ROTATION-INVARIANT LOCAL FEATURE MATCHING WITH COMPLEX WAVELETS

Nick Kingsbury

Signal Processing Group, Dept. of Engineering, University of Cambridge, Cambridge, CB2 1PZ, U.K.
phone: + (44) 1223 338514 email: ngk@eng.cam.ac.uk web: www.eng.cam.ac.uk/~ngk

ABSTRACT

This paper describes a technique for using dual-tree complex wavelets to obtain rich feature descriptors of keypoints in images. The main aim has been to develop a method for retaining the full phase and amplitude information from the complex wavelet coefficients at each scale, while presenting the feature descriptors in a form that allows for arbitrary rotations between the candidate and reference image patches.

1. INTRODUCTION

An important problem in image analysis is that of finding similar objects in sets of images, where the objects are often at different locations, scales and orientations in the various images. Partial occlusion of objects is also quite common. An effective general approach to this problem is first to find a relatively large number (typically several thousand) of key feature points in each image, using for example the Harris corner and edge detector [1], and then to develop a more detailed descriptor for each keypoint, which allows points from different images to be compared and matched to create candidate pairings. Often a reference object is taken from one image and then other instances of the object are searched for in the remaining images, so the number of reference keypoints is quite small (10 - 100), but the number of candidate keypoints can be very large (10^5 - 10^7). Hence it is important to develop keypoint descriptors which allow efficient comparison of pairs of keypoints (reference-to-candidate), and this is the main topic of this paper.

One of the most popular recent algorithms for this application has been Lowe's Scale Invariant Feature Transform (SIFT) [2]. In SIFT, keypoints are located by detecting extrema in a 3-D scale-space, formed by differences on a Gaussian pyramid of multi-scale lowpass filters. Each keypoint is thus located spatially and in scale before its descriptor is calculated. The next step is to calculate dominant orientations based on local image gradient directions, and for each orientation a descriptor is generated based on orientation histograms, accumulated over 4×4 subregions around the keypoint. Here we propose an alternative descriptor which does not require the dominant orientation(s) to be computed first because it allows efficient matching of descriptor pairs in a rotationally invariant way.

Our descriptor is based on the outputs of the Dual-Tree Complex Wavelet Transform (DTCWT) of the input image. The DTCWT was developed by the author [5, 6] in order to provide a multi-scale decomposition of images (and other multi-dimensional data) which overcomes the problems of poor directional selectivity and strong shift dependence exhibited by the Discrete Wavelet Transform (DWT), at the expense of a moderate level of redundancy, 4:1 for 2-D data. Like the DWT, the DTCWT is a multi-scale transform with decimated subbands, but instead of three subbands per scale in 2-D, the DTCWT has six, and each coefficient is complex (i.e. it has a real and imaginary part). Figure 1(a) shows the real and imaginary parts of the 2-D impulse responses that define the six subbands at a given scale (level 4 in this case). We see that these responses are similar to those of a 6-directional Gabor transform with orientations of $\{15°, 45°, 75°, 105°, 135°, 165°\}$, as labelled. An alternative similar transform that could be used here is Simoncelli's Steerable Pyramid [3] which has the attractive property of approximate rotational symmetry, but here we concentrate on the DTCWT because of its lower redundancy and greater computational efficiency.

A key feature of the DTCWT is that it is approximately *shift invariant*, which means that the z-transfer function, through any given subband of a forward and inverse DTCWT in tandem, is invariant to spatial shifts, and that aliasing effects due to decimation within the transform are small enough to be neglected for most image processing purposes [5]. A corollary of this is that the complex wavelet coefficients within any given subband are sufficiently bandlimited that we can interpolate between them in order to calculate coefficients that correctly correspond to any desired sampling location or pattern of locations. Hence for a given keypoint location we may calculate the coefficients for an arbitrary sampling pattern centred on that location. To obtain circular symmetry consistent with our 6 subband orientations, we have chosen the 13-point sampling pattern of fig. 2.

