Keypoint matching with Complex wavelets

PASHMINA Z. BENDALE AND NICK G. KINGSBURY

UNIVERSITY OF CAMBRIDGE

pb397@cam.ac.uk, ngk@eng.cam.ac.uk

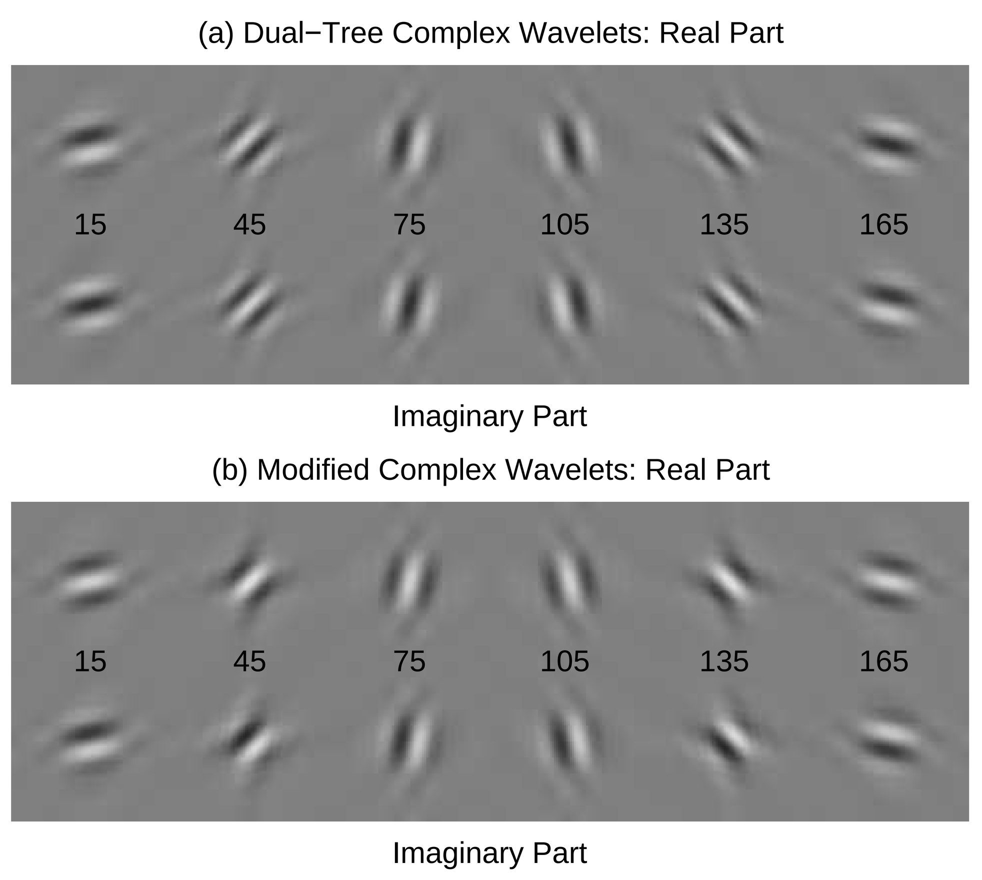
I. Objective

Match corresponding locations in two images of an object in a rotation-invariant way

II. DTCWT

The Dual Tree Complex Wavelet Transform [DTCWT] provides multi-scale (N levels), six-directional gradient information for images using an efficient $\mathcal{O}(N)$ filter bank. It offers the following key advantages over other methods:

1. Better directionality (Complex filters can separate adjacent quadrants of the 2D frequency spectrum)
2. Better robustness (Smoothly varying, Repeatable)
3. Shift invariance: Each subband can be interpolated independently
4. Geometry of the image features is retained in the phase
5. Complex envelope produces single tractable maximum per discontinuity (unlike the LoG,DoG)
6. Lower redundancy and greater computational efficiency as compared to Steerable Pyramid



For better rotational symmetry, the DTCWT filters are modified such that

1. The centre frequencies of all subbands are equidistant from the origin of the frequency space
2. The filters have an even-symmetric real part and odd-symmetric imaginary part

