Outline

Applied Harmonic Analysis and Imaging Sciences
  Anisotropic Phenomena
  Goal for Today

Compactly Supported Shearlet Systems
  Review: Discrete Shearlets in 2D
  Shearlet Systems in 3D
  Optimal Sparse Approximation
  Frame Properties

Applications
  Denoising
  Image Separation

Conclusions
Anisotropic Phenomena in Multivariate Data

Many important multivariate problem classes are governed by **anisotropic features**, which require efficient encoding strategies.

**The anisotropic structure can be given...**

- **...explicitly.**
  - Image Processing: Edges.
  - Seismology: Layers of earth.

- **...implicitly.**
  - Transport Equations: Shock fronts.
Modern Imaging

Some important tasks:

- Denoising.
- Inpainting.
- Feature Detection/Extraction.
- ...

Imaging Sciences using Applied Harmonic Analysis:
Exploit a carefully designed representation system \((\psi_\lambda)_{\lambda}\):

\[
\text{Image} = \sum_{\lambda} \langle \text{Image}, \psi_\lambda \rangle \psi_\lambda.
\]

- Sparse coefficients!
- Approximation properties!
What is an Image?
What is an Image?
What is an Image?

![Image 1](image1.jpg) → ![Image 2](image2.jpg)

![Image 3](image3.jpg) → ![Image 4](image4.jpg)
Reasonable Model for Images

Definition (Donoho; 2001):
The set of cartoon-like 2D images $\mathcal{E}_2^2(\nu)$ is defined by

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

where $f_0, f_1 \in C^2(\mathbb{R}^2)$ with $\text{supp} f_i \subset [0, 1]^2$ and $B \subset [0, 1]^2$ with $\partial B$ a closed $C^2$-curve whose curvature is bounded by $\nu$. 
Reasonable Model for Images

Definition (Donoho; 2001):
The set of cartoon-like 2D images $E^2_2(\nu)$ is defined by

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

where $f_0, f_1 \in C^2(\mathbb{R}^2)$ with $\text{supp} f_i \subset [0, 1]^2$ and $B \subset [0, 1]^2$ with $\partial B$ a closed $C^2$-curve whose curvature is bounded by $\nu$.

Theorem (Donoho; 2001):
Let $(\psi_\lambda)_{\lambda} \subseteq L^2(\mathbb{R}^2)$. Allowing only polynomial depth search, the optimal asymptotic approximation error of $f \in E^2_2(\nu)$ is

$$\| f - f_N \|_2^2 \leq C \cdot N^{-2}, \quad N \to \infty,$$

where $f_N = \sum_{\lambda \in I_N} c_\lambda \psi_\lambda$. 
What can Wavelets do?

Problem:

- For $\mathcal{E}_2^2(\nu)$, Wavelets only achieve $\| f - f_N \|_2^2 \leq C \cdot N^{-1}$, $N \to \infty$.

- Isotropic structure of wavelets:

  $$\{ \psi_{j,k} = 2^{-j/2} \psi(2^{-j} \cdot -k) : j \in \mathbb{Z}, k \in \mathbb{Z}^2 \}$$

- Wavelets cannot sparsely represent curves.

Intuitive explanation:
Main Goal in Geometric Multiscale Analysis

Design a representation system which...

- ...is generated by one ‘mother function’,
- ...provides **optimally sparse approximation** of cartoons,
- ...allows for **compactly supported** analyzing elements,
- ...is associated with **fast decomposition algorithms**, 
- ...treats the **continuum and digital ‘world’** uniformly.

*Vision: Introduce a system for 2D data as powerful as wavelets for 1D data!*
Previous Approaches

Non-exhaustive list:

- Directional wavelets (Antoine, Murenzi, Vandergheynst; 1999)
- Ridgelets (Candès and Donoho; 1999)
- Curvelets (Candès and Donoho; 2002)
- Contourlets (Do and Vetterli; 2002)
- Bandlets (LePennec and Mallat; 2003)
- Wavelets with composite dilations (Guo, Labate, Lim, Weiss, and Wilson; 2004)
- Shearlets (K, Labate, Lim, and Weiss; 2005), (Guo, K, and Labate; 2006)
- ...
Most Approaches are for 2D Data

Question:

Why is the 3D situation such crucial?
Most Approaches are for 2D Data

**Question:**
Why is the 3D situation such crucial?

**Answer:**
- Our world is 3-dimensional.
  - 3D data is essential for Biology, Seismology, ...
- Transition $2D \rightarrow 3D$ is unique.
  - Anisotropic features occur in 3D for the first time in different dimensions.
Goal for Today

Goal: Introduce shearlets for 3D data which...

