# The Monogenic Framework:
A Systematic Approach to Image Processing and Computer Vision

Michael Felsberg

Computer Vision Laboratory
Linköping University

Sept 2, 2003



## Coherent Framework

- Phase based image processing
- In scale-space
- Allows to derive applications in a systematic way



## Phase Based Processing

- Why?
- When?
- How?



## Phase Based Processing in nD

- Generalized quadrature filter?
- Generalized Hilbert transform?
- Generalized phase definition?

## Scale-Space

- Scale relative image processing
- Efficiency
- Correctness



## Systematic Derivations

- Complex IP & CV systems typically contain heuristic *and* systematic approaches
- Reduce the amount of heuristics
  - improve accuracy
  - reduce computational complexity
  - stability and robustness
  - correctness

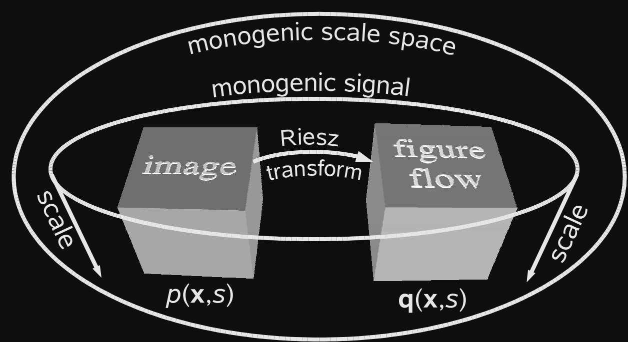


## Monogenic Framework

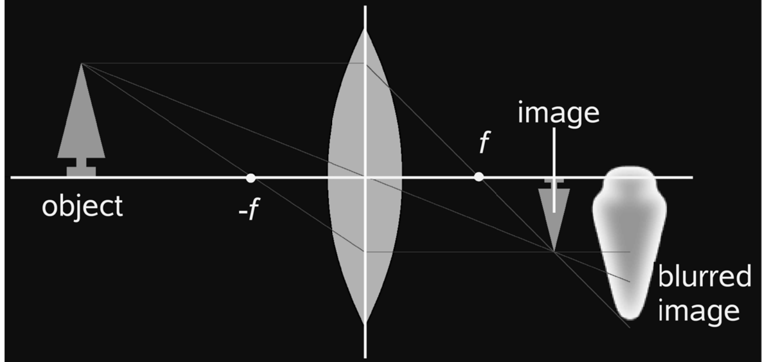
- New quadrature filter combine properties of
  - Gaussian derivatives (n-jets) and
  - Gabor filters
- The 2D phase approach allows to compute phase congruency which is
  - isotropic (no sampled orientation)
  - differential (no sampled scales)
- Relation of amplitude and phase
  - reconstruction from phase and amplitude
  - comparison of phase congruency and amplitude maxima



## Monogenic Scale Space

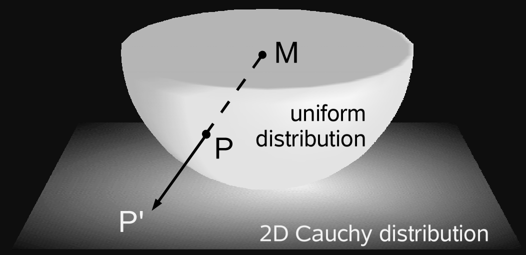


## Motivation Scale Space






## Gnomonic Projection




Cauchy distribution corresponds to Poisson kernel



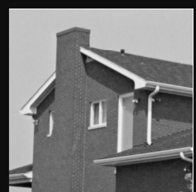
## Poisson Scale Space

- Poisson kernel: $P_s(\mathbf{x}) = \frac{s}{\omega_n(|\mathbf{x}|^2 + s^2)^{(n+1)/2}}$
- frequency response: $\hat{P}_s(\mathbf{u}) = \exp(-2\pi|\mathbf{u}|s)$
- Poisson scale space:

$$p(\mathbf{x}, s) = (b * P_s)(\mathbf{x})$$

- corresponding flux?