III. Method: Keypoint localisation

1. Look for points that are characterised by a large geometric mean wavelet response

$$M_\pi\{k\}(x,y) = \left(\prod_{d=1}^{6} H\{k\}(x,y,d)\right)^{1/6} , \quad (1)$$

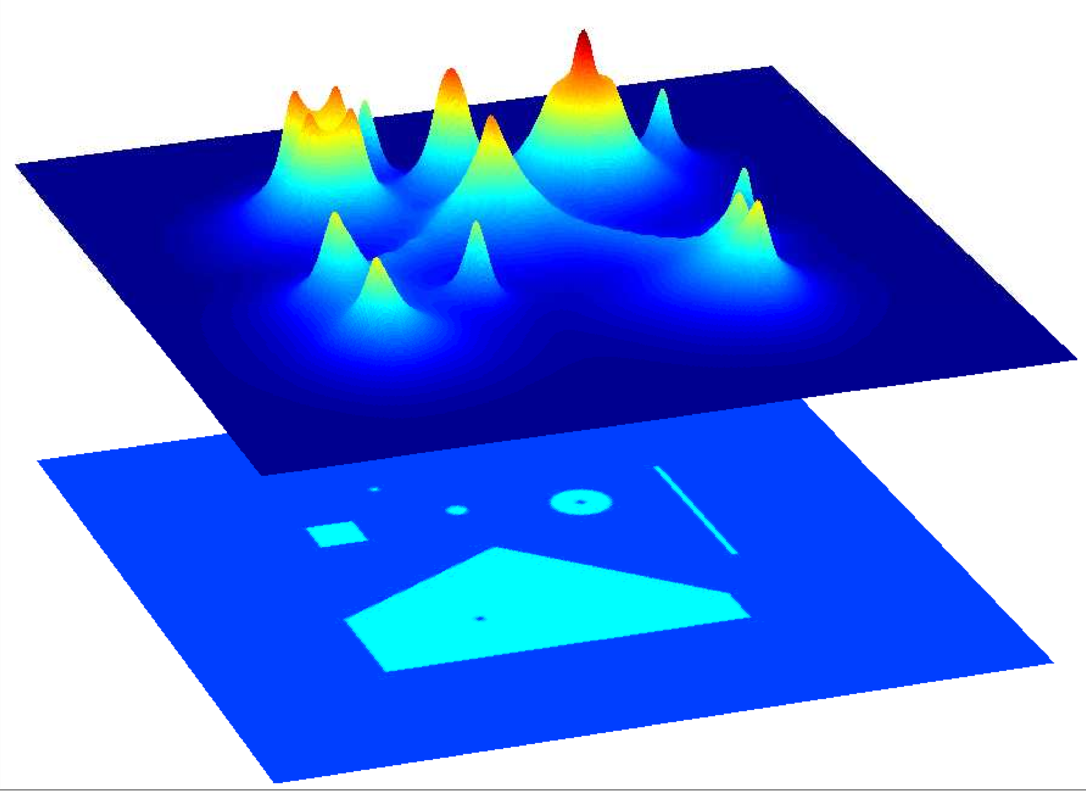
forming a multiscale discontinuity map, M_π.

2. Form accumap, A_π, a combined representation of the multiscale discontinuity map, by summing over a subset of levels, $k_s \subseteq S, S = \{1,2,\ldots N\}$

$$A_\pi(x,y) = \sum_{k\in k_s} \uparrow M_\pi\{k\}(x,y) \quad (2)$$

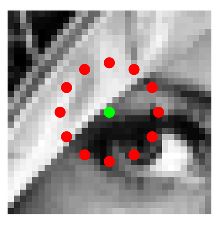
where $\uparrow$ denotes interpolation by a factor of 2 before adding to the next finer level.

