

# Dual-Tree Wavelets and their application to Vision Systems: key-point detectors and descriptors with polar matching.

Nick Kingsbury

Signal Processing and Communications Laboratory  
Department of Engineering, University of Cambridge, UK.  
email: [ngk@eng.cam.ac.uk](mailto:ngk@eng.cam.ac.uk)    web: [www.eng.cam.ac.uk/~ngk](http://www.eng.cam.ac.uk/~ngk)

Part 2, UDRC Short Course, 21 March 2013



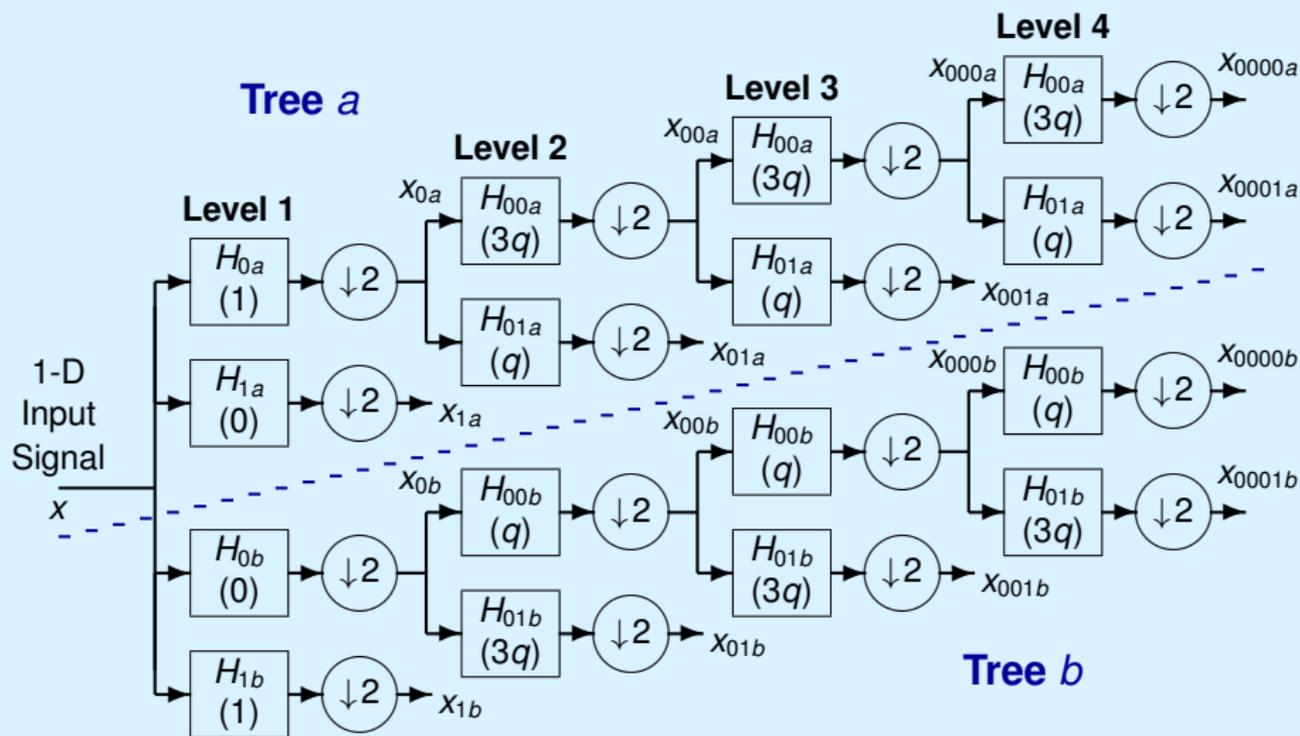
## Dual-Tree Wavelets and their application to Vision Systems

- Properties of Dual-Tree Complex Wavelets (DT CWT) and why we use them
- Multiscale Keypoint Detection using Complex Wavelets (Julien Fauqueur and Pashmina Bendale)
- Rotation-invariant Local Feature Matching (an alternative to Lowe's popular SIFT system):
  - Modifications to DT CWT to improve rotational symmetry
  - Resampling using bandpass interpolation
  - The Polar Matching Matrix descriptor
  - Efficient Fourier-based matching
  - Similar to log-polar mapping of Fourier domain, but much more localised.
- Enhancements for resilience to keypoint location errors and scale estimates.

## Features of the Dual Tree Complex Wavelet Transform (DT CWT)

- Good **shift invariance** = **negligible aliasing**. Hence transfer function through each subband is independent of shift **and** wavelet coefs can be interpolated within each subband, independent of all other subbands.
- Good **directional selectivity** in 2-D, 3-D etc. – derives from **analyticity** in 1-D (ability to separate positive from negative frequencies).
- Similarity to **Primary (V1) Cortex filters** of the human visual system.
- **Perfect reconstruction** with short support filters.
- **Limited redundancy** – 2:1 in 1-D, 4:1 in 2-D etc.
- **Low computation** – much less than the undecimated (à trous) DWT.

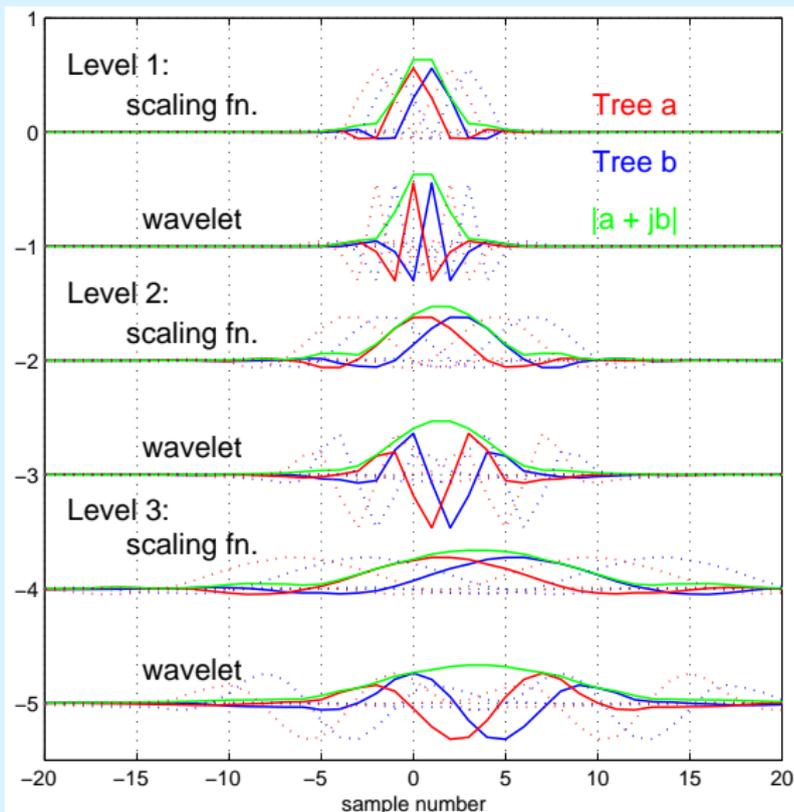
## Q-shift Dual Tree Complex Wavelet Transform in 1-D



