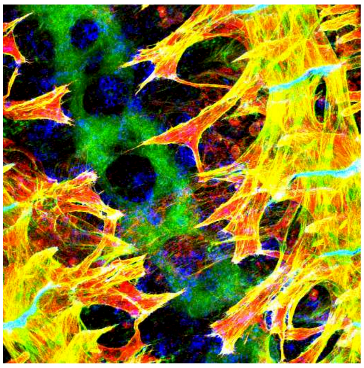
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Recent advances in biomedical imaging and signal analysis



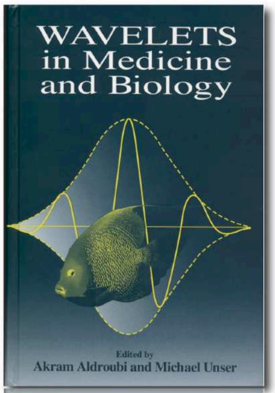
Michael Unser

Biomedical Imaging Group (BIG)
EPFL, Lausanne, Switzerland

Inaugural Lecture, EUSIPCO'10, Aalborg, Denmark, August 23-27, 2010

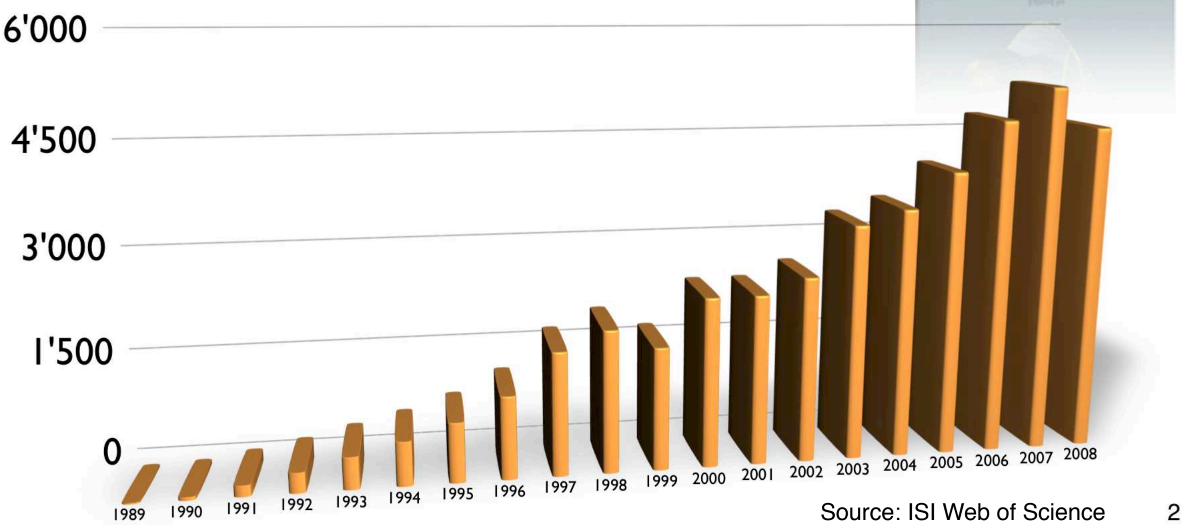
Wavelets in bioimaging

- Importance of wavelets in #publications
  - Overview articles:
    - Unser and Aldroubi, *Proc IEEE*, 1996
    - Laine, *Annual Rev Biomed Eng*, 2000
    - Special issue, *IEEE Trans Med Im*, 2003
    - Van De Ville et al., *IEEE EMB Mag,* 2006



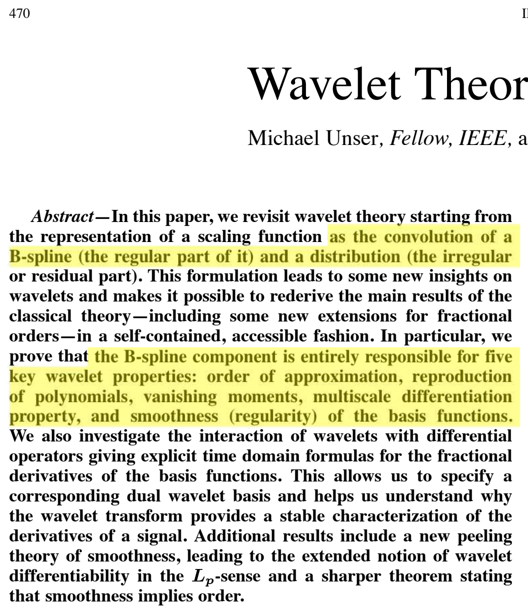
WAVELETS
in Medicine
and Biology

Edited by
Akram Aldroubi and Michael Unser



Source: ISI Web of Science

What about splines ?



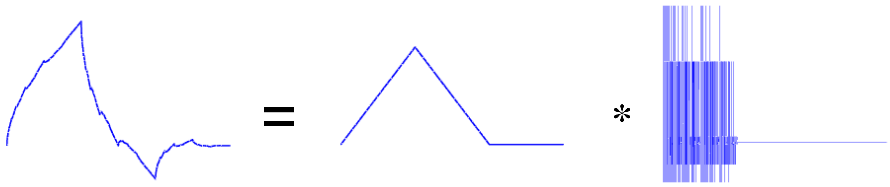
470 IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 51, NO. 2, FEBRUARY 2003

Wavelet Theory Demystified

Michael Unser, *Fellow, IEEE*, and Thierry Blu, *Member, IEEE*

***Abstract*—In this paper, we revisit wavelet theory starting from the representation of a scaling function as the convolution of a B-spline (the regular part of it) and a distribution (the irregular or residual part). This formulation leads to some new insights on wavelets and makes it possible to rederive the main results of the classical theory—including some new extensions for fractional orders—in a self-contained, accessible fashion. In particular, we prove that the B-spline component is entirely responsible for five key wavelet properties: order of approximation, reproduction of polynomials, vanishing moments, multiscale differentiation property, and smoothness (regularity) of the basis functions. We also investigate the interaction of wavelets with differential operators giving explicit time domain formulas for the fractional derivatives of the basis functions. This allows us to specify a corresponding dual wavelet basis and helps us understand why the wavelet transform provides a stable characterization of the derivatives of a signal. Additional results include a new peeling theory of smoothness, leading to the extended notion of wavelet differentiability in the L_p-sense and a sharper theorem stating that smoothness implies order.**

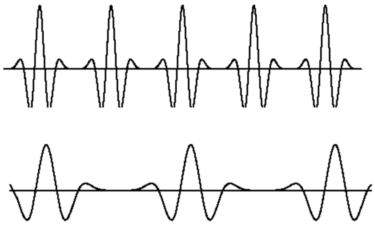
Wavelets are B-splines convolved by (nasty) distributions



CONTENT

- Wavelets and sparsity
  - Image denoising
  - Wavelet-regularized image reconstruction
- Wavelets revisited (in multiple dimensions)
  - Non-separable
  - Directional, **steerable**
  - Derivative-like (gradient, Hessian, ...)
  - Shape diversity, signal-adaptation

Wavelet basis of L_2



- Family of wavelet templates (basis functions)

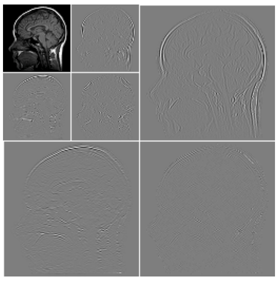
$$\psi_{i,k}(x) = 2^{-i/2}\psi\left(\frac{x-2^i k}{2^i}\right)$$

- Orthogonal wavelet basis

$$\langle \psi_{i,k}, \psi_{j,l} \rangle = \delta_{i-j,k-l} \qquad \Leftrightarrow \qquad \mathbf{W}^{-1} = \mathbf{W}^T$$

Analysis: $w_i[k] = \langle f, \psi_{i,k} \rangle$ (wavelet coefficients)

Reconstruction: $\forall f(x) \in L_2(\mathbb{R}), \quad f(x) = \sum_{i\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} w_i[k]\ \psi_{i,k}(x)$



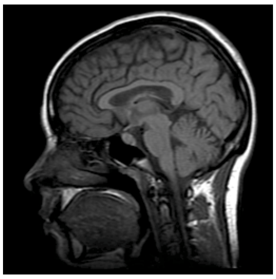
- Vector/matrix notation

Discrete signal: $\mathbf{f} = (\cdots, c[0], c[1], c[2], \cdots)$

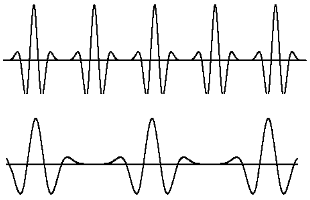
Wavelet coefficients: $\mathbf{w} = (\cdots, w_1[0], w_1[1], \cdots, w_2[0], \cdots)$

Analysis formula: $\mathbf{w} = \mathbf{W}^T\mathbf{f}$

Synthesis formula: $\mathbf{f} = \mathbf{W}\mathbf{w} = \sum_n w_n \boldsymbol{\psi}_n$

