Applied Harmonic Analysis Methods in Imaging Science
Introduction

Gitta Kutyniok
(Technische Universität Berlin)

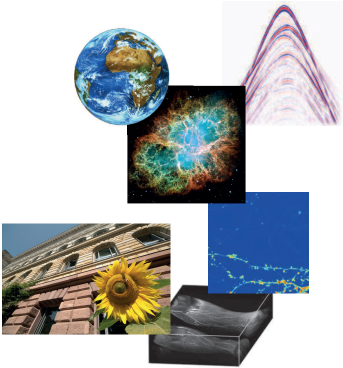
SIAM Conference on Imaging Science
Albuquerque, May 23 – 26, 2016

Imaging Science Today

Due to the data deluge, the area of imaging science is of tremendous importance in today's world.

Main Tasks

- Acquisition
- Preprocessing
  - Denoising, Inpainting, ...
- Analysis
  - Feature Detection, ...
- Storing
  - Compression, ...

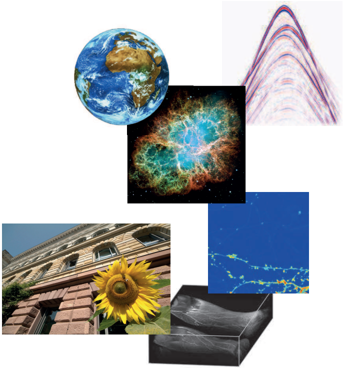


Imaging Science Today

Due to the data deluge, the area of imaging science is of tremendous importance in today's world.

Main Tasks

- Acquisition
- Preprocessing
  - Denoising, Inpainting, ...
- Analysis
  - Feature Detection, ...
- Storing
  - Compression, ...



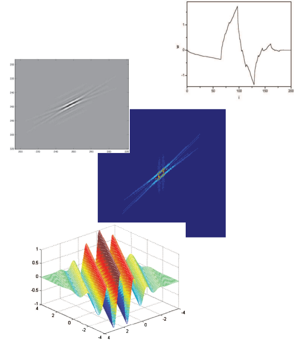
What has Applied Harmonic Analysis to offer?

Applied Harmonic Analysis

Representation systems designed by Applied Harmonic Analysis concepts have established themselves as a standard tool in applied mathematics, computer science, and engineering.

Examples:

- Wavelets.
- Ridgelets.
- Curvelets.
- Shearlets.
- ...



Key Property:

Fast Algorithms combined with Sparse Approximation Properties!

An Applied Harmonic Analysis Viewpoint

Exploit a carefully designed representation system $(\psi_\lambda)_{\lambda \in \Lambda} \subseteq \mathcal{H}$:

$$\mathcal{H} \ni f \longrightarrow (\langle f, \psi_\lambda \rangle)_{\lambda \in \Lambda} \longrightarrow \sum_{\lambda \in \Lambda} \langle f, \psi_\lambda \rangle \, \psi_\lambda = f.$$

Desiderata:

- Special features encoded in the "large" coefficients $| \langle f, \psi_\lambda \rangle |$.
- Efficient representations:

$$f \approx \sum_{\lambda \in \Lambda_N} \langle f, \psi_\lambda \rangle \, \psi_\lambda, \quad \#(\Lambda_N) \text{ small}$$

Goals:

- Modification of the coefficients according to the task.
- Derive high compression by considering only the "large" coefficients.

Two Main Viewpoints

Decomposition:

$$\mathcal{H} \ni f \longrightarrow (\langle f, \psi_\lambda \rangle)_{\lambda \in \Lambda}.$$

- Preprocessing (e.g. denoising).
- Analysis (e.g. feature detection).
- Clustering/Classification.
- ...

Efficient Representations:

$$f = \sum_{\lambda \in \Lambda} c_\lambda \psi_\lambda.$$

- Compression.
- Regularization of inverse problems.
- Ansatz functions for PDE solvers.
- ...

Sparsity

Novel Paradigm:

For each class of data, there exists a sparsifying system!

Sparsity

Novel Paradigm:

For each class of data, there exists a sparsifying system!

Two Viewpoints of 'Sparsifying System':

Let $\mathcal{C} \subseteq \mathcal{H}$ and $(\psi_\lambda)_\lambda \subseteq \mathcal{H}$.

- Decay of Coefficients. Consider the decay for $n \to \infty$ of the sorted sequence of coefficients

$$(|\langle x, \psi_{\lambda_n} \rangle|)_n \quad \text{for all } x \in \mathcal{C}.$$

- Approximation Properties. Consider the decay for $N \to \infty$ of the error of best N-term approximation, i.e.,

$$\inf_{\#\Lambda_N = N, (c_\lambda)_\lambda} \left\| x - \sum_{\lambda \in \Lambda_N} c_\lambda \psi_\lambda \right\| \quad \text{for all } x \in \mathcal{C}.$$

Sparsifying System

Functional Analytic Properties:

- $(\psi_\lambda)_\lambda$ can be an orthonormal basis.
- $(\psi_\lambda)_\lambda$ can form a frame, i.e., there exist $0 < A \leq B < \infty$ with

$$A\|x\|^2 \leq \sum_\lambda |\langle x, \psi_\lambda \rangle|^2 \leq B\|x\|^2 \quad \text{for all } x \in \mathcal{H}.$$

Sparsifying System

Functional Analytic Properties:

- $(\psi_\lambda)_\lambda$ can be an orthonormal basis.
- $(\psi_\lambda)_\lambda$ can form a frame, i.e., there exist $0 < A \le B < \infty$ with

$$A\|x\|^2 \le \sum_\lambda |\langle x, \psi_\lambda\rangle|^2 \le B\|x\|^2 \quad \text{for all } x \in \mathcal{H}.$$

Basic Facts about Frames:

- The frame operator $S : \mathcal{H} \to \mathcal{H}$, $Sx = \sum_\lambda \langle x, \psi_\lambda\rangle \psi_\lambda$ is invertible.
- The dual frame $(\tilde{\psi}_\lambda)_\lambda := (S^{-1}\psi_\lambda)_\lambda$ yields

$$x = \sum_\lambda \langle x, \psi_\lambda\rangle \tilde{\psi}_\lambda = \sum_\lambda \langle x, \tilde{\psi}_\lambda\rangle \psi_\lambda \quad \text{for all } x \in \mathcal{H}.$$

Sparsifying System

Functional Analytic Properties:

- $(\psi_\lambda)_\lambda$ can be an orthonormal basis.
- $(\psi_\lambda)_\lambda$ can form a frame, i.e., there exist $0 < A \leq B < \infty$ with

$$A\|x\|^2 \leq \sum_\lambda |\langle x, \psi_\lambda \rangle|^2 \leq B\|x\|^2 \quad \text{for all } x \in \mathcal{H}.$$

Basic Facts about Frames:

- The frame operator $S : \mathcal{H} \to \mathcal{H}$, $Sx = \sum_\lambda \langle x, \psi_\lambda \rangle \psi_\lambda$ is invertible.
- The dual frame $(\tilde{\psi}_\lambda)_\lambda := (S^{-1}\psi_\lambda)_\lambda$ yields

$$x = \sum_\lambda \langle x, \psi_\lambda \rangle \tilde{\psi}_\lambda = \sum_\lambda \langle x, \tilde{\psi}_\lambda \rangle \psi_\lambda \quad \text{for all } x \in \mathcal{H}.$$

Some Advantages of Redundancy:

- Flexibility in expansions $x = \sum_\lambda c_\lambda \psi_\lambda$.
- Robustness against loss of coefficients $\langle x, \psi_\lambda \rangle$.

Notion of Optimality

Two Viewpoints of Optimality of $(\psi_\lambda)_\lambda$: Let $\mathcal{C} \subseteq \mathcal{H}$.

- Decay of Coefficients. $\beta > 0$ is largest (for all systems) with

$$|\langle x, \psi_{\lambda_n}\rangle| \lesssim n^{-\beta} \text{ as } n \to \infty, \quad \text{for all } x \in \mathcal{C}.$$

- Approximation Properties. $\gamma > 0$ is largest (for all systems) with

$$\inf_{\#\Lambda_N = N, (c_\lambda)_\lambda} \Big\| x - \sum_{\lambda \in \Lambda_N} c_\lambda \psi_\lambda \Big\| \lesssim N^{-\gamma} \text{ as } N \to \infty, \quad \text{for all } x \in \mathcal{C}.$$

Notion of Optimality

Two Viewpoints of Optimality of $(\psi_\lambda)_\lambda$: Let $\mathcal{C} \subseteq \mathcal{H}$.

- Decay of Coefficients. $\beta > 0$ is largest (for all systems) with

$$|\langle x, \psi_{\lambda_n} \rangle| \lesssim n^{-\beta} \text{ as } n \to \infty, \quad \text{for all } x \in \mathcal{C}.$$

- Approximation Properties. $\gamma > 0$ is largest (for all systems) with

$$\inf_{\#\Lambda_N = N, (c_\lambda)_\lambda} \Big\| x - \sum_{\lambda \in \Lambda_N} c_\lambda \psi_\lambda \Big\| \lesssim N^{-\gamma} \text{ as } N \to \infty, \quad \text{for all } x \in \mathcal{C}.$$

Situation of an ONB: For the best N-term approximation x_N of x, we have

$$\|x - x_N\|^2 = \sum_{\lambda \notin \Lambda_N} |c_\lambda|^2 = \sum_{n > N} |\langle x, \psi_{\lambda_n} \rangle|^2$$

Notion of Optimality

Two Viewpoints of Optimality of $(\psi_\lambda)_\lambda$: Let $\mathcal{C} \subseteq \mathcal{H}$.

- Decay of Coefficients. $\beta > 0$ is largest (for all systems) with

$$|\langle x, \psi_{\lambda_n}\rangle| \lesssim n^{-\beta} \text{ as } n \to \infty, \quad \text{for all } x \in \mathcal{C}.$$

- Approximation Properties. $\gamma > 0$ is largest (for all systems) with

$$\inf_{\#\Lambda_N = N, (c_\lambda)_\lambda} \left\| x - \sum_{\lambda \in \Lambda_N} c_\lambda \psi_\lambda \right\| \lesssim N^{-\gamma} \text{ as } N \to \infty, \quad \text{for all } x \in \mathcal{C}.$$

Situation of an ONB: For the best N-term approximation x_N of x, we have

$$\|x - x_N\|^2 = \sum_{\lambda \notin \Lambda_N} |c_\lambda|^2 = \sum_{n > N} |\langle x, \psi_{\lambda_n}\rangle|^2$$

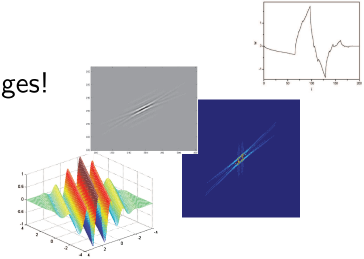
Situation of a Frame: For the N-term approximation $x_N = \sum_{\lambda \in \Lambda_N} \langle x, \psi_\lambda\rangle \tilde{\psi}_\lambda$ of x consisting of the N largest coefficients $|\langle x, \psi_\lambda\rangle|$, we *only* have

$$\|x - x_N\|^2 \leq \frac{1}{A} \sum_{n > N} |\langle x, \psi_{\lambda_n}\rangle|^2 .$$

Applied Harmonic Analysis

Desiderata:

- Multiscale representation system.
- Convenient structure: Operators applied to one generating function.
- Partition of Fourier domain.
- Space/frequency localization.
- Fast algorithms: $x \mapsto (\langle x, \psi_\lambda \rangle)_\lambda \rightsquigarrow x$.
- Optimality for the considered class.
  $\rightsquigarrow$ In this Talk: Modeling natural images!



Continuous versus Discrete

Continuous World:

- Continuous index sets.
- Resolution of Singularities/Wavefront sets.
- More flexibility in scale $\to 0$.
- Allows strong theoretical results.

Discrete World:

- Discrete index sets.
- (Sparse) approximation properties.
- More efficient numerical realization.

Outline

1. Continuous World
   - Resolution of Singularities
   - Continuous Wavelet Transform
   - Continuous Shearlet Transform
   - Applications: Edge Detection, ...

2. Discrete World
   - Sparse Approximations
   - Discrete Wavelets
   - Directional Representation Systems: Curvelets, Shearlets,...
   - Applications: Inpainting, Magnetic Resonance Imaging, ...

Applied Harmonic Analysis Methods in Imaging Science Part II

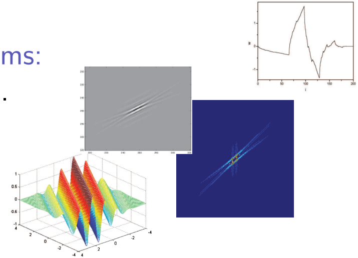
Gitta Kutyniok

(Technische Universität Berlin)

SIAM Conference on Imaging Science
Albuquerque, May 23 – 26, 2016

Applied Harmonic Analysis Approach

Selection of different Representation Systems:
Wavelets, Ridgelets, Curvelets, Shearlets,...



Main Desiderata:

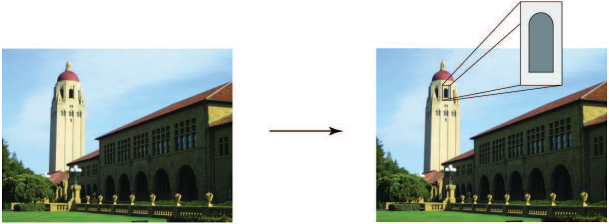
- Multiscale representation system.
- Partition of Fourier domain.
- Fast decomposition and reconstruction algorithm.
- Optimally sparse approximation of the considered class.
  $\rightsquigarrow$ Here: Modeling natural images!

Outline

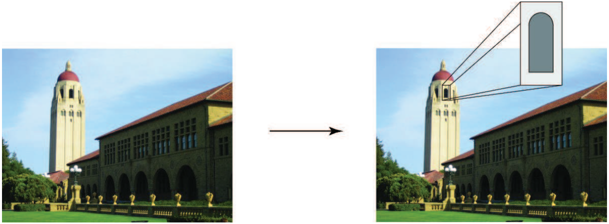
1. Sparse Approximation of Images
   - Model Situation
   - Benchmark Result
2. Wavelets
3. Shearlets
4. Applications
   - Denoising
   - Feature Extraction
   - Inpainting
   - Magnetic Resonance Imaging
5. 3D Shearlets

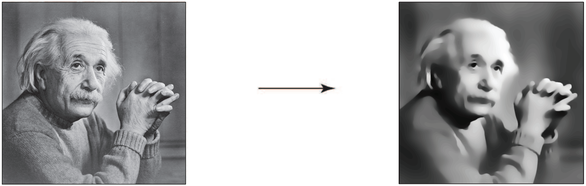
What is an Image?

What is an Image?

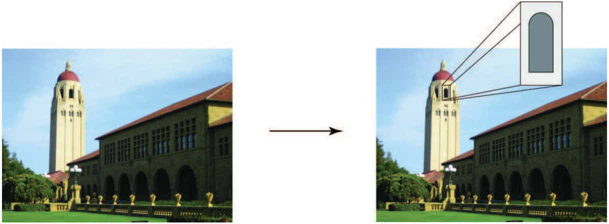


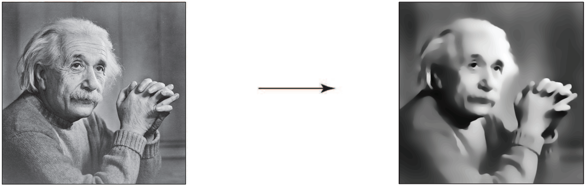
What is an Image?



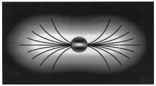


What is an Image?





- Intuitively edges are main structure.
- Justified by neurophysiology.



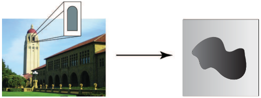
Field et al., 1993

Fitting Model

Definition (Donoho; 2001):
The set of cartoon-like functions $\mathcal{E}^2(\mathbb{R}^2)$ is defined by

$$\mathcal{E}^2(\mathbb{R}^2) = \{f \in L^2(\mathbb{R}^2) : f = f_0 + f_1 \cdot \chi_B\},$$

where $\emptyset \neq B \subset [0,1]^2$ simply connected with C^2-boundary and bounded curvature, and $f_i \in C^2(\mathbb{R}^2)$ with supp $f_i \subseteq [0,1]^2$ and $\|f_i\|_{C^2} \leq 1$, $i = 0, 1$.



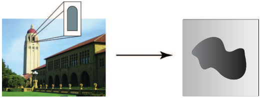
Fitting Model

Definition (Donoho; 2001):

The set of cartoon-like functions $\mathcal{E}^2(\mathbb{R}^2)$ is defined by

$$\mathcal{E}^2(\mathbb{R}^2) = \{f \in L^2(\mathbb{R}^2) : f = f_0 + f_1 \cdot \chi_B\},$$

where $\emptyset \neq B \subset [0,1]^2$ simply connected with C^2-boundary and bounded curvature, and $f_i \in C^2(\mathbb{R}^2)$ with supp $f_i \subseteq [0,1]^2$ and $\|f_i\|_{C^2} \leq 1$, $i = 0, 1$.



Theorem (Donoho; 2001):

Let $(\psi_\lambda)_\lambda \subseteq L^2(\mathbb{R}^2)$. Allowing only polynomial depth search, we have the following optimal behavior for $f \in \mathcal{E}^2(\mathbb{R}^2)$:

$$\|f - f_N\|_2^2 \asymp N^{-2} \quad \text{and} \quad |\langle f, \psi_{\lambda_n}\rangle| \lesssim n^{-\frac{3}{2}} \quad \text{as } N, n \to \infty.$$

Review of 2-D Wavelets

Definition (1D): Let $\phi \in L^2(\mathbb{R})$ be a scaling function and $\psi \in L^2(\mathbb{R})$ be a wavelet. Then the associated wavelet system is defined by

$$\{\phi(x - m) : m \in \mathbb{Z}\} \cup \{2^{j/2}\,\psi(2^j x - m) : j \geq 0, m \in \mathbb{Z}\}.$$

Review of 2-D Wavelets

Definition (1D): Let $\phi \in L^2(\mathbb{R})$ be a scaling function and $\psi \in L^2(\mathbb{R})$ be a wavelet. Then the associated wavelet system is defined by

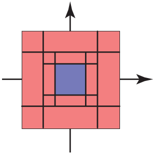
$$\{\phi(x-m) : m \in \mathbb{Z}\} \cup \{2^{j/2}\,\psi(2^j x - m) : j \geq 0, m \in \mathbb{Z}\}.$$

Definition (2D): A wavelet system is defined by

$$\{\phi^{(1)}(x-m) : m \in \mathbb{Z}^2\} \cup \{2^j \psi^{(i)}(2^j x - m) : j \geq 0, m \in \mathbb{Z}^2, i = 1,2,3\},$$

where

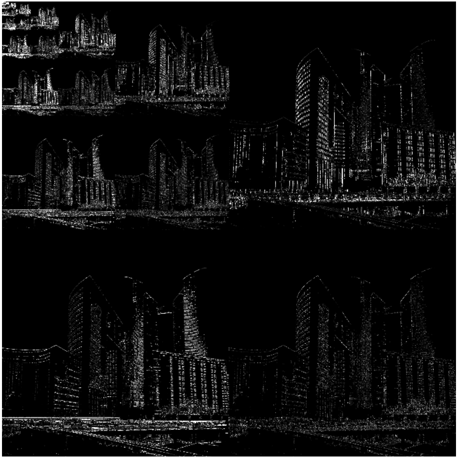
$$\phi^{(1)}(x) = \phi(x_1)\phi(x_2) \quad \text{and} \quad \begin{array}{rcl} \psi^{(1)}(x) & = & \phi(x_1)\psi(x_2), \\ \psi^{(2)}(x) & = & \psi(x_1)\phi(x_2), \\ \psi^{(3)}(x) & = & \psi(x_1)\psi(x_2). \end{array}$$



Theorem: Wavelets provide optimally sparse approximations for functions $f \in L^2(\mathbb{R}^2)$, which are C^2 apart from point singularities:

$$\|f - f_N\|_2^2 \asymp N^{-1}, \quad N \to \infty.$$

Wavelet Decomposition: JPEG2000



Wavelet Decomposition: JPEG2000

Original

25% Compression

5% Compression

What can Wavelets do?

