1}^{2} A_k \cos(\langle \mathbf{z}, \mathbf{o}_k \rangle + \phi), \tag{12}$$

and is calculated using the 0th to 2nd-order RTs.

The signal multi-vector model consists of two sinusoids without restriction,

$$f(\mathbf{z}) = \sum_{k=1}^{2} A_k \cos(\langle \mathbf{z}, \mathbf{o}_k \rangle + \phi_k), \tag{13}$$

and is calculated using the 0th to 3rd-order RTs.

In the case of the structure multi-vector and signal multi-vector, including higher-order RTs leads to more orthogonal responses than model parameters (responses to RT orders above zero are complex, and therefore have two dimensions). The structure multi-vector deals with this by projecting the RT responses on to four basis functions, from which the model parameters are calculated, while the signal multi-vector projects onto six [57]. Expanding the model to include extra sinusoids also makes use of the extra information from higher-order responses.

Problems exist with the 2D analytic signal and signal multi-vector solutions. For the 2D analytic signal, when $|\mathcal{R}^2 f| > |\mathcal{R}^0 f|$, (60) in [57] gives a complex value for the apex angle, $\theta_1 - \theta_2$, violating the model. This occurs at image structures with a higher 2nd-order RT response, such as saddles. For the signal multi-vector, an image structure consisting of two equal amplitude sinusoids with phases $\{0, \pi\}$ and opposite orientations $\{-\theta, \theta\}$ gives $\mathcal{R}^0 f = \mathcal{R}^1 f = \mathcal{R}^3 f = 0$ and $\mathcal{R}^2 f = 2A(e^{\mathrm{i}2\theta} + e^{-\mathrm{i}2\theta}) = 2A\cos(2\theta)$. Thus $\mathcal{R}^2 f$ is real valued and the model parameters cannot be found due to having more unknowns than knowns. The addition of the 4th-order RT response may be required. Furthermore, the method of calculating orientation given by (130-131) in [57] uses only odd-order RT responses, and thus again the orientation estimate will be poor near even structures in the presence of noise. In contrast, the structure multi-vector uses the 2nd-order RT in the orientation estimation and thus does resolve this problem. However, the maximum order RT used is still limited to three.

Other approaches that explicitly use RTs include the monogenic curvature signal [60] and the boundary tensor [19] but these do not use a sinusoidal model. A phase-based representation can be obtained using the monogenic curvelet transform [50]; however, it uses the 1st-order RT of an even filter at discrete orientations.

3 Proposed Model

Each of the higher-order sinusoidal models previously described are limited in the number of RT orders used. Furthermore, the 2D analytic signal and signal multi-vector do not improve on the orientation estimation of the monogenic signal. Addressing this problem was the original motivation behind this work. Specifically, we wished to obtain a robust phase-invariant orientation estimate from any number of higher-order RT responses. Preliminary attempts can be found in our earlier work [24,26].

To achieve this, we have previously proposed a multi-sinusoidal image model that consists of multiple oriented sinusoids with differing amplitude, phase and orientation parameters plus a residual component [28]. Local image structure at a point of interest, $\mathbf{z} = \mathbf{0}$, is modelled by

$$f(\mathbf{z}) = \sum_{k=1}^{K} \underbrace{A_k \cos(\langle \mathbf{z}, \mathbf{o}_k \rangle + \phi_k)}_{f_{\mathrm{S}}(\mathbf{z})} + f_\epsilon(\mathbf{z}), \tag{14}$$

where $f_{\mathrm{S}}(\mathbf{z})$ is a single sinusoidal model component with amplitude A, phase ϕ and orientation θ such that $\mathbf{o} = [\cos\theta, \sin\theta]$, and $f_\epsilon(\mathbf{z})$ is a *residual* component. In this paper, we shall restrict the model to a single sinusoid, that is, $K = 1$, and multiple sinusoids will be discussed in future work. The model is

$$f(\mathbf{z}) = \underbrace{A\cos(\langle \mathbf{z}, \mathbf{o} \rangle + \phi)}_{f_{\mathrm{S}}(\mathbf{z})} + f_\epsilon(\mathbf{z}). \tag{15}$$

Adding the residual component allows for the inclusion of higher-order RT responses up to any order, as it then becomes possible to choose model parameters that satisfy the RT sinusoidal response equations,

$$\mathcal{R}^n f = \begin{cases} A e^{\mathrm{i}n\theta} \cos(\phi) + \mathcal{R}^n f_\epsilon, & n \text{ is even}, \\ A e^{\mathrm{i}n\theta}\, \mathrm{i}\sin(\phi) + \mathcal{R}^n f_\epsilon, & n \text{ is odd}, \end{cases} \tag{16}$$

for arbitrary images.

The task now is to choose appropriate values of amplitude, phase and orientation. To do this we shall minimise the local energy of the residual component so that the sinusoidal model explains as much of the local image structure as possible. In the next section, we describe a method of finding the parameters of a general image model from a vector of higher-order RT responses. The method can then be applied to the sinusoidal image model.

4 Circular Harmonic Wavelets

The RT has properties of steerability, translation invariance, scale-invariance and inner-product preservation [10,51]. Of particular relevance is that the RT is norm-preserving, $\|\mathcal{R} f\| = \|f\|$, and invertible, $\mathcal{R}^{-n}\{\mathcal{R}^n f\}(\mathbf{z}) = f(\mathbf{z})$, if $f(\mathbf{z})$ has zero mean. These properties allow the generation

of monogenic *wavelets* [17,36] in multiple dimensions [54] and for colour images [48], as well as monogenic versions of existing quadrature wavelets that give a directional decomposition into amplitude and phase components [50].

The higher-order RT can also be used to construct 2D steerable wavelet frames [51,52]. These consist of a set of wavelets constructed from the linear combination of higher-order RTs of a suitable primary isotropic wavelet frame. The wavelets can be rotated, shifted and linearly combined to construct an image. Apart from the strong mathematical basis in the literature, wavelets enable a multi-scale approach that encompasses multi-scale filter banks, pyramidal decompositions and image reconstruction [18,51]. As such, we shall formulate the sinusoidal model solution in this context.

Specifically, our approach is to treat the vector of 2D steerable *circular harmonic* (CH) wavelet responses as the primary descriptor of local image structure. Given a set of wavelets that match the components of particular signal model, we can split the CH vector into a model component that describes the structures of interest, and a residual component that describes the remainder. This general approach is developed in this section.

4.1 2D Steerable Wavelets

A general framework for 2D steerable wavelet transforms is laid out in [51]. Restating proposition 4.1 from [51], if $h(\omega)$ is a radial frequency profile satisfying the following conditions:

$$h(\omega) = 0, \quad \forall\, \omega > \pi, \tag{17}$$

$$\sum_{i\in\mathbb{Z}} |h(2^i\omega)|^2 = 1, \tag{18}$$

$$\left.\frac{\mathrm{d}^n h(\omega)}{\mathrm{d}\omega^n}\right|_{\omega=0} = 0, \quad \text{for } n = 0,\ldots,N, \tag{19}$$

then the isotropic wavelet mother wavelet ψ with spectrum $\hat{\psi}(\boldsymbol{\omega}) = h(\|\boldsymbol{\omega}\|)$ generates a tight wavelet frame of $L_2(\mathbb{R}^2)$ whose basis functions, $\psi_{i,\mathbf{k}} = \psi_i(\mathbf{z} - 2^i\mathbf{k})$ with $\psi_i(\mathbf{z}) = 2^{-2i}\psi(\mathbf{z}/2^i)$, are isotropic with vanishing moments up to order N. Given a primary isotropic wavelet frame $\{\psi_{i,\mathbf{k}}\}_{i\in\mathbb{Z},\mathbf{k}\in\mathbb{Z}^2}$ that satisfies the above conditions, the higher-order RT can be used create a steerable wavelet frame $\{\psi^{(m)}_{i,\mathbf{k}}\}_{m\in\mathbb{N}^+_M,i\in\mathbb{Z},\mathbf{k}\in\mathbb{Z}^2}$ of $L_2(\mathbb{R}^2)$ by [51]

$$\psi^{(m)}_{i,\mathbf{k}} = \sum_{|n|\in\mathbb{N}_N} u_{m,n}\mathcal{R}^n\psi_{i,\mathbf{k}}, \tag{20}$$

where $\mathbf{U}_{M,N}$ is a complex-valued shaping matrix of size $M \times (2N+1)$, $M \geq 1$. Note, $\mathbb{N}^+_M$ means the set of integers from 1 to M, $\mathbb{N}_M$ is the set from 0 to M.

Of the three conditions, the second results in an energy-preserving partition of the frequency spectrum, creating a tight wavelet frame that gives exact image reconstruction. Scaling by 2^i allows for sub-sampling to create pyramidal decompositions, although alternative partitions that are more narrowly spaced can be used, such as in [17]. The third condition requires the primary wavelet to have at least N vanishing moments to account for the singularity of the RT at the origin, and thus for the wavelets to have sufficient spatial decay [51,56]. If reconstruction or pyramidal decompositions are not of interest, the second condition can be abandoned, and an image can be analysed using an isotropic filter bank that preferably satisfies the third condition.

4.2 Circular Harmonic Vector

Details of various shaping matrices that generate tight wavelet frames are given in [51]. In particular, it is shown that a pyramidal decomposition with exact reconstruction is possible for directional wavelets, using a frame consisting of $2N+1$ equiangular copies of the wavelet, or $N+1$ copies over the half circle if the wavelet is linearly symmetric along some axis. However, rather than using a specific wavelet, our approach starts with the circular harmonic (CH) wavelet frame, which is a basis for other 2D steerable wavelet frames.

The CH frame, $\{\psi^n_{i,\mathbf{k}}\}_{|n|\in\mathbb{N}_N,i\in\mathbb{Z},\mathbf{k}\in\mathbb{Z}^2}$, consists of wavelets given by the $-N$-th to N-th order RTs of a primary isotropic wavelet ($\mathbf{U} = \mathbf{I}_{2N+1}$ [51]), and are also known as CH functions [18]. The CH frame is similar to the set of SQFs constructed from an isotropic band-pass filter [7,10] used for image analysis in our previous work [25]. Each wavelet is given by

$$\psi^n_{i,\mathbf{k}} = \mathcal{R}^n\psi_{i,\mathbf{k}}. \tag{21}$$

The CH wavelets have equal norm with frame bounds of $2N+1$.

Figure 2 shows an example of these wavelets generated using the 0th- to 3rd-order RTs. It can be seen that the number of axes of symmetry increases with RT order. Thus the 0th-order wavelet would respond to isotropic structures, the

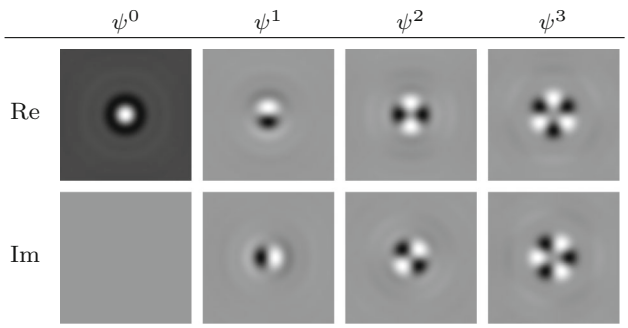


Fig. 2 Real and imaginary parts of the 0th- to 3rd-order circular harmonic wavelets

1st-order wavelet to odd structures such as edges, the 2nd-order wavelet to even structures such as lines, and so on. Importantly, each wavelet is orthogonal and has an increasingly complex structure. Correlation with the set of wavelets thus gives a set of independent responses that describe the local image structure.

With this in mind, we propose a novel approach to local image representation. We collect the $-N$-th to N-th order CH responses into a vector, which shall be referred to as a *CH vector*. Applying the CH wavelet frame to an image $f \in L_2(\mathbb{R}^2)$, the CH vector of correlation coefficients (responses to the CH wavelets) at each scale and location is given by

$$\mathbf{f}_{i,\mathbf{k}} = \left[\left\langle f, \psi^{-N}_{i,\mathbf{k}}\right\rangle, ..., \left\langle f, \psi^{N}_{i,\mathbf{k}}\right\rangle\right]^T. \tag{22}$$

It is proposed that the CH vector is a *local descriptor* of image structure:

1. The amplitude of the n-th component, $|(\mathbf{f}_{i,\mathbf{k}})_n|$, represents the magnitude of the n-th order rotational symmetry of the local image structure.
2. The argument of the n-th component, $\arg((\mathbf{f}_{i,\mathbf{k}})_n)$, represents the orientation offset of the n-th order symmetry.
3. The magnitude of the entire vector, $\|\mathbf{f}_{i,\mathbf{k}}\|$, is a measure of local energy.
4. The scale of the basis wavelet, i, indicates the size of the local structure.
5. The CH wavelets are a basis for other 2D steerable wavelets and thus the vector of responses for a particular frame is $\mathbf{U}^H\mathbf{f}_{i,\mathbf{k}}$.
6. Rotation of the image causes a rotation of the CH vector components (8). For a rotation of the image axes by θ, the CH vector becomes $\mathbf{S}_{-\theta}\mathbf{f}_{i,\mathbf{k}}$ where $\mathbf{S}_\theta = \text{diag}[e^{\text{i}N\theta}, ..., e^{-\text{i}N\theta}]$ is a diagonal rotation matrix.

CH wavelets as a local descriptor have also been investigated in [18] and used for tasks such as edge enhancement and pattern recognition [35].

Because the CH frame bounds are $2N+1$, to obtain exact reconstruction (frame bounds = 1) the CH vector coefficients must be weighted. Given a real-valued diagonal weighting matrix $\mathbf{W} = \text{diag}[w_{-N}, ..., w_N]$ where $w_{-n} = w_n$ and $\sum_{|n|\in\mathbb{N}_N} w_n^2 = 1$, the set of weighted CH wavelets $\{w_n\psi^n_{i,\mathbf{k}}\}_{|n|\in\mathbb{N}_N, i\in\mathbb{Z}, \mathbf{k}\in\mathbb{Z}^2}$ will have a frame bound of 1. Exact reconstruction is then possible using the *weighted* CH vector, $\mathbf{Wf}$, by

$$f(\mathbf{z}) = \sum_{i,\mathbf{k}} \sum_{|n|\in\mathbb{N}_N} (\mathbf{Wf}_{i,\mathbf{k}})_n w_n \psi^n_{i,\mathbf{k}}. \tag{23}$$

4.3 Comparing CH Vectors

Weighting each order differently also allows us to control how each order contributes to the magnitude of the CH vector, $\|\mathbf{Wf}\|$, which is a measure of local energy. For example, for $N = 1$ with weights $\mathbf{w} = [1/\sqrt{2}, 1/2, 1/\sqrt{2}]$, the CH vector is equivalent to the monogenic signal vector in [10], where the odd and even orders are equally weighted. The normalised weighted CH vector, $\mathbf{Wf}/\|\mathbf{Wf}\|$, gives an illumination-invariant description of the local image structure.

One may wish to compare different image structures using their corresponding CH vectors. A simple measure is the distance between the vectors,

$$\text{d}(\mathbf{Wf}, \mathbf{Wg}) = \|\mathbf{Wf} - \mathbf{Wg}\|. \tag{24}$$

However, the distance varies with the magnitude of the vectors, a measure of feature strength. It is often desirable to compare structures by feature shape alone, so to introduce illumination invariance we may normalise the vectors,

$$\text{d}_{\text{norm}}(\mathbf{Wf}, \mathbf{Wg}) = \left\| \frac{\mathbf{Wf}}{\|\mathbf{Wf}\|} - \frac{\mathbf{Wg}}{\|\mathbf{Wg}\|} \right\|, \tag{25}$$

or use the angle difference between them,

$$\gamma = \cos^{-1} \frac{\langle \mathbf{Wf}, \mathbf{Wg} \rangle}{\|\mathbf{Wf}\|\|\mathbf{gW}\|}. \tag{26}$$

4.4 Solving Models

Consider a general image model given by the linear combination of a set of M structures of interest, $\{u_m(\mathbf{z})\}m_{\in\mathbb{N}_M}$, rotated to different orientations, plus a residual component, $f_\epsilon(\mathbf{z})$. At a point of interest, $\mathbf{z} = \mathbf{0}$, the model is

$$f(\mathbf{z}) = \sum_m \lambda_m u_m(\mathbf{R}_{\theta_m}\mathbf{z}) + f_\epsilon(\mathbf{z}), \tag{27}$$

where λ_m and θ_m are the amplitude and orientation parameters that need to be found.

