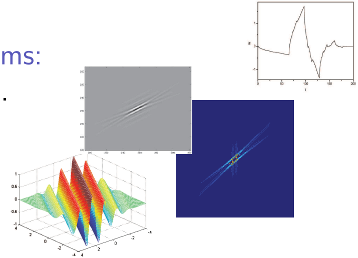
Applied Harmonic Analysis Methods in Imaging Science Part II

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,...

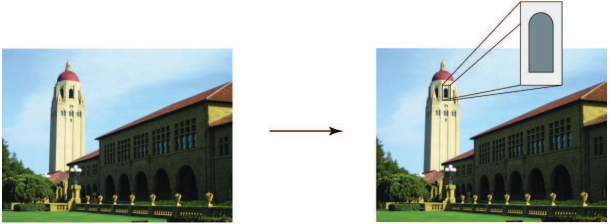
Main Desiderata:

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

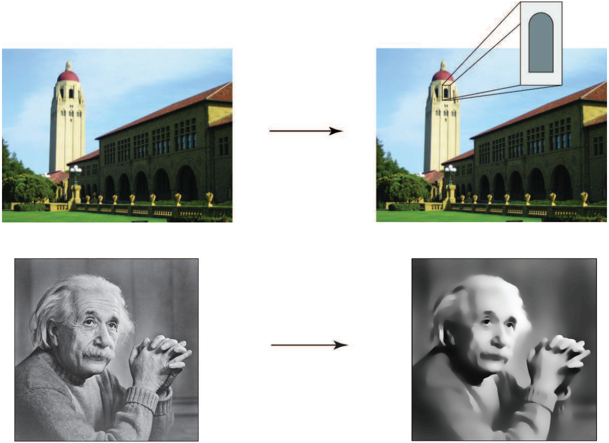
Outline

What is an Image?

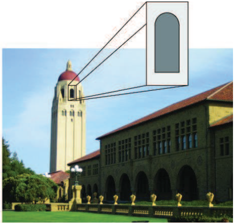
What is an Image?

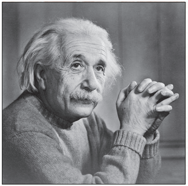


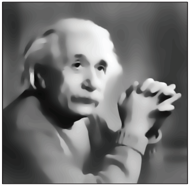
What is an Image?



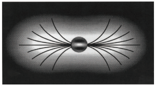
What is an Image?







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



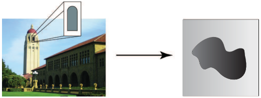
Field et al., 1993

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 $\|f_i\|_{C^2} \leq 1$, $i = 0, 1$.



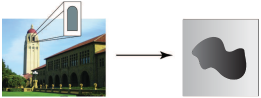
Fitting Model

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

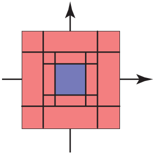
$$\{\phi(x - m) : m \in \mathbb{Z}\} \cup \{2^{j/2}\,\psi(2^j x - m) : j \geq 0, m \in \mathbb{Z}\}.$$

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

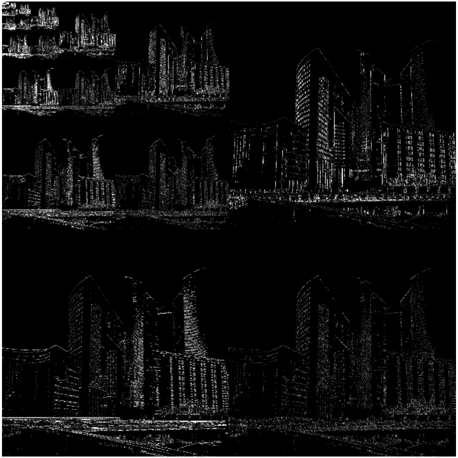
$$\phi^{(1)}(x) = \phi(x_1)\phi(x_2) \quad \text{and} \quad \begin{aligned} \psi^{(1)}(x) &= \phi(x_1)\psi(x_2), \\ \psi^{(2)}(x) &= \psi(x_1)\phi(x_2), \\ \psi^{(3)}(x) &= \psi(x_1)\psi(x_2). \end{aligned}$$



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:

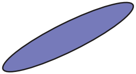
- Ridgelets (Candès and Donoho; 1999)
- Curvelets (Candès and Donoho; 2002)
- Contourlets (Do and Vetterli; 2002)
- Bandlets (LePennec and Mallat; 2003)
- Shearlets (K and Labate; 2006)

What is a Shearlet?