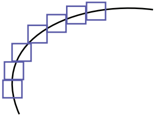
Problem:

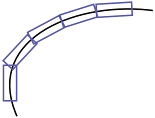
- For $f \in \mathcal{E}^2(\mathbb{R}^2)$, wavelets only achieve $\|f - f_N\|_2^2 \asymp N^{-1}$, $N \to \infty$.
- Isotropic structure of wavelets:

$$\{2^j \psi(\begin{pmatrix} 2^j & 0 \\ 0 & 2^j \end{pmatrix} x - m) : j \geq 0, m \in \mathbb{Z}^2\}.$$

- Wavelets cannot sparsely represent cartoon-like functions.

Intuitive explanation:





Main Goal

Design a Representation System which...

- ...fits into the framework of affine systems,
- ...provides an optimally sparsifying system for cartoons,
- ...allows for compactly supported analyzing elements,
- ...is associated with fast decomposition algorithms,
- ...treats the continuum and digital ‘world’ uniformly.

Main Goal

Design a Representation System which...

- ...fits into the framework of affine systems,
- ...provides an optimally sparsifying system for cartoons,
- ...allows for compactly supported analyzing elements,
- ...is associated with fast decomposition algorithms,
- ...treats the continuum and digital 'world' uniformly.

Non-Exhaustive List of Approaches:

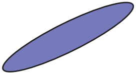
- Ridgelets (Candès and Donoho; 1999)
- Curvelets (Candès and Donoho; 2002)
- Contourlets (Do and Vetterli; 2002)
- Bandlets (LePennec and Mallat; 2003)
- Shearlets (K and Labate; 2006)

What is a Shearlet?

Scaling and Orientation

Parabolic scaling ('width $\approx$ length^2'):

$$A_{2^j} = \begin{pmatrix} 2^j & 0 \\ 0 & 2^{j/2} \end{pmatrix}, \quad j \in \mathbb{Z}.$$



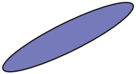
Historical remark:

- 1970's: Fefferman und Seeger/Sogge/Stein.

Scaling and Orientation

Parabolic scaling ('width ≈ length^2'):

$$A_{2^j} = \begin{pmatrix} 2^j & 0 \\ 0 & 2^{j/2} \end{pmatrix}, \quad j \in \mathbb{Z}.$$



Historical remark:

- 1970's: Fefferman und Seeger/Sogge/Stein.

Orientation via shearing:

$$S_k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}, \quad k \in \mathbb{Z}.$$

Advantage:

- Shearing leaves the digital grid $\mathbb{Z}^2$ invariant.
- Uniform theory for the continuum and digital situation.

Shearlet Systems

Affine systems:

$$\{|\det M|^{1/2}\psi(M\ \cdot\ -m) : M \in G \subseteq GL_2,\ m \in \mathbb{Z}^2\}.$$

Definition (K, Labate; 2006):
For $\psi \in L^2(\mathbb{R}^2)$, the associated shearlet system is defined by

$$\{2^{\frac{3j}{4}}\psi(S_k A_{2^j}\ \cdot\ -m)\ :\ j,k \in \mathbb{Z}, m \in \mathbb{Z}^2\}.$$

$\rightsquigarrow$ *Can be regarded as discretization of continuous shearlet systems!*

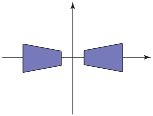
Remarks:

- Advantage: Generated by a unitary representation of the locally compact group $(\mathbb{R}^+ \times \mathbb{R}) \ltimes \mathbb{R}^2$.
- Disadvantage: Non-uniform treatment of directions.

Example of Classical (Band-Limited) Shearlet

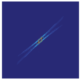
Let $\psi \in L^2(\mathbb{R}^2)$ be defined by

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\,\hat{\psi}_2(\tfrac{\xi_2}{\xi_1}),$$



where

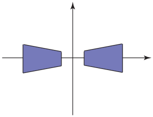
- ψ_1 wavelet, $\mathrm{supp}(\hat{\psi}_1) \subseteq [-2, -\frac{1}{2}] \cup [\frac{1}{2}, 2]$ and $\hat{\psi}_1 \in C^\infty(\mathbb{R})$.
- $\mathrm{supp}(\hat{\psi}_2) \subseteq [-1, 1]$ and $\hat{\psi}_2 \in C^\infty(\mathbb{R})$.



Example of Classical (Band-Limited) Shearlet

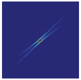
Let $\psi \in L^2(\mathbb{R}^2)$ be defined by

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\,\hat{\psi}_2(\tfrac{\xi_2}{\xi_1}),$$

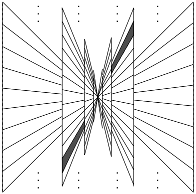


where

- ψ_1 wavelet, $\mathrm{supp}(\hat{\psi}_1) \subseteq [-2, -\frac{1}{2}] \cup [\frac{1}{2}, 2]$ and $\hat{\psi}_1 \in C^\infty(\mathbb{R})$.
- $\mathrm{supp}(\hat{\psi}_2) \subseteq [-1, 1]$ and $\hat{\psi}_2 \in C^\infty(\mathbb{R})$.



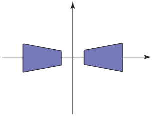
Induced tiling of Fourier domain:



Example of Classical (Band-Limited) Shearlet

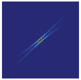
Let $\psi \in L^2(\mathbb{R}^2)$ be defined by

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\, \hat{\psi}_2(\tfrac{\xi_2}{\xi_1}),$$

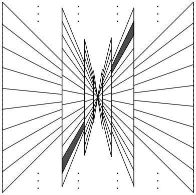


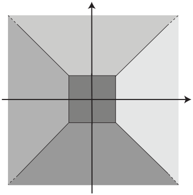
where

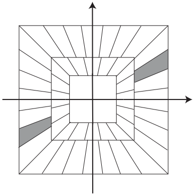
- ψ_1 wavelet, $\mathrm{supp}(\hat{\psi}_1) \subseteq [-2, -\frac{1}{2}] \cup [\frac{1}{2}, 2]$ and $\hat{\psi}_1 \in C^\infty(\mathbb{R})$.
- $\mathrm{supp}(\hat{\psi}_2) \subseteq [-1, 1]$ and $\hat{\psi}_2 \in C^\infty(\mathbb{R})$.



Induced tiling of Fourier domain:







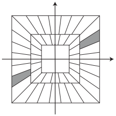
(Cone-adapted) Shearlet Systems

Definition (K, Labate; 2006):
The (cone-adapted) shearlet system $\mathcal{SH}(c;\phi,\psi,\tilde{\psi})$, $c>0$, generated by $\phi \in L^2(\mathbb{R}^2)$ and $\psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ is the union of

$$\{\phi(\cdot - cm) : m \in \mathbb{Z}^2\},$$

$$\{2^{3j/4}\psi(S_k A_{2^j} \cdot - cm) : j \geq 0, |k| \leq \lceil 2^{j/2} \rceil, m \in \mathbb{Z}^2\},$$

$$\{2^{3j/4}\tilde{\psi}(\tilde{S}_k \tilde{A}_{2^j} \cdot - cm) : j \geq 0, |k| \leq \lceil 2^{j/2} \rceil, m \in \mathbb{Z}^2\}.$$



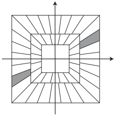
(Cone-adapted) Shearlet Systems

Definition (K, Labate; 2006):
The (cone-adapted) shearlet system $\mathcal{SH}(c;\phi,\psi,\tilde{\psi})$, $c>0$, generated by $\phi \in L^2(\mathbb{R}^2)$ and $\psi,\tilde{\psi} \in L^2(\mathbb{R}^2)$ is the union of

$$\{\phi(\cdot - cm) : m \in \mathbb{Z}^2\},$$

$$\{2^{3j/4}\psi(S_k A_{2^j}\,\cdot\, - cm) : j \geq 0, |k| \leq \lceil 2^{j/2} \rceil, m \in \mathbb{Z}^2\},$$

$$\{2^{3j/4}\tilde{\psi}(\tilde{S}_k \tilde{A}_{2^j}\,\cdot\, - cm) : j \geq 0, |k| \leq \lceil 2^{j/2} \rceil, m \in \mathbb{Z}^2\}.$$



Theorem (K, Labate, Lim, Weiss; 2006):
For $\psi,\tilde{\psi}$ classical shearlets, $\mathcal{SH}(1;\phi,\psi,\tilde{\psi})$ is a Parseval frame for $L^2(\mathbb{R}^2)$:

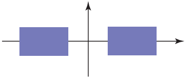
$$A\|f\|_2^2 \leq \sum_{\sigma \in \mathcal{SH}(\phi,\psi,\tilde{\psi})} |\langle f,\sigma\rangle|^2 \leq B\|f\|_2^2 \quad \text{for all } f \in L^2(\mathbb{R}^2)$$

holds for $A = B = 1$.

Compactly Supported Shearlets

Theorem (Kittipoom, K, Lim; 2012):

Let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ be compactly supported, and let $\hat{\phi}$, $\hat{\psi}$, $\hat{\tilde{\psi}}$ satisfy certain decay conditions. Then there exists c_0 such that $\mathcal{SH}(c; \phi, \psi, \tilde{\psi})$ forms a shearlet frame with controllable frame bounds for all $c \leq c_0$.

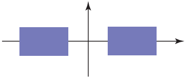


Remark: Exemplary class with $B/A \approx 4$.

Compactly Supported Shearlets

Theorem (Kittipoom, K, Lim; 2012):

Let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ be compactly supported, and let $\hat{\phi}$, $\hat{\psi}$, $\hat{\tilde{\psi}}$ satisfy certain decay conditions. Then there exists c_0 such that $\mathcal{SH}(c; \phi, \psi, \tilde{\psi})$ forms a shearlet frame with controllable frame bounds for all $c \leq c_0$.

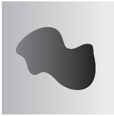


Remark: Exemplary class with $B/A \approx 4$.

Theorem (Guo, Labate; 2007)(K, Lim; 2011):

Let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ be compactly supported, and let $\hat{\phi}$, $\hat{\psi}$, $\hat{\tilde{\psi}}$ satisfy certain decay conditions. Then $\mathcal{SH}(c; \phi, \psi, \tilde{\psi}) = (\sigma_\eta)_\eta$ provides an optimally sparsifying system for $f \in \mathcal{E}^2(\mathbb{R}^2)$, i.e., for $N, n \to \infty$,

$$\|f - f_N\|_2^2 \lesssim N^{-2}(\log N)^3 \text{ and } |\langle f, \sigma_{\eta_n} \rangle| \lesssim n^{-\frac{3}{2}}(\log n)^{\frac{3}{2}}.$$

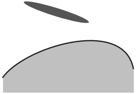


Heuristic Argument

Estimate:

$$\|f - f_N\|_2^2 \lesssim \sum_{n>N} (|\langle f, \sigma_{\eta_n} \rangle|)^2 \lesssim \sum_{n>N} (n^{-\frac{3}{2}})^2 \lesssim N^{-2}.$$

Case 1:



$|\langle f, \sigma_\eta \rangle|$ negligible!

Case 2:

$|\langle f, \sigma_\eta \rangle|$ negligible!

Case 3:

$|\langle f, \sigma_\eta \rangle| \leq \|f\|_\infty \|\sigma_\eta\|_1 \lesssim 2^{-\frac{3}{4}j}$

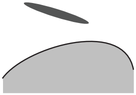
$\rightsquigarrow |\langle f, \sigma_{\eta_n} \rangle| \lesssim n^{-\frac{3}{2}}$

Heuristic Argument

Estimate:

$$\|f - f_N\|_2^2 \lesssim \sum_{n>N} (|\langle f, \sigma_{\eta_n} \rangle|)^2 \lesssim \sum_{n>N} (n^{-\frac{3}{2}})^2 \lesssim N^{-2}.$$

Case 1:



$|\langle f, \sigma_\eta \rangle|$ negligible!

Case 2:

$|\langle f, \sigma_\eta \rangle|$ negligible!

Case 3:

$$|\langle f, \sigma_\eta \rangle| \leq \|f\|_\infty \|\sigma_\eta\|_1 \lesssim 2^{-\frac{3}{4}j}$$

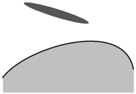
$$\rightsquigarrow |\langle f, \sigma_{\eta_n} \rangle| \lesssim n^{-\frac{3}{2}}$$

Heuristic Argument

Estimate:

$$\|f - f_N\|_2^2 \lesssim \sum_{n>N} (|\langle f, \sigma_{\eta_n} \rangle|)^2 \lesssim \sum_{n>N} (n^{-\frac{3}{2}})^2 \lesssim N^{-2}.$$

Case 1:



$|\langle f, \sigma_\eta \rangle|$ negligible!

Case 2:

$|\langle f, \sigma_\eta \rangle|$ negligible!

Case 3:

$|\langle f, \sigma_\eta \rangle| \leq \|f\|_\infty \|\sigma_\eta\|_1 \lesssim 2^{-\frac{3}{4}j}$

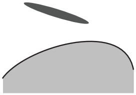
$\rightsquigarrow |\langle f, \sigma_{\eta_n} \rangle| \lesssim n^{-\frac{3}{2}}$

Heuristic Argument

Estimate:

$$\|f - f_N\|_2^2 \lesssim \sum_{n>N} (|\langle f, \sigma_{\eta_n} \rangle|)^2 \lesssim \sum_{n>N} (n^{-\frac{3}{2}})^2 \lesssim N^{-2}.$$

Case 1:



$|\langle f, \sigma_\eta \rangle|$ negligible!

Case 2:

$|\langle f, \sigma_\eta \rangle|$ negligible!

Case 3:

$$|\langle f, \sigma_\eta \rangle| \le \|f\|_\infty \|\sigma_\eta\|_1 \lesssim 2^{-\frac{3}{4}j}$$

$$\rightsquigarrow |\langle f, \sigma_{\eta_n} \rangle| \lesssim n^{-\frac{3}{2}}$$

Curvelets

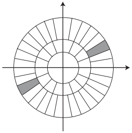
Definition (Candès, Donoho; 2002):

Let

- $W \in C^\infty(\mathbb{R})$ be a wavelet with $\mathrm{supp}(W) \subseteq \left(\frac{1}{2}, 2\right)$,
- $V \in C^\infty(\mathbb{R})$ be a 'bump function' with $\mathrm{supp}(V) \subseteq (-1, 1)$.

Then the curvelet system $(\gamma_{(j,l,k)})_{(j,l,k)}$ is defined by

$$\hat{\gamma}_{(j,0,0)}(r,\omega) := 2^{-3j/4} W\left(2^{-j} r\right) V(2^{\lfloor j/2 \rfloor}\omega)$$



and

$$\gamma_{(j,l,k)}(\cdot) := \gamma_{(j,0,0)}(R_{\theta_{(j,l,k)}}(\cdot - x_{(j,l,k)})).$$

Theorem (Candès, Donoho; 2002):

The Parseval frame of curvelets provides optimally sparse approximations of $f \in \mathcal{E}^2(\mathbb{R}^2)$, i.e.,

$$\|f - f_N\|_2^2 \leq C \cdot N^{-2} \cdot (\log N)^3, \quad N \to \infty.$$

Framework for Sparse Approximation Results

General Framework:

- Parabolic Molecules (Grohs, K; 2013) (Flinth; 2013)
  $\rightsquigarrow$ *includes curvelets, shearlets, ...*
- α-Molecules (Grohs, Keiper, K, Schäfer; 2016)
  $\rightsquigarrow$ *includes ridgelets, wavelets, curvelets, shearlets, ...*

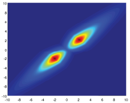


Illustration (“α = degree of anisotropy”):

$\alpha = 0$ — $\frac{1}{2}$ — 1

Ridgelets — Curvelets/Shearlets — Wavelets

Theorem (Grohs, Keiper, K, Schäfer; 2016):

“Sparse approximation results for appropriate function classes can be derived in the very general setting of α-molecules.”

Recent Approaches to Fast Shearlet Transforms

`www.ShearLab.org`:

- Separable Shearlet Transform *(Lim; 2009)*
- Digital Shearlet Transform *(K, Shahram, Zhuang; 2011)*
- 2D&3D (parallelized) Shearlet Transform *(K, Lim, Reisenhofer; 2013)*

Additional Code:

- Filter-based implementation *(Easley, Labate, Lim; 2009)*
- Fast Finite Shearlet Transform *(Häuser, Steidl; 2014)*
- Shearlet Toolbox 2D&3D *(Easley, Labate, Lim, Negy; 2014)*

Theoretical Approaches:

- Adaptive Directional Subdivision Schemes *(K, Sauer; 2009)*
- Shearlet Unitary Extension Principle *(Han, K, Shen; 2011)*
- Gabor Shearlets *(Bodmann, K, Zhuang; 2013)*

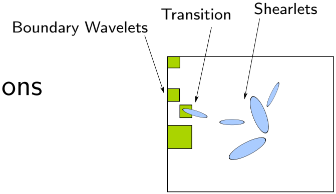
Boundary Shearlets

Definition (Grohs, K, Ma, and Petersen; 2016):
For $t \in \mathbb{N}$, $\mathcal{W}$ a biorthogonal wavelet basis, $(\sigma_\eta)_\eta$ a shearlet system, and $\mathcal{W}_0 := \{\omega_{j,m} \in \mathcal{W} : d(\operatorname{supp} \omega_{j,m}, \partial\Omega) < 2^{-\frac{j-t}{2}}\}$, the boundary shearlet system with offset t is defined as

$$\{\sigma_\eta : \operatorname{supp} \sigma_\eta \subseteq \Omega\} \cup \mathcal{W}_0$$

Some Results (Grohs, K, Ma, Petersen, and Raslan; 2016):
Boundary shearlet systems...

- ...form a frame for $L^2(\Omega)$.
- ...provide optimally sparse approximations for adapted cartoon-like functions.
- ...characterize Sobolev spaces.
- ...can be designed to provide Sobolev frames.



Selected Applications...