$$\partial_s p(\mathbf{x}, s) = -\nabla \cdot \mathbf{q}(\mathbf{x}, s)$$





## 2D Riesz Transform

- Generalizes the Hilbert transform to 2D
- Consists of two components
- Kernel:

$$\mathbf{h}(\mathbf{x}) = \frac{\mathbf{x}}{2\pi|\mathbf{x}|^3}$$

- Frequency response:

$$\mathbf{H}(\mathbf{u}) = -i\frac{\mathbf{u}}{|\mathbf{u}|}$$

$$\mathbf{m}(\mathbf{x}, s) = \begin{bmatrix} \mathbf{q}(\mathbf{x}, s) \\ p(\mathbf{x}, s) \end{bmatrix}$$

$$\mathbf{q}(\mathbf{x}, s) = (p(\cdot, s) * \mathbf{h})(\mathbf{x})$$

# nD Riesz Transform

- Generalizes the 1D analytic signal to nD
- Consists of n components
- Kernel:

$$\mathbf{h}(\mathbf{x}) = \frac{\mathbf{x}}{\omega_n |\mathbf{x}|^{n+1}}$$

Frequency response:

$$\mathbf{H}(\mathbf{u}) = -i \frac{\mathbf{u}}{|\mathbf{u}|}$$

$$\mathbf{m}(\mathbf{x}, s) = \begin{bmatrix} \mathbf{q}(\mathbf{x}, s) \\ p(\mathbf{x}, s) \end{bmatrix}$$

$$\mathbf{q}(\mathbf{x}, s) = (p(\cdot, s) * \mathbf{h})(\mathbf{x})$$



# PDE Formulation

1D: **m** corresponds to an analytic function

= 2D harmonic field

$\nabla_{\mathbf{x},s} \times \mathbf{m}(\mathbf{x}, s) = \mathbf{0}$ zero curl

$\nabla_{\mathbf{x},s} \cdot \mathbf{m}(\mathbf{x}, s) = 0$ zero divergence

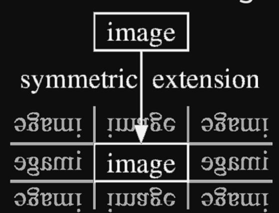
if $s > 0$

$p(\mathbf{x}, 0) = b(\mathbf{x})$ boundary condition



# Finite Domain Solution

- no intensity should enter or leave the image (zero Neumann boundary condition)
- effective filter kernel: reflection at the boundary
- alternative: extend the image





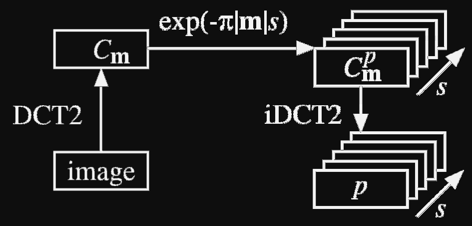

# Eigensystem

- solution by eigensystem
  - eigenfunctions: $I_{mn}(x, y) = 2\cos(\pi x m)\cos(\pi y n)$
  - eigenvalues of the generator: $\pi\sqrt{m^2+n^2}$
  - 'frequency response': $\exp(-\pi\sqrt{m^2+n^2}\, s)$

# Discrete Implementation

$$C_{mn}=\alpha_{mn}\sum_{x=0}^{X-1}\sum_{y=0}^{Y-1} b_{xy}\cos(\pi(x+\tfrac{1}{2})m/X)\cos(\pi(y+\tfrac{1}{2})n/Y)$$

$$0\leqslant m\leqslant X-1$$
$$0\leqslant n\leqslant Y-1$$






# Finite Domain Implementation II

- DCT coefficients of potential

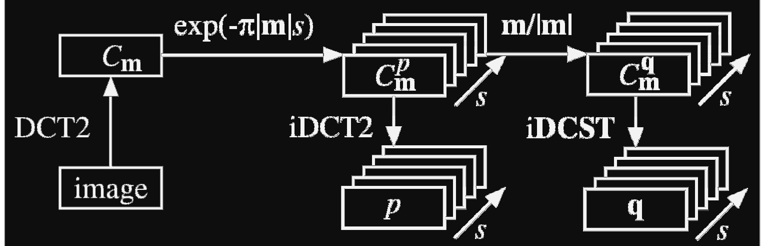
$$C^{\varphi}_{\mathbf{m},s}=\frac{C^{p}_{\mathbf{m},s}}{-\pi|\mathbf{m}|} \qquad \exp(-\pi|\mathbf{m}|s)$$

