Imaging Science meets Compressed Sensing

Gitta Kutyniok
(TU Berlin)

joint with: David Donoho (Stanford Univ.) & Wang-Q Lim (TU Berlin)

BMS Friday Colloquium
January 6, 2012
1. The Separation Problem
   - Motivating Problems
   - Goal for Today

2. Imaging Science
   - Models for Image Data
   - Mathematical Approaches

3. Compressed Sensing
   - Compressed Sensing and Component Separation
   - Avalanche of Recent Work

4. Separation of Points and Curves
   - Wavelets and Shearlets
   - Algorithm and Asymptotic Separation Result

5. Conclusions
General Challenge in Data Analysis

Modern Data in general is often composed of two or more morphologically distinct constituents, and we face the task of separating those components given the composed data.

Examples include...

- Audio data: Different instruments.

- Imaging data: Cartoon and texture.

- High-dimensional data: Lower-dimensional structures of different dimensions.
Separating Artifacts in Images, I

(Source: J. L. Starck, M. Elad, D. L. Donoho; 2005 (Artificial Data))
Separating Artifacts in Images, II

(Source: J. L. Starck, M. Elad, D. L. Donoho; 2006)
Separating Artifacts in Images, III

(Source: J. L. Starck, M. Elad, D. L. Donoho; 2006)
Separating Artifacts in Images, IV

Problem from Neurobiology

Alzheimer Research:

- Detection of characteristics of Alzheimer.
- Separation of spines and dendrites.

(Confocal-Laser Scanning-Microscopy)
Numerical Result

(Source: Brandt, K, Lim, Sündermann; 2010)
Goal for Today

Neurobiological Data:

Observed signal $x = x_1 + x_2$.

- $x_1$ = Point structures.
- $x_2$ = Curvilinear structures.

Challenges for Today:

- Mathematical methodology to derive the empirical results!
- Fundamental mathematical concept behind the empirical results!
What is

Modern Imaging Science?
Numerous Tasks in Imaging Science

- Denoising.
- Deblurring.
- Inpainting.
- Component Separation.
- Superresolution.
- ...
Examples for Modeling of Image Data

Digital Model:
- \( A \in \mathbb{R}^{N \times N} \).

Continuum Model:
- \( f \in L^2([0, 1]^2) \).
- \( f \in \mathcal{D}'(\mathbb{R}^2) \).
- ... 

\( \rightsquigarrow \) What is a ‘natural’ image?
Exploit a carefully designed representation system \((\psi_\lambda)_{\lambda \in \Lambda} \subseteq L^2(\mathbb{R}^2)\):

\[
L^2(\mathbb{R}^2) \ni f \rightarrow ((\langle f, \psi_\lambda \rangle)_{\lambda \in \Lambda} \rightarrow \sum_{\lambda \in \Lambda} \langle f, \psi_\lambda \rangle \psi_\lambda = f.
\]

**Desiderata:**

- Special features encoded in the “large” coefficients \(|\langle f, \psi_\lambda \rangle|\).
- Efficient representations:
  \[
f \approx \sum_{\lambda \in \Lambda_N} \langle f, \psi_\lambda \rangle \psi_\lambda, \quad #(\Lambda_N) \text{ small}
\]

**Methodology:**

- Modification of the coefficients according to the task.
Other Approaches to Imaging Science

PDE-based Methods:

- Given an image \( f \in L^2(\mathbb{R}^2) \).
- Let \( g : [0, \infty) \times \mathbb{R}^2 \to \mathbb{R}, \ g(0, x) = f(x). \)
- Solve

\[
F(t, x, g, \partial_1 g, ...) = 0, \quad g(0, x) = f(x).
\]
Other Approaches to Imaging Science

PDE-based Methods:

- Given an image \( f \in L^2(\mathbb{R}^2) \).
- Let \( g : [0, \infty) \times \mathbb{R}^2 \rightarrow \mathbb{R}, \; g(0, x) = f(x) \).
- Solve
  \[
  F(t, x, g, \partial_1 g, \ldots) = 0, \quad g(0, x) = f(x).
  \]

Variational Methods:

- Given an image \( f \in L^2(\mathbb{R}^2) \).
- Introduce functionals \( \Phi, \Psi : L^2(\mathbb{R}^2) \rightarrow \mathbb{R} \).
- Solve
  \[
  \min_g \Phi(f - g) + \mu \Psi(g).
  \]
How does Compressed Sensing help with Component Separation?
Model for 2 Components:

- Observe a signal $x$ composed of two subsignals $x_1$ and $x_2$:
  \[ x = x_1 + x_2. \]

- Extract the two subsignals $x_1$ and $x_2$ from $x$, if only $x$ is known.
Model for 2 Components:

- Observe a signal $x$ composed of two subsignals $x_1$ and $x_2$:
  $$x = x_1 + x_2.$$ 
- Extract the two subsignals $x_1$ and $x_2$ from $x$, if only $x$ is known.