Image Denoising, I

Original

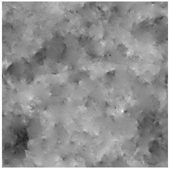
Noisy Version (20.17dB)

Curvelets (28.70dB, 7.22sec)

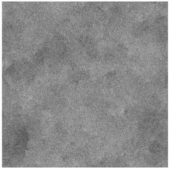
Shearlets (29.20dB, 5.56sec)

(Source: W.-Q Lim; 2011)

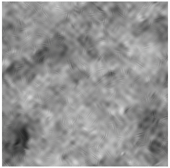
Image Denoising, II



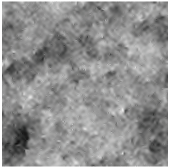
Original



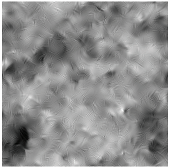
Noisy Version (6.5dB)



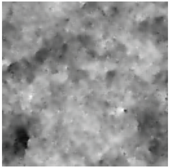
Curvelets (22.10dB)



Wavelets (23.68dB)



Shearlets (24.45dB)



Dict.Lear. (24.70dB)

(Source: S. Beckouche; 2012)

Regularization of Inverse Problems

Generalized Tikhonov Regularization:

Given an ill-posed inverse problem $Kx = y$, where $K : X \to Y$, an approximate solution $x^\alpha \in X$, $\alpha > 0$, can be determined by minimizing

$$\tilde{J}_\alpha(x) := \|Kx - y\|^2 + \alpha \mathcal{P}(x), \quad x \in X.$$

$\rightsquigarrow$ *The penalty term $\mathcal{P}$ incorporates properties of the solution!*

Some Examples for $\mathcal{P}$:

$$\|x\|_{TV}, \ \ \|x\|_{H^s}, \ \ \|(\langle x, \psi_\lambda \rangle)_\lambda\|_1, \ldots$$

Regularization of Inverse Problems

Generalized Tikhonov Regularization:

Given an ill-posed inverse problem $Kx = y$, where $K : X \to Y$, an approximate solution $x^\alpha \in X$, $\alpha > 0$, can be determined by minimizing

$$\tilde{J}_\alpha(x) := \|Kx - y\|^2 + \alpha \mathcal{P}(x), \quad x \in X.$$

$\rightsquigarrow$ *The penalty term $\mathcal{P}$ incorporates properties of the solution!*

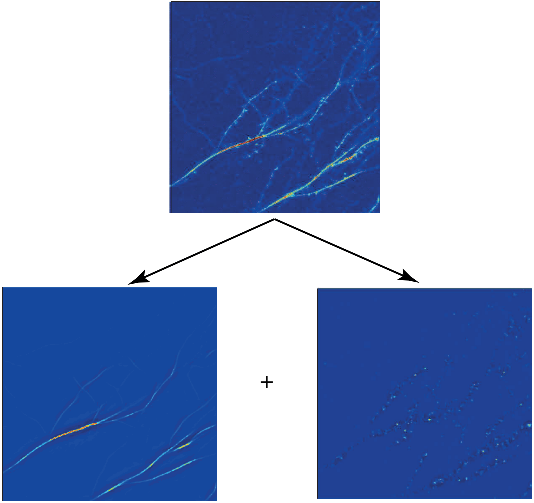
Some Examples for $\mathcal{P}$:

$$\|x\|_{TV}, \ \|x\|_{H^s}, \ \|(\langle x, \psi_\lambda \rangle)_\lambda\|_1, \ldots$$

Some Earlier Footprints in Inverse Problems:

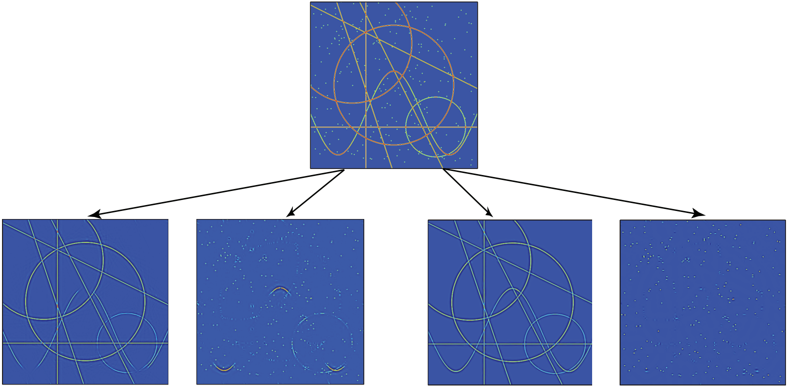
- Donoho (1995): Wavelet-Vaguelette decomposition.
- Chambolle, DeVore, Lee, Lucier (1998): Penalty on the Besov norm.
- Daubechies, Defries, De Mol (2004): General sparsity constraints.

Numerical Results of Feature Extraction, I



(Source: Brandt, K, Lim, Sündermann; 2011)

Numerical Results of Feature Extraction, II



MCALab 120 (52.74 sec)   ShearLab (33.75 sec)

Feature Extraction and ℓ_1 Minimization

Key Idea: Let $x = x_1 + x_2$. Let Φ_1 and Φ_2 be sparsifying frames for x_1 and x_2, respectively, but not conversely, and consider

$$(x_1^*, x_2^*) = \operatorname{argmin}_{\tilde{x}_1, \tilde{x}_2} \|\Phi_1^T \tilde{x}_1\|_1 + \|\Phi_2^T \tilde{x}_2\|_1 \text{ subject to } x = \tilde{x}_1 + \tilde{x}_2.$$

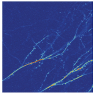
Feature Extraction and ℓ_1 Minimization

Key Idea: Let $x = x_1 + x_2$. Let Φ_1 and Φ_2 be sparsifying frames for x_1 and x_2, respectively, but not conversely, and consider

$$(x_1^*, x_2^*) = \mathrm{argmin}_{\tilde{x}_1, \tilde{x}_2} \|\Phi_1^T \tilde{x}_1\|_1 + \|\Phi_2^T \tilde{x}_2\|_1 \text{ subject to } x = \tilde{x}_1 + \tilde{x}_2.$$

Model: For τ a closed C^2-curve,

$$f = \mathcal{P} + \mathcal{C} = \sum_{i=1}^{P} |x - x_i|^{-3/2} + \int \delta_{\tau(t)} dt.$$



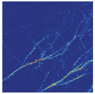
Feature Extraction and ℓ_1 Minimization

Key Idea: Let $x = x_1 + x_2$. Let Φ_1 and Φ_2 be sparsifying frames for x_1 and x_2, respectively, but not conversely, and consider

$$(x_1^*, x_2^*) = \text{argmin}_{\tilde{x}_1,\tilde{x}_2} \|\Phi_1^T \tilde{x}_1\|_1 + \|\Phi_2^T \tilde{x}_2\|_1 \text{ subject to } x = \tilde{x}_1 + \tilde{x}_2.$$

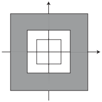
Model: For τ a closed C^2-curve,

$$f = \mathcal{P} + \mathcal{C} = \sum_{i=1}^{P} |x - x_i|^{-3/2} + \int \delta_{\tau(t)} dt.$$



Subband Decomposition:

$$f_j = \mathcal{P}_j + \mathcal{C}_j, \quad \mathcal{P}_j = \mathcal{P} \star F_j \text{ and } \mathcal{C}_j = \mathcal{C} \star F_j.$$



Two Sparsifying Systems:

$$\text{Wavelets } (\psi_\lambda)_\lambda \quad \text{and} \quad \text{Shearlets } (\sigma_\eta)_\eta.$$

Analysis of Feature Extraction

ℓ_1-Decomposition:

$$(\mathcal{P}_j^*, \mathcal{C}_j^*) = \mathrm{argmin}_{\tilde{\mathcal{P}}_j, \tilde{\mathcal{C}}_j} \|(\langle \tilde{\mathcal{P}}_j, \psi_\lambda \rangle)_\lambda\|_1 + \|(\langle \tilde{\mathcal{C}}_j, \sigma_\eta \rangle)_\eta\|_1 \text{ s.t. } f_j = \tilde{\mathcal{P}}_j + \tilde{\mathcal{C}}_j$$

Analysis of Feature Extraction

ℓ_1-Decomposition:

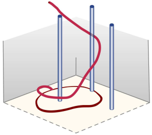
$$(\mathcal{P}_j^*, \mathcal{C}_j^*) = \text{argmin}_{\tilde{\mathcal{P}}_j, \tilde{\mathcal{C}}_j} \|(\langle \tilde{\mathcal{P}}_j, \psi_\lambda \rangle)_\lambda\|_1 + \|(\langle \tilde{\mathcal{C}}_j, \sigma_\eta \rangle)_\eta\|_1 \text{ s.t. } f_j = \tilde{\mathcal{P}}_j + \tilde{\mathcal{C}}_j$$

Theorem (Donoho, K; 2013):

$$\frac{\|\mathcal{P}_j^* - \mathcal{P}_j\|_2 + \|\mathcal{C}_j^* - \mathcal{C}_j\|_2}{\|\mathcal{P}_j\|_2 + \|\mathcal{C}_j\|_2} \to 0, \qquad j \to \infty.$$

Idea of Proof:

- Relative sparsity and cluster coherence.
- Analyze wavefront sets of $\mathcal{P}$ and $\mathcal{C}$ in phase space.



Analysis of Feature Extraction

ℓ_1-Decomposition:

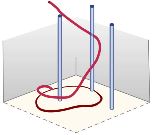
$$(\mathcal{P}_j^*, \mathcal{C}_j^*) = \text{argmin}_{\tilde{\mathcal{P}}_j, \tilde{\mathcal{C}}_j} \|(\langle \tilde{\mathcal{P}}_j, \psi_\lambda \rangle)_\lambda\|_1 + \|(\langle \tilde{\mathcal{C}}_j, \sigma_\eta \rangle)_\eta\|_1 \text{ s.t. } f_j = \tilde{\mathcal{P}}_j + \tilde{\mathcal{C}}_j$$

Theorem (Donoho, K; 2013):

$$\frac{\|\mathcal{P}_j^* - \mathcal{P}_j\|_2 + \|\mathcal{C}_j^* - \mathcal{C}_j\|_2}{\|\mathcal{P}_j\|_2 + \|\mathcal{C}_j\|_2} \to 0, \qquad j \to \infty.$$

Idea of Proof:

- Relative sparsity and cluster coherence.
- Analyze wavefront sets of $\mathcal{P}$ and $\mathcal{C}$ in phase space.

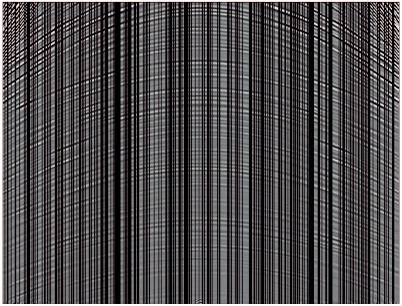


Theorem (K; 2014):

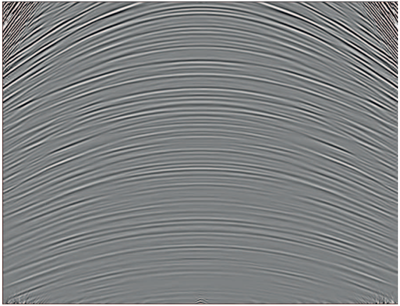
Using One-Step-Thresholding, we also have

$$WF(\sum_j F_j \star \mathcal{P}_j^*) = WF(\mathcal{P}) \quad \text{and} \quad WF(\sum_j F_j \star \mathcal{C}_j^*) = WF(\mathcal{C}).$$

Numerical Results of Inpainting, I



Undersampled seismic data



Reconstructed image

(Source: K, Lim; 2012)

Numerical Results of Inpainting, II

(Source: Kutyniok, Lim; 2014)

Analysis of Inpainting

Key Idea:

Let Φ be a sparsifying frame for x in $\mathcal{H} = \mathcal{H}_M \oplus \mathcal{H}_K$. Solve

$$x^* = \text{argmin}_{\tilde{x}} \|\Phi^T \tilde{x}\|_1 \text{ subject to } P_{\mathcal{H}_K} x = P_{\mathcal{H}_K} \tilde{x}.$$

Analysis of Inpainting

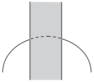
Key Idea:

Let Φ be a sparsifying frame for x in $\mathcal{H} = \mathcal{H}_M \oplus \mathcal{H}_K$. Solve

$$x^* = \mathrm{argmin}_{\tilde{x}} \|\Phi^T \tilde{x}\|_1 \text{ subject to } P_{\mathcal{H}_K} x = P_{\mathcal{H}_K} \tilde{x}.$$

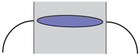
Observed Object:

$$f = 1_{\mathbb{R}^2 \setminus M_h} \cdot \mathcal{C}.$$



ℓ_1-Inpainting:

$$\mathcal{C}_j^* = \mathrm{argmin}_{\tilde{\mathcal{C}}_j} \|(\langle \tilde{\mathcal{C}}_j, \sigma_\eta \rangle)_\eta\|_1 \text{ s.t. } 1_{\mathbb{R}^2 \setminus M_{h_j}} \cdot (\mathcal{C} \star F_j) = 1_{\mathbb{R}^2 \setminus M_{h_j}} \cdot \tilde{\mathcal{C}}_j$$



Analysis of Inpainting

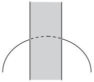
Key Idea:

Let Φ be a sparsifying frame for x in $\mathcal{H} = \mathcal{H}_M \oplus \mathcal{H}_K$. Solve

$$x^* = \text{argmin}_{\tilde{x}} \|\Phi^T \tilde{x}\|_1 \text{ subject to } P_{\mathcal{H}_K} x = P_{\mathcal{H}_K} \tilde{x}.$$

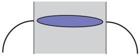
Observed Object:

$$f = 1_{\mathbb{R}^2 \setminus M_h} \cdot \mathcal{C}.$$



ℓ_1-Inpainting:

$$\mathcal{C}_j^* = \text{argmin}_{\tilde{\mathcal{C}}_j} \|(\langle \tilde{\mathcal{C}}_j, \sigma_\eta \rangle)_\eta\|_1 \text{ s.t. } 1_{\mathbb{R}^2 \setminus M_{h_j}} \cdot (\mathcal{C} \star F_j) = 1_{\mathbb{R}^2 \setminus M_{h_j}} \cdot \tilde{\mathcal{C}}_j$$



Theorem (King, K, Zhuang; 2014)(Genzel, K; 2015)

For $h_j = o(2^{-j/2})$ as $j \to \infty$,

$$\frac{\|\mathcal{C}_j^* - \mathcal{C}_j\|_2}{\|\mathcal{C}_j\|_2} \to 0, \qquad j \to \infty.$$

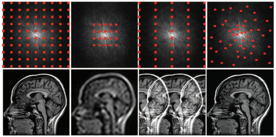
Application to MRI

Model Situation:

Reconstruct $f \in L^2(\mathbb{R}^2)$ from Fourier samples $\hat{f}(\lambda)$, $\lambda \in \Lambda \subseteq \mathbb{R}^2$.

Goals:

- Fast acquisition $\longleftrightarrow$ Small set Λ
- Optimality result



Initial idea with wavelets: Lustig, Donoho, Pauly; 2007

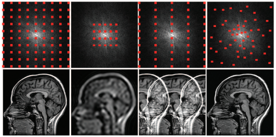
Application to MRI

Model Situation:

Reconstruct $f \in L^2(\mathbb{R}^2)$ from Fourier samples $\hat{f}(\lambda)$, $\lambda \in \Lambda \subseteq \mathbb{R}^2$.

Goals:

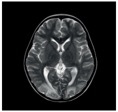
- Fast acquisition $\longleftrightarrow$ Small set Λ
- Optimality result



Initial idea with wavelets: Lustig, Donoho, Pauly; 2007

General Idea (K, Ma, and Lim; 2014):

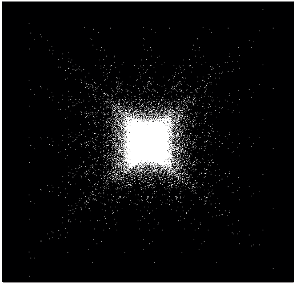
- Model for f: Cartoon-like functions.
- (Dualizable) shearlets as sparsifying system $(\sigma_\eta)_\eta$.
- Directional (random) sampling scheme Λ.
- Algorithmic approach:



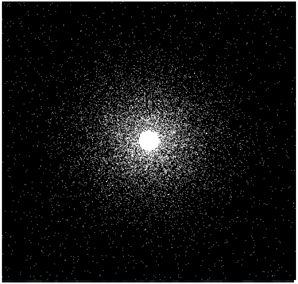
$$\min_f \|(\langle f, \sigma_\eta \rangle)_\eta\|_1 \quad \text{subject to} \quad \|(\hat{f}(\lambda) - g_\lambda)_{\lambda \in \Lambda}\|_2 \le \varepsilon$$

Asymptotic Optimality of Shearlet Scheme

Sampling Schemes:



Directional Sampling Scheme



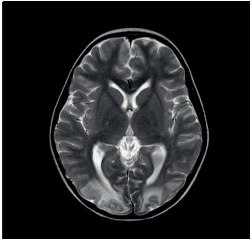
Variable Density Sampling Scheme

Theorem (K, Lim; 2015):

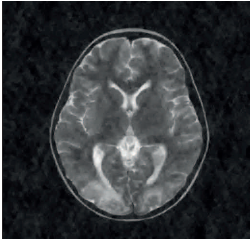
"Using the directional sampling schemes $(\Delta_M)_M$, $\#\Delta_M = M$, and $M \to \infty$ in combination with dualizable shearlets, this reconstruction scheme $\mathcal{R}$ is asymptotically optimal in the sense that, for all $f \in \mathcal{E}^2(\mathbb{R}^2)$,

$$\|f - \mathcal{R}(f, \Delta_M)\|_2^2 \lesssim M^{-2+\delta} \text{ as } M \to \infty."$$

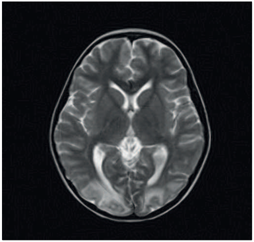
Numerical Results for 512x512 MRI Image



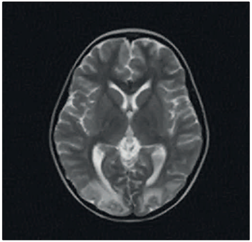
Original



Wavelets + Variable Density Sampling
(5% sampling rate, 25.00dB)



Shearlet Scheme
(5% sampling rate, 32.28dB)



Wavelets + Directional Sampling
(5% sampling rate, 29.81dB)

From 2D to 3D...

Question:

Why is the 3D situation such crucial?

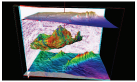
2D $\longrightarrow$ 3D

Question:

Why is the 3D situation such crucial?

Obvious answer:

- 3D data is essential for Astronomy, Biology, Seismology,...



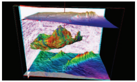
2D $\longrightarrow$ 3D

Question:

Why is the 3D situation such crucial?

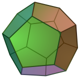
Obvious answer:

- 3D data is essential for Astronomy, Biology, Seismology,...