- gradient of potential

$$\mathbf{q}_{\mathbf{x},s}=\nabla\varphi_{\mathbf{x},s}=$$
$$-\pi\sum_{m=0}^{X-1}\sum_{n=0}^{Y-1}C^{\varphi}_{\mathbf{m},s}\begin{bmatrix} m\sin(\pi(x+\frac{1}{2})m/X)\cos(\pi(y+\frac{1}{2})n/Y) \\ n\cos(\pi(x+\frac{1}{2})m/X)\sin(\pi(y+\frac{1}{2})n/Y)\end{bmatrix}$$



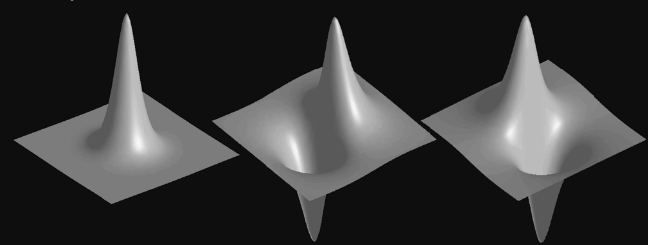
# Algorithm Monogenic Scale Space






# Spherical Quadrature Filter

- combine radial bandpass with its Riesz transform
- orthogonal like Gaussian derivatives / n-jets
- in quadrature like Gabor filters



## Truncated DOP Filter

- truncate DOP filter at a certain energy of lp

$$lp_1(\mathbf{x}) = \frac{s_1}{2\pi(|\mathbf{x}|^2 + s_1^2)^{3/2}} \qquad lp_2(\mathbf{x}) = \frac{s_2}{2\pi(|\mathbf{x}|^2 + s_2^2)^{3/2}}$$

- ensure zero DC

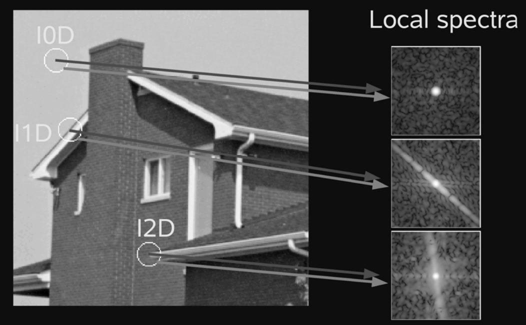
$$bp(\mathbf{x}) = \frac{lp_1(\mathbf{x})}{\sum_{\mathbf{x}} lp_1(\mathbf{x})} - \frac{lp_2(\mathbf{x})}{\sum_{\mathbf{x}} lp_2(\mathbf{x})}$$

- ensure same energy of components

$$bp(\mathbf{x}) \Rightarrow \sqrt{\sum_{\mathbf{x}} \mathbf{cbp}(\mathbf{x})^2 / \sum_{\mathbf{x}} bp(\mathbf{x})^2}\, bp(\mathbf{x})$$

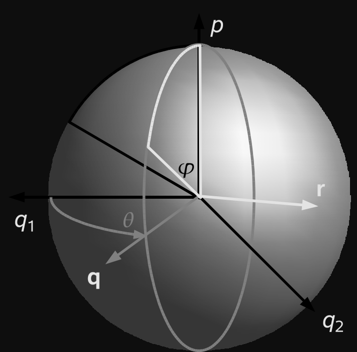


## Intrinsic Dimensionality






## Monogenic Phase

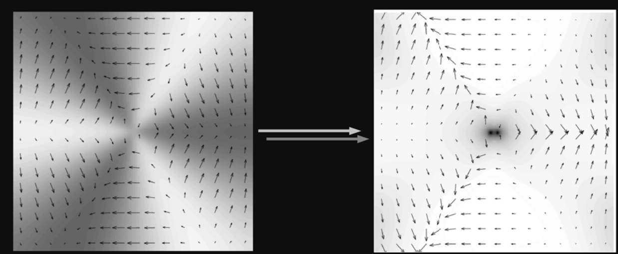



attenuation: logarithm of Euclidean norm



## Attenuation and Phase

$$A = \log(\sqrt{|\mathbf{q}|^2 + p^2}) = \frac{1}{2}\log(|\mathbf{q}|^2 + p^2) \qquad \mathbf{r} = \frac{\mathbf{q}}{|\mathbf{q}|}\arctan\left(\frac{|\mathbf{q}|}{p}\right)$$



