Abstract—In this paper we describe a novel method of feature detection and classification using the maximal response of a set of spherical quadrature filters to either a line-segment or wedge-segment signal type. This is achieved via a rotation and illumination invariant distance function. The development of the method is described, and some experimental results are provided to demonstrate the usefulness of the technique.

I. INTRODUCTION

A first step in many image analysis routines is the detection of low-level image features, such as lines, edges, corners and junctions, which consequently guides higher level processes such as object recognition and image registration.

One approach to detection is to apply a set of local linear operators to a signal, and from the responses calculate a scalar value of how similar the local signal is to a structure of interest. For example, Canny [1] used the magnitude of the response to a Gaussian derivative filter as a measure of edge strength, while derivatives can also be used to detect corners [2]. Extending from this, a parametrised signal model may be estimated from the operator responses. Deriving signal-model parameters along with detection can give rich information such as feature orientation and structure type (e.g. [3]).

Of particular interest are operators that estimate signal models in a phase and rotationally invariant way. The 1-D analytic signal, given by

$$f_A(x) = f(x) - i(h*f)(x),$$

where $h(x) = \frac{1}{\pi x}$ is the Hilbert transform kernel, can be expressed as the complex sinusoid $A(x)e^{i\phi(x)}$. Amplitude, $A(x)$, corresponds to feature strength [4], and phase, $\phi(x)$, corresponds to feature type (e.g. rising, falling, peak, trough). This split of identity is a key property of phase-based methods [5].

Much effort has been put into extending this concept to images, e.g. [6], [7], [8], however the monogenic signal [9] was the first rotationally-invariant extension. It expresses local structure in terms of the amplitude, phase and orientation of a 2D sinusoid, with the locations of lines and edges indicated by the phase value. Subsequent developments include the monogenic curvature signal [3] and signal multi-vector [10], which can describe multiple structure symmetries often present in images, and therefore detect locations of interest, such as junctions and corners. These higher-dimensional signals are all obtained via the generalised Hilbert transform (GHT) [10].

Jacob & Unser [11] take a different approach to detection. Rather than using a single signal model that encompasses a range of feature types, the maximal response of a steerable Gaussian filter set [12] is analytically determined for a particular set of types, specifically edge, line and wedge. For detection, a measure of similarity to a particular type is given by evaluating the linear combination of filter responses in the proportion corresponding to the maximal response to that type. To obtain rotational invariance, these are steered according to an estimated orientation. Higher-order derivatives are used to obtain more directionally selective detection, as well as the estimation of a wedge angle parameter.

In this paper we apply elements of the feature detector design method described in [11], but instead of a steerable Gaussian filter set, we propose using basis filters constructed from the higher-order generalised Hilbert transform [10]. Two classes of signal types are proposed, line-segment and wedge-segment, of which line and edge are particular cases. The filter set obtained via the GHT is steerable, and rotationally-invariant [13]. We analytically determine the maximal response to either the line- or wedge-segment types. Furthermore, a novel rotation estimation method is described, which is maximal when the local signal structure matches that of the signal type being detected.

The structure of this paper is as follows. In the next section we give a brief background on spherical quadrature filters before proposing a novel feature vector in section III. Details of our proposed rotation estimation procedure is given in section IV. A definition of line and wedge signal types follows in section V, followed by the derivation of maximal spherical quadrature filter responses in section VI, and detection in section VII. Section VIII contains details and results of experiments on: rotation parameter estimation, feature vector size, image feature detection, and finally the classification of features on a bee-wing vein image.
II. BACKGROUND: SPHERICAL QUADRATURE FILTERS

The 2-D \(n\)-th order generalised Hilbert transform (GHT) [10], also known as the higher-order Reisz transform [13], is a bounded linear operator. It can be expressed as a Fourier multiplier, \(\tilde{h}^{(n)}(u)\), or a convolution kernel in the spatial domain, \(h^{(n)}(z)\),

\[
\tilde{q}^{(n)}(u) = \hat{p}(u) \cdot \left( \frac{u}{\|u\|} \right)^n \tag{1}
\]

\[
q^{(n)}(x, y) = p(x, y) \ast \left( \frac{n}{2\pi} \frac{-(x+iy)^n}{\|x+iy\|^{n+2}} \right) \tag{2}
\]

where \(x, y \in \mathbb{R}\), \(u \in \mathbb{R}^2\) and \(n \in \mathbb{N}^*\).