The wavelet, ψ_{u_m}, with CH vector $\mathbf{Wu}_m$, that *matches* a particular feature for a given weighting is described by the normalised CH vector of that image structure. That is,

$$\mathbf{Wu}_m = \mathbf{Wg}/\|\mathbf{Wg}\|, \tag{28}$$

where $\mathbf{g}$ is the CH vector at the centre of the feature. Consider the set of M 2D steerable wavelets, $\{\psi^{(m)}\}_{m\in\mathbb{N}_M}$, that match the particular set of image structures we are interested in, and let $\{\mathbf{Wu}_m\}_{m\in\mathbb{N}_M}$ be the corresponding set of weighted and normalised CH vectors. The local image structure CH vector, $\mathbf{Wf}$, can be written as the sum of individually scaled

and rotated versions of each model wavelet CH vector, plus the residual component, $\mathbf{W}\boldsymbol{\epsilon}$. The relationship for a particular scale i and location $\mathbf{k}$ is

$$\mathbf{W}\mathbf{f} = \sum_m \lambda_m \mathbf{S}_{\theta_m} \mathbf{W}\mathbf{u}_m + \mathbf{W}\boldsymbol{\epsilon}. \tag{29}$$

Since we can reconstruct the image exactly from the CH vector, it can be expressed as the sum of separate model and residual reconstructions,

$$f(\mathbf{z}) = f_\psi(\mathbf{z}) + f_\epsilon(\mathbf{z}), \tag{30}$$

where $f_\psi(\mathbf{z})$ is the image synthesised from the model wavelets,

$$f_\psi(\mathbf{z}) = \sum_{i,\mathbf{k}} \sum_m \sum_{|n|\in\mathbb{N}_N} (\lambda_m)_{i,\mathbf{k}} (\mathbf{S}_{\theta_{i,\mathbf{k}}} \mathbf{W}\mathbf{u}_m)_n \, w_n \psi^n_{i,\mathbf{k}}, \tag{31}$$

and $f_\epsilon(\mathbf{z})$ is the image synthesised from the residual component,

$$f_\epsilon(\mathbf{z}) = \sum_{i,\mathbf{k}} \sum_m \sum_{|n|\in\mathbb{N}_N} (\mathbf{W}\boldsymbol{\epsilon}_{i,\mathbf{k}})_n w_n \psi^n_{i,\mathbf{k}}. \tag{32}$$

The residual component is the missing part of the local image structure that is not correlated with the model wavelets but is needed for exact reconstruction of the image. By adding the residual we can use arbitrary model wavelets and still have exact reconstruction.

To solve the model, we choose values of λ_m and θ_m that minimise the residual component at each scale and location, such that the model wavelets explain as much of the image as possible. Since the CH wavelets $\{\psi^n\}_{|n|\in\mathbb{N}_N}$ are orthogonal, the L_2-norm of the residual wavelet $\psi_{\epsilon_{i,\mathbf{k}}}$ is proportional to the l_2-norm of the residual vector $\mathbf{W}\boldsymbol{\epsilon}_{i,\mathbf{k}}$ [51]. Therefore

$$\|\psi_\epsilon\| \propto \|\mathbf{W}\boldsymbol{\epsilon}\| \tag{33}$$

$$\propto \left\| \mathbf{W}\mathbf{f} - \sum_m \lambda_m \mathbf{S}_{\theta_m} \mathbf{W}\mathbf{u}_m \right\|. \tag{34}$$

Letting $\boldsymbol{\lambda} = [\lambda_1, ..., \lambda_M]$ and $\boldsymbol{\theta} = [\theta_1, ..., \theta_M]$, the linear scale and rotation parameters that minimise the residual are given by

$$\boldsymbol{\lambda}, \boldsymbol{\theta} = \arg\min_{\boldsymbol{\lambda},\boldsymbol{\theta}} \left\| \mathbf{W}\mathbf{f} - \sum_m \lambda_m \mathbf{S}_{\theta_m} \mathbf{W}\mathbf{u}_m \right\|. \tag{35}$$

Finding an exact solution for the above equation is difficult due to the multiple non-linear parts from multiple orientation parameters. If we constrain the problem such that all the model CH vectors are linearly independent and each model wavelet has the same orientation, only the single non-linear part needs to be solved. For simplicity and without loss of generality we drop the weighting matrix $\mathbf{W}$, that is $\mathbf{W}\mathbf{f} \to \mathbf{f}$ and so on. Collecting all the normalised model CH vectors, $\{\mathbf{u}_m\}_{m\in\mathbb{N}_M}$ into the columns of a $2N+1 \times M$ matrix $\mathbf{U}$, the problem is reduced to

$$\boldsymbol{\lambda}, \theta = \arg\min_{\boldsymbol{\lambda},\theta} \|\mathbf{f} - \mathbf{S}_\theta \mathbf{U}\boldsymbol{\lambda}\| \tag{36}$$

$$= \arg\min_{\boldsymbol{\lambda},\theta} \|\mathbf{S}_{-\theta}\mathbf{f} - \mathbf{U}\boldsymbol{\lambda}\|. \tag{37}$$

By holding θ constant, this becomes the classic linear least-squares problem. Using the properties $\|\mathbf{a}\|^2 = \langle \mathbf{a}, \mathbf{a}\rangle = \mathbf{a}^H\mathbf{a}$ and $(\mathbf{AB})^H = \mathbf{B}^H\mathbf{A}^H$, and since $\mathbf{S}_\theta\mathbf{S}_\theta^H = \mathbf{I}_N$ and $\mathbf{S}_{-\theta} = \mathbf{S}_\theta^H$, we have

$$\min_{\boldsymbol{\lambda},\theta} \|\mathbf{S}_{-\theta}\mathbf{f} - \mathbf{U}\boldsymbol{\lambda}\| \tag{38}$$

$$= \min_{\boldsymbol{\lambda},\theta} \; -2\boldsymbol{\lambda}^H\mathbf{U}^H\mathbf{S}_\theta^H\mathbf{f} + \boldsymbol{\lambda}^H\mathbf{U}^H\mathbf{U}\boldsymbol{\lambda}. \tag{39}$$

At the minimum, the derivative with respect to $\boldsymbol{\lambda}$ will be 0, which gives

$$0 = -2\mathbf{U}^H\mathbf{S}_\theta^H\mathbf{f} + 2\mathbf{U}^H\mathbf{U}\boldsymbol{\lambda}, \tag{40}$$

$$\boldsymbol{\lambda} = (\mathbf{U}^H\mathbf{U})^{-1}\mathbf{U}^H\mathbf{S}_\theta^H\mathbf{f} \tag{41}$$

$$= \mathbf{U}^+\mathbf{S}_\theta^H\mathbf{f}, \tag{42}$$

where $\mathbf{U}^+$ is the pseudo-inverse of $\mathbf{U}$. If all the model CH vectors are orthogonal as well as normal, $\mathbf{U}^+ = \mathbf{U}^H$. To solve for θ we substitute back into (39) to get

$$\theta = \min_\theta \; -2\boldsymbol{\lambda}_\theta^H\mathbf{U}^H\mathbf{S}_\theta^H\mathbf{f} + \boldsymbol{\lambda} - \theta^H\mathbf{U}^H\mathbf{U}\boldsymbol{\lambda}_\theta \tag{43}$$

$$= \min_\theta \; -2\mathbf{f}^H\mathbf{S}_\theta(\mathbf{U}^+)^H\mathbf{U}^H\mathbf{S}_\theta^H\mathbf{f} \tag{44}$$

$$+ \mathbf{f}^H\mathbf{S}_\theta(\mathbf{U}^+)^H\mathbf{U}^H\mathbf{U}\mathbf{U}^+\mathbf{S}_\theta^H\mathbf{f}. \tag{45}$$

Since $\mathbf{U}\mathbf{U}^+$ is hermitian, $(\mathbf{U}^+)^H\mathbf{U}^H = (\mathbf{U}\mathbf{U}^+)^H = \mathbf{U}\mathbf{U}^+$ and $\mathbf{U}^+\mathbf{U} = \mathbf{I}_M$ and the above equation simplifies to

$$\theta = \max_\theta \; (\mathbf{U}^H\mathbf{S}_\theta^H\mathbf{f})(\mathbf{U}^+\mathbf{S}_\theta^H\mathbf{f}). \tag{46}$$

Letting $\lambda_m(\theta) = \mathbf{u}_m^+\mathbf{S}_\theta^H\mathbf{f}$ and $\delta_m(\theta) = \mathbf{u}_m\mathbf{S}_\theta^H\mathbf{f}$ we arrive at

$$\theta = \max_\theta \sum_m \delta_m(\theta)\lambda_m(\theta). \tag{47}$$

When the model CH vectors are orthonormal, $\delta_m(\theta) = \lambda_m(\theta)$. Both of these functions are trigonometric polynomials with order $2N$, that is

$$\lambda_m(\theta) = \sum_{|n|\in\mathbb{N}_N} w_n \bar{u}_{m_n} w_n f_n \, e^{-\mathrm{i}n\theta} \tag{48}$$

and thus $p(\theta) = \sum_m \delta_m(\theta)\lambda_m(\theta)$ is a trigonometric polynomial with degree $4N$. At the maximum, the derivative of $p(\theta)$ equals 0, therefore finding the roots of $p'(\theta)$ gives us $4N$ candidate values for θ. Substituting back into $p(\theta)$ we choose the value that gives the maximum, from which λ_m can be calculated for each wavelet.

4.5 Summary

The CH vector is a descriptor of local image structure that generalises previous signal vectors such as the monogenic signal and signal multi-vector. The main advantage of the proposed method is that we can design a set of arbitrary wavelets to model a particular set of image structures, where the wavelets do not necessarily form a frame, and yet still have exact reconstruction with the addition of the residual component. Furthermore, no structural information is lost; the residual vector provides a description of the non-model part of the local image structure, which can be used for further analysis.

5 Sinusoidal Image Model

We shall now apply the CH vector method to finding the amplitude, phase and orientation of the sinusoidal model.

5.1 Matched Wavelets

The proposed sinusoidal model of local image structure is locally represented by an oriented sinusoid as

$$f_{\mathrm{S}}(\mathbf{z}) = A\cos(\omega_0 \langle \mathbf{z}, \mathbf{o}\rangle + \phi), \tag{49}$$

where $\mathbf{o} = [\cos\theta, \sin\theta]$ and ω_0 is the sinusoid frequency.

The first step is to find the wavelet that matches the sinusoid for a particular amplitude, A, phase, ϕ and orientation, θ. Let $\{\psi^n\}_{|n|\leq N}$ be the set of CH wavelets up to order N, generated from an isotropic wavelet ψ. For simplicity, let the sinusoid frequency be located at the centre of the wavelet passband $h(\omega)$ such that $|h(\omega_0)| = 1$. The value of the sinusoid CH vector $\mathbf{f}_{\mathrm{S}}$ at the origin is therefore given by

$$f_{\mathrm{S}_n} = \langle f_{\mathrm{S}}, \psi^n \rangle \tag{50}$$

$$= \mathcal{R}^{-n}(f_S * \psi)(\mathbf{0}) \tag{51}$$

$$= \begin{cases} Ae^{-\mathrm{i}n\theta}\cos(\phi), & n \text{ is even}, \\ Ae^{-\mathrm{i}n\theta}\,\mathrm{i}\sin(\phi), & n \text{ is odd}. \end{cases} \tag{52}$$

The sinusoidal image CH vector can be written as a function of amplitude, phase and orientation,

$$\mathbf{f}_{\mathrm{S}}(A, \phi, \theta) = A\mathbf{S}_\theta \cos\phi\, \mathbf{s}_e + A\mathbf{S}_\theta \sin\phi\, \mathbf{s}_o, \tag{53}$$

where $\mathbf{s}_e$ and $\mathbf{s}_o$ are orthogonal CH vectors given by

$$s_{e_n} = 1 \quad \text{if } n \text{ even, 0 otherwise}, \tag{54}$$

$$s_{o_n} = -\mathrm{i} \quad \text{if } n \text{ odd, 0 otherwise}. \tag{55}$$

The vector $\mathbf{s}_e$ has only even orders and therefore represents an even wavelet, while $\mathbf{s}_o$ has only odd orders and therefore represents an odd wavelet. The sinusoidal image consists of the linear combination of these two wavelets, rotated to the same orientation.

We thus create two matched wavelets that correlate with the even and odd parts of the signal. Setting orientation to 0, the two matched wavelets for a given weighting matrix $\mathbf{W}$ are

$$\mathbf{W}\mathbf{f}_e = \mathbf{W}\mathbf{s}_e/\sqrt{W_e}, \tag{56}$$

$$\mathbf{W}\mathbf{f}_o = \mathbf{W}\mathbf{s}_o/\sqrt{W_o}, \tag{57}$$

where W_e and W_o are the sum of the even and odd weights, respectively,

$$W_e = \sum_{n \text{ even}, |n|\in\mathbb{N}_N} w_n^2, \tag{58}$$

$$W_o = \sum_{n \text{ odd}, |n|\in\mathbb{N}_N} w_n^2. \tag{59}$$

The model sinusoidal CH vector can thus be expressed as the scaled and rotated sum of the model wavelets,

$$\mathbf{W}\mathbf{f}_{\mathrm{S}}(A, \phi, \theta) = \lambda_e \mathbf{S}_\theta \mathbf{W}\mathbf{f}_e + \lambda_o \mathbf{S}_\theta \mathbf{W}\mathbf{f}_o, \tag{60}$$

where

$$\lambda_e = \sqrt{W_e}\, A\cos\phi, \tag{61}$$

$$\lambda_o = \sqrt{W_o}\, A\sin\phi. \tag{62}$$

An example of the two types of model wavelets is shown in Fig. 3, generated from a Simoncelli-type primary wavelet for different values of N. The wavelets for $N = 1$ are the monogenic signal wavelets. As N increases the wavelets elongate along the axis perpendicular to the sinusoid orientation.