A different viewpoint:

- Anisotropic features occur in 3D for the first time in different dimensions.
- Transition 2D $\rightarrow$ 3D is unique.



Extended Model for 3D Images

Definition:

Let $1 < \alpha \leq 2$. The set of 3D images $\mathcal{E}^2(\mathbb{R}^3)$ is defined by

$$\mathcal{E}_2^\alpha(\mathbb{R}^3) = \{f \in L^2(\mathbb{R}^3) \,:\, f = f_0 + f_1\,\chi_B\},$$

where $f_i \in C^2$, $\operatorname{supp} f_i \subset [0,1]^3$ and $B \subset [0,1]^3$ with ∂B a closed C^2-surface whose principal curvatures are bounded by ν.

Extended Model for 3D Images

Definition:

Let $1 < \alpha \leq 2$. The set of 3D images $\mathcal{E}^2(\mathbb{R}^3)$ is defined by

$$\mathcal{E}_2^{\alpha}(\mathbb{R}^3) = \{f \in L^2(\mathbb{R}^3) \, : \, f = f_0 + f_1 \, \chi_B\},$$

where $f_i \in C^2$, $\operatorname{supp} f_i \subset [0,1]^3$ and $B \subset [0,1]^3$ with ∂B a closed C^2-surface whose principal curvatures are bounded by ν.

Theorem (K, Lemvig, Lim; 2011):

Let $(\psi_\lambda)_\lambda \subset L^2(\mathbb{R}^3)$. Allowing only polynomial depth search, the optimal asymptotic approximation error of $f \in \mathcal{E}^2(\mathbb{R}^3)$ is

$$\|f - f_N\|_2^2 \asymp N^{-1}, \quad N \to \infty.$$

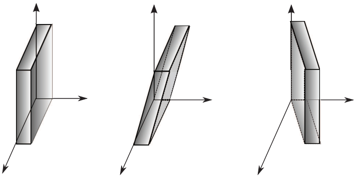
3D Shearlets

- Anisotropic scaling A_j:

$$A_j = \begin{pmatrix} 2^j & 0 & 0 \\ 0 & 2^{j/2} & 0 \\ 0 & 0 & 2^{j/2} \end{pmatrix}$$

- Shearing $S_k, k = (k_1, k_2)$:

$$S_k = \begin{pmatrix} 1 & k_1 & k_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$



Pyramid-adapted Shearlet Systems

Definition:

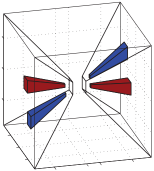
The pyramid-adapted shearlet system $\mathcal{SH}(\phi, \psi, \tilde{\psi}, \breve{\psi}; c)$ generated by $\phi \in L^2(\mathbb{R}^3)$ and $\psi, \tilde{\psi}, \breve{\psi} \in L^2(\mathbb{R}^3)$ is

$$\{\phi(\cdot - cm) : m \in \mathbb{Z}^3\}$$

$$\cup\{2^j \psi(S_k A_j \cdot - cm) : (j, k, m) \in \Lambda_{pyramid}\}$$

$$\cup\{2^j \tilde{\psi}(\tilde{S}_k \tilde{A}_j \cdot - cm) : (j, k, m) \in \Lambda_{pyramid}\}$$

$$\cup\{2^j \breve{\psi}(\breve{S}_k \breve{A}_j \cdot - cm) : (j, k, m) \in \Lambda_{pyramid}\},$$



where

$$\Lambda_{pyramid} = \{(j, k, m) : j \geq 0, |k_1|, |k_2| \leq \lceil 2^{j/2} \rceil, m \in \mathbb{Z}^2\}, \quad c > 0.$$

Optimal Sparse Approximation

Theorem (K, Lemvig, Lim; 2011):

Let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^3)$ be compactly supported, and let $\hat{\phi}$, $\hat{\psi}$, $\hat{\tilde{\psi}}$ satisfy certain decay conditions. Then $\mathcal{SH}(\phi, \psi, \tilde{\psi}; c) = (\sigma_\eta)_\eta$ provides an optimally sparsifying system for $f \in \mathcal{E}^2(\mathbb{R}^3)$, i.e., for $N \to \infty$,

$$\|f - f_N\|_2^2 \lesssim N^{-1}(\log N)^2.$$

Optimal Sparse Approximation

Theorem (K, Lemvig, Lim; 2011):

Let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^3)$ be compactly supported, and let $\hat{\phi}$, $\hat{\psi}$, $\hat{\tilde{\psi}}$ satisfy certain decay conditions. Then $\mathcal{SH}(\phi, \psi, \tilde{\psi}; c) = (\sigma_\eta)_\eta$ provides an optimally sparsifying system for $f \in \mathcal{E}^2(\mathbb{R}^3)$, i.e., for $N \to \infty$,

$$\|f - f_N\|_2^2 \lesssim N^{-1}(\log N)^2.$$

Extended Model containing 0D, 1D & 2D Features:

- Does the optimal approximation rate change?
- Do we require additional 3D shearlets?

Optimal Sparse Approximation

Theorem (K, Lemvig, Lim; 2011):

Let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^3)$ be compactly supported, and let $\hat{\phi}$, $\hat{\psi}$, $\hat{\tilde{\psi}}$ satisfy certain decay conditions. Then $\mathcal{SH}(\phi, \psi, \tilde{\psi}; c) = (\sigma_\eta)_\eta$ provides an optimally sparsifying system for $f \in \mathcal{E}^2(\mathbb{R}^3)$, i.e., for $N \to \infty$,

$$\|f - f_N\|_2^2 \lesssim N^{-1}(\log N)^2.$$

Extended Model containing 0D, 1D & 2D Features:

- Does the optimal approximation rate change?
- Do we require additional 3D shearlets?

Theorem (K, Lemvig, Lim; 2011):

(i) The optimal approximation rate remains the same for cartoon-like 3D images with only piecewise smooth C^2.

(ii) The shearlet approximation rate remains the same for cartoon-like 3D images with only piecewise smooth C^2.

Video Denoising, I

original

noisy (σ = 40 PSNR = 16.06)

$SL3D_1$ (PSNR = 26.17)

$SL3D_2$ (PSNR = 27.14)

NSST (PSNR = 25.68)

SURF (PSNR = 25.91)

Video Denoising, II

	$\sigma = 10$	20	30	40	50
$SL3D_1$	33.13	29.46	27.51	26.17	25.18
$SL3D_2$	**33.81**	**30.28**	**28.41**	**27.14**	**26.17**
$NSST$	32.59	29.00	27.05	25.68	24.63
$SURF$	30.86	28.26	26.87	25.91	25.18

- $SL3D_1$: $SL3D_1$ with $13, 13$ and 49 directions on scales one, two and three (K, Lim, and Reisenhofer, 2016).
- $SL3D_2$: $SL3D_2$ with $49, 49$ and 193 directions on scales one, two and three (K, Lim, and Reisenhofer, 2016).
- $NSST$: Nonsubsampled Shearlet Transform (Negi and Labate; 2013).
- $SURF$: Surfacelet Transform (Do and Lu; 2007).

Let's conclude...

What to take Home...?

- Applied Harmonic Analysis provides various representation systems such as wavelets, ridgelets, curvelets, and shearlets.
- They provide sparse approximation for certain classes of images, leading to
  - Efficient decompositions for, e.g., the analysis/processing of images, in particular for regularization of inverse problems.
  - Sparse representations for, e.g., compression of images.
- Continuous and discrete systems/frames and associated transforms are available.
- Some applications using wavelets and shearlets for regularization:
  - Edge Detection.
  - Feature extraction.
  - Inpainting.
  - Magnetic Resonance Imaging.

THANK YOU!

References available at:

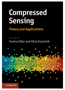
www.math.tu-berlin.de/∼kutyniok

Code available at:

www.ShearLab.org

Related Books:

- Y. Eldar and G. Kutyniok
  Compressed Sensing: Theory and Applications
  Cambridge University Press, 2012.
- G. Kutyniok and D. Labate
  Shearlets: Multiscale Analysis for Multivariate Data
  Birkhäuser-Springer, 2012.



Applied Harmonic Analysis Methods in Imaging Science
Part I

Demetrio Labate
(University of Houston)

SIAM Conference on Imaging Science
Albuquerque, May 23 – 26, 2016

Outline

1. Continuous Wavelet Transform
2. Continuous Shearlet Transform
   - Shearlet analysis of singularities
3. Applications
   - Edge analysis and detection
   - Soma detection in neuronal images
   - Classification with scattering transform

The Continuous Wavelet Transform

The classical continuous wavelet transform on $\mathbb{R}$ is associated with the **affine systems** of functions

$$\{\psi_{a,t}(x) = a^{-\frac{1}{2}}\psi(a^{-1}(x-t)) : \, a > 0, t \in \mathbb{R}\},$$

where $\psi \in L^2(\mathbb{R})$.

The Continuous Wavelet Transform

The classical continuous wavelet transform on $\mathbb{R}$ is associated with the **affine systems** of functions

$$\{\psi_{a,t}(x) = a^{-\frac{1}{2}}\psi(a^{-1}(x-t)) :\ a > 0, t \in \mathbb{R}\},$$

where $\psi \in L^2(\mathbb{R})$.

Provided that ψ satisfies the **admissibility condition** [Calderón, 1964]

$$\int_{a>0} |\psi(a\xi)|^2 \frac{da}{a} = 1, \qquad \text{for a.e. } \xi \in \mathbb{R},$$

the **continuous wavelet transform** of f

$$\mathcal{W}_\psi :\ f \to \mathcal{W}_\psi f(a,t) = \langle f, \psi_{a,t} \rangle, \qquad \text{for } a > 0, t \in \mathbb{R}^d,$$

is a linear isometry (from $L^2(\mathbb{R})$ to $L^2(\mathbb{A})$).

The Continuous Wavelet Transform

That is,

$$\|f\|^2_{L^2(\mathbb{R})} = \int_{\mathbb{R}} \int_{a>0} |\mathcal{W}_\psi f(a,t)|^2 \, \frac{da}{a} dt,$$

The Continuous Wavelet Transform

That is,

$$\|f\|^2_{L^2(\mathbb{R})} = \int_{\mathbb{R}} \int_{a>0} |\mathcal{W}_\psi f(a,t)|^2 \, \frac{da}{a} dt,$$

or

$$f(x) = \int_{\mathbb{R}} \int_{a>0} \langle f, \psi_{a,t} \rangle \; \psi_{a,t}(x) \, \frac{da}{a} dt.$$

$d\lambda(a,t) = \frac{da}{a} dt$ is the *left Haar measure* on the affine group.

The Continuous Wavelet Transform

That is,

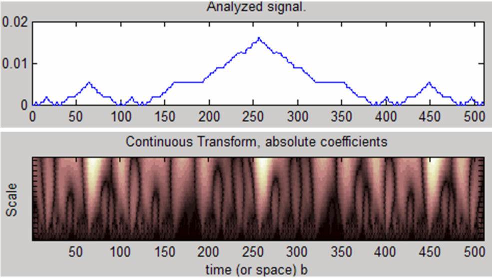
$$\|f\|^2_{L^2(\mathbb{R})} = \int_{\mathbb{R}} \int_{a>0} |\mathcal{W}_\psi f(a,t)|^2 \, \frac{da}{a} dt,$$

or

$$f(x) = \int_{\mathbb{R}} \int_{a>0} \langle f, \psi_{a,t} \rangle \; \psi_{a,t}(x) \, \frac{da}{a} dt.$$

$d\lambda(a,t) = \frac{da}{a} dt$ is the *left Haar measure* on the affine group.

- $\mathcal{W}_\psi f(a,t)$ measures the content of f at **scale** a and **location** t.



The Continuous Wavelet Transform

• The continuous wavelet transform has a special ability to deal with point singularities.

The Continuous Wavelet Transform

• The continuous wavelet transform has a special ability to deal with point singularities.

If f is singular at location t_0, $\mathcal{W}_\psi f(a,t)$ signals the location t_0 through its asymptotic decay at fine scales, $a \to 0$.

The Continuous Wavelet Transform

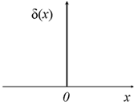
• The continuous wavelet transform has a special ability to deal with point singularities.

If f is singular at location t_0, $\mathcal{W}_\psi f(a,t)$ signals the location t_0 through its asymptotic decay at fine scales, $a \to 0$.

• This property is a manifestation of the *sparsity and locality* of the wavelet representation and it is critical in multiple signal/image processing applications.

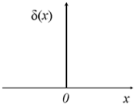
Example: Dirac Delta

Let ψ be a well-localized wavelet (e.g., Schwartz class) on $\mathbb{R}$, and δ be the Dirac delta.



Example: Dirac Delta

Let ψ be a well-localized wavelet (e.g., Schwartz class) on $\mathbb{R}$, and δ be the Dirac delta.

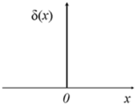


We have:

$$\mathcal{W}_\psi \delta(a, t) = \langle \delta, \psi_{a,t} \rangle = \psi_{a,t}(0).$$

Example: Dirac Delta

Let ψ be a well-localized wavelet (e.g., Schwartz class) on $\mathbb{R}$, and δ be the Dirac delta.



We have:

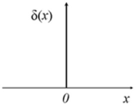
$$\mathcal{W}_\psi \delta(a,t) = \langle \delta, \psi_{a,t} \rangle = \psi_{a,t}(0).$$

- If $t = 0$, then

$$\mathcal{W}_\psi \delta(a,0) = \psi_{a,0}(0) = a^{-\frac{1}{2}}\, \psi(0) \sim O(a^{-\frac{1}{2}}).$$

Example: Dirac Delta

Let ψ be a well-localized wavelet (e.g., Schwartz class) on $\mathbb{R}$, and δ be the Dirac delta.



We have:

$$\mathcal{W}_\psi \delta(a, t) = \langle \delta, \psi_{a,t} \rangle = \psi_{a,t}(0).$$

- If $t = 0$, then

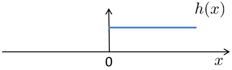
$$\mathcal{W}_\psi \delta(a, 0) = \psi_{a,0}(0) = a^{-\frac{1}{2}}\, \psi(0) \sim O(a^{-\frac{1}{2}}).$$

- If $t \neq 0$, then, for each $k \in \mathbb{N}$, there is a constant C_k such that

$$|\mathcal{W}_\psi \delta(a, t)| = |\psi_{a,t}(0)| \leq C_k\, a^k, \quad a \to 0.$$

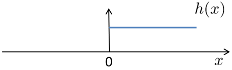
Example: Heaviside function

Let ψ be a well-localized wavelet (e.g., Schwartz class) on $\mathbb{R}$, and $h(x) = 1$ if $x \geq 0$, $h(x) = 0$ if $x < 0$.



Example: Heaviside function

Let ψ be a well-localized wavelet (e.g., Schwartz class) on $\mathbb{R}$, and $h(x) = 1$ if $x \geq 0$, $h(x) = 0$ if $x < 0$.

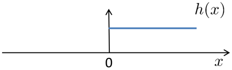


We have:

$$\mathcal{W}_\psi h(a,t) \; = \; \left\langle \hat{h}, \hat{\psi}_{a,t} \right\rangle \; = \sqrt{a} \int_{\mathbb{R}} \frac{1}{2\pi i \xi} \, \overline{\hat{\psi}(a\xi)} \, e^{-2\pi i \xi t} \, d\xi$$

Example: Heaviside function

Let ψ be a well-localized wavelet (e.g., Schwartz class) on $\mathbb{R}$, and $h(x) = 1$ if $x \geq 0$, $h(x) = 0$ if $x < 0$.

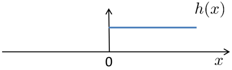


We have:

$$\begin{aligned} \mathcal{W}_\psi h(a,t) \;=\; \left\langle \hat{h}, \hat{\psi}_{a,t} \right\rangle \quad &= \sqrt{a} \int_{\mathbb{R}} \frac{1}{2\pi i \xi} \, \overline{\hat{\psi}(a\xi)} \, e^{-2\pi i \xi t} \, d\xi \\ \text{(set } \hat{\gamma}(\eta) = \frac{1}{2\pi i \eta} \overline{\hat{\psi}(\eta)}\text{)} \quad &= \sqrt{a} \int_{\mathbb{R}} \hat{\gamma}(\eta) \, e^{-2\pi i \eta \frac{t}{a}} \, d\eta \end{aligned}$$

Example: Heaviside function

Let ψ be a well-localized wavelet (e.g., Schwartz class) on $\mathbb{R}$, and $h(x) = 1$ if $x \geq 0$, $h(x) = 0$ if $x < 0$.

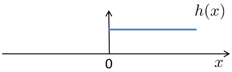


We have:

$$\begin{aligned}
\mathcal{W}_\psi h(a,t) \; = \; \left\langle \hat{h}, \hat{\psi}_{a,t} \right\rangle \quad & = \sqrt{a} \int_{\mathbb{R}} \frac{1}{2\pi i \xi} \, \overline{\hat{\psi}(a\xi)} \, e^{-2\pi i \xi t} \, d\xi \\
\left(\text{set } \hat{\gamma}(\eta) = \frac{1}{2\pi i \eta} \overline{\hat{\psi}(\eta)}\right) \quad & = \sqrt{a} \int_{\mathbb{R}} \hat{\gamma}(\eta) \, e^{-2\pi i \eta \frac{t}{a}} \, d\eta \\
& = \sqrt{a} \, \gamma(-t/a)
\end{aligned}$$

Example: Heaviside function

Let ψ be a well-localized wavelet (e.g., Schwartz class) on $\mathbb{R}$, and $h(x) = 1$ if $x \geq 0$, $h(x) = 0$ if $x < 0$.



We have:

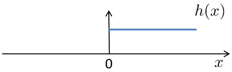
$$\begin{aligned} \mathcal{W}_\psi h(a,t) \; = \; \left\langle \hat{h}, \hat{\psi}_{a,t} \right\rangle \; &= \sqrt{a} \int_{\mathbb{R}} \frac{1}{2\pi i \xi} \, \overline{\hat{\psi}(a\xi)} \, e^{-2\pi i \xi t} \, d\xi \\ \left(\text{set } \hat{\gamma}(\eta) = \frac{1}{2\pi i \eta} \overline{\hat{\psi}(\eta)}\right) \quad &= \sqrt{a} \int_{\mathbb{R}} \hat{\gamma}(\eta) \, e^{-2\pi i \eta \frac{t}{a}} \, d\eta \\ &= \sqrt{a} \, \gamma(-t/a) \end{aligned}$$

- If $t = 0$, provided $\int \hat{\gamma}(\eta) d\eta \neq 0$, then

$$|\mathcal{W}_\psi h(a,0)| \approx \sqrt{a}.$$

Example: Heaviside function

Let ψ be a well-localized wavelet (e.g., Schwartz class) on $\mathbb{R}$, and $h(x) = 1$ if $x \geq 0$, $h(x) = 0$ if $x < 0$.