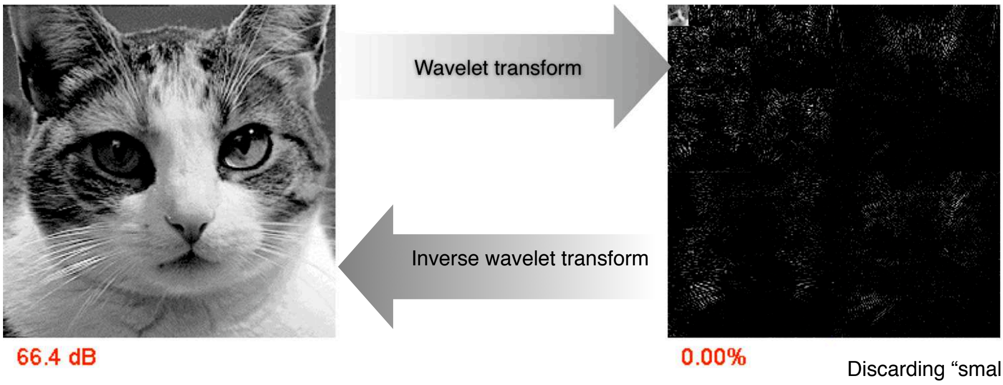


Sparsity of wavelet decomposition: example



Space-domain representation: $\mathbf{f}$

Wavelet-domain representation: $\mathbf{w} = \mathbf{W}^T\mathbf{f}$

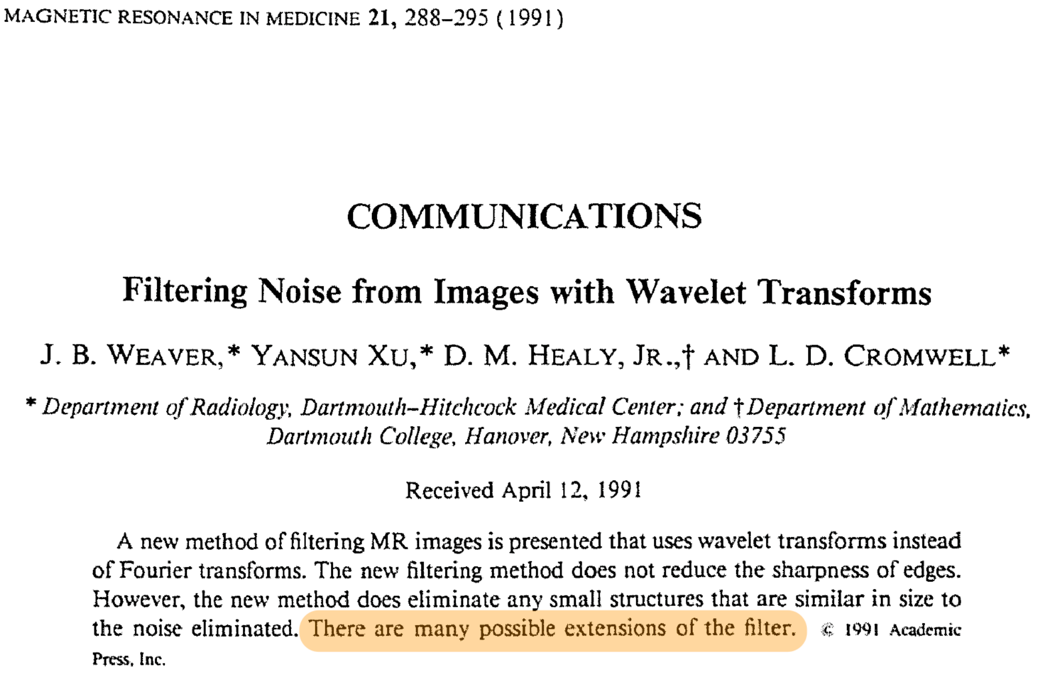


Discarding “small coefficients”

Reconstruction: $\mathbf{f}_N = \mathbf{W}\mathbf{w}_N$

Thresholding: $\mathbf{w} \rightarrow \mathbf{w}_N$

First published paper on biomedical applications



MAGNETIC RESONANCE IN MEDICINE **21**, 288–295 (1991)

COMMUNICATIONS

Filtering Noise from Images with Wavelet Transforms

J. B. WEAVER,* YANSUN XU,* D. M. HEALY, JR.,† AND L. D. CROMWELL*

** Department of Radiology, Dartmouth–Hitchcock Medical Center; and †Department of Mathematics, Dartmouth College, Hanover, New Hampshire 03755*

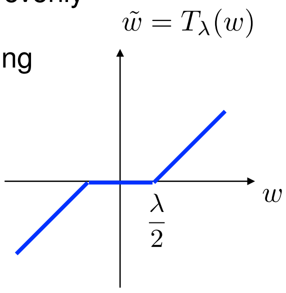
Received April 12, 1991

A new method of filtering MR images is presented that uses wavelet transforms instead of Fourier transforms. The new filtering method does not reduce the sharpness of edges. However, the new method does eliminate any small structures that are similar in size to the noise eliminated. There are many possible extensions of the filter. © 1991 Academic Press, Inc.

Denoising by wavelet thresholding

- Basic idea
  - Orthogonal WT: white noise → white noise
  - Signal is concentrated in few coefficients, while noise is spread-out evenly

$\Rightarrow$ Noise attenuation is achieved by simple wavelet shrinkage/thresholding



- Who gets credit ?
  - The celebrated statistician

    D.L. Donoho, “De-noising by soft-thresholding,” *IEEE Trans. Information Theory*, vol. 41, no. 3, pp. 613-627, May 1995. (>2000 ISI citations)
  - The pioneers

    B. Weaver, X. Yansun, D.M. Healy Jr., and L.D. Cromwell, “Filtering noise from images with wavelet transforms,” *Magnet. Reson. in Med.*, vol. 21, no. 2, pp. 288-295, 1991.

Denoising and wavelet regularization

- Measurement model

Space domain　　Wavelet domain

$$\mathbf{y} = \mathbf{f} + \mathbf{n} \quad \Leftrightarrow \quad w_i[\boldsymbol{k}] = \underset{\nwarrow\ \text{signal}}{s_i[\boldsymbol{k}]} + \underset{\nwarrow\ \text{noise}}{n_i[\boldsymbol{k}]}$$

(additive white noise)

- Signal estimation
  - Reconstruction formula: $\tilde{\mathbf{f}} = \mathbf{W}\tilde{\mathbf{w}}$ (inverse wavelet transform)
  - Data term: $\|\mathbf{y} - \tilde{\mathbf{f}}\|_2^2 = \|\mathbf{w} - \tilde{\mathbf{w}}\|_2^2$ (Parseval)
  - Regularization functional: $R(\tilde{f}) = R(\tilde{\mathbf{w}}) = \|\tilde{\mathbf{w}}\|_{\ell_1} = \sum_i \sum_{\boldsymbol{k}} |\tilde{w}_i[\boldsymbol{k}]| \sim \|\tilde{f}\|_{B_1^1(L_1(\mathbb{R}^2))}$

Optimization problem: $\tilde{\mathbf{w}}_0 = \arg\min_{\tilde{\mathbf{w}}} \left\{ \|\mathbf{w} - \tilde{\mathbf{w}}\|_2^2 + \lambda \|\tilde{\mathbf{w}}\|_{\ell_1} \right\}$

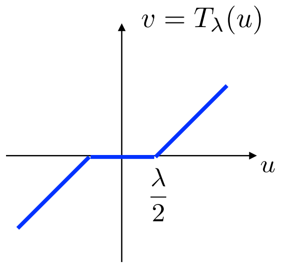
Wavelet-domain solution

- Equivalent convex optimization problem (decoupled)

$$\tilde{w}_0 = \arg\min_{\tilde{w}} \left\{ \sum_i \sum_{\boldsymbol{k}} |w_i[\boldsymbol{k}] - \tilde{w}_i[\boldsymbol{k}]|^2 + \lambda_i |\tilde{w}_i[\boldsymbol{k}]| \right\}$$

- Basic scalar optimization problem

minimize $J(u,v) = (v-u)^2 + \lambda |v|$



- Soft-threshold solution

Moreau's proximal operator: $\mathrm{Prox}_\varphi(u) = \arg\min_{v \in \mathbb{R}} \left\{ \frac{1}{2}(u-v)^2 + \varphi(v) \right\}$

$$v = \mathrm{Prox}_{2\lambda|\cdot|}(u) = T_\lambda(u) = \begin{cases} u - \lambda/2, & \lambda/2 < u \\ 0, & |u| \leq \lambda/2 \\ u + \lambda/2, & u < -\lambda/2 \end{cases}$$

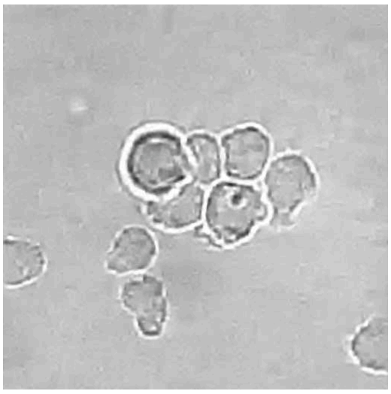
(Moreau, 1965; Chambolle et al., *IEEE Trans. Image Proc.*, 1998; Combette, 2005)

BIG extension: SURE-LET

- Key features of SURE-LET wavelet denoising algorithm
  - Generalized non-linearities: Linear Expansion of Thresholds:
    $T_\lambda(u) \quad \rightarrow \quad \sum_{k=1}^{K} a_k f_k(u)$
  - Optimizes thresholding parameters a_k from noisy data using Stein's Unbiased Risk Estimate (SURE)
  - Incorporates inter-scale dependencies via prediction tree
  - Improved performance:
    - 1 to 1.5 dB better than basic soft thresholding
    - Very close to oracle performance
    - Outperforms standard Wiener filter

(Luisier et al., *IEEE Trans. Image Proc.* , 2007)