5.2 Model Solution

The proposed sinusoidal model of an arbitrary image, f, with the point of interest located at $\mathbf{z} = \mathbf{0}$, and localised by an isotropic wavelet, ψ_i, is

$$(f * \psi_i)(\mathbf{z}) = \underbrace{A\cos(\langle \mathbf{z}, \mathbf{o}\rangle + \phi)}_{f_{\mathrm{S}}} + f_\epsilon(\mathbf{z}). \tag{63}$$

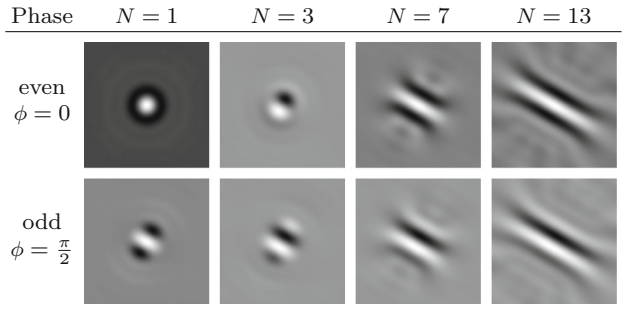


Fig. 3 Sinusoidal matched wavelets for different N, even or odd phase, and $\theta = \pi/3$. Other phase values are obtained by a linear combination of both wavelets. For $N = 1$ the even wavelet has no directionality

Using the sinusoidal model wavelets previously derived, we may now write the image CH vector as the sum of model and residual components,

$$\mathbf{Wf} = \mathbf{Wf}_{\mathrm{S}}(A, \phi, \theta) + \mathbf{W}\boldsymbol{\epsilon} \tag{64}$$

$$= \lambda_e \mathbf{S}_\theta \mathbf{Wf}_e + \lambda_o \mathbf{S}_\theta \mathbf{Wf}_o + \mathbf{W}\boldsymbol{\epsilon}. \tag{65}$$

The model parameters are solved for by minimising the residual, $\|\mathbf{W}\boldsymbol{\epsilon}\|$. Since the model vectors are orthonormal and have the same orientation the solution is given by the method in Sect. 4.4. Orientation is thus

$$\theta = \max_\theta \lambda_e(\theta)^2 + \lambda_o(\theta)^2 \tag{66}$$

where

$$\lambda_e(\theta) = \mathbf{Wf}_e^H \mathbf{S}_\theta^H \mathbf{Wf} \tag{67}$$

$$= \langle \mathbf{Wf}, \mathbf{S}_\theta \mathbf{Wf}_e \rangle, \tag{68}$$

$$\lambda_o(\theta) = \langle \mathbf{Wf}, \mathbf{S}_\theta \mathbf{Wf}_o \rangle. \tag{69}$$

The function $p(\theta) = \lambda_e(\theta)^2 + \lambda_o(\theta)^2$ is a trigonometric polynomial with degree $4N$. However, since $\mathbf{f}_e$ only has non-zero even orders, and $\mathbf{f}_o$ only has non-zero odd orders, $p(\theta)$ will only have non-zero even coefficients. Therefore it can be written as a degree $2N$ trigonometric polynomial in 2θ.

$$p(2\theta) = \lambda_e(\theta)^2 + \lambda_o(\theta)^2 \tag{70}$$

Solving for θ thus gives estimates in the range $[0, \pi)$. The method used depends on the maximum order, N. When $N = 1$, the only possible values for W where $W_e = W_o$ are $w_0 = 1/\sqrt{2}$ and $w_1 = 1/2$. The resulting wavelets are the monogenic wavelets [51] and the sinusoidal model parameters can be derived analytically without root finding, as follows:

$$A = \sqrt{2}\|\mathbf{Wf}\|, \tag{71}$$

$$\phi = \arg(f_0 + \mathrm{i}\,|f_1|), \tag{72}$$

$$\theta = -\arg(-\mathrm{i}\, f_1), \tag{73}$$

where $\phi \in [0, \pi)$ and $\theta \in [-\pi, \pi)$. Note that for a sinusoidal model, a rotation of π radians is equivalent to a sign change of the phase. For example, a sinusoid with parameters $\{\phi = \pi/2, \theta = 0\}$ is equivalent to one with $\{\phi = -\pi/2, \theta = \pi\}$. Therefore two ranges for phase and orientation can be used interchangeably for the model; one can either restrict orientation to the half circle, or restrict phase to the half circle.

For larger N, finding the maximum typically involves finding the roots of the derivative of $p(2\theta)$. The roots are candidate values for the orientation which corresponds to the maximum. For $N = 2$, the polynomial is fourth degree and the roots can be solved for analytically using quartic solvers. For larger orders a numerical solution is required in most cases.

Once θ has been found, we have

$$A\cos\phi = \frac{\lambda_e(\theta)}{\sqrt{W_e}}, \tag{74}$$

$$A\sin\phi = \frac{\lambda_o(\theta)}{\sqrt{W_o}}, \tag{75}$$

and thus amplitude and phase are given by

$$A = \sqrt{\frac{\lambda_e(\theta)^2}{W_e} + \frac{\lambda_o(\theta)^2}{W_o}}, \tag{76}$$

$$\phi = \arg\left(\frac{\lambda_e(\theta)}{\sqrt{W_e}} + \mathrm{i}\frac{\lambda_o(\theta)}{\sqrt{W_o}}\right), \tag{77}$$

where $A \in \mathbb{R}^+$ and $\phi \in [-\pi, \pi)$. Finally, the residual vector is given by

$$\mathbf{W}\boldsymbol{\epsilon} = \mathbf{Wf} - \mathbf{Wf}_{\mathrm{S}}(A, \phi, \theta). \tag{78}$$

We may synthesise the image separately from the model and residual coefficients by

$$f(\mathbf{z}) = f_{\mathrm{S}}(\mathbf{z}) + f_\epsilon(\mathbf{z}), \tag{79}$$

where

$$f_{\mathrm{S}}(\mathbf{z}) = \sum_{i,\mathbf{k}} \sum_{|n| \in \mathbb{N}_N} (\mathbf{Wf}_{\mathrm{S}_{i,\mathbf{k}}})_n w_n \psi^n_{i,\mathbf{k}}, \tag{80}$$

$$f_\epsilon(\mathbf{z}) = \sum_{i,\mathbf{k}} \sum_{|n| \in \mathbb{N}_N} (\mathbf{W}\boldsymbol{\epsilon}_{i,\mathbf{k}})_n w_n \psi^n_{i,\mathbf{k}}. \tag{81}$$

5.3 Example Solution

A pyramidal decomposition of the *Pentagon* image into four scales plus a low-pass component is shown in Fig. 4. The steerable wavelet frame was generated using a Simoncelli-type isotropic wavelet [40] and the 0th to 7th-order RTs.

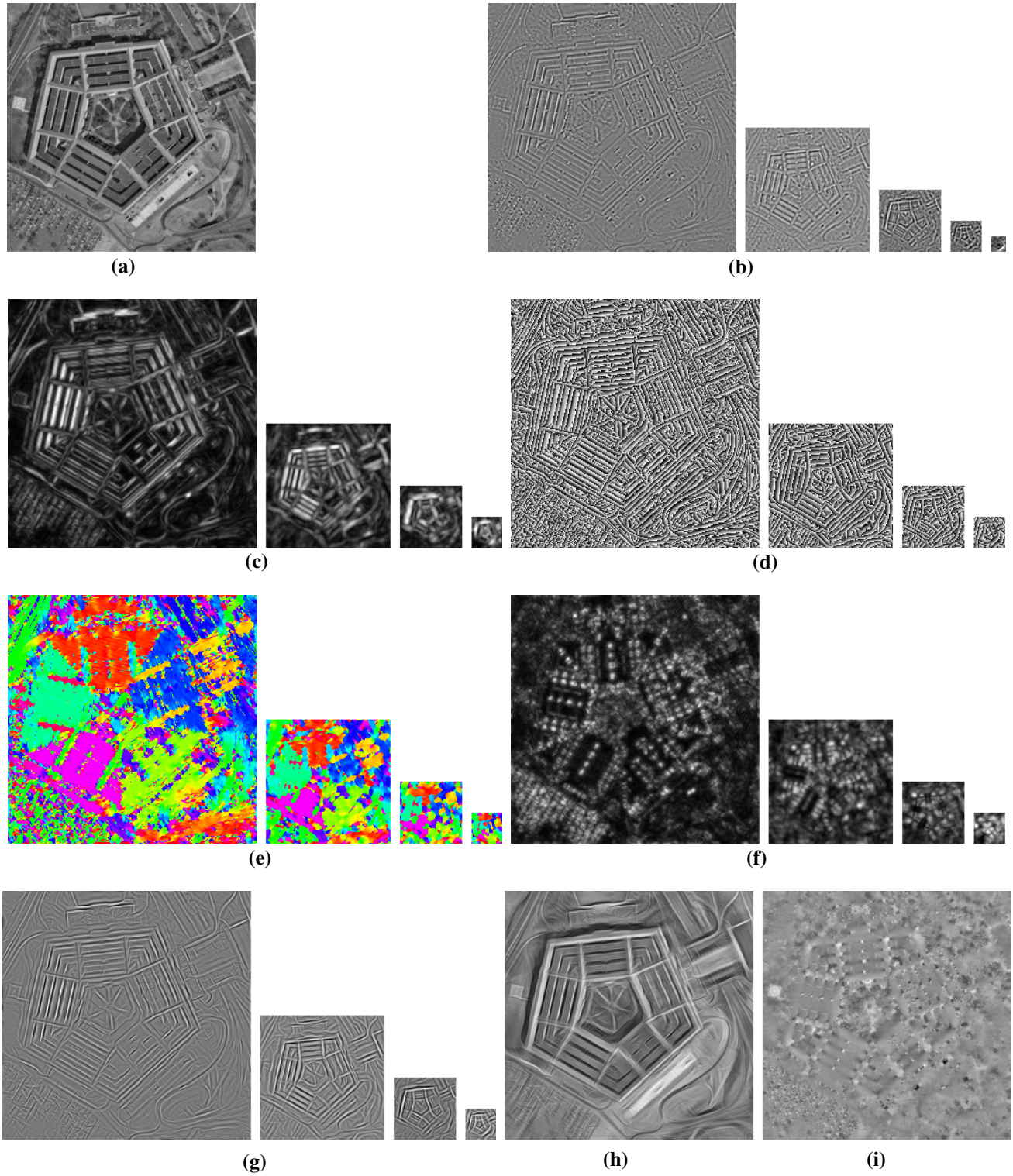


Fig. 4 Decomposition of a 512 × 512 pixel version of the *Pentagon* image into amplitude, phase, orientation and residual components over four scales using a pyramidal Simoncelli-type wavelet frame and $N = 7$. **a** Original image. **b** Isometric wavelet response (4 scales + low-pass). **c** Amplitude. **d** Phase. **e** Orientation. **f** Norm of residual. **g** Model Reconstruction for each scale. **h** Model reconstruction. **i** Residual reconstruction

Each scale was sub-sampled by two, and the odd and even orders were each equally weighted such that $W_e = W_o$ (98). The amplitude at each scale is high at the location of linear features, indicating that these are well described by the sinusoidal model. The phase value describes the symmetry of the local image structure, independently of the amplitude. The orientation shows the main axis of symmetry regardless of the phase or amplitude values. The *spilt of identity* property [10] of phase-based image representations is preserved.

The image was also reconstructed from the model for each scale (Fig. 4g), as well as separately from the model and residual components (Fig. 4h, i). Reconstruction from the model acts as a wide-band rotation-invariant line and edge filter. In contrast, the residual reconstruction contains features which have multiple axes of symmetry, such as corners and junctions, therefore these are not well-represented by a sinusoidal model.

5.4 Choice of Basis Filter

Three choices must be made when applying the model:

- The primary isotropic basis filter, $\psi(\omega)$.
- The number of RT orders, N.
- The values for the weights, $\mathbf{W}$.

To construct CH wavelet frames, the primary isotropic basis filter must have at least N vanishing moments [51]. Wavelets such as the Simoncelli, Papadakis and variance optimised wavelets (VOW) [37] satisfy these conditions; however, they contain discontinuities in the frequency domain that lead to a slow decay in the spatial domain [56]. In contrast, the second Meyer wavelet (Example 2, Sect. III C in [6]) is smooth and may therefore be a better choice.

For filter banks, the basis filter should also have the minimum number of vanishing moments and a smooth frequency profile to ensure fast decay. The log-Gabor filter [12] is often used for quadrature filters, as it is possible to construct with a large bandwidth frequency response and zero mean. In [4] it is shown that the difference-of-Gaussian (DoG) and Cauchy ($h(\omega) = n_c\,\omega^a e^{-\sigma\omega}$) quadrature (Hilbert transform) filters are better for edge detection [4]. However, the DoG filter has only one vanishing moment and a large minimum bandwidth, while the number of vanishing moments of the Cauchy filter is dependent on its bandwidth. The log-Gabor filter has infinite vanishing moments and thus remains an suitable choice for RT-derived filters. An extension on the log-Gabor filter is,

$$h(\omega) = \exp\left(-\left|\frac{\log^a(\omega/\omega_0)}{a\log^a(\sigma)}\right|\right), \tag{82}$$

where increasing a gives a more compact frequency response and shorter tail. The normal log-Gabor filter is given by $a = 2$.

5.5 Effect of N

Higher-order CH wavelets have a higher order of rotational symmetry. Therefore, increasing N increases the complexity of the local signal structure that the CH vector can represent. However, higher-order CH wavelets also have a larger spatial extent, increasing the size of the local image patch under consideration. This is because the magnitude of the radial frequency response remains constant due to the RT.

The model wavelets thus also increase in size with increasing N. An example of the two types of sinusoidal model wavelets for different values of N is shown in Fig. 3. It shows the even monogenic wavelet ($N = 1$) has no directionality, hence the problem with resolving orientation near even structures. As N increases, the wavelets become elongated along the axis perpendicular to their orientation, becoming more orientation selective due to a narrower angular profile but only responding to longer linear features.

5.5.1 Noise

To quantify the effect of increasing N on sinusoidal model accuracy, the amplitude, phase and orientation were calculated for a zero mean, 512×512 pixel sinusoidal image with different levels of additive white Gaussian noise. An all-pass basis filter was used and the CH vector was weighted using the phase-invariant equal weighting scheme (98). Figure 5 shows the mean error in estimated model parameters compared to the phase of the sinusoid for a 3dB signal-to-noise ratio (SNR). Increasing N decreased the average error for all parameters. However, both the amplitude and phase errors appear to reach a plateau around $N = 13$, after which increasing N gives little improvement. In contrast, the orientation estimate is particularly improved, with a ten-times reduction in error between $N = 3$ and $N = 13$. As expected, the orientation error is high for the monogenic signal ($N = 1$) at even locations, regardless of noise. For a SNR greater than 3dB, the errors vary proportionally with the noise standard deviation.

5.5.2 Qualitative Image Results

The effect of increasing N for a natural image is shown in Fig. 6. The sinusoidal model was calculated for the *Pentagon* image using a log-Gabor primary filter with wavelength 8 pixels, $\sigma = 0.65$, and $N \in \{1, 3, 7, 13, 21\}$. For the monogenic signal ($N = 1$), the amplitude is large and the phase is the same for both isometric features (blobs) and lines, and thus they cannot be differentiated from the model parameters

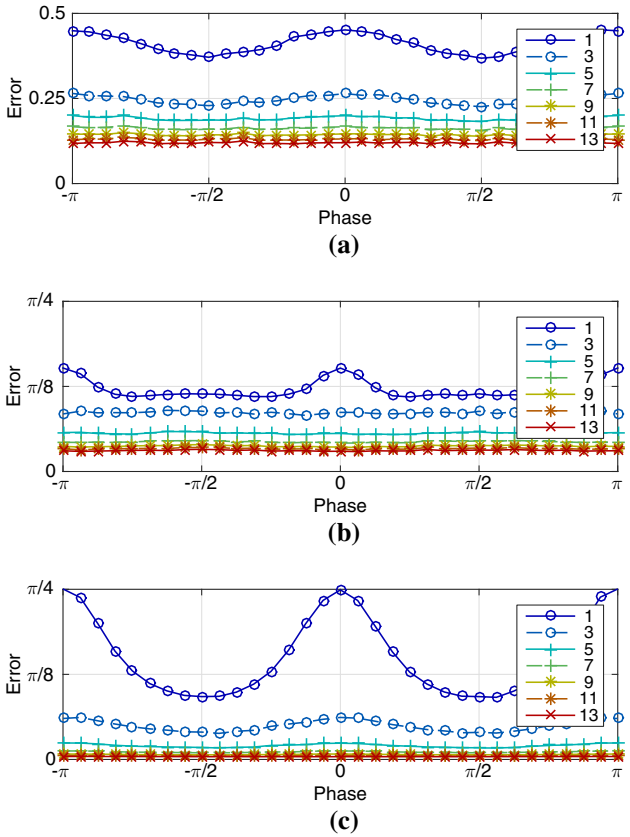


Fig. 5 Average model error for a sinusoidal image with added Gaussian noise (SNR: 3dB) for different N (*shown in legend*) compared to the original image phase. The phase-invariant equal weighting scheme was used. **a** Amplitude error. **b** Phase error. **c** Orientation error

alone. As N increases, the model becomes more selective for longer linear features due to the increasing elongation of the model wavelets. This is particularly noticeable going from $N = 7$ to $N = 21$, as the roof edges are no longer broken up. Blobs also have a reduced amplitude response; those in the lower right quadrant of the image have almost disappeared by $N = 13$ and instead the amplitude is large only at the location of linear features. Likewise, the residual norm is large at the location of corners and junctions which have multiple linear symmetries.