# Applications of the Framework

- Flux gives an *unbiased* orientation estimate
  - regularized derivative operator (“derivator”)
  - all frequency components are weighted the same
- Superior edge detection by phase congruency
  - differential phase congruency
  - no sampling of the orientation necessary (fast)
- Disparity estimation from local phase
  - no need for an a priori fixed orientation
  - optimal SNR
- Reconstruction from local amplitude and phase
  - extremly simple
  - fast and accurate



# Orientation Estimation

- noisy image with different local frequencies

original signal | orientation from Riesz transform | ground truth



# Edge & Line Detection by PC

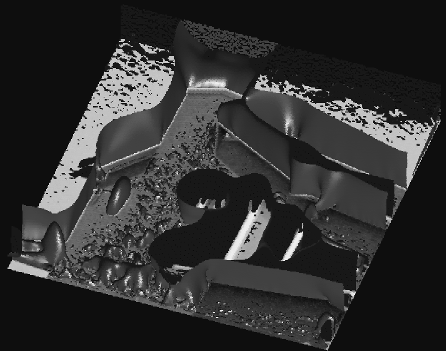
Zeros of $\partial_s \mathbf{r}$



# Filter for PC Estimation

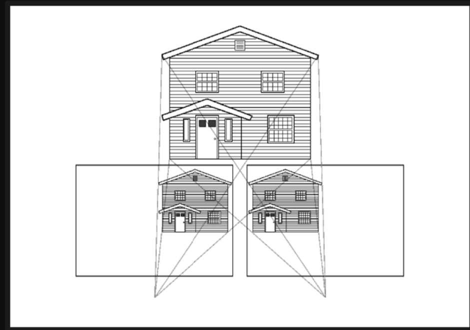
$$\begin{aligned}
\partial_s \mathbf{r} &= \partial_s \frac{\mathbf{q}}{|\mathbf{q}|} \arctan\left(\frac{|\mathbf{q}|}{p}\right) \\
&= \frac{\mathbf{q}}{|\mathbf{q}|} \frac{1}{1+\frac{|\mathbf{q}|^2}{p^2}} \frac{p\,\partial_s|\mathbf{q}| - |\mathbf{q}|\,p_s}{p^2} \\
&= \frac{p\,\mathbf{q}_s - \mathbf{q}\,p_s}{|\mathbf{q}|^2 + p^2} \\
&= \frac{(bp*b)(\mathbf{cbp}_s*b) - (\mathbf{cbp}*b)(bp_s*b)}{|\mathbf{cbp}*b|^2 + (bp*b)^2}
\end{aligned}$$

## Phase Congruency through Scale



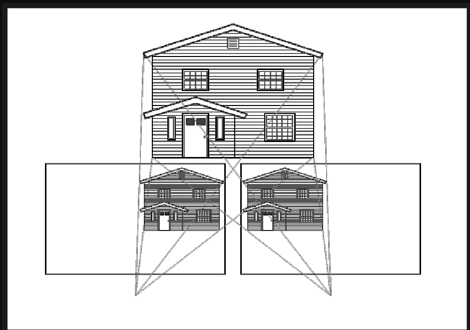


## Disparity Estimation





## Disparity Estimation



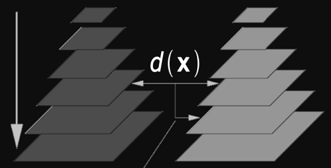


## Disparity from Local Phase

- Faster than other appearance based approaches
  (e.g. correlation based)
- Feature / model based approaches produce more ambiguities
- Sub-pixel accuracy

# Disparity Algorithm

- Compute scale pyramid
- For each scale (C2F)
  - estimate disparity
  - add to old estimate
  - upsample and CS
  - compensate disparity in next scale
- Add final estimate






# Taylor Expansion of Phase

- Taylor series expansion for phase vector ($\mathbf{d} \| \mathbf{r}$)

$$\mathbf{r}(\mathbf{x} \pm \frac{\mathbf{d}}{2}) = \mathbf{r}(\mathbf{x}) \pm \frac{\mathbf{d}}{2} \cdot \nabla \mathbf{r}(\mathbf{x}) + O(|\mathbf{d}|^2)$$

- Solve for normal disparity

$$\mathbf{d}(\mathbf{x}) \approx 2 \frac{\mathbf{r}^l(\mathbf{x}) - \mathbf{r}^r(\mathbf{x})}{\nabla \cdot \mathbf{r}^l(\mathbf{x}) + \nabla \cdot \mathbf{r}^r(\mathbf{x})}$$