The main innovative feature of this paper is the technique for assembling complex coefficients from the 13 sampling locations, 6 subband orientations, and one or more scales, such that they form a 'polar' matching matrix P, in which a rotation of the image about the centre of the sampling pattern corresponds to a cyclic shift of the columns of P. This is described in section 3. Before that, in section 2, we briefly discuss two modifications which improve the rotational symmetry of the 6 subbands at each scale of the DTCWT and hence improve the performance of the complete system.

The cyclic shift property of the matrix P, when rotation occurs, means that Fourier transform methods are appropriate for performing correlations between two matrices P_r and P_s from the reference and search images respectively. In section 4 we show that this correlation may be performed efficiently in the Fourier domain, followed by a single low-complexity inverse FFT to recover the correlation result as a function of rotation θ. The peak of this result is the required rotation-invariant similarity measure between P_r and P_s. A key aspect is that phase information from the complex coefficients can be fully preserved in this whole process. Prior work in this area by Hill [4] has used only the DTCWT co-

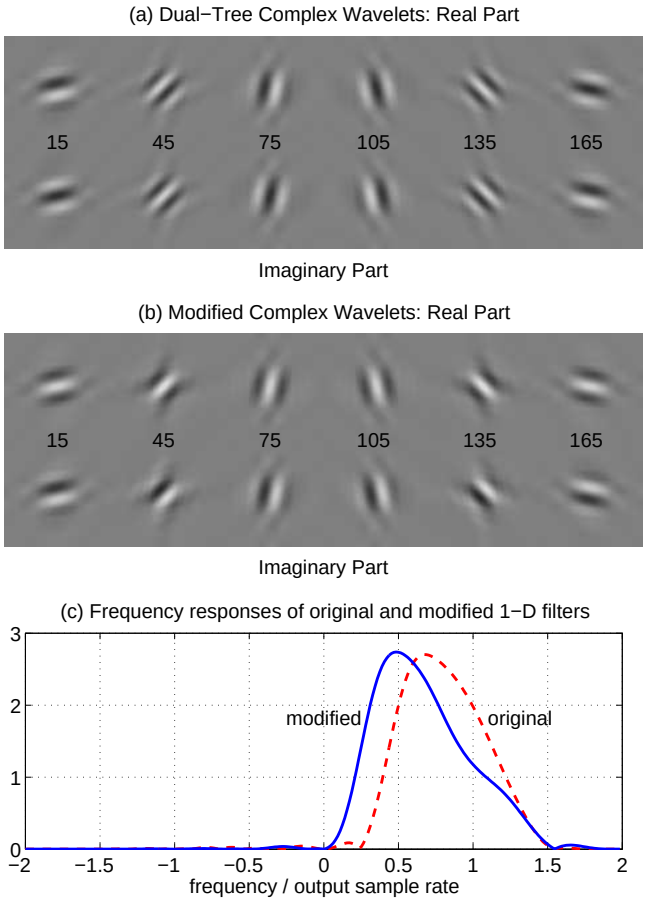

Figure 1: *(a) 2-D impulse responses of dual-tree complex wavelets at level 4; (b) 2-D impulse responses of the complex wavelets, modified to have lower centre frequencies (reduced by* $1/\sqrt{1.8}$*) in the* 45° *and* 135° *subbands and to have zero phase at the mid-point of each response; (c) frequency responses of the 1-D complex filters used to create the original and the modified 2-D responses, .*

efficient amplitudes.

In section 5 we discuss how the technique may be extended to include information from multiple scales and to be approximately scale-invariant, as well as rotation invariant, based on the dominant scale associated with each detected keypoint. Finally section 6 shows results and presents conclusions.

In parallel with this work, we have been developing keypoint detectors based on the DTCWT [7], which are alternatives to the isotropic Gaussian pyramid methods of SIFT and take advantage of the orientation selectivity of the DTCWT subbands. However feature detectors are beyond the scope of this paper and are not discussed further here.