SURE-LET Demo

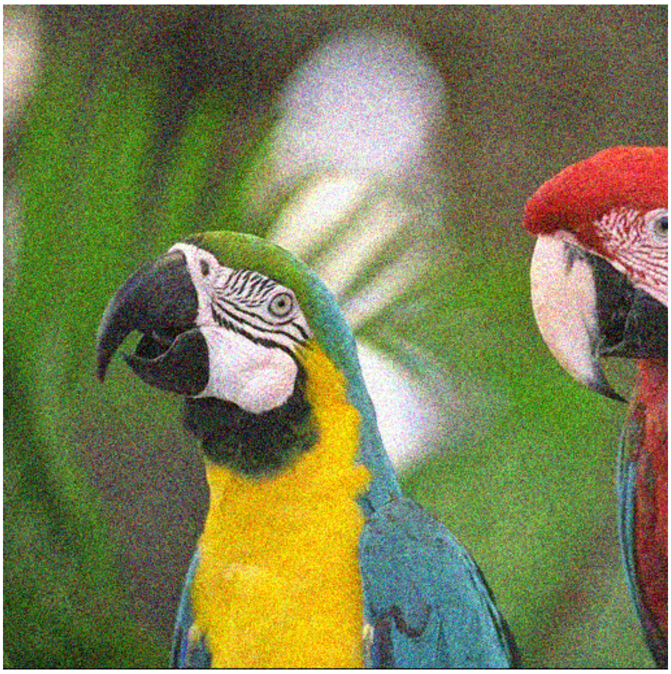


SNR improvement: + 15.73 dB

IEEE

2009 Young Author Best Paper Award
IEEE Signal Processing Society

Standard Color Image



Input PSNR=18.59 dB

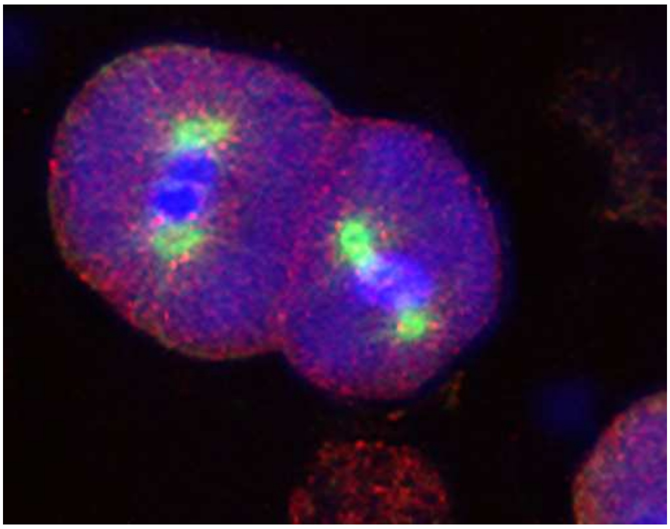
Denoised with OWT SURE-LET

Output PSNR = 31.91 dB

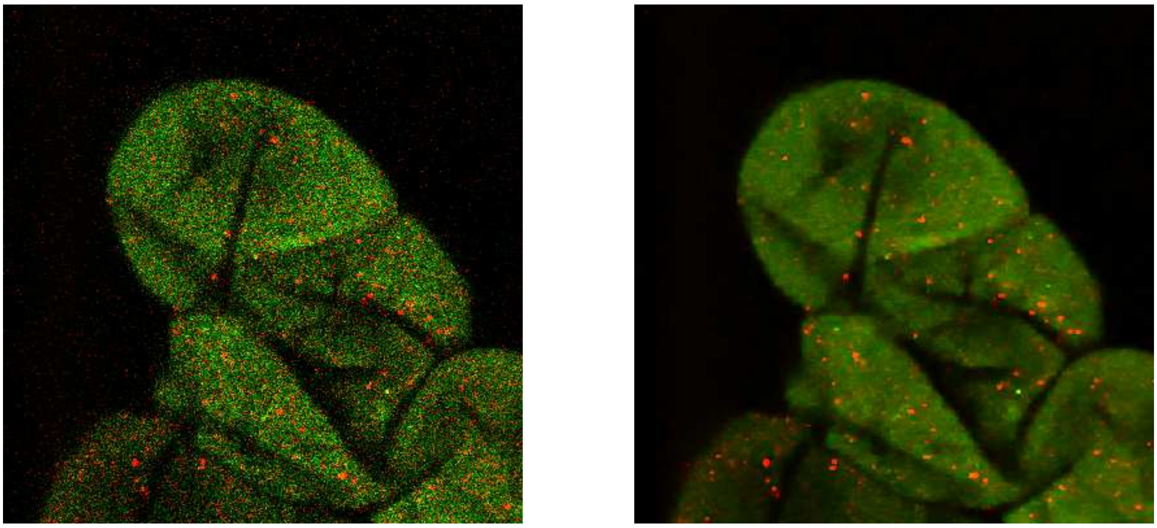
Denoised with UWT SURE-LET

Output PSNR = 33.27 dB

2D SURE-LET denoising (UWT): C-elegance embryo

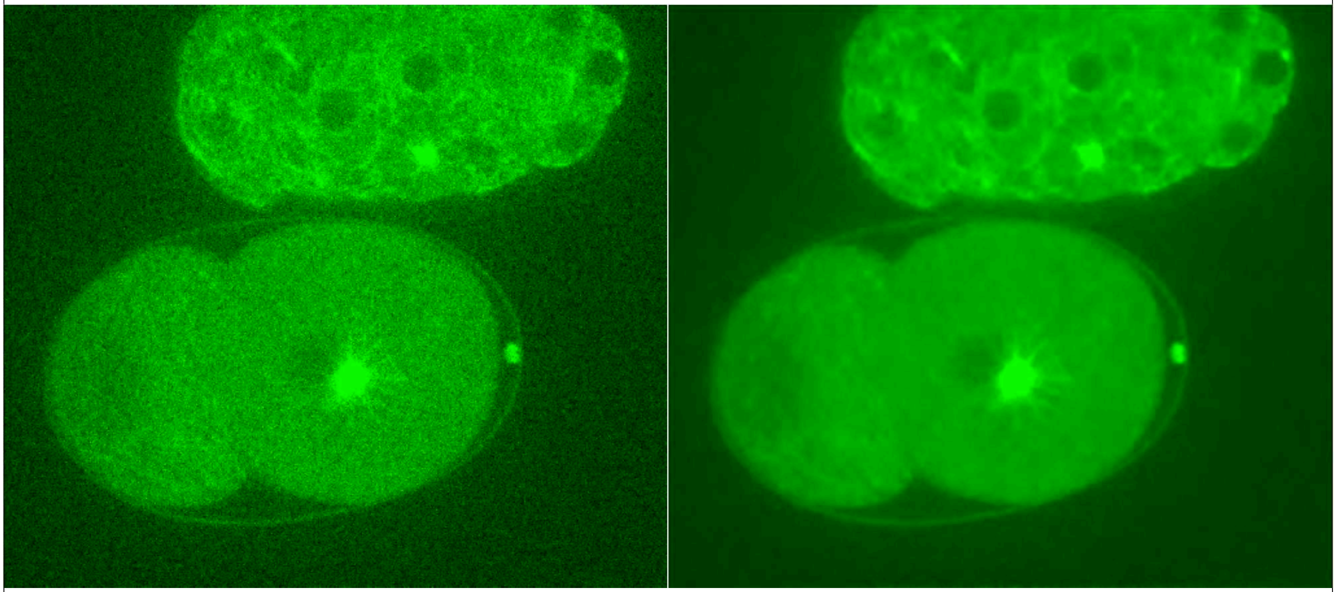


2D SURE-LET denoising (UWT): Tobacco cells



Ground truth
(average over 500 acquisitions)

2D + time SURE-LET denoising (DWT) : C-elegance embryo



Wavelet-regularized image reconstruction

- Space-domain measurement model

  $\mathbf{y} = \mathbf{H}\mathbf{f} + \mathbf{n}$

$\mathbf{H}$: system matrix (e.g., convolution)

$\mathbf{n}$: additive noise component

- Wavelet-regularized signal recovery
  - Wavelet expansion of signal: $\tilde{\mathbf{f}} = \mathbf{W}\tilde{\mathbf{w}}$
  - Data term: $\|\mathbf{y} - \mathbf{H}\tilde{\mathbf{f}}\|_2^2 = \|\mathbf{y} - \mathbf{H}\mathbf{W}\tilde{\mathbf{w}}\|_2^2$
  - Wavelet-domain sparsity constraint: $\|\tilde{\mathbf{w}}\|_{\ell_1} \leq C_1$

Convex optimization problem

$$\tilde{\mathbf{w}} = \arg\min_{\mathbf{w}} \left\{ \|\mathbf{y} - \mathbf{A}\mathbf{w}\|_2^2 + \lambda \|\mathbf{w}\|_{\ell_1} \right\} \quad \text{with } \mathbf{A} = \mathbf{H}\mathbf{W}$$

or

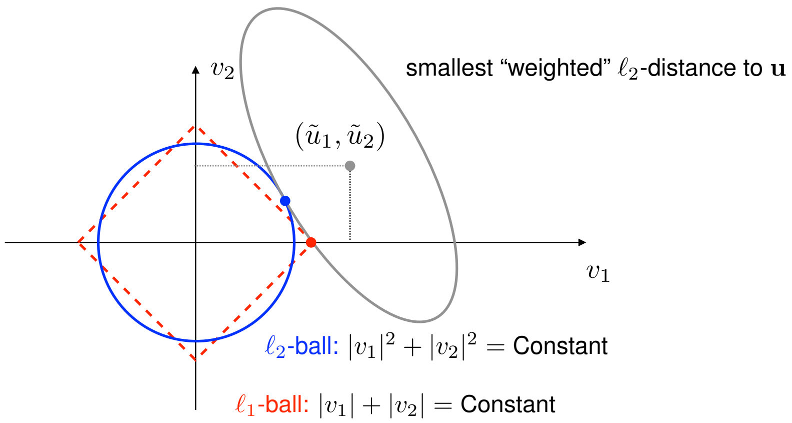
$$\tilde{\mathbf{f}} = \arg\min_{\mathbf{f}} \left\{ \|\mathbf{y} - \mathbf{H}\mathbf{f}\|_2^2 + \lambda \|\mathbf{W}^T\mathbf{f}\|_{\ell_1} \right\}$$

Sparsity and l_1-minimization

- Prototypical inverse problem

$$\min_{\mathbf{v}} \left\{ \|\mathbf{u} - \mathbf{A}\mathbf{v}\|_{\ell_2}^2 + \lambda \|\mathbf{v}\|_{\ell_2}^2 \right\} \quad \Leftrightarrow \quad \min_{\mathbf{v}} \|\mathbf{u} - \mathbf{A}\mathbf{v}\|_{\ell_2}^2 \text{ subject to } \|\mathbf{v}\|_{\ell_2} = C_2$$

$$\min_{\mathbf{v}} \left\{ \|\mathbf{u} - \mathbf{A}\mathbf{v}\|_{\ell_2}^2 + \lambda \|\mathbf{v}\|_{\ell_1} \right\} \quad \Leftrightarrow \quad \min_{\mathbf{v}} \|\mathbf{u} - \mathbf{A}\mathbf{v}\|_{\ell_2}^2 \text{ subject to } \|\mathbf{v}\|_{\ell_1} = C_1$$


v_2
smallest "weighted" ℓ_2-distance to $\mathbf{u}$
$(\tilde{u}_1, \tilde{u}_2)$
v_1
ℓ_2-ball: $|v_1|^2 + |v_2|^2 =$ Constant
ℓ_1-ball: $|v_1| + |v_2| =$ Constant

Elliptical norm: $\|\mathbf{u} - \mathbf{A}\mathbf{v}\|_2^2 = (\mathbf{v} - \tilde{\mathbf{u}})^T \mathbf{A}^T \mathbf{A} (\mathbf{v} - \tilde{\mathbf{u}})$ with $\tilde{\mathbf{u}} = \mathbf{A}^{-1}\mathbf{u}$

Alternating minimization: ISTA algorithm

- Convex cost functional: $\mathcal{C}(\mathbf{f}) = \|\mathbf{y} - \mathbf{H}\mathbf{f}\|_2^2 + \lambda \|\mathbf{W}^T\mathbf{f}\|_1$

- Special cases
  - Classical least squares: $\lambda = 0 \quad \Rightarrow \quad \mathbf{f} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{y}$

    Landweber algorithm: $\mathbf{f}_{n+1} = \mathbf{f}_n + \gamma \mathbf{H}^T(\mathbf{y} - \mathbf{H}\mathbf{f}_n)$ (steepest descent)
  - Pure denoising: $\mathbf{H} = \mathbf{I} \quad \Rightarrow \quad \mathbf{f} = \mathbf{W}\, T_\lambda\{\mathbf{W}^T\mathbf{y}\}$

- Iterative Soft Thresholding Algorithm (ISTA)

(Figueiredo-Nowak, 2003)

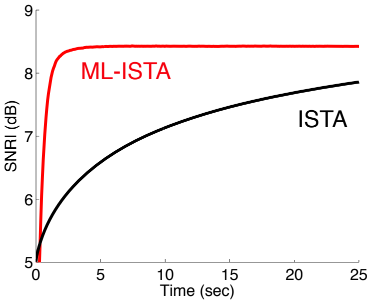
1. Initialization $(n \leftarrow 0), \mathbf{f}_0 = \mathbf{y}$
2. Landweber update: $\mathbf{z} = \mathbf{f}_n + \gamma \mathbf{H}^T(\mathbf{y} - \mathbf{H}\mathbf{f}_n)$
3. Wavelet denoising: $\mathbf{w} = \mathbf{W}^T\mathbf{z}$, $\quad \tilde{\mathbf{w}} = T_{\gamma\lambda}\{\mathbf{w}\}$ (soft threshold)
4. Signal update: $\mathbf{f}_{n+1} \leftarrow \mathbf{W}\tilde{\mathbf{w}}$ and repeat from Step 2 until convergence

