Complex Wavelets:
What are they and what can they do?

Nick Kingsbury

Signal Processing and Communications Group, Dept. of Engineering
University of Cambridge, Cambridge CB2 1PZ, UK.
ngk@eng.cam.ac.uk
www.eng.cam.ac.uk/~ngk

June 2008

UNIVERSITY OF CAMBRIDGE

Complex Wavelets: What are they and what can they do?

- Basic form of the DT CWT
- Shift invariance of subband transfer functions
- DT CWT in 2-D – directional selectivity
- DT CWT in 3-D
- Denoising
- Image Registration
- Accumulated maps for keypoint detection
- Rotation-invariant local feature matching

Features of the (Real) Discrete Wavelet Transform (DWT)

- **Good compression** of signal energy.
- **Perfect reconstruction** with short support filters.
- **No redundancy**.
- **Very low computation** – order-N only.

But

- **Severe shift dependence**.
- **Poor directional selectivity** in 2-D, 3-D etc.

The DWT is normally implemented with a tree of highpass and lowpass filters, separated by $2:1$ decimators.

Real Discrete Wavelet Transform (DWT) in 1-D

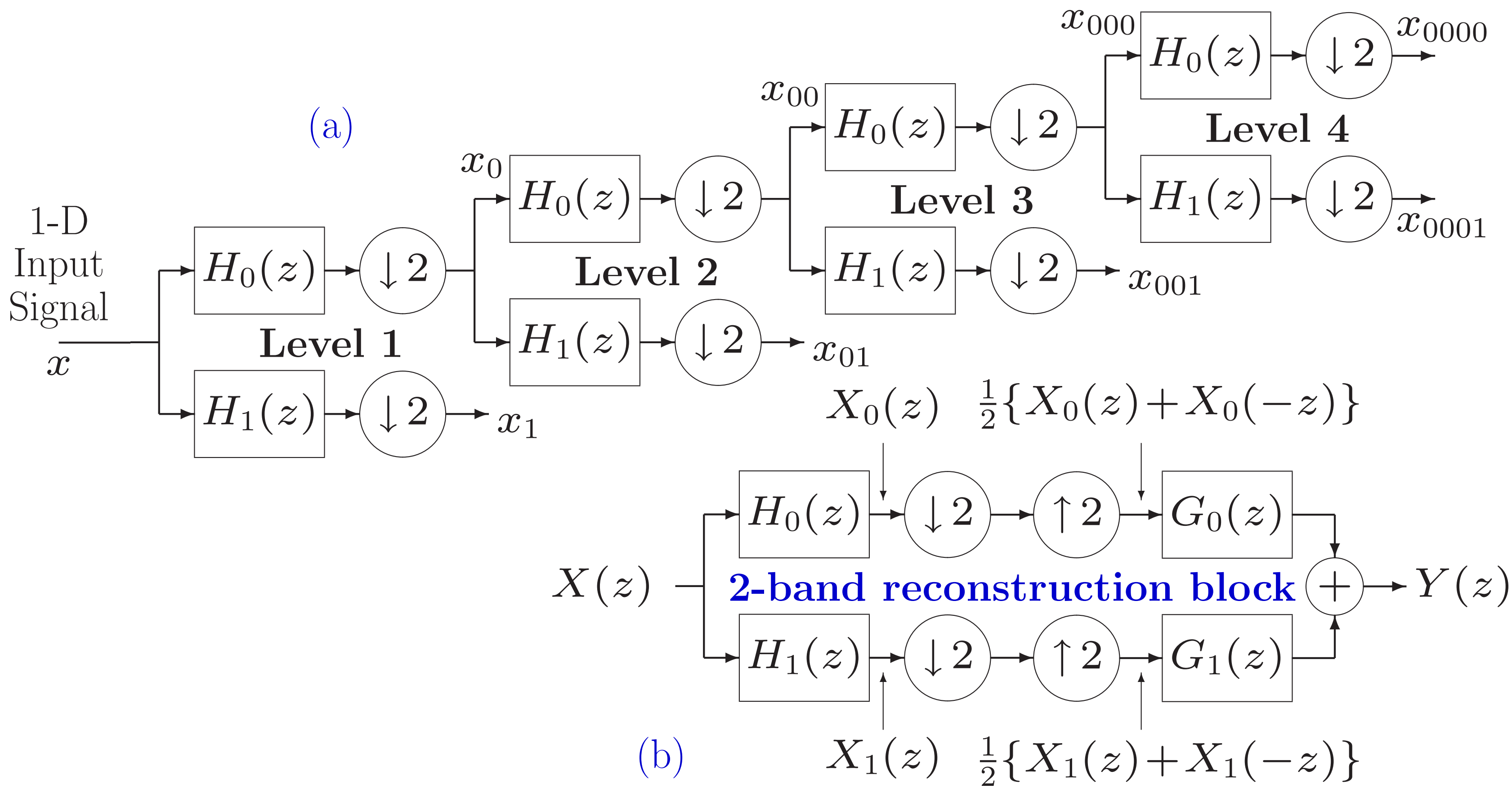


Figure 1: (a) Tree of real filters for the DWT. (b) Reconstruction filter block for 2 bands at a time, used in the inverse transform.

Visualising Shift Invariance

- Apply a standard input (e.g. unit step) to the transform for a **range of shift positions**.
- Select the transform coefficients from **just one wavelet level** at a time.
- Inverse transform each set of selected coefficients.
- Plot the component of the reconstructed output for each shift position at each wavelet level.
- Check for **shift invariance** (similarity of waveforms).

See Matlab demonstration.

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).

- **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.

Each tree contains purely real filters, but the two trees produce the **real and imaginary parts** respectively of each complex wavelet coefficient.

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

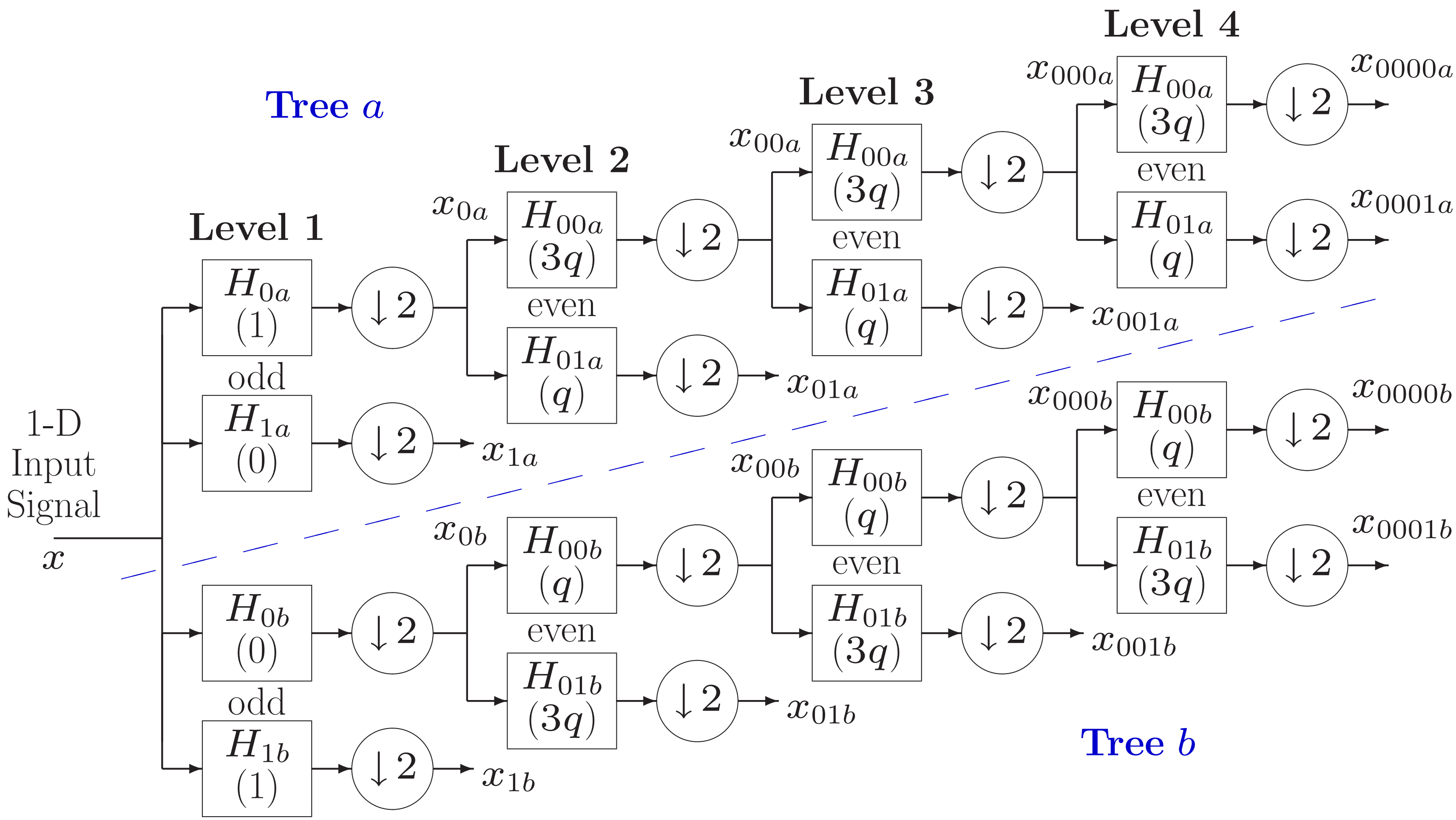


Figure 2: Dual tree of real filters for the Q-shift CWT, giving real and imaginary parts of complex coefficients from tree a and tree b respectively. Figures in brackets indicate the approximate delay for each filter, where $q = \frac{1}{4}$ sample period.

Features of the Q-shift Filters

Below level 1:

- Half-sample delay difference is obtained with filter delays of $\frac{1}{4}$ and $\frac{3}{4}$ of a sample period (instead of 0 and $\frac{1}{2}$ a sample for our original DT CWT).
- This is achieved with an **asymmetric even-length** filter $H(z)$ and its time reverse $H(z^{-1})$.
- Due to the asymmetry (like Daubechies filters), these may be designed to give an **orthonormal perfect reconstruction** wavelet transform.
- Tree **b** filters are the **reverse** of tree **a** filters, and reconstruction filters are the reverse of analysis filters, so **all filters** are from the **same orthonormal set**.
- Both trees have the **same frequency responses**.
- The combined **complex** impulse responses are **conjugate symmetric** about their mid points, even though the separate responses are asymmetric. Hence **symmetric extension** still works at image edges.