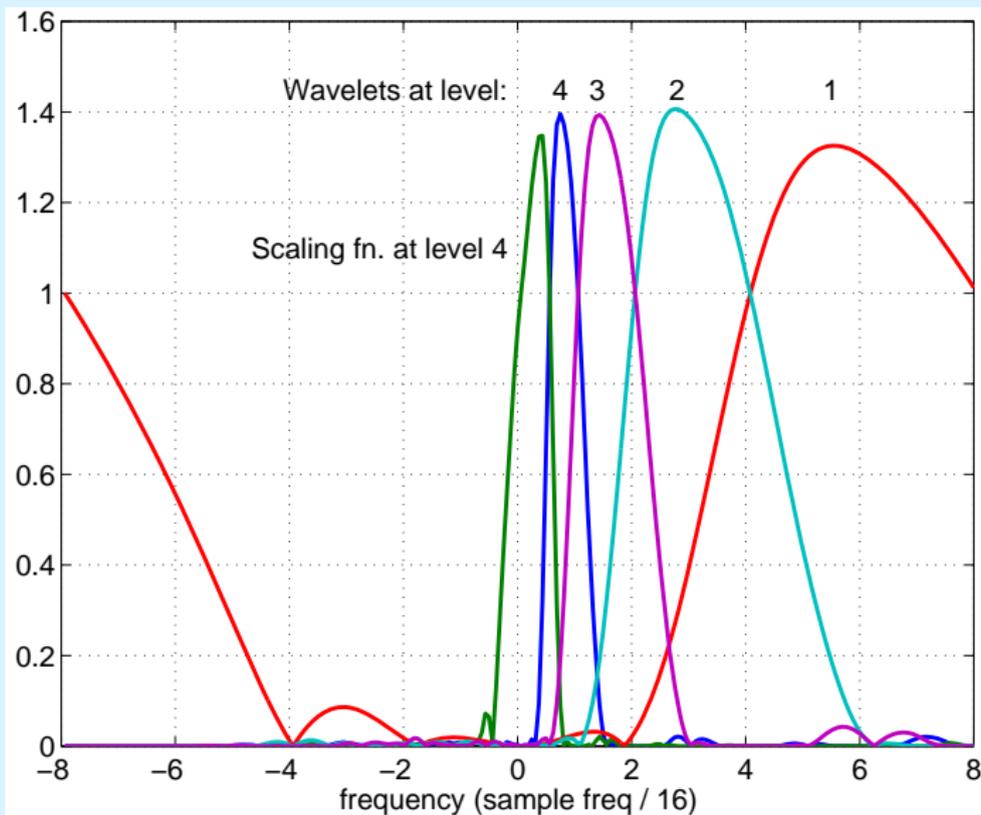
**Figure:** Dual tree of real filters for the Q-shift CWT, giving real and imaginary parts of complex coefficients from tree a and tree b respectively.

## Q-shift DT CWT Basis Functions – Levels 1 to 3

Basis functions for adjacent sampling points are shown dotted.



# Frequency Responses of 18-tap Q-shift filters

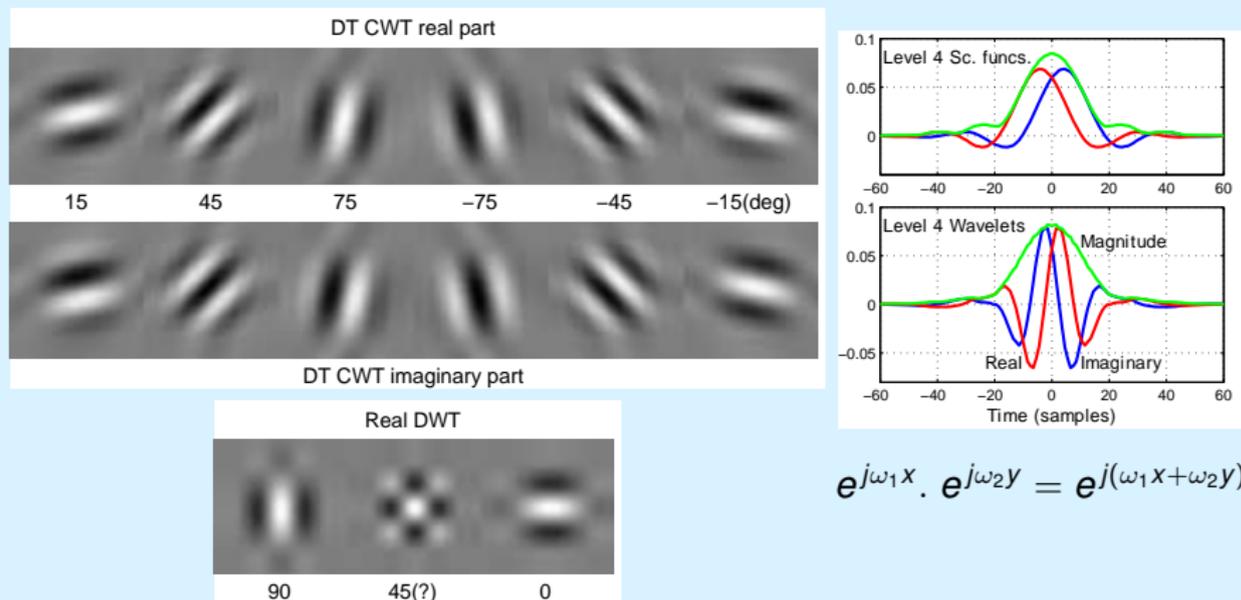


## The DT CWT in 2-D

When the DT CWT is applied to 2-D signals (images), it has the following features:

- It is performed separably, using 2 trees for the rows of the image and 2 trees for the columns – yielding a **Quad-Tree** structure (4:1 redundancy).
- The 4 quad-tree components of each coefficient are combined by simple sum and difference operations to yield a **pair of complex coefficients**. These are part of two separate subbands in adjacent quadrants of the 2-D spectrum.
- This produces **6 directionally selective subbands** at each level of the 2-D DT CWT. Fig 2 shows the basis functions of these subbands at level 4, and compares them with the 3 subbands of a 2-D DWT.
- The DT CWT is directionally selective (see fig ??) because the complex filters can **separate positive and negative frequency components** in 1-D, and hence **separate adjacent quadrants** of the 2-D spectrum. Real separable filters cannot do this!