Scaling and Orientation

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

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



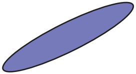
Historical remark:

- 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!*

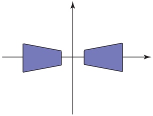
Remarks:

- 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

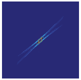
Let $\psi \in L^2(\mathbb{R}^2)$ be defined by

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



where

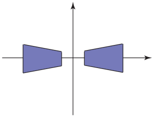
- ψ_1 wavelet, $\mathrm{supp}(\hat{\psi}_1) \subseteq [-2, -\frac{1}{2}] \cup [\frac{1}{2}, 2]$ and $\hat{\psi}_1 \in C^\infty(\mathbb{R})$.
- $\mathrm{supp}(\hat{\psi}_2) \subseteq [-1, 1]$ and $\hat{\psi}_2 \in C^\infty(\mathbb{R})$.



Example of Classical (Band-Limited) Shearlet

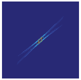
Let $\psi \in L^2(\mathbb{R}^2)$ be defined by

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

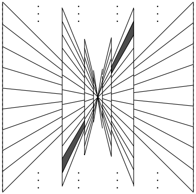


where

- 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 $\hat{\psi}_2 \in C^\infty(\mathbb{R})$.



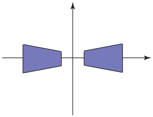
Induced tiling of Fourier domain:



Example of Classical (Band-Limited) Shearlet

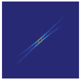
Let $\psi \in L^2(\mathbb{R}^2)$ be defined by

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

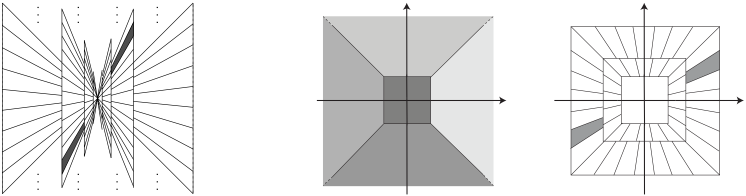


where

- 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:



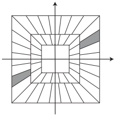
Definition (K, Labate; 2006):

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

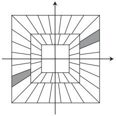
Definition (K, Labate; 2006):

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 $L^2(\mathbb{R}^2)$:

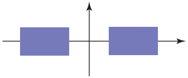
$$A\|f\|_2^2 \leq \sum_{\sigma \in \mathcal{SH}(\phi,\psi,\tilde{\psi})} |\langle f, \sigma \rangle|^2 \leq B\|f\|_2^2 \quad \text{for all } f \in L^2(\mathbb{R}^2)$$

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 $c \leq c_0$.

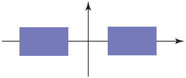


Remark: Exemplary class with $B/A \approx 4$.

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 $c \leq c_0$.

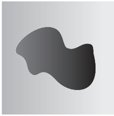


Remark: Exemplary class with $B/A \approx 4$.