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

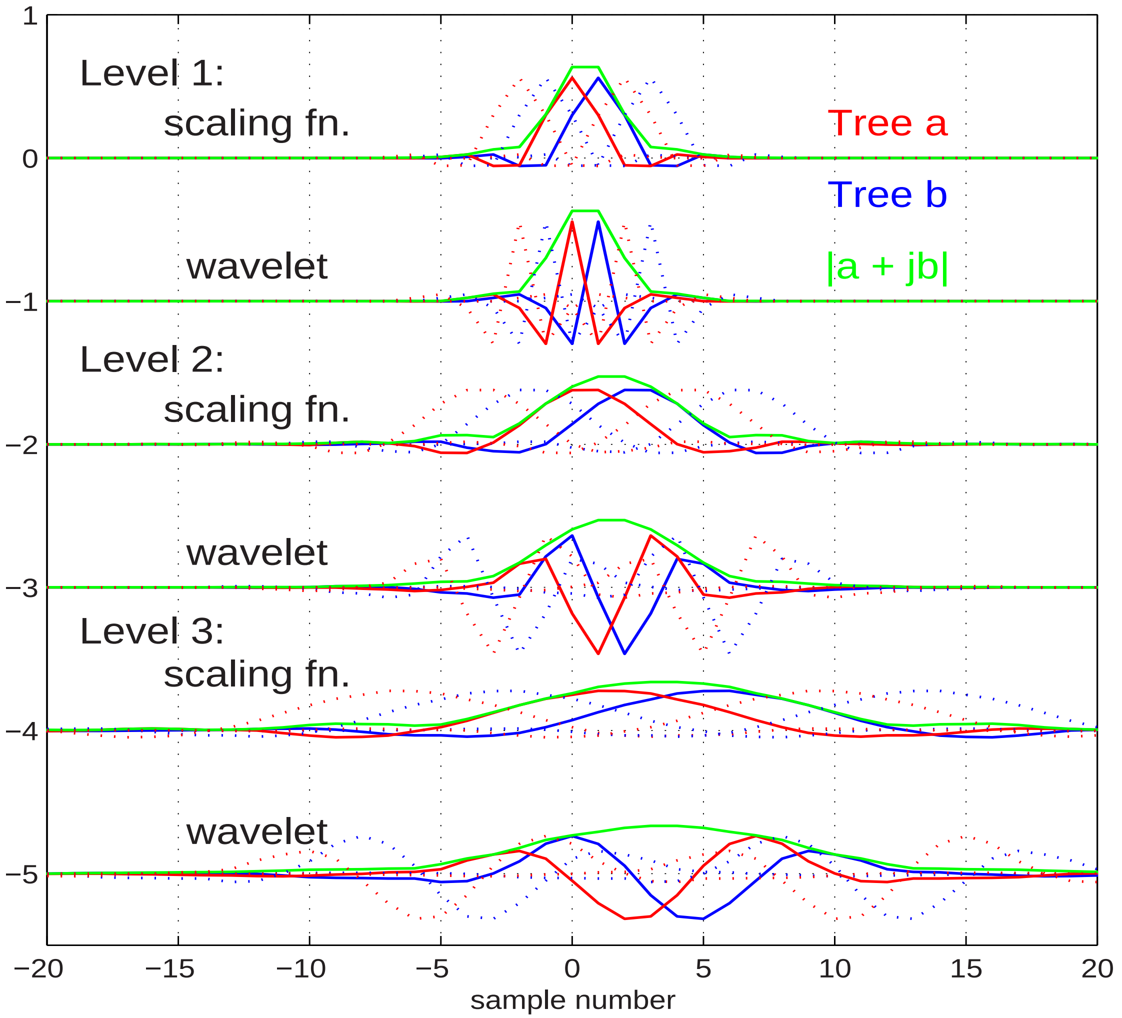
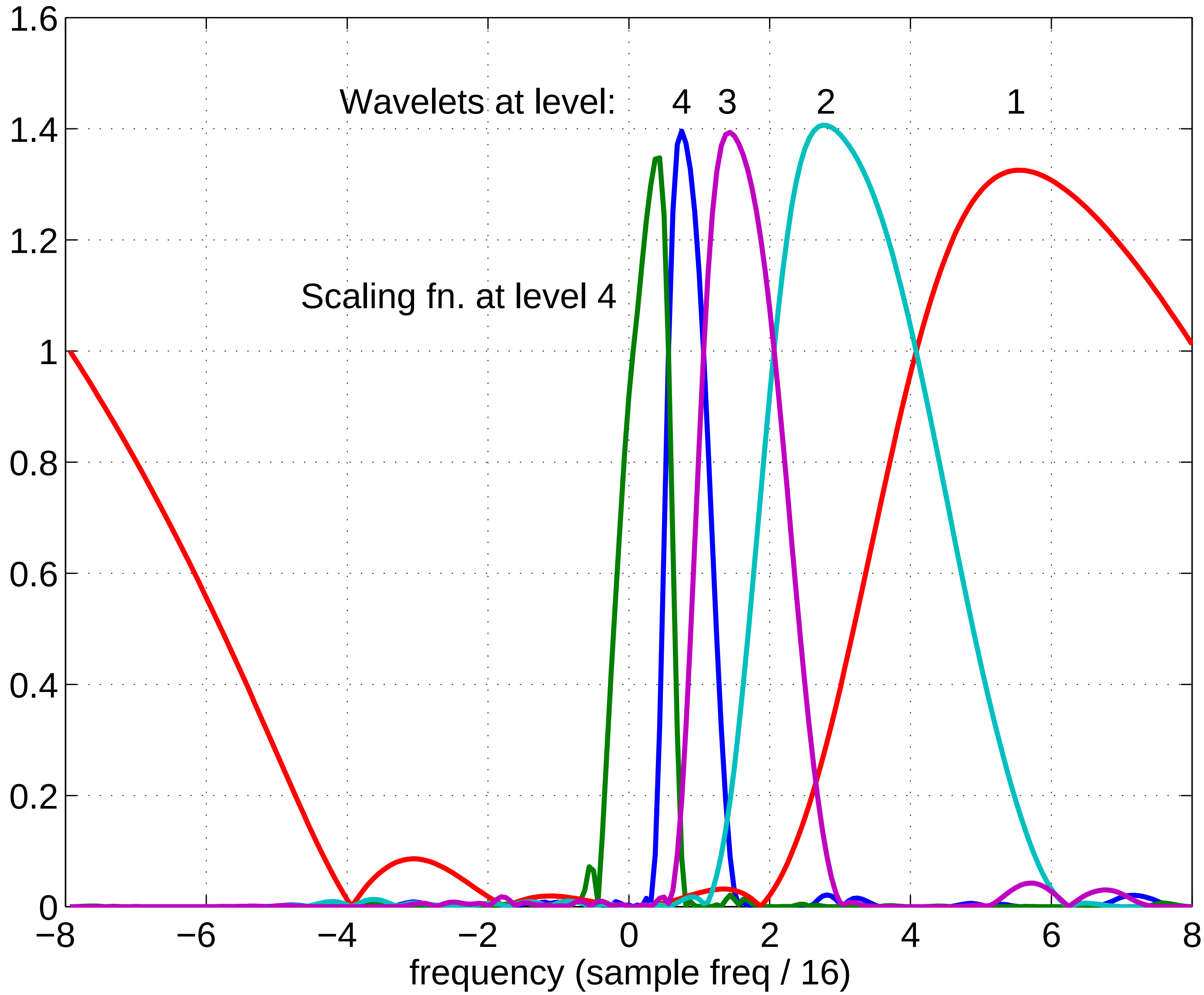
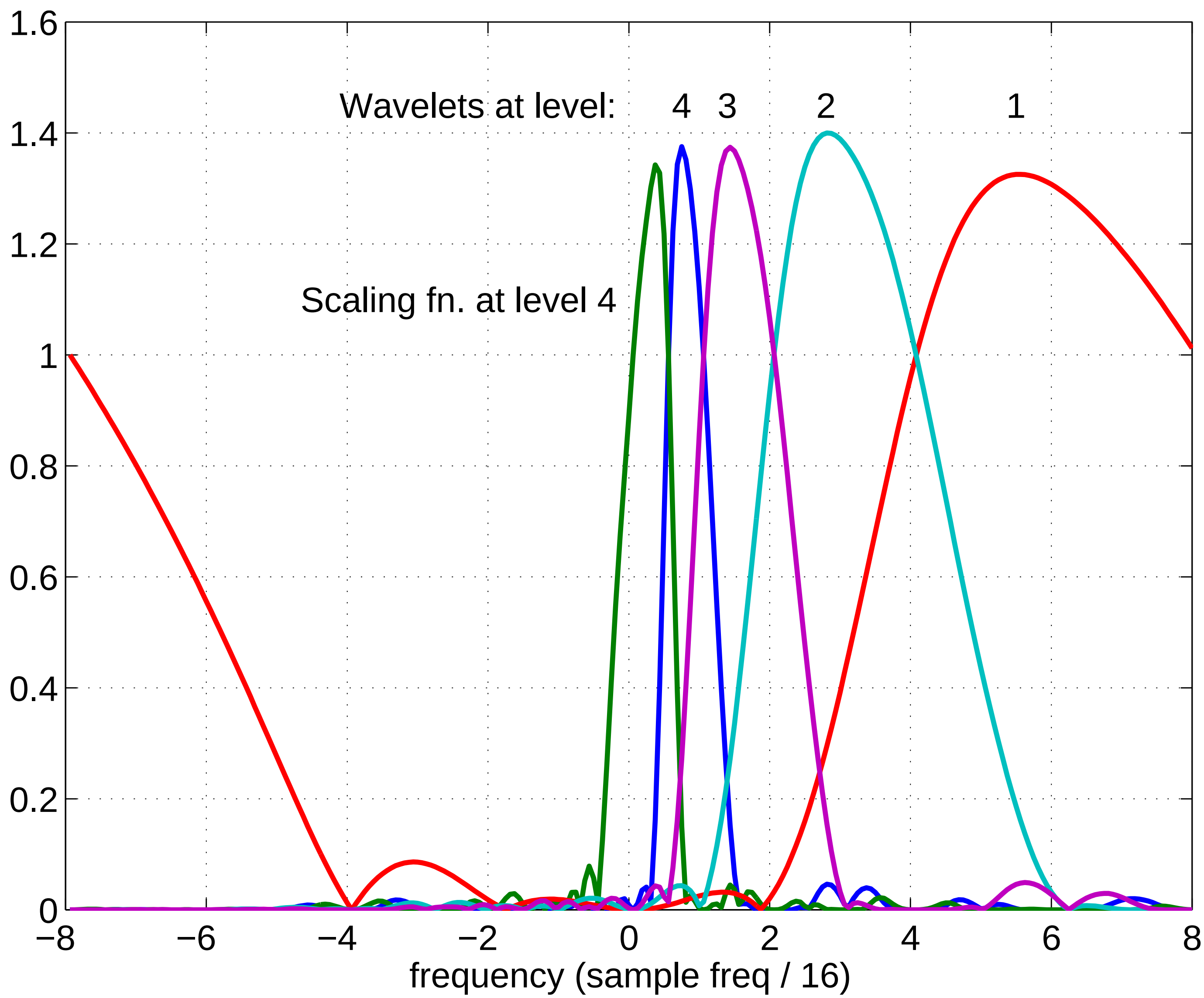


Figure 3: Basis functions for adjacent sampling points are shown dotted.

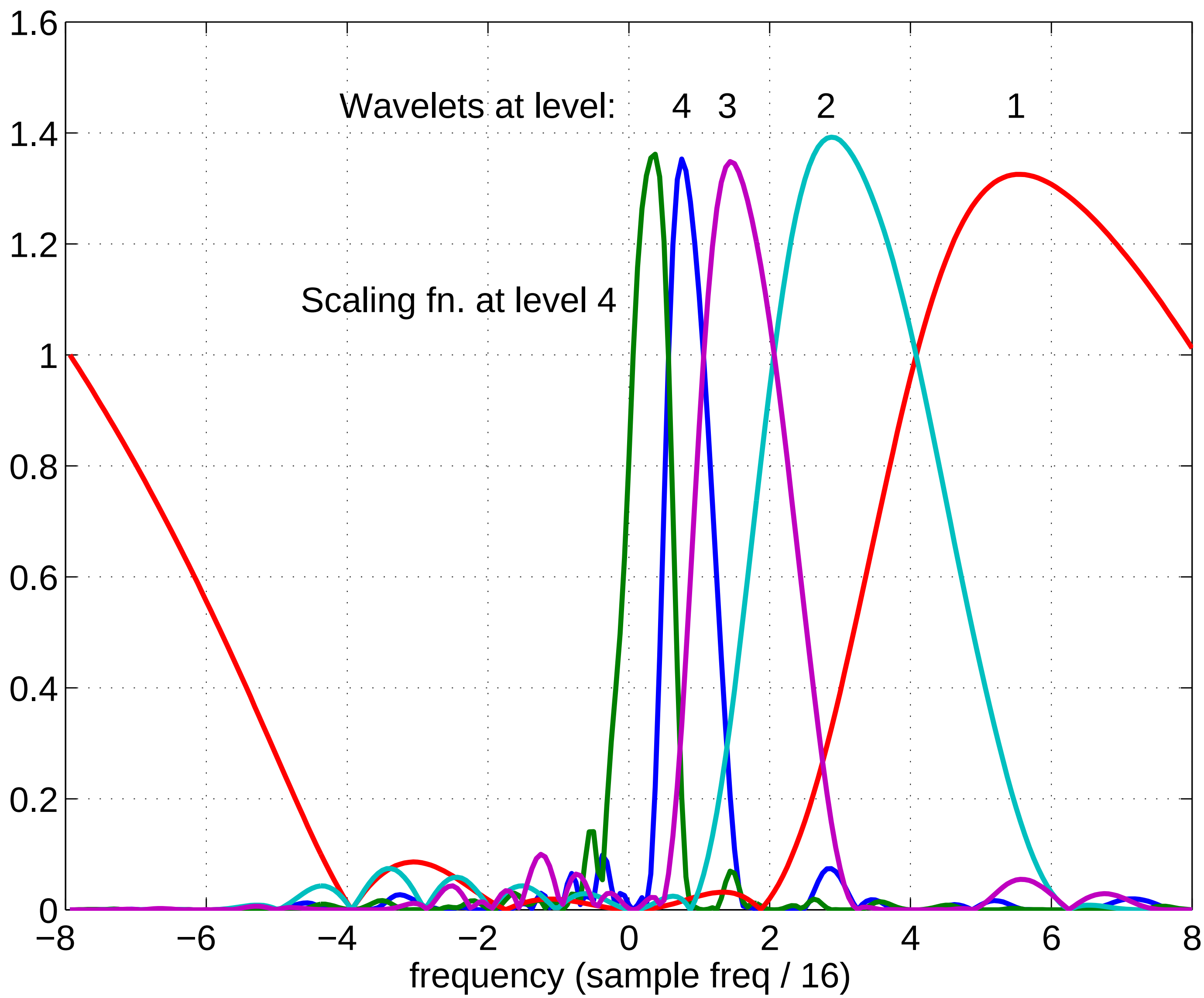
Frequency Responses of 18-tap Q-shift filters



Frequency Responses of 14-tap Q-shift filters



Frequency Responses of 6-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:

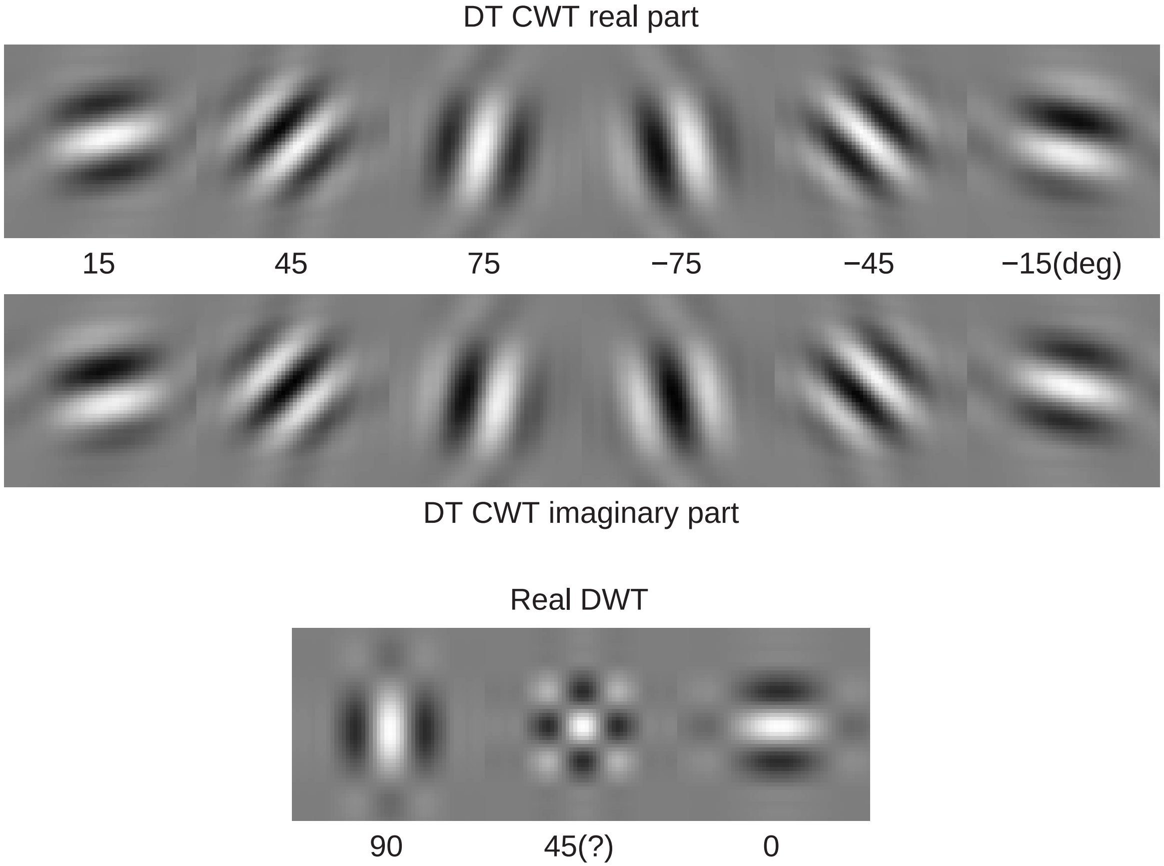
- It is performed separably, with 2 trees used for the rows of the image and 2 trees for the columns – yielding a **Quad-Tree** structure (4:1 redundancy).