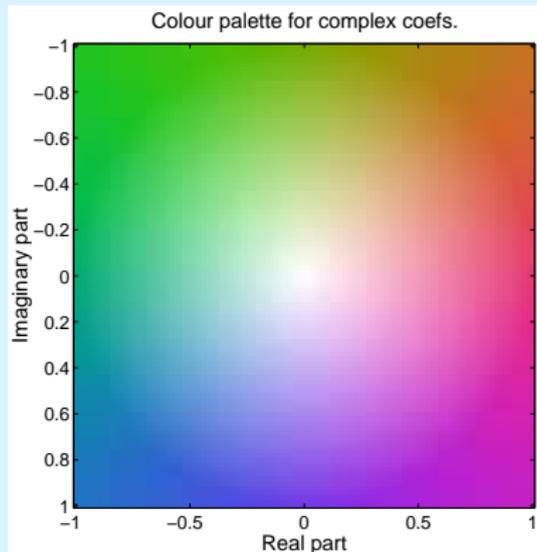
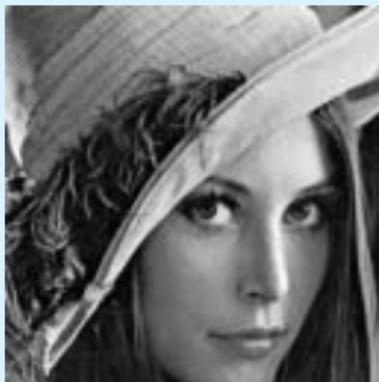
## 2-D Basis Functions at level 4



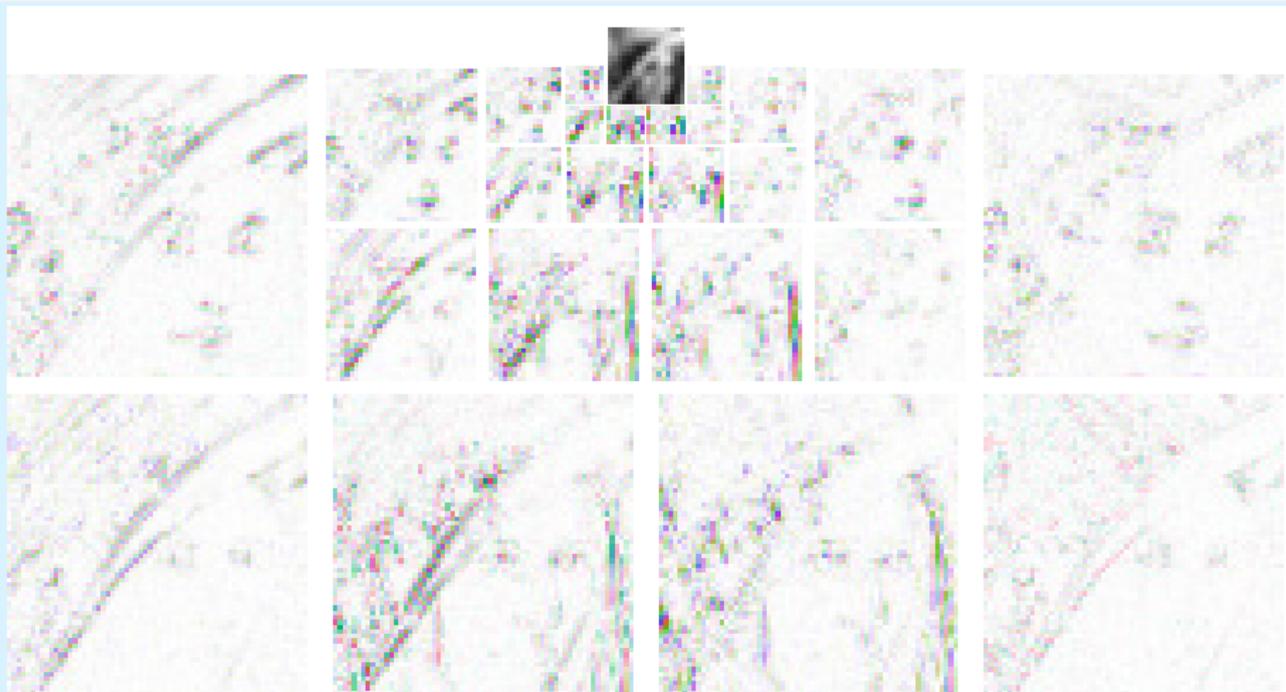
$$e^{j\omega_1 x} \cdot e^{j\omega_2 y} = e^{j(\omega_1 x + \omega_2 y)}$$

**Figure:** Basis functions of 2-D Q-shift complex wavelets (top), and of 2-D real wavelet filters (bottom), all illustrated at level 4 of the transforms. The complex wavelets provide 6 directionally selective filters, while real wavelets provide 3 filters, only two of which have a dominant direction. The 1-D bases, from which the 2-D complex bases are derived, are shown to the right.

# Test Image and Colour Palette for Complex Coefficients

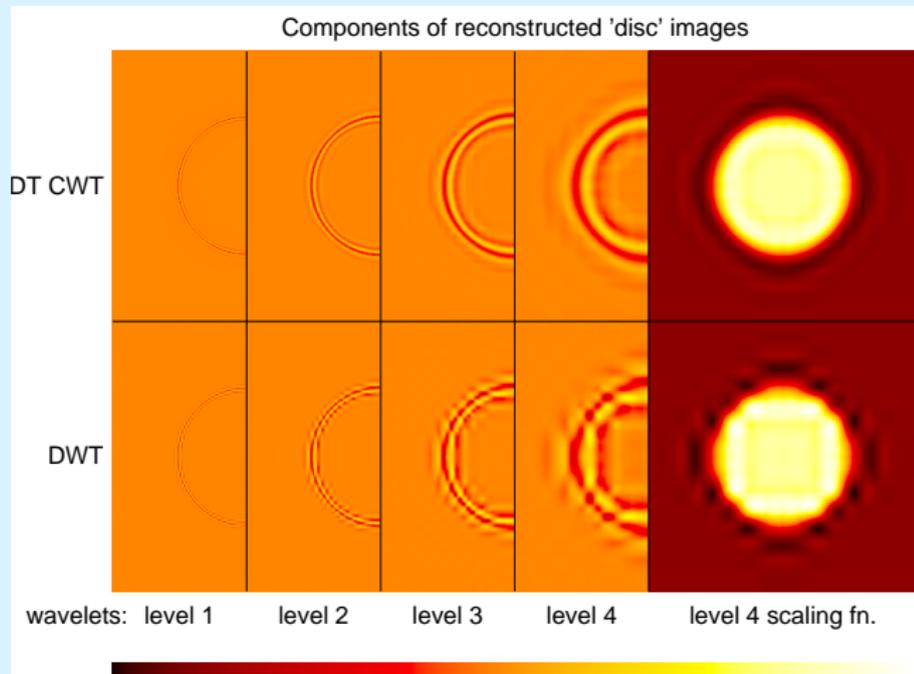
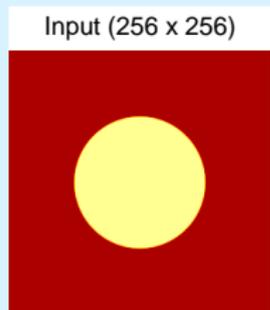


## 2-D DT CWT Decomposition into Subbands



**Figure:** Four-level DT CWT decomposition of **Lena** into 6 subbands per level (only the central  $128 \times 128$  portion of the image is shown for clarity). A colour-wheel palette is used to display the complex wavelet coefficients.

