Multi-Scale Displacement Estimation and Registration for 2-D and 3-D Datasets

Nick Kingsbury

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

July 2005

UNIVERSITY OF CAMBRIDGE

The Problem:

Efficient Displacement Estimation / Registration of Noisy Data

Applications:

- Registration of medical datasets taken some time apart and correction for patient movement
- Conversion from low-quality video to high-quality still images – e.g. correction of fluctuations in atmospheric refraction (heat shimmer)
- Motion estimation for non-rigid objects and fluids
- Registration of multi-look images affected by speckle, usually due to illumination from coherent sources such as lasers or synthetic aperture radar (SAR).

Displacement estimation usually involves measuring **gradients, derivatives or differences**. High noise levels mean that registration algorithms must be **robust to noise** if the noise is uncorrelated between images.

Key Features of Robust Registration Algorithms

- Edge-based methods are more robust than point-based ones.
- Bandlimited multiscale (wavelet) methods allow spatially adaptive denoising.
- Phase-based bandpass methods can give rapid convergence and immunity to illumination changes between images (**but** we have to be careful about 2π ambiguities) .
- If the displacement field is smooth, a wider-area parametric (affine) model of the field is likely to be more robust than a highly-local translation-only model.

Note: Biological vision systems have evolved to use multiscale directional bandpass filters as their front-end process (e.g. the V1 cortical filters in humans / mammals).

Selected Methods

- Dual-tree Complex Wavelet Transform (DT CWT):
  - efficiently synthesises a multiscale directional shift-invariant filterbank, with perfect reconstruction;
  - provides complex coefficients whose phase shift depends approximately linearly on displacement;
  - allows each subband of coefficients to be interpolated (shifted) independently of other subbands (because of shift invariance of the subband $H(z)$).