Note, in subsequent equations we shall use \(z = [x, y]\). The zero-th order GHT is the identity operator. A bandpass filter, such as a log-normal filter [14], can be applied to the image to restrict analysis to a particular scale. The convolution of an isotropic filter and GHT kernel is known as a spherical quadrature filter (SQF) [9], or is alternatively expressed as a 2D monomial filter set [15].

Labelling the real and imaginary parts of the \(n\)-th order SQF, \(q^{(n)}(z)\), convolved with an image signal, \(f \in L^2(\mathbb{R}^2)\), as,

\[
f_a^{(n)}(z) = \mathbb{R} \{ q^{(n)} \ast f \}(z) \tag{3}
\]

\[
f_b^{(n)}(z) = \mathbb{I} \{ q^{(n)} \ast f \}(z), \tag{4}
\]

the response may be expressed in the form of the complex exponential, \(A^{(n)}(z) e^{i\theta^{(n)}(z)}\), where the phase-invariant amplitude is given by

\[
A^{(n)}(z) = \left| f_a^{(n)}(z) + if_b^{(n)}(z) \right|, \tag{5}
\]

and orientation by,

\[
\theta^{(n)}(z) = \text{Arg} \left( f_a^{(n)}(z) + if_b^{(n)}(z) \right). \tag{6}
\]

The monogenic signal is constructed from the 0 to 1st order SQF responses, while the monogenic curvature signal and signal multi-vector use 0 to 3rd order SQFs.

An example of 1st to 4th order log-normal SQFs is shown in Figure 1. It can be inferred that odd order filters would respond to odd order signals, and even order filters to even signals. Higher-order SQFs encode more axes of symmetry, and would also respond most strongly at locations in signals with the same order rotational symmetry. This gives a hint as to the maximum order SQF required to construct an adequately discriminatory feature detector. For example, an ‘X’ junction detector would need the 4th order SQF. The higher-order SQF kernels are also larger spatially, as the frequency response of each is the same. This is an important property not necessarily present in filters constructed using spherical harmonics or polynomials applied via the spatial domain.

III. FEATURE VECTOR

In this section, we propose to use a filter set of different order spherical quadrature filters to derive the maximal responses to our proposed signal types.

Consider a filter set made up of the 0 to \(n\)-th order SQFs. If the SQFs are formed using an isotropic filter, they will be steerable and orthogonal [13]. This motivates embedding the response of an image, \(f(z)\), to the filter set into the \(n\)-dimensional, complex valued, euclidean space. That is, each point in the image will be represented by a vector \(f^n \in \mathbb{R} \times \mathbb{C}^n\), where \(f_k^n = f_a^{(k)} + if_b^{(k)}\) (from (3) and (4)), and \(n\) representing the maximum order SQF used. In, for example, [5], [16], [17], geometric algebra is used for this embedding, however we do not perform this here. Thus, each point in the image is represented as a complex valued feature vector, \(\hat{f}^n(z)\). A split of identity is possible by expressing this feature vector in terms of scalar amplitude and phase-vector components, e.g.

\[
A(z) = \|\hat{f}^n(z)\|, \tag{7}
\]

\[
\hat{f}^n(z) = f^n(z)/A(z). \tag{8}
\]