- ...are generated by one ‘mother function’,
- ...provide optimally sparse approximation of cartoons,
- ...allow for compactly supported analyzing elements,
- ...are associated with fast decomposition algorithms,
- ...treat the continuum and digital ‘world’ uniformly.

Remark: Even new in 2D:

- Construction of compactly supported shearlets.
- Optimal sparse approximation of such shearlets.
In harmonic analysis there have been a number of important applications of decompositions based on parabolic scaling

\[ f_a(x_1, x_2) = f(a^{\frac{1}{2}}x_1, ax_2), \]

which leaves invariant the parabola \( x_2 = x_1^2 \).

History:

- 1970’s: Fefferman and later Seeger, Sogge, and Stein studied boundedness of certain operators.
Which is the Correct Measure for Directions…?

Angle versus slope:

- It seems that the appropriate measure for directions is the angle (rotation).
Which is the Correct Measure for Directions…?

Angle versus slope:

- It seems that the appropriate measure for directions is the angle (rotation).

- However, in today’s digital world slope (shearing) is much more appropriate.
  - Example: Parametrization by slope is more suitable for measurements in seismology.
  - Supports the treating of the digital setting:

\[
\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \mathbb{Z}^2 = \mathbb{Z}^2 \quad \text{for all } k \in \mathbb{Z}.
\]
Construction of 2D Discrete Shearlets

Idea:

- Consider (affine) systems:
  $$\left\{ |\det M|^{1/2} \psi(Mx - m) : M \in G \subseteq GL_n, \ m \in \mathbb{Z}^2 \right\}.$$  

- Recall: For wavelets, $G = \{\text{diag}(2^j, 2^j) : j \in \mathbb{Z}\}$.  

- Construct $G$ to be a special 2-parameter dilation group by:

  $$A_j = \begin{pmatrix} 2^j & 0 \\ 0 & 2^{j/2} \end{pmatrix} \quad \text{and} \quad S_k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}, \quad j, k \in \mathbb{Z}. $$

  $A_j$: Parabolic scaling matrix, $S_k$: Shear matrix.

- Dilation matrices: $G = \{S_kA_j : j, k \in \mathbb{Z}\}$.  

Gitta Kutyniok, Jakob Lemvig, & Wang-Q Lim
Compactely Supported Shearlets
Discrete Shearlet Systems

Definition:
Let \( \psi \in L^2(\mathbb{R}^2) \). Then the associated discrete shearlet system is defined by

\[
\{ 2^{3j/4} \psi(S_kA_j \cdot -m) : j, k \in \mathbb{Z}, m \in \mathbb{Z}^2 \}.
\]

Remarks:
- Advantage: Associated shearlet group \( \mathcal{S} = (\mathbb{R}^+ \times \mathbb{R}) \times \mathbb{R}^2 \).
- Disadvantage: Biased treatment of directions.
Discrete Shearlet Systems

Induced Tiling of the Frequency Domain:
Cone-adapted Discrete Shearlet Systems

Induced Tiling of the Frequency Domain:
Cone-adapted Discrete Shearlet Systems

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

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

$$\cup \left\{2^{\frac{3j}{4}} \psi(S_k A_j \cdot - cm) : (j, k, m) \in \Lambda_{cone}\right\}$$

$$\cup \left\{2^{\frac{3j}{4}} \tilde{\psi}(S^T_k \tilde{A}_j \cdot - cm) : (j, k, m) \in \Lambda_{cone}\right\},$$

where

$$\Lambda_{cone} = \{(j, k, m) : j \geq 0, |k| \leq \lfloor 2^{j/2} \rfloor, m \in \mathbb{Z}^2\}, \quad c > 0.$$
Example of Shearlet