2. IMPROVING THE ROTATIONAL SYMMETRY OF THE DT CWT

Since the aim of this work was to create a rotation-invariant similarity measure using outputs from the DTCWT, it is clear that the six directional subbands at a given scale need to be rotationally similar (apart from their different orientations). However we see from fig. 1(a) that the 45° and 135° subband

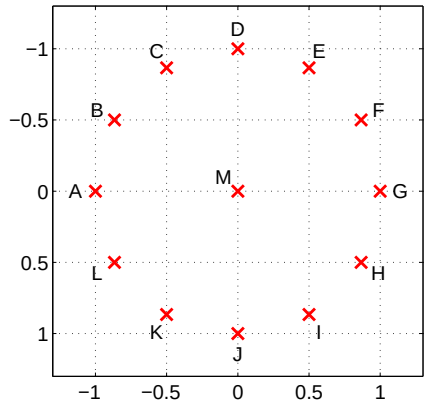

Figure 2: *The 13-point circular sampling pattern for DTCWT coefficients at each keypoint location. (The y-axis is inverted because image matrices have an inverted vertical axis.)*

responses comprise waves of rather higher frequency than the other four subbands, and this translates to a pass-band centre frequency which is further from the origin in the 2-D frequency plane. We can explain this by noting that 2-D wavelet filters comprise various combinations of *Lo* and *Hi* (lowpass and highpass) 1-D filters. The centre of the *Hi* filter pass-band is approximately three times that of the *Lo* pass-band (because they are half-band filters), so the *Hi-Hi* 2-D filters (at 45° and 135°) will be at a distance of $\sqrt{3^2+3^2} = \sqrt{18}$ units from the origin, whereas the *Hi-Lo* and *Lo-Hi* filters (at $15°, 75°, 105°$ and $165°$) will only be $\sqrt{3^2+1^2} = \sqrt{10}$ units from the origin. This ratio of $\sqrt{1.8}$ is the factor by which the 45° and 135° band centres (wave frequencies) are higher than the others.

Now, for feature description, the wavelet transform does not need to provide perfect reconstruction, so we are free to modify the 1-D filters (or add new filters). Since we need the existing *Lo* and *Hi* filters to form the *Hi-Lo* and *Lo-Hi* 2-D filters, the simplest way to reduce the centre frequency of the *Hi-Hi* filters is to add a *Ba* (bandpass) 1-D filter, which is applied to both rows and columns at each scale, and to create *Ba-Ba* 2-D filters that are used instead of the *Hi-Hi* filters.

Figure 1(c) shows the frequency response of the original *Hi* filter and the modified *Ba* filter that has been designed to achieve this. To keep the same impulse response envelope but with lower underlying wave frequency, the frequency response has been shifted down so that its mid-band frequency is reduced by the required factor $1/\sqrt{1.8}$. To preserve computational efficiency, the *Ba* filter operates at the same sampling frequency as the *Hi* filter, and this corresponds to 2 units on the x-axis of fig. 1(c). Hence the filter still has to rely on the lowpass filter at the next scale up to define its cut-off above 1 unit (half the sampling rate of the filter). This explains why it has not been possible to shift the upper edge of the *Ba* filter down as much as we would wish. We have however been able to preserve the Hilbert transform properties of the new filter and its time reverse, so that the complex response has negligible gain at negative frequencies. In 2-D, fig. 1(b) shows that the new filter is able to create 45° and 135° subband responses which are much closer in wave frequency to those of the other bands. The extra computation for the 2-D DTCWT with the new filters is about 9% if they

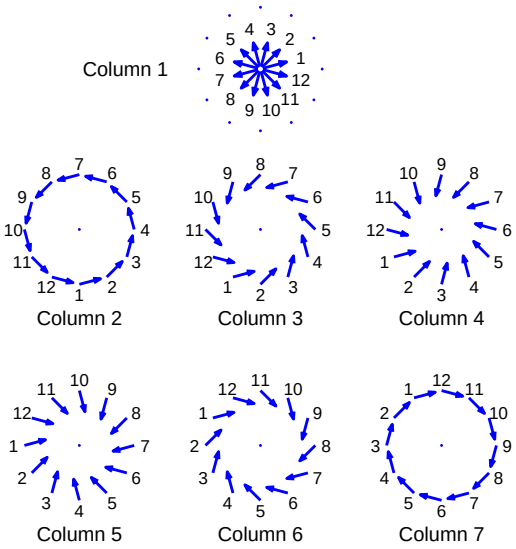

Figure 3: *Shows how each column of the polar matching matrix P is comprised of a set of rotationally symmetric samples from the 6 subbands and their conjugates, whose orientations are shown by the arrows. Numbers give the row indices in P.*

are only used below level 1, and is about 18% if they are used at all levels.