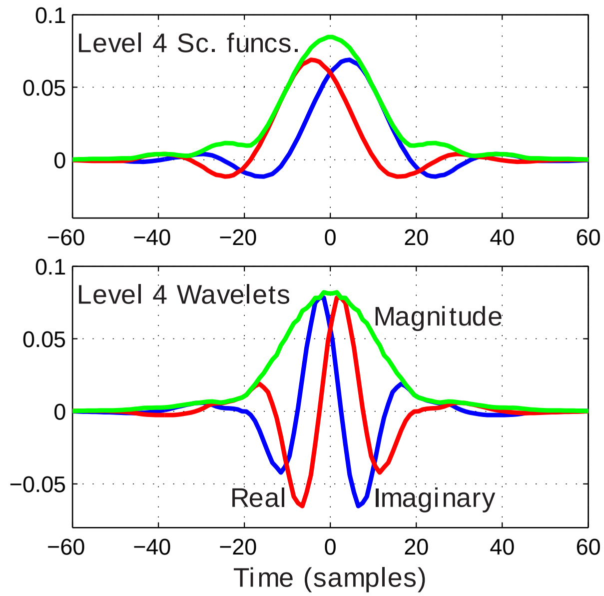
- The 4 quad-tree components of each coefficient are combined by simple sum and difference operations to yield a **pair of complex coefficients**. These are part of two separate subbands in adjacent quadrants of the 2-D spectrum.

- This produces **6 directionally selective subbands** at each level of the 2-D DT CWT. Fig 4 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 6) 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

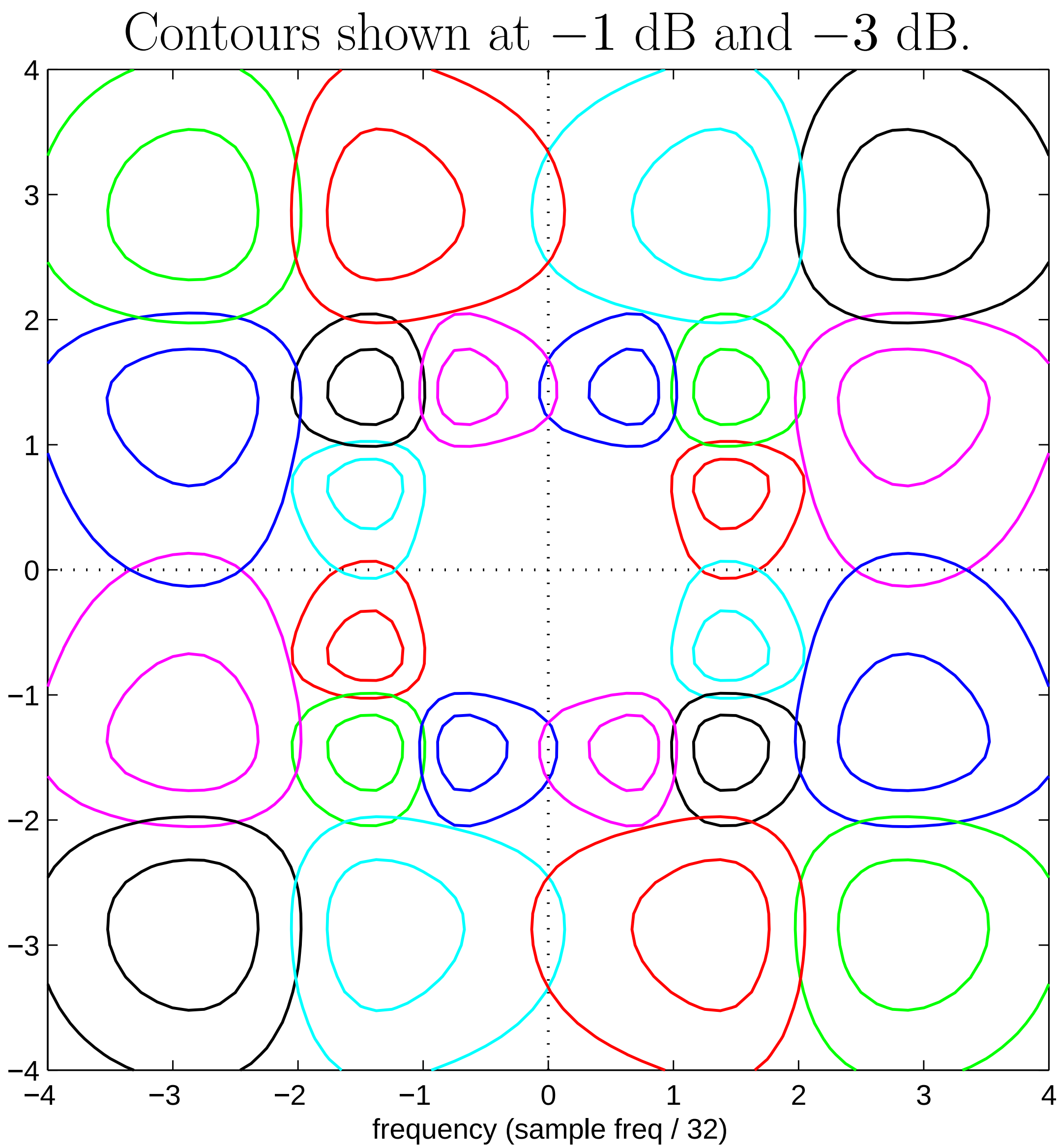




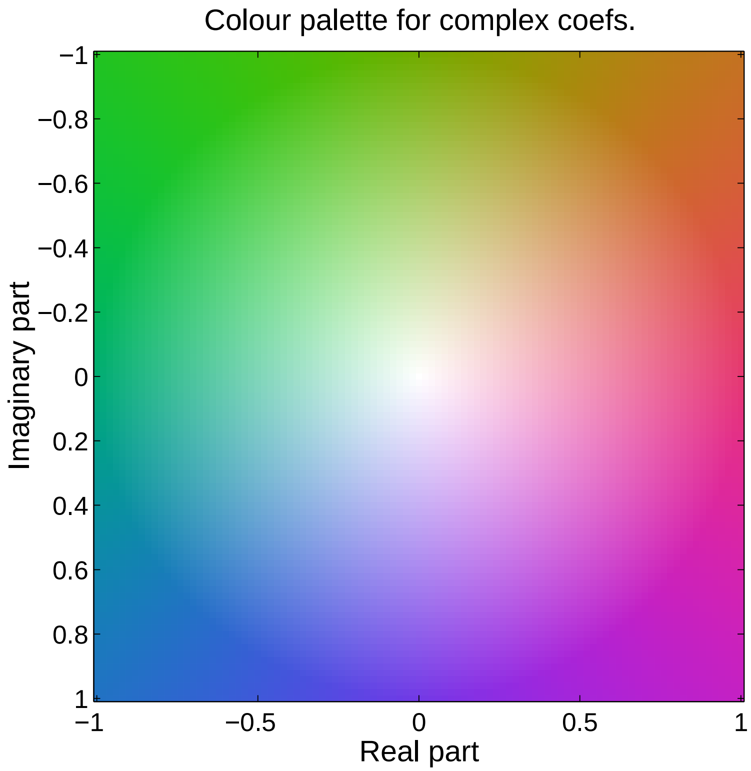
$$e^{j\omega_1 x}\ e^{j\omega_2 y} = e^{j(\omega_1 x + \omega_2 y)}$$

Figure 4: 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.

FREQUENCY RESPONSES OF 2-D Q-SHIFT FILTERS AT LEVELS 3 AND 4



Test Image and Colour Palette for Complex Coefficients



2-D DT-CWT Decomposition into Subbands

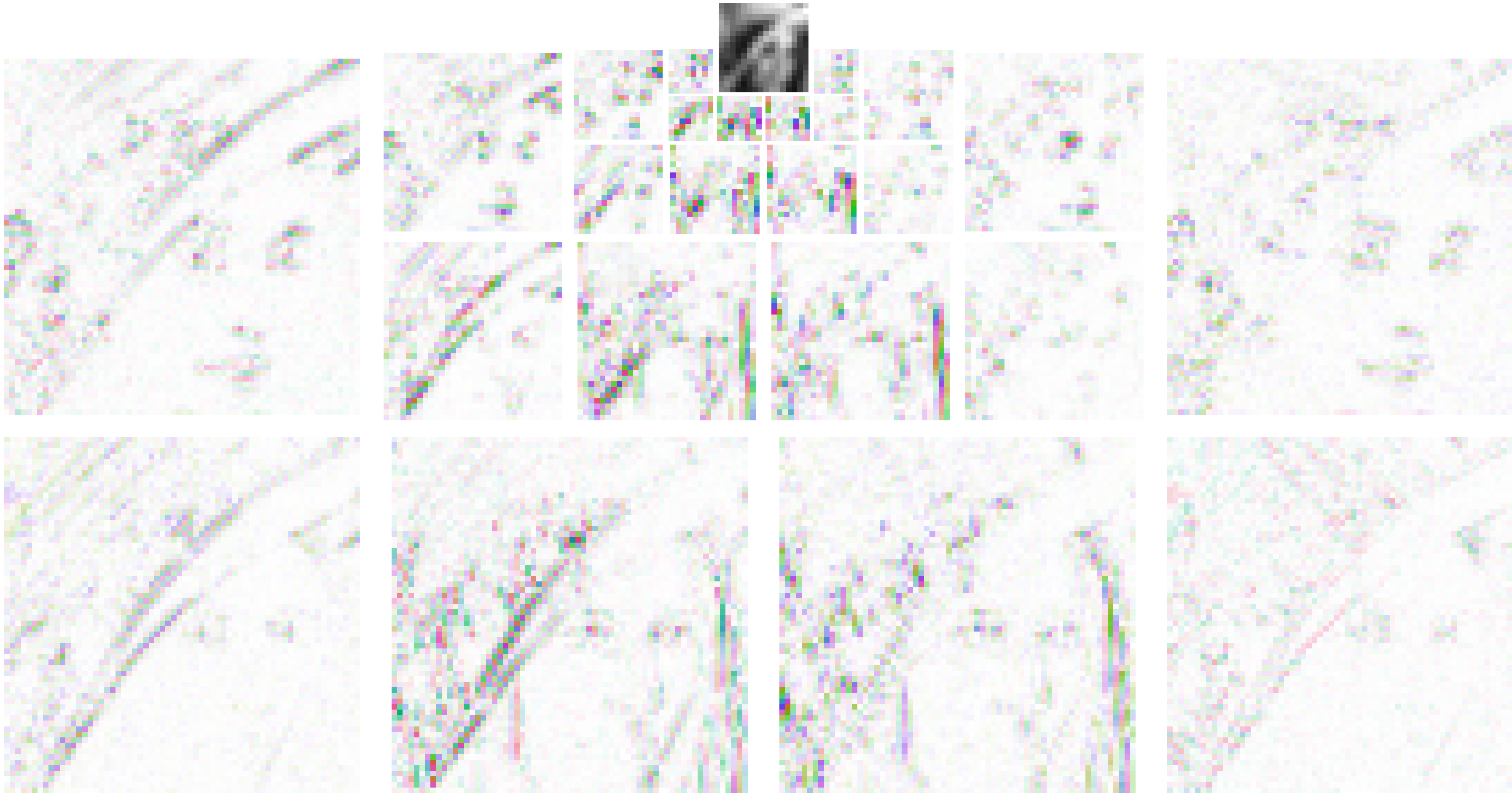


Figure 5: Four-level DT-CWT decomposition of *Lenna* into 6 subbands per level (only the central 128×128 portion of the image is shown for clarity). A colour-wheel palette is used to display the complex wavelet coefficients.

2-D DT-CWT Reconstruction Components from Each Subband

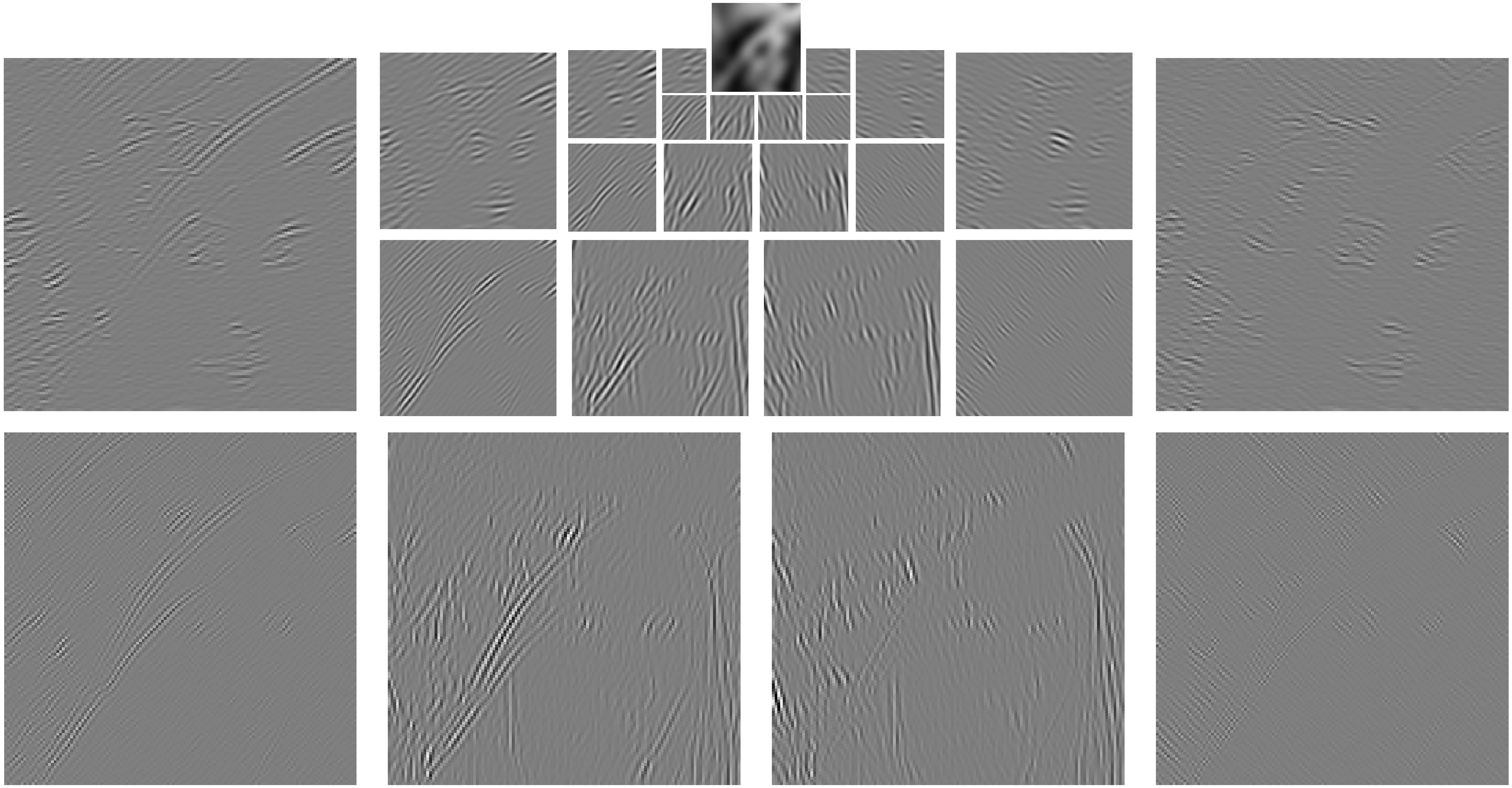


Figure 6: Components from each subband of the reconstructed output image for a 4-level DT-CWT decomposition of *Lenna* (central 128×128 portion only).