3. Keypoints are detected at stable maxima in A_π.

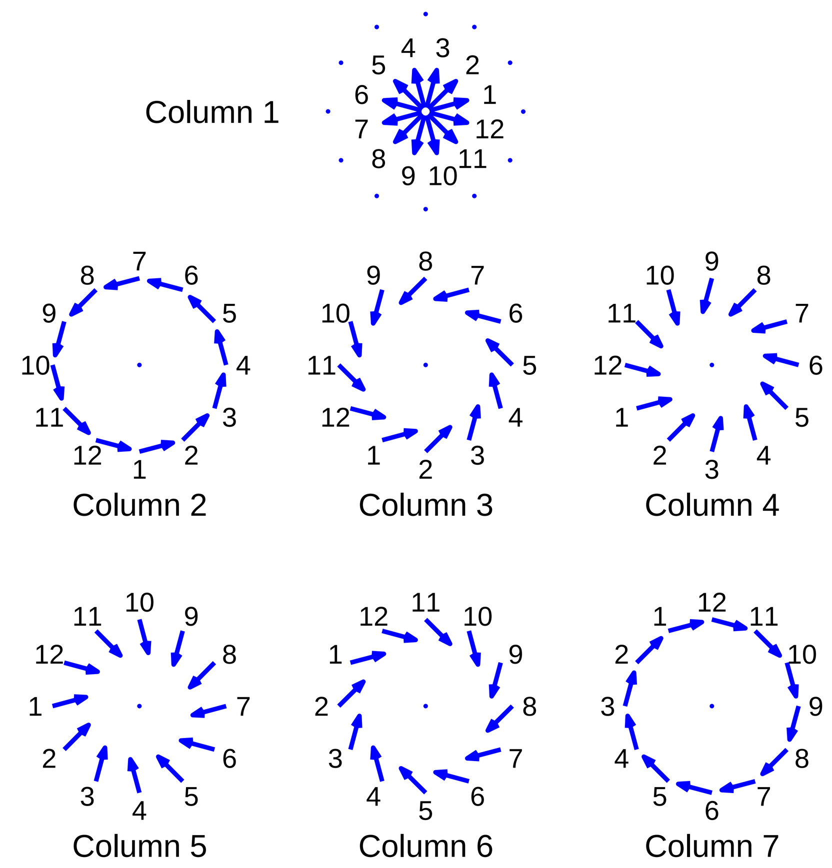


IV. P-matrix descriptor

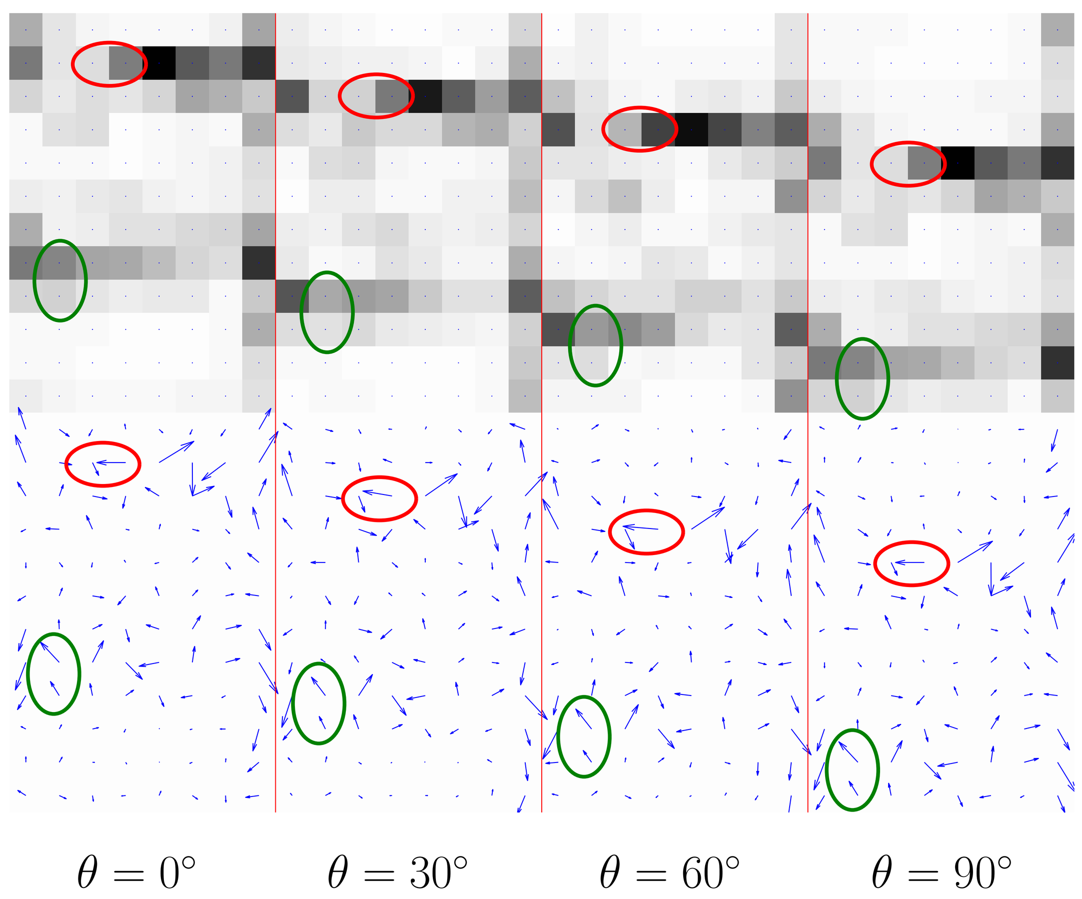
Aim:
Form a local feature descriptor from the DTCWT coefficients in a rotation-invariant way.