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 $N, n \to \infty$,

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

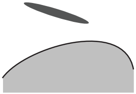


Heuristic Argument

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:

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

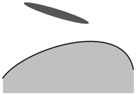
$\rightsquigarrow |\langle f, \sigma_{\eta_n} \rangle| \lesssim n^{-\frac{3}{2}}$

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:

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

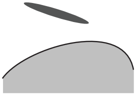
$$\rightsquigarrow |\langle f, \sigma_{\eta_n} \rangle| \lesssim n^{-\frac{3}{2}}$$

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:

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

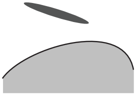
$$\rightsquigarrow |\langle f, \sigma_{\eta_n} \rangle| \lesssim n^{-\frac{3}{2}}$$

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

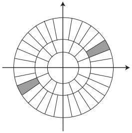
Definition (Candès, Donoho; 2002):

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, ...*

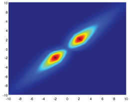
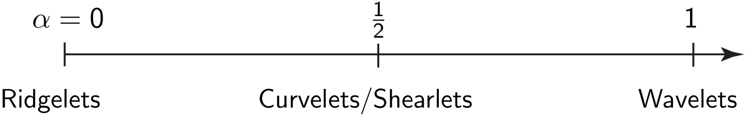


Illustration (“$\alpha =$ degree of anisotropy”):



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)*

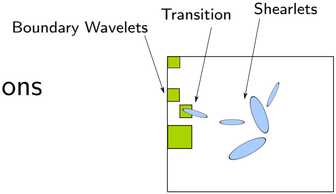
Boundary Shearlets

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

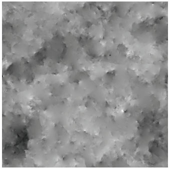
Noisy Version (20.17dB)

Curvelets (28.70dB, 7.22sec)

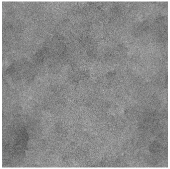
Shearlets (29.20dB, 5.56sec)

(Source: W.-Q Lim; 2011)

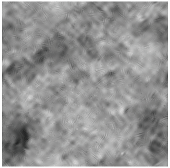
Image Denoising, II



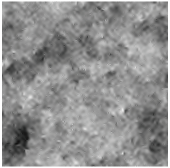
Original



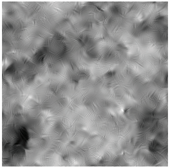
Noisy Version (6.5dB)



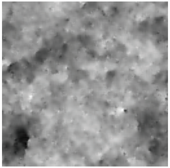
Curvelets (22.10dB)



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!*

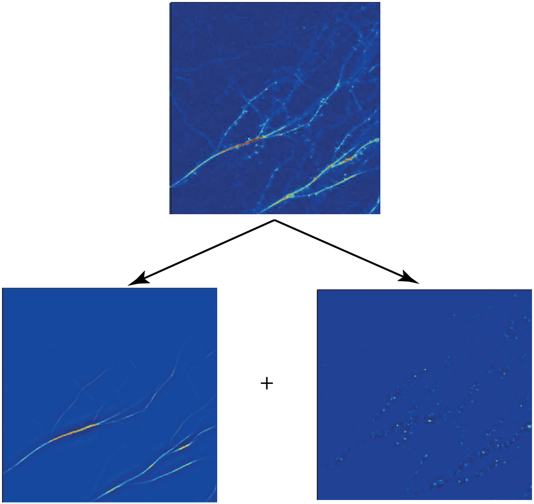
Some Examples for $\mathcal{P}$:

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

Some Earlier Footprints in Inverse Problems:

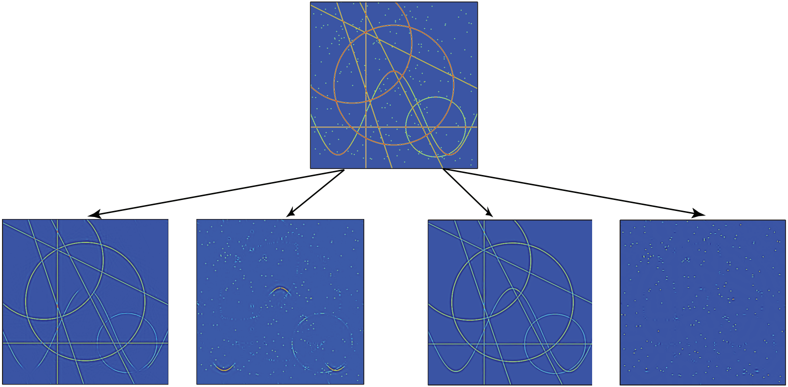
- Donoho (1995): Wavelet-Vaguelette decomposition.
- Chambolle, DeVore, Lee, Lucier (1998): Penalty on the Besov norm.
- Daubechies, Defries, De Mol (2004): General sparsity constraints.