2-D Shift Invariance of DT CWT vs DWT

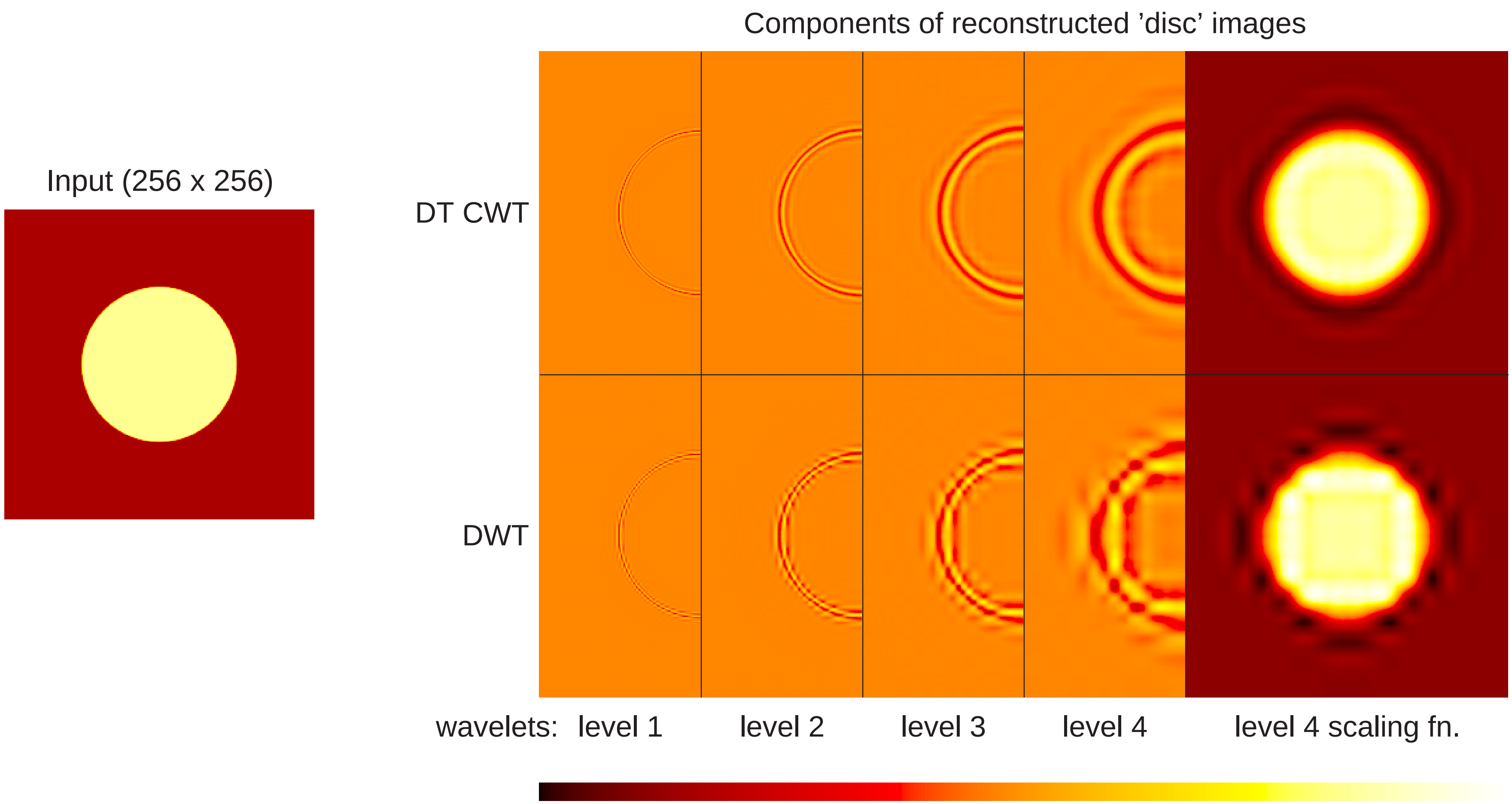


Figure 7: Wavelet and scaling function components at levels 1 to 4 of an image of a light circular disc on a dark background, using the 2-D DT CWT (upper row) and 2-D DWT (lower row). Only half of each wavelet image is shown in order to save space.

THE DT CWT IN 3-D

When the DT CWT is applied to 3-D signals (eg medical MRI or CT datasets), it has the following features:

- It is performed separably, with 2 trees used for the rows, 2 trees for the columns and 2 trees for the slices of the 3-D dataset – yielding an **Octal-Tree** structure (8:1 redundancy).

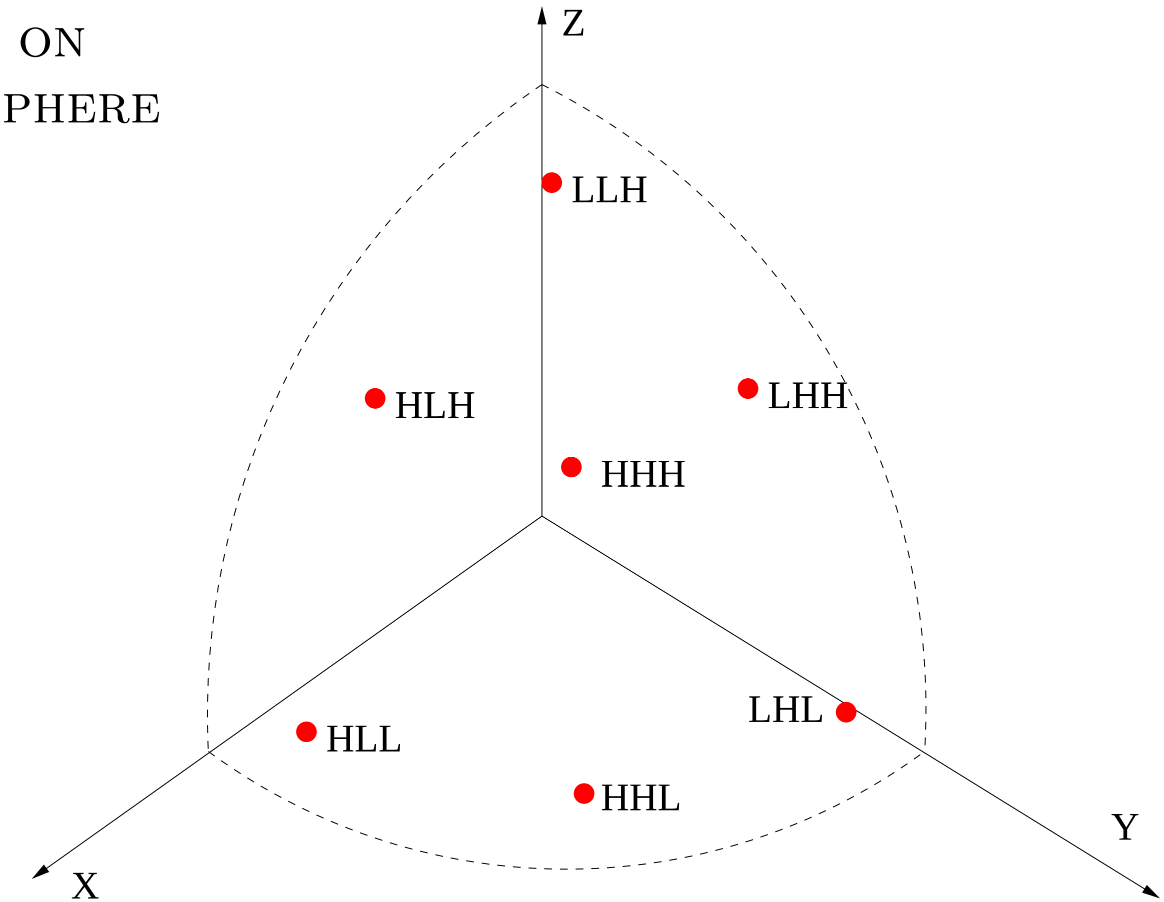
- The 8 octal-tree components of each coefficient are combined by simple sum and difference operations to yield a **quad of complex coefficients**. These are part of 4 separate subbands in adjacent octants of the 3-D spectrum.

- This produces **28 directionally selective subbands** $(4 \times 8 - 4)$ at each level of the 3-D DT CWT. The subband basis functions are now **planar waves** of the form $e^{j(\omega_1 x + \omega_2 y + \omega_3 z)}$, modulated by a 3-D Gaussian envelope.

- Each subband responds to approximately flat surfaces of a particular orientation. There are 7 orientations on each quadrant of a hemisphere.

3D subband orientations on one quadrant of a hemisphere

3D frequency domain:



3D Gabor-like basis functions:

$$h_{k1,k2,k3}(x,y,z) \simeq e^{-(x^2+y^2+z^2)/2\sigma^2} \times e^{j(\omega_{k1}\,x + \omega_{k2}\,y + \omega_{k3}\,z)}$$

These are **28 planar waves** (7 per quadrant of a hemisphere) whose orientation depends on $\omega_{k1} \in \{\omega_L, \omega_H\}$ and $\omega_{k2}, \omega_{k3} \in \{\pm\omega_L, \pm\omega_H\}$, where $\omega_H \simeq 3\omega_L$.

APPLICATIONS OF THE DT CWT

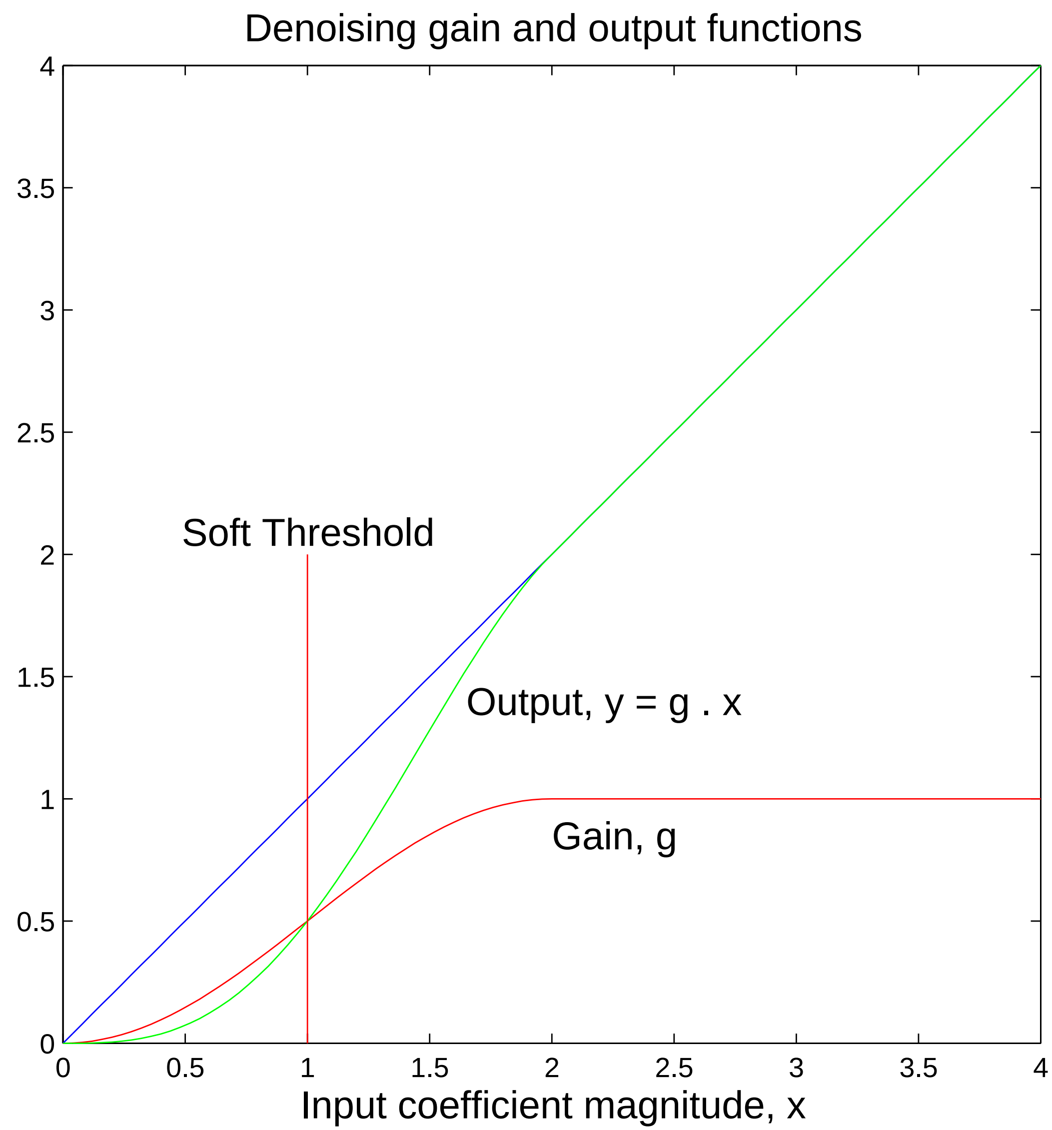
- **Motion estimation** [Magarey 98] and **compensation**
- **Registration** [Kingsbury 02]
- **Denoising** [Choi 00, Miller 06] and **deconvolution** [Jalobeanu 00, De Rivaz 01, J Ng 07]
- **Texture analysis** [Hatipoglu 99] and **synthesis** [De Rivaz 00]
- **Segmentation** [De Rivaz 00, Shaffrey 02]
- **Classification** [Romberg 00] and **image retrieval** [Kam & T T Ng 00, Shaffrey 03]
- **Watermarking of images** [Loo 00] and **video** [Earl 03]
- **Compression / Coding** [Reeves 03]
- **Seismic analysis** [van Spaendonck & Fernandes 02, Miller 05]
- **Diffusion Tensor MRI visualisation** [Zymnis 04]
- **Object matching & recognition** [Anderson & Fauqueur 06]

De-Noising – Method:

- Transform the noisy input image to **compress the image energy** into as few coefs as possible, leaving the noise well distributed.
- Suppress lower energy coefs (mainly noise).
- Inverse transform to recover de-noised image.