For detection, a comparison measure between two feature vectors is necessary; one vector will be describing the local signal structure, and one will be describing the signal type being detected. Let \(\mathbf{a}\) and \(\mathbf{b}\) be two such feature vectors. Euclidean distance shall be used to give a scalar measure of dissimilarity,

\[
d(\mathbf{a}, \mathbf{b}) = \|\mathbf{a} - \mathbf{b}\|_2. \tag{9}
\]

Taking advantage of the split of identity, we can compare two feature vectors without regard to their strength (see (6)), by finding the distance between two unit vector phase components,

\[
d_{\|\cdot\|}(\mathbf{a}, \mathbf{b}) = \|\hat{\mathbf{a}} - \hat{\mathbf{b}}\|_2. \tag{10}
\]

However, this does not take into account rotation of the type. A rotationally-invariant measure is given by

\[
d_{\theta}(\mathbf{a}, \mathbf{b}) = \min_{\theta \in (-\pi, \pi]} \|\mathbf{r}(\theta) \cdot \hat{\mathbf{a}} - \hat{\mathbf{b}}\|_2. \tag{11}
\]

where \(\mathbf{r}(\theta)\) is the complex valued rotation vector \(\{r_k(\theta) = e^{ik\theta} | k \in \mathbb{N}^n\}\). Performing an iterative search over the range of rotations would be time consuming. Accordingly, an estimation of the rotation angle is needed.
IV. Rotation Estimation

Previous methods of estimating this orientation involve numerically searching over multiple orientations, solving a polynomial equation analytically, for \( n \leq 3 \) [11], or using the orientation of a single SQF component. We propose to estimate the rotation angle, \( \theta \), from the orientation of all the SQF components. Equation (7) gives,

\[
\theta = \arg \max_{\theta \in (-\pi, \pi)} \sum_{k=0}^{n} |\hat{a}_k| |\hat{b}_k| \cos(\alpha_k - \beta_k + k\theta)
\]

where \( \alpha = \text{Arg}(\hat{a}_k) \) and \( \beta = \text{Arg}(\hat{b}_k) \). This has a simple solution only in a small number of cases, where there are few non-zero components.

We note that the minima for the \( k \)-th term occur when \( k\theta \equiv \beta_k - \alpha_k \mod 2\pi \), and since we must estimate \( \theta \) over the interval \([0, 2\pi)\), there exist \( k \) possible solutions for \( \theta \). Accordingly, these solutions are represented by,

\[
R_{k,j} = \exp(i\frac{2\pi j + \beta_k - \alpha_k}{k}),
\]

where \( j \in \mathbb{N}_k \).

We propose to estimate \( \theta \) using the vector sum of possible solutions, one chosen from each order, weighted by the order value. The permutation with the greatest amplitude is used to estimate the correct orientation. i.e.

\[
\{m_1, \ldots, m_n\} = \arg \max_{m_k \in \mathbb{N}_k} \left| \sum_{k=1}^{n} k|\hat{a}_k| |\hat{b}_k| R_{k,m_k} \right|
\]

\[
\theta = \text{Arg} \left( \sum_{k=1}^{n} k|\hat{a}_k| |\hat{b}_k| R_{k,m_k} \right)
\]

Using this method an iterative search or polynomial solving need not be applied. Furthermore, if the vectors being compared are simply rotated versions of one another, the estimated angle will be correct.

V. Signal Types

No one signal model can encompass all possible image structures that one may want to investigate. So instead of performing detection by calculating signal parameters, we shall follow a similar approach to [11] and use the maximal responses of a filter set. Our novel approach is to use a set of different order SQFs, with responses embedded into a feature vector, as described in the previous section. An expanded set of signal types will be described, along with their maximal response to the SQF filter set.