Proof of convergence: (Daubechies, Defrise, De Mol, 2004)

Fast multilevel wavelet-regularized deconvolution

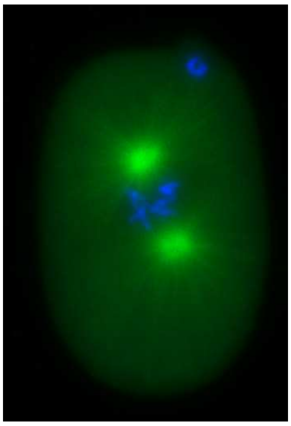
- Key features of multilevel wavelet deconvolution algorithm (ML-ISTA)
  - Subband **adaptive steps** (optimized for fast convergence)
  - Acceleration by one order of magnitude with respect to state-of-the art algorithm (ISTA) (multigrid iteration strategy)
  - Applicable in 2D or 3D:
    first wavelet attempt for the deconvolution of 3D fluorescence micrographs
  - Typically outperforms oracle Wiener solution (best linear algorithm)

9/7 wavelets

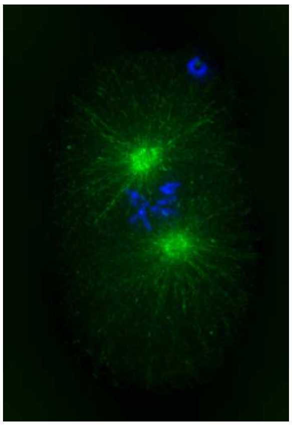


(Vonesch-U., *IEEE-IP*, 2009)

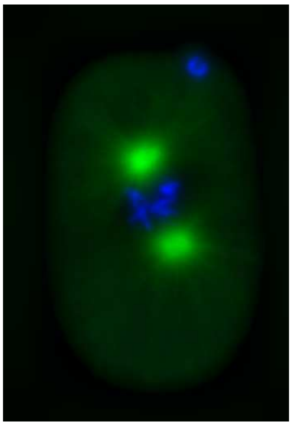
3D fluorescence microscopy experiment



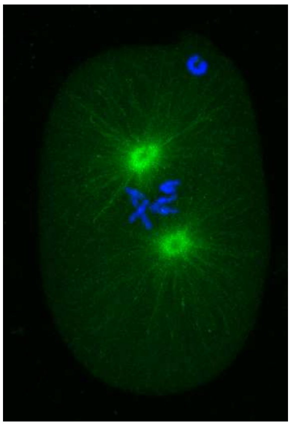
Input data
(open pinhole)



ML-ISTA 15 iterations



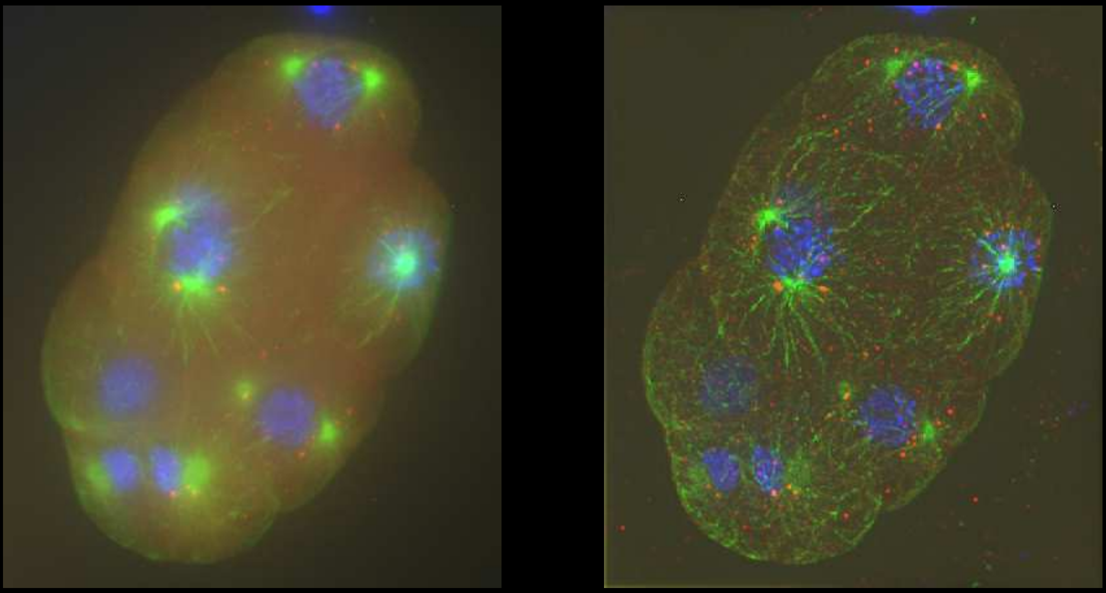
ISTA 15 iterations



Confocal reference

Maximum-intensity projections of 512×352×96 image stacks;
Zeiss LSM 510 confocal microscope with a 63× oil-immersion objective;
C. Elegans embryo labeled with Hoechst, Alexa488, Alexa568;
each channel processed separately; computed PSF based on diffraction-limited model;
separable orthonormalized linear-spline/Haar basis.

3D deconvolution of widefield stack



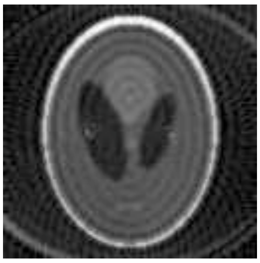
Maximum intensity projections of 384×448×260 image stacks;
Leica DM 5500 widefield epifluorescence microscope with a 63× oil-immersion objective;
C. Elegans embryo labeled with Hoechst, Alexa488, Alexa568;
each channel processed separately; computed PSF based on diffraction-limited model;
Haar basis, 3 decomposition levels for X-Y, 2 decomposition levels for Z.

Reconstruction results with parallel MRI

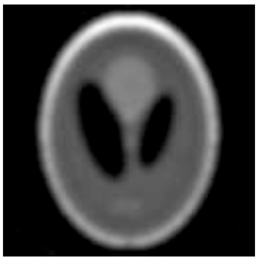
Simulated parallel MRI experiment (M. Guerquin-Kern, BIG)

Shepp-Logan brain phantom
4 coils, undersampled spiral acquisition, 15dB noise

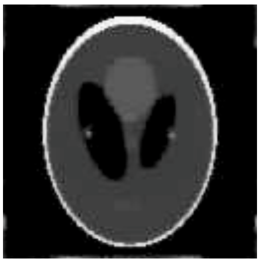
Space



Backprojection



L_2 regularization (CG)

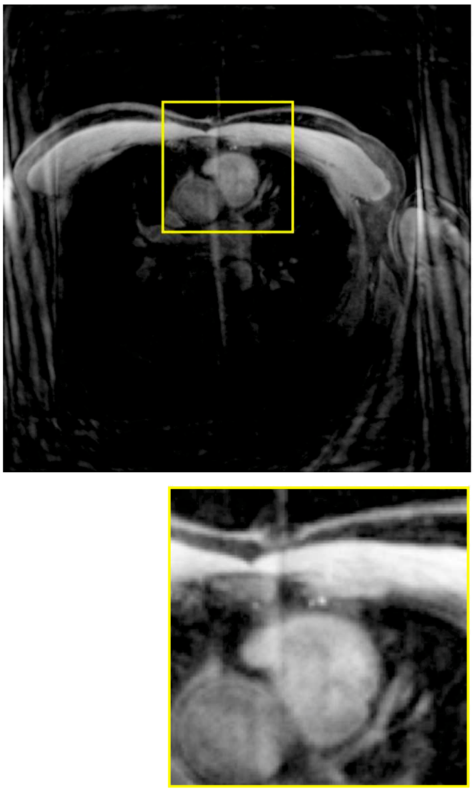


ℓ_1 wavelet regularization

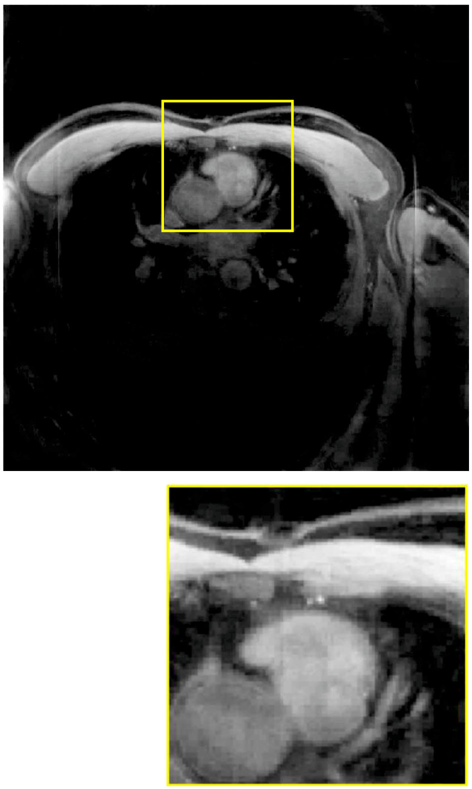
NCCBI collaboration with K. Prüssmann, ETHZ

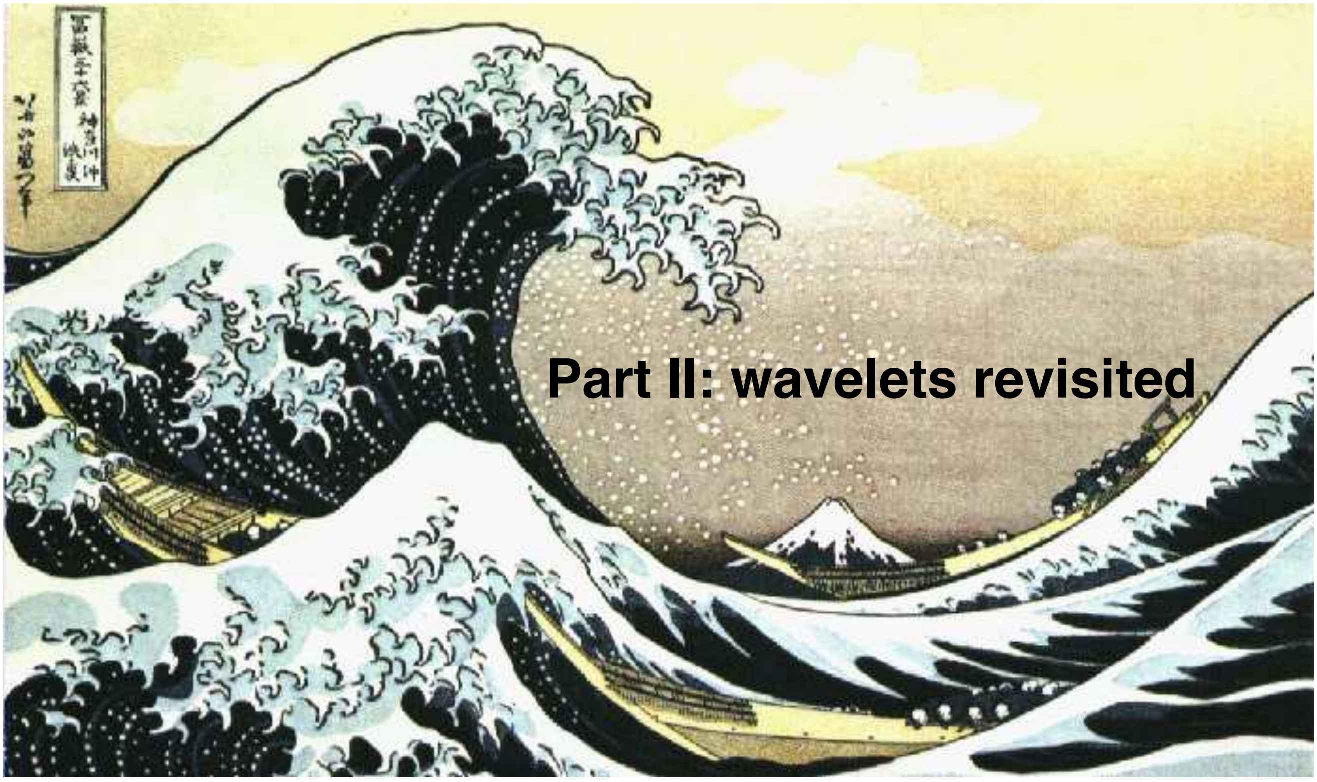
Try at ISMRM reconstruction challenge

L_2 regularization (Laplacian)