What is the Optimum Transform ?

- **DWT** is better than **DCT** or **DFT** for compressing image energy.
- But DWT is **shift dependent** – Is a coef small because there is no signal energy at that scale and location, **or** because it is sampled near a zero-crossing in the wavelet response?
- The **undecimated DWT** can solve this problem but at **significant cost** – redundancy (and computation) is increased by $3M : 1$, where M is no. of DWT levels.
- The **DT CWT** has only $4 : 1$ redundancy, is directionally selective, and works well.


Denoising gain and output functions
4
3.5
3
2.5
2
1.5
1
0.5
0
0
0.5
1
1.5
2
2.5
3
3.5
4
Soft Threshold
Output, y = g . x
Gain, g
Input coefficient magnitude, x

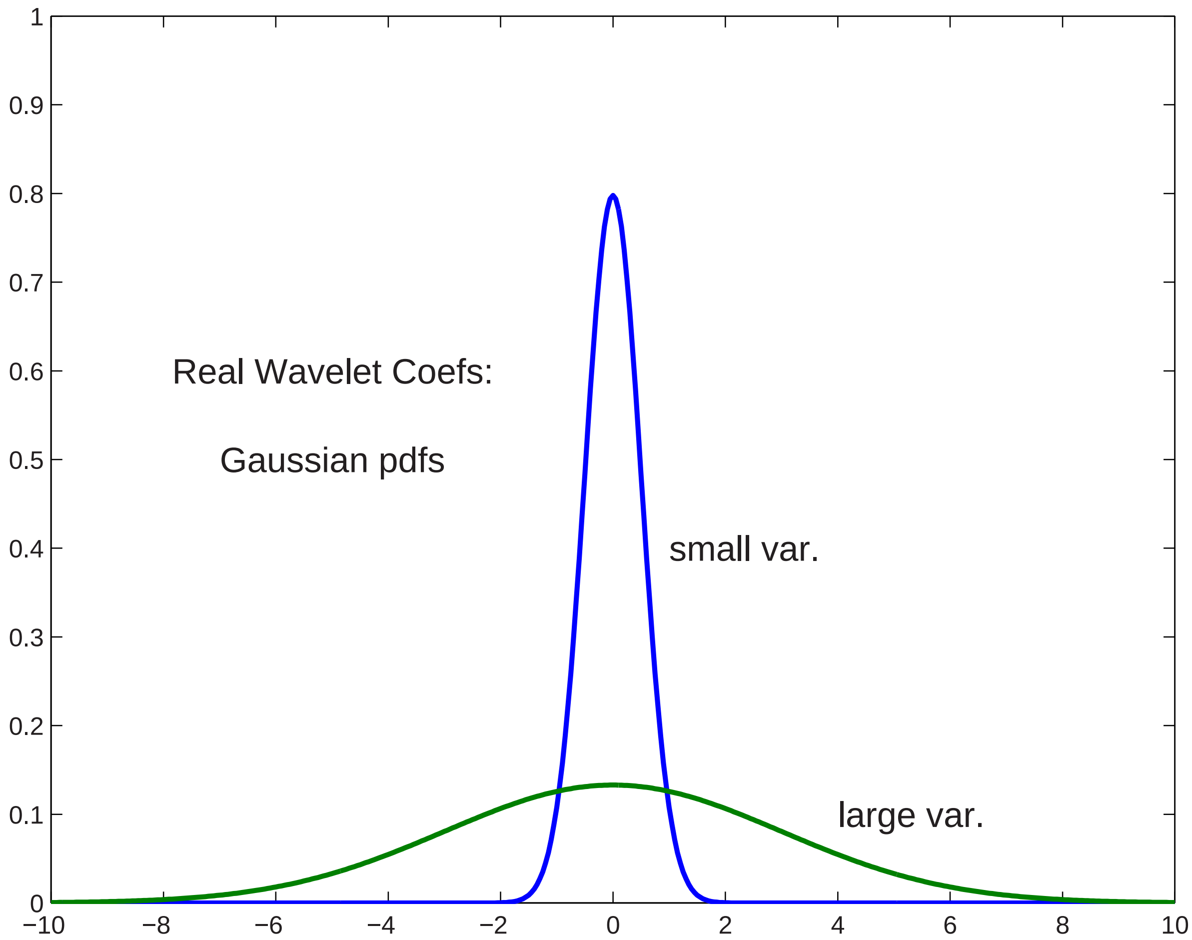


Figure 8: Probability density functions (pdfs) of small and large variance Gaussian distributions, typical for modelling **real and imaginary parts** of complex wavelet coefficients.

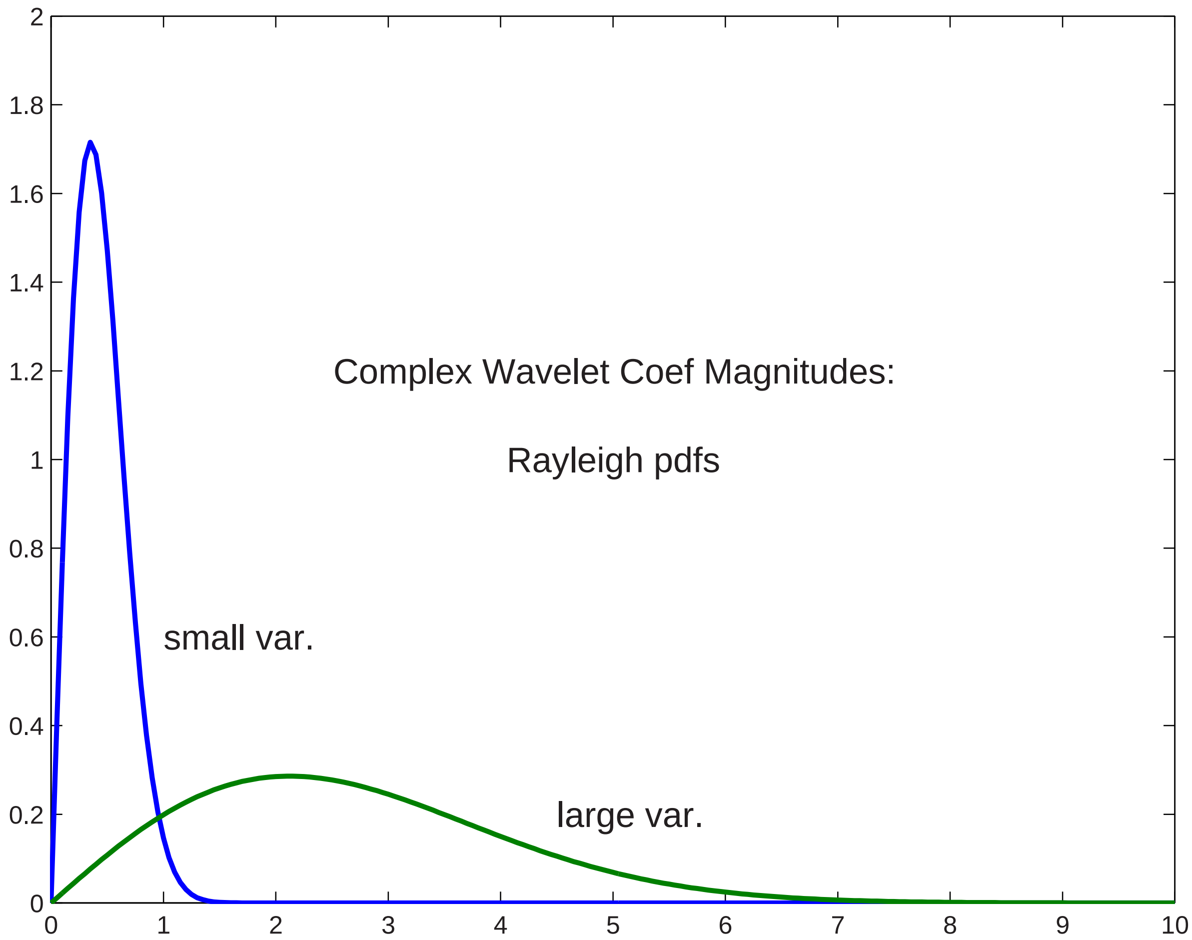


Figure 9: Probability density functions (pdfs) of small and large variance Rayleigh distributions, typical for modelling **magnitudes** of complex wavelet coefficients.

Image Denoising with different Wavelet Transforms - Lenna

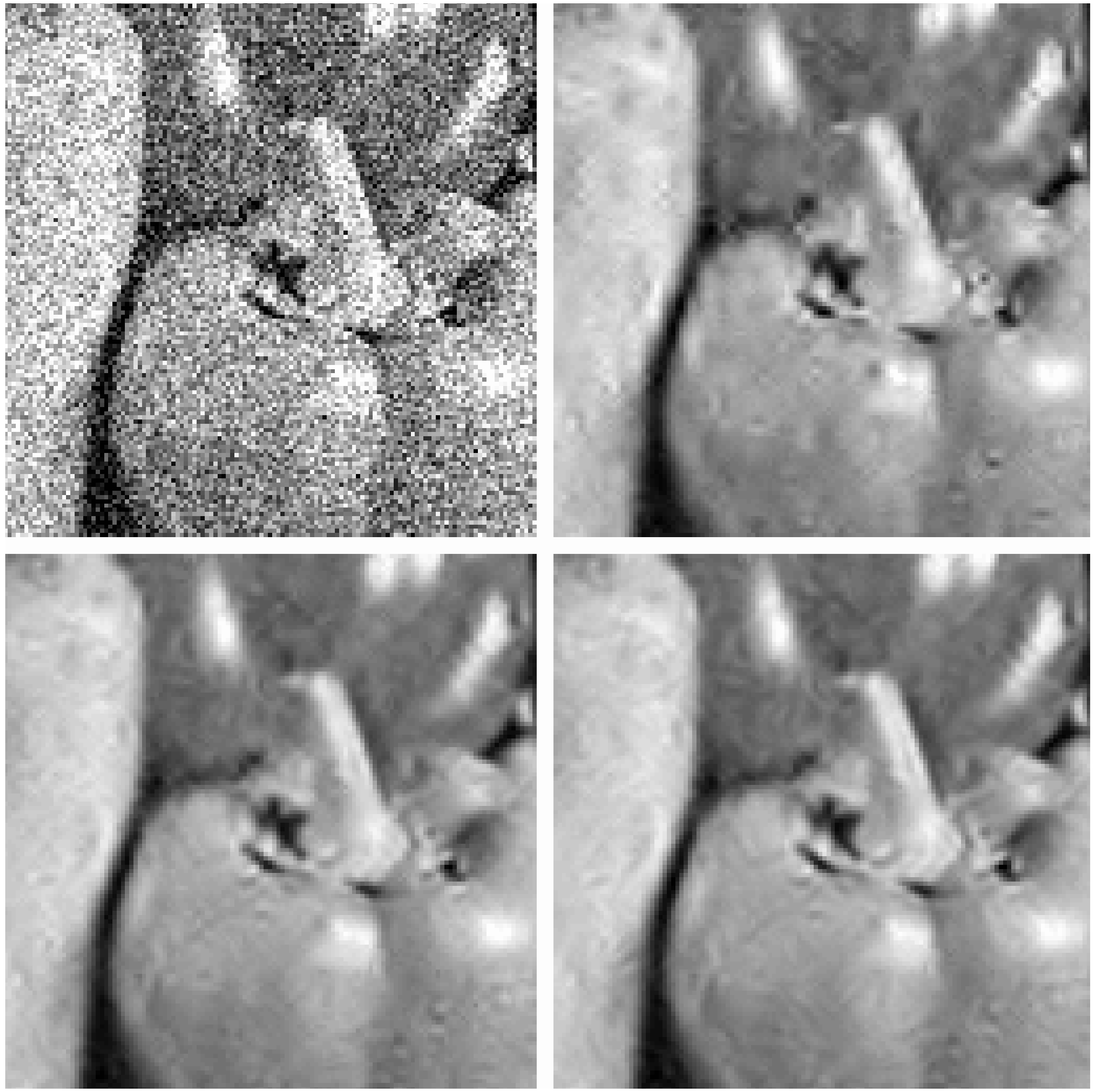
AWGN
SNR =
3.0 dB

Real DWT
SNR =
11.67 dB

Undec. WT
SNR =
12.82 dB

DT CWT
SNR =
12.99 dB

Image Denoising with different Wavelet Transforms - Peppers



AWGN
SNR =
3.0 dB

Real DWT
SNR =
12.24 dB

Undec. WT
SNR =
13.45 dB

DT CWT
SNR =
13.51 dB

Heirarchical Denoising with Gaussian Scale Mixtures (GSMs)

Non-heir.
DT CWT
SNR =
12.99 dB

Heirarchical
DT CWT
SNR =
13.51 dB

Non-heir.
DT CWT
SNR =
13.51 dB

Heirarchical
DT CWT
SNR =
13.85 dB

Denoising a 3-D dataset

e.g. Medical 3-D MRI or helical CT scans.

Method:

- Perform 3-D DT CWT on the dataset.
- Attenuate smaller coefficients, based on their magnitudes, as for 2-D denoising. (Heirarchical methods are also quite feasible.)
- Perform inverse 3-D DT CWT to recover the denoised dataset.

A Matlab example shows denoising of an ellipsoidal surface, buried in Gaussian white noise.

Image Registration

Key Features of Robust Registration Algorithms

- Edge-based methods are more robust than point-based ones.
- Must be automatic (no human picking of correspondence points) in order to achieve sub-pixel accuracy in noise.
- Bandlimited multiscale (wavelet) methods will allow spatially adaptive denoising.
- Phase-based bandpass methods can give rapid convergence and immunity to illumination changes between images.
- Displacement field should be smooth, so use of a wide-area parametric (affine) model is preferable to local translation-only models.

Selected Method

- Dual-tree Complex Wavelet Transform (DT CWT):
  - provides complex coefficients whose phase shift depends approximately linearly with displacement;
  - allows each subband of coefficients to be interpolated independently of other subbands (because of shift invariance).