The orientation estimate also becomes smoother with increased N, and eventually larger features begin to dominate. In contrast, the orientation of curved lines appear to become more disjoint. The increased size of the wavelet, and thus local image patch, means that curves are less well modelled by a sinusoid at larger N. The residual norm images confirm this, showing a larger magnitude for curved structures as N increases. Overall, less of the image is well described by the sinusoidal model with larger N, and the residual norm increases for most locations. This effect can be seen in the separate reconstructions from the model and residual components, with more of the structure identifiable in the residual reconstruction. A qualitative assessment suggests that, for this image, $N = 7$ provides a good balance between the resolution of linear features and too much energy in the residual component.

5.5.3 Computation Time

Polynomial root finding is computationally expensive for large values of N, increasing the time to calculate the model parameters. Table 2 shows the calculation times for the *Pentagon* image and different values of N, implemented in MATLAB on a single core of a 2.5 GHz Intel Core i7 processor. Calculating the model takes much longer than a typical filtering operation, for example, 33.5 seconds for $N = 7$. To speed up the process, the root finding can be split across parallel cores, or a less accurate method used to estimate the polynomial maximum. We have previously proposed a quick approximation method in [25] and an updated version is described in Appendix A. It is an iterative process that makes q estimates of the orientation per order, and gives reasonably accurate results with much lower computation time, as shown in Table 2. For example, for $N = 7$ the orientation of 98% of the image can be estimated with better than 3 degrees accuracy in approximately 1.1 seconds.

For $N = 2$ the model parameters can be obtained analytically using a quartic solver, in under a second. Since $N = 2$ is the minimum number of orders to estimate orientation from both even and odd structures, it provides a good alternative to using the monogenic signal when speed is important.

5.6 Choosing Weights

The weighting matrix scales each RT order in the image CH vector, and therefore different weightings affect the CH vector magnitude and values of the model parameters for a given image structure. We shall choose the weights to fit a particular angular profile for the sinusoidal model.

The magnitude of the CH vector of a purely sinusoidal image is

$$\|\mathbf{W}\mathbf{f}_{\mathrm{S}}(A,\phi,\theta)\| \tag{83}$$

$$= A\|\cos\phi\mathbf{W}\mathbf{S}_\theta\,\mathbf{s}_e + \sin\phi\mathbf{W}\mathbf{S}_\theta\,\mathbf{s}_o\| \tag{84}$$

$$= A\sqrt{\cos^2\phi\,W_e + \sin^2\phi\,W_o} \tag{85}$$

$$= \begin{cases} A\sqrt{W_e}, & \phi = 0,\pi \\ A\sqrt{W_o}, & \phi = \pm\pi/2. \end{cases} \tag{86}$$

If $W_e \neq W_o$ the magnitude of the CH vector is affected by the phase. If we choose $W_e = W_o = 1/2$, the odd and even components are each weighted equally and the magnitude is invariant to phase. That is,

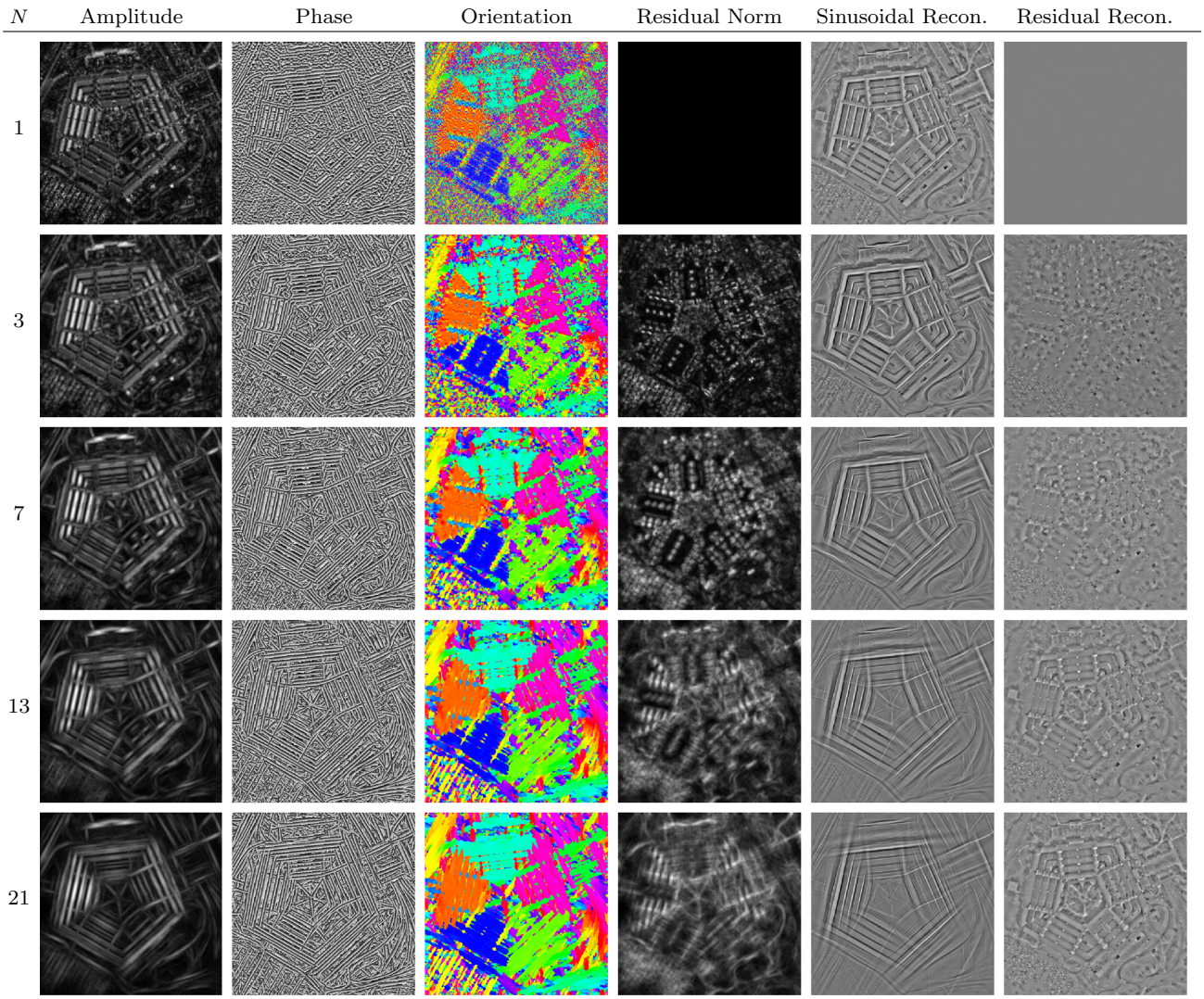


Fig. 6 Sinusoidal model parameters for the second scale of the *Pentagon* image for different values of N. Also shown is reconstruction from the sinusoidal model and reconstruction from the residual component, using four scales and not including the low-pass response

$$\|\mathbf{W}\mathbf{f}_S(A,\phi,\theta)\| = A/\sqrt{2}. \tag{87}$$

This is desirable as it preserves the invariance properties of previous approaches, for example, the monogenic signal vector magnitude is invariant to phase.

The weighting also determines the angular response of the model wavelets in the frequency domain, which is given by the trigonometric polynomial

$$h_u(\theta) = \sum_{|n|\in\mathbb{N}_N} w_n u_n e^{-in\theta}, \tag{88}$$

where **Wu** is the weighted model wavelet CH vector.

To pick the coefficients of **W** it is proposed to maximise the energy of the angular response of the sinusoidal wavelets inside of a window $h(\theta)$ by adapting the method described in [42,43,51] for designing *prolate spheroidal wavelets*. Each sinusoidal wavelet is 2nd-order rotationally symmetric or anti-symmetric and therefore has its angular response concentrated at two points π radians apart. The same energy window can be used for each. Let $v(2\theta)$ be a positive window function, symmetric at both $\theta = 0$ and $\theta = \pi$. That is,

$$v(\theta) = v(-\theta), \tag{89}$$

$$v(\theta) = v(\theta - \pi), \tag{90}$$

and let $u(\theta)$ describe the angular response of a wavelet, **u**, that is

$$u(\theta) = \sum_{n=-N}^{N} u_n u_n e^{-in\theta}. \tag{91}$$

Table 2 Orientation calculation setup and computation time (seconds) for the MATLAB roots method (and quartic solver* for $N = 2$), versus our approximation method for $q \in [1, \ldots, 5]$, for the 512 × 512 *Pentagon* image

N	Setup	Roots*	Quick method				
			q=1	q=2	q=3	q=4	q=5
2	0.3	0.4	0.1	0.2			
			(94)	(94)	(94)	(94)	(94)
3	0.4	18.9	0.2	0.3	0.6		
			(94)	(96)	(97)	(97)	(97)
5	0.7	24.2	0.2	0.4	0.8	2.8	19.1
			(93)	(97)	(97)	(98)	(98)
7	0.8	33.5	0.3	0.6	1.1	3.0	19.1
			(92)	(96)	(98)	(98)	(99)
9	1.4	44.3	0.4	0.7	1.1	3.0	18.4
			(90)	(96)	(97)	(98)	(99)
11	1.4	57.9	0.4	0.8	1.4	3.3	19.0
			(89)	(95)	(97)	(98)	(99)
13	1.6	73.6	0.5	1.0	1.6	3.6	19.2
			(88)	(94)	(96)	(97)	(98)

Percentage of errors under 3 degrees is shown in parentheses

Then the energy within the window is given by [42,43,51]

$$E[w] = \int_{-\pi}^{\pi} u(\theta)^2 v(\theta)\, d\theta \tag{92}$$

$$= \sum_{n'=-N}^{N} \sum_{n=-N}^{N} \bar{u}_{n'} u_n \int_{-\pi}^{\pi} v(\theta)\, d\theta \tag{93}$$

$$= \mathbf{u}^H \mathbf{V} \mathbf{u}, \tag{94}$$

where $V_{n,n'} = \int_{-\pi}^{\pi} v(\theta)\, d\theta$ since $\mathbf{u}^H \mathbf{u} = 1$.

Two types of orthogonal symmetric functions that fit this window are an even function with extrema of the same sign at 0 and π, and an odd function with extrema of opposite signs at 0 and π, where $u(\theta) = -u(\theta - \pi)$. The eigenvectors corresponding to the largest two eigenvalues of $\mathbf{V}$ thus describe the even and odd wavelets in the sinusoidal model. Let $\mathbf{u}_1$ and $\mathbf{u}_2$ be these eigenvectors. Each either has only odd orders, or only even orders. The final weighting is given by the absolute value of each order scaled by $\sqrt{2}$, since both are of unit norm. That is,

$$w_n = \frac{|u_{1_n} + u_{2_n}|}{\sqrt{2}}, \tag{95}$$

and thus $\|\mathbf{w}\| = 1$ and the weighting matrix is therefore $\mathbf{W} = \mathrm{diag}(\mathbf{w})$.

A simple window function consists of two rectangular functions with angular width B separated by π,

$$v(\theta) = \mathrm{rect}\left(\frac{\theta}{B}\right) + \mathrm{rect}\left(\frac{\theta + \pi}{B}\right). \tag{96}$$

The values of $\mathbf{V}$ for this function are

$$V_{n,n'} = \begin{cases} 2B, & n - n' = 0, \\ \frac{2\sin(B(n-n'))}{n-n'}, & n - n' \text{ is even}, \\ 0, & n - n' \text{ is odd}, \end{cases} \tag{97}$$

where B is the width of the rectangle in radians. When B approaches 0 the even components each become equally weighted, as do the odd, so that $W_e = W_o$. This equal weighting scheme is given by

$$w_n = \begin{cases} \frac{1}{\sqrt{2(N+1)}} & \text{if } (N - n) \text{ is even}, \\ \frac{1}{\sqrt{2N}} & \text{if } (N - n) \text{ is odd}. \end{cases} \tag{98}$$

Figure 7 shows an example of the even and odd angular profiles for different values of B and $N = 7$. For smaller B, the angular response has a narrower peak but larger oscillations. For larger B, the response is smoother and wider, but

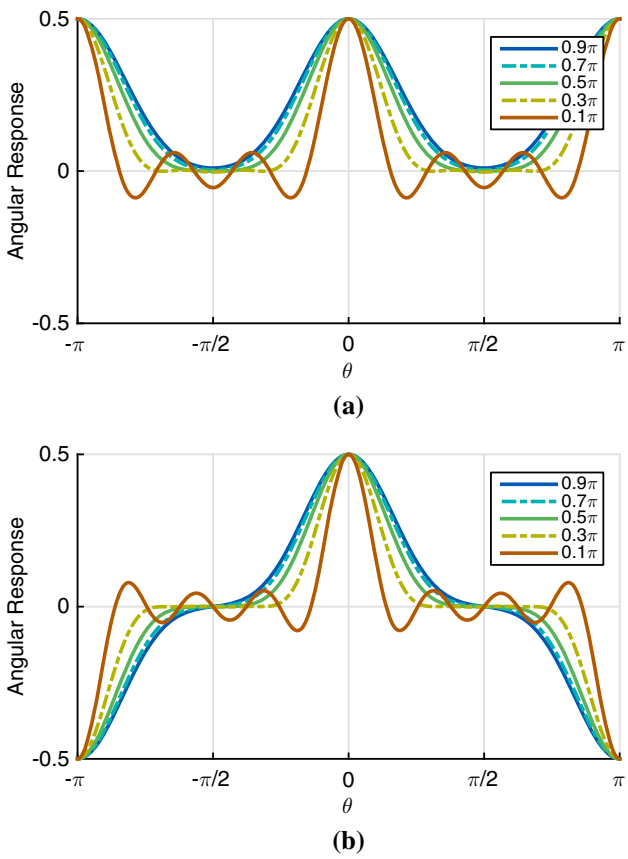


Fig. 7 Angular response of odd and even sinusoidal model wavelets ($N = 7$) for different window widths, B (*shown in legend*). **a** Even wavelet ($\phi = 0$). **b** Odd wavelet ($\phi = \pi/2$)

less orientation selective. To quantify the amount of oscillation, an experiment was performed to measure the ratio of the energy under the side lobes to the total energy for different values of N and B. The ratio was less than 0.1% (indicating small oscillations) when $B > 5.64/N - 6.57/N^2$. For $N = 7$ this equals approximately 0.21π.

Repeating the noisy sinusoidal image experiment (Sect. 5.5.1) for different values of B we found the equal weighting scheme ($B \approxeq 0$) gave the best results. This is likely due the narrowness of the main lobe, and the even distribution of noise energy in the side lobes negating their effect on the response. However, in natural images the local structure can have multiple discrete elements. In that case, it may be advantageous to have smaller oscillations (higher value of B) so that these extra parts interfere less with the main response. Furthermore, if one wished to model using multiple sinusoids, a larger B would be helpful to reduce the correlation between the model wavelets at different orientations.

6 Intrinsic Dimension

The advantage of a sinusoidal model is the parametrisation of the local image structure into amplitude, phase and orientation values which can be separately analysed. Deriving the model starting with the CH vector is different to other quadrature filter type methods in that we are left with a residual vector that describes the non-model part of the local image structure. From the split into model and residual components a representation of the local *intrinsic dimension* can be developed.