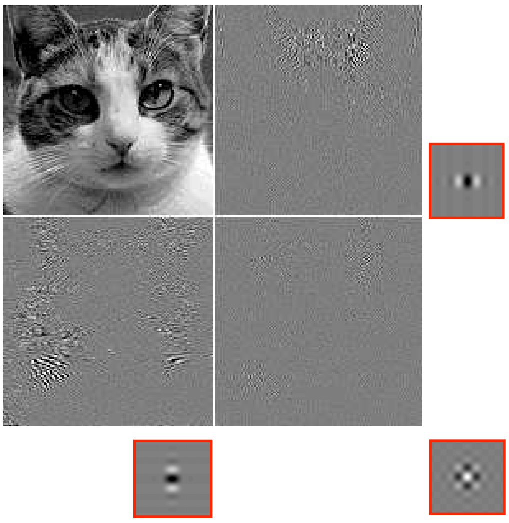
ℓ_1 wavelet regularization





Part II: wavelets revisited

Beyond separable wavelet representations



- Limitations of separable wavelets
  - Limited amount of invariance (in particular, to rotation)
  - Poor handling of directional features
  - Lack of proper differential interpretation
- Multidimensional alternatives
  - Wavelet frames for better shift, scale and rotation invariance
  - Curvelets, bandelets, contourlets, . . .
  - **Steerable pyramid**

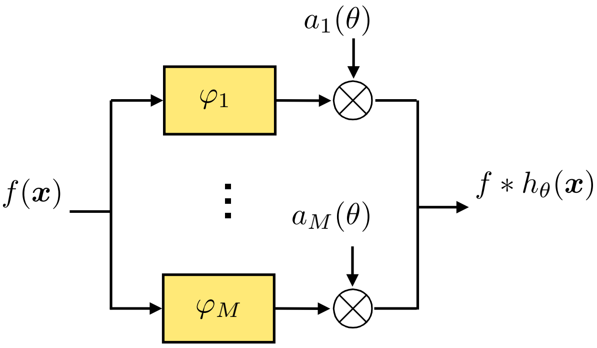
Steerable filters

(Freeman & Adelson, 1991)

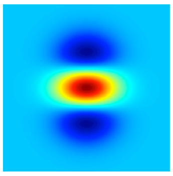
Definition. A 2D filter $h(\boldsymbol{x})$, $\boldsymbol{x} \in \mathbb{R}^2$ is steerable of order M iff. there exist some basis filters $\varphi_m(\boldsymbol{x})$ and coefficients $a_m(\theta)$ such that

$$\forall\theta \in [-\pi, \pi], \quad h_\theta(\boldsymbol{x}) := h(\boldsymbol{R}_\theta \boldsymbol{x}) = \sum_{m=1}^{M} a_m(\theta)\, \varphi_m(\boldsymbol{x})$$

- Fast filterbank implementation



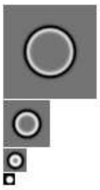
Optimized ridge detector (*M*=3)



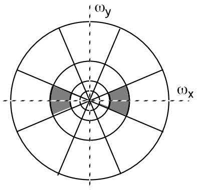
(Jacob-U., IEEE-PAMI, 2004)

Simoncelli’s steerable pyramid (1995)

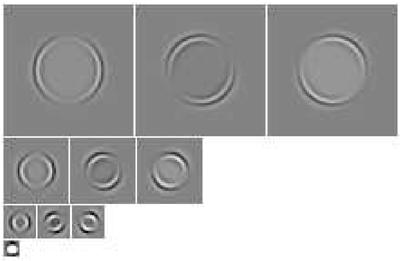
Isotropic wavelet pyramid

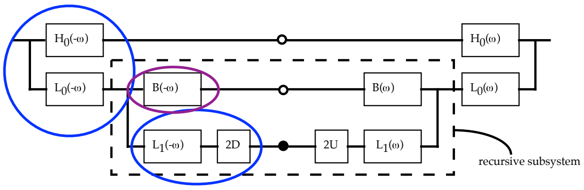


Multichannel polar filtering



Directional wavelet coefficients





- Many successful applications
  - Contour detection
  - Image filtering and denoising
  - Orientation analysis
  - Texture analysis and synthesis
- Limitations
  - **Fixed design**
  - Purely discrete framework (no functional counterpart)
  - **Does not extend to dimensions higher than two**

Riesz transform

- Definition: $$\boldsymbol{\mathcal{R}} f(\boldsymbol{x}) = \begin{pmatrix} \mathcal{R}_1 f(\boldsymbol{x}) \\ \vdots \\ \mathcal{R}_d f(\boldsymbol{x}) \end{pmatrix} \quad \xleftrightarrow{\mathcal{F}} \quad -j \frac{\boldsymbol{\omega}}{\|\boldsymbol{\omega}\|} \hat{f}(\boldsymbol{\omega})$$

Multi-dimensional Fourier transform

$$\hat{f}(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} f(\boldsymbol{x}) e^{-j\langle \boldsymbol{\omega}, \boldsymbol{x} \rangle} \mathrm{d}x_1 \cdots \mathrm{d}x_d$$

with $\boldsymbol{\omega} = (\omega_1, \dots, \omega_d) \in \mathbb{R}^d$

- Multi-channel convolution

$$\mathcal{R}_n f(\boldsymbol{x}) = (h_n * f)(\boldsymbol{x}) \quad \text{with} \quad h_n = \mathcal{R}_n\{\delta\} \quad \xleftrightarrow{\mathcal{F}} \quad -j \frac{\omega_n}{\|\boldsymbol{\omega}\|}$$

- Riesz transform and partial derivatives

$$\boldsymbol{\mathcal{R}} f(\boldsymbol{x}) = (-1)(-\Delta)^{-\frac{1}{2}} \boldsymbol{\nabla} f(\boldsymbol{x})$$

“Smoothed version of gradient”

$$\boldsymbol{\nabla} f(\boldsymbol{x}) = -\boldsymbol{\mathcal{R}}(-\Delta)^{\frac{1}{2}} f(\boldsymbol{x})$$

Reversibility of the Riesz transform

- Adjoint operator

$$\boldsymbol{\mathcal{R}}^*\boldsymbol{r}(\boldsymbol{x}) = \mathcal{R}_1^* r_1(\boldsymbol{x}) + \cdots + \mathcal{R}_d^* r_d(\boldsymbol{x}) \quad \xleftrightarrow{\mathcal{F}} \quad j\frac{\boldsymbol{\omega}^T}{\|\boldsymbol{\omega}\|}\hat{\boldsymbol{r}}(\boldsymbol{\omega})$$

- Self-reversibility

$$\boldsymbol{\mathcal{R}}^*\boldsymbol{\mathcal{R}}f(\boldsymbol{x}) = \sum_{i=1}^{d} \mathcal{R}_i^*\mathcal{R}_i f(\boldsymbol{x}) = f(\boldsymbol{x})$$

- What about iterating ?
  - Combining Nth-order components of the form $\mathcal{R}_{i_1}\mathcal{R}_{i_2}\cdots\mathcal{R}_{i_N} f$ with $i_1, \cdots i_N \in \{1, \cdots, d\}$
  - n-fold iteration: $\mathcal{R}_i^n = \mathcal{R}_i\mathcal{R}_i^{n-1}$ with $\mathcal{R}_i^0 = \mathrm{Id}$

Higher-order Riesz transform

Theorem (Decomposition of the identity)

$$\sum_{\substack{n_1,\dots,n_d \geq 0 \\ n_1+\cdots+n_d = N}} \frac{N!}{n_1!\,n_2!\cdots n_d!} \left(\mathcal{R}_1^{n_1}\cdots\mathcal{R}_d^{n_d}\right)^* \left(\mathcal{R}_1^{n_1}\cdots\mathcal{R}_d^{n_d}\right) = \mathrm{Id}$$

(U. - Van De Ville, *TIP* 2010)

- Proper definition of Nth-order transform

$M = \binom{N+d-1}{d-1}$ distinct Riesz components with $n_1 + \cdots + n_d = N$

$$\boldsymbol{\mathcal{R}}^{(N)} f(\boldsymbol{x}) = \begin{pmatrix} \mathcal{R}^{(N,0,\cdots,0)} f(\boldsymbol{x}) \\ \vdots \\ \mathcal{R}^{(n_1,\cdots,n_d)} f(\boldsymbol{x}) \\ \vdots \\ \mathcal{R}^{(0,\cdots,0,N)} f(\boldsymbol{x}) \end{pmatrix} \quad \text{where} \quad \mathcal{R}^{(n_1,\dots,n_d)} = \sqrt{\frac{N!}{n_1!\cdots n_d!}}\,\mathcal{R}_1^{n_1}\cdots\mathcal{R}_d^{n_d}$$

Multi-index notation

Multi-index: $\boldsymbol{n} = (n_1, \ldots, n_d)$ with $n_1, \ldots, n_d \in \mathbb{Z}^+$

- Sum of components: $|\mathbf{n}| = \sum_{i=1}^{d} n_i = N$
- Factorial: $\mathbf{n}! = n_1!\, n_2! \cdots n_d!$
- Exponentiation of a vector $\boldsymbol{z} = (z_1, \ldots, z_d) \in \mathbb{C}^d$: $\boldsymbol{z}^{\mathbf{n}} = z_1^{n_1} \cdots z_d^{n_d}$