- Parametric model of displacement field, whose solution is based on local edge-based motion constraints (Hemmendorf et al., IEEE Trans Medical Imaging, Dec 2002):
  - derives straight-line contraints from directional subbands of DT CWT;
  - solves for model parameters which minimise constraint error energy over multiple directions and scales.

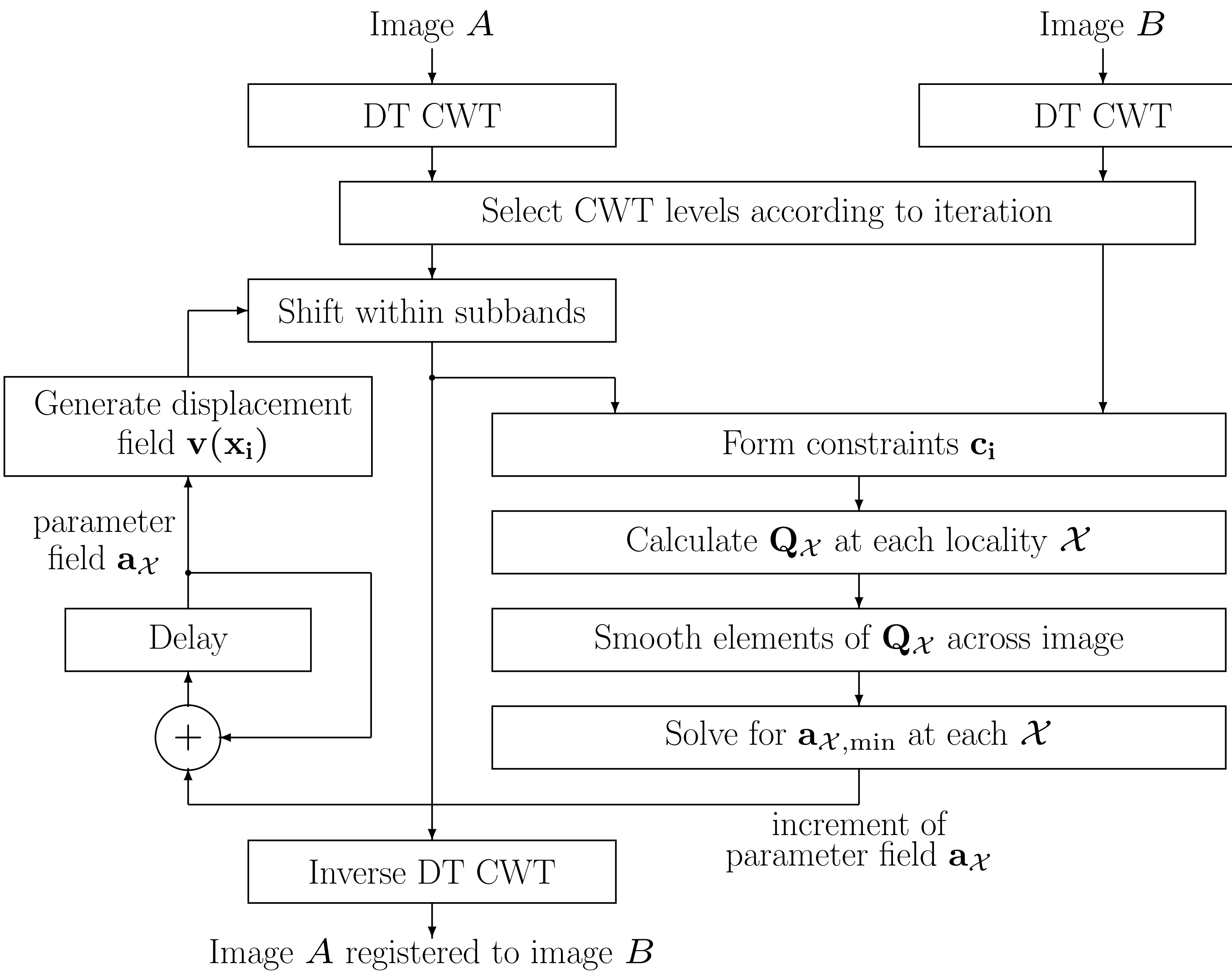


Image A
Image B
DT CWT
DT CWT
Select CWT levels according to iteration
Shift within subbands
Generate displacement field $\mathbf{v}(\mathbf{x_i})$
Form constraints $\mathbf{c_i}$
parameter field $\mathbf{a}_{\mathcal{X}}$
Calculate $\mathbf{Q}_{\mathcal{X}}$ at each locality $\mathcal{X}$
Delay
Smooth elements of $\mathbf{Q}_{\mathcal{X}}$ across image
Solve for $\mathbf{a}_{\mathcal{X},\mathrm{min}}$ at each $\mathcal{X}$
increment of parameter field $\mathbf{a}_{\mathcal{X}}$
Inverse DT CWT
Image A registered to image B

BASIC LINEAR FLOW MODEL

Key Assumption for local translation model:

- Time derivative of the phase θ of each complex wavelet coefficient depends **approximately linearly** on the local velocity vector $\mathbf{v}$.

This can be expressed as a flow equation in time and spatial derivatives:

$$\frac{\partial \theta}{\partial t} = \nabla_{\mathbf{x}}\, \theta \,.\, \mathbf{v}$$

We can rearrange this to be in the form:

$$\nabla_{\mathbf{x}}\, \theta \,.\, \mathbf{v} - \frac{\partial \theta}{\partial t} = 0$$

or

$$\begin{bmatrix} \nabla_{\mathbf{x}}\, \theta \\ -\frac{\partial \theta}{\partial t} \end{bmatrix}^T \tilde{\mathbf{v}} = 0 \qquad \text{where} \qquad \tilde{\mathbf{v}} = \begin{bmatrix} \mathbf{v} \\ 1 \end{bmatrix}$$

Parametric Model: Constraint equations

Let the displacement vector at the i^{th} location $\mathbf{x}_i$ be $\mathbf{v}(\mathbf{x}_i)$; and let $\tilde{\mathbf{v}}_i = \begin{bmatrix} \mathbf{v}(\mathbf{x}_i) \\ 1 \end{bmatrix}$.

A straight-line constraint on $\mathbf{v}(\mathbf{x}_i)$ can be written

$$\mathbf{c}_i^T \, \tilde{\mathbf{v}}_i = 0 \quad \text{or} \quad c_{1,i}v_{1,i} + c_{2,i}v_{2,i} + c_{3,i} = 0$$

For a phase-based system in which wavelet coefficients at $\mathbf{x}_i$ in images A and B have phases θ_A and θ_B, approximate phase linearity means that

$$\mathbf{c}_i = C_i \begin{bmatrix} \nabla_{\mathbf{x}} \, \theta(\mathbf{x}_i) \\ \theta_B(\mathbf{x}_i) - \theta_A(\mathbf{x}_i) \end{bmatrix}$$

In practise we compute this by averaging finite differences at the centre of a $2 \times 2 \times 2$ block of coefficients from images A and B.

C_i is a constant which does not affect the line defined by the constraint, but which is important later.

Parameters of the Model

We can define an affine parametric model for $\mathbf{v}$ such that

$$\mathbf{v}(\mathbf{x}) = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} a_3 & a_5 \\ a_4 & a_6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

or in a more useful form

$$\mathbf{v}(\mathbf{x}) = \begin{bmatrix} 1 & 0 & x_1 & 0 & x_2 & 0 \\ 0 & 1 & 0 & x_1 & 0 & x_2 \end{bmatrix} . \begin{bmatrix} a_1 \\ \vdots \\ a_6 \end{bmatrix} = \mathbf{K}(\mathbf{x}) \ . \ \mathbf{a}$$

Affine models can synthesise translation, rotation, constant zoom, and shear.

A quadratic model, which allows for linearly changing zoom (approx perspective), requires up to 6 additional parameters and columns in $\mathbf{K}$ of the form

$$\begin{bmatrix} \ldots & x_1 x_2 & 0 & x_1^2 & 0 & x_2^2 & 0 \\ \ldots & 0 & x_1 x_2 & 0 & x_1^2 & 0 & x_2^2 \end{bmatrix}$$

Solving for the Model Parameters

Let $\tilde{\mathbf{K}}_i = \begin{bmatrix} \mathbf{K}(\mathbf{x}_i) & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix}$ and $\tilde{\mathbf{a}} = \begin{bmatrix} \mathbf{a} \\ 1 \end{bmatrix}$ so that $\tilde{\mathbf{v}}_i = \tilde{\mathbf{K}}_i \, \tilde{\mathbf{a}}$.

Ideally for a given image locality $\mathcal{X}$, we wish to find the parametric vector $\tilde{\mathbf{a}}$ such that

$$\mathbf{c}_i^T \, \tilde{\mathbf{v}}_i = 0 \quad \text{when} \quad \tilde{\mathbf{v}}_i = \tilde{\mathbf{K}}_i \, \tilde{\mathbf{a}} \quad \text{for all } i \text{ such that } \mathbf{x}_i \in \mathcal{X}.$$

In practise this is an overdetermined set of equations, so we find the LMS solution, the value of $\mathbf{a}$ which minimises the squared error

$$\begin{aligned} \mathcal{E}_{\mathcal{X}} &= \sum_{i \in \mathcal{X}} ||\mathbf{c}_i^T \, \tilde{\mathbf{v}}_i||^2 = \sum_{i \in \mathcal{X}} ||\mathbf{c}_i^T \, \tilde{\mathbf{K}}_i \, \tilde{\mathbf{a}}||^2 = \sum_{i \in \mathcal{X}} (\tilde{\mathbf{a}}^T \, \tilde{\mathbf{K}}_i^T \, \mathbf{c}_i)(\mathbf{c}_i^T \, \tilde{\mathbf{K}}_i \, \tilde{\mathbf{a}}) \\ &= \tilde{\mathbf{a}}^T \, \tilde{\mathbf{Q}}_{\mathcal{X}} \, \tilde{\mathbf{a}} \qquad \text{where } \tilde{\mathbf{Q}}_{\mathcal{X}} = \sum_{i \in \mathcal{X}} (\tilde{\mathbf{K}}_i^T \, \mathbf{c}_i \, \mathbf{c}_i^T \, \tilde{\mathbf{K}}_i) \end{aligned}$$

Solving for the Model Parameters (cont.)

Since $\tilde{\mathbf{a}} = \begin{bmatrix} \mathbf{a} \\ 1 \end{bmatrix}$ and $\tilde{\mathbf{Q}}_{\mathcal{X}}$ is symmetric, we define $\tilde{\mathbf{Q}}_{\mathcal{X}} = \begin{bmatrix} \mathbf{Q} & \mathbf{q} \\ \mathbf{q}^T & q_0 \end{bmatrix}_{\mathcal{X}}$ so that

$$\mathcal{E}_{\mathcal{X}} = \tilde{\mathbf{a}}^T \; \tilde{\mathbf{Q}}_{\mathcal{X}} \; \tilde{\mathbf{a}} = \mathbf{a}^T \; \mathbf{Q} \; \mathbf{a} + 2 \; \mathbf{a}^T \mathbf{q} + q_0$$

$\mathcal{E}_{\mathcal{X}}$ is minimised when $\nabla_{\mathbf{a}} \, \mathcal{E}_{\mathcal{X}} = 2 \; \mathbf{Q} \; \mathbf{a} + 2 \; \mathbf{q} = \mathbf{0}$, so $\mathbf{a}_{\mathcal{X},\min} = - \; \mathbf{Q}^{-1} \; \mathbf{q}$.

The choice of locality $\mathcal{X}$ will depend on application:

- If it is expected that the affine (or quadratic) model will apply accurately to the whole image, then $\mathcal{X}$ can be the whole image and maximum robustness will be achieved.

- If not, then $\mathcal{X}$ should be a smaller region, chosen to optimise the tradeoff between robustness and model accuracy. A good way to produce a smooth field is to make $\mathcal{X}$ fairly small (e.g. a 32×32 pel region) and then to apply a smoothing filter across all the $\tilde{\mathbf{Q}}_{\mathcal{X}}$ matrices, element by element, before solving for $\mathbf{a}_{\mathcal{X},\min}$ in each region.

CONSTRAINT WEIGHTING FACTORS

Returning to the equation for the constraint vectors, $\mathbf{c}_i = C_i \begin{bmatrix} \nabla_{\mathbf{x}}\, \theta(\mathbf{x}_i) \\ \theta_B(\mathbf{x}_i) - \theta_A(\mathbf{x}_i) \end{bmatrix}$,

the constant gain parameter C_i will determine how much weight is given to each constraint in $\tilde{\mathbf{Q}}_{\mathcal{X}} = \sum_{i \in \mathcal{X}} (\tilde{\mathbf{K}}_i^T\, \mathbf{c}_i\, \mathbf{c}_i^T\, \tilde{\mathbf{K}}_i)$.

Hemmendorf proposes some quite complicated heuristics for computing C_i, but for the DT CWT, we find the following works well:

$$C_i = \frac{|d_{AB}|^2}{\sum_{k=1}^{4} |u_k|^3 + |v_k|^3} \quad \text{where} \quad d_{AB} = \sum_{k=1}^{4} u_k^*\, v_k$$

and $\begin{bmatrix} u_1 & u_2 \\ u_3 & u_4 \end{bmatrix}$ and $\begin{bmatrix} v_1 & v_2 \\ v_3 & v_4 \end{bmatrix}$ are 2×2 blocks of wavelet coefficients centred on $\mathbf{x}_i$ in images A and B respectively.