A. Line-Segment Signal Types

We define our line-segment signal type to be one or more occluding lines of width, \( w \), radiating from a point, as shown in Figure 2. In polar co-ordinates this could be represented for \( m \) line segments as,

\[
t_L(r, \theta) = \sum_{k=1}^{m} l_k(r, \theta)
\]

where

\[
l_k(r, \theta) = \begin{cases} 1, & |\theta - \psi_k| < \sin^{-1} \frac{w}{r} \\ 0, & \text{otherwise} \end{cases}
\]

If we set \( \int l_k(r, \theta) d\theta = 1 \) then as \( w \to 0 \) we obtain,

\[
t_L(r, \theta) = \sum_{k=1}^{m} \delta(\theta - \psi_k)
\]

For example, an isometric ‘Y’ signal type would correspond to \( \psi = [0, \frac{\pi}{3}, \frac{2\pi}{3}] \).

B. Wedge-Segment Signal Types

The wedge-segment signal type replaces line segments with wedges of different angular widths, and we define it as,

\[
t_E(r, \theta) = \sum_{k=1}^{m} e_k(r, \theta)
\]

where

\[
e_k(r, \theta) = \begin{cases} 1, & \theta \in [\psi_{2k-1}, \psi_{2k}] \\ 0, & \text{otherwise} \end{cases}
\]

For example, a solid corner corresponds to \( \psi = [0, \frac{\pi}{2}] \).

VI. Maximal SQF Response

We now derive the response of each signal type to a particular order SQF, \( q^{(n)}(z) \), where \( P(u) \) is the spectral response of the band-pass filter from which it was derived. Expanding on the derivation in [10] we have

\[
q^{(n)}(z) = \int_{\mathbb{R}^2 \setminus B_r(0)} \left( \frac{u}{||u||} \right)^n P(u) e^{i2\pi(u,z)} du
\]

Substituting \( u = [k \cos \phi, k \sin \phi] \) and \( z = [r \cos \theta, k \sin \theta] \) this simplifies to,

\[
q^{(n)}(r, \theta) = \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} P(k) \left(ie^{i\theta}\right)^n e^{-ikr \cos(\phi - \theta)} d\phi dk
\]

\[
= \frac{i^n}{2\pi} e^{in\theta} \int_0^\infty \int_0^{2\pi} P(k) e^{i\phi} e^{-ikr \cos \phi} d\phi dk
\]

Since the Bessel function is defined as

\[
J_n(x) = \frac{1}{2\pi} \int_{-\pi}^\pi e^{-i(n\tau - x \sin \tau)} d\tau
\]

\[
= \frac{1}{2\pi} \int_0^{2\pi} i^n e^{in\tau} e^{-ix \cos \tau} d\tau
\]

The original function then becomes,

\[
q^{(n)}(r, \theta) = \left(-e^{i\theta}\right)^n \int_0^\infty P(k) J_n(kr) k dk
\]
A. Line-Segment Response

For the line-segment type, the response for a single line-segment to an \( n \)-th order SQF at the local origin is,

\[
t^{(n)}_k(0) = \left\langle I_k(z), q^{(n)}(-z) \right\rangle = \int_{r=0}^{\infty} \int_{\theta=0}^{2\pi} \frac{\delta(\theta - \psi_k)}{r} \left( (e^{i\theta})^n \int_{0}^{\infty} P(k) J_n(kr) \, dk \right) \, d\theta \, dr = e^{in\psi_k} \int_{0}^{\infty} P(k) \, dk
\]

Thus for an \( m \) line-segment type, the response to an \( n \)-th order SQF at the origin is,

\[
t^{(n)}_L(0) = C \sum_{k=1}^{m} e^{in\psi_k}
\]

where \( C \) is a constant depending on \( P(k) \). The feature vector corresponding to this type is thus \( t^n = [t^{(0)}_L, t^{(1)}_L, ..., t^{(n)}_L] \)

B. Wedge-Segment Response

For a single wedge segment, the response is,

\[
e^{(n)}_k(0) = \left\langle e_k(z), q^{(n)}(-z) \right\rangle = \frac{i}{n} (e^{in\psi_{2k-1}} - e^{in\psi_{2k}}), \int_{r=0}^{\infty} \int_{0}^{\infty} P(k) J_n(kr) \, kr \, dk \, dr
\]