Intrinsic dimension described the linear symmetry of the local image structure. Flat areas with constant intensity are intrinsically 0D (i0D) as they can be described by a single value. Linear features, such as lines and edges, are intrinsically 1D (i1D), as they vary along a single axis and can be represented by a 1D function. More complex structures, such as corners and junctions, have multiple symmetries and are intrinsically 2D (i2D) [10,21,59,61]. The strict definition is,

$$f(\mathbf{x}) \in \begin{cases} \{\text{i0D}\} & \text{if } f(\mathbf{z}) = \text{constant}, \\ \{\text{i1D}\} & \text{if } f(\mathbf{z}) = f(\langle \mathbf{z}, \mathbf{o} \rangle), \mathbf{o} = [\cos\theta, \sin\theta], \\ \{\text{i2D}\} & \text{otherwise.} \end{cases} \tag{99}$$

However, typical images do not have perfectly linear structures and are often corrupted by noise, meaning that most structures would be classified as i2D. Instead, a continuous representation of intrinsic dimensionality is necessary, which in turn requires a local structure descriptor that is able to discriminate classes. A classic example is the structure tensor [3,13] whose eigenvalues describe the strength of the local symmetry along two main axes. If only one eigenvalue is large, the local structure is i1D, if both are large it is i2D, otherwise it is i0D. This the basis behind the popular Harris corner detector [16]. In [9] a continuous representation was introduced that plots the eigenvalues onto a bounded triangle, whose barycentric co-ordinates give a probability of belonging to each class. The representation can be applied to the output of other descriptors that discriminate between i1D and i2D structures.

A drawback of the structure tensor is that it is only computed from 1st-order derivatives and therefore can have a double response for roof edges (thick lines) [19]. A improvement proposed was the boundary tensor [19], which gives a boundary energy value consisting of line / edge (i1D) and junction (i2D) energies, using the 0th- to 2nd-order RTs. In fact, the boundary energy is equal to the square of the CH vector norm for $N = 2$ with weighting $\mathbf{w} = \sqrt{[1/2, 1, 1, 1, 1/2]/4}$. Likewise, the boundary tensor i1D energy measure roughly corresponds to the energy of the sinusoidal model vector, and the boundary tensor junction energy to the energy of the residual vector for the same weighting and N. The monogenic curvature tensor [47] uses the same basis functions as the boundary tensor but adds phase and curvature measurements to give a richer description of the local structure.

The sinusoidal model appeared more descriptive of i1D structures and the residual component was higher around i2D locations for larger N for the *Pentagon* image (Fig. 6). This suggests that the model and residual components can be used to discriminate between these classes. Therefore we shall use the proposed sinusoidal model calculated using higher-order RTs to develop a representation of intrinsic dimension.

6.1 Model Response

The first step is to determine the sinusoidal model response to an i1D structure. Consider an image, f, that is locally i1D at a point of interest centred at $\mathbf{z} = \mathbf{0}$ when filtered by an isotropic wavelet, ψ. The local structure can be represented as 1D function,

$$(f * \psi)(\mathbf{z}) = f_{\text{i1D}}(x), \tag{100}$$

where $x = \langle \mathbf{z}, \mathbf{o} \rangle$ and $\mathbf{o} = [\cos\theta, \sin\theta]$ with θ being the orientation of symmetry. According to the Fourier slice theorem, the Fourier transform of the image patch will have all non-zero coefficients concentrated along a line through the origin. Therefore, the local image structure can be exactly modelled as a sum of sinusoids

$$(f * \psi)(\mathbf{z}) = \sum_k \alpha_k \cos(\omega_k x + \phi_k), \tag{101}$$

which can also be expressed as a single sinusoid,

$$(f * \psi)(\mathbf{z}) = A(x)\cos(\phi(x)), \tag{102}$$

which is equivalent to the analytic signal representation of $f_{\mathrm{i1D}}(x)$, with local amplitude, A, and phase, ϕ. It follows that the overall RT responses are given by the sum of the RT responses of the individual sinusoids, and thus

$$\mathcal{R}^n(f * \psi)(\mathbf{z}) = \begin{cases} A(x)e^{\mathrm{i}n\theta}\cos(\phi(x)) & n \text{ is even}, \\ A(x)e^{\mathrm{i}n\theta}\,\mathrm{i}\sin(\phi(x)) & n \text{ is odd}. \end{cases} \tag{103}$$

The magnitudes of all the even-order responses are equal, as are all the odd-order responses. Let $\mathbf{f}_{\mathrm{i1D}}$ be the CH vector generated for this image structure at $\mathbf{z} = \mathbf{0}$, using ψ as the primary isotropic wavelet. The structure is completely described by the sinusoidal model, that is,

$$\mathbf{f}_{\mathrm{i1D}} = \mathbf{f}_{\mathrm{S}}(A, \phi, \theta) \tag{104}$$

$$= A\mathbf{S}_\theta(\cos\phi\,\mathbf{s}_e + \sin\phi\,\mathbf{s}_o). \tag{105}$$

It follows that the magnitude of the residual component is zero,

$$\min_{A,\phi,\theta} \|\boldsymbol{\epsilon}\| = \min_{A,\phi,\theta} \|\mathbf{f}_{\mathrm{i1D}} - \mathbf{f}_{\mathrm{S}}\| \tag{106}$$

$$= 0, \tag{107}$$

which means that an i1D signal can be completely reconstructed from the sinusoidal model wavelets rotated to the same orientation.

6.2 Complex Exponential Representation

Since the sinusoidal model can completely represent an i1D signal, it follows that the residual component represents the other parts of the local image structure with a different shape or orientation. The ratio of the residual vector norm to the model vector norm is therefore a measure of where the overall local structure lies on a i1D to i2D scale. Furthermore, it is invariant to the magnitude of the local structure (CH vector norm). We can represent this relationship in the form of a complex exponential, $\mathbf{d}_0$, given by

$$\mathbf{d}_0 = \|\mathbf{W}\mathbf{f}_{\mathrm{S}}\| + \mathrm{i}\|\mathbf{W}\boldsymbol{\epsilon}\| \tag{108}$$

$$= \|\mathbf{W}\mathbf{f}\|e^{\mathrm{i}\gamma_0}, \tag{109}$$

where $\gamma_0 = \tan^{-1}\frac{\|\mathbf{W}\boldsymbol{\epsilon}\|}{\|\mathbf{W}\mathbf{f}_{\mathrm{S}}\|}$ is the angle representing the ratio between $\|\mathbf{W}\boldsymbol{\epsilon}\|$ and $\|\mathbf{W}\mathbf{f}_{\mathrm{S}}\|$.

While it is possible to have a zero residual response, it is not possible to have a zero sinusoidal model response, as the model wavelets will always positively correlate with some part of the local image structure. This means that the upper bound of possible values of γ_0 will always be less than $\pi/2$, and can change according to the number of orders and weighting scheme. For a purely sinusoidal signal, we have $\mathbf{W}\mathbf{f} = \mathbf{W}\mathbf{f}_{\mathrm{S}}$ and $\|\mathbf{W}\boldsymbol{\epsilon}\| = 0$, and therefore $\gamma_0 = 0$ is the lower bound.

By finding this upper bound, we can adjust γ_0 so that the range is always $[0, \pi/2)$ regardless of weighting, and by extension, N. Consider an image CH vector, $\mathbf{f}$, that is zero for every order except for n, and without loss of generality let $f_{|n|} = 1$. Calculating the sinusoidal signal model for a weighting scheme, $\mathbf{W}$, we obtain

$$\|\mathbf{W}\mathbf{f}_{\mathrm{S}}\| = \begin{cases} 2w_n^2/\sqrt{W_e} & n \neq 0, \text{even} \\ 2w_n^2/\sqrt{W_o} & n \neq 0, \text{odd} \\ w_n^2/\sqrt{W_e} & n = 0. \end{cases} \tag{110}$$

Since $\|\mathbf{W}\boldsymbol{\epsilon}\| = \sqrt{\|\mathbf{W}\mathbf{f}\|^2 - \|\mathbf{W}\mathbf{f}_{\mathrm{S}}\|^2}$,

$$\gamma = \begin{cases} \tan^{-1}\sqrt{\frac{W_e}{2w_n^2} - 1} & n \neq 0, \text{even} \\ \tan^{-1}\sqrt{\frac{W_o}{2w_n^2} - 1} & n \neq 0, \text{odd} \\ \tan^{-1}\sqrt{\frac{W_e}{w_n^2} - 1} & n = 0. \end{cases} \tag{111}$$

The upper bound, $\gamma_{\max}$, is the maximum value of γ in the above equation for all n. Note that as N increases, w_n tends to get smaller, and therefore $\gamma_{\max}$ gets closer to $\pi/2$. The new intrinsic dimension representation is thus

$$\mathbf{d}_1 = \|\mathbf{W}\mathbf{f}\|e^{\mathrm{i}\gamma_1} \tag{112}$$

$$\text{where } \gamma_1 = \frac{\gamma_0}{\gamma_{\max}}\frac{\pi}{2}. \tag{113}$$

6.3 i2D Detection

The norm of the residual component is large around corners and junctions (Fig. 6). Thus the next step is to use the i2D part of the intrinsic dimension representation as a corner and junction detector, in the same way the junction energy is used for the boundary tensor [19]. The proposed detection measure is thus the imaginary part of the intrinsic dimension:

$$\det_{\mathrm{i2D}}(\mathbf{d}) = \mathrm{Im}\{\mathbf{d}\} \tag{114}$$

$$= \|\mathbf{W}\mathbf{f}\|\sin(\gamma). \tag{115}$$

However, common i2D features can have a large i1D component. For example, a T junction will give a large sinusoid amplitude representing the top bar that a Y junction will not. The angle, γ_1, will therefore be different at the centre of these two junctions. To compensate we shall rescale γ_1 using a sigmoidal function so that the angle values are closer together. The new representation is

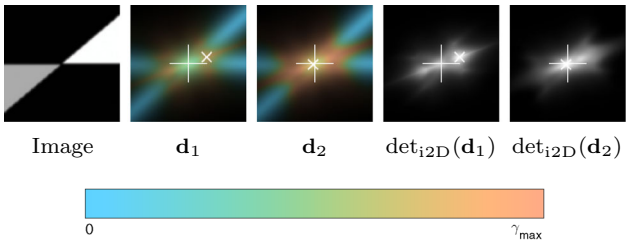


Fig. 8 Intrinsic dimension representation and i2D detection measure before and after rescaling γ_1 for $N = 7$. In the 2nd and 3rd images, brightness corresponds to $\|\mathbf{d}\|$ and colour to γ. The actual feature centre is indicated by a plus; the maximum of the detection measure is indicated by an X

$$\mathbf{d}_2 = \|\mathbf{W}\mathbf{f}\| e^{\mathrm{i}\gamma_2}, \tag{116}$$

where $\gamma_2 = \pi/2 \times \eta(2\gamma/\pi, h, s)$ and $\eta(x, h, s)$ is a sigmoidal function given by the regularised incomplete gamma function as

$$\eta(x, h, s) = \begin{cases} I_\gamma(x, s, s/h - s)) & h \leq 0.5, \\ 1 - \eta(1 - x, 1 - h, s) & h > 0.5, \end{cases} \tag{117}$$

with $x \in [0, 1]$ and $h \in [0, 1]$. The gamma function was chosen because the slope and position of the curve can be easily manipulated and the output values cover the complete range from 0 to 1. In the above equation, h roughly corresponds to the halfway point, that is, for $x = h$, $g(x, h, s) = 0.5$. Increasing the value of s increases the steepness of the slope at the halfway point. The effect of rescaling γ_1 for a chequer feature is shown in Fig. 8. The middle of the feature is less i2D than the surrounding areas, resulting in an off-centre detection. After rescaling γ_1, the detection energy is concentrated more towards the centre of the feature.

6.4 Corner and Junction Response

The intrinsic dimension representation, $\mathbf{d}_2$, was calculated for 15 corner and junction features using $N \in \{2, 3, 7, 13\}$, and sigmoidal function parameters $h = 1/3$ and $s = 2.4$, and is shown in Fig. 9. A depiction of intrinsic dimension using the structure tensor is given for comparison, using the largest eigenvalue as the i1D component, the smallest as the i2D component, and $\gamma_{\max} = \pi/4$. The images are 128×128 pixels in size, and were constructed from either line or wedge segments radiating from the centre pixel, followed by some Gaussian blurring. A log-Gabor filter with wavelength 32 pixels and $\sigma = 0.6$ was used to localise the model response, and a Gaussian filter with $\sigma = 6$ was used for the structure tensor. Brightness is equal to $\|\mathbf{d}_2\|$ which is the same as the norm of the image CH vector. Colour describes γ_2, with blue indicating i1D and red indicating i2D. An iso-luminant colour map from [20] was used to ensure correct perception.

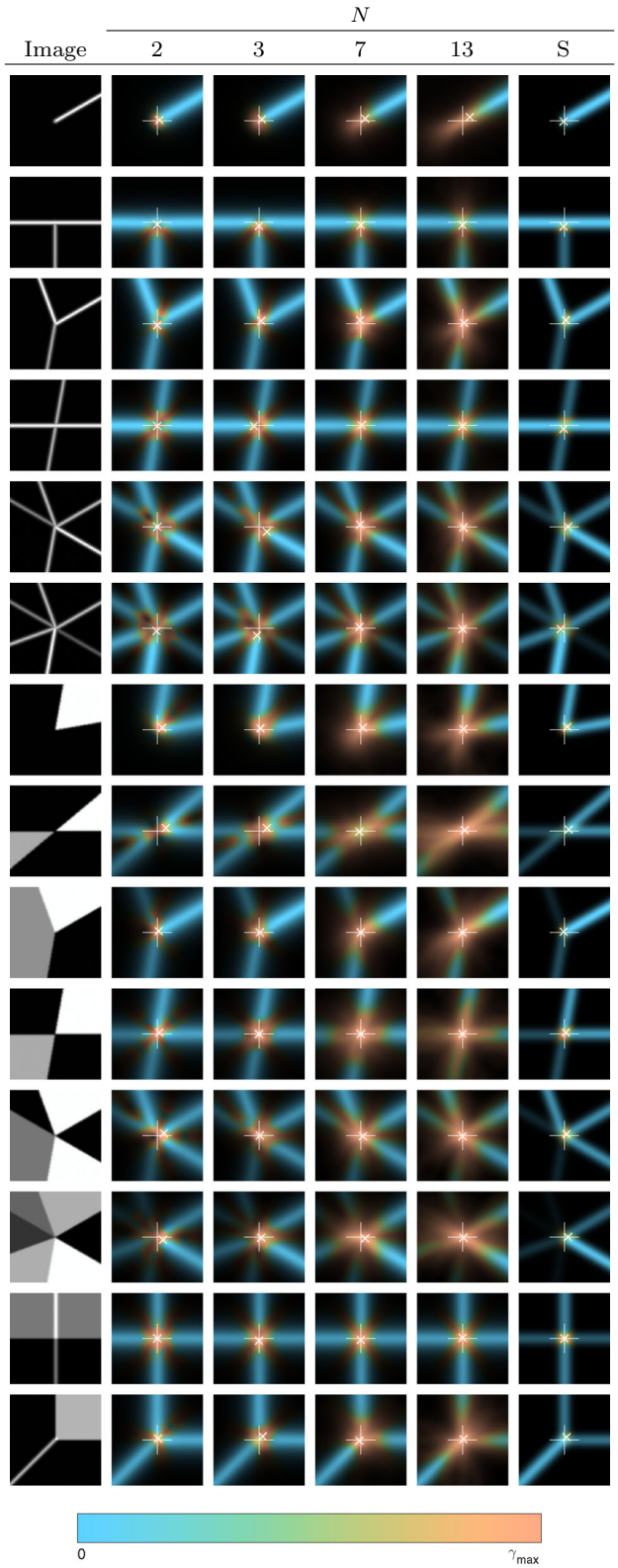


Fig. 9 Examples of image features along with their intrinsic dimension representation for different N and the structure tensor (S). Brightness represents magnitude, $\|\mathbf{d}_2\|$, colour represents angle, γ_2. Centre of the feature is indicated by a plus, location of the detection point is indicated by an X

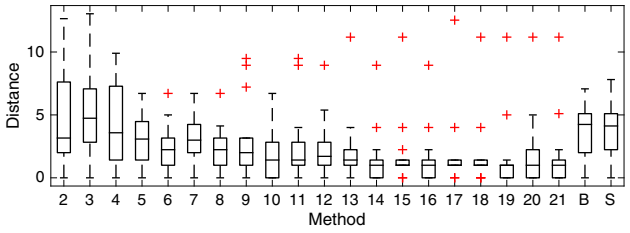


Fig. 10 Range of detection location errors (pixels) for the feature set in Fig. 9 using the sinusoidal model method for different N (numbered), compared to the boundary tensor (B) and the Harris corner detector from the structure tensor (S). The large outlier for higher values of N is due to the single line segment feature

The magnitude of response appears phase-invariant, as both line and edge features have similar patterns of magnitude and intrinsic dimension angle. It can be observed that for smaller N, both the magnitude and the extent of the i2D region is concentrated more towards the centre of the features, due to the smaller size of the wavelets. However, for smaller N there are also regions of low magnitude near the centre of features with more than two segments. This is particularly visible in the fifth and sixth images for $N = 2$. As N increases, the magnitude becomes more uniform, showing that a larger N is required for boundary estimation of complex features. However, this also causes a smearing of the magnitude response along the direction of line and edge segments, particularly noticeable in the seventh image as well as in Fig. 4. Again this is due to the increased size of the wavelets.