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

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

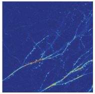
Feature Extraction and ℓ_1 Minimization

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

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

Model: For τ a closed C^2-curve,

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



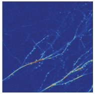
Feature Extraction and ℓ_1 Minimization

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

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

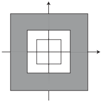
Model: For τ a closed C^2-curve,

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



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:

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

Analysis of Feature Extraction

ℓ_1-Decomposition:

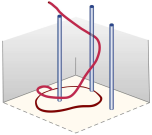
$$(\mathcal{P}_j^*, \mathcal{C}_j^*) = \text{argmin}_{\tilde{\mathcal{P}}_j, \tilde{\mathcal{C}}_j} \|(\langle \tilde{\mathcal{P}}_j, \psi_\lambda \rangle)_\lambda\|_1 + \|(\langle \tilde{\mathcal{C}}_j, \sigma_\eta \rangle)_\eta\|_1 \text{ s.t. } f_j = \tilde{\mathcal{P}}_j + \tilde{\mathcal{C}}_j$$

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:

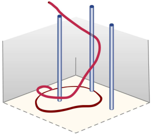
$$(\mathcal{P}_j^*, \mathcal{C}_j^*) = \text{argmin}_{\tilde{\mathcal{P}}_j, \tilde{\mathcal{C}}_j} \|(\langle \tilde{\mathcal{P}}_j, \psi_\lambda \rangle)_\lambda\|_1 + \|(\langle \tilde{\mathcal{C}}_j, \sigma_\eta \rangle)_\eta\|_1 \text{ s.t. } f_j = \tilde{\mathcal{P}}_j + \tilde{\mathcal{C}}_j$$

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.

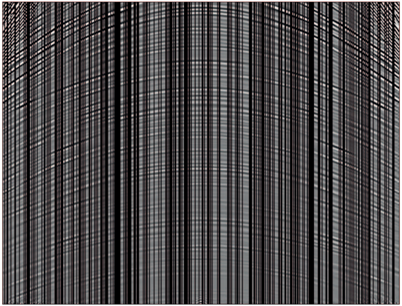


Theorem (K; 2014):

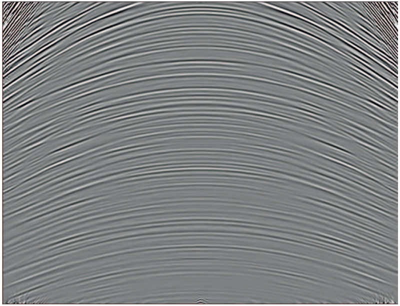
Using One-Step-Thresholding, we also have

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

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

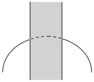
Key Idea:

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

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

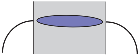
Observed Object:

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



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

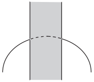
Key Idea:

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

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

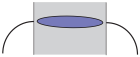
Observed Object:

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



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 $j \to \infty$,

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

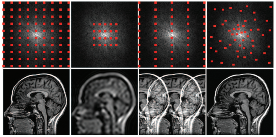
Application to MRI

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



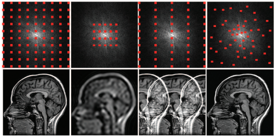
Initial idea with wavelets: Lustig, Donoho, Pauly; 2007

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:

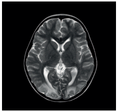
- Fast acquisition $\longleftrightarrow$ Small set Λ
- Optimality result