Classical (Band-Limited) Shearlet:

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

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

where

- \( \psi_1 \) wavelet, \( \text{supp}(\hat{\psi}_1) \subseteq [-2, -\frac{1}{2}] \cup [\frac{1}{2}, 2] \), and \( \hat{\psi}_1 \in C^\infty(\mathbb{R}) \).
- \( \psi_2 \) ‘bump’ function, \( \text{supp}(\hat{\psi}_2) \subseteq [-1, 1] \), and \( \hat{\psi}_2 \in C^\infty(\mathbb{R}) \).

Illustration:
Results for Band-Limited 2D Shearlets

Hypothesis:

- Let $\phi$ be a scaling function.
- Let $\psi$ be a classical shearlet, and $\tilde{\psi}$ defined analogously.

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

Theorem (Guo, Labate; 2007):
$\mathcal{SH}(\phi, \psi, \tilde{\psi}; 1)$ provides (almost) optimally sparse approximations of functions $f \in \mathcal{E}_2^2(\nu)$ with asymptotic error

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

Definition (Donoho; 2001):
The set of 2D images $\mathcal{E}_2^2(\nu)$ is defined by

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

where $f_1 \in C^2(\mathbb{R}^2)$ with $\text{supp} f_i \subset [0,1]^2$ and $B \subset [0,1]^2$ with $\partial B$ a closed $C^2$-curve whose curvature is bounded by $\nu$.

Theorem (Donoho; 2001):
Let $(\psi_\lambda)_\lambda \subseteq L^2(\mathbb{R}^2)$. Allowing only polynomial depth search, the optimal asymptotic approximation error of $f \in \mathcal{E}_2^2(\nu)$ is

$$\| f - f_N \|_2^2 \leq C \cdot N^{-2}, \quad N \to \infty.$$
Reasonable Model for 3D Images

Definition:
The set of 3D images $\mathcal{E}_3^2(\nu)$ is defined by

$$\mathcal{E}_3^2 = \{ f \in L^2(\mathbb{R}^3) : f = f_0 + f_1 \chi_B \},$$

where $f_i \in C^2(\mathbb{R}^3)$ with $\text{supp } f_i \subset [0, 1]^3$ and $B \subset [0, 1]^3$ with $\partial B$ a closed $C^2$-surface whose principal curvatures are bounded by $\nu$. 
Reasonable Model for 3D Images

Definition:
The set of 3D images $\mathcal{E}_3^2(\nu)$ is defined by

$$\mathcal{E}_3^2 = \{ f \in L^2(\mathbb{R}^3) : f = f_0 + f_1 \chi_B \},$$

where $f_i \in C^2(\mathbb{R}^3)$ with $\text{supp } f_i \subset [0, 1]^3$ and $B \subset [0, 1]^3$ with $\partial B$ a closed $C^2$-surface whose principal curvatures are bounded by $\nu$.

Theorem (K, Lemvig, Lim; 2010):
Let $(\psi_\lambda)_\lambda \subset L^2(\mathbb{R}^3)$. Allowing only polynomial depth search, the optimal asymptotic approximation error of $f \in \mathcal{E}_3^2(\nu)$ is

$$\| f - f_N \|_2^2 = O(N^{-1}), \quad N \to \infty.$$
**Definition:**

Let $1 < \alpha, \beta \leq 2$. The set of 3D images $\mathcal{E}^{\alpha,\beta}_3(\nu)$ is defined by

$$
\mathcal{E}^{\alpha,\beta}_3 = \{ f \in L^2(\mathbb{R}^3) : f = f_0 + f_1 \chi_B \},
$$

where $f_i \in C^\beta \cap H^\beta(\mathbb{R}^3)$, $\text{supp} \ f_i \subset [0, 1]^3$ and $B \subset [0, 1]^3$ with $\partial B$ a closed $C^\alpha$-surface whose principal curvatures are bounded by $\nu$. 
More Extensive Model for 3D Images

Definition:
Let $1 < \alpha, \beta \leq 2$. The set of 3D images $\mathcal{E}^{\alpha,\beta}_3(\nu)$ is defined by

$$\mathcal{E}^{\alpha,\beta}_3 = \left\{ f \in L^2(\mathbb{R}^3) : f = f_0 + f_1 \chi_B \right\},$$

where $f_i \in C^\beta \cap H^\beta(\mathbb{R}^3)$, supp $f_i \subset [0, 1]^3$ and $B \subset [0, 1]^3$ with $\partial B$ a closed $C^\alpha$-surface whose principal curvatures are bounded by $\nu$.