Isn’t this impossible?

- There are two unknowns for every datum.
Model for 2 Components:

- Observe a signal $x$ composed of two subsignals $x_1$ and $x_2$:
  \[ x = x_1 + x_2. \]

- Extract the two subsignals $x_1$ and $x_2$ from $x$, if only $x$ is known.

Isn’t this impossible?

- There are two unknowns for every datum.

But we have additional Information:

- The two components are geometrically different.
Problem: \[ x = x_1^0 + x_2^0 = \Phi_1 c_1^0 + \Phi_2 c_2^0 = [\Phi_1 \mid \Phi_2] \begin{bmatrix} c_1^0 \\ c_2^0 \end{bmatrix}, \]

Composition of Sinusoids and Spikes sampled at \( n \) points:

- \( x, c_1^0, \) and \( c_2^0 \in \mathbb{R}^n. \)
- \( \Phi_1 \) is the \( n \times n \)-Fourier matrix \((\Phi_1)_{t,k} = e^{2\pi i tk/n}\).
- \( \Phi_2 \) is the \( n \times n \)-Identity matrix.
Compressed Sensing

Observation:
Let $A$ be an $n \times N$-matrix, $n << N$. In many situations the sought solution $c^0$ of $x = Ac^0$ is sparse, i.e.,

$$\|c^0\|_0 = \#\{i : c_i^0 \neq 0\} \text{ is ‘small’}.$$
Compressed Sensing

Observation:
Let $A$ be an $n \times N$-matrix, $n \ll N$. In many situations the seeked solution $c^0$ of $x = Ac^0$ is sparse, i.e.,

$$\|c^0\|_0 = \# \{ i : c^0_i \neq 0 \} \text{ is ‘small’}.$$ 

First idea: Solve...

$$(P_0) \min_c \|c\|_0 \text{ such that } x = Ac$$
Compressed Sensing

Observation:
Let $A$ be an $n \times N$-matrix, $n << N$. In many situations the seeked solution $c^0$ of $x = Ac^0$ is sparse, i.e.,

$$\|c^0\|_0 = \#\{i : c_i^0 \neq 0\}$$

is ‘small’.

First idea: Solve...

$$(P_0) \min_c \|c\|_0 \text{ such that } x = Ac$$

$\leadsto$ This problem is NP-hard!
Compressed Sensing

Observation:
Let $A$ be an $n \times N$-matrix, $n << N$. In many situations the seeked solution $c^0$ of $x = Ac^0$ is sparse, i.e.,

$$\|c^0\|_0 = \#\{i : c_i^0 \neq 0\} \text{ is 'small'}. $$

First idea: Solve...

$$(P_0) \quad \min_c \|c\|_0 \text{ such that } x = Ac$$

$\leadsto$ This problem is NP-hard!

Basis Pursuit (Chen, Donoho, Saunders; 1998)

$$(P_1) \quad \min_c \|c\|_1 \text{ such that } x = Ac$$
Intuition

\[ \{ c : x = Ac \} \]
Exact Recovery by $\ell_1$ Minimization

Meta-Result: If

- $\|c^0\|_0$ is sufficiently small,
- $A$ is sufficiently incoherent,

then

$$c^0 = \arg\min_c \|c\|_1 \quad \text{such that} \quad x = Ac.$$
Exact Recovery by $\ell_1$ Minimization

Meta-Result: If

- $\|c^0\|_0$ is sufficiently small,
- $A$ is sufficiently incoherent,

then

$$c^0 = \arg\min_c \|c\|_1 \text{ such that } x = Ac.$$ 

Exemplary Result (Donoho, Elad; 2003)
Let $A$ be an $n \times N$-matrix with normalized columns, $n \ll N$, and let $c^0 \in \mathbb{R}^N$ satisfy

$$\|c^0\|_0 < \frac{1}{2} \left( 1 + \frac{1}{\mu(A)} \right),$$

where the coherence $\mu(A)$ is defined by $\mu(A) = \max_{i \neq j} |\langle a_i, a_j \rangle|$. Then

$$c^0 = \arg\min_c \|c\|_1 \text{ such that } x = Ac.$$
Composition of Sinusoids and Spikes sampled at $n$ points:

$$x = x_1^0 + x_2^0 = \Phi_1 c_1^0 + \Phi_2 c_2^0 = [\Phi_1 \mid \Phi_2] \begin{bmatrix} c_1^0 \\ c_2^0 \end{bmatrix}.$$ 