We have:

$$
\begin{aligned}
\mathcal{W}_\psi h(a,t) \;=\; \left\langle \hat{h}, \hat{\psi}_{a,t} \right\rangle \quad &= \sqrt{a} \int_{\mathbb{R}} \frac{1}{2\pi i \xi} \, \overline{\hat{\psi}(a\xi)} \, e^{-2\pi i \xi t} \, d\xi \\
\left(\text{set } \hat{\gamma}(\eta) = \frac{1}{2\pi i \eta} \, \overline{\hat{\psi}(\eta)}\right) \quad &= \sqrt{a} \int_{\mathbb{R}} \hat{\gamma}(\eta) \, e^{-2\pi i \eta \frac{t}{a}} \, d\eta \\
&= \sqrt{a} \, \gamma(-t/a)
\end{aligned}
$$

- If $t = 0$, provided $\int \hat{\gamma}(\eta) d\eta \neq 0$, then

$$|\mathcal{W}_\psi h(a,0)| \approx \sqrt{a}.$$

- If $t \neq 0$, for any $k \in \mathbb{N}$,

$$|\mathcal{W}_\psi h(a,0)| \leq C_k \, a^k, \quad a \to 0.$$

The Continuous Wavelet Transform

In general...

The Continuous Wavelet Transform

In general...

For an appropriate well-localized wavelet ψ:

The Continuous Wavelet Transform

In general...

For an appropriate well-localized wavelet ψ:

- $\mathcal{W}_\psi f(a,t) \to 0$, rapidly as $a \to 0$, for t away from singularities;

The Continuous Wavelet Transform

In general...

For an appropriate well-localized wavelet ψ:

- $\mathcal{W}_\psi f(a,t) \to 0$, rapidly as $a \to 0$, for t away from singularities;
- $\mathcal{W}_\psi f(a,t_0) \to 0$, "slowly" (not rapidly), as $a \to 0$, if $t = t_0$ is a singularity.

The Continuous Wavelet Transform

In general...

For an appropriate well-localized wavelet ψ:

- $\mathcal{W}_\psi f(a,t) \to 0$, rapidly as $a \to 0$, for t away from singularities;
- $\mathcal{W}_\psi f(a,t_0) \to 0$, "slowly" (not rapidly), as $a \to 0$, if $t = t_0$ is a singularity.

Locations $t = t_0$ of "slow" (not-rapid) asymptotic decay of $\mathcal{W}_\psi f(a,t)$, as $a \to 0$, are exactly those points where f is singular.

The Continuous Wavelet Transform

In general...

For an appropriate well-localized wavelet ψ:

- $\mathcal{W}_\psi f(a, t) \to 0$, rapidly as $a \to 0$, for t away from singularities;
- $\mathcal{W}_\psi f(a, t_0) \to 0$, "slowly" (not rapidly), as $a \to 0$, if $t = t_0$ is a singularity.

Locations $t = t_0$ of "slow" (not-rapid) asymptotic decay of $\mathcal{W}_\psi f(a, t)$, as $a \to 0$, are exactly those points where f is singular.

The continuous wavelet transform resolves the singular support

The Continuous Wavelet Transform

In higher dimensions...

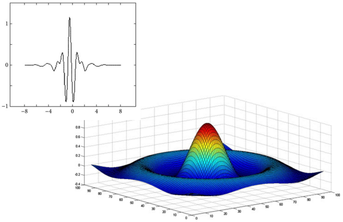
The Continuous Wavelet Transform

In higher dimensions...

The simplest way to extend the continuous wavelet transform to $\mathbb{R}^d$ is by considering the affine systems

$$\{\psi_{a,t}(x) = a^{-\frac{d}{2}}\psi(a^{-1}(x-t)) : \ a > 0, t \in \mathbb{R}^d\},$$

where $\psi \in L^2(\mathbb{R}^d)$.



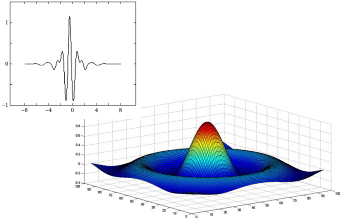
The Continuous Wavelet Transform

In higher dimensions...

The simplest way to extend the continuous wavelet transform to $\mathbb{R}^d$ is by considering the affine systems

$$\{\psi_{a,t}(x) = a^{-\frac{d}{2}}\psi(a^{-1}(x-t)) : \; a > 0, t \in \mathbb{R}^d\},$$

where $\psi \in L^2(\mathbb{R}^d)$.



Similar to the 1D case, it can detect point-singularities and resolve the singular support.

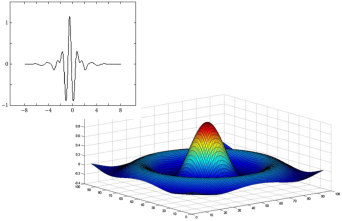
The Continuous Wavelet Transform

In higher dimensions...

The simplest way to extend the continuous wavelet transform to $\mathbb{R}^d$ is by considering the affine systems

$$\{\psi_{a,t}(x) = a^{-\frac{d}{2}}\psi(a^{-1}(x-t)) :\ a > 0, t \in \mathbb{R}^d\},$$

where $\psi \in L^2(\mathbb{R}^d)$.



Similar to the 1D case, it can detect point-singularities and resolve the singular support.

However, it provides very limited information about the geometry of singularities of multivariate functions and distributions.

Continuous Shearlet Transform (D=2)

Shearlets are derived from the framework of affine systems.

Continuous Shearlet Transform (D=2)

Shearlets are derived from the framework of affine systems.

The **full affine group of motions on** $\mathbb{R}^2$ is the set

$$\mathbb{A} = \{(M, t) : \ M \in GL_2(\mathbb{R}), t \in \mathbb{R}^2\}$$

with group operation $(M, t) \cdot (M', t') = (MM', t + Mt')$.

Continuous Shearlet Transform (D=2)

Shearlets are derived from the framework of affine systems.

The **full affine group of motions on** $\mathbb{R}^2$ is the set

$$\mathbb{A} = \{(M,t) : \; M \in GL_2(\mathbb{R}), t \in \mathbb{R}^2\}$$

with group operation $(M,t) \cdot (M',t') = (MM', t + Mt')$.
We consider subgroups $\mathbb{A}_G$ of $\mathbb{A}$ of the form

$$\mathbb{A}_G = \{(M,t) : \; M \in G \subset GL_2(\mathbb{R}), t \in \mathbb{R}^2\}$$

where G is referred to as the **dilation subgroup**.

Continuous Shearlet Transform (D=2)

Shearlets are derived from the framework of affine systems.

The **full affine group of motions on** $\mathbb{R}^2$ is the set

$$\mathbb{A} = \{(M,t) : \; M \in GL_2(\mathbb{R}), t \in \mathbb{R}^2\}$$

with group operation $(M,t)\cdot(M',t') = (MM', t + Mt')$.
We consider subgroups $\mathbb{A}_G$ of $\mathbb{A}$ of the form

$$\mathbb{A}_G = \{(M,t) : \; M \in G \subset GL_2(\mathbb{R}), t \in \mathbb{R}^2\}$$

where G is referred to as the **dilation subgroup**.
The **affine system** generated by $\psi \in L^2(\mathbb{R}^2)$ and $\mathbb{A}_G$ is

$$\left\{\psi_{M,t}(x) = |\det M|^{-1/2}\psi(M^{-1}(x-t)) : \; (M,t) \in \mathbb{A}_G\right\}.$$

Continuous Shearlet Transform (D=2)

Under appropriate admissibility conditions on ψ, it may be possible to define a (generalized) continuous wavelet transform associated with $\mathbb{A}_G$.
(Note: not all $\mathbb{A}_G$ *have admissible functions)*

Continuous Shearlet Transform (D=2)

Under appropriate admissibility conditions on ψ, it may be possible to define a (generalized) continuous wavelet transform associated with $\mathbb{A}_G$. *(Note: not all $\mathbb{A}_G$ have admissible functions)*

In this case, the continuous wavelet transform associated with $\mathbb{A}_G$

$$\mathcal{W}_\psi : f \to \mathcal{W}_\psi f(M,t) = \langle f, \psi_{M,t} \rangle, \quad \text{for } (M,t) \in \mathbb{A}_G,$$

is a linear isometry from $L^2(\mathbb{R}^2)$ into $L^2(\mathbb{A}_G)$.

Continuous Shearlet Transform (D=2)

Under appropriate admissibility conditions on ψ, it may be possible to define a (generalized) continuous wavelet transform associated with $\mathbb{A}_G$. *(Note: not all $\mathbb{A}_G$ have admissible functions)*

In this case, the continuous wavelet transform associated with $\mathbb{A}_G$

$$\mathcal{W}_\psi : f \to \mathcal{W}_\psi f(M,t) = \langle f, \psi_{M,t} \rangle, \quad \text{for } (M,t) \in \mathbb{A}_G,$$

is a linear isometry from $L^2(\mathbb{R}^2)$ into $L^2(\mathbb{A}_G)$.
For all $f \in L^2(\mathbb{R}^2)$

$$f(x) = \int_{\mathbb{R}^2} \int_G \langle f, \psi_{M,t} \rangle \, \psi_{M,t}(x) \, d\lambda(M) \, dt,$$

where λ is the left Haar measure on G.

Example: G = isotropic dilations

- **Isotropic dilations.** The dilation group G is:

$$G = \left\{ M = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} : a > 0 \right\}$$

Example: G = isotropic dilations

- **Isotropic dilations.** The dilation group G is:

$$G = \left\{ M = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} : a > 0 \right\}$$

Admissibility is given by the classical Calderón condition.
This group is associated with the conventional continuous wavelet systems

$$\left\{ \psi_{a,t}(x) = a^{-1}\, \psi(a^{-1}(x-t)) : \ a > 0, t \in \mathbb{R}^2 \right\}$$

Example: G = shearlet group [K,Labate,2009],[Dahlke et al,2008]

- **Shearlet group.** The dilation group G is:

$$G = \left\{ M_{as} = \begin{pmatrix} a & -\sqrt{a}\,s \\ 0 & \sqrt{a} \end{pmatrix}, \quad a > 0,\ s \in \mathbb{R} \right\}$$

Example: G = shearlet group [K,Labate,2009],[Dahlke et al,2008]

- **Shearlet group.** The dilation group G is:

$$G = \left\{ M_{as} = \begin{pmatrix} a & -\sqrt{a}\, s \\ 0 & \sqrt{a} \end{pmatrix}, \quad a > 0,\, s \in \mathbb{R} \right\}$$

We have the factorization

$$M_{as} = \begin{pmatrix} a & -\sqrt{a}\, s \\ 0 & \sqrt{a} \end{pmatrix} = \begin{pmatrix} 1 & -s \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}$$

into **anisotropic dilation** $\begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}$ and **shear transformation** $\begin{pmatrix} 1 & -s \\ 0 & 1 \end{pmatrix}$

NOTE: $\sqrt{a}$ can be replaced by a^α, $0 < \alpha < 1$.

Example: G = shearlet group [K,Labate,2009],[Dahlke et al,2008]

- **Shearlet group.** The dilation group G is:

$$G = \left\{ M_{as} = \begin{pmatrix} a & -\sqrt{a}\, s \\ 0 & \sqrt{a} \end{pmatrix}, \quad a > 0,\ s \in \mathbb{R} \right\}$$

We have the factorization

$$M_{as} = \begin{pmatrix} a & -\sqrt{a}\, s \\ 0 & \sqrt{a} \end{pmatrix} = \begin{pmatrix} 1 & -s \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}$$

into **anisotropic dilation** $\begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}$ and **shear transformation** $\begin{pmatrix} 1 & -s \\ 0 & 1 \end{pmatrix}$

NOTE: $\sqrt{a}$ can be replaced by a^{α}, $0 < \alpha < 1$.

A system associated with this group is a **continuous shearlet system**

$$\left\{ \psi_{a,s,t}(x) = a^{-3/4}\, \psi(M_{as}^{-1}(x-t)) :\ a \in \mathbb{R}^+, s \in \mathbb{R}, t \in \mathbb{R}^2 \right\}$$

Construction of Continuous Shearlets

There are many admissible shearlets.

Band-limited shearlets [Guo,Kutyniok,L, 2006]. We choose:

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\, \hat{\psi}_2(\frac{\xi_2}{\xi_1}),$$

where

Construction of Continuous Shearlets

There are many admissible shearlets.

Band-limited shearlets [Guo,Kutyniok,L, 2006]. We choose:

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\, \hat{\psi}_2(\frac{\xi_2}{\xi_1}),$$

where

- ψ_1 is a continuous wavelet with $\hat{\psi}_1 \in C^\infty(\mathbb{R})$ and supp $\hat{\psi}_1 \subset [-2, -\frac{1}{2}] \cup [\frac{1}{2}, 2]$.

Construction of Continuous Shearlets

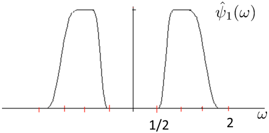
There are many admissible shearlets.

Band-limited shearlets [Guo,Kutyniok,L, 2006]. We choose:

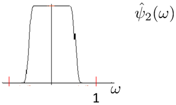
$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\, \hat{\psi}_2\Big(\frac{\xi_2}{\xi_1}\Big),$$

where

- ψ_1 is a continuous wavelet with $\hat{\psi}_1 \in C^\infty(\mathbb{R})$ and supp $\hat{\psi}_1 \subset [-2, -\frac{1}{2}] \cup [\frac{1}{2}, 2]$.



- ψ_2 satisfies $\hat{\psi}_2 \in C^\infty(\mathbb{R})$, supp $\hat{\psi}_2 \subset [-1, 1]$ and $\|\psi_2\| = 1$.



Construction of Continuous Shearlets

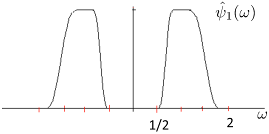
There are many admissible shearlets.

Band-limited shearlets [Guo,Kutyniok,L, 2006]. We choose:

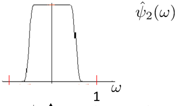
$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\,\hat{\psi}_2\Big(\frac{\xi_2}{\xi_1}\Big),$$

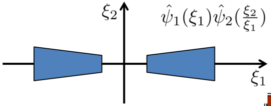
where

- ψ_1 is a continuous wavelet with $\hat{\psi}_1 \in C^\infty(\mathbb{R})$ and supp $\hat{\psi}_1 \subset [-2, -\frac{1}{2}] \cup [\frac{1}{2}, 2]$.



- ψ_2 satisfies $\hat{\psi}_2 \in C^\infty(\mathbb{R})$, supp $\hat{\psi}_2 \subset [-1, 1]$ and $\|\psi_2\| = 1$.





Hence ψ is a smooth bandlimited function.

Construction of Continuous Shearlets

Alternatively...

Compactly supported shearlets
[Lim,Kutyniok,2011] [Kutyniok,Petersen,2015]. We choose:

$$\psi(x_1, x_2) = \psi_1(x_1)\phi(x_2)$$

where

Construction of Continuous Shearlets

Alternatively...

Compactly supported shearlets
[Lim,Kutyniok,2011] [Kutyniok,Petersen,2015]. We choose:

$$\psi(x_1, x_2) = \psi_1(x_1)\phi(x_2)$$

where

- ψ_1 is a compactly supported continuous wavelet.

Construction of Continuous Shearlets

Alternatively...

Compactly supported shearlets
[Lim,Kutyniok,2011] [Kutyniok,Petersen,2015]. We choose:

$$\psi(x_1, x_2) = \psi_1(x_1)\phi(x_2)$$

where

- ψ_1 is a compactly supported continuous wavelet.
- $\phi \in C^2([-r, r])$ satisfies $\phi(0) = 0$, $\phi'(0) \neq 0$, $\int \phi(x)dx > C > 0$.

Construction of Continuous Shearlets

Alternatively...

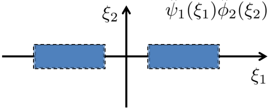
Compactly supported shearlets
[Lim,Kutyniok,2011] [Kutyniok,Petersen,2015]. We choose:

$$\psi(x_1, x_2) = \psi_1(x_1)\phi(x_2)$$

where

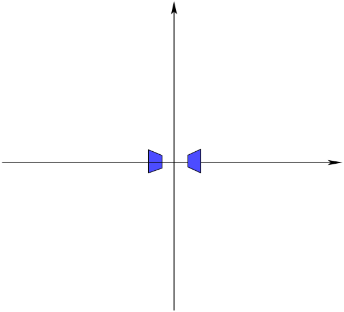
- ψ_1 is a compactly supported continuous wavelet.
- $\phi \in C^2([-r, r])$ satisfies $\phi(0) = 0$, $\phi'(0) \neq 0$, $\int \phi(x)dx > C > 0$.

Hence ψ is a compactly supported function.



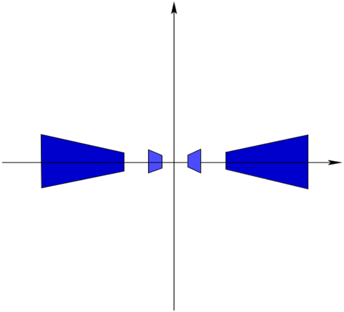
Construction of Continuous Shearlets

The elements of a shearlet system $\{\psi_{a,s,t}\}$ are a well localized waveforms, with **orientation** controlled by the shear parameter s, and increasingly **elongated** at fine scales ($a \to 0$).



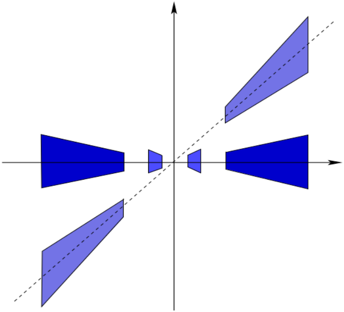
Construction of Continuous Shearlets

The elements of a shearlet system $\{\psi_{a,s,t}\}$ are a well localized waveforms, with orientation controlled by the shear parameter s, and increasingly thin at fine scales ($a \to 0$).



Construction of Continuous Shearlets

The elements of a shearlet system $\{\psi_{a,s,t}\}$ are a well localized waveforms, with orientation controlled by the shear parameter s, and increasingly thin at fine scales ($a \to 0$).



Construction of Continuous Shearlets

Choosing an admissible function ψ, the **Continuous Shearlet Transform**

$$\mathcal{SH}_\psi : f \to \mathcal{SH}_\psi f(a,s,t) = \langle f, \psi_{a,s,t} \rangle ,$$

is a linear isometry from $L^2(\mathbb{R}^2)$ to $L^2(\mathbb{A}_G)$.

Construction of Continuous Shearlets

Choosing an admissible function ψ, the **Continuous Shearlet Transform**

$$\mathcal{SH}_\psi : f \to \mathcal{SH}_\psi f(a,s,t) = \langle f, \psi_{a,s,t} \rangle \,,$$

is a linear isometry from $L^2(\mathbb{R}^2)$ to $L^2(\mathbb{A}_G)$. Hence

$$\|f\|^2 = \int_{\mathbb{R}^2} \int_{\mathbb{R}^+} \int_0^\infty |\mathcal{SH}_\psi f(a,s,t)|^2 \, \frac{da}{a^3} \, ds \, dt.$$

Construction of Continuous Shearlets

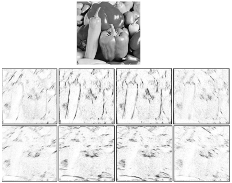
Choosing an admissible function ψ, the **Continuous Shearlet Transform**

$$\mathcal{SH}_\psi : f \to \mathcal{SH}_\psi f(a,s,t) = \langle f, \psi_{a,s,t} \rangle ,$$

is a linear isometry from $L^2(\mathbb{R}^2)$ to $L^2(\mathbb{A}_G)$. Hence

$$\|f\|^2 = \int_{\mathbb{R}^2} \int_{\mathbb{R}^+} \int_0^\infty |\mathcal{SH}_\psi f(a,s,t)|^2 \, \frac{da}{a^3} \, ds \, dt.$$

$\mathcal{SH}_\psi f(a,s,t)$ measures the content of f as a function of the **scale** a, the **shear** s and the **location** t.