Apart for the single line segment image, detection position improves with increasing N (Fig. 9). Figure 10 shows the range of detection position errors for the same set of features for different N, compared to those using the boundary tensor junction energy and Harris corner detector from the structure tensor. The filter sizes were the same as for Fig. 9. As N increases, the position error decreases, and appears to plateau after about $N = 11$. Detecting the actual centre of the i2D feature is important for methods that use steerable filters to find the orientation of component line or edge segments, such as in [34] and our method in [27]. If the position is off-centre, the calculated orientations can be affected. Therefore, applying the proposed detection measure using larger N should be useful for these methods.

6.5 Test Set Evaluation

To gauge the repeatability of the detector under different image transformations, a grey-scale version of the test set from [30] was used. This test set has seen popular use for the evaluation of local descriptors and interest point detectors, such as in [2,31,45]. It consists of eight subsets of images, each with an original image and five or six transformed images. The transforms are viewpoint change (subset: *graffiti*, *bricks*), scale and rotation (subset: *boat*, *bark*), blur (subset: *bikes*, *trees*), illumination (subset: *cars*) and JPEG compression (subset: *ubc*).

The sinusoidal model was calculated over four scales using a log-Gabor filter with $\sigma = 0.6$ and wavelength of $\{4, 8, 16, 32\}$ pixels. The i2D detection measure without angle rescaling (113) was calculated for each scale and added together to give a final detection score. Adding the scores was found to give better results than choosing the maximum from each scale. Candidate detection points were chosen as the locations of the local maxima in a three pixel radius area. Computational load was almost wholly taken up with calculating the sinusoidal model, and thus the time to calculate the detections was approximately four times longer than that given in Table 2 for each N.

The detections from the original image in each subset were compared to each of the transformed images. Any points that were not within the area common to each image were discarded. Each detection point was considered matched if there was a corresponding detection point in the transformed image within a distance of three pixels. However multiple correspondences were not allowed. Repeatability was calculated as the average of the fraction of matched points in the first image and the fraction of matched points in the transformed image, as in [2]. These values were then averaged for each transformation type to give an overall score. The results are shown in Fig. 11a–e. Different levels of white Gaussian noise were also added to the original image from each subset and the repeatability calculated. The results for the top 100 points averaged across all subsets is shown in Fig. 11f.

The repeatability varied with increasing N, the type of transformation and number of points. For the viewpoint, scale and rotation, and illumination subsets, the middle range value of $N = 7$ had the highest repeatability for smaller numbers of points, whereas low range N values had the highest repeatability for larger numbers of points. For the blur subsets, any increase in N reduced the scores dramatically, and $N = 2$ gave the best results. In contrast, increasing N increased repeatability consistently for the JPEG subset, for less than 300 points. This is due to the increased size of the wavelets averaging out the block-like compression artefacts. Repeatability increased with N for the noise experiment for the same reason.

Using a value of $N = 7$ appears to give the best all-round score. Qualitative analysis of the detection images revealed that for larger values of N, detection of curved lines started to increase, due to the lower correlation of these features with the model wavelets because of their increased size. Using $N = 7$ appears to be a good compromise between having enough RT orders to discriminate more complex junctions and corners from i1D features, yet having compact enough model wavelets to follow the curved lines in the image sets.

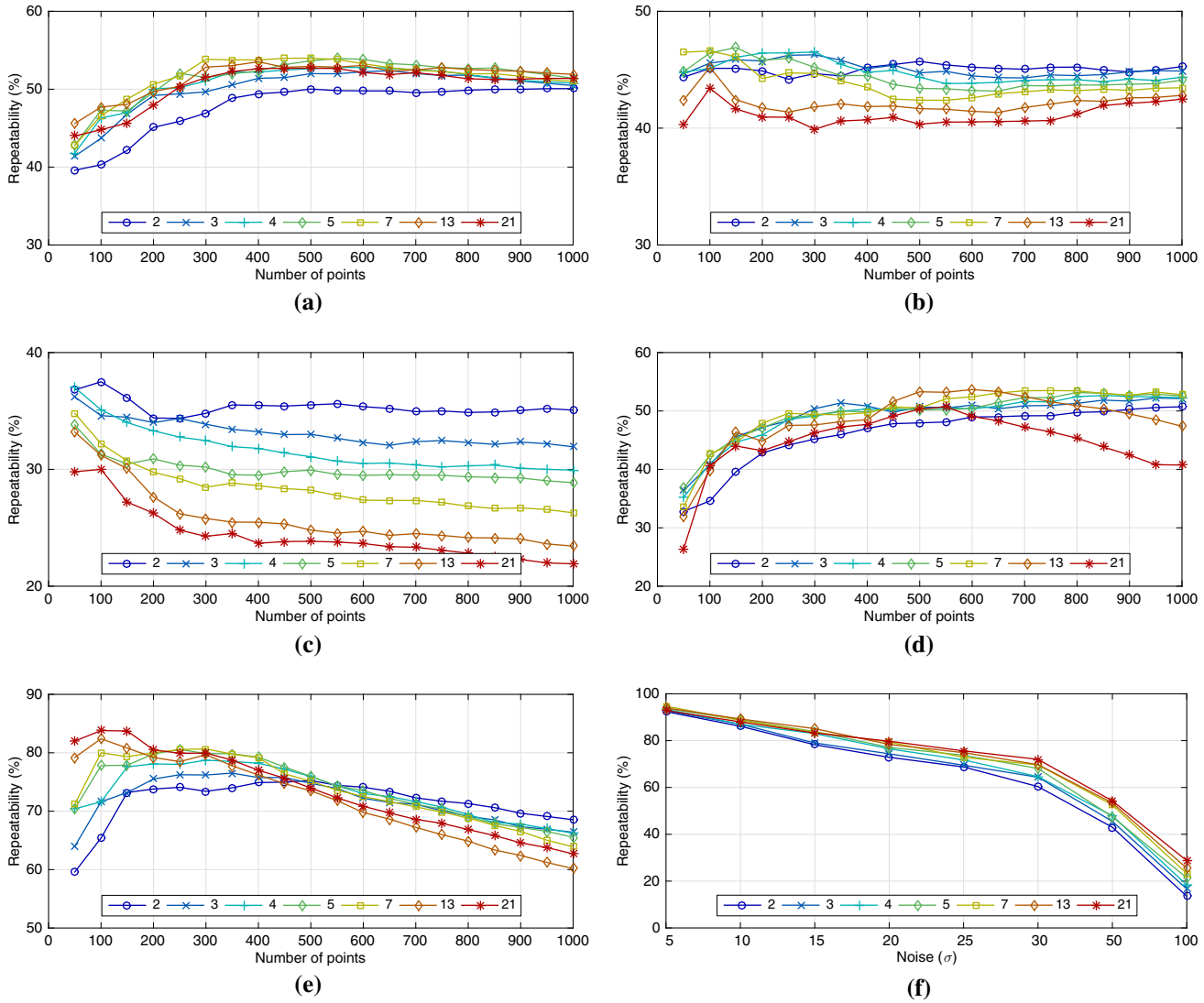


Fig. 11 Average i2D detection repeatability for different values of N (*shown in legend*) and number of detection points, evaluated for various image transformations (**a**–**e**) and additive Gaussian noise (**f**) using the test set from [30]

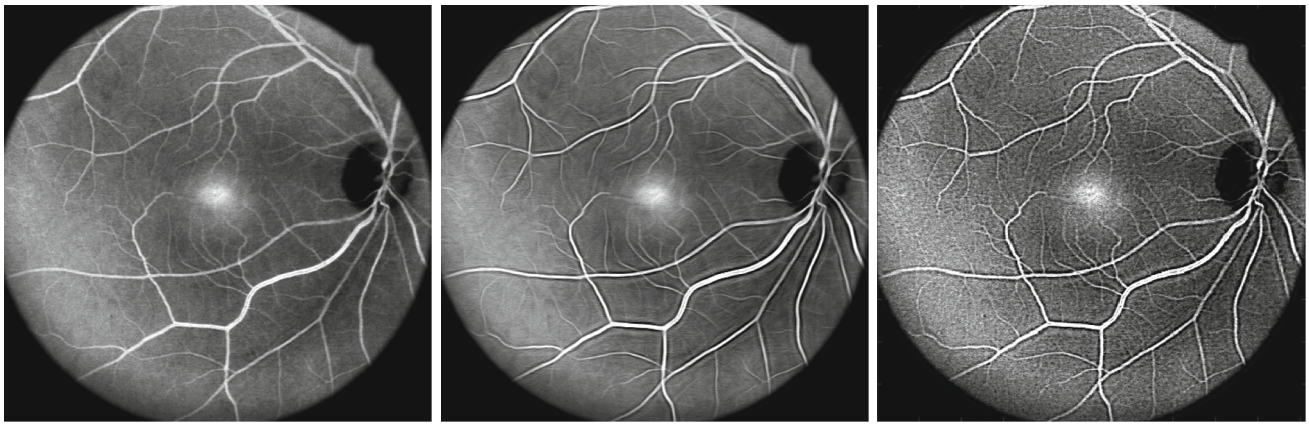
We also found that if only one scale (wavelength = 8 pixels) is used, repeatability remains generally the same except for a slight decrease in the blur subset. This shows that performing detection at a single scale is sufficient to capture most of the interest points in the test set, and has the benefit of reducing computation time.

Contour-based corner detectors along with the Harris-Laplace [31] and Laplacian-of-Guassian [23] detectors were tested in [2] on the same set of images. The results for $N = 7$ and 200 points were compared to that of the best detector in each subset in [2]. For our detection method, repeatability was approximately 20% better for the viewpoint subset, 15% better for the scale and rotation subset, same for the blur and JPEG subsets, and 20% worse for the illumination subset. The results are encouraging, and the detection method warrants further investigation.

7 Model-Driven Reconstruction

The previous section showed how splitting the local image structure into sinusoidal model and residual components can be used to discriminate i1D and i2D features, and the CH vector magnitude is a general boundary energy measure. The phase the sinusoidal model allow us to differentiate the symmetry of i1D structures, while the orientation indicates the direction of symmetry in a phase-invariant manner (Fig. 6).

One of the benefits of to using the RT is that the CH wavelets form a 2D steerable wavelet frame, and thus the image can be reconstructed from the CH vector. Furthermore, reconstructing an image from only the sinusoidal model components acts like a wide-band linear (i1D) filter, while reconstructing from only the residual components filters for i2D structures (Fig. 6). Modulating the model parameters



(a) **(b)** **(c)**

Fig. 12 *Retina* image (**a**) with linear features in the first two scales (**b**) and all features enhanced (**c**) using the second Meyer wavelet and $N = 7$

beforehand can be used to perform various image processing tasks. For example, in [17] brightness equalisation was performed by setting the amplitude of the monogenic signal to unity. In this section we shall demonstrate some useful applications of the model and its wavelet embedded for linear feature enhancement, orientation separation and image denoising.

7.1 Linear Feature Enhancement

Linear features can be enhanced by modulating the sinusoid and residual components separately for each scale. For example, the *Retina* image shown in Fig. 12a has small i1D features corresponding to blood vessels. To enhance these features, the model components of the first and second scales were amplified according to the intrinsic dimension angle of the second scale, and the residual components were attenuated (Fig. 12b). The formula applied was

$$\mathbf{Wf} = 3\cos(\gamma_1)\mathbf{Wf}_\mathrm{S} + 0.5\mathbf{W}\boldsymbol{\epsilon}, \tag{118}$$

with the model calculated using the second Meyer wavelet [6] and $N = 7$ (Fig. 12b). Amplifying the model and attenuating the residual enhances the i1D features without a corresponding increase in noise. In contrast, simply amplifying the first two scales by three (Fig. 12c) increases the noise in flat areas. Adding the intrinsic dimension term further reduces noise, as it ensures only model components in predominately i1D locations are amplified. The images are shown clamped to the original zero to 255 intensity range.

7.2 Orientation Separation

Selective reconstruction guided by the model parameters can be used to isolate different parts of the image. In particular, one can modulate the model component according to difference between the model orientation and a fixed orientation, θ_filter, by

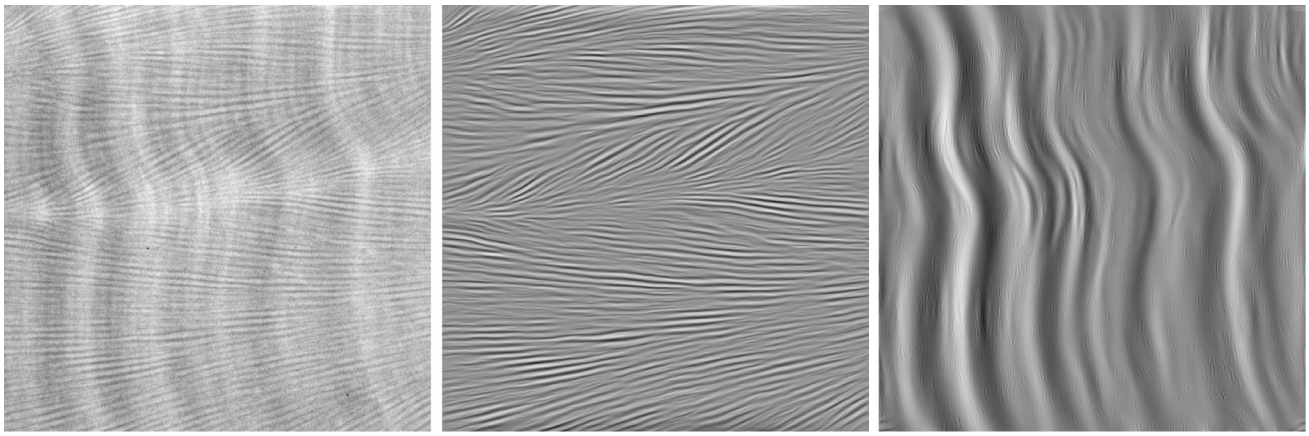
$$\mathbf{Wf} = |\theta - \theta_\mathrm{filter}|^\alpha \, \mathbf{Wf}_\mathrm{S}, \tag{119}$$

where a higher value for α penalises off-axis model components more. For example, the coral core x-ray image in Fig. 13a has both horizontal and vertical i1D features. An estimate of the orientation of the horizontal features was obtained by smoothing the sinusoidal model orientation at the second scale with a 20×20 pixel median filter. This estimate was used as θ_filter in (119) with $\alpha = 16$ to modulate the model components of the first five scales of the image, from which the image was reconstructed to isolate the horizontal components (Fig. 13b). The same estimate plus $\pi/2$ radians was also used to reconstruct the vertical components (Fig. 13c). The process acts like a wide-band orientation-adaptive filter.