Decomposition of the identity: $\forall \psi \in L_2(\mathbb{R}^d), \; \sum_{|\mathbf{n}|=N} (\mathcal{R}^{\mathbf{n}})^* \mathcal{R}^{\mathbf{n}} \psi = \psi$

$$\mathcal{R}^{(n_1,\ldots,n_d)}\psi(\boldsymbol{x}) = \mathcal{R}^{\mathbf{n}}\psi(\boldsymbol{x}) \quad \overset{\mathcal{F}}{\longleftrightarrow} \quad \sqrt{\frac{|\mathbf{n}|!}{\mathbf{n}!}}\, \frac{(-j\boldsymbol{\omega})^{\mathbf{n}}}{\|\boldsymbol{\omega}\|^{|\mathbf{n}|}}\, \hat{\psi}(\boldsymbol{\omega})$$

Properties of higher-order Riesz transform

- Shift invariance: $\forall \boldsymbol{x}_0 \in \mathbb{R}^d, \quad \boldsymbol{\mathcal{R}}^{(N)}\{f(\cdot - \boldsymbol{x}_0)\}(\boldsymbol{x}) = \boldsymbol{\mathcal{R}}^{(N)}\{f(\cdot)\}(\boldsymbol{x} - \boldsymbol{x}_0)$
- Scale invariance: $\forall a \in \mathbb{R}^+, \quad \boldsymbol{\mathcal{R}}^{(N)}\{f(\cdot/a)\}(\boldsymbol{x}) = \boldsymbol{\mathcal{R}}^{(N)}\{f(\cdot)\}(\boldsymbol{x}/a)$
- Parseval-like identity: $\forall f, \phi \in L_2(\mathbb{R}^d)$

$$\begin{aligned} \langle \boldsymbol{\mathcal{R}}^{(N)} f, \boldsymbol{\mathcal{R}}^{(N)} \phi \rangle_{L_2} &= \sum_{|\mathbf{n}|=N} \langle \mathcal{R}^{\mathbf{n}} f, \mathcal{R}^{\mathbf{n}} \phi \rangle_{L_2} \\ &= \langle f, \phi \rangle_{L_2} \end{aligned}$$

Energy conservation: $\|\boldsymbol{\mathcal{R}}^{(N)} f\|_{L_2}^2 = \sum_{|\mathbf{n}|=N} \|\mathcal{R}^{\mathbf{n}} f\|_{L_2}^2 = \|f\|_{L_2}^2$

Steerability of higher-order Riesz transform

$\mathcal{R}^{\mathbf{n}}\{\delta\}(\boldsymbol{x})$: impulse response of $\mathbf{n}$-component Riesz operator

$\mathbf{R} = (\mathbf{r}_1 \cdots \mathbf{r}_d)^T$: $d \times d$ spatial rotation matrix e.g., $\mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ for $d = 2$

- Steerability of Nth-order Riesz transform
  - Rotated version of $\mathbf{n}$-component impulse response

$$\mathcal{R}^{\mathbf{n}}\{\delta\}(\mathbf{R}\boldsymbol{x}) = \sum_{|\mathbf{m}|=N} s_{\mathbf{n},\mathbf{m}}(\mathbf{R})\, \mathcal{R}^{\mathbf{m}}\{\delta\}(\boldsymbol{x})$$

steering coefficients

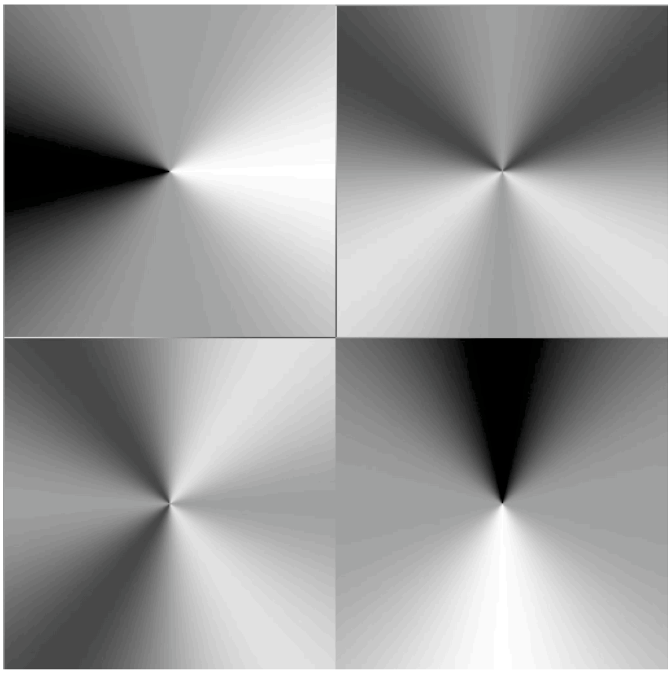
unrotated impulse responses of nth-order Riesz transform

  - Explicit form of steering coefficients

$$s_{\mathbf{n},\mathbf{m}}(\mathbf{R}) = \sqrt{\frac{\mathbf{m}!}{\mathbf{n}!}} \sum_{|\mathbf{k}_1|=n_1} \cdots \sum_{|\mathbf{k}_d|=n_d} \delta_{\mathbf{k}_1+\cdots+\mathbf{k}_d,\mathbf{m}} \frac{\mathbf{n}!}{\mathbf{k}_1! \cdots \mathbf{k}_d!} \mathbf{r}_1^{\mathbf{k}_1} \cdots \mathbf{r}_d^{\mathbf{k}_d}$$

  - The steering coefficients specify a group of orthogonal matrices of size $M = \binom{N+d-1}{d-1}$

Visualization in the frequency domain



$N = 3$

Frame = redundant extension of a basis

- Definition
  - A family of functions $\{\psi_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\mathbb{Z}^d}$ is called a frame of $L_2(\mathbb{R}^d)$ iff.

    $$\forall f \in L_2(\mathbb{R}^d), \quad A\,\|f\|_{L_2}^2 \leq \sum_{\boldsymbol{k}\in\mathbb{Z}^d} |\langle \psi_{\boldsymbol{k}}, f\rangle_{L_2}|^2 \leq B\,\|f\|_{L_2}^2$$
  - Tight frame: $A = B$
  - Parseval frame: $A = B = 1$

- Analysis/synthesis formula
  - $\forall f \in L_2(\mathbb{R}^d), \quad f = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} \langle \psi_{\boldsymbol{k}}, f\rangle_{L_2}\, \tilde{\psi}_{\boldsymbol{k}}$
  - $\{\tilde{\psi}_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\mathbb{Z}^d}$: dual frame (minimum-norm inverse)
  - Parseval frame: $\tilde{\psi}_{\boldsymbol{k}} = \psi_{\boldsymbol{k}}$

Construction of steerable wavelet frames

(U. - Van De Ville, 2010)

- Wavelet frame of $L_2(\mathbb{R}^d)$

$$\forall f \in L_2(\mathbb{R}^d), \quad f(\boldsymbol{x}) = \sum_{i\in\mathbb{Z}} \sum_{\boldsymbol{k}\in\mathbb{Z}^d} \langle f, \psi_{i,\boldsymbol{k}}\rangle_{L_2}\, \tilde{\psi}_{i,\boldsymbol{k}}(\boldsymbol{x})$$

Wavelet property: $\psi_{i,\boldsymbol{k}}(\boldsymbol{x}) = 2^{-\frac{id}{2}}\,\psi_{0,\boldsymbol{k}}(\boldsymbol{x}/2^i)$

Multi-index: $\boldsymbol{n} = (n_1, \dots, n_d)$

Theorem

Let $\{\psi_{i,\boldsymbol{k}}\}$ be a primal wavelet frame of $L_2(\mathbb{R}^d)$. Then, $\{\psi^{\boldsymbol{n}}_{i,\boldsymbol{k}} = \mathcal{R}^{\boldsymbol{n}}\psi_{i,\boldsymbol{k}}\}_{|\boldsymbol{n}|=N}$ and $\{\tilde{\psi}^{\boldsymbol{n}}_{i,\boldsymbol{k}} = \mathcal{R}^{\boldsymbol{n}}\tilde{\psi}_{i,\boldsymbol{k}}\}_{|\boldsymbol{n}|=N}$ form a dual set of wavelet frames such that

$$\forall f \in L_2(\mathbb{R}^d), \quad f(\boldsymbol{x}) = \sum_{i\in\mathbb{Z}} \sum_{\boldsymbol{k}\in\mathbb{Z}^d} \sum_{|\boldsymbol{n}|=N} \langle f, \psi^{\boldsymbol{n}}_{i,\boldsymbol{k}}\rangle_{L_2}\, \tilde{\psi}^{\boldsymbol{n}}_{i,\boldsymbol{k}}(\boldsymbol{x})$$

Justification

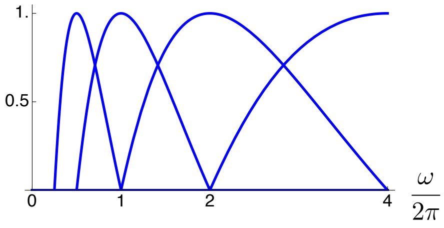
Inner product preservation $\Rightarrow$ $\langle \psi_{i,\boldsymbol{k}}, \psi_{i',\boldsymbol{k}'}\rangle_{L_2} = \langle \boldsymbol{\mathcal{R}}^{(N)}\psi_{i,\boldsymbol{k}}, \boldsymbol{\mathcal{R}}^{(N)}\psi_{i',\boldsymbol{k}'}\rangle_{L_2}$

Shift and scale invariance $\Rightarrow$ $\mathcal{R}^{\boldsymbol{n}}\psi_{i,\boldsymbol{k}}(\boldsymbol{x}) = 2^{-\frac{id}{2}}\,\psi^{\boldsymbol{n}}(\boldsymbol{x}/2^i - \boldsymbol{k})$ with $\psi^{\boldsymbol{n}} = \mathcal{R}^{\boldsymbol{n}}\psi$

Backbone: primal isotropic wavelet pyramid

- Frequency domain design of band-limited wavelets

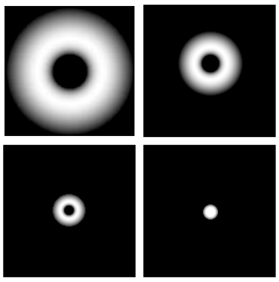
Radial wavelet filters

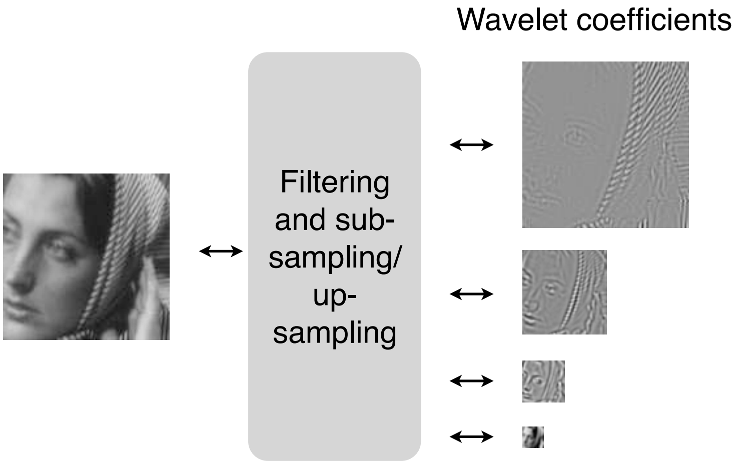


Tight frame property:

$$\sum_{i\in\mathbb{Z}} |\hat{\psi}(\omega/2^i)|^2 = 1$$

2D frequency view





Wavelet coefficients

Filtering and sub-sampling/ up-sampling

Perfect isotropy $\psi(\boldsymbol{x}) = \psi(\|\boldsymbol{x}\|)$ no preferred direction