## 2-D Shift Invariance of DT CWT vs DWT



**Figure:** Wavelet and scaling function components at levels 1 to 4 of an image of a light circular disc on a dark background, using the 2-D DT CWT (upper row) and 2-D DWT (lower row). Only half of each wavelet image is shown in order to save space.

## Multi-scale Keypoint Detection using Accumulated Maps

### Basic Method:

- Detect keypoint energy as locations  $\mathbf{x}$  at scale  $k$  where complex wavelet energy exists in multiple directions  $d$  using:

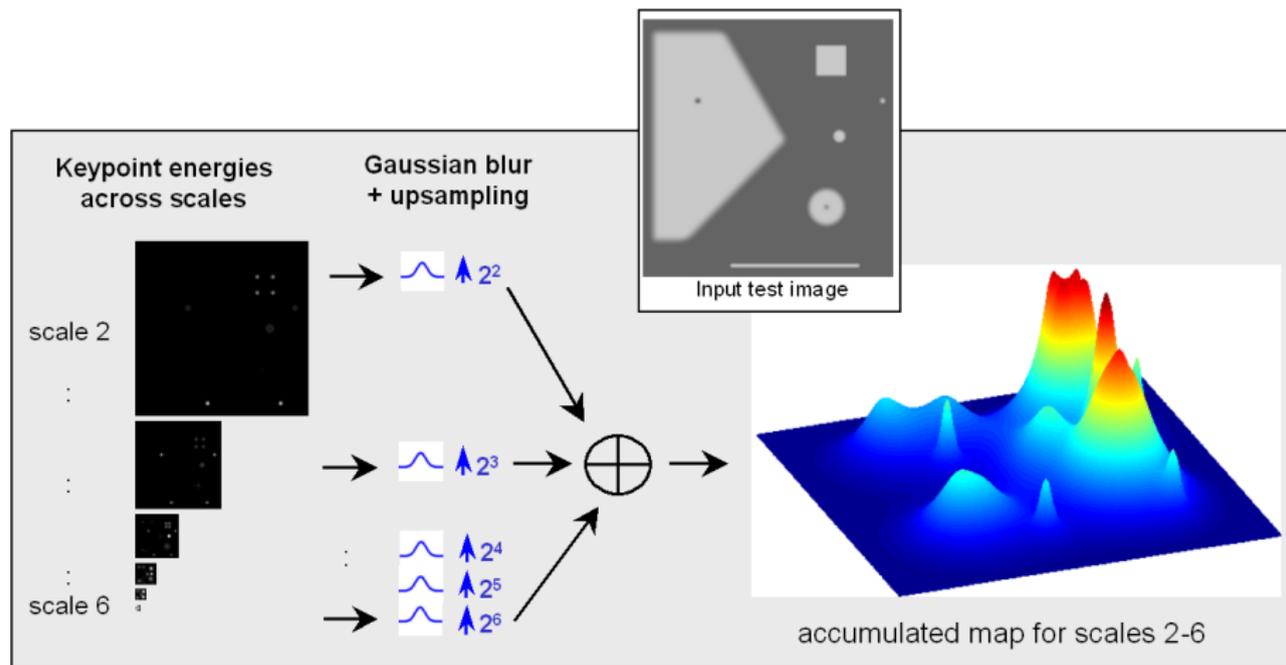
$$E_{kp}(k, \mathbf{x}) = \min_{d=1\dots 6} |w_{k,d}(\mathbf{x})|$$

$$\text{or } E_{kp}(k, \mathbf{x}) = \left( \prod_{d=1}^6 |w_{k,d}(\mathbf{x})| \right)^{\frac{1}{6}}$$

$$\text{or } E_{kp}(k, \mathbf{x}) = \frac{\frac{1}{3} \sum_{d=1}^3 |w_{k,d}(\mathbf{x})| \cdot |w_{k,d+3}(\mathbf{x})|}{\sqrt{\frac{1}{6} \sum_{d=1}^6 |w_{k,d}(\mathbf{x})|^2}}$$

- Interpolate and accumulate  $E_{kp}(k, \mathbf{x})$  across groups of relevant scales  $k$  to create the **Accumulated Map**.
- Pick keypoints as locations of maxima in the Accumulated Map.

## Multi-scale Keypoint Detection using Accumulated Maps



# Comparison of keypoint detectors

MANUAL 37 kp



DTCWT 15 of 31 kp close to manual kp



SIFT: 8 of 50 kp close to manual kp



Harris: 8 of 44 kp close to manual kp



## Rotation-Invariant Local Feature Matching

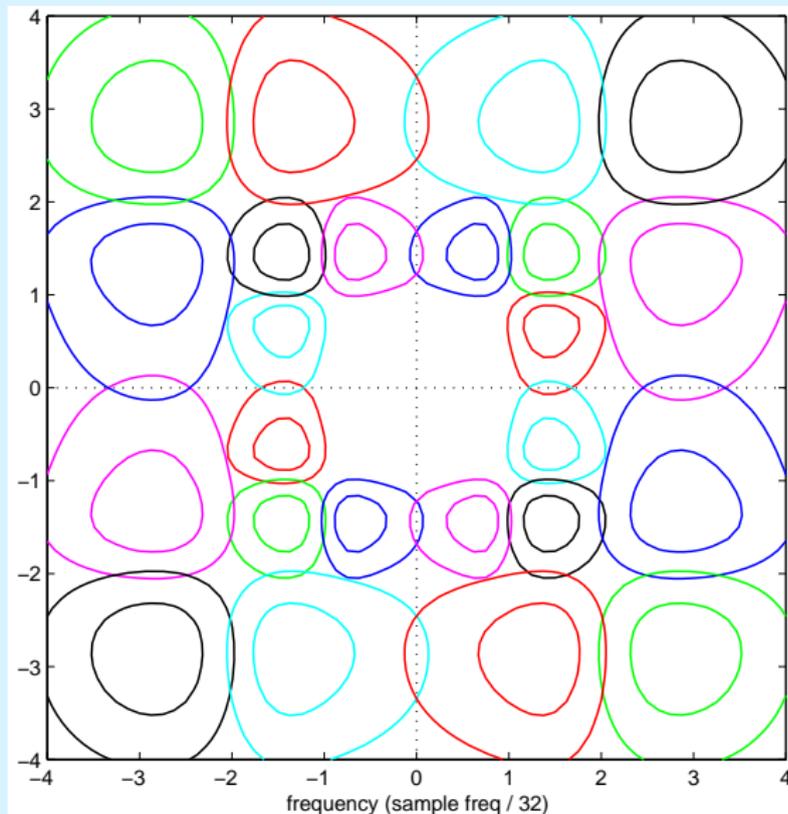
### Aims:

- To derive a **local feature descriptor** for the region around a detected keypoint, so that keypoints from similar objects may be **matched reliably**.
- Matching must be performed in a **rotationally invariant** way if all rotations of an object are to be matched correctly.
- The feature descriptor must have **sufficient complexity** to give good detection reliability and low false-alarm rates.
- The feature descriptor must be **simple enough** to allow rapid pairwise comparisons of keypoints.
- Raw DT CWT coefficients provide multi-resolution local feature descriptors, but they are tied closely to a **rectangular sampling** system (as are most other multi-resolution decompositions).

Hence we first need **better rotational symmetry** for the DT CWT.

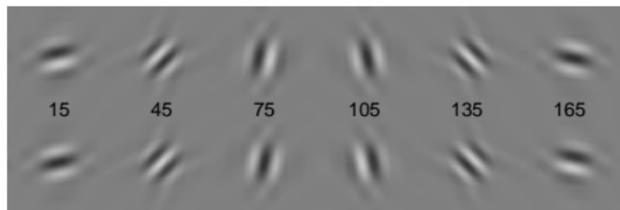
## Frequency Responses of 2-D Q-shift filters at levels 3 and 4

Contours shown at  
-1 dB and -3 dB.



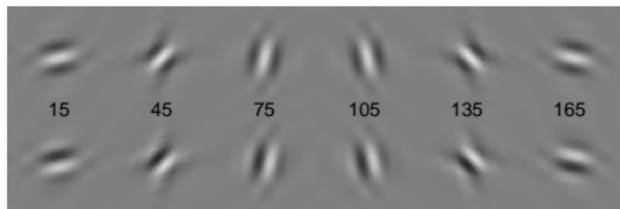
# Modification of $45^\circ$ and $135^\circ$ subband responses for improved rotational symmetry (shown at level 4).

(a) Dual-Tree Complex Wavelets: Real Part



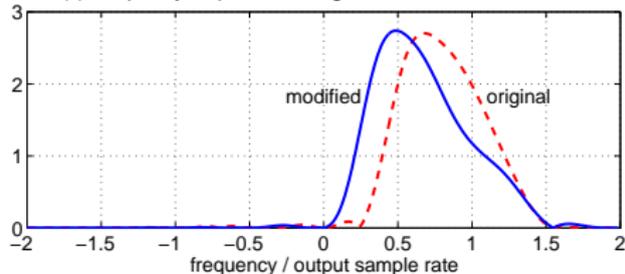
Imaginary Part

(b) Modified Complex Wavelets: Real Part



Imaginary Part

(c) Frequency responses of original and modified 1-D filters



(a) Original 2-D impulse responses;

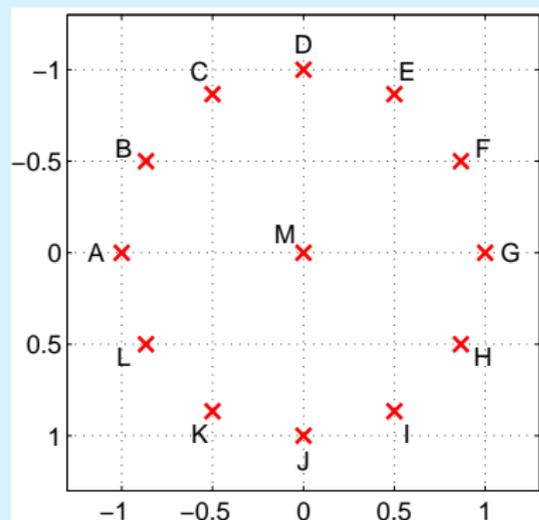
(b) 2-D responses, modified to have lower centre frequencies (reduced by  $1/\sqrt{1.8}$ ) in the  $45^\circ$  and  $135^\circ$  subbands, and even / odd symmetric real / imaginary parts;

(c) Original and modified 1-D filters.

**Better rotational symmetry** is achieved,  
but **we have lost Perfect Reconstruction**.

## 13-point circular pattern for sampling DT CWT coefs at each keypoint location

M is a precise keypoint location, obtained from the keypoint detector.

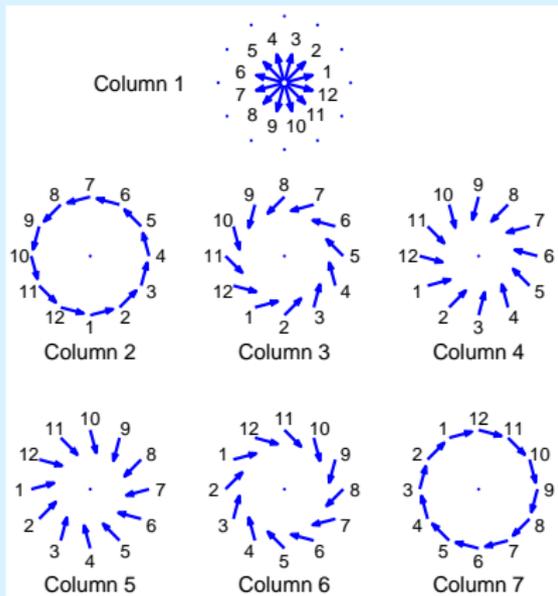


**Bandpass interpolation** calculates the required samples and can be performed on each subband independently because of the shift-invariance of the transform:

1. Shift by  $\{-\omega_1, -\omega_2\}$  down to zero frequency (i.e. multiply by  $e^{-j(\omega_1 x_1 + \omega_2 x_2)}$  at each point  $\{x_1, x_2\}$ );
2. Lowpass interpolate to each new point (spline / bi-cubic / bi-linear);
3. Shift up by  $\{\omega_1, \omega_2\}$  (multiply by  $e^{j(\omega_1 y_1 + \omega_2 y_2)}$  at each new point  $\{y_1, y_2\}$ ).

Form the Polar Matching Matrix  $P$ 

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



Each column of  $P$  comprises a set of **rotationally symmetric** samples from the 6 subbands and their conjugates (\*), whose orientations are shown by the arrows and locations in the 13-point pattern by the letters.

Numbers for each arrow give the row indices in  $P$ .

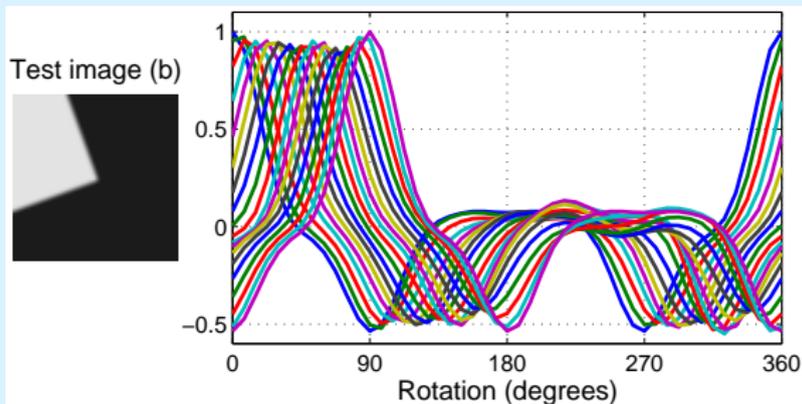
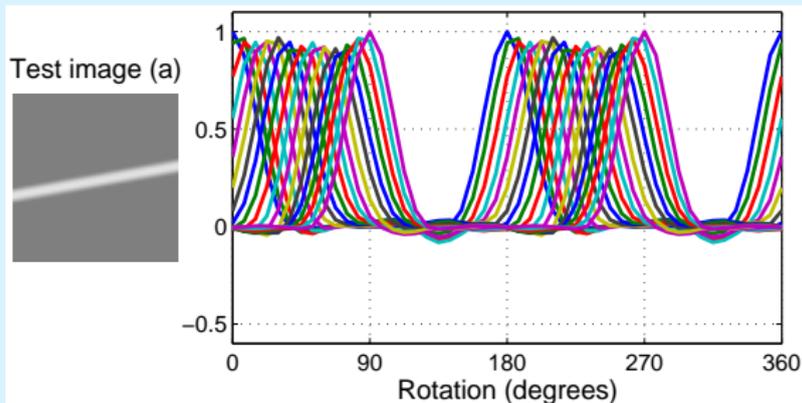
## Efficient Fourier-based Matching

Columns of  $P$  **shift cyclically with rotation** of the object about keypoint  $M$ . Hence we perform correlation matching in the **Fourier** domain, as follows:

- First, take **12-point FFT** of each column of  $P_k$  at every keypoint  $k$  to give  $\bar{P}_k$  and **normalise** each  $\bar{P}_k$  to unit total energy.
- Then, for each pair of keypoints  $(k, l)$  to be matched:
  - **Multiply**  $\bar{P}_k$  by  $\bar{P}_l^*$  element-by-element to give  $\bar{S}_{k,l}$ .
  - **Accumulate** the 12-point columns of  $\bar{S}_{k,l}$  into a 48-element spectrum vector  $\bar{\mathbf{s}}_{k,l}$  (to give a 4-fold extended frequency range and hence finer correlation steps). Different columns of  $\bar{S}_{k,l}$  are bandpass signals with differing centre frequencies, so optimum interpolation occurs if zero-padding is introduced over the part of the spectrum which is likely to contain least energy in each case.
  - Take the real part of the **inverse FFT** of  $\bar{\mathbf{s}}_{k,l}$  to obtain the 48-point correlation result  $\mathbf{s}_{k,l}$ .
  - The **peak** in  $\mathbf{s}_{k,l}$  gives the **rotation and value** of the best match.

Extra columns can be added to  $P$  for multiple scales or colour components.

## Correlation plots for two simple images

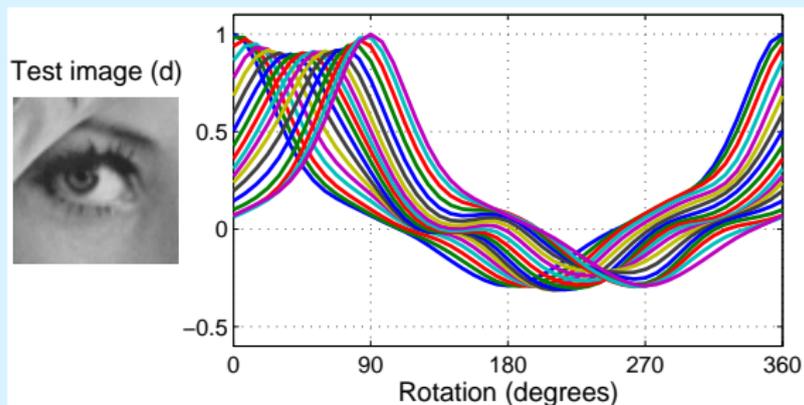
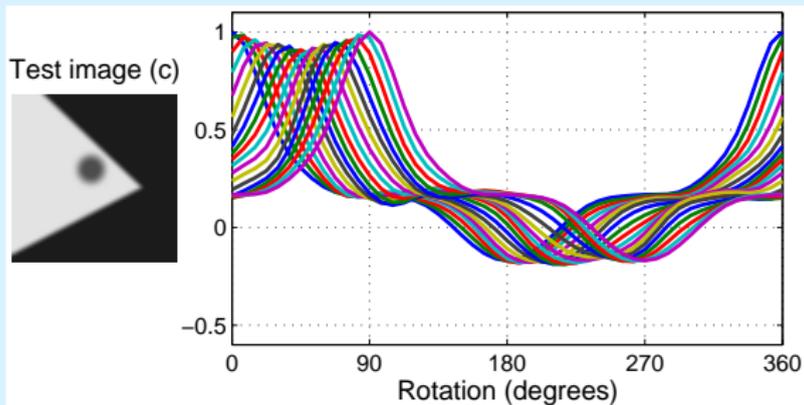


Each set of curves shows the output of the normalised correlator for 48 angles in  $7.5^\circ$  increments, when the test image is rotated in  $5^\circ$  increments from  $0^\circ$  to  $90^\circ$ .

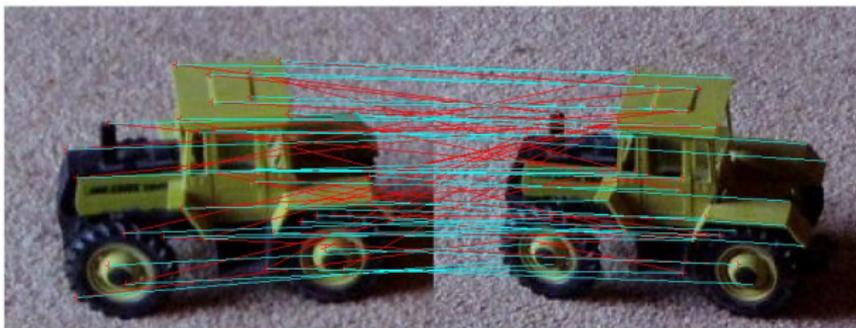
Levels 4 and 5 of the DT CWT were used in an 8-column  $P$  matrix format.