7.3 Denoising

As a final example, we demonstrate wavelet denoising using an equiangular 2D steerable wavelet frame compared to using the model and residual components. One of the state-of-the-art approaches to wavelet-based image denoising is the *Bayesian least-squares—Gaussian scale mixture* (BLS-GSM) method [41]. This works by modelling the coefficients of each channel of a wavelet frame in a local area as the produce of a Gaussian random vector and a scalar multiplier. The model is compared to noise statistics known beforehand and the wavelet coefficients shrunk accordingly. BLS-GSM can be used with a frame consisting of $N + 1$ copies of either an even or odd 2D steerable wavelet rotated around the half circle at equiangular spacing [51].

The method cannot be used with the magnitude of the model and residual components, as it assumes zero-mean random vectors. Instead we propose to denoise the image



(a) **(b)** **(c)**

Fig. 13 *Coral* image (**a**) reconstructed from horizontal components (**b**) and vertical components (**c**). Five scales of the second Meyer wavelet with phase-invariant equal weighting and $N = 7$ were used

by shrinking the model and residual components separately, thus reconstructing from a modified image CH vector given by

$$\mathbf{Wf}_{i,\mathbf{k}}^{\text{denoised}} = \varphi(\|\mathbf{Wf}_{\mathrm{S}_{i,\mathbf{k}}}\|, \delta_{S_i})\mathbf{Wf}_{\mathrm{S}_{i,\mathbf{k}}} \tag{120}$$

$$+ \varphi(\|\mathbf{Wf}_{\mathrm{S}_{i,\mathbf{k}}}\|, \delta_{\epsilon_i})\mathbf{W}\boldsymbol{\epsilon}_{i,\mathbf{k}}, \tag{121}$$

where φ is a smooth threshold function dependent on the magnitude of the vector. It is given by

$$\varphi(x, \delta) = x - xe^{-(x/\delta)^8}. \tag{122}$$

There is a separate noise threshold for the sinusoidal component, δ_{S_i}, and the residual, δ_{ϵ_i}, for each scale. The value depends on the type and magnitude of the noise, choice of primary wavelet frame, CH vector weighting and number of RT orders.

The values are calculated from the CH vector magnitude and sinusoidal model magnitude distribution from a noisy image. Assuming additive white Gaussian noise, let g be a noisy image with standard deviation $\sigma = 1$. The image CH vectors, $\mathbf{Wg}_{i,\mathbf{k}}$, are calculated for a specific CH wavelet frame, weighting scheme and number of RT orders. The magnitude distribution of these vectors is the square root of the weighted sum of Rayleigh random variables (Nakagami distribution). Decomposing into model and residual vectors, $\mathbf{Wg}_{i,\mathbf{k}} = \mathbf{Wg}_{S_{i,\mathbf{k}}} + \mathbf{Wg}_{\epsilon_{i,\mathbf{k}}}$, we numerically estimate the mean, μ_i, and standard deviation, σ_i, of the image CH vector magnitudes for each scale. We also calculate the mean μ_{S_i} and standard deviation σ_{S_i} of the model CH vector magnitudes.

The model threshold chosen for a particular scale, i, and Gaussian noise level, σ_{noise} is then obtained as a function of the model magnitude noise distribution, while the residual is obtained from the CH vector magnitude distribution, as

$$\delta_{\mathrm{S}_i} = \sigma_{\text{noise}}(\mu_i + T\sigma_i), \tag{123}$$

$$\delta_{\epsilon_i} = \sigma_{\text{noise}}(\mu_{\mathrm{S}_i} + T\sigma_{\mathrm{S}_i}), \tag{124}$$

where T is the sole tuning parameter.

To incorporate information from the local area, rather than just at a point, the n-th largest magnitude value in the local area (ordinal filtering) is used in the shrinking function (121). This ensures that that there must be a larger number of high magnitude coefficients in the local area for the response to be passed, and also that smaller coefficients in the vicinity of larger ones are also passed more easily.

The proposed approach to denoising was compared BLS-GSM for three 512 × 512 pixel test images. Five scale of the VOW wavelets from [37] were used with $N = 7$ and phase-invariant equal weighting. The approaches compared were:

- BLS-GSM in a 3 × 3 pixel neighbourhood using the odd sinusoidal wavelet at eight equiangular orientations.
- BLS-GSM in a 3 × 3 pixel neighbourhood using both the odd and even sinusoidal wavelets at eight equiangular orientations.
- Shrinking using thresholding of the CH vector magnitude.
- Shrinking using thresholding of the model and residual components separately (121).
- Shrinking using thresholding of the model and residual components separately (121) and ordinal filtering by choosing the third largest value in a 3 × 3 pixel neighbourhood.

For shrinkage using the model, the tuning parameter, T, was varied and the best result selected. The best performing approach was BLS-GSM using both sinusoidal wavelets

Table 3 Denoising performance (PSNR) for three images using BLS-GSM with either an equiangular sinusoidal pair wavelet frame or odd sinusoidal wavelet frame, and soft-thresholding of the model and residual components, the model and residual components with ordinal filtering, or the CH vector magnitude

σ_{noise} / PSNR	15/24.61	25/20.17	50/14.15
Pentagon			
BLS-GSM sinusoidal	**30.88**	**28.50**	**25.85**
BLS-GSM odd	30.81	28.46	25.81
Model/resid	30.66 (2.6)	28.27 (2.6)	25.68 (3.2)
Model/resid + ord. filt.	30.66 (2.8)	28.27 (3.0)	25.75 (3.4)
CH vector	30.09 (3.6)	27.94 (3.6)	25.50 (4.0)
Barbara			
BLS-GSM sinusoidal	**31.72**	**29.06**	25.49
BLS-GSM odd	31.65	29.05	**25.51**
Model/resid	31.23 (2.8)	28.44 (2.6)	24.64 (2.8)
Model/resid + ord. filt.	31.32 (3.2)	28.59 (3.0)	24.77 (3.0)
CH vector	29.41 (4.0)	26.65 (3.8)	23.87 (3.6)
Boats			
BLS-GSM sinusoidal	**31.69**	**29.36**	**26.29**
BLS-GSM odd	31.61	29.29	26.25
Model/resid	31.18 (2.8)	28.97 (2.8)	25.93 (3.2)
Model/resid + ord. filt.	31.18 (3.2)	29.01 (3.2)	26.01 (3.4)
CH vector	30.65 (4.0)	28.51 (3.8)	25.74 (4.0)

Tuning parameter, T, is shown in parentheses. Bold numbers show best denoising performance (PSNR) for each image. Five scales of the VOW wavelet frame were used with $N = 7$ and phase-invariant equal weighting

(Table 3). Of the CH vector approaches, denoising the model and residual separately performed better than denoising using the CH vector as a whole. Ordinal filtering further improved the model results, and subsequent experiments found that increasing the size of the filtering neighbourhood also improved the score. Overall, this approach was within 0.51 dB of the results of the BLS-GSM method. The results are encouraging, considering that only the two model and residual components are used, compared to eight or 16 channels when using BLS-GSM with the odd or sinusoidal wavelets, respectively. It shows that much of the information in an image is represented by the sinusoidal model component.

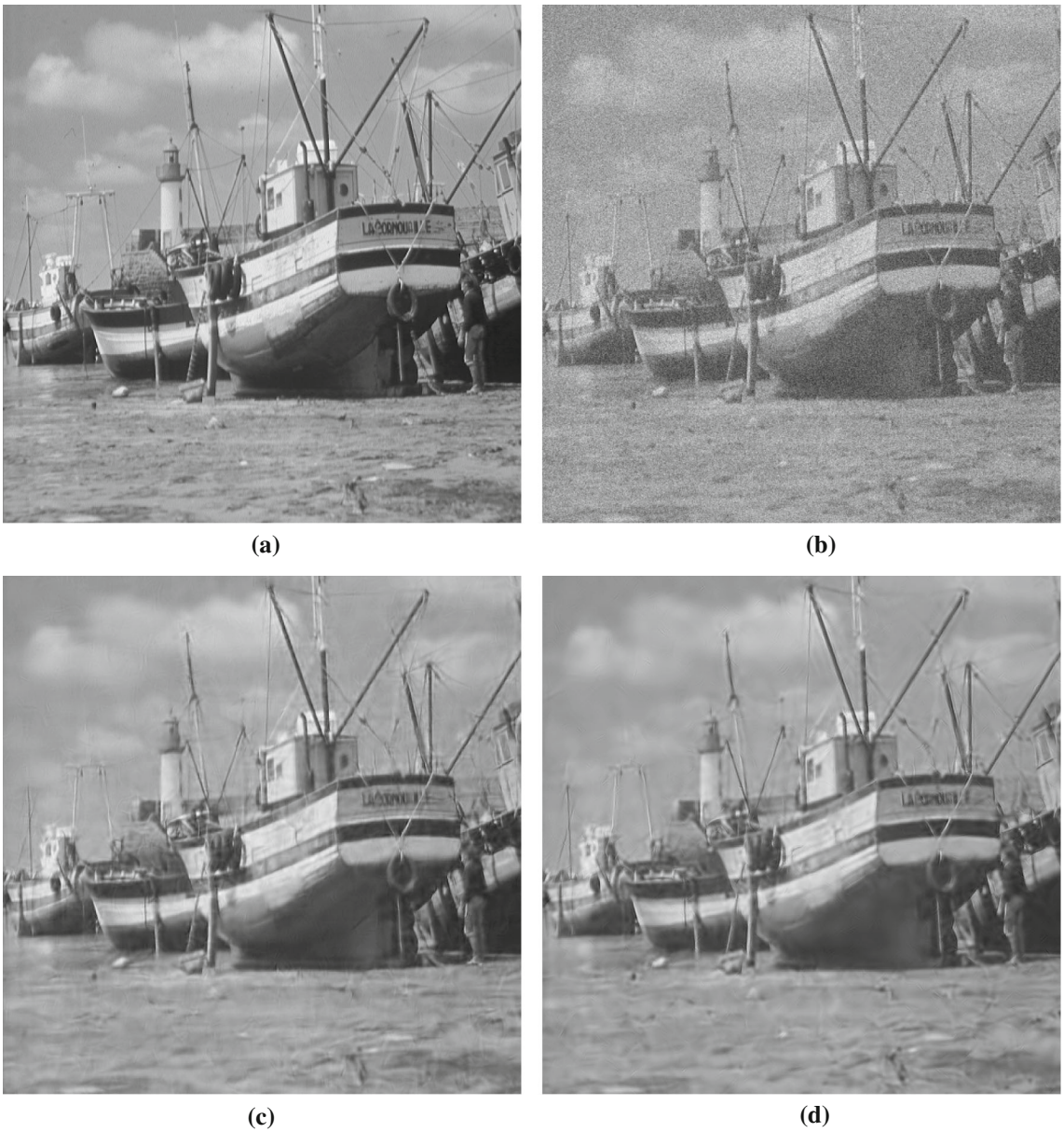
The difference in performance between BLS-GSM using the sinusoidal wavelet frame and shrinking the model and residual wavelets is demonstrated for the *Boats* image in Fig. 14. The sinusoidal model approach appears to have lower noise adjacent to linear features, such as the boat masts, compared to BLS-GSM with the wavelet frame. This may be because the wavelets are steered to the best orientation for the feature when using the sinusoidal model, whereas the feature energy can be split between two wavelets with a fixed-angle frame. Since ordinal filtering improves the results, a more principled method of including neighbouring pixel information may deliver further improvements. Another research direction would be to include extra sinusoids in the model to see if further splitting of the CH vector can improve performance.

8 Conclusion

The CH vector describes the local image structure in terms of its rotational symmetry. We have presented a method to solve a general signal model by splitting the CH vector into model components that represent structures of interest, and a residual component. The method is useful for image analysis, as it allows for the part of the local image structure that is *not* well-represented by a model to also be investigated. In the context of 2D steerable wavelets, it allows for exact reconstruction of the image when the model wavelets do not form a tight frame.

The CH vector method was used to derive the parameters of a sinusoidal image model. The sinusoid part of the model describes the strength (amplitude), symmetry (phase) and orientation of i1D features. Increasing the number of RT orders improves the estimation of these parameters in the presence of noise. The residual part of the model describes i2D features consisting of multiple symmetries. Both the model and residual components were used to construct a complex-valued representation of intrinsic dimension, which performs well for detecting corners and junctions.

Many applications that use the monogenic signal for phase and orientation estimation should be improved by using the sinusoidal model derived from higher-order RTs. As a minimum, adding the 2nd-order RT gives an estimation of orientation from both odd and even structures, is quick to



(a) **(b)** **(c)** **(d)**

Fig. 14 Denoised images of the *Boats* image (**a**) with additive white Gaussian noise (**b**) using BLS-GSM with a equiangular sinusoidal wavelet frame (**c**) and soft-thresholding of the sinusoidal model and residual components (**d**). Five scales of the VOW wavelet frame were used with $N = 7$ and phase-invariant equal weighting

compute, as the model parameters can be found analytically using quartic solvers, and also gives an intrinsic dimension representation.

The CH vector can also be used to solve for other image models. If the model components are not linearly independent or have different orientations, an iterative process can be used, whereby one component is initially solved using the CH vector, and subsequent components are solved using the residual vector. This is the subject of future investigation into an image model consisting of multiple sinusoids.

Acknowledgements This research was funded by scholarships from the Commonwealth Scientific and Industrial Research Organisation (CSIRO) and James Cook University (JCU).

Appendix A: Quick Trigonometric Polynomial Maximum Estimation

A quick approximation method for locating the maximum of a trigonometric polynomial is described below.

Consider a trigonometric polynomial of degree $2N$ given by

$$p(\theta) = \sum_{|n| \in \mathbb{N}_N} c_n e^{\mathrm{i}n\theta}. \tag{125}$$

The argument of n-th coefficient, c_n, gives an estimate of θ in the range $[0, 2\pi/n)$, and therefore n possible estimates for θ over the entire range $[0, 2\pi)$, while the magnitude of c_n is the strength of each estimate. We can think of these as representing n equiangular complex vectors for each positive order, n, as

$$b_{n_m} = \delta(n)\,|c_n| \exp\left(\mathrm{i}\frac{2\pi m + \arg(c_n)}{n}\right), \tag{126}$$

where $m \in \mathbb{N}_{n-1}$ is the index of the estimate, and $\delta(n)$ is a weighting function reflecting that higher orders are more sensitive to orientation changes and are therefore better estimates.

Next we choose one vector from each order and sum them to give a combined vector,

$$v_{\mathbf{m}} = \sum_{n=1}^{N} b_{n_{m_n}}, \tag{127}$$

where $\mathbf{m} \in \mathbf{M}^{(N)}$ and

$$\mathbf{M}^{(N)} = \left\{ M_n^{(N)} \in \mathbb{N}_{n-1} \mid n = [1, ..., N] \right\} \tag{128}$$

is the set of all possible indices of the estimates up to order N. There are thus $N!$ combinations we can create. Out of these $N!$ we choose the $v_{\mathbf{m}_N}$ with the greatest magnitude. Its argument is the final estimate for θ:

$$\theta = \arg v_{\mathbf{m}}, \quad \text{where } \mathbf{m} = \max_{\mathbf{m} \in \mathbf{M}_N} |v_{\mathbf{m}}|. \tag{129}$$

However, $N!$ possible combinations to search through becomes very large for large N, and defeats the purpose of a quick algorithm.

Instead, we propose the following hybrid scheme. Firstly, all combinations are calculated for each of the first q orders, giving $q!$ vectors, from which we choose the q vectors with the largest magnitude. For each of these q vectors, the vectors for the next order are added and the combined vector with the largest magnitude is kept. The number of vectors remains at q. This is repeated for the remaining orders, giving q final vectors, from which θ is chosen as the argument of the vector with largest magnitude. Using $\delta(n) = n^2$ for the order weighting function was found to give good results.