The remaining integral only converges if the filter envelop, \( P(k) \), converges fast enough as \( k \to \infty \). We have not yet derived a solution, but from numerical experimentation have found the following relationship for typical band-pass filters such as the log-normal filter

\[
e^{(n)}_k \approx \frac{D}{n} (e^{in\psi_{2k-1}} - e^{in\psi_{2k}})
\]

where \( D \) is a constant dependant on \( P(k) \). Thus for an \( m \) wedge-segment type, the response to an \( n \)-th order SQF at the origin is,

\[
t^{(n)}_E(0) = \frac{D}{n} \sum_{k=1}^{m} e^{in\psi_{2k-1}} - e^{in\psi_{2k}}
\]

with corresponding feature vector \( t^n = [t^{(0)}_E, t^{(1)}_E, ..., t^{(n)}_E] \)

C. Inverse Types

The responses to inverse forms of each signal type can be obtained by multiplying each component by \( e^{in} \).

VII. Detection

We have so far described the construction of an SQF feature vector, a method to measure dissimilarity (distance) between feature vectors, and the maximal response to line-segment and wedge-segment signal types.

Detection of a particular signal type, \( t(z) \), is performed by measuring the distance between the corresponding feature vector, \( t^n \), and the feature vector, \( f^n(z) \), at each location in an image, as follows:

\[
d(z) = d_\theta(f^n(z), t^n)
\]

Detection using this distance measurement is therefore rotationally invariant, due to the use of SQFs, and illumination invariant, as the feature vectors are normalised. Therefore one need not consider the absolute values of the signal when thresholding.

Two other useful quantities arising from the split of identity are the amplitude of the image feature vector, \( A(z) \), which gives an indication of feature strength, and the estimated rotation parameter, which gives the orientation of the portion of the local signal structure matching the type being detected. Factors to be considered are the basis band-pass filter (see [14]) and maximum order SQF required, the effect of the latter shall be shown in the results.

A. Multiple Scales

Phase-congruency based methods [18] show that at locations of strong image features, the feature type is similar across scale. We propose a multi-scale detection method where the dissimilarity measure is added across scales,

\[
d(z; s) = \sum_{k} d_\theta(f^n(z; s_k), t^n) \quad (8)
\]

where \( f^n(z; s_k) \) is the feature vector calculated for some scale parameter \( s_k \in s \). Again, since each distance is calculated using a normalised feature vector, this measure is also invariant to illumination.

B. Classification

Given a set of different types, each point in an image can be classified according to which type gives the lowest distance score.

VIII. EXPERIMENTAL RESULTS

A. Rotation Parameter Estimation

The distance measure calculated using the rotation parameter estimation (7) was compared to that obtained using a search over all possible rotations. Twenty-thousand random pairs of feature vectors were constructed, using 0 to 2nd- \( (F^2) \), 0 to 3rd- \( (F^3) \), and 0 to 4th-order \( (F^4) \) SQFs. The amplitude of each vector component was selected from a zero-mean Gaussian distribution, and the orientation from a uniform distribution over \([−\pi, \pi]\). The distance between each pair was calculated using both the proposed method and the full search, and the distribution of the difference is shown in Figure 3.
Fig. 2: Different signal types and the combination of SQF filters corresponding to the maximal solution, \( f^n \), where \( n \) indicates the 0 to \( n \)-th order SQFs were used.

Fig. 3: Histogram of error between distance calculated using rotation parameter estimation method and a full search, for 20000 random feature vectors.

The difference was well approximated by a gamma distribution; mean difference was 0.01, maximum was 0.25, and 95% of values were less than 0.05. The number of SQF filters used does not appear to affect the distribution. These results show the proposed method performs well at estimating the minimum distance between feature vectors, regardless of orientation.