Assemble DTCWT responses in 12 directions at the keypoint and at 12 points on a circle around the keypoint, to form the **P**-matrix.



1. Each row of the **P**-matrix corresponds to a spatial direction.
2. There are $[m + 6\times n]$ columns. (m centre points at m scales, n rings at any scale)



Rotation of an object produces a cyclic shift in columns of **P**-matrix, at a known frequency.
Cyclic shift in the columns of the **P**-matrix is manifested as a phase shift in $\overline{\mathbf{P}} = \mathrm{FFT}_c\{\mathbf{P}\}$. Magnitude spectrum of $\overline{\mathbf{P}}$ is approximately rotation-invariant.
FFT_c denotes a column-wise FFT operation.

V. Algorithm: FFT based correlation

Aim:
Obtain a matching score for a pair of keypoints for every relative orientation in a single cross-correlation operation.
Given:
$\mathbf{P}_{r,i}$ = **P**-matrix of i^{th} reference keypoint (i=1:M)
$\mathbf{P}_{s,j}$ = **P**-matrix of j^{th} search keypoint (j=1:N)
For each keypoint,

1: $\overline{\mathbf{P}}_{r,i} = \mathrm{FFT}_c\{\mathbf{P}_{r,i}\}$
2: $\overline{\mathbf{P}}_{s,j} = \mathrm{FFT}_c\{\mathbf{P}_{s,j}\}$