One other problem with the responses of fig. 1(a) is that the phases of the central point of each response are not all the same. So in fig. 1(b) we have chosen to make these phases all zero, by multiplying the 6 band outputs at each scale by $\{j,-j,j,-1,1,-1\}$ respectively; and thus, at the centre of each response, the real part has a white peak and the imaginary part has a transition from black to white.

3. THE POLAR MATCHING MATRIX

We now get to the fundamental part of this algorithm. Like the technique of polar mapping, the aim is to sample the six directional subbands at a given scale on a grid, centred on the desired keypoint, and then to map the data to a matrix P, such that rotations of the image about the keypoint are converted into linear cyclic shifts down the columns of P. The sampling grid that is centred on the keypoint is shown in fig. 2. It is circularly symmetric and the sampling interval is chosen to be 30°, to match that of the directional subbands. Hence we have 12 samples around the circle (A ... L) and one at its centre (M). The radius of the circle is equal to the sampling interval of the DTCWT subbands at the given scale, as this is an appropriate interval to avoid aliasing and yet provide a rich description of the keypoint locality.

Our technique for obtaining samples on the circular grid around each keypoint is to use bandpass interpolation. The information contained in a given directional complex subband is bandlimited to a particular region of the 2-D frequency space, which has a centre frequency $\{\omega_1,\omega_2\}$. Bandpass interpolation may be implemented by

1. a frequency shift by $\{-\omega_1,-\omega_2\}$ down to zero frequency (i.e. a multiplication of the complex subband coefficients by $\exp[-j(\omega_1 x_1+\omega_2 x_2)]$ at each sampling point $\{x_1,x_2\}$),
2. a conventional lowpass spline or bi-cubic interpolation to each new grid point (e.g. using Matlab function *interp2*),
3. an inverse frequency shift up by $\{\omega_1,\omega_2\}$ (a multiplication by $\exp[j(\omega_1 y_1+\omega_2 y_2)]$ at each grid point $\{y_1,y_2\}$).

To simplify notation for the mapping to matrix P, for a given keypoint locality $\{A,B,\ldots M\}$ in fig. 2 we denote the 13 subband coefficients by $\{a_d,b_d,\ldots m_d\}$, where $d=1\ldots6$ indicates the direction of the subband at $(30d-15)^\circ$ in fig. 1(b). We also note from fig. 1(b) that if the impulse response of a subband is rotated by 180°, then its real part is unaltered and its imaginary part is negated. This is equivalent to taking the conjugate of the complex coefficients, so the 13 outputs of a (virtual) subband at $(30d-15+180)^\circ$ will be $\{a_d^*,b_d^*,\ldots m_d^*\}$. The 12×7 matrix P is then formed from the 13×6 coefficients and their conjugates as follows:

$$P=\begin{bmatrix}
m_1 & j_1 & k_1 & l_1 & a_1 & b_1 & c_1\\
m_2 & i_2 & j_2 & k_2 & l_2 & a_2 & b_2\\
m_3 & h_3 & i_3 & j_3 & k_3 & l_3 & a_3\\
m_4 & g_4 & h_4 & i_4 & j_4 & k_4 & l_4\\
m_5 & f_5 & g_5 & h_5 & i_5 & j_5 & k_5\\
m_6 & e_6 & f_6 & g_6 & h_6 & i_6 & j_6\\
m_1^* & d_1^* & e_1^* & f_1^* & g_1^* & h_1^* & i_1^*\\
m_2^* & c_2^* & d_2^* & e_2^* & f_2^* & g_2^* & h_2^*\\
m_3^* & b_3^* & c_3^* & d_3^* & e_3^* & f_3^* & g_3^*\\
m_4^* & a_4^* & b_4^* & c_4^* & d_4^* & e_4^* & f_4^*\\
m_5^* & l_5^* & a_5^* & b_5^* & c_5^* & d_5^* & e_5^*\\
m_6^* & k_6^* & l_6^* & a_6^* & b_6^* & c_6^* & d_6^*
\end{bmatrix}$$