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

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

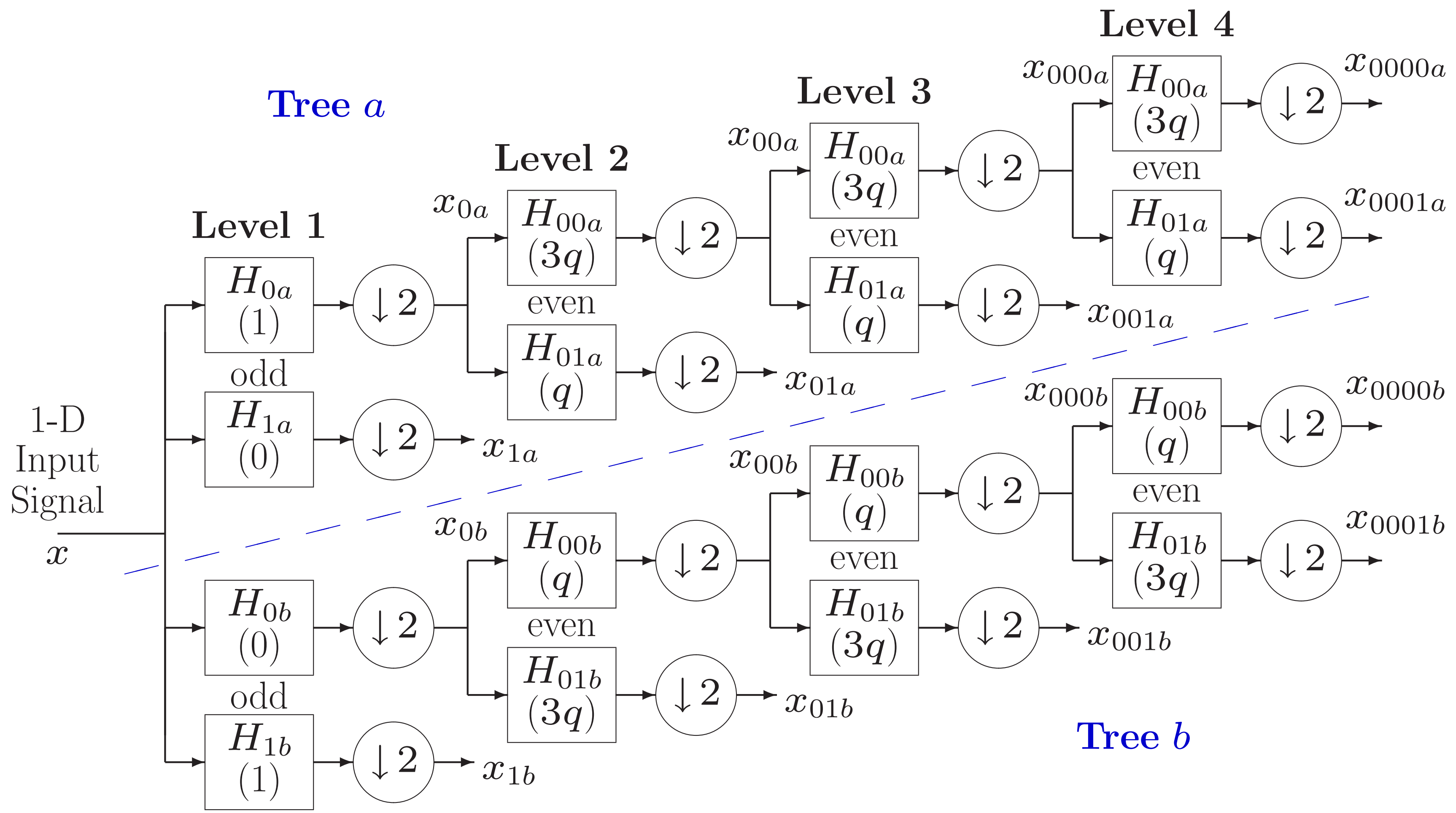


Figure 1: Dual tree of real filters for the Q-shift CWT, giving real and imaginary parts of complex coefficients from tree a and tree b respectively. Figures in brackets indicate the approximate delay for each filter, where $q = \frac{1}{4}$ sample period.