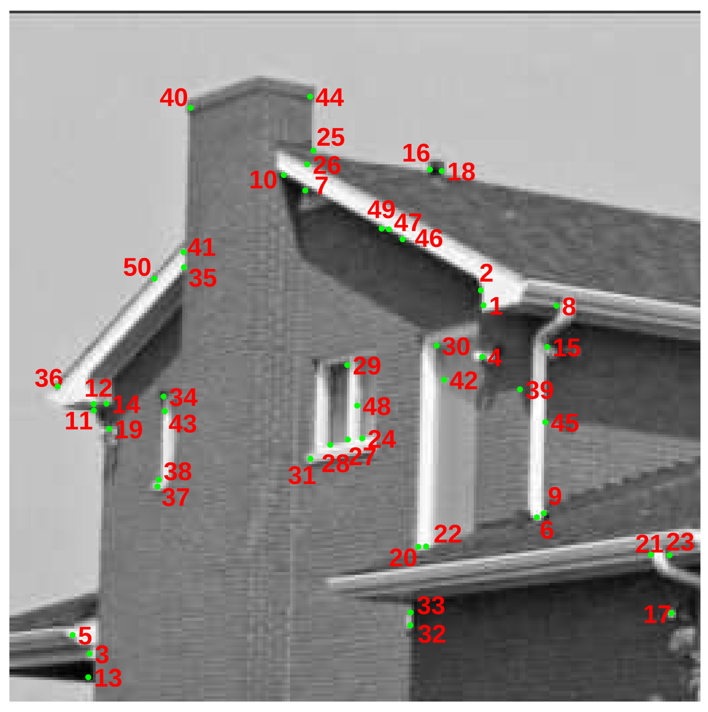
End
For each pair of keypoints,

1: $\overline{\mathbf{C}}(i,j) = \overline{\mathbf{P}}_{r,i} \cdot \{\overline{\mathbf{P}}^*_{s,j}\}$
2: Zero-pad the low-energy parts of $\overline{\mathbf{C}}(i,j)$ to get $\bar{\mathbf{c}}(i,j)$
3: $\mathbf{c}(i,j) = \sum \mathrm{IFFT}_c\{\bar{\mathbf{c}}(i,j)\}$ {$\sum$ along rows}
4: $\mathrm{MatchScore}(i,j) = \max_\theta[\mathbf{c}(i,j,\theta)]$
5: $\mathrm{MatchAngle}(i,j) = \arg\max_\theta[\mathbf{c}(i,j,\theta)]$

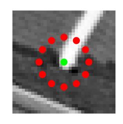
End
Cost of one correlation = 350 complex multiply-adds
The value and location of the peak in the 48-point correlation curve $\mathbf{c}(i,j)$ gives the strength and relative orientation of the best match.
Correlation scores lie within the range $[-1,1]$ and may be transformed using an inverse sigmoid function.

VII. Results

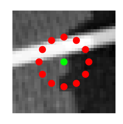
50 most dominant keypoints detected in the House image.



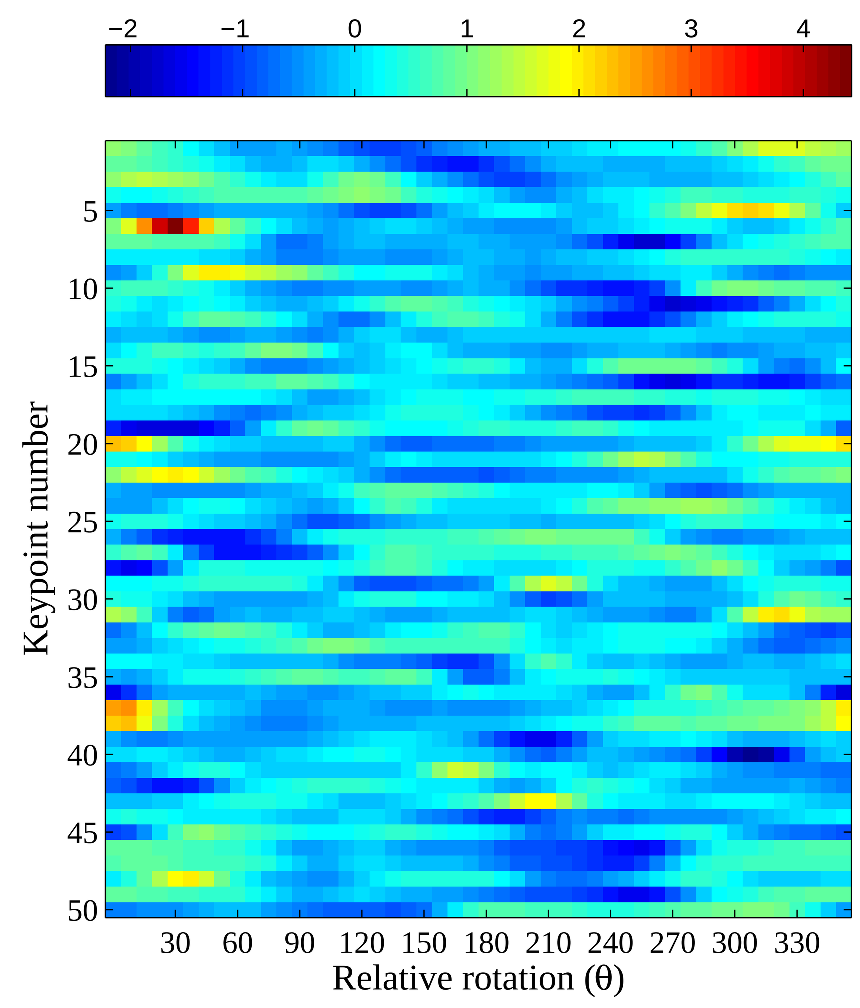
Query 1 : Keypoint region 6 rotated by 30°