---

Differential interpretation of Riesz wavelets

- Frequency-domain wavelet formula:

$$\widehat{\psi^{\mathbf{n}}}(\boldsymbol{\omega}) = \sqrt{\frac{N!}{\mathbf{n}!}} \frac{(-j\boldsymbol{\omega})^{\mathbf{n}}}{\|\boldsymbol{\omega}\|^N} \hat{\psi}(\boldsymbol{\omega}) \quad \propto \quad (j\omega_1)^{n_1} \cdots (j\omega_d)^{n_d} \frac{\hat{\psi}(\boldsymbol{\omega})}{\|\boldsymbol{\omega}\|^N}$$

- Isotropic smoothing kernel: $\phi_N(\boldsymbol{x}) = (-\Delta)^{-\frac{N}{2}} \psi(\boldsymbol{x}) = \mathcal{F}^{-1}\left\{\frac{\hat{\psi}(\boldsymbol{\omega})}{\|\boldsymbol{\omega}\|^N}\right\}$

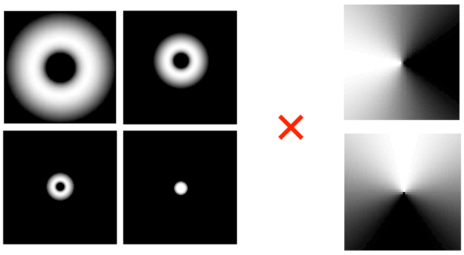
- Space-domain wavelet formula:

$$\psi^{\mathbf{n}}(\boldsymbol{x}) = \mathcal{R}^{\mathbf{n}}\psi(\boldsymbol{x}) \quad \propto \quad \frac{\partial^N}{\partial x_1^{n_1} \cdots \partial x_d^{n_d}} \phi_N(\boldsymbol{x}),$$

$$\langle f, \psi^{\mathbf{n}}(\cdot - \boldsymbol{x})\rangle \quad \propto \quad \frac{\partial^N}{\partial x_1^{n_1} \cdots \partial x_d^{n_d}} (f * \phi_N)(\boldsymbol{x})$$

First-order Riesz wavelet transform

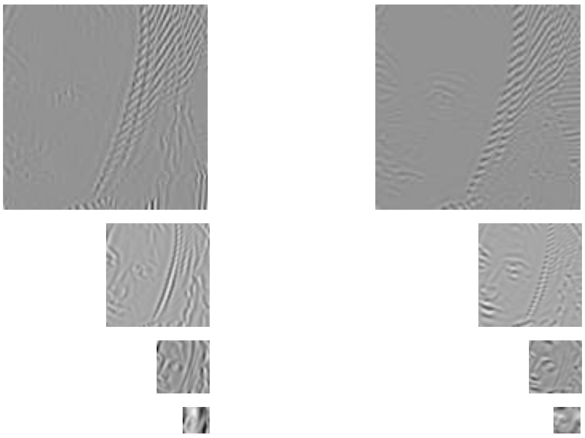
2D frequency view



isotropic pyramid

Riesz transform

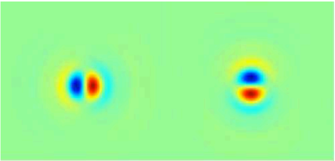
Riesz wavelet coefficients



vertical features selectivity

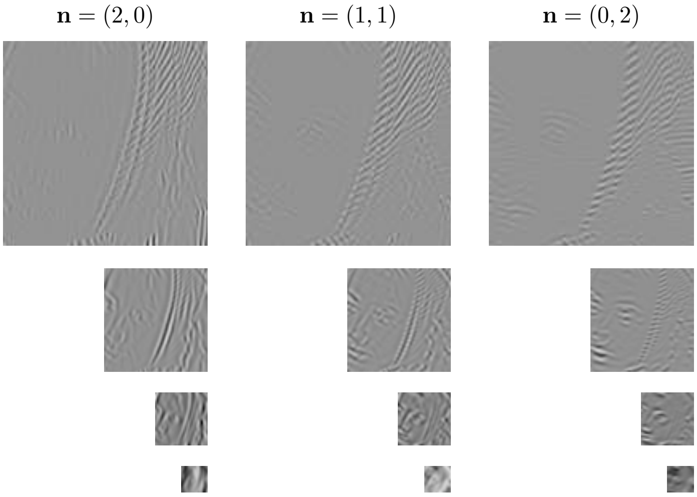
horizontal features selectivity

Steerable Gradient-like wavelets

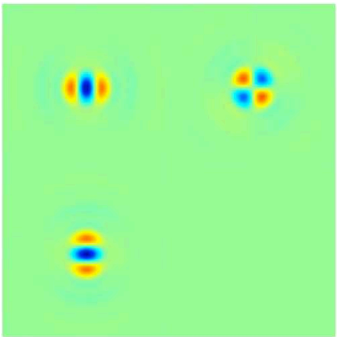


$$\left(\psi^{(1,0)} = \frac{\partial \phi_1}{\partial x_1}, \quad \psi^{(0,1)}(\boldsymbol{x}) = \frac{\partial \phi_1}{\partial x_2}\right)$$

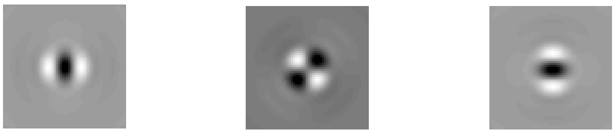
Hessian-like Riesz wavelet transform



Steerability



Second-order Riesz wavelets



$$\left(\psi^{(2,0)} = \frac{\partial^2 \phi_2}{\partial x_1^2}, \quad \psi^{(1,1)} = \sqrt{2}\frac{\partial^2 \phi_2}{\partial x_1 \partial x_2}, \quad \psi^{(0,2)} = \frac{\partial^2 \phi_2}{\partial x_2^2}\right)$$

Generalized Riesz wavelets

- Steerable wavelet subspace
  - ψ: primary isotropic bandlimited wavelet
  - The Riesz wavelets $\{\mathcal{R}^{\mathbf{n}}\psi\}_{|\mathbf{n}|=N}$ span a steerable subspace of dimension $M = \binom{N+d-1}{d-1}$
  - There are many other wavelet bases that spans the same subspace

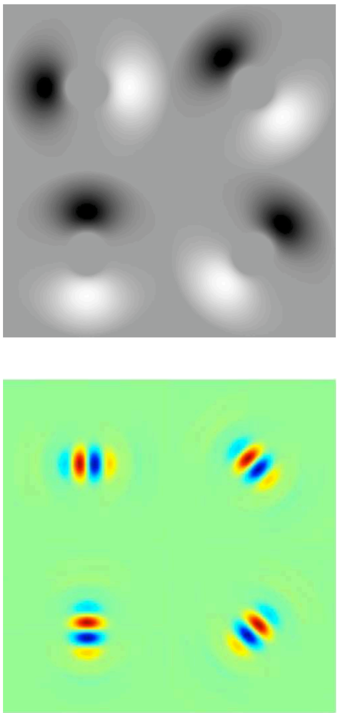
- Generalized Riesz wavelets
  - Parametrized by a $M \times M$ non-singular shaping matrix $\mathbf{U}$
  - Generalized $\mathbf{n}$-component wavelet: $\tilde{\psi}^{\mathbf{n}}_{i,\boldsymbol{k}} = \sum_{|\mathbf{m}|=N} u_{\mathbf{n},\mathbf{m}} \mathcal{R}^{\mathbf{m}} \psi_{i,\boldsymbol{k}}$

- Special case: Simoncelli's equiangular design (2-D only)

$$[\mathbf{U}_{\text{Simon}}]_{m+1,n+1} = \sqrt{\binom{N}{m}} \cos\left(\frac{\pi n}{N+1}\right)^{m} \sin\left(\frac{\pi n}{N+1}\right)^{N-m}$$

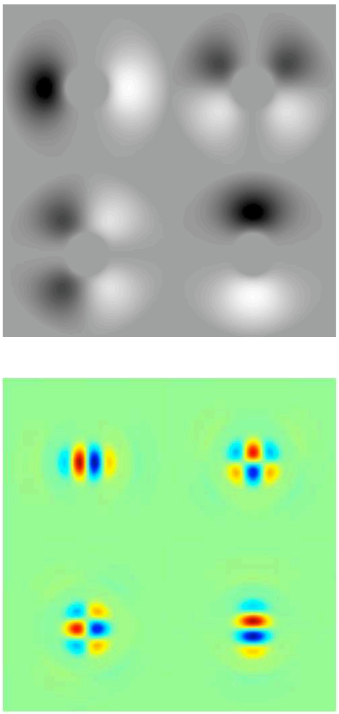
$$m, n \in \{0, \dots, N\}$$

Equi-angular vs. Riesz wavelets



Frequency domain

Space domain

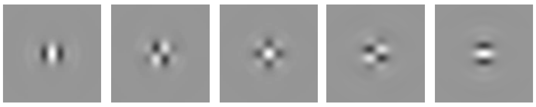


Simoncelli's 4-channel steerable pyramid

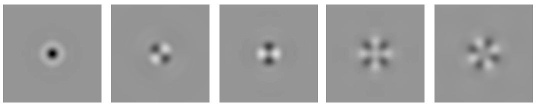
Riesz wavelets $(N = 3)$

Riesz and equalized PCA wavelets

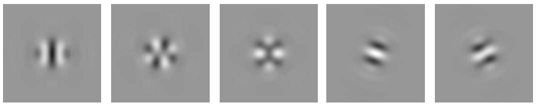
(a) Riesz wavelets



(b) Equalized Riesz wavelets

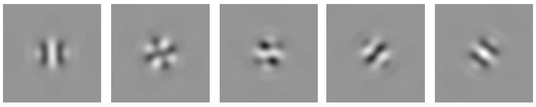


(c) Lena (N=4)



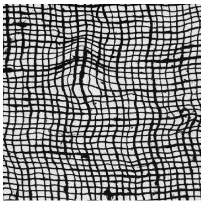
72% 12% 6.3% 5% 4.7%

(d) Texture D104 (N=4)

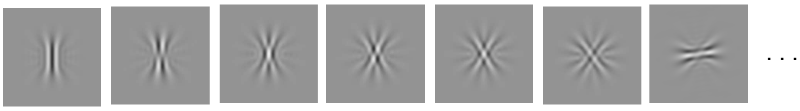


50% 15.6% 13.7% 11.4% 9.4%

D104 fragment



(f) Barbara (N=14)