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

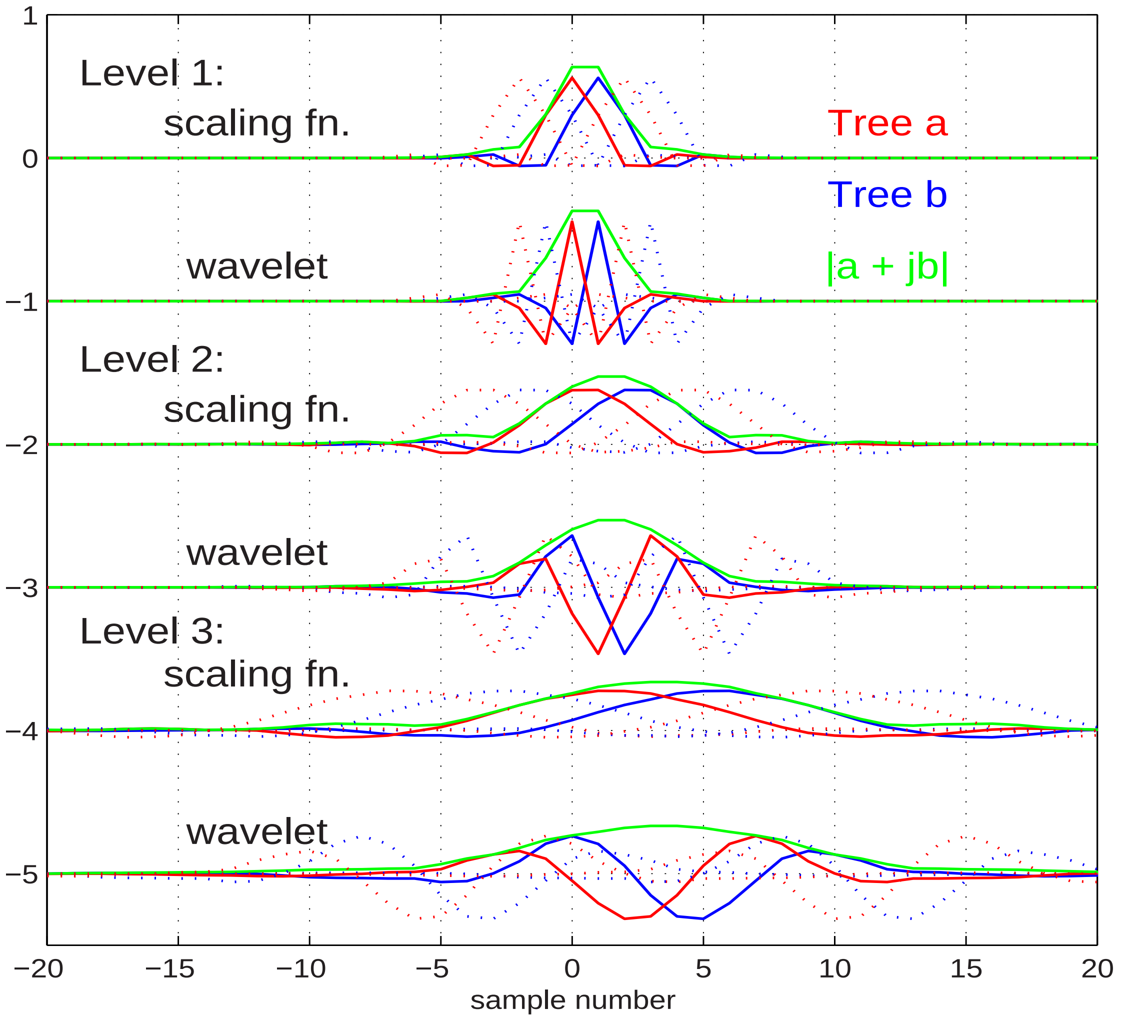
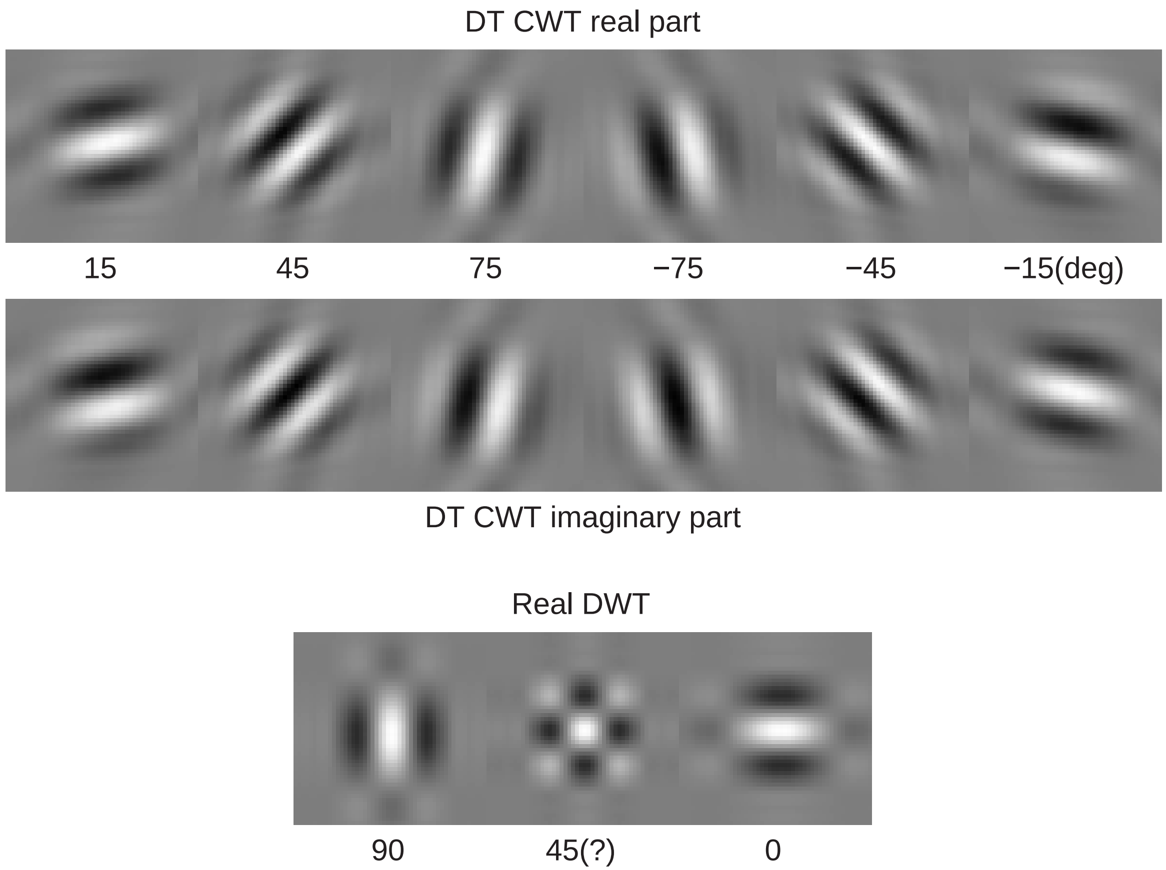
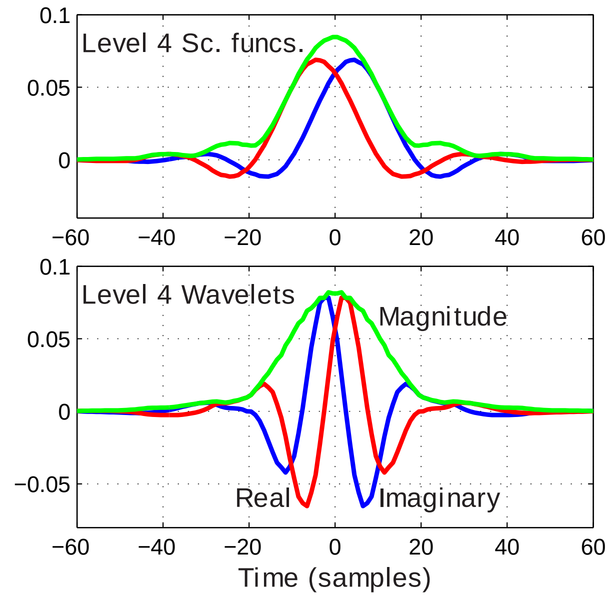


Figure 2: Basis functions for adjacent sampling points are shown dotted.