The diameter of the 13-point sampling pattern is half the width of the subimages shown.

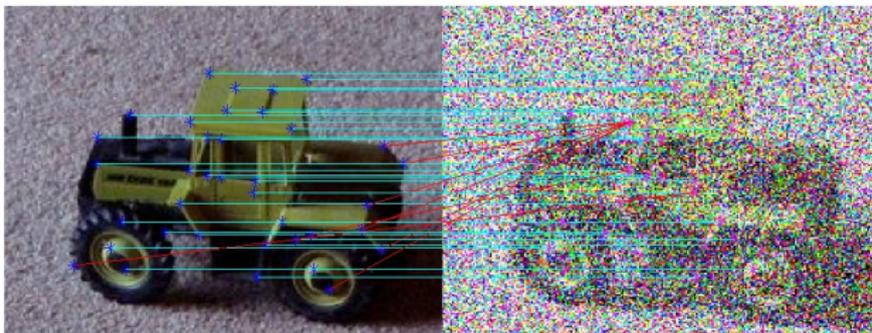
## Correlation plots for more complicated images



## Matching hand-picked keypoints across different 3D views



## Matching hand-picked keypoints in heavy noise



## Improving resilience to errors in keypoint location and scale

The basic  $P$ -matrix normalised correlation measure is **highly resilient to changes in illumination, contrast and rotation**.

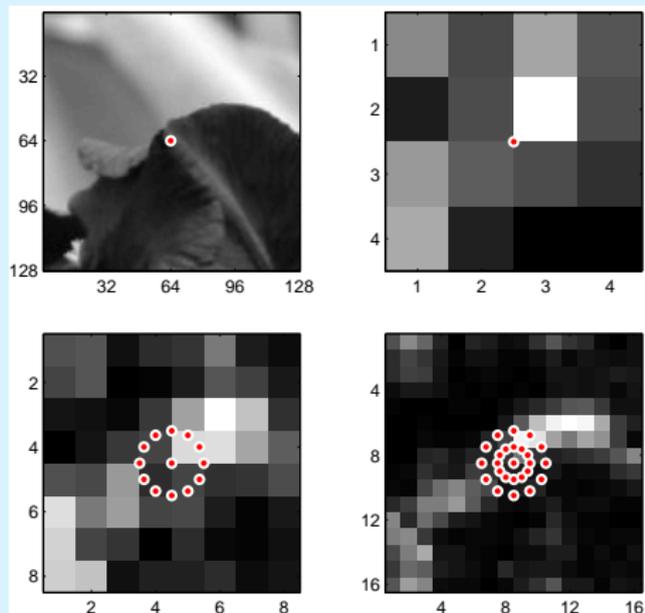
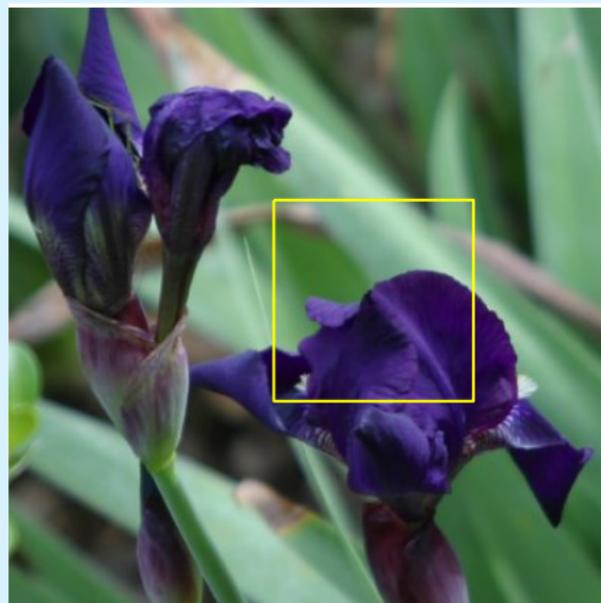
**BUT** it is still rather sensitive to discrepancies in **keypoint location and estimated dominant scale**.

To correct for small errors (typically a few pixels) in keypoint location, we modify the algorithm as follows:

- Measure **derivatives** of  $\bar{P}_k$  with respect to shifts  $\mathbf{x}$  in the sampling circle.
- Using the derivatives, calculate the shift vectors  $\mathbf{x}_i$  which maximise the normalised correlation measures  $\mathbf{s}_{k,i}$  at each of the 48 rotations  $i$  (using LMS methods with approximate adjustments for normalised vectors).
- By regarding the 48-point IFFT as a sparse matrix multiplication, the computation load is only **3 times** that of the basic algorithm.

We do the same for small scale errors using a derivative of  $\bar{P}_k$  wrt scale dilation,  $\psi$ , increasing computation to **4 times** that of the basic algorithm.

## A keypoint and its corresponding sampling circles over levels 5, 4 and 3

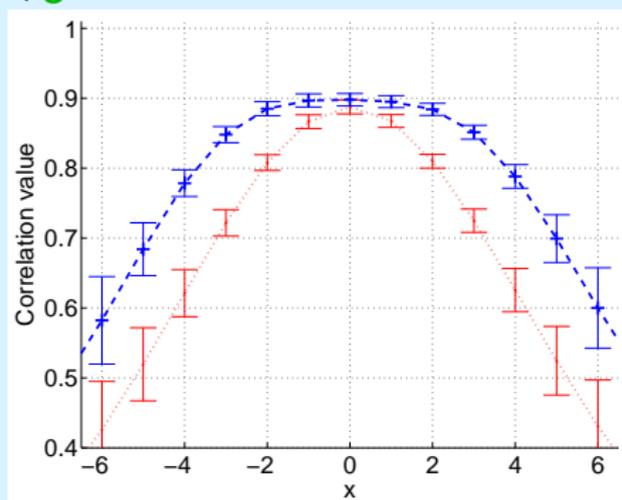
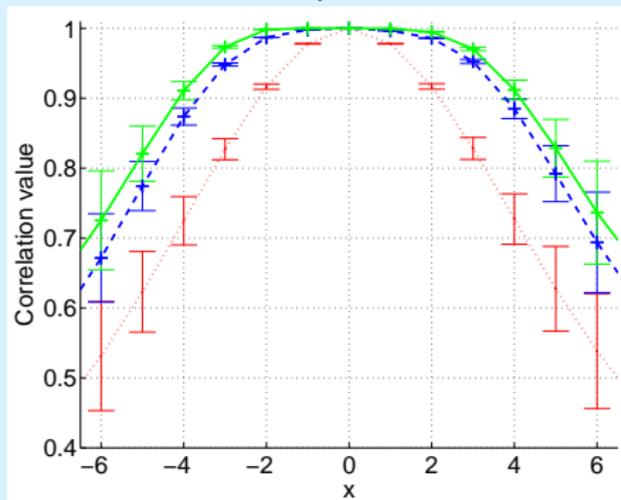


Background images show magnitudes of DT-CWT coefs for subband 1, oriented for edges at 15° above the horizontal.