Theorem (K, Lemvig, Lim; 2010):
Let $(\psi_\lambda)_\lambda \subset L^2(\mathbb{R}^3)$. Allowing only polynomial depth search, the optimal asymptotic approximation error of $f \in \mathcal{E}^{\alpha,\beta}_3(\nu)$ is

$$\|f - f_N\|_2^2 = O(N^{-\min\{\alpha/2, 2\beta/3\}}), \quad N \to \infty.$$
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** $S_k$, $k = (k_1, k_2)$ (direction parameter $\leftrightarrow$ rotations):

  $$S_k = \begin{pmatrix} 1 & k_1 & k_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
Anisotropic scaling $A_j$ ($\alpha \in (1, 2]$):

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

Shearing $S_k, k = (k_1, k_2)$ (direction parameter $\leftrightarrow$ rotations):

$$S_k = \begin{pmatrix}
1 & k_1 & k_2 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix}.$$
3D Action of Anisotropic Scaling and Shearing

Spatial Domain:

Frequency Domain:
Tiling of Frequency Domain Induced by Shearlets

2D Shearlets:

3D Shearlets:
Pyramid-adapted Shearlet Systems

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

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

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

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

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

where

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

Definition:
The pyramid-adapted shearlet system $\mathcal{S}\mathcal{H}(\phi, \psi, \tilde{\psi}, \check{\psi}; c; \alpha)$ generated by $\phi \in L^2(\mathbb{R}^3)$ and $\psi, \tilde{\psi}, \check{\psi} \in L^2(\mathbb{R}^3)$ is

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

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

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

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

where

$$\Lambda_{pyramid} = \{(j, k, m) : j \geq 0, |k_1|, |k_2| \leq \lceil 2^{(\alpha-1)j/2} \rceil, m \in \mathbb{Z}^2\}, \quad c > 0.$$
Optimal Sparse Approximation with Shearlets

Theorem (K, Lemvig, Lim; 2010):
Let $\alpha, \beta \in (1, 2]$. Let $\phi, \psi, \tilde{\psi}, \check{\psi} \in L^2(\mathbb{R}^3)$ be compactly supported. For $\psi$ suppose that:

(i) $|\hat{\psi}(\xi)| \leq C \cdot \min(1, |\xi_1|^\delta) \cdot \prod_{i=1}^3 \min(1, |\xi_i|^{\gamma})$

(ii) $\left| \frac{\partial}{\partial \xi_2} \hat{\psi}(\xi) \right| \leq |h(\xi_1)| \cdot \left(1 + \frac{\xi_2}{|\xi_1|}\right)^{-\gamma}$,

(iii) $\left| \frac{\partial}{\partial \xi_3} \hat{\psi}(\xi) \right| \leq |h(\xi_1)| \cdot \left(1 + \frac{\xi_3}{|\xi_1|}\right)^{-\gamma}$,

where $\delta > 6 + \beta$, $\gamma \geq 4$, $h \in L^1(\mathbb{R})$, and similar for $\tilde{\psi}$ and $\check{\psi}$.

Further, suppose that $\mathcal{SH}(\phi, \psi, \tilde{\psi}, \check{\psi}; c; \alpha)$ forms a frame for $L^2(\mathbb{R}^3)$. Then, for $f \in \mathcal{E}_3^{\alpha, \beta}(\nu)$,

$$\|f - f_N\|_2^2 = O(N^{-\min\{\alpha/2, 2\beta/3\}} (\log N)^2), \quad N \to \infty.$$
Idea of Proof

Heuristic Argument:
We consider three cases of coefficients $\langle f, \psi_{j,k,m} \rangle$:
(a) Shearlets whose support do not overlap with the boundary $\partial B$.
(b) Shearlets whose support overlap with $\partial B$ and are nearly tangent.
(c) Shearlets whose support overlap with $\partial B$, but tangentially.
extension of cartoon-like model

Is our model sufficient?