2-D Basis Functions at level 4

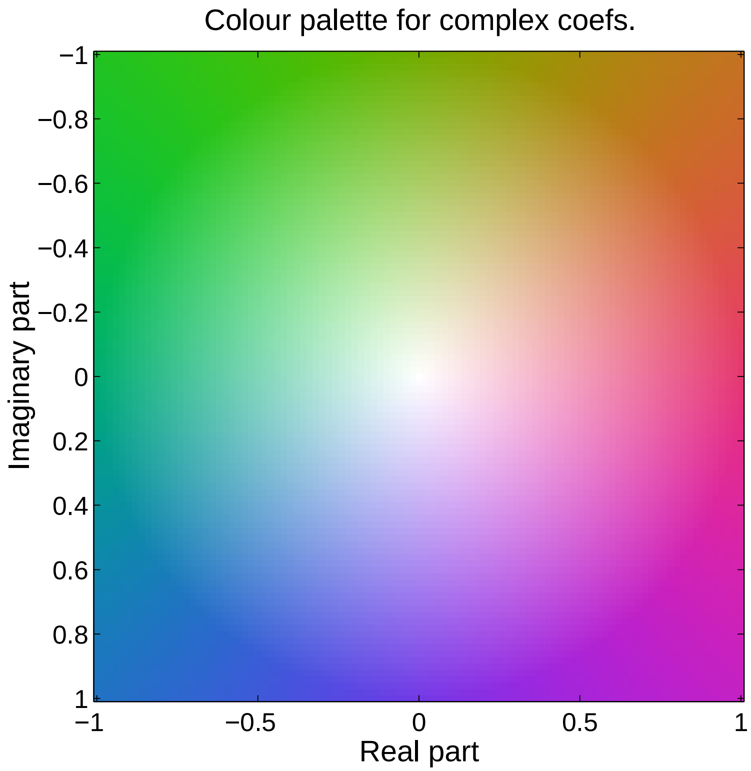




$$e^{j\omega_1 x}\ e^{j\omega_2 y} = e^{j(\omega_1 x + \omega_2 y)}$$

Figure 3: 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

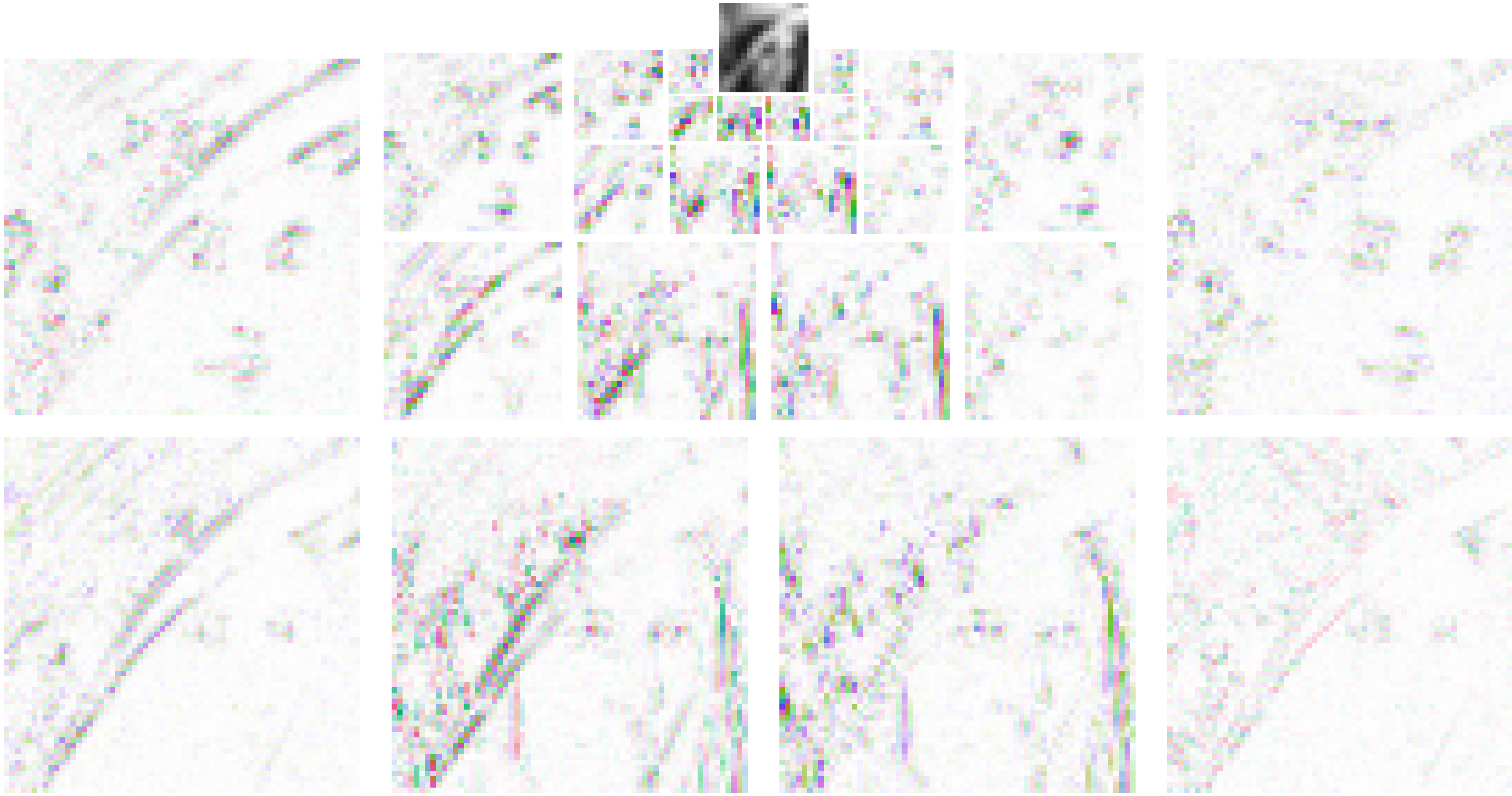
