Applied Harmonic Analysis Methods in Imaging Science
Introduction

Gitta Kutyniok
(Technische Universität Berlin)

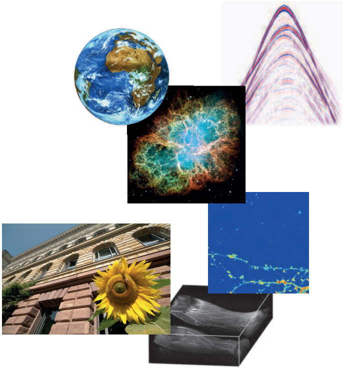
SIAM Conference on Imaging Science
Albuquerque, May 23 – 26, 2016

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, ...

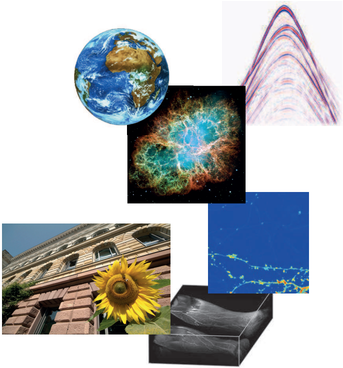


Imaging Science Today

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

Main Tasks

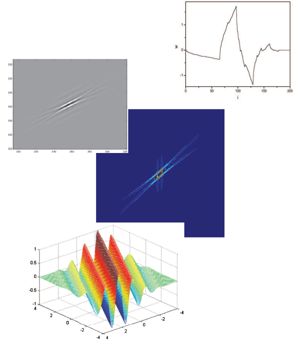
- Acquisition
- Preprocessing
  - Denoising, Inpainting, ...
- Analysis
  - Feature Detection, ...
- Storing
  - Compression, ...



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 $0 < A \leq B < \infty$ 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

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

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

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

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

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

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

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

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

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

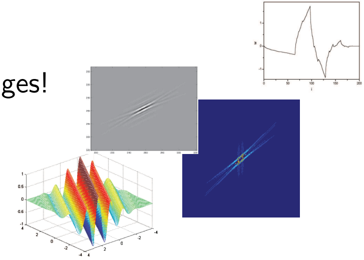
Situation of a Frame: For the N-term approximation $x_N = \sum_{\lambda \in \Lambda_N} \langle x, \psi_\lambda \rangle \tilde{\psi}_\lambda$ of x consisting of the N largest coefficients $|\langle x, \psi_\lambda \rangle|$, we *only* have

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

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, ...