- The local frequency (divergence of the phase vector) can be computed without using inverse trigonometric functions (see PC)

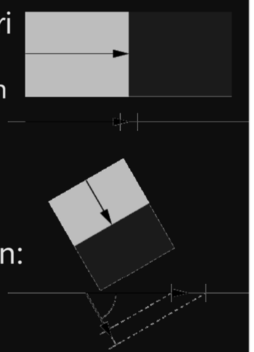


# Advantages

- Disparity orientation $\mathbf{e}$ is a priori known:
  - Disparity is obtained by projection

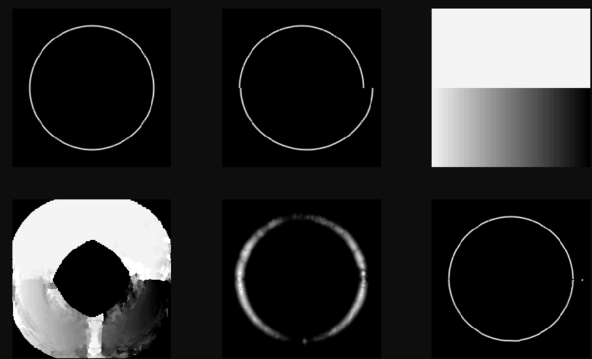
$$d(\mathbf{x}) = \frac{|\mathbf{d}(\mathbf{x})|^2}{\mathbf{e} \cdot \mathbf{d}(\mathbf{x})}$$

  - Correct confidence can be computed explicitly
- Disparity orientation is unknown:
  - Estimate $\mathbf{e}$ from the normal disparities
  - Continue as in the first case

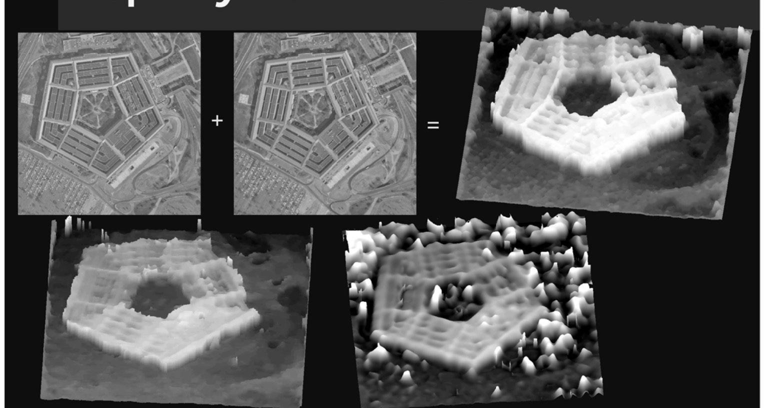




# Synthetic Experiment

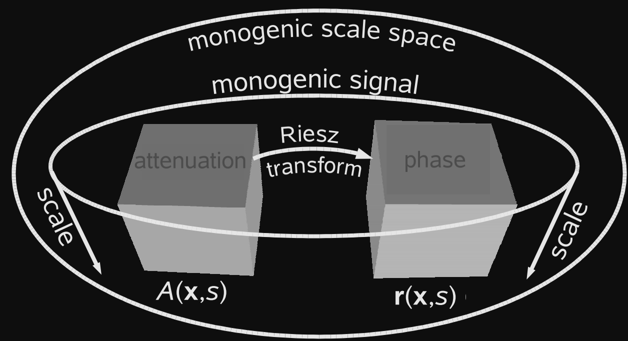


# Disparity from Phase






# Attenuation-Phase Scale Space






# Attenuation & Phase

- Known: complex derivative of complex logarithm of an analytic function is analytic

$$\partial_z \log(f(z)) = \frac{\partial_z f(z)}{f(z)}$$

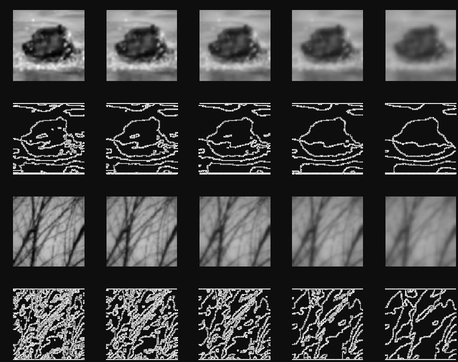
- Generalizes to 2D:

$$\begin{bmatrix} \mathbf{r} \\ A \end{bmatrix} = \text{'log'} \begin{bmatrix} \mathbf{q} \\ p \end{bmatrix}$$

is a harmonic field



# Zeros in Monogenic Scale-Space



## Fundamental Relations

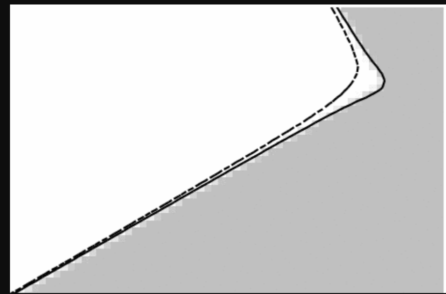
$$\left.\begin{array}{r}\partial_x r_2 - \partial_y r_1 = 0 \\ \nabla \cdot \mathbf{r} + A_s = 0\end{array}\right] \longrightarrow \omega = \nabla \cdot \mathbf{r} = -A_s$$

$$\left.\begin{array}{r}\partial_x A - \partial_s r_1 = 0 \\ \partial_y A - \partial_s r_2 = 0\end{array}\right] \longrightarrow \text{local maxima = PC}$$



## PC & Local Amplitude Maxima

➲ phase congruency = local maxima of amplitude





## Fundamental Relations

$$\left.\begin{array}{r}\partial_x r_2 - \partial_y r_1 = 0 \\ \nabla \cdot \mathbf{r} + A_s = 0\end{array}\right] \longrightarrow \omega = \nabla \cdot \mathbf{r} = -A_s$$

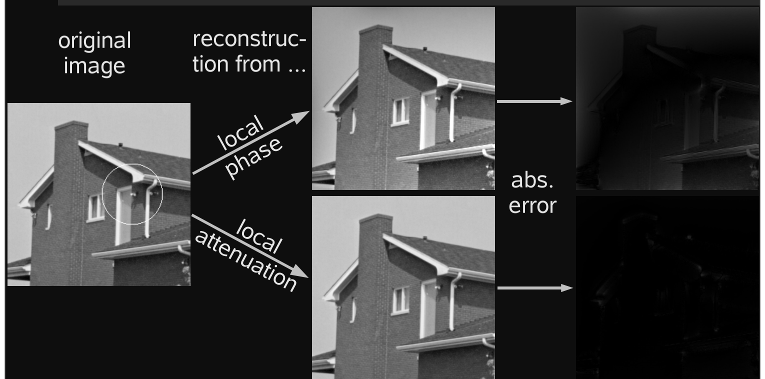
$$\left.\begin{array}{r}\partial_x A - \partial_s r_1 = 0 \\ \partial_y A - \partial_s r_2 = 0\end{array}\right] \longrightarrow \text{local maxima = PC}$$

$\boxed{\mathbf{r} = \mathbf{h} * A}$ ⟶ reconstruction from attenuation / phase

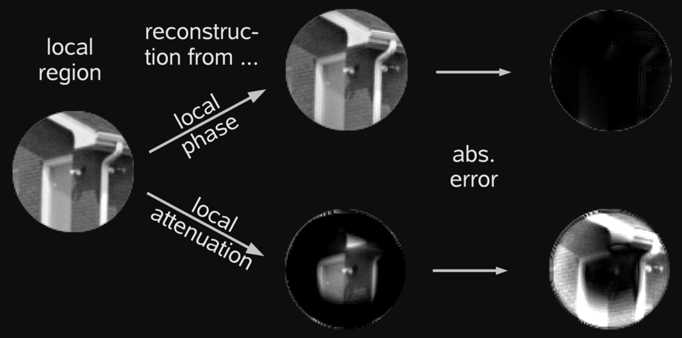
$$b \approx \exp(A)\cos(|\mathbf{r}|)$$



## Global Reconstruction



## Local Reconstruction






## Conclusion

- The monogenic scale space approach can be implemented
  - globally by DCT and
  - locally by SQFs in the spatial domain.
  - PC and local frequency can be computed from implicit derivatives.
- Future topics include
  - further applications (e.g. optical flow)
  - JAVA demo / implementation
  - extension to higher dimensions (straightforward).