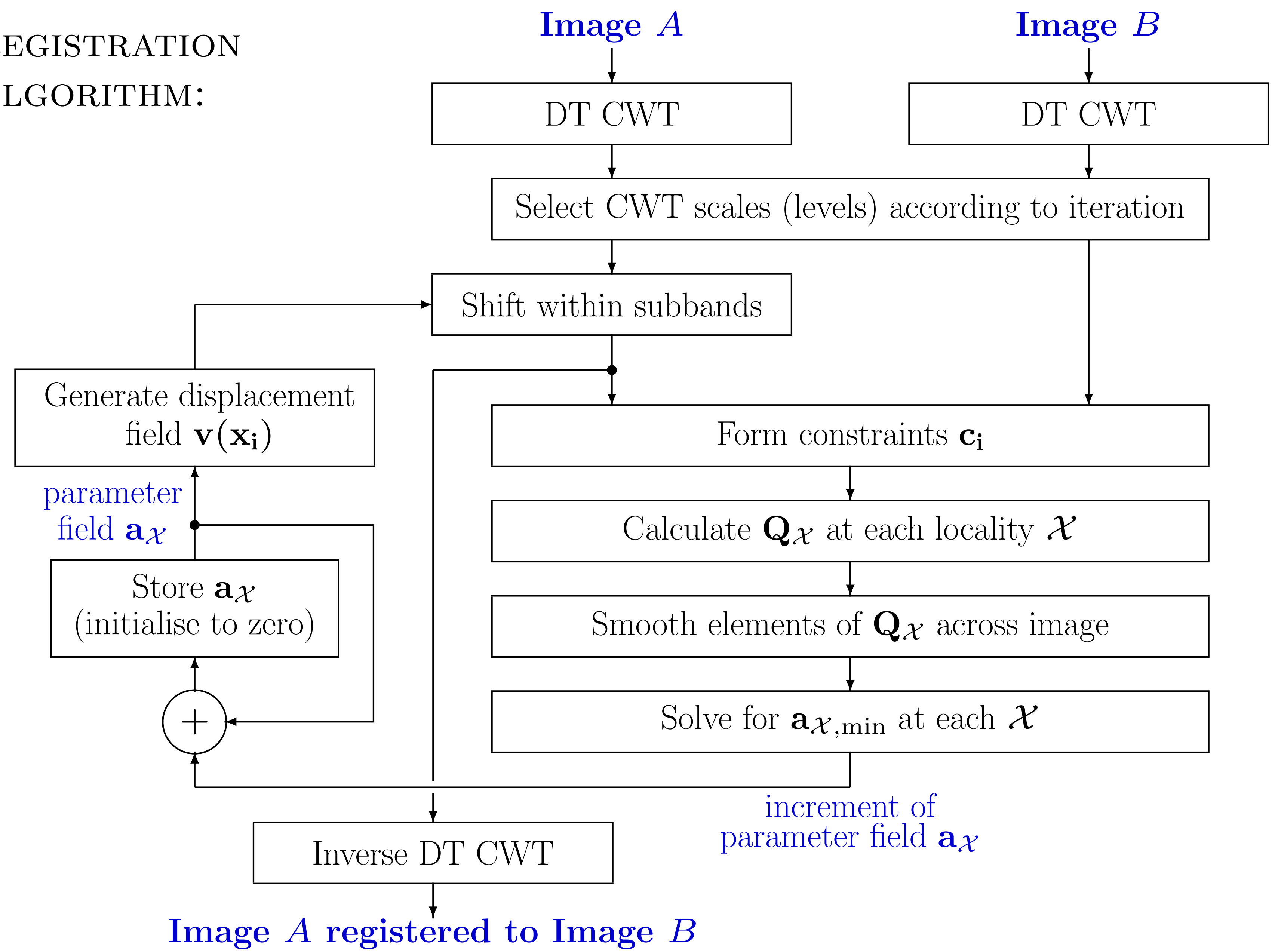


Figure 4: Four-level DT-CWT decomposition of *Lenna* into 6 subbands per level (only the central **128 × 128** portion of the image is shown for clarity). A colour-disc palette (see previous slide) is used to display the complex wavelet coefficients.