The rational for choosing this mapping can be understood from fig. 3, which shows each of the columns of P in diagramatic form using arrows on the grid of fig. 2 to represent the direction of each subband. Hence all the samples in column 1 of P are taken at the midpoint M and correspond to the 6 subbands and their conjugates taken in sequence. The arrow labelled '1' is from the 15° subband, arrow '2' is from the 45° subband, arrow '7' is from the conjugate of the 15° subband (i.e. the 195° subband), and so on. The circle of arrows for column 2 shows the location and subband from which each element in column 2 of P is taken, and this is also shown for the remaining columns. Thus we see that each column of P represents a particular pattern of rotationally symmetric combinations of sampling location and subband orientation, such that if an object is rotated clockwise about the centre of the sampling pattern by $k\times30^\circ$ (k integer), then each column of P will be cyclically shifted k places downwards.

We note that, since the patterns are cyclic, the start and end point for each column of P are arbitrary; so for simplicity we have chosen each row of P to contain samples from the same subband (i.e. to be of the same orientation) and the subbands are in numerical order down each column. The conjugated wavelet coefficients representing subbands with orientations between 180° and 360° are in the lower half of P. All coefficients from the 13-point sampling pattern in the 6 subband orientations are included in P in either their normal or conjugated form, and those from the midpoint M are included in both forms. P is thus a rich descriptor of the band-limited pixel intensities in the locality defined by the 13-point pattern, convolved with the impulse responses of the wavelet subbands. The diameter of each subband response in fig. 1(b) to its half-power points is approximately equal to the radius of the 13-point pattern for that scale.

4. FOURIER-BASED MATCHING

In order to perform rotation-invariant object detection, a matching technique is required which measures the correlation between a candidate locality in the search image and all possible rotations of a reference object in an efficient way. We assume (in the same way that SIFT does [2]) that candidate locations and scales for the search have been defined by locating extrema in scale-space in both the search and reference images. Hence shift-invariance is not needed in our matching process, although it may be desirable to include some insensitivity to shift to allow for inaccuracies in the keypoint locations and for modest shape distortions between the two regions being matched. This is discussed later.

The Fourier transform is well-known to be a useful aid to performing cyclic correlations, and in conjunction with the mapping to the P matrix, as above, it turns out to be effective at performing rotational correlations too. The basic idea is to form matrices $P_{r,i}$ at every keypoint i in the reference image, and to form matrices $P_{s,j}$ at all candidate keypoints j in the search image. For a given reference and candidate keypoint pair $\{i, j\}$, we wish to calculate the correlation between $P_{r,i}$ and $P_{s,j}$ at all possible cyclic shifts of their columns, including fractional sample shifts at some relatively fine spacing. There are 12 samples in each column and if, typically, we require $\frac{1}{4}$-sample intervals (equivalent to $30/4 = 7.5^\circ$ rotational spacing) to get accurate estimation of correlation peaks, then 48 correlation samples are needed on each of the 7 matrix columns for each pair $\{i, j\}$. To compare N pairs requires of the order of $N \times 48 \times 12 \times 7 = 4032N$ complex multiply-and-add operations.