Construction of Continuous Shearlets

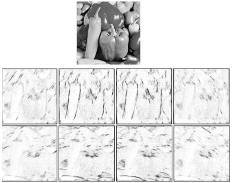
Choosing an admissible function ψ, the **Continuous Shearlet Transform**

$$\mathcal{SH}_\psi : f \to \mathcal{SH}_\psi f(a,s,t) = \langle f, \psi_{a,s,t} \rangle \,,$$

is a linear isometry from $L^2(\mathbb{R}^2)$ to $L^2(\mathbb{A}_G)$. Hence

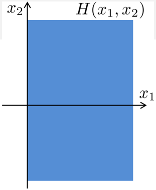
$$\|f\|^2 = \int_{\mathbb{R}^2} \int_{\mathbb{R}^+} \int_0^\infty |\mathcal{SH}_\psi f(a,s,t)|^2 \, \frac{da}{a^3} \, ds \, dt.$$

$\mathcal{SH}_\psi f(a,s,t)$ measures the content of f as a function of the **scale** a, the **shear** s and the **location** t.



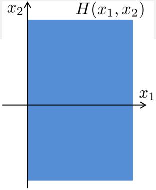
It is able to resolve both the location and orientation of singularities.

Example: Heaviside function (2D)



Let $H(x_1, x_2) = \chi_{x_1>0}(x_1, x_2)$.

Example: Heaviside function (2D)

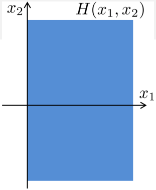


Let $H(x_1, x_2) = \chi_{x_1>0}(x_1, x_2)$.

Then:

$$\mathcal{SH}_\psi H(a, s, t) = \int_{\mathbb{R}^2} \hat{H}(\xi)\,\overline{\hat{\psi}_{a,s,t}}(\xi)\,d\xi \;=\; a^{\frac{3}{4}} \int_{\mathbb{R}} \frac{\overline{\hat{\psi}_1}(a\,\xi_1)}{2\pi i\xi_1}\,\overline{\hat{\psi}_2}(a^{-\frac{1}{2}}s)e^{2\pi i\xi_1 t_1}\,d\xi_1$$

Example: Heaviside function (2D)



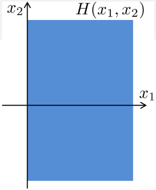
Let $H(x_1, x_2) = \chi_{x_1>0}(x_1, x_2)$.

Then:

$$\mathcal{SH}_\psi H(a,s,t) = \int_{\mathbb{R}^2} \hat{H}(\xi)\, \overline{\hat{\psi}_{a,s,t}}(\xi)\, d\xi \;= a^{\frac{3}{4}} \int_{\mathbb{R}} \frac{\overline{\hat{\psi}_1}(a\,\xi_1)}{2\pi i \xi_1}\, \overline{\hat{\psi}_2}(a^{-\frac{1}{2}} s) e^{2\pi i \xi_1 t_1}\, d\xi_1$$

$$\text{(set } \hat{\gamma}(\eta) = \frac{1}{2\pi i \eta}\, \overline{\hat{\psi}_1(\eta)}) \qquad = a^{\frac{3}{4}}\, \overline{\hat{\psi}_2}(a^{-1/2} s) \int_{\mathbb{R}} \hat{\gamma}(\eta)\, e^{2\pi i \eta \frac{t_1}{a}}\, d\eta$$

Example: Heaviside function (2D)



Let $H(x_1, x_2) = \chi_{x_1>0}(x_1, x_2)$.

Then:

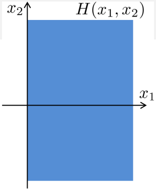
$$\mathcal{SH}_\psi H(a,s,t) = \int_{\mathbb{R}^2} \hat{H}(\xi)\, \overline{\hat{\psi}_{a,s,t}}(\xi)\, d\xi \;=\; a^{\frac{3}{4}} \int_{\mathbb{R}} \frac{\overline{\hat{\psi}_1}(a\,\xi_1)}{2\pi i \xi_1}\, \overline{\hat{\psi}_2}(a^{-\frac{1}{2}}s) e^{2\pi i \xi_1 t_1}\, d\xi_1$$

$$\left(\text{set } \hat{\gamma}(\eta) = \frac{1}{2\pi i \eta}\, \overline{\hat{\psi}_1(\eta)}\right) \qquad = a^{\frac{3}{4}}\, \overline{\hat{\psi}_2}(a^{-1/2}s) \int_{\mathbb{R}} \hat{\gamma}(\eta)\, e^{2\pi i \eta \frac{t_1}{a}}\, d\eta$$

- If $t_1 \neq 0$, since $\hat{\psi}_1 \in C_c^\infty(\mathbb{R})$, for any $k \in \mathbb{N}$

$$\mathcal{SH}_\psi H(a,s,t) \leq C_k\, a^k, \quad \text{as } a \to 0.$$

Example: Heaviside function (2D)



Let $H(x_1, x_2) = \chi_{x_1>0}(x_1, x_2)$.

Then:

$$\mathcal{SH}_\psi H(a,s,t) = \int_{\mathbb{R}^2} \hat{H}(\xi)\,\overline{\hat{\psi}_{a,s,t}}(\xi)\,d\xi \;=\; a^{\frac{3}{4}} \int_{\mathbb{R}} \frac{\overline{\hat{\psi}_1}(a\,\xi_1)}{2\pi i \xi_1}\,\overline{\hat{\psi}_2}(a^{-\frac{1}{2}}s)e^{2\pi i \xi_1 t_1}\,d\xi_1$$

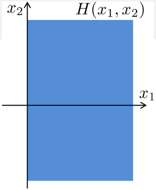
$$\left(\text{set } \hat{\gamma}(\eta) = \frac{1}{2\pi i \eta}\,\overline{\hat{\psi}_1(\eta)}\right) \qquad = a^{\frac{3}{4}}\,\overline{\hat{\psi}_2}(a^{-1/2}s)\int_{\mathbb{R}} \hat{\gamma}(\eta)\,e^{2\pi i \eta \frac{t_1}{a}}\,d\eta$$

- If $t_1 \neq 0$, since $\hat{\psi}_1 \in C_c^\infty(\mathbb{R})$, for any $k \in \mathbb{N}$

$$\mathcal{SH}_\psi H(a,s,t) \le C_k\,a^k, \quad \text{as } a \to 0.$$

- If $t_1 = 0$ and $s \neq 0$, the term $\overline{\hat{\psi}_2}(a^{-1/2}s)$ will vanish as $a \to 0$.

Example: Heaviside function (2D)



Let $H(x_1, x_2) = \chi_{x_1>0}(x_1, x_2)$.

Then:

$$\mathcal{SH}_\psi H(a,s,t) = \int_{\mathbb{R}^2} \hat{H}(\xi)\,\overline{\hat{\psi}_{a,s,t}}(\xi)\,d\xi \;=\; a^{\frac{3}{4}} \int_{\mathbb{R}} \frac{\overline{\hat{\psi}_1(a\,\xi_1)}}{2\pi i\xi_1}\,\overline{\hat{\psi}_2}(a^{-\frac{1}{2}}s)e^{2\pi i \xi_1 t_1}\,d\xi_1$$

$$\left(\text{set } \hat{\gamma}(\eta) = \frac{1}{2\pi i\eta}\,\overline{\hat{\psi}_1(\eta)}\right) \qquad = a^{\frac{3}{4}}\,\overline{\hat{\psi}_2}(a^{-1/2}s)\int_{\mathbb{R}} \hat{\gamma}(\eta)\,e^{2\pi i\eta\frac{t_1}{a}}\,d\eta$$

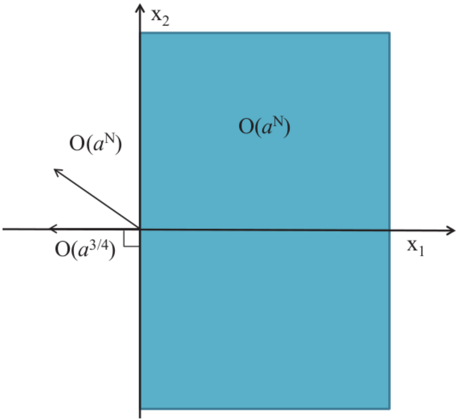
- If $t_1 \neq 0$, since $\hat{\psi}_1 \in C_c^\infty(\mathbb{R})$, for any $k \in \mathbb{N}$

$$\mathcal{SH}_\psi H(a,s,t) \le C_k\,a^k, \quad \text{as } a \to 0.$$

- If $t_1 = 0$ and $s \neq 0$, the term $\overline{\hat{\psi}_2}(a^{-1/2}s)$ will vanish as $a \to 0$.
- If $t_1 = 0$ and $s = 0$, provided $\hat{\psi}_2(0) \neq 0$ and $\int_{\mathbb{R}} \hat{\gamma}(\eta)\,d\eta \neq 0$, we have

$$\mathcal{SH}_\psi H(a,0,(0,t_2)) = O(a^{\frac{3}{4}}).$$

Example: Heaviside function (2D)



$\mathcal{SH}_\psi H(a, s, t)$ decays rapidly for all values of s and $t = (t_1, t_2)$, except for $s = 0$ and $t_2 = 0$

Resolution of the Wavefront Set

The **Continuous Shearlet Transform** of f

$$\mathcal{SH}_\psi s(a,s,t) = \langle f, \psi_{a,s,t} \rangle\,, \quad a \in \mathbb{R}^+, s \in \mathbb{R}, t \in \mathbb{R}^2$$

describes the **geometry of the singularities** of f through its decay at fine scales.

Resolution of the Wavefront Set

The **Continuous Shearlet Transform** of f

$$\mathcal{SH}_\psi s(a,s,t) = \langle f, \psi_{a,s,t} \rangle, \quad a \in \mathbb{R}^+, s \in \mathbb{R}, t \in \mathbb{R}^2$$

describes the **geometry of the singularities** of f through its decay at fine scales.

- $\mathcal{SH}_\psi f$ characterizes the **wavefront set** of a distribution f through its decay at fine scales [Kutyniok,L,2009], [Grohs, 2011].

Resolution of the Wavefront Set

The **Continuous Shearlet Transform** of f

$$\mathcal{SH}_\psi s(a,s,t) = \langle f, \psi_{a,s,t} \rangle\,, \quad a \in \mathbb{R}^+, s \in \mathbb{R}, t \in \mathbb{R}^2$$

describes the **geometry of the singularities** of f through its decay at fine scales.

- $\mathcal{SH}_\psi f$ characterizes the **wavefront set** of a distribution f through its decay at fine scales [Kutyniok,L,2009], [Grohs, 2011].
- The *continuous curvelet transform* has similar properties [Candès,Donoho,2005].

Resolution of the Wavefront Set

The **Continuous Shearlet Transform** of f

$$\mathcal{SH}_\psi s(a,s,t) = \langle f, \psi_{a,s,t} \rangle, \quad a \in \mathbb{R}^+, s \in \mathbb{R}, t \in \mathbb{R}^2$$

describes the **geometry of the singularities** of f through its decay at fine scales.

- $\mathcal{SH}_\psi f$ characterizes the **wavefront set** of a distribution f through its decay at fine scales [Kutyniok,L,2009], [Grohs, 2011].
- The *continuous curvelet transform* has similar properties [Candès,Donoho,2005].
- $\mathcal{SH}_\psi f$ provides a precise description of the **geometry** of **piecewise-smooth edges** of f through its asymptotic decay at fine scales [Guo,L,2008-2015]. This holds also in 3D.

Resolution of edges using the CST ($d = 2$)

Theorem [Guo,L] *Let $B = \chi_S$, $S \subset \mathbb{R}^2$ compact, and ∂S is piecewise smooth.*

Resolution of edges using the CST ($d = 2$)

Theorem [Guo,L] *Let $B = \chi_S$, $S \subset \mathbb{R}^2$ compact, and ∂S is piecewise smooth.*

(i) *If $t \notin S$ or if $t \in \partial S$ and s does not correspond to the normal direction of ∂S at t then*

$$\lim_{a \to 0^+} a^{-N} \mathcal{SH}_\psi B(a, s, t) = 0, \qquad \text{for all } N > 0.$$

Resolution of edges using the CST ($d = 2$)

Theorem [Guo,L] *Let $B = \chi_S$, $S \subset \mathbb{R}^2$ compact, and ∂S is piecewise smooth.*

(i) *If $t \notin S$ or if $t \in \partial S$ and s does not correspond to the normal direction of ∂S at t then*

$$\lim_{a \to 0^+} a^{-N} \mathcal{SH}_\psi B(a, s, t) = 0, \qquad \textit{for all } N > 0.$$

(ii) *If $s = s_0$ corresponds to the normal direction of ∂S at t then*

$$0 < \lim_{a \to 0^+} a^{-\frac{3}{4}} \left| \mathcal{SH}_\psi B(a, s_0, t) \right| < \infty.$$

Resolution of edges using the CST ($d = 2$)

Theorem [Guo,L] *Let $B = \chi_S$, $S \subset \mathbb{R}^2$ compact, and ∂S is piecewise smooth.*

(i) *If $t \notin S$ or if $t \in \partial S$ and s does not correspond to the normal direction of ∂S at t then*

$$\lim_{a \to 0^+} a^{-N} \mathcal{SH}_\psi B(a, s, t) = 0, \qquad \text{for all } N > 0.$$

(ii) *If $s = s_0$ corresponds to the normal direction of ∂S at t then*

$$0 < \lim_{a \to 0^+} a^{-\frac{3}{4}} \left|\mathcal{SH}_\psi B(a, s_0, t)\right| < \infty.$$

Resolution of edges using the CST ($d = 2$)

Theorem [Guo,L] *Let $B = \chi_S$, $S \subset \mathbb{R}^2$ compact, and ∂S is piecewise smooth.*

(i) *If $t \notin S$ or if $t \in \partial S$ and s does not correspond to the normal direction of ∂S at t then*

$$\lim_{a \to 0^+} a^{-N} \mathcal{SH}_\psi B(a, s, t) = 0, \qquad \textit{for all } N > 0.$$

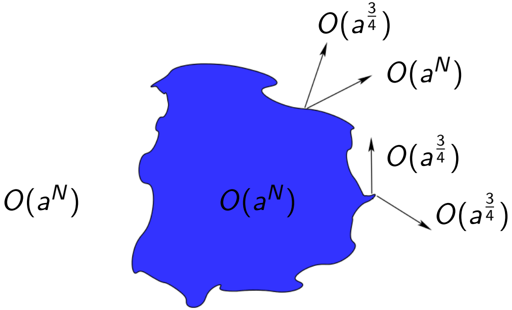
(ii) *If $s = s_0$ corresponds to the normal direction of ∂S at t then*

$$0 < \lim_{a \to 0^+} a^{-\frac{3}{4}} \left|\mathcal{SH}_\psi B(a, s_0, t)\right| < \infty.$$

That is, $\mathcal{SH}_\psi B$ has slow asymptotic decay only at the edge points for normal orientations, where

$$\mathcal{SH}_\psi B(a, s_0, t) = O(a^{\frac{3}{4}}) \quad \text{as } a \to 0$$

Resolution of Edges (D=2)



At the **regular points** t on an edge, for normal orientation, the shearlet transform decays as $O(a^{\frac{3}{4}})$. For all other values of s, the decay is as fast as $O(a^N)$, for any $N \in \mathbb{N}$.

At the **corner points**, the shearlet transform decays as $O(a^{\frac{3}{4}})$ for both normal orientations.

Extensions and generalizations

The shearlet analysis of discontinuities extends to:

Extensions and generalizations

The shearlet analysis of discontinuities extends to:

Extensions and generalizations

The shearlet analysis of discontinuities extends to:

- Functions $f = \sum f_i \chi_{S_i}$ where f_i are smooth functions and the boundary sets ∂S_i may contain corner points [Guo,L,2009,2015].

Extensions and generalizations

The shearlet analysis of discontinuities extends to:

- Functions $f = \sum f_i \chi_{S_i}$ where f_i are smooth functions and the boundary sets ∂S_i may contain corner points [Guo,L,2009,2015].
- Characterization of edges with uniform decay estimates [Kutyniok,Petersen,2015].

Extensions and generalizations

The shearlet analysis of discontinuities extends to:

- Functions $f = \sum f_i \chi_{S_i}$ where f_i are smooth functions and the boundary sets ∂S_i may contain corner points [Guo,L,2009,2015].
- Characterization of edges with uniform decay estimates [Kutyniok,Petersen,2015].
- Characterization of edge curvature and flatness [Guo,L,2015].

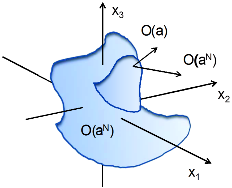
Extensions and generalizations

The shearlet analysis of discontinuities extends to higher dimensions:

Extensions and generalizations

The shearlet analysis of discontinuities extends to higher dimensions:

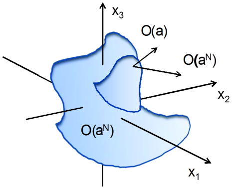
- Functions $f = \chi_S$ where $S \subset \mathbb{R}^3$ and the boundary set ∂S is a piecewise smooth boundary which may contain wedges. [Guo,L,2011],[Guo,L,2012].



Extensions and generalizations

The shearlet analysis of discontinuities extends to higher dimensions:

- Functions $f = \chi_S$ where $S \subset \mathbb{R}^3$ and the boundary set ∂S is a piecewise smooth boundary which may contain wedges. [Guo,L,2011],[Guo,L,2012].

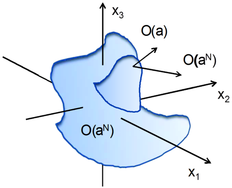


- Analysis of 3D edges and corners [Kutyniok,Petersen,2015].

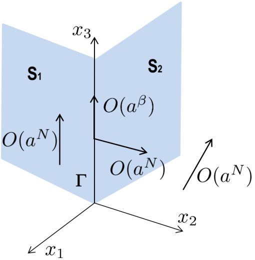
Extensions and generalizations

The shearlet analysis of discontinuities extends to higher dimensions:

- Functions $f = \chi_S$ where $S \subset \mathbb{R}^3$ and the boundary set ∂S is a piecewise smooth boundary which may contain wedges. [Guo,L,2011],[Guo,L,2012].



- Analysis of 3D edges and corners [Kutyniok,Petersen,2015].
- Analysis of one-dimensional manifolds, such as the curve of intersection of 2 surfaces. [Houska,L,2015] [Guo,L,2015]



Analysis of singularities: geometric separation

A related problem is the **geometric separation** of singularities.

Analysis of singularities: geometric separation

A related problem is the **geometric separation** of singularities.