Registration Algorithm:



PARAMETRIC MODEL: LINEAR CONSTRAINT EQUATIONS

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

Note that, as well as $\mathbf{x}_i$, the locator i also specifies a subband direction d_i $(1 \ldots 6)$ and a scale (level) s_i. 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 $\{\mathbf{x}_i, d_i, s_i\}$ in images A and B have phases $\theta_{A,i}$ and $\theta_{B,i}$, approximate **linearity of phase θ vs. displacement $\mathbf{v}(\mathbf{x}_i)$** means that

$$\mathbf{c}_i^T \; \tilde{\mathbf{v}}_i \approx 0 \quad \text{if} \quad \mathbf{c}_i = C_i \begin{bmatrix} \nabla_{\mathbf{x}}\, \theta_i \\ \theta_{B,i} - \theta_{A,i} \end{bmatrix}$$

In practise we compute this by averaging finite differences at the centre $\mathbf{x}_i$ of a $2 \times 2 \times 2$ block of coefficients from a given subband $\{d_i, s_i\}$ of images A and B.

Note: C_i is a constant which does not affect the line defined by the constraint, but it is important as a weight for combining constraint errors (see later).

Parameters of the Model

We can define a **6-term affine parametric model a** for **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 **K** of the form

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

Solving for the Model Parameters

Using techniques (due to Hemmendorff et al) similar to homogeneous coordinates:

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 scale-space locality $\mathcal{X}$, we wish to find the parametric vector $\tilde{\mathbf{a}}$ such that

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

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

$$\mathcal{E}_{\mathcal{X}} = \sum_{i \in \mathcal{X}} ||\mathbf{c}_i^T \, \tilde{\mathbf{v}}_i||^2 = \sum_{i \in \mathcal{X}} ||\mathbf{c}_i^T \, \tilde{\mathbf{K}}_i \, \tilde{\mathbf{a}}||^2 = \sum_{i \in \mathcal{X}} \tilde{\mathbf{a}}^T \, \tilde{\mathbf{K}}_i^T \, \mathbf{c}_i \, \mathbf{c}_i^T \, \tilde{\mathbf{K}}_i \, \tilde{\mathbf{a}} = \tilde{\mathbf{a}}^T \, \tilde{\mathbf{Q}}_{\mathcal{X}} \, \tilde{\mathbf{a}}$$

where $\tilde{\mathbf{Q}}_{\mathcal{X}} = \sum_{i \in \mathcal{X}} \tilde{\mathbf{K}}_i^T \, \mathbf{c}_i \, \mathbf{c}_i^T \, \tilde{\mathbf{K}}_i$.

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 (including all directions d and all selected scales s) 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, $\mathbf{c}_i = C_i \begin{bmatrix} \nabla_{\mathbf{x}}\, \theta(\mathbf{x}_i) \\ \theta_B(\mathbf{x}_i) - \theta_A(\mathbf{x}_i) \end{bmatrix}$,