The amount of computation can be greatly reduced if the columns of all the P matrices are transformed with a 12-point FFT before the pairwise matching processes start. The computation load for these FFTs can be largely ignored since typically the number of P matrices $\ll N$. The pairwise correlation process for each transformed matrix pair $\overline{P}_{r,i}$ and $\overline{P}_{s,j}$ then becomes:

1. Multiply each Fourier component of $\overline{P}_{s,j}$ with the conjugate of the equivalent Fourier component of $\overline{P}_{r,i}$ to get a matrix $\overline{S}_{i,j}$ ($12 \times 7 = 84$ complex multiplies).
2. Accumulate the $12 \times 7 = 84$ elements of $\overline{S}_{i,j}$ into a 48-element spectrum vector $\overline{\mathbf{s}}_{i,j}$ (84 complex adds, see below for details).
3. Take the real part of the inverse FFT of $\overline{\mathbf{s}}_{i,j}$ to obtain the 48-point correlation result $\mathbf{s}_{i,j}$ ($\leq 48 \times \log_2(48) = 270$ complex multiply-an-adds).

The computation for N pairwise comparisons with the Fourier method will thus be $\leq (84 + 270)N = 354N$ complex multiply-an-adds, i.e. less than one tenth of that for the direct method.

An additional advantage of processing in the Fourier domain is that the interpolation of the correlation result from 12 up to 48 samples can be performed efficiently by padding the spectra with zeros before the inverse FFT is taken. However this must be done rather carefully because different columns of $\overline{S}_{i,j}$ are bandpass signals with differing centre frequencies. Optimum interpolation is achieved if the zero-padding is introduced over the part of the spectrum which is likely to contain least energy, for each column of $\overline{S}_{i,j}$, as follows.

Consider forming a P matrix from the complex wavelet coefficients of a simple object that is a single step edge which is rotated slowly through 360° from the horizontal and is centred on our 13-point pattern. Looking carefully at figs. 2 and 3, we note that the edge at any angle will pass through the centre of point M so the phases of the coefficients in column 1 of P will remain constant as the edge rotates, while the amplitude peak will gradually shift from one subband to the next. Hence column 1 is a lowpass function of rotation angle θ, so its centre frequency will be zero and the zero-padding should be performed equally on the higher positive and negative regions of the spectrum.

To illustrate this we drop the $\{i, j\}$ subscripts, label the 48 elements of vector $\overline{\mathbf{s}}_{i,j}$ as $\{s_{-24} \ldots s_{-1}, s_0 \ldots s_{23}\}$, and assume that the zero frequency (dc) component is s_0 (i.e. use the Matlab *fftshift* convention). Similarly we label the elements of column v of matrix $\overline{S}_{i,j}$ as $\{S_{-6,v} \ldots S_{-1,v}, S_{0,v} \ldots S_{5,v}\}$. Hence for column 1:

$$s_u = \begin{cases} S_{u,1} & \text{for } -6 \leq u \leq 5 \\ 0 & \text{elsewhere.} \end{cases}$$

Now consider columns 4 and 5 of the matrix, where the subband directions in fig. 3 are almost radial (in fact they are at $\pm 15^\circ$ to the radial direction). As our single edge slowly rotates with angle θ, it will slide past each subband wave on the circumference of the 13-point pattern, causing a fairly rapid change in phase ϕ of the nearest wavelet coefficient. For the DTCWT with 14-tap Q-shift filters ([5], table 2), $d\phi/d\theta \approx 4$ in these columns. In columns 3 and 6 of the matrix, the subbands are at $\pm 45^\circ$ to the radial direction and the rotating edge cuts the subband waves more slowly so that $d\phi/d\theta \approx 3$. Finally in columns 2 and 7, the subbands are at $\pm 75^\circ$ to the radial direction and $d\phi/d\theta \approx 1$. For subbands at ψ to the radial direction, the general formula is $d\phi/d\theta \approx 4.2\cos(\psi)$.