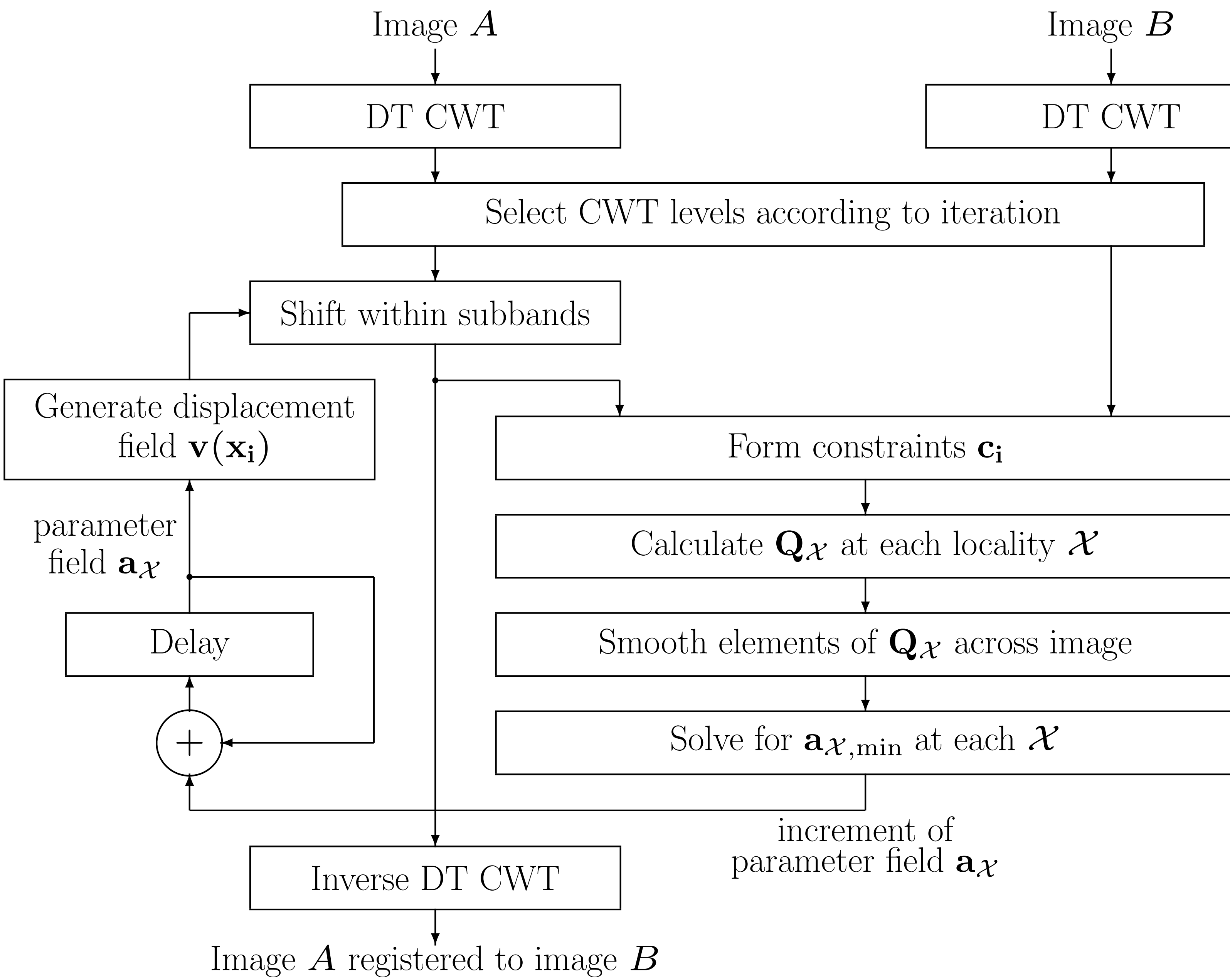


Image A
Image B
DT CWT
DT CWT
Select CWT levels according to iteration
Shift within subbands
Generate displacement field $\mathbf{v}(\mathbf{x_i})$
Form constraints $\mathbf{c_i}$
parameter field $\mathbf{a}_{\mathcal{X}}$
Calculate $\mathbf{Q}_{\mathcal{X}}$ at each locality $\mathcal{X}$
Smooth elements of $\mathbf{Q}_{\mathcal{X}}$ across image
Delay
Solve for $\mathbf{a}_{\mathcal{X},\mathrm{min}}$ at each $\mathcal{X}$
increment of parameter field $\mathbf{a}_{\mathcal{X}}$
Inverse DT CWT
Image A registered to image B

Demonstration of Registration and Image Fusion

- House on a hillside, viewed on a video camera with telephoto lens through air with significant heat turbulence (due to a hot runway).

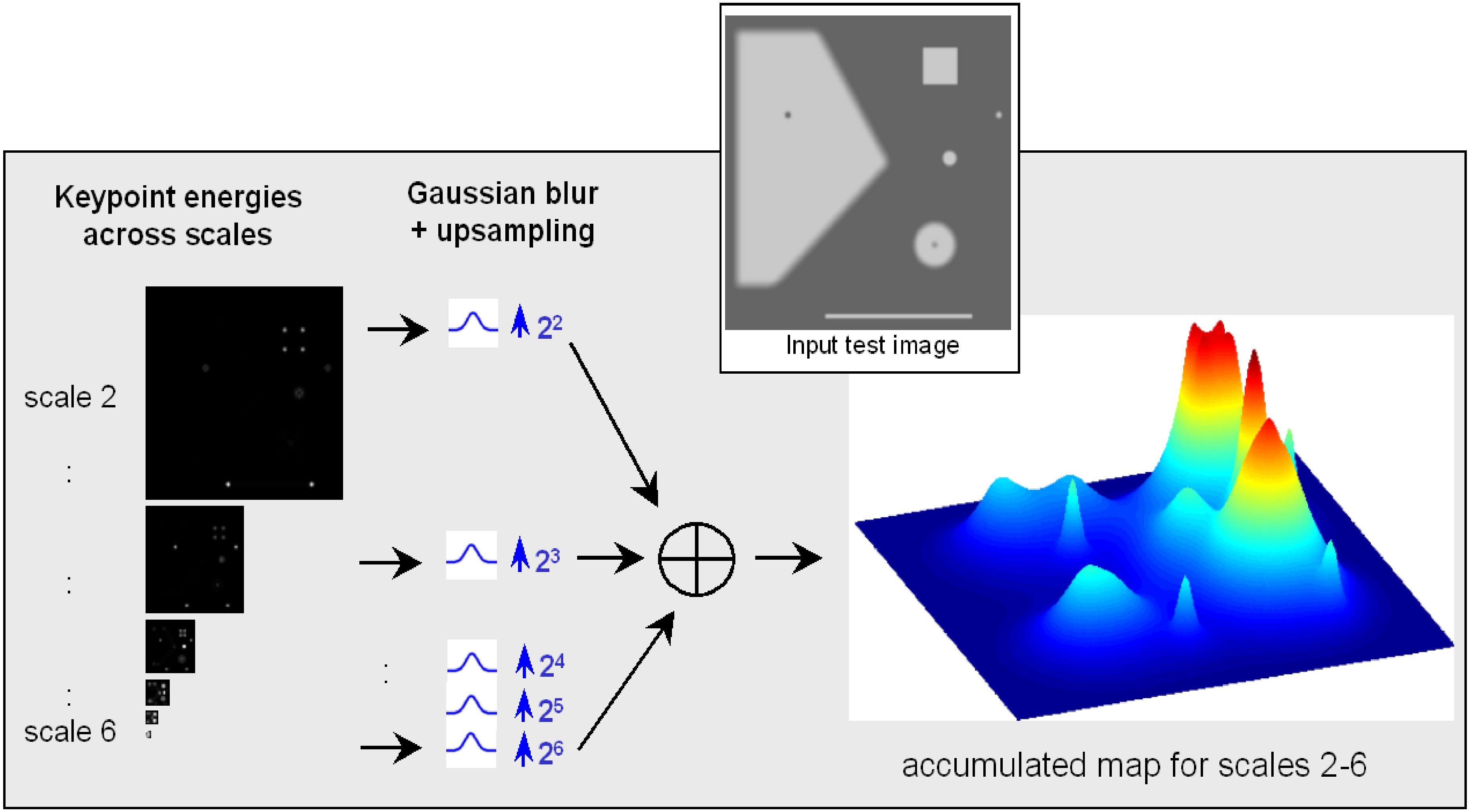
- **Aim:** to recover the best still image from the jittery video sequence of 75 frames.

- Video sequence is courtesy of Don Fraser, Australian Defence Forces Academy, Canberra.

- **Fusion:** based on max of each wavelet coefficient magnitude across the 75 frames, combined with the mean of each coefficient's phase.

Multi-scale Keypoint Detection using Accumulated Maps

Subject of work by **Julien Fauqueur**.



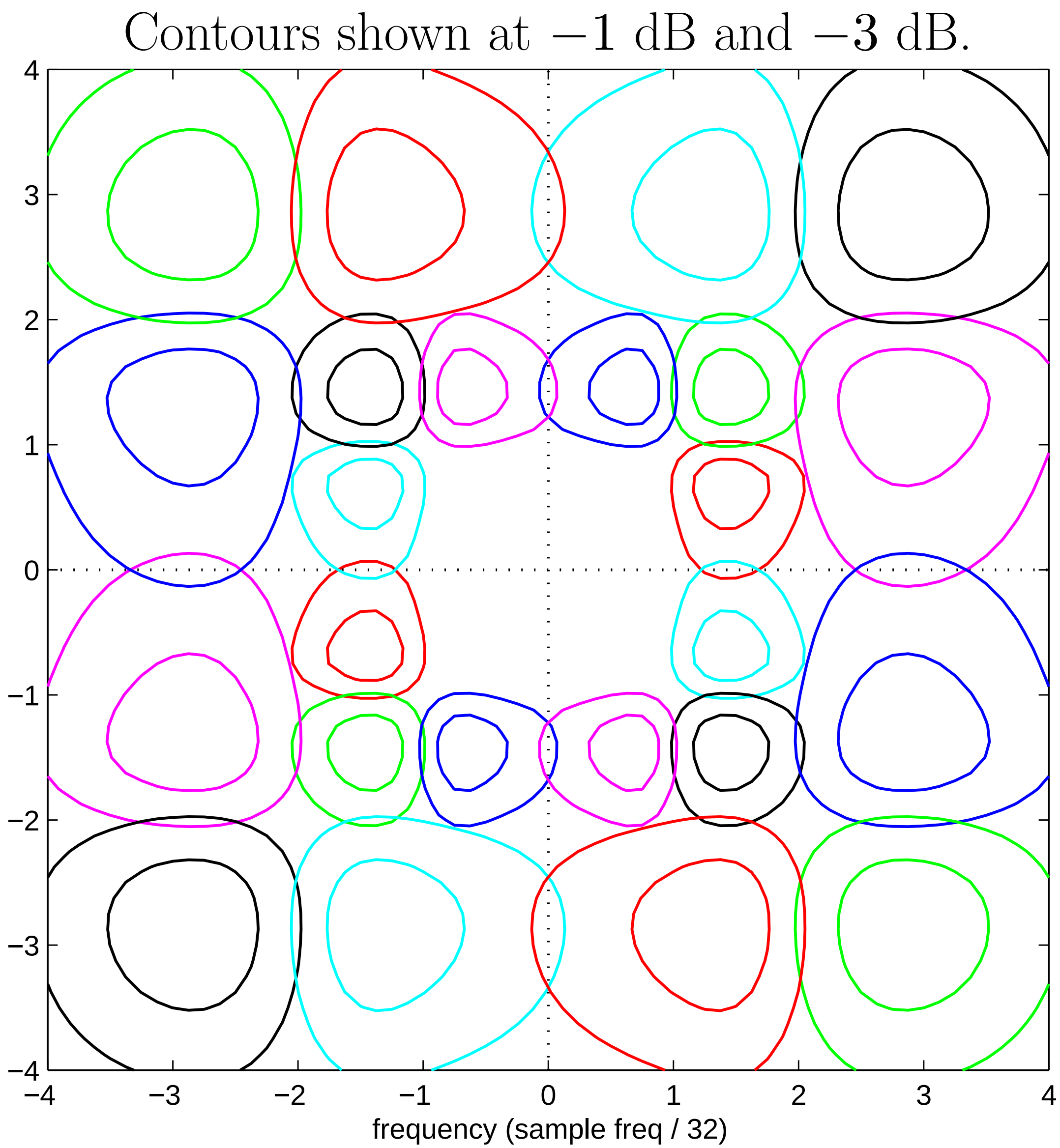
Rotation-Invariant Local Feature Matching

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 DTCWT coefficients provide multi-resolution local feature descriptors, but they are tied closely to a **rectangular sampling** system (as are most other multi-resolution decompositions).

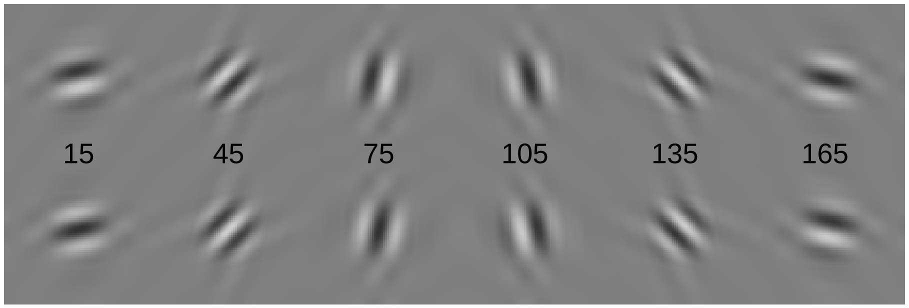
Hence we first need **better rotational symmetry** for the DTCWT.

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



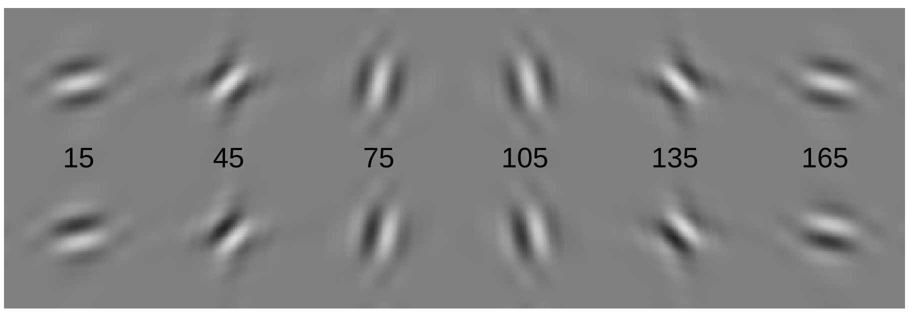
MODIFICATION OF 45° AND 135° SUBBAND RESPONSES FOR IMPROVED ROTATIONAL SYMMETRY (SHOWN AT LEVEL 4).