Coherence of $[\Phi_1 \mid \Phi_2]$:

$$\mu([\Phi_1 \mid \Phi_2]) = \mu([F \mid I]) = \frac{1}{\sqrt{n}}.$$
Composition of \textbf{Sinusoids} and \textbf{Spikes} sampled at \(n\) points:

\[
x = x_1^0 + x_2^0 = \Phi_1 c_1^0 + \Phi_2 c_2^0 = \begin{bmatrix} \Phi_1 & \Phi_2 \end{bmatrix} \begin{bmatrix} c_1^0 \\ c_2^0 \end{bmatrix}.
\]

Coherence of \([\Phi_1|\Phi_2]\):

\[
\mu([\Phi_1|\Phi_2]) = \mu([F|I]) = \frac{1}{\sqrt{n}}.
\]

\textbf{Theorem (Donoho, Huo; 2001)}

If \(\#(\text{Sinusoids}) + \#(\text{Spikes}) = \| (c_1^0) \|_0 + \| (c_2^0) \|_0 < \frac{(1 + \sqrt{n})}{2}\), then

\[
(c_1^0, c_2^0) = \text{argmin}(\| c_1 \|_1 + \| c_2 \|_1) \quad \text{subject to} \quad x = \Phi_1 c_1 + \Phi_2 c_2.
\]
Let $x$ be a signal composed of two subsignals $x_1^0$ and $x_2^0$:

$$x = x_1^0 + x_2^0.$$ 

Desiderata for two orthonormal bases $\Phi_1$ and $\Phi_2$:

- $x_i^0 = \Phi_i c_i^0$ with $\|c_i^0\|_0$ small, $i = 1, 2 \leadsto$ Sparsity!
- $\mu([\Phi_1|\Phi_2])$ small $\leadsto$ Morphological Difference!

Solve

$$(c_1^*, c_2^*) = \operatorname{argmin}(\|c_1\|_1 + \|c_2\|_1) \text{ subject to } x = \Phi_1 c_1 + \Phi_2 c_2$$

and derive the approximate components

$$x_i^0 \approx x_i^* = \Phi_i c_i^*, \quad i = 1, 2.$$
Two Paths
Problem: Solve $x = Ac^0$ with $A$ an $n \times N$-matrix ($n < N$).

Results using structured matrices $A$:
- $A$ is often to some extent given by the application.
- When can $c^0$ still be recovered and how fast?
- Contributors: Candès, Donoho, Elad, Rauhut, Temlyakov, Tropp, ...

Results using random matrices $A$:
- The ‘best’ $A$ is a random matrix.
- What is maximally possible if $A$ can be freely chosen?
- Contributors: Candès, Donoho, Pajor, Romberg, Tanner, Tao, ...

Remark: Matheon-Talk by Emmanuel Candès (June 20th).
How can these Ideas be applied to Separation of Points and Curves?
Two morphologically distinct components:
- Points
- Curves

Choose suitable representation systems which provide optimally sparse representations of
- pointlike structures $\rightarrow$ Wavelets
- curvelike structures $\rightarrow$ Shearlets

Minimize the $\ell_1$ norm of the coefficients.

This forces
- the pointlike objects into the wavelet part of the expansion
- the curvelike objects into the shearlet part.
Empirical Separation of Spines and Dendrites

Wavelet Expansion

Shearlet Expansion

(Source: Brandt, K, Lim, Sündermann; 2010)
Wavelets

Definition:
The wavelet system associated with $\psi \in L^2(\mathbb{R}^2)$ is defined by

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

Theorem:
Let $f \in C^2(\mathbb{R}^2)$ except finitely many point singularities. Then wavelets provide an optimally sparse approximation of $f$, i.e.,

$$\|f - f_N\|_2^2 \leq C \cdot N^{-1}, \quad N \to \infty,$$

where $f_N = \sum_{\lambda \in \Lambda_N} c_\lambda \psi_\lambda$. 
Beyond Wavelets...