In the Fourier domain a rate of phase rotation, $d\phi/d\theta = k$, corresponds to k cycles of phase shift for each complete cycle of rotation of the object, and hence shifts the centre frequency of the subband down by k frequency bins. Thus the rules for accumulating columns $v = 2$ to 7 of $\overline{S}$ into the vector $\overline{\mathbf{s}}$ with frequency shifts of k bins are:

$$\begin{array}{ll} v = 4,5 \\ k = 4: \end{array} \qquad s_u = \begin{cases} s_u + S_{u+12,v} & \text{for } -10 \leq u \leq -7 \\ s_u + S_{u,v} & \text{for } -6 \leq u \leq 1 \end{cases}$$

$$\begin{array}{ll} v = 3,6 \\ k = 3: \end{array} \qquad s_u = \begin{cases} s_u + S_{u+12,v} & \text{for } -9 \leq u \leq -7 \\ s_u + S_{u,v} & \text{for } -6 \leq u \leq 2 \end{cases}$$

$$\begin{array}{ll} v = 2,7 \\ k = 1: \end{array} \qquad s_u = \begin{cases} s_u + S_{u+12,v} & \text{for } u = -7 \\ s_u + S_{u,v} & \text{for } -6 \leq u \leq 4 \end{cases}$$

Note that in the above accumulation, it is not the elements of $\overline{S}$ which are shifted, but just the start and end of the zero padding, because the underlying 48-point spectrum (before 'padded' elements are set to zero) is 4 cyclic repeats of the original 12-point spectrum, from Nyquist sampling theory.

5. EXTENSION TO MULTIPLE SCALES

For simplicity, the above technique has been described only for subband coefficients from a single scale of the DTCWT. However it can readily be extended to multiple scales by adding extra columns to P. The simplest extension is to include coefficients from the midpoint M of the next coarser scale as an 8^{th} column, with the coefficients ordered exactly as in column 1. Column 8 of the correlation matrix $\overline{S}$ is then accumulated into $\overline{\mathbf{s}}$ with the same zero padding as column 1 as it also is a lowpass signal. The image area covered by

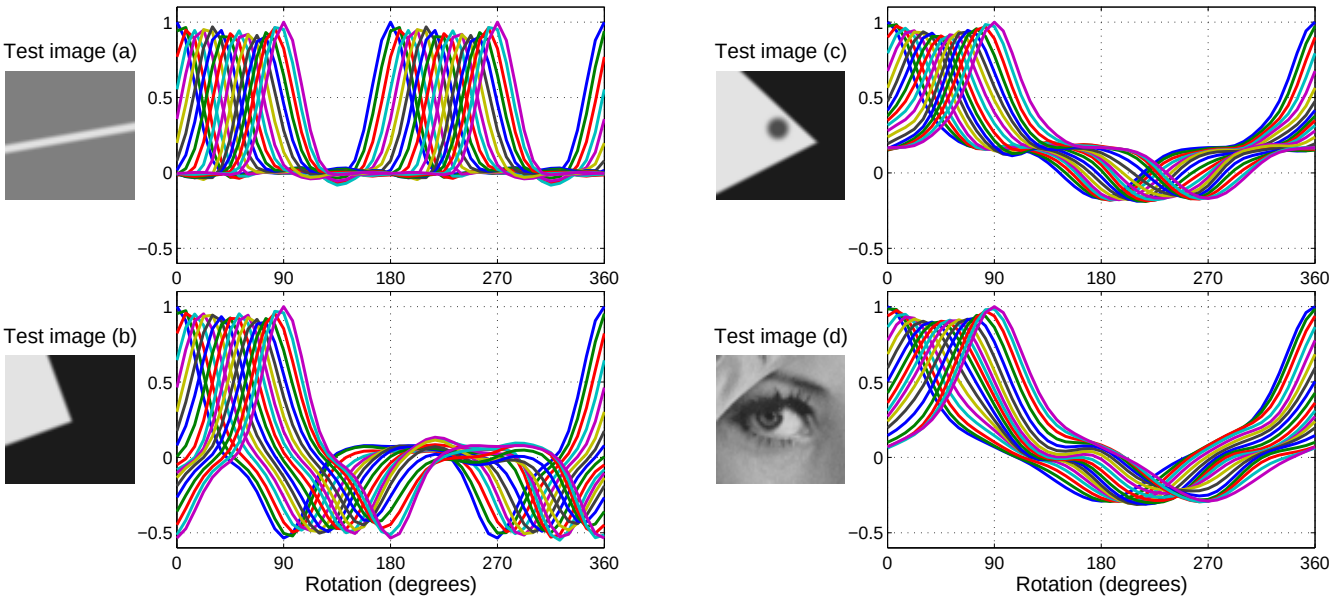

Figure 4: *Correlation results for 4 test images: bar, 90° corner, corner+blob, eye of 'Lena'. Each set of curves shows the output of the normalised correlator for 48 angles in 7.5° increments, when the test image is rotated in 5° increments from 0° to 90°. Levels 4 and 5 of the DTCWT were used in an 8-column P matrix format. The diameter of the 13-point sampling pattern is half the width of the subimages shown.*

just the centre coefficient at the next coarser level approximately equals that covered by the 13-point pattern at the current level.