(a) Dual−Tree Complex Wavelets: Real Part



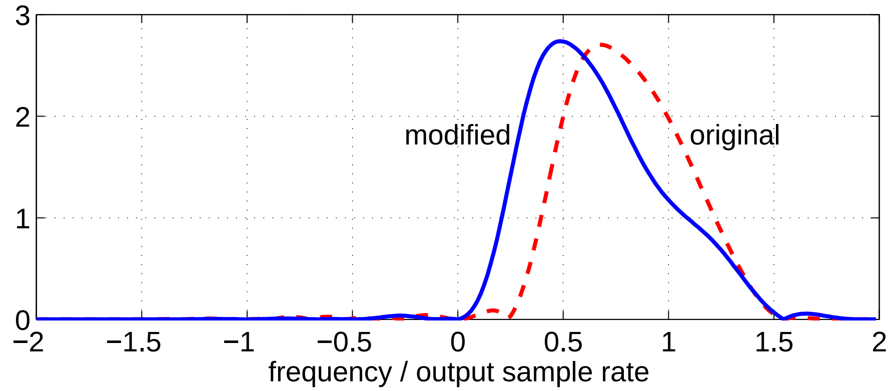
Imaginary Part

(b) Modified Complex Wavelets: Real Part



Imaginary Part

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



(a) Original 2-D impulse responses;

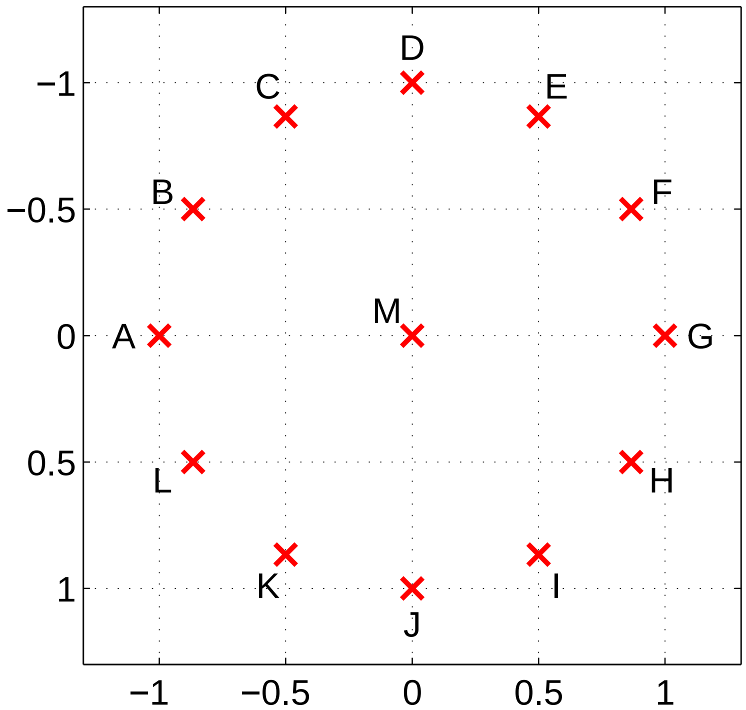
(b) 2-D responses, modified to have lower centre frequencies (reduced by $1/\sqrt{1.8}$) in the $45°$ and $135°$ subbands, and even / odd symmetric real / imaginary parts;

(c) Original and modified 1-D filters.

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

13-point circular pattern for sampling DTCWT 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:

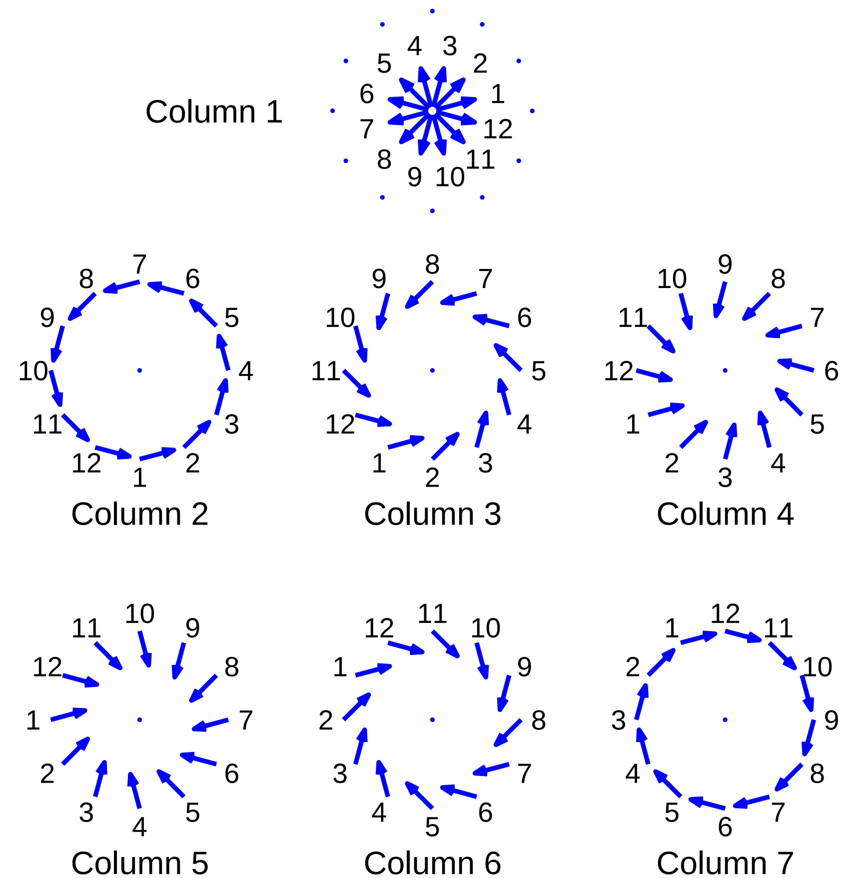
1. Shift by $\{-\omega_1, -\omega_2\}$ down to zero frequency (i.e. multiply by $e^{-j(\omega_1 x_1 + \omega_2 x_2)}$ at each point $\{x_1, x_2\}$);

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.

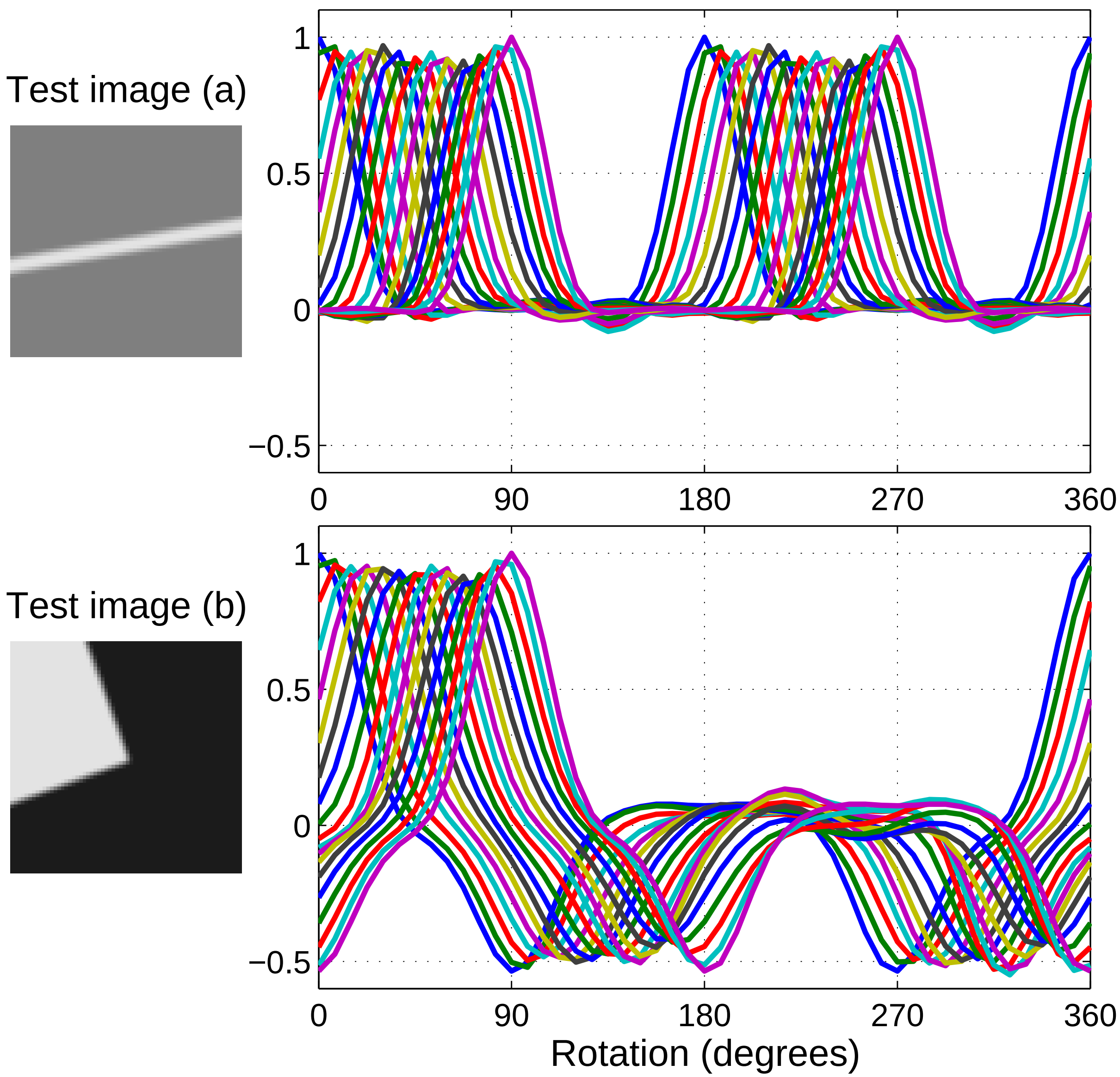
Numbers for each arrow give the row indices in P.

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 $\overline{P}_k$.
- Then, for each pair of keypoints (k, l) to be matched:
  - **Multiply** $\overline{P}_k$ by $\overline{P}_l^*$ element-by-element to give $\overline{S}_{k,l}$.
  - **Accumulate** the 12-point columns of $\overline{S}_{k,l}$ 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 $\overline{S}_{k,l}$ 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.

Correlation plots for two simple images

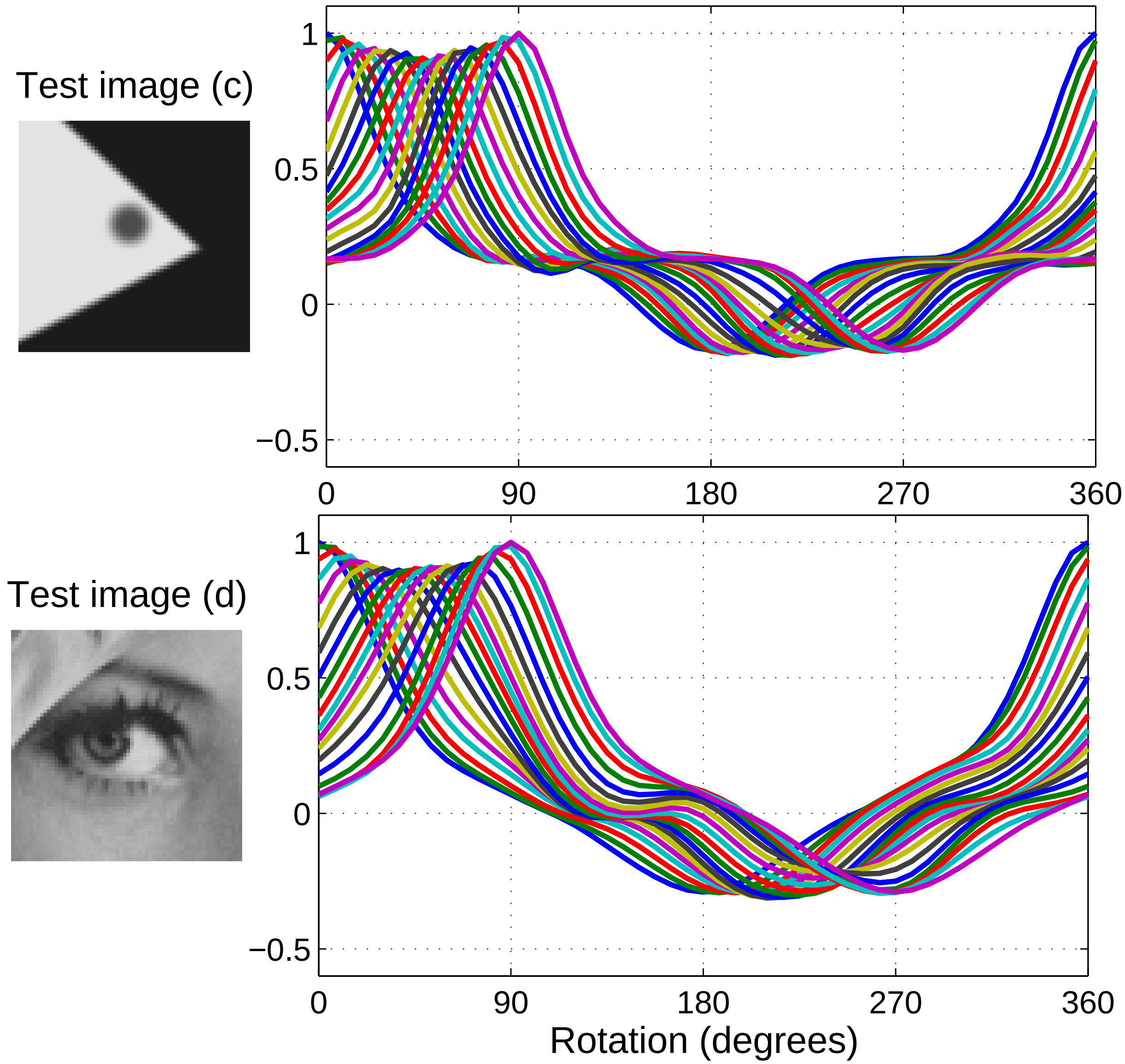


Each set of curves shows the output of the normalised correlator for 48 angles in 7.5° increments, when the test image is rotated in 5° increments from 0° to 90°.

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.

Correlation plots for more complicated images



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 $\overline{P}_k$ 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 $\mathbf{s}_{k,l}$ 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 propose to do the same for small scale errors using a derivative of $\overline{P}_k$ wrt scale.

CONCLUSIONS

The Dual-Tree Complex Wavelet Transform provides **shift invariance** and **orientation selectivity**, in addition to the usual properties of the DWT. We have shown how to apply the DTCWT in the following areas:

- **Denoising** of images and 3D data to achieve performance that equals or exceeds other approaches requiring much more computation.
- **Image Registration** with an efficient multi-resolution iterative algorithm - particularly suited to non-rigid motion.
- **Rotation-invariant local feature matching** at detected keypoints for object detection and keypoint matching applications.

Papers on complex wavelets are available at:

http://www.eng.cam.ac.uk/~ngk/

A Matlab DTCWT toolbox is available on request from:

ngk@eng.cam.ac.uk