Observation:
- Wavelets can not approximate curvilinear singularities optimally sparse.
- Reason: Isotropic structure of wavelets:

\[ 2^j \psi\left( \begin{pmatrix} 2^j & 0 \\ 0 & 2^j \end{pmatrix} x - m \right) \]

Intuitive explanation:
Shearlets

Parabolic scaling:

\[
A_j = \begin{pmatrix} 2^j & 0 \\ 0 & 2^{j/2} \end{pmatrix}, \quad j \in \mathbb{Z}.
\]

Orientation via shearing:

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

Definition (K, Labate, Lim; 2006):

For \( \psi \in L^2(\mathbb{R}^2) \), the associated shearlet system is defined by

\[
\mathcal{S}\mathcal{H}(\psi) = \{ 2^{\frac{3j}{4}} \psi(S_k A_j \cdot -m) : j, k \in \mathbb{Z}, m \in \mathbb{Z}^2 \}.
\]
Compactly Supported Shearlets

Theorem (Kittipoom, K, Lim; 2010):
Let \( \psi \in L^2(\mathbb{R}^2) \) be compactly supported, and let \( \hat{\psi} \) satisfy certain decay conditions. Then \( \mathcal{SH}(\psi) = (\sigma_\eta)_\eta \) forms a frame with controllable frame bounds, i.e.,

\[
A \|f\|_2^2 \leq \sum_\eta |\langle f, \sigma_\eta \rangle|^2 \leq B \|f\|_2^2 \quad \text{for all } f \in L^2(\mathbb{R}^2).
\]

Theorem (K, Lim; 2010):
Let \( \psi \in L^2(\mathbb{R}^2) \) be compactly supported, and let \( \hat{\psi} \) satisfy certain decay conditions. Then \( \mathcal{SH}(\psi) \) provides an optimally sparse approximation of \( f \), i.e.,

\[
\|f - f_N\|_2^2 \leq C \cdot N^{-2} \cdot (\log(N))^3, \quad N \to \infty.
\]
Optimal for Pointlike Structures:
Orthonormal Wavelets are a basis with perfectly isotropic generating elements at different scales.

Optimal for Curvelike Structures:
Shearlets (K, Labate, Lim; 2006) are a highly directional frame with increasingly anisotropic elements at fine scales (→ www.ShearLab.org).
Separation Algorithm

Observed Object:

\[ f = \mathcal{P}^0 + \mathcal{C}^0. \]

Subband Decomposition:
Wavelets and shearlets use the same scaling subbands!

\[ f_j = \mathcal{P}_j^0 + \mathcal{C}_j^0, \quad \mathcal{P}_j^0 = \mathcal{P}^0 \ast F_j \quad \text{and} \quad \mathcal{C}_j^0 = \mathcal{C}^0 \ast F_j. \]

\( \ell_1 \)-Decomposition:

\[(\mathcal{P}_j^*, \mathcal{C}_j^*) = \text{argmin} \| (\langle \mathcal{P}_j, \psi_\lambda \rangle_\lambda \|_1 + \| (\langle \mathcal{C}_j, \sigma_\eta \rangle_\eta \|_1 \quad \text{s.t.} \quad f_j = \mathcal{P}_j + \mathcal{C}_j \]

Gitta Kutyniok (TU Berlin)  
Imaging Science meets Compressed Sensing  
BMS Friday Colloquium
Empirical Separation of Spines and Dendrites

Wavelet Expansion

Shearlet Expansion

(Source: Brandt, K. Lim, Sündermann; 2010)
Microlocal Model

Neurobiological Geometric Mixture in 2D:

Point Singularity:

$$\mathcal{P}^0(x) = \sum_{i=1}^{P} |x - x_i|^{-3/2}$$

Curvilinear Singularity:

$$\mathcal{C}^0 = \int \delta_{\tau(t)} dt, \quad \tau \text{ a closed } C^2\text{-curve.}$$

Observed Signal:

$$f = \mathcal{P}^0 + \mathcal{C}^0$$
Asymptotic Separation

Theorem (Donoho, K; 2010)

\[
\frac{\|P_j^* - P_j^0\|_2 + \|C_j^* - C_j^0\|_2}{\|P_j^0\|_2 + \|C_j^0\|_2} \to 0, \quad j \to \infty.
\]

At all sufficiently fine scales, nearly-perfect separation is achieved!
Microlocal Analysis Heuristics

Singular Support and Wavefront Set of $\mathcal{P}^0$ and $\mathcal{C}^0$:

Phase Space Portrait of Wavelets and Shearlets:
Let’s conclude…
What to take Home...?

- One main task in imaging science: Component Separation.
- One approach to imaging science: Applied Harmonic Analysis.
- Compressed Sensing allows exact solution of underdetermined linear systems of equations if the solution is sparse and the matrix is incoherent.
- Separation of point- and curvelike structures:
  - Wavelets sparsify points and shearlets sparsify curves.
  - Morphological distance encoded in incoherence.
  - Solution: $\ell_1$ minimization.
THANK YOU!

References available at:

page.math.tu-berlin.de/~kutyniok

Related Books:

- Y. Eldar and G. Kutyniok
  *Compressed Sensing: Theory and Applications*

- G. Kutyniok and D. Labate
  *Shearlets: Multiscale Analysis for Multivariate Data*