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

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

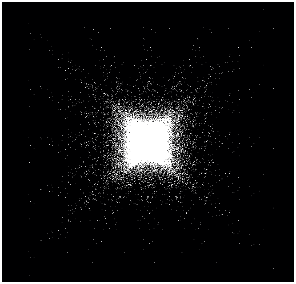
- Model for f: Cartoon-like functions.
- (Dualizable) shearlets as sparsifying system $(\sigma_\eta)_\eta$.
- Directional (random) sampling scheme Λ.
- Algorithmic approach:



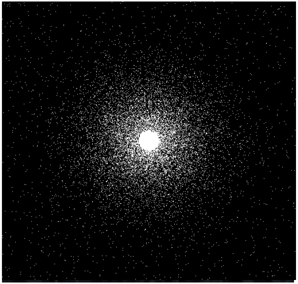
$$\min_f \|(\langle f, \sigma_\eta \rangle)_\eta\|_1 \quad \text{subject to} \quad \|(\hat{f}(\lambda) - g_\lambda)_{\lambda \in \Lambda}\|_2 \leq \varepsilon$$

Asymptotic Optimality of Shearlet Scheme

Sampling Schemes:



Directional Sampling Scheme



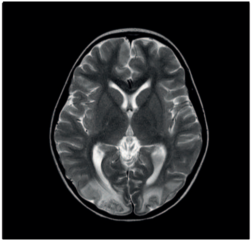
Variable Density Sampling Scheme

Theorem (K, Lim; 2015):

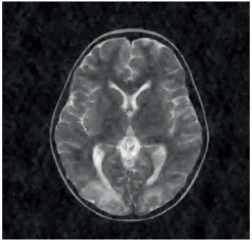
"Using the directional sampling schemes $(\Delta_M)_M$, $\#\Delta_M = M$, and $M \to \infty$ in combination with dualizable shearlets, this reconstruction scheme $\mathcal{R}$ is asymptotically optimal in the sense that, for all $f \in \mathcal{E}^2(\mathbb{R}^2)$,

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

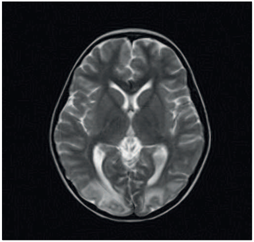
Numerical Results for 512x512 MRI Image



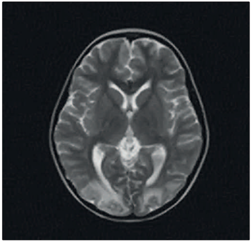
Original



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...

2D $\longrightarrow$ 3D

Question:

Why is the 3D situation such crucial?

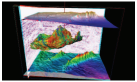
2D $\longrightarrow$ 3D

Question:

Why is the 3D situation such crucial?

Obvious answer:

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



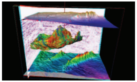
2D $\longrightarrow$ 3D

Question:

Why is the 3D situation such crucial?

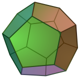
Obvious answer:

- 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

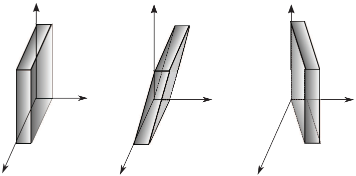
$$\|f - f_N\|_2^2 \asymp N^{-1}, \quad N \to \infty.$$

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

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

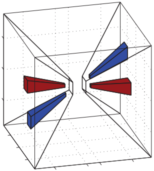


Pyramid-adapted Shearlet Systems

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

$$\begin{aligned} &\{\phi(\cdot - cm) : m \in \mathbb{Z}^3\} \\ \cup &\{2^j \psi(S_k A_j \cdot - cm) \,:\, (j,k,m) \in \Lambda_{pyramid}\} \\ \cup &\{2^j \tilde{\psi}(\tilde{S}_k \tilde{A}_j \cdot - cm) \,:\, (j,k,m) \in \Lambda_{pyramid}\} \\ \cup &\{2^j \breve{\psi}(\breve{S}_k \breve{A}_j \cdot - cm) \,:\, (j,k,m) \in \Lambda_{pyramid}\}, \end{aligned}$$



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.

Applied Functional Analysis Group
Technische Universität Berlin

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.