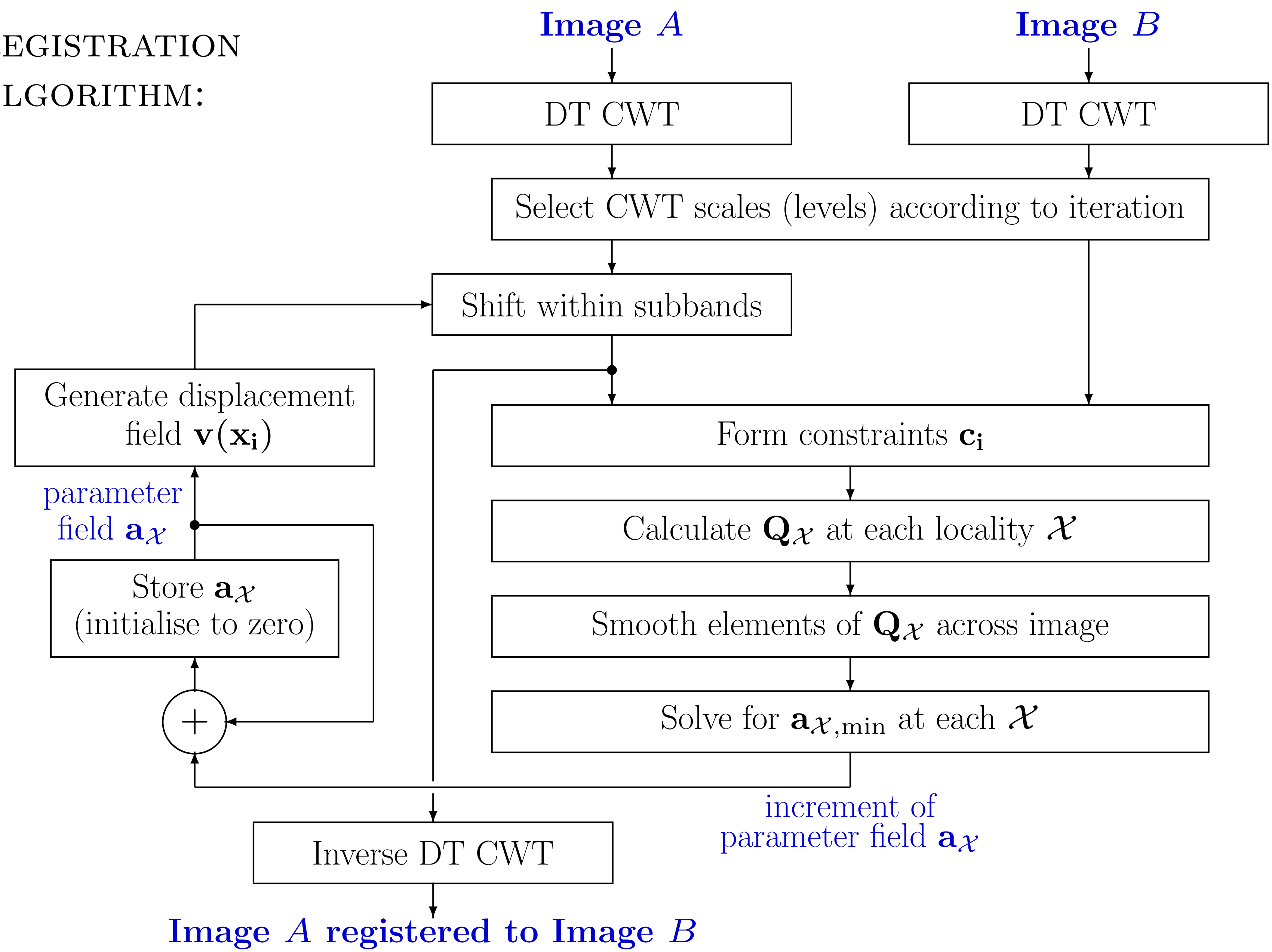
the constant gain parameter C_i will determine how much weight is given to each constraint in $\tilde{\mathbf{Q}}_{\mathcal{X}} = \sum_{i \in \mathcal{X}} \tilde{\mathbf{K}}_i^T\, \mathbf{c}_i\, \mathbf{c}_i^T\, \tilde{\mathbf{K}}_i$.

Hemmendorf proposes some quite complicated heuristics for computing C_i, but for our work, we find the following gives **maximum weight to consistent sets of wavelet coefficients** and 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.

Registration Algorithm:



Demonstrations

- Registration of CT scans
  - Two scans of the abdomen of the same patient, taken at different times with significant differences in position and contrast.
  - Task is to register the two images as well as possible, despite the differences.

- Enhancement of video corrupted by atmospheric turbulence.
  - 75 frames of video of a house on a distant hillside, taken through a high-zoom lens with significant turbulence of the intervening atmosphere due to rising hot air.
  - Task is to register each frame to a 'mean' image from the sequence, and then to reconstruct a high-quality still image from the registered sequence.

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:

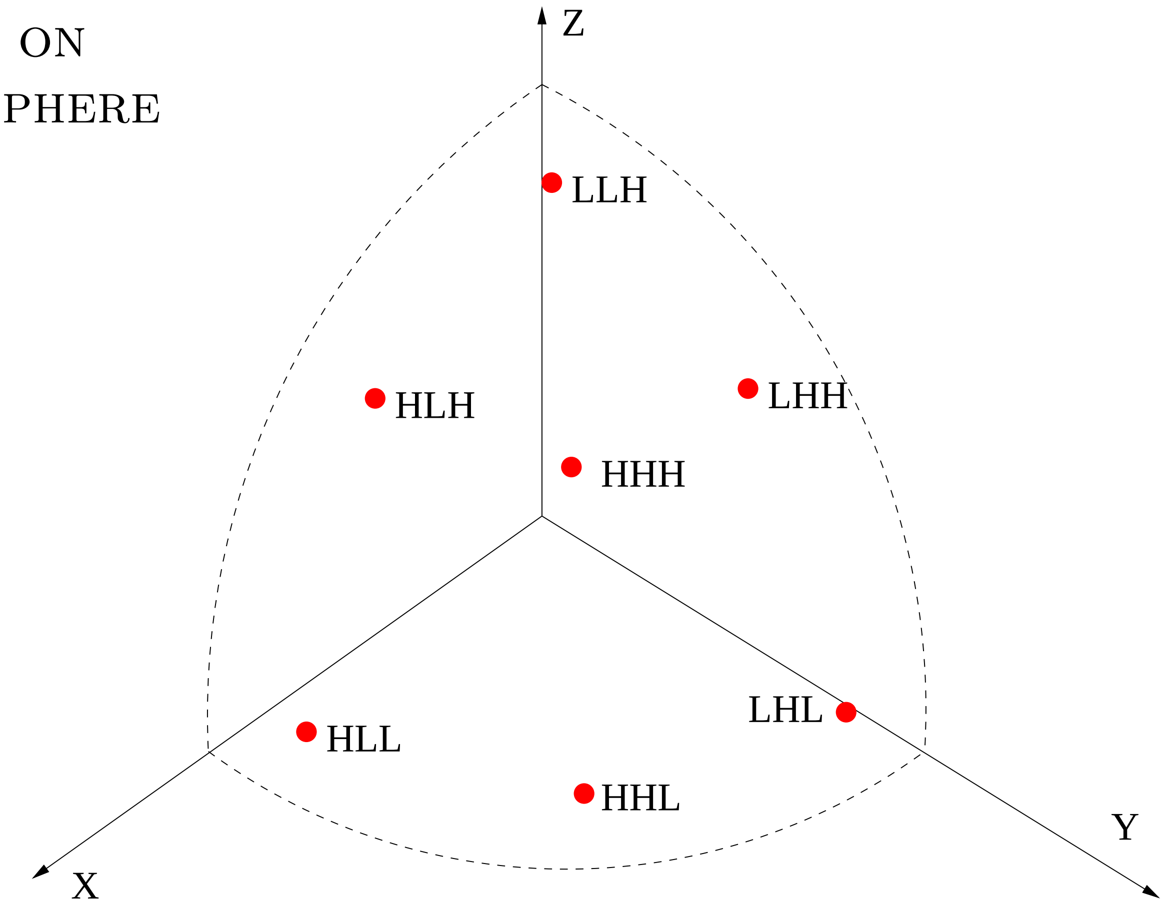
- It is performed separably, with 2 trees used for the rows, 2 trees for the columns and 2 trees for the slices of the 3-D dataset – yielding an **Octal-Tree** structure (8:1 redundancy).

- 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 $\omega_H \simeq 3\omega_L$.

3-D Implications for the Phase-based Parametric Method

- $\mathbf{x}_i$ and $\mathbf{v}(\mathbf{x}_i)$ become 3-element vectors, so $\mathbf{c}_i$ and $\tilde{\mathbf{v}}_i$ become 4-vectors.

- For a 3-D affine model, $\mathbf{K}$ becomes a 3×12 matrix, so that:

$$\mathbf{v}(\mathbf{x}) = \begin{bmatrix} 1 & 0 & 0 & x_1 & 0 & 0 & x_2 & 0 & 0 & x_3 & 0 & 0 \\ 0 & 1 & 0 & 0 & x_1 & 0 & 0 & x_2 & 0 & 0 & x_3 & 0 \\ 0 & 0 & 1 & 0 & 0 & x_1 & 0 & 0 & x_2 & 0 & 0 & x_3 \end{bmatrix} . \begin{bmatrix} a_1 \\ \vdots \\ a_{12} \end{bmatrix} = \mathbf{K}(\mathbf{x}) \ . \ \mathbf{a}$$

  and $\tilde{\mathbf{K}}$ becomes a 4×13 matrix.

- Hence $\tilde{\mathbf{Q}}_{\mathcal{X}} = \sum_{i \in \mathcal{X}} \tilde{\mathbf{K}}_i^T \ \mathbf{c}_i \ \mathbf{c}_i^T \ \tilde{\mathbf{K}}_i$ becomes a 13×13 symmetric matrix, containing $13 \times 7 = 91$ distinct elements per locality $\mathcal{X}$. At each selected scale s_i and spatial location $\mathbf{x}_i$ in $\mathcal{X}$, there are now 28 subband directions d_i.

Conclusions

Our proposed algorithm for **robust registration** effectively combines

- **The Dual-Tree Complex Wavelet Transform**
  - Linear phase vs. shift behaviour
  - Easy shiftability of subbands
  - Directional filters select edge-like structures
  - Good denoising of input images
- **Hemmendorf's phase-based parametric method**
  (Hemmendorff et al, IEEE Trans Medical Imaging, Dec 2002)
  - Finds LMS fit of parametric model to edges in images
  - Allows simple filtering of $\mathbf{Q}_{\mathcal{X}}$ to fit more complex motions
  - Integrates well with multiscale DT CWT structure

Papers on complex wavelets are available at:

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

Points for Discussion:

- Why has the **human / mammalian visual system** evolved to use directional multiscale bandpass filters as its front end?
- Are directional multiscale complex bandpass filters the **optimum** approach to detecting displacement / motion?
- What is the real meaning of **Hilbert Transform** in 2-D and 3-D spaces?
- Are directional (multiscale?) bandpass filters the key to giving meaning to 2-D and 3-D Hilbert Transforms?