Let $f = P + C$ where P is a collection of point-like singularities and C is a cartoon-like image.

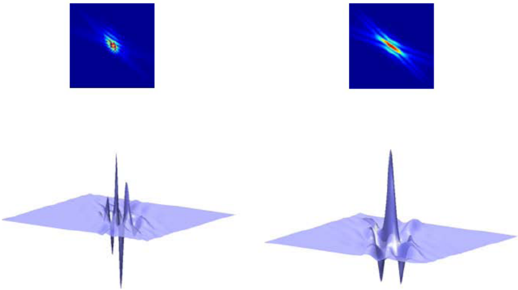
Analysis of singularities: geometric separation

A related problem is the **geometric separation** of singularities.

Let $f = P + C$ where P is a collection of point-like singularities and C is a cartoon-like image.

It is possible to separate, in a precise sense, point and curvilinear singularities in 2D [Donoho,Kutyniok, 2013] or points and piecewise linear singularities (polyhedral singularities) in 3D [Guo & L, 2014].

Applications



Some image processing applications

The microlocal properties of the Continuous Shearlet Transform and related tranforms are useful to derive efficient algorithms for

Some image processing applications

The microlocal properties of the Continuous Shearlet Transform and related tranforms are useful to derive efficient algorithms for

- *Edge and boundary detection (2D/3D)*

Some image processing applications

The microlocal properties of the Continuous Shearlet Transform and related tranforms are useful to derive efficient algorithms for

- *Edge and boundary detection (2D/3D)*
- *Estimation of edge/boundary orientation*

Some image processing applications

The microlocal properties of the Continuous Shearlet Transform and related tranforms are useful to derive efficient algorithms for

- *Edge and boundary detection (2D/3D)*
- *Estimation of edge/boundary orientation*
- *Identification of geometric features (corners, junctions,...)*

Some image processing applications

The microlocal properties of the Continuous Shearlet Transform and related tranforms are useful to derive efficient algorithms for

- *Edge and boundary detection (2D/3D)*
- *Estimation of edge/boundary orientation*
- *Identification of geometric features (corners, junctions,...)*
- *Segmentation, classification*

Some image processing applications

The microlocal properties of the Continuous Shearlet Transform and related tranforms are useful to derive efficient algorithms for

- *Edge and boundary detection (2D/3D)*
- *Estimation of edge/boundary orientation*
- *Identification of geometric features (corners, junctions,...)*
- *Segmentation, classification*
- ...

Edge Detection

Several **edge analysis and detection** algorithms based on multiscale methods and shearlets were proposed. For example:

Edge Detection

Several **edge analysis and detection** algorithms based on multiscale methods and shearlets were proposed. For example:

- [Mallat,Zhong,1992] introduce a multiscale wavelet-based algorithm for edge detection.

Edge Detection

Several **edge analysis and detection** algorithms based on multiscale methods and shearlets were proposed. For example:

- [Mallat,Zhong,1992] introduce a multiscale wavelet-based algorithm for edge detection.
- [Easley,Labate,Yi,2008], [Duval,Odone,De Vito,2015] use a multiscale shearlet-based algorithm that reinforces true edges and suppresses noise.

Edge Detection

Several **edge analysis and detection** algorithms based on multiscale methods and shearlets were proposed. For example:

- [Mallat,Zhong,1992] introduce a multiscale wavelet-based algorithm for edge detection.
- [Easley,Labate,Yi,2008], [Duval,Odone,De Vito,2015] use a multiscale shearlet-based algorithm that reinforces true edges and suppresses noise.
- [Schug,Easley,O'Leary,2011] extend idea above to 3D for surface detection.

Edge Detection

Several **edge analysis and detection** algorithms based on multiscale methods and shearlets were proposed. For example:

- [Mallat,Zhong,1992] introduce a multiscale wavelet-based algorithm for edge detection.
- [Easley,Labate,Yi,2008], [Duval,Odone,De Vito,2015] use a multiscale shearlet-based algorithm that reinforces true edges and suppresses noise.
- [Schug,Easley,O'Leary,2011] extend idea above to 3D for surface detection.
- [King at al.,2015] apply complex shearlets to detect edges in flame front images recorded by planar laser-induced fluorescence

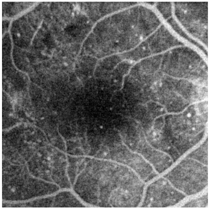
Edge Detection

Several **edge analysis and detection** algorithms based on multiscale methods and shearlets were proposed. For example:

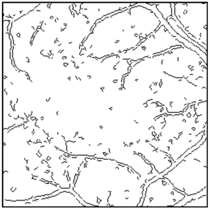
- [Mallat,Zhong,1992] introduce a multiscale wavelet-based algorithm for edge detection.
- [Easley,Labate,Yi,2008], [Duval,Odone,De Vito,2015] use a multiscale shearlet-based algorithm that reinforces true edges and suppresses noise.
- [Schug,Easley,O'Leary,2011] extend idea above to 3D for surface detection.
- [King at al.,2015] apply complex shearlets to detect edges in flame front images recorded by planar laser-induced fluorescence
- . . .

Edge Detection

Shearlet-based **edge detection** on retina images [Easley,L,Yi,2008].



Original retina image



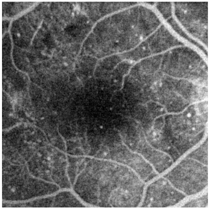
Wavelet/Canny result
FOM=0.27

Shearlet result
FOM=0.45

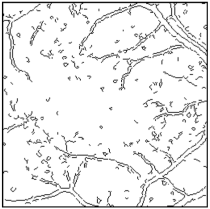
The Figure Of Merit (FOM) measures the closeness of reconstruction to the true edge map (the higher the better).

Edge Detection

Shearlet-based **edge detection** on retina images [Easley,L,Yi,2008].



Original retina image



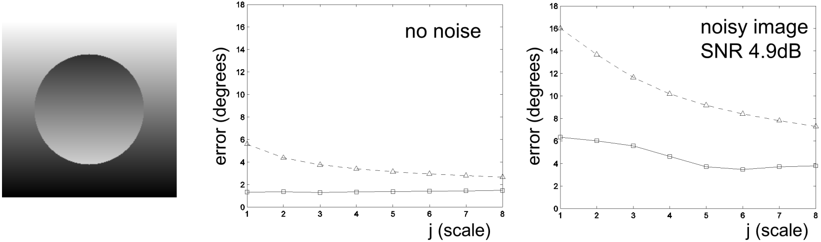
Wavelet/Canny result
FOM=0.27

Shearlet result
FOM=0.45

The Figure Of Merit (FOM) measures the closeness of reconstruction to the true edge map (the higher the better).
Shearlet-based methods yield extremely competitive results.

Edge Orientation

With respect to conventional multiscale methods, shearlets enable more accurate and robust estimation of **edge orientation**.



Average error (degrees) in estimating edge orientation using a wavelet method (dashed line) versus a shearlet method (solid line), as a function of the scale 2^{-j}.

Feature Extraction

Multiscale methods can be very useful to extract **features and landmarks** in images. For example:

Feature Extraction

Multiscale methods can be very useful to extract **features and landmarks** in images. For example:

- [Lee,Sun,Chen,1992], [Quddus,Gabbouj,2002] multiscale corner detection using wavelet transform.

Feature Extraction

Multiscale methods can be very useful to extract **features and landmarks** in images. For example:

- [Lee,Sun,Chen,1992], [Quddus,Gabbouj,2002] multiscale corner detection using wavelet transform.
- [Easley,Labate,Yi,2008], [Duval,Odone,De Vito,2015] shearlet-based corner and junction detection.

Feature Extraction

Multiscale methods can be very useful to extract **features and landmarks** in images. For example:

- [Lee,Sun,Chen,1992], [Quddus,Gabbouj,2002] multiscale corner detection using wavelet transform.
- [Easley,Labate,Yi,2008], [Duval,Odone,De Vito,2015] shearlet-based corner and junction detection.
- [Shui,Zhang,2013] corner detection using directional representations

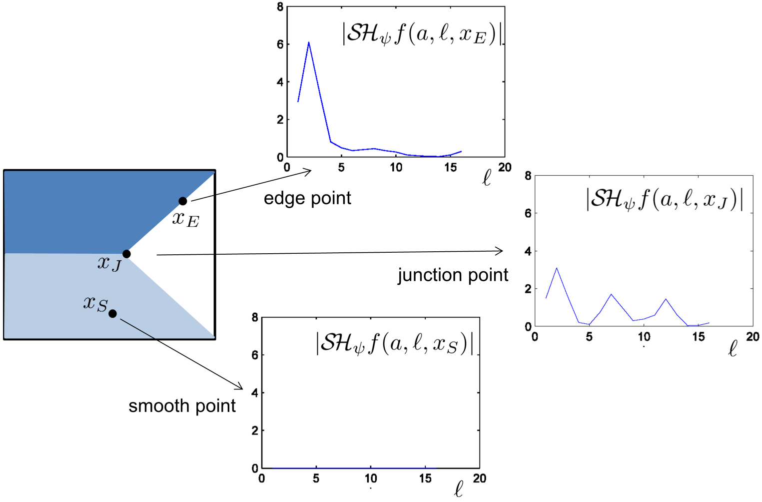
Feature Extraction

Multiscale methods can be very useful to extract **features and landmarks** in images. For example:

- [Lee,Sun,Chen,1992], [Quddus,Gabbouj,2002] multiscale corner detection using wavelet transform.
- [Easley,Labate,Yi,2008], [Duval,Odone,De Vito,2015] shearlet-based corner and junction detection.
- [Shui,Zhang,2013] corner detection using directional representations
- ...

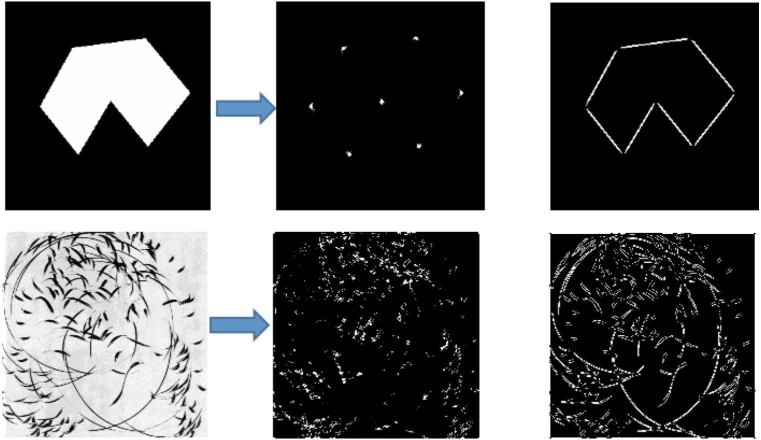
Feature Extraction

Single-scale shearlet analysis of **corners and junctions** [Easley,Labate,Yi,2008]



Feature Extraction

This idea can be used to **classify smooth regions, edges, corner points** [Easley,Labate,Yi,2008].



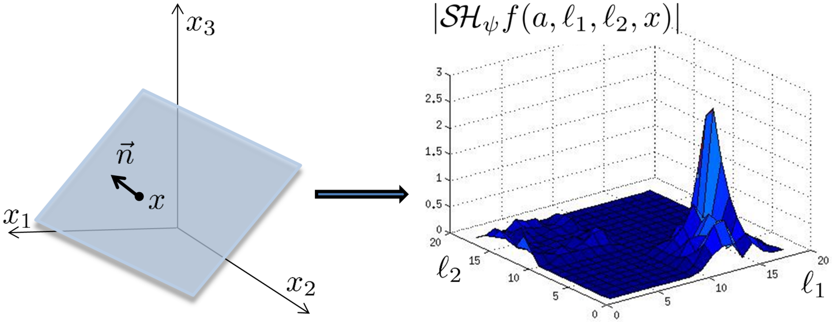
Feature Extraction

A multiscale variant of this idea can be used to define a **corner detector** that is stable to viewpoint and illumination change, and robust to blur and noise [Duval,Odone,De Vito,2015].

Shearlet multiscale corner detection: j = 0 (Blue); j = 1 (Green); j = 2 (Red); j = 3 (Magenta).

Surface Orientation

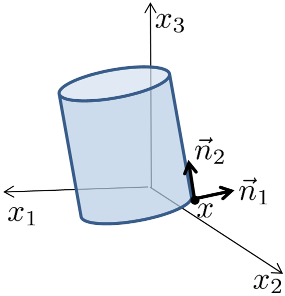
Same idea extends to 3D. The 3D shearlet transform can be used to estimate the **local surface orientation** [L,Negi,2013].

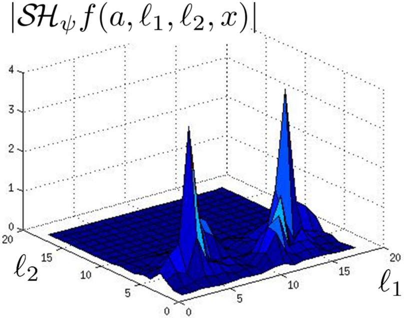


The magnitude of the continuous shearlet transform signals the local orientation of the surface of a solid

Surface Orientation

It can also be useful to detect **wedges and corners**.





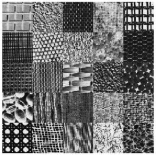
The magnitude of the continuous shearlet transform signals the local orientation of the surface of a solid

Classification

Due to their ability to capture singularities over multiple scales, multiscale representations are useful to generate highly **informative features** for problems of **classification**.

Classification

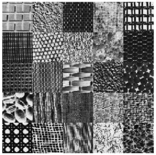
Due to their ability to capture singularities over multiple scales, multiscale representations are useful to generate highly **informative features** for problems of **classification**.



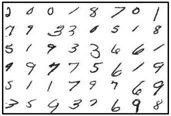
- Wavelet methods for texture classification and segmentation [Unser,1995],[Laine,Fan,1993,1996],. . .

Classification

Due to their ability to capture singularities over multiple scales, multiscale representations are useful to generate highly **informative features** for problems of **classification**.



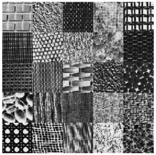
- Wavelet methods for texture classification and segmentation [Unser,1995],[Laine,Fan,1993,1996],...



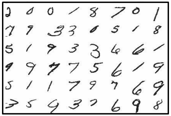
- Rotation, scale invariance using wavelet packets [Pun,Lee,2003], ridgelets [Chen,Bui,2005], contourlets [Chen,Kegl,2010], ...

Classification

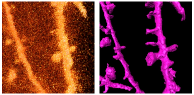
Due to their ability to capture singularities over multiple scales, multiscale representations are useful to generate highly **informative features** for problems of **classification**.



- Wavelet methods for texture classification and segmentation [Unser,1995],[Laine,Fan,1993,1996],...



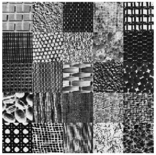
- Rotation, scale invariance using wavelet packets [Pun,Lee,2003], ridgelets [Chen,Bui,2005], contourlets [Chen,Kegl,2010], ...



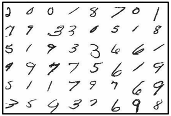
- Wavelet, shearlet filters in combination with SVM [Chen,Xie,2007],[Jimenez,L,Papadakis,2015],...

Classification

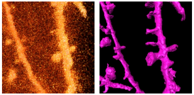
Due to their ability to capture singularities over multiple scales, multiscale representations are useful to generate highly **informative features** for problems of **classification**.



- Wavelet methods for texture classification and segmentation [Unser,1995],[Laine,Fan,1993,1996],. . .



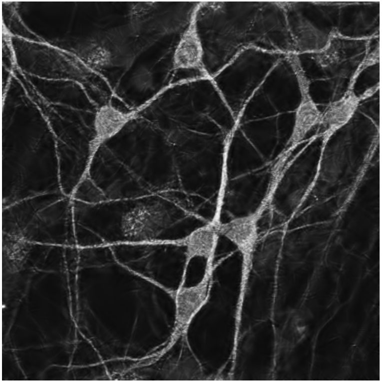
- Rotation, scale invariance using wavelet packets [Pun,Lee,2003], ridgelets [Chen,Bui,2005], contourlets [Chen,Kegl,2010], . . .



- Wavelet, shearlet filters in combination with SVM [Chen,Xie,2007],[Jimenez,L,Papadakis,2015],. . .
- . . .

Problem: Soma Extraction

In neuroscience imaging, it is useful to automatically separate somas from dendrites in fluorescent images of neurons.



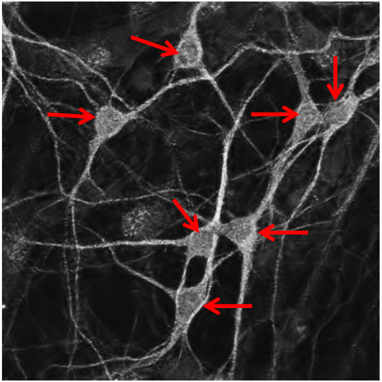
It may be challenging to accurately **detect** and **extract** somas due to large variations in size and shape and irregularities of fluorescence intensity.

Naive methods based on intensity thresholding or standard morphological filters are not reliable and often yield vey inaccuarate results.

Confocal image of neuronal culture (maximum projection view)

Problem: Soma Extraction

In neuroscience imaging, it is useful to automatically separate somas from dendrites in fluorescent images of neurons.



It may be challenging to accurately **detect** and **extract** somas due to large variations in size and shape and irregularities of fluorescence intensity.

Naive methods based on intensity thresholding or standard morphological filters are not reliable and often yield vey inaccuarate results.

Confocal image of neuronal culture (maximum projection view)

Soma extraction

Shearlets and similar directional multiscale representations can be used to detects regions of **local isotropy**.

Soma extraction

Shearlets and similar directional multiscale representations can be used to detects regions of **local isotropy**.

Definition. Let $f = \chi_A$, where $A \subset \mathbb{R}^2$. If $x \in A$ we say that f *is locally isotropic at* x *and at scale* $s > 0$ *if* $B(x, s/2) \subseteq A$.

Soma extraction

Shearlets and similar directional multiscale representations can be used to detects regions of **local isotropy**.

Definition. Let $f = \chi_A$, where $A \subset \mathbb{R}^2$. If $x \in A$ we say that f *is locally isotropic at* x *and at scale* $s > 0$ *if* $B(x, s/2) \subseteq A$.

Due to its *directional sensitivity*, the shearlet transform will exhibit a very different behavior at points of local isotropy (inside soma) vs. points of local anisotropy (inside dendrites)

Directionality Ratio

We define the **directionality ratio** of an image $f \in L^2(\mathbb{R}^2)$ at scale $a > 0$ and location $t \in \mathbb{R}^2$ as the quantity

$$\mathcal{D}_a f(t) = \frac{\inf_s\{|S_\psi f(a,s,t)|\}}{\sup_s\{|S_\psi f(a,s,t)|\}}$$

Directionality Ratio

We define the **directionality ratio** of an image $f \in L^2(\mathbb{R}^2)$ at scale $a > 0$ and location $t \in \mathbb{R}^2$ as the quantity

$$\mathcal{D}_a f(t) = \frac{\inf_s\{|S_\psi f(a,s,t)|\}}{\sup_s\{|S_\psi f(a,s,t)|\}}$$

- It measures the **strength of anisotropy** at a location t and a scale a.