75.6% 8.3% 4.9% 2.8% 1.8% 1.2% 0.9%

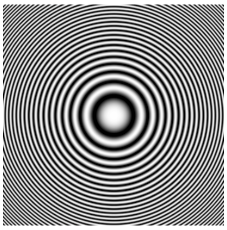
Basic denoising benchmark

Wavelet domain soft-thresholding with optimized λ for max SNR
Steering is the same in all cases

	order	Barbara			Lena			zoneplate		
σ initial PSNR		10 28.11	20 22.10	30 18.59	10 28.13	20 22.11	30 18.59	10 28.14	20 22.12	30 18.59
Equi-Angular	2	31.43	27.33	25.22	33.63	30.24	28.47	34.03	29.02	26.08
	3	31.57	27.47	25.38	33.68	30.30	28.54	34.56	29.54	26.58
	4	31.72	27.60	25.50	33.77	30.37	28.58	34.92	29.86	26.91
	5	31.81	27.69	25.57	33.76	30.38	28.58	35.08	30.10	27.14
Riesz	2	31.68	27.44	25.29	33.68	30.27	28.47	35.06	29.83	28.47
	3	31.86	27.67	25.48	33.76	30.34	28.53	35.44	30.22	26.78
	4	32.03	27.86	25.66	33.89	30.47	28.64	35.79	30.55	27.16
	5	32.09	27.95	25.74	33.88	30.46	28.63	35.94	30.72	27.49
Equalized Riesz	2	30.85	26.63	24.58	33.14	29.57	27.83	32.79	27.51	24.83
	3	31.09	26.94	24.64	33.25	30.06	28.19	33.18	27.76	24.74
	4	31.02	26.83	24.76	33.28	29.71	27.95	32.95	27.69	25.07
	5	31.06	26.97	24.70	33.37	30.12	28.23	33.19	27.95	24.86
PCA	2	31.65	27.33	25.14	33.59	30.14	28.33	35.00	29.70	26.62
	3	31.75	27.44	25.19	33.58	30.11	28.30	35.27	29.96	26.86
	4	31.86	27.53	25.27	33.64	30.15	28.32	35.46	30.12	27.03
	5	31.87	27.55	25.25	33.59	30.10	28.28	35.54	30.20	27.11

Equalized PCA	2	**31.80**	**27.61**	**25.41**	**33.75**	**30.32**	**28.50**	**35.32**	**30.07**	**27.03**
	3	**32.05**	**27.92**	**25.70**	**33.84**	**30.40**	**28.56**	**35.87**	**30.64**	**27.58**
	4	**32.23**	**28.17**	**25.94**	**33.99**	**30.53**	**28.67**	**36.25**	**31.10**	**28.01**
	5	**32.31**	**28.29**	**26.04**	**33.98**	**30.52**	**28.66**	**36.40**	**31.33**	**28.29**

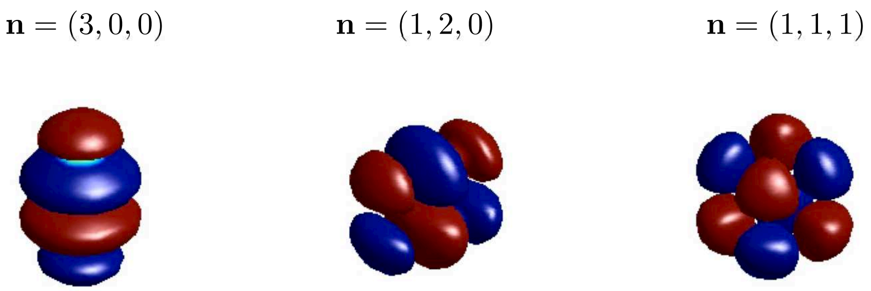
+0.2 - 0.6 dB +0.05 - 0.2 dB +1.0 - 1.3 dB



zoneplate

Examples of 3-D steerable wavelets

- Third-order wavelets in 3-D



$\mathbf{n} = (3, 0, 0)$ $\mathbf{n} = (1, 2, 0)$ $\mathbf{n} = (1, 1, 1)$

iso-surface representation of wavelets in space domain

3-D work in progress

CONCLUSION

- **Sparsity** as a powerful design paradigm
  - Denoising by simple wavelet-domain processing (non-linear)
  - Compressed sensing / sparse signal recovery
  - Wavelet-regularized image reconstruction
- General operator-based design of **steerable** wavelets
  - Decoupled **multiresolution** and **multiorientation** properties
  - Simplicity of implementation (FFT, multirate filterbank)
  - Tight frame property
  - Extended class of partial derivative/Riesz wavelets
  - Adaptivity
- Novel perspectives for wavelet-domain image processing
  - Rotation-invariant processing/feature extraction
  - Learning the wavelet dictionary
  - Steerable wavelets in 3D
    Good potential for biomedical imaging (MRI, confocal microscopy)

Acknowledgments

Many thanks to

- Prof. Thierry Blu
- Prof. Dimitri Van De Ville
- Dr. Daniel Sage
- Dr. Cédric Vonesch
- Mathieu Guerquin-Kern
- Dr. Florian Luisier
- Prof. Klaas Prüssmann
- Dr. Nicolas Chenouard

+ many other researchers, and graduate students

EPFL's Biomedical Imaging Group

- Preprints and demos: http://bigwww.epfl.ch/

Bibliography

http://bigwww.epfl.ch/

- Wavelet denoising and restoration
  - F. Luisier, T. Blu, M. Unser, "A New SURE Approach to Image Denoising: Interscale Orthonormal Wavelet Thresholding," *IEEE Trans. Image Processing*, vol. 16, no. 3, pp. 593-606, March 2007.
  - I. Daubechies, M. Defrise, and C. De Mol, "An Iterative Thresholding Algorithm for Linear Inverse Problems with a Sparsity Constraint," *Comm. Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, August 2004.
  - C. Vonesch, M. Unser, "A Fast Multilevel Algorithm for Wavelet-Regularized Image Restoration," *IEEE Trans. Image Processing*, vol. 18, no. 3, pp. 509-523, 2009.
- Steerable wavelets
  - E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable multiscale transforms," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 587–607, 1992.
  - M. Unser, D. Van De Ville, "Wavelet steerability and the higher-order Riesz transform," *IEEE Trans. Image Processing*, vol. 19, no. 3, pp. 636-652, March 2010.
  - M. Unser, N. Chenouard, and D. Van De Ville, "Steerable pyramids and tight wavelet frames of $L_2(\mathbb{R}^d)$," under review.
- Review and tutorial
  - M. Unser, A. Aldroubi, "A Review of Wavelets in Biomedical Applications," *Proc. of the IEEE*, vol. 84, no. 4, pp. 626-638, April 1996.
  - M. Unser, T. Blu, "Wavelet Theory Demystified," *IEEE Trans. Sig. Proc.*, vol. 51, no. 2, pp. 470-483, 2003.
  - M. Unser, M. Unser, "Wavelet Games," *Wavelet Digest*, vol. 11, no. 4, April 1, 2003. **Lego connection at**: http://www.wavelet.org/phpBB2/viewtopic.php?t=5129

First-order steering mechanism (max energy)

- Gradient-like wavelet transform

  $$\mathbf{w}_i[\boldsymbol{k}] = \left(\langle f, \psi_{i,\boldsymbol{k}}^{(1,0,\cdots,0)}\rangle, \ldots, \langle f, \psi_{i,\boldsymbol{k}}^{(0,\cdots,0,1)}\rangle\right)$$

  Wavelet projection along unit vector $\mathbf{u}$: $w_{\mathbf{u},i}[\boldsymbol{k}] = \langle \mathbf{u}, \mathbf{w}_i[\boldsymbol{k}]\rangle$

- Pointwise orientation: $\mathbf{u} = \dfrac{\mathbf{w}_i[\boldsymbol{k}]}{\|\mathbf{w}_i[\boldsymbol{k}]\|} \quad \Leftrightarrow \quad w_{\mathbf{u},i}^2[\boldsymbol{k}]$ maximum

- Orientation within a neighborhood
  - Local Gaussian-like window: $v[\boldsymbol{k}] \geq 0$
  - Local wavelet energy at $(i, \boldsymbol{k}_0)$ along direction $\mathbf{u}$

    $$E_{\mathbf{u}} = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} v[\boldsymbol{k}-\boldsymbol{k}_0]\; w_{\mathbf{u},i}^2[\boldsymbol{k}] = \mathbf{u}^T \mathbf{J}_{i,\boldsymbol{k}_0}\mathbf{u}$$

  - Max energy orientation: $\mathbf{u}_1 = \arg \max_{\mathbf{u}\in\mathbb{R}^d, \|\mathbf{u}\|=1} \{\mathbf{u}^T \mathbf{J}_{i,\boldsymbol{k}_0}\mathbf{u}\}$

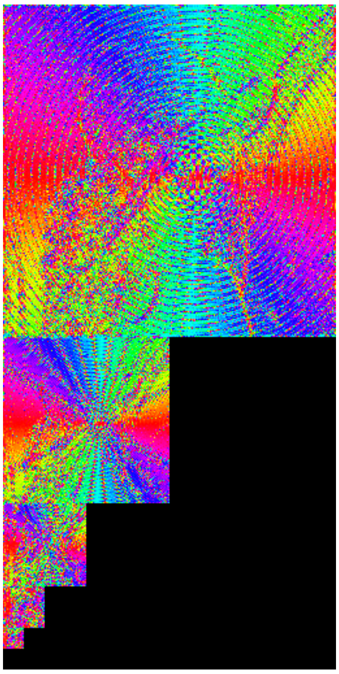
    $\Rightarrow$ $\mathbf{u}_1$ = first eigenvector of $\mathbf{J}_{i,\boldsymbol{k}_0}$

Wavelet structure tensor

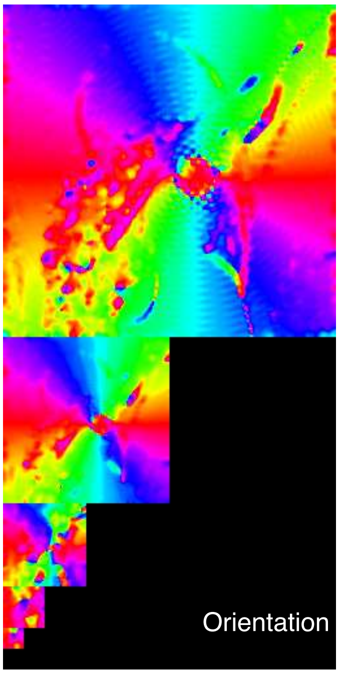
$$\mathbf{J}_{i,\boldsymbol{k}_0} = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} v[\boldsymbol{k}-\boldsymbol{k}_0]\,\mathbf{w}_i[\boldsymbol{k}]\mathbf{w}_i^T[\boldsymbol{k}]$$

Pointwise vs. tensor-based steering

Psychedelic Lena



Pointwise orientation



tensor orientation