## Mean correlation surfaces with no change of scale

Results averaged over 73 images from Caltech *PP Toys 03* dataset, using the centre of the image as the keypoint.

**red** – no tolerance; **blue** – shift tolerance; **green** – shift + scale tolerance.



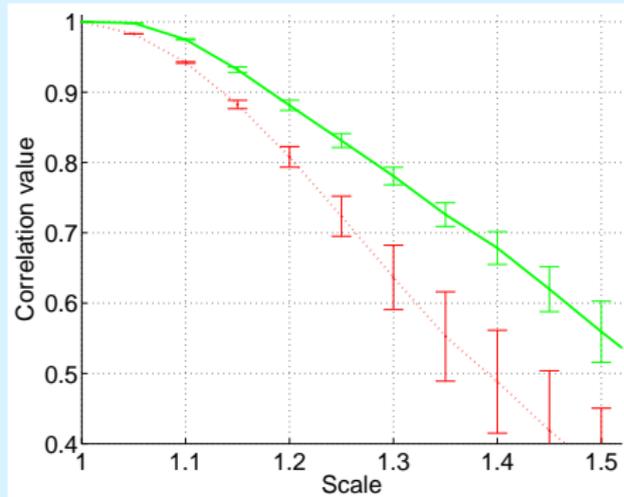
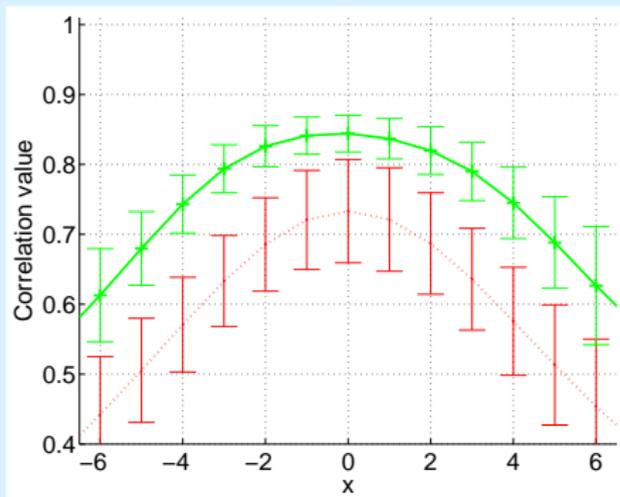
Left) No change of orientation;

Right) Average over rotations from  $0^\circ$  to  $90^\circ$  in steps of  $7.5^\circ$ .

## Mean correlation surfaces with change of scale (dilation)

Results averaged over 73 images from Caltech *PP Toys 03* dataset, using the centre of the image as the keypoint.

**red** – no tolerance; **green** – shift + scale tolerance.

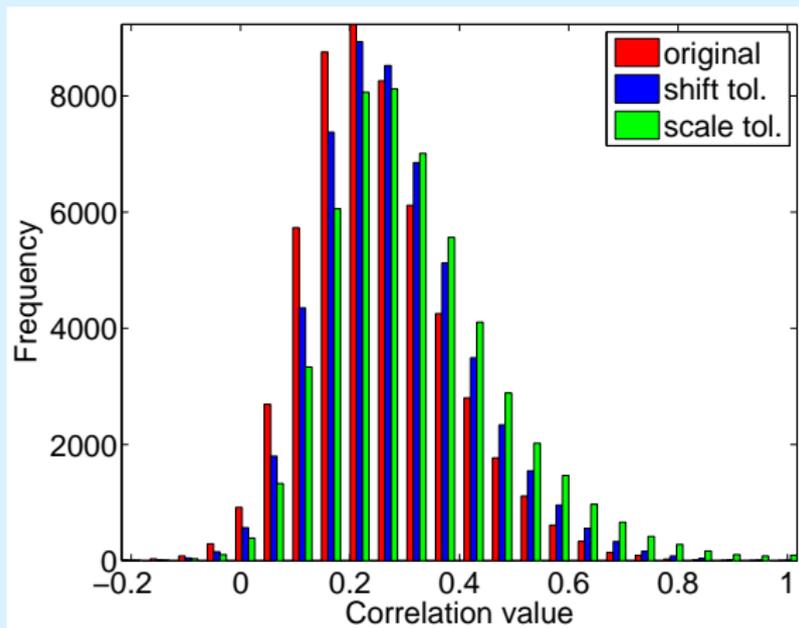


Left) Average over dilations from 1.0 to 1.4, with respect to shift;

Right) Mean correlation values with respect to scale change (dilation), no shift.

## Histograms of non-target correlation scores

**red** – no tolerance; **blue** – shift tolerance; **green** – shift + scale tolerance.



The histograms demonstrate the increase in false positive detection rate when shift and scale tolerances are introduced.

## Methods for handling colour and lighting changes

- For **colour keypoint detection**, we add the wavelet energies from the DT CWT of each R, G and B component, so that we get the total wavelet response to any changes in *RGB*-space; and use this to form a single Accumulated Map. Alternatively we could do this in a perceptually more uniform space such as *Lab*.
- We perform **contrast equalisation** over local image regions to handle variations between lighted / shaded parts of the image and between images.
- For **colour feature matching**, we form 3 polar-matching matrices into a single matrix  $[P_R \ P_G \ P_B]$  and perform matching with this.
- Prior to matching, we normalise each composite P-matrix with a **single scaling factor for all 3 colours**, so that relative amplitudes of colour variations are maintained, despite lighting changes.

## Conclusions

New local **Keypoint Detectors** and **Feature Descriptors** have been proposed for use when comparing the similarity of **detected objects** in images. They have the following properties:

- They are based on a modified form of the efficient **Dual-Tree Complex Wavelet Transform**. The modifications improve the rotational symmetry of the 6 directionally selective subbands at each scale.
- The **12-point circular sampling patterns** and the **Polar Matching Matrix  $P$**  are defined such that arbitrary rotations of an object can be accommodated efficiently in the matching process.
- **Shift-tolerant** and **scale-tolerant** extensions of the basic  $P$ -matrix method give greater robustness at the expense of some increase in false-positive rate.
- There is scope for applying **non-linear pre-processing** (e.g. magnitude compression and phase adjustment) to the complex wavelet coefficients prior to matching, to improve resilience to illumination and viewpoint changes.

Papers on complex wavelets and applications are available at:

[www.eng.cam.ac.uk/~ngk/](http://www.eng.cam.ac.uk/~ngk/)