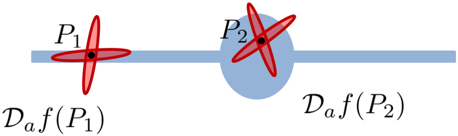
Directionality Ratio

We define the **directionality ratio** of an image $f \in L^2(\mathbb{R}^2)$ at scale $a > 0$ and location $t \in \mathbb{R}^2$ as the quantity

$$\mathcal{D}_a f(t) = \frac{\inf_s \{|S_\psi f(a,s,t)|\}}{\sup_s \{|S_\psi f(a,s,t)|\}}$$

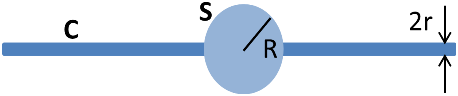
- It measures the **strength of anisotropy** at a location t and a scale a. •

The directionality ratio $\mathcal{D}_a f(t)$ will be very different depending on t being a point of local isotropy of f or not.



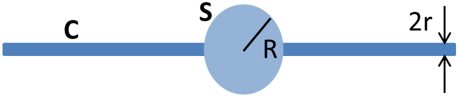
Soma Extraction

Theorem [Labate,Negi,Ozcan,Papadakis,2014]: Let $f = \chi_N$, where N is the union of two subsets: a ball S with radius $R > 0$ and a cylinder C of size $2r \times L$, where $r > 0$, $L \gg R$.



Soma Extraction

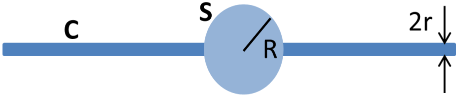
Theorem [Labate,Negi,Ozcan,Papadakis,2014]: Let $f = \chi_N$, where N is the union of two subsets: a ball S with radius $R > 0$ and a cylinder C of size $2r \times L$, where $r > 0$, $L \gg R$.



Then, for $4r \leq a \leq 1/4$, there exists a threshold τ such that, for all $y \in C$, the directionality ratio yields: $\mathcal{D}_a f(y) \leq \tau$.

Soma Extraction

Theorem [Labate,Negi,Ozcan,Papadakis,2014]: Let $f = \chi_N$, where N is the union of two subsets: a ball S with radius $R > 0$ and a cylinder C of size $2r \times L$, where $r > 0$, $L \gg R$.

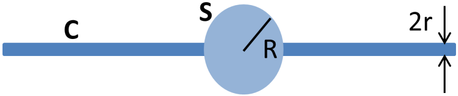


Then, for $4r \leq a \leq 1/4$, there exists a threshold τ such that, for all $y \in C$, the directionality ratio yields: $\mathcal{D}_a f(y) \leq \tau$.

That is, the directionality ratio of f is **small on the cylinder** C.

Soma Extraction

Theorem [Labate,Negi,Ozcan,Papadakis,2014]: Let $f = \chi_N$, where N is the union of two subsets: a ball S with radius $R > 0$ and a cylinder C of size $2r \times L$, where $r > 0$, $L \gg R$.



Then, for $4r \le a \le 1/4$, there exists a threshold τ such that, for all $y \in C$, the directionality ratio yields: $\mathcal{D}_a f(y) \le \tau$.

That is, the directionality ratio of f is **small on the cylinder** C.

On the other hand, the directionality ratio of f is **large (close to 1) inside the ball** S.

Soma Extraction. Segmentation

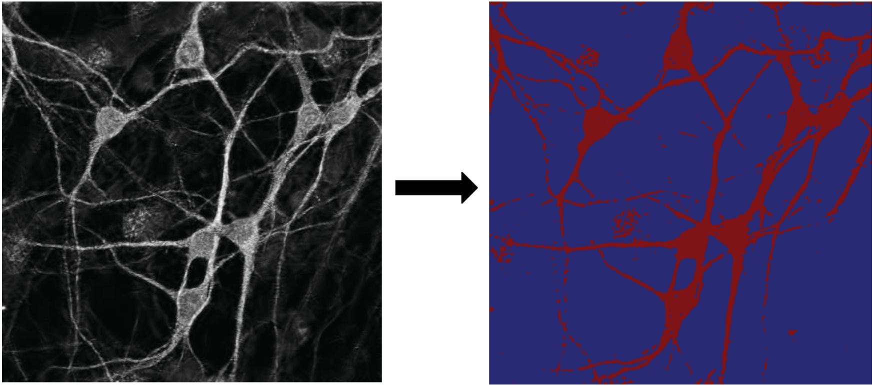
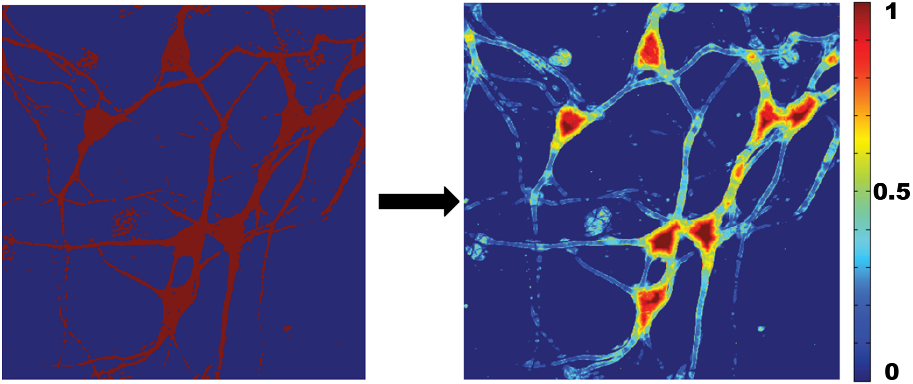


Image segmentation (SVM based)

Soma Extraction. Directionality ratio



Computation of directionality ratio

Soma Extraction

Large values of directionality ratio only identify a region *strictly inside* the soma, not entire soma.

Soma Extraction

Large values of directionality ratio only identify a region *strictly inside* the soma, not entire soma.

To complete the soma, we apply the **level set method**. We compute:

Soma Extraction

Large values of directionality ratio only identify a region *strictly inside* the soma, not entire soma.

To complete the soma, we apply the **level set method**. We compute:

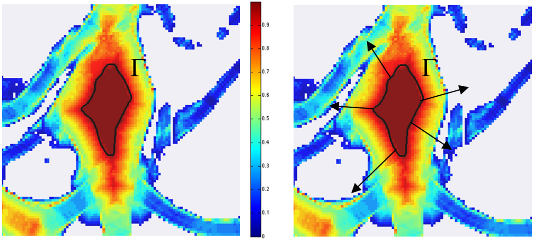
- initial curve = boundary of set where directionality ratio is large;

Soma Extraction

Large values of directionality ratio only identify a region *strictly inside* the soma, not entire soma.

To complete the soma, we apply the **level set method**. We compute:

- initial curve = boundary of set where directionality ratio is large;
- a force field from the gradient of directionality ratio.

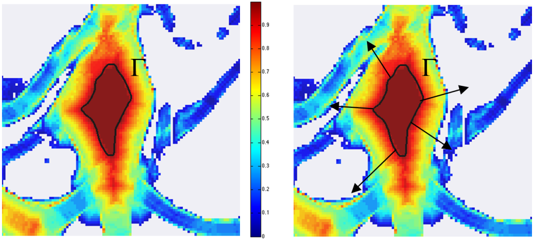


Soma Extraction

Large values of directionality ratio only identify a region *strictly inside* the soma, not entire soma.

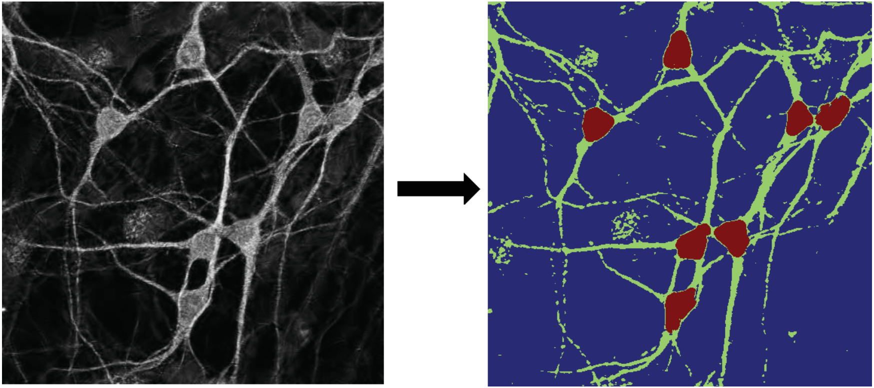
To complete the soma, we apply the **level set method**. We compute:

- initial curve = boundary of set where directionality ratio is large;
- a force field from the gradient of directionality ratio.



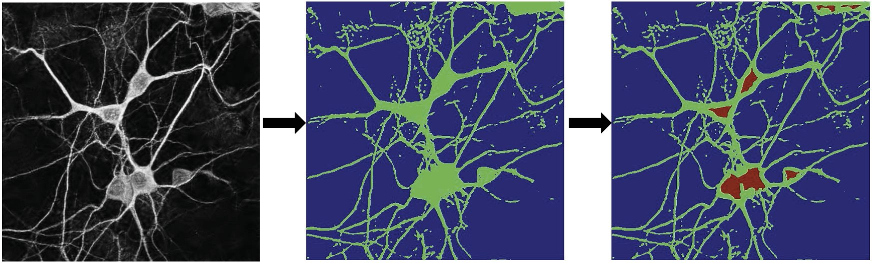
We also use this method to separate clustered somas.

Soma Extraction



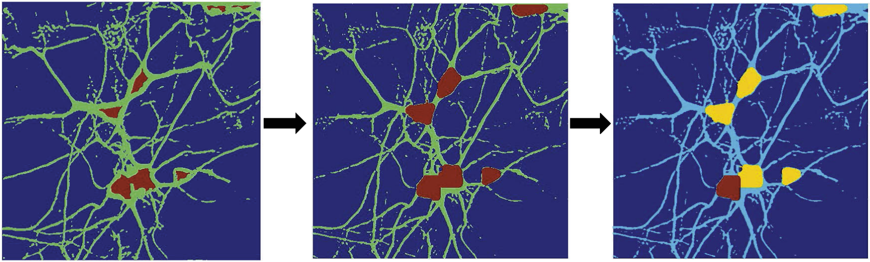
Directionality ratio + level set: soma detection

Soma Extraction. Another example



Identification of somas

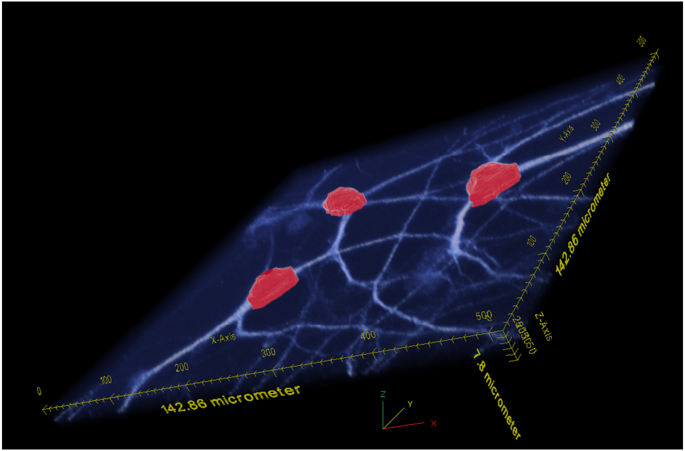
Soma Extraction. Another example



Identification of somas and separation of clustered ones

Soma Extraction (3D)

Method extends to 3D where soma detection can be combined with the extraction of soma morphology [Bozcan,L,Laezza,Negi,Papadakis,2014]



Scattering Transform

The **scattering** transform [Mallat,2012, Mallat,Bruna,2013] computes data representations targeted to problems of pattern recognition.
Key features:

Scattering Transform

The **scattering** transform [Mallat,2012, Mallat,Bruna,2013] computes data representations targeted to problems of pattern recognition.
Key features:

- It extracts locally invariant, stable, highly informative features.

Scattering Transform

The **scattering** transform [Mallat,2012, Mallat,Bruna,2013] computes data representations targeted to problems of pattern recognition.
Key features:

- It extracts locally invariant, stable, highly informative features.
- It is implemented through a cascade of **wavelet filters** and **modulus operators** over multiple layers (deep convolution network).

Scattering Transform

The **scattering** transform [Mallat,2012, Mallat,Bruna,2013] computes data representations targeted to problems of pattern recognition.
Key features:

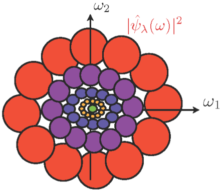
- It extracts locally invariant, stable, highly informative features.
- It is implemented through a cascade of **wavelet filters** and **modulus operators** over multiple layers (deep convolution network).

Dilated wavelets are also rotated with elements $r \in G$:

$$\psi_\lambda(x) = a^{-1}\psi(a^{-1}rx)$$

with $\lambda = (a, r)$, $a > 0, r \in G$.

$$\mathcal{W}_\psi : f \mapsto \mathcal{W}_\psi f(a, t) = f * \psi_\lambda(t)$$



Scattering Transform

By taking the magnitude and then averaging with a low-pass function ϕ, one defines **locally translation invariant** coefficients

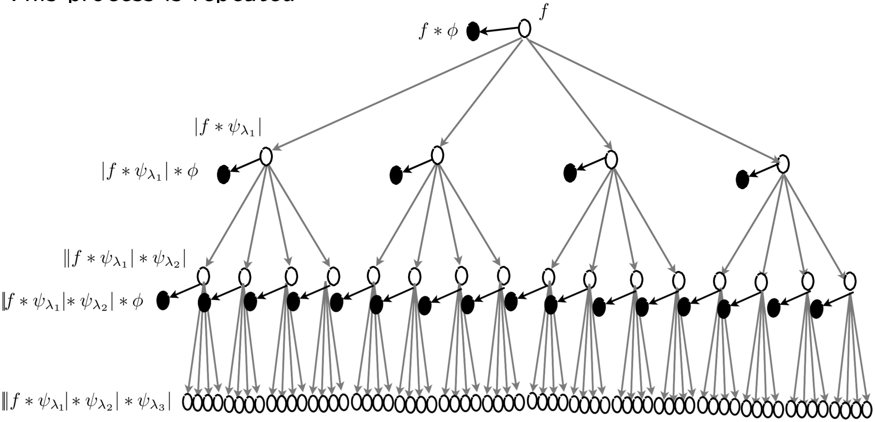
$$S_1 f(x, \lambda) = |f * \psi_\lambda(x)| * \phi(x).$$

Scattering Transform

By taking the magnitude and then averaging with a low-pass function ϕ, one defines **locally translation invariant** coefficients

$$S_1 f(x,\lambda) = |f * \psi_\lambda(x)| * \phi(x).$$

This process is repeated



Scattering Transform

The scattering transform builds coefficients

Scattering Transform

The scattering transform builds coefficients

- invariant to local translations and stable to small deformations;

Scattering Transform

The scattering transform builds coefficients

- invariant to local translations and stable to small deformations;
- other invariances can be built into this approach.

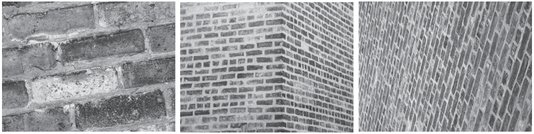
Scattering Transform

The scattering transform builds coefficients

- invariant to local translations and stable to small deformations;
- other invariances can be built into this approach.

Multiple applications including:

- texture classification [Sifre,Mallat,2014]



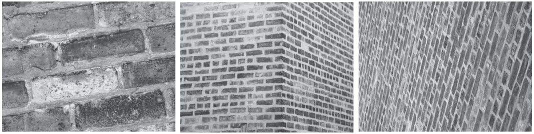
Scattering Transform

The scattering transform builds coefficients

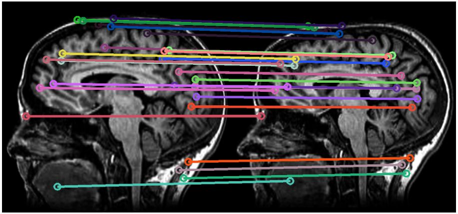
- invariant to local translations and stable to small deformations;
- other invariances can be built into this approach.

Multiple applications including:

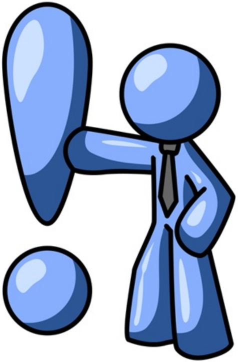
- texture classification [Sifre,Mallat,2014]



- image registration [Easley,Mc-Innis,L,2015]



Conclusion



Conclusion

- Methods from **applied harmonic analysis**s offer powerful tools to capture the structure of imaging data.

Conclusion

- Methods from **applied harmonic analysis** offer powerful tools to capture the structure of imaging data.
  - Wavelets, curvelets, shearlets, scattering wavelets,. . .

Conclusion

- Methods from **applied harmonic analysis** offer powerful tools to capture the structure of imaging data.
  - Wavelets, curvelets, shearlets, scattering wavelets,...
- Shearlets and related multiscale representations enable a precise **geometrical description** of the **singularities** of multivariate functions and distributions.

Conclusion

- Methods from **applied harmonic analysis** offer powerful tools to capture the structure of imaging data.
  - Wavelets, curvelets, shearlets, scattering wavelets,...
- Shearlets and related multiscale representations enable a precise **geometrical description** of the **singularities** of multivariate functions and distributions.
- These properties are useful to extract essential **image features**

Conclusion

- Methods from **applied harmonic analysis** offer powerful tools to capture the structure of imaging data.
  - Wavelets, curvelets, shearlets, scattering wavelets,...
- Shearlets and related multiscale representations enable a precise **geometrical description** of the **singularities** of multivariate functions and distributions.
- These properties are useful to extract essential **image features**
  - edge analysis, edge/boundary and corner detection, local isotropy, ...

Conclusion

- Methods from **applied harmonic analysis** offer powerful tools to capture the structure of imaging data.
  - Wavelets, curvelets, shearlets, scattering wavelets,...
- Shearlets and related multiscale representations enable a precise **geometrical description** of the **singularities** of multivariate functions and distributions.
- These properties are useful to extract essential **image features**
  - edge analysis, edge/boundary and corner detection, local isotropy, ...
- Building on this low-level image processing capabilities, one can construct improved methods for **pattern recognition and classification**

Conclusion

- Methods from **applied harmonic analysis** offer powerful tools to capture the structure of imaging data.
  - Wavelets, curvelets, shearlets, scattering wavelets,...
- Shearlets and related multiscale representations enable a precise **geometrical description** of the **singularities** of multivariate functions and distributions.
- These properties are useful to extract essential **image features**
  - edge analysis, edge/boundary and corner detection, local isotropy, ...
- Building on this low-level image processing capabilities, one can construct improved methods for **pattern recognition and classification**
  - soma detection, texture classification, ...

References + codes at:

`www.math.uh.edu\~dlabate`

Research supported in part by NSF (DMS 1320910) and by Norman Hackerman Advanced Research Program