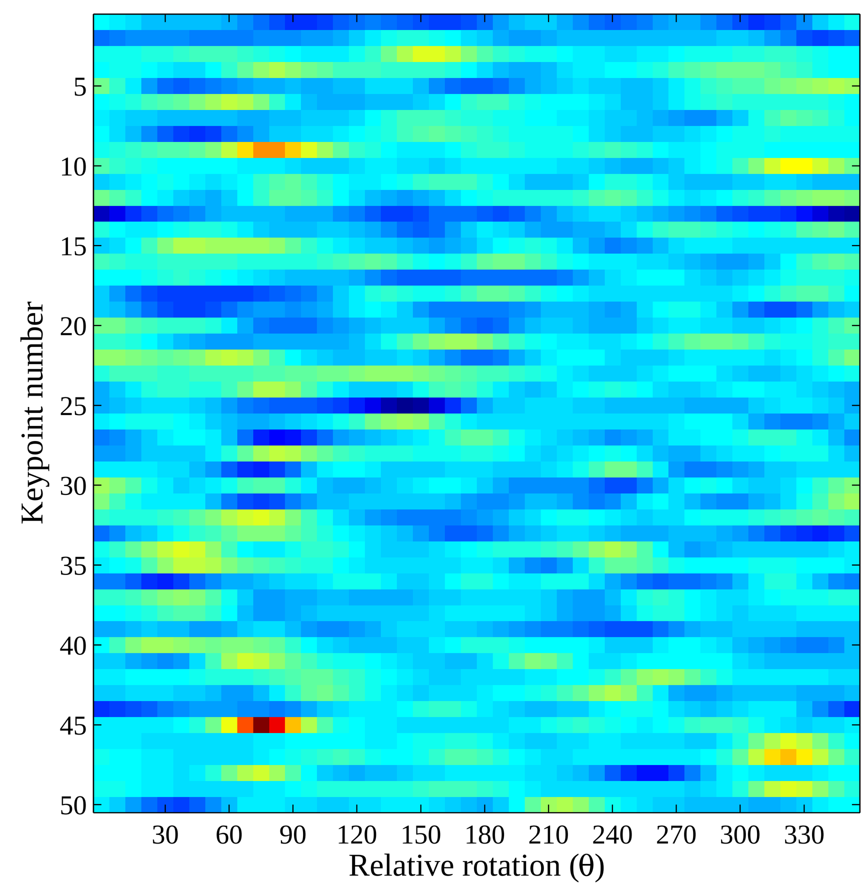
Query 2 : Keypoint region 45 rotated by 75°



Result 1: Result of cross-correlation of keypoint in Query 1 with each keypoint in the House image using the **P**-matrix descriptor.



Result 2: Result of cross-correlation of keypoint in Query 2 with each keypoint in the House image using the **P**-matrix descriptor.



In each case, the best match is found at the correct relative location and correct relative orientation.

VII. Conclusions

1. DTCWT is an efficient energy filtering tool useful for keypoint detection and description
2. A rotation-invariant local feature descriptor of arbitrary richness can be built using DTCWT coefficients owing to their interpolability
3. Match strength and relative orientation of keypoints can be efficiently inferred from FFT-based rotational correlations of the **P**-matrix descriptor

N. G. Kingsbury, *Complex wavelets for shift invariant analysis and filtering of signals*, Journal of Applied and Computational Harmonic Analysis 2001
J. Fauqueur, N. Kingsbury and R. Anderson, *Multiscale keypoint detection using the Dual-Tree Complex Wavelet Transform*, ICIP 2006
N. G. Kingsbury, *Rotation-invariant local feature matching with complex wavelets*, EUSIPCO 2006

Pashmina Z Bendale is supported by a *Gates Cambridge Scholarship*