The method was evaluated for solving the sinusoidal image model. A histogram of the orientation error of the approximation method compared to root finding is shown in Fig. 15. Orientation was calculated for the first scale of the *Pentagon* image using $N = 7$, phase-invariant equal weighting, and different values of q. The errors were split into two classes depending on whether the local structure had a larger sinusoidal model component ($\|\mathbf{W}\mathbf{f}_S\| > \|\mathbf{W}\boldsymbol{\epsilon}\|$) or larger residual component. For the sinusoidal-like structures, over half the errors were below 0.01 degrees, and increasing q gave only a marginal improvement in error. For non-sinusoidal structures the errors were higher, but still mostly less than 3 degrees. Increasing q reduced the number of large errors (above 30 degrees). These errors are due to structures that can be modelled as two or more sinusoids of similar magnitude, as this gives multiple similar value local maxima in the orientation polynomial, $p(2\theta)$. If one is only interested in analysing sinusoidal structures, then choosing $q = 1$ gives fast and accurate results.

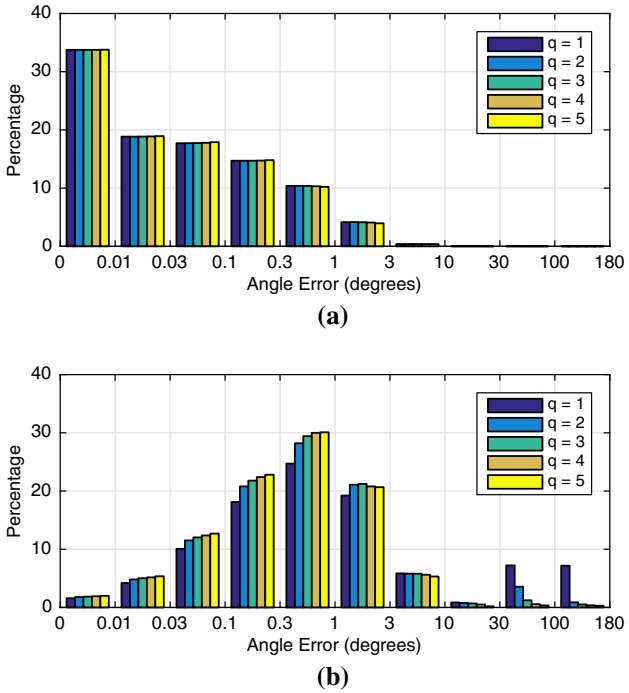


Fig. 15 Histogram of orientation errors using the quick approximation method for the *Pentagon* image with $N = 7$ and phase-invariant equal weighting. **a** Error in θ for predominantly sinusoidal structures. $\|\mathbf{W}\mathbf{f}_S\| > \|\mathbf{W}\boldsymbol{\epsilon}\|$. **b** Error in θ for predominantly non-sinusoidal structures $\|\mathbf{W}\mathbf{f}_S\| < \|\mathbf{W}\boldsymbol{\epsilon}\|$

Computation time for the approximation method versus MATLAB roots was calculated for the first scale of the *Pentagon* image, for q from 1 to 5, N ranging from 2 to 13, equally weighted CH vector, and using a single core of a 2.5Ghz Intel Core i7 processor. The results are shown in earlier in Table 2 along with the percentage of errors under 3 degrees. The setup time required to perform the correlation with the sinusoidal model vectors and square the polyno-

mials $\lambda_e(\theta)$ and $\lambda_o(\theta)$ is also shown. The approximation method is over 100 times faster for $q = 1$; however, jumping from $q = 4$ to $q = 5$ results in a significant slow down. The time complexity for the initial q estimates and the subsequent search is $\mathcal{O}(q! + q\sum_{q+1}^{N} k)$ which makes $q > 5$ longer than the MATLAB roots method and therefore impractical. Choosing $q = 3$ is a good trade off between speed and accuracy.

References

1. Ali, R., Gooding, M., Christlieb, M., Brady, M.: Advanced phase-based segmentation of multiple cells from brightfield microscopy images. In: IEEE International Symposium on Biomed. Imaging, pp. 181–184. IEEE (2008)
2. Awrangjeb, M., Lu, G., Fraser, C.S.: Performance comparisons of contour-based corner detectors. IEEE Trans. Image Process. **21**(9), 4167–4179 (2012)
3. Bigun, J., Granlund, G.: Optimal orientation detection of linear symmetry. In: IEEE International Conference on Computer Vision, vol. 54, p. 258 (1987)
4. Boukerroui, D., Noble, J., Brady, M.: On the choice of band-pass quadrature filters. J. Math. Imaging Vis. **21**(1), 53–80 (2004)
5. Bülow, T., Sommer, G.: Hypercomplex signals: a novel extension of the analytic signal to the multidimensional case. IEEE Trans. Signal Process. **49**(11), 2844–2852 (2001)
6. Daubechies, I., Grossmann, A., Meyer, Y.: Painless nonorthogonal expansions. J. Math. Phys. **27**(5), 1271 (1986)
7. Felsberg, M.: Low-level image processing with the structure multivector. Ph.D. thesis (2002)
8. Felsberg, M.: Optical flow estimation from monogenic phase. In: Lecture Notes in Computer Science, pp. 1–13. Springer, Berlin (2007)
9. Felsberg, M., Kalkan, S., Krüger, N.: Continuous dimensionality characterization of image structures. Image Vis. Comput. **27**(6), 628–636 (2009)
10. Felsberg, M., Sommer, G.: The monogenic signal. IEEE Trans. Signal Process. **49**(12), 3136–3144 (2001)
11. Felsberg, M., Sommer, G.: Image features based on a new approach to 2D rotation invariant quadrature filters. In: Lecture Notes in Computer Science, vol. 2350, pp. 369–383. Springer, Heidelberg (2002)
12. Field, D.: Relations between the statistics of natural images and the response properties of cortical cells. J. Opt. Soc. Am. A **4**(12), 2379–2394 (1987)
13. Forstner, W., Gulch, E.: A fast operator for detection and precise location of distinct points, corners and centres of circular features. In: ISPRS Intercommission Workshop (1987)
14. Freeman, W., Adelson, E.: The design and use of steerable filters. IEEE Trans. Pattern Anal. **13**(9), 891–906 (1991)
15. Hahn, S.: Multidimensional complex signals with single-orthant spectra. Proc. IEEE **80**(8), 1287–1300 (1992)
16. Harris, C., Stephens, M.: A combined corner and edge detector. In: Proceedings of the fourth Alvey Vision conference, vol. 15, pp. 147–151. Manchester (1988)
17. Held, S., Storath, M., Massopust, P., Forster, B.: Steerable wavelet frames based on the Riesz transform. IEEE Trans. Image Process. **19**(3), 653–667 (2010)
18. Jacovitti, G., Neri, A.: Multiresolution circular harmonic decomposition. IEEE Trans. Signal Process. **48**(11), 3242–3247 (2000)
19. Köthe, U.: Low-level feature detection using the boundary tensor. In: Weickert, J., Hagen, H. (eds.) Visualization and Processing of Tensor Fields, pp. 63–79. Springer, Berlin (2006)
20. Kovesi, P.: Good colour maps: How to design them. CoRR arXiv:1509.03700 (2015)
21. Krieger, G., Zetzsche, C.: Nonlinear image operators for the evaluation of local intrinsic dimensionality. IEEE Trans. Image Process. **5**(6), 1026–1042 (1996)
22. Larkin, K.G., Bone, D.J., Oldfield, M.A.: Natural demodulation of two-dimensional fringe patterns. I. General background of the spiral phase quadrature transform. J. Opt. Soc. Am. A **18**(8), 1862–1870 (2001)
23. Lindeberg, T.: Feature detection with automatic scale selection. Int. J. Comput. Vis. **30**(2), 79–116 (1998)
24. Marchant, R., Jackway, P.: Generalised Hilbert transforms for the estimation of growth direction in coral cores. In: International Conference on Digital Image Computing: Techniques and Applications, pp. 660–665 (2011)
25. Marchant, R., Jackway, P.: Feature detection from the maximal response to a spherical quadrature filter set. In: International Conference on Digital Image Computing: Techniques and Applications. Perth (2012)
26. Marchant, R., Jackway, P.: Local feature analysis using a sinusoidal signal model derived from higher-order Riesz transforms. In: IEEE Image Processing, pp. 3489–3493 (2013)
27. Marchant, R., Jackway, P.: Modelling line and edge features using higher-order Riesz transforms. Lecture Notes in Computer Science **8192**, 438–449 (2013)
28. Marchant, R., Jackway, P.: Using super-resolution methods to solve a novel multi-sinusoidal signal model. In: International Conference on Digital Image Computing: Techniques and Applications, pp. 1–8. IEEE (2013)
29. Mellor, M., Brady, M.: Phase mutual information as a similarity measure for registration. Med. Image Anal. **9**(4 SPEC. ISS.), 330–343 (2005)
30. Mikolajczyk, K., Schmid, C.: Scale & affine invariant interest point detectors. Int. J. Comput. Vis. **60**(1), 63–86 (2004)
31. Mikolajczyk, K., Schmid, C.: Performance evaluation of local descriptors. IEEE Trans. Pattern Anal. **27**(10), 1615–1630 (2005)
32. Morrone, M., Burr, D.: Feature detection in human vision: a phase-dependent energy model. In: Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society (Great Britain) **235**(1280), 221–245 (1988)
33. Morrone, M., Owens, R.: Feature detection from local energy. Pattern Recogn. Lett. **6**(5), 303–313 (1987)
34. Mühlich, M., Friedrich, D., Aach, T.: Design and implementation of multisteerable matched filters. IEEE Trans. Pattern Anal. **34**(2), 279–291 (2012)
35. Neri, A.: Circular harmonic functions: a unifying mathematical framework for image restoration, enhancement, indexing, retrieval and recognition. In: Proceedings of the EUVIP, pp. 269–276. IEEE (2010)
36. Olhede, S., Metikas, G.: The monogenic wavelet transform. IEEE Trans. Signal Process. **57**(9), 3426–3441 (2009)
37. Pad, P., Uhlmann, V., Unser, M.: VOW: variance-optimal wavelets for the steerable pyramid. In: IEEE Image Processing, pp. 2973–2977 (2014)
38. Perona, P., Malik, J.: Detecting and localizing edges composed of steps, peaks and roofs. IEEE International Conference Computer Vision (1990)
39. Perona, P., Padova, U.: Deformable kernels for early vision. IEEE Trans. Pattern Anal. **17**(5), 488–499 (1995)
40. Portilla, J., Simoncelli, E.: A parametric texture model based on joint statistics of complex wavelet coefficients. Int. J. Comput. Vis. **40**(1), 49–71 (2000)
41. Portilla, J., Strela, V., Wainwright, M., Simoncelli, E.: Image denoising using scale mixtures of Gaussians in the wavelet domain. IEEE Trans. Image Process. **12**(11), 1338–1351 (2003)

42. Puspoki, Z., Uhlmann, V., Vonesch, C., Unser, M.: Design of Steerable Wavelets to Detect Multifold Junctions. IEEE Trans. Image Process. **25**(2), 643–657 (2016)
43. Puspoki, Z., Vonesch, C., Unser, M.: Detection Of symmetric junctions in biological images using 2-D steerable wavelet transforms. In: International Symposium on Biomedical Imaging, pp. 1488–1491 (2013)
44. Ronse, C.: On idempotence and related requirements in edge detection. IEEE Trans. Pattern Anal. **15**(5), 484–491 (1993). doi:10.1109/34.211468
45. Rosten, E., Porter, R., Drummond, T.: Faster and better: a machine learning approach to corner detection. IEEE Trans. Pattern Anal. **32**(1), 105–119 (2010)
46. Simoncelli, E.P.: Shiftable multiscale transforms. IEEE Trans. Inf. Theory **38**(2), 587–607 (1992)
47. Sommer, G., Wietzke, L., Zang, D.: Monogenic curvature tensor as image model. Vis. Process. Tensor Fields **2**(357), 281–301 (2009)
48. Soulard, R., Carre, P., Fernandez-Maloigne, C.: Vector Extension of Monogenic Wavelets for Geometric Representation of Color Images. IEEE Trans. Image Process. **22**(3), 1070–1083 (2013)
49. Stark, H.: An extension of the Hilbert transform product theorem. Proc. IEEE **59**(9), 1359–1360 (1971)
50. Storath, M.: Directional multiscale amplitude and phase decomposition by the monogenic curvlet transform. SIAM J. Imaging Sci. **4**(1), 57–78 (2011)
51. Unser, M., Chenouard, N.: A unifying parametric framework for 2D steerable wavelet transforms. SIAM J. Imaging Sci. **6**(1), 102–135 (2012)
52. Unser, M., Chenouard, N., Van de Ville, D.: Steerable pyramids and tight wavelet frames in L2(R(d)). IEEE Trans. Image Proc. **20**(10), 2705–2721 (2011)
53. Unser, M., Sage, D., Van De Ville, D.: Multiresolution monogenic signal analysis using the Riesz-Laplace wavelet transform. IEEE Trans. Image Process. **18**(11), 2402–2418 (2009)
54. Unser, M., Van De Ville, D.: Wavelet steerability and the higher-order Riesz transform. IEEE Trans. Image Process. **19**(3), 636–652 (2010)
55. Venkatesh, S., Owens, R.: On the classification of image features. Pattern Recogn. Lett. **11**(5), 339–349 (1990)
56. Ward, J., Chaudhury, K., Unser, M.: Decay properties of Riesz transforms and steerable wavelets. SIAM J. Imaging Sci. **6**(2), 984–998 (2013)
57. Wietzke, L., Sommer, G.: The signal multi-vector. J. Math. Imaging Vis. **37**(2), 132–150 (2010)
58. Wietzke, L., Sommer, G., Fleischmann, O.: The geometry of 2D image signals. In: Proceedings of the CVPR IEEE, pp. 1690–1697. IEEE (2009)
59. Zang, D., Sommer, G.: Detecting intrinsically two-dimensional image structures using local phase. Pattern Recogn. **357**(357), 222–231 (2006)
60. Zang, D., Sommer, G.: The monogenic curvature scale-space. In: Klette, R., Zunic, J. (eds.) Combinatorial Image Analysis, pp. 320–332. Springer, Berlin (2006)
61. Zetzsce, C., Barth, E.: Fundamental limits of linear filters in the visual processing of two dimensional signals. Vis. Res. **30**(7), 1111–1117 (1990)

Ross Marchant received a degree in electrical engineering from the University of Queensland, Australia, in 2002, and is currently working toward the Ph.D. degree at James Cook University, Australia. He also works as a Research Officer in the Marine Geophysics Laboratory, James Cook University, designing new instrumentation for oceanographic and other research. His research interests include phase-based image processing, low-level feature analysis, and embedded electronics. Mr. Marchant has received Ph.D. scholarships from James Cook University and the Commonwealth Scientific and Industrial Research Organisation (CSIRO).

Paul Jackway studied Electronics Engineering followed by a Masters in Applied Statistics and Mathematical Modelling at Royal Melbourne Institute of Technology (RMIT) in the 1980's. After various engineering jobs in Telstra he moved to Brisbane in 1991 to undertake graduate studies at Queensland University of Technology (QUT) and was awarded his PhD in 1995 for a thesis on Mathematical Morphology and Scale space. Paul was then appointed to a Research Fellowship in the Cooperative Research Centre for Sensor Signal and Information Processing (CSSIP) based at the University of Queensland, in Brisbane, where he led a research group working on automation of Pap smear cytology for cervical cancer screening. In 2002, Paul took up a position as a Principal Research Scientist in the Commonwealth Scientific and Industrial Research Organisation (CSIRO) in Brisbane. At CSIRO he contributed to and led a range of applied imaging, image analysis, and pattern recognition projects, while leading the groups strategic push into automated microscopy. Projects he worked on included automated species discovery in insects and the aerial quantification of wheat and other field crops for scientific research. In late 2013 Paul left CSIRO and now continues his research at a more leisurely pace in a private capacity.