B. Feature Vector Size and Number of Scales

Figure 2 shows the combination of SQF filters corresponding to the maximal solution for some selected line-segment and edge-segment types. As the maximum order of SQF used increases, the kernel corresponding to the maximal solution increases in size and orientation selectivity. These results are similar to that of [11] for the line and edge types. For the crossed line type (four line segments spaced 90 degrees apart) it can be seen that the 4th-order SQF is needed to distinguish this structure from an isometric ‘blob’. This can be interpreted as a structure with large degree of rotational symmetry of a particular order requires the corresponding order SQF to be present in the filter set.

The effects on the distance measurement from increasing the number of scales and increasing the size of the feature vector are shown for a neuron image in Figure 4. The signal type used was a line, \( \psi = [0, \pi] \). As the maximum order of SQF used increases, the distance measurement becomes more selective to the line type, and more localised to the centre of the line. Furthermore, response to noise, which is often a problem in phase-based detection methods, is reduced. Increasing the number of scales also isolates the response to the line type further, and noise response in the local area of a feature is diminished.
Fig. 4: The effect of number of scales and feature vector order on the line-type distance measurement for a neuron image. Lighter colour indicates smaller distance. Image taken from [19].

Of particular note, some locations of neurite outgrowth that are imperceptible to the eye have been enhanced. For example, the outgrowth extending from the left side of the image (indicated by a black arrow) can be seen to connect to the main structure of the neuron in sub-figures e) through to j). This demonstrates one of the strengths of the approach, that detection is based on feature type, not amplitude.

C. Detection and Classification

Figure 5 shows the result of the proposed method for identifying the veins and junctions in a bee-wing image. Five signal types were used, based on the type of vein structures observed. These were blob, line, \( \psi = [0, \pi] \); ‘Y’, \( \psi = [0, \pi/3, 2\pi/3] \); ‘T’, \( \psi = [0, \pi/2, \pi] \); ‘X’, \( \psi = [0, \pi/2, \pi, 3\pi/2] \); and edge, \( \psi = [0, \pi] \). The 0 to 4-th order SQF responses \( f^4(z) \) were used, over 3 scales given by a log-normal filters with
wavelength 16, 32 and 64 pixels, and sigma 0.65.

Distance to each type was calculated according to (8), and each point in the image classified according to the type with minimum distance. The results show that the distance appears invariant to rotation and illumination. For example, the fainter veins are strongly identified as line type (red), allowing for more precise location of these features. Vein junctions are clearly identified according to their type, however there are many noisy locations. These potential junctions can be pruned by considering the amplitude response. Classifying every point in the image according to the nearest signal type, while treating the amplitude separately, shows promise in realising Marr’s idea of a primal sketch [20] in the one framework.

Running time for a 512 x 512 pixel image was approximately 15s (per scale) using MATLAB 2011 on a dual-core i7 2.2GHz processor.

IX. CONCLUSION

In this paper we have presented a method of feature detection using the maximal response to a particular signal type, following a similar procedure to [11]. Our novel approach is to use spherical quadrature filters as basis filters, and to define signal types from two classes, multiple line-segments and multiple wedge-segments. By using spherical quadrature filters, the choice of band-pass filter used to select scale is not limited, as any suitably compact functions will give a similar (normalised) maximal response. Defining signal types as segments radiating from a point of origin allows for more complexity. Furthermore, to obtain illumination and rotation invariance, instead of calculating the magnitude of the response to the maximal filter set, a rotation and illumination invariant distance measure was introduced. This method requires orientation to be estimated, and the estimation procedure described was shown to give a quite accurate values. A drawback of this method is that the signal type is fixed. One may wish to detect the type under a set of transformations, for example an ‘X’ type that has been stretched due to perspective distortion. This is the current focus of research.

Formal evaluations of the proposed method compared to alternatives, on specific problems or classes of problems, have yet to be completed. However, preliminary results show usefulness in classification problems such as character recognition and handwriting author recognition.

REFERENCES