To extend to the next finer scale as well, 7 more columns can be added to P, just like the first 7 but from the finer scale subbands; and to cover an equivalent area if required, a further ring of 12 coefficients at a radius of 2 units can be generated to produce 6 more columns of P. Since this doubles $d\phi/d\theta$, for these last 6 columns the frequency shifts for the zero-padding must also be doubled. The descriptor is by now getting very detailed, and so, for many applications, the 8-column 2-scale version of P with $8 \times 12 = 96$ complex elements is likely to be sufficient.

To provide better interpolation of scale, SIFT derives several intermediate scales in each octave. This is also possible with the DTCWT by appropriate scaling of the input image, each with a new DTCWT. The appropriate scale at which to form each keypoint descriptor is obtained from the keypoint detector, and using this (as in SIFT) the complete object descriptor can be made approximately scale invariant.

6. RESULTS AND CONCLUSIONS

Here we have concentrated on the theory of our technique, as it is quite complicated, so there is space only for limited results. For 4 shapes of increasing complexity, we show in fig. 4 the result of correlating rotated copies of the object with the original, and plot the normalised correlation vector, covering 0 to 360°, in each case. The peak of the vector indicates the degree of match and its location gives the object orientation. If the whole process were completely rotation invariant, all the correlation peaks should be unity and the curves be the same shape. In practise, with the improvements of section 2, all the peaks exceed 0.896; without these improvements to the diagonal subbands, the peaks dip to 0.66. The highest *cross* correlation between these images is 0.397.

In this paper, we have shown how rotational correlations may be performed using interpolated complex samples from the DTCWT, utilising both phase and amplitude information. There is considerable scope for extending these ideas to increase the robustness to typical image distortions (e.g. due to change of viewpoint or lighting) and small mis-registration of keypoints. For example, each row of the P matrix uses samples from a given direction of subband, so it would be possible to perform phase corrections using inter-scale prediction to achieve some invariance to shifts and small geometric distortions, as proposed in [8]. This is a current topic of research in our group.

REFERENCES

[1] C Harris and M Stephens, "A combined corner and edge detector, " in *Proc. Alvey Vision Conference*, pp. 147–151, 1988.

[2] D G Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[3] E P Simoncelli and W T Freeman, "The Steerable Pyramid: a flexible architecture for multi-scale derivative computation," in *Proc. ICIP 1995*, Washington DC, USA, Oct. 1995, vol. 3, pp. 444–447.

[4] P R Hill, D R Bull and C N Canagarajah, "Rotationally invariant texture features using the dual-tree complex wavelet transform," in *Proc. ICIP 2000*, Vancouver BC, Canada, Sept. 10–13, 2000, vol. 3, pp. 901–904.

[5] N G Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Journal of Applied and Computational Harmonic Analysis*, vol. 10, no. 3, pp. 234–253, May 2001.

[6] I W Selesnick, R G Baraniuk and N G Kingsbury, "The Dual-Tree Complex Wavelet Transform," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 123–151, Nov. 2005.

[7] J Fauqueur, N G Kingsbury and R Anderson, "Multiscale keypoint detetction using the Dual-Tree Complex Wavelet Transform," submitted to *ICIP 2006*, Atlanta, USA, Oct. 8–11, 2006.

[8] R Anderson, N Kingsbury and J Fauqueur "Coarse level object recognition using interlevel products of complex wavelets," in *Proc. ICIP 2005*, Genoa, Italy, Sept. 11–14, 2005.