- Typically point, curve and surface singularities are present.
Is our model sufficient?

- Typically point, curve and surface singularities are present.

Theorem (K, Lemvig, Lim; 2010):
The optimal sparse approximation result extends to a class of cartoon-like 3D images with only piecewise smooth $C^\alpha$ boundaries separating $C^\beta \cap H^\beta$ functions.
Two Surprises

**Surprise 1:**
- The optimal rate is still the same when introducing 0- and 1-dimensional singularities.

**Surprise 2:**
- Pyramid-adapted shearlets do achieve the optimal approximation rate.
- ‘Needle-like’ shearlets are not necessary for optimal sparse approximation in 3D.
Compactely Supported Generators

Dream:

- Generators of the form

\[ \psi(x_1, x_2, x_3) = \eta(x_1) \phi(x_2) \phi(x_3) \]

for the pyramid-adapted shearlets, where \( \eta \) is a 1D wavelet and \( \phi \) is a scaling/bump function.

- These generators support fast decomposition algorithms.
Compactly Supported Shearlet Frames

Theorem (K, Lim; 2010):
Let $\phi, \psi, \tilde{\psi}, \breve{\psi} \in L^2(\mathbb{R}^3)$ and for $\psi$ suppose that:

$$|\hat{\psi}(\xi)| \leq C \cdot \min(1, |\xi_1|^\delta) \cdot \prod_{i=1}^3 \min(1, |\xi_i|^{-\gamma}),$$

where $\delta > 2\gamma > 6$ and

$$\sum_{j,k} |\hat{\psi}(S_k^T A^j \xi)|^2 \geq A > 0,$$

and similar for $\tilde{\psi}$ and $\breve{\psi}$.

Then there exists a sampling constant $c_0 > 0$ such that $\mathcal{SH}(\phi, \psi, \tilde{\psi}, \breve{\psi}; c; \alpha)$ forms a frame for $L^2(\mathbb{R}^3)$ for $c \leq c_0$. 

Gitta Kutyniok, Jakob Lemvig, & Wang-Q Lim
Recent Approaches to Fast 2D Shearlet Transforms

- Filter-based implementation (Easley, Labate, Lim; 2009)
- Separable Shearlet Transform (Lim; 2009)
  \[\text{www.ShearLab.org}\]
- Decomposition based on Adaptive Subdivision Schemes (K, Sauer; 2009)
- Shearlet Unitary Extension Principle (Han, K, Shen; 2009)
- Rationally designed Digital Shearlet Transform (Donoho, K, Shahram, Zhuang; 2010)
  \[\text{www.ShearLab.org}\]
Image Denoising

Original

Noisy (20.17dB)
Image Denoising

Curvelets (28.70dB / 7.22 sec)  Shearlets (29.20dB / 5.56 sec)

(L. Demanet, L. Ying; CurveLab 2.1.2, 2008)
Image Separation: Points + Curves
Image Separation: Curves

MCALab (42.74sec)
(J. Fadili, J. Starck; MCALab 120, 2009)

ShearLab (33.75sec)
Image Separation: Points

MCALab

ShearLab
Image Separation: Points (Spines) + Curves (Dendrites)
Image Separation: Points (Spines) + Curves (Dendrites)

(Source: Brandt, K, Lim, Sundermann; 2010)

Theory: Geometric Separation (Donoho, K; 2009)
Anisotropic features in multivariate data require special efficient encoding strategies.

The shearlet theory is perfectly suited to this problem and is already very well established in the band-limited case.

One main advantage of shearlets is that they provide a unified treatment of the continuum and digital setting.

We introduced a theory for compactly supported shearlets for 2D and 3D data encompassing:

- Compactly supported shearlet frames,
- Explicit estimates for frame bounds,
- Optimal sparse approximation of cartoon-like images.
THANK YOU!

References available at: