Harmonic Analysis and Big Data: Introduction

Wojciech Czaja

Institute for Mathematics and Its Applications 2015 PI Summer Graduate Program
College Park, August 5–7, 2015

Introduction

- There is an abundance of available data. This data is often large, high-dimensional, noisy, and complex, e.g., geospatial imagery.
- Typical problems associated with such data are to cluster, classify, or segment it; and to detect anomalies or embedded targets.
- Our proposed approach to deal with these problems is by combining techniques from harmonic analysis and machine learning:
  - **Harmonic Analysis** is the branch of mathematics that studies the representation of functions and signals.
  - **Machine Learning** is the branch of computer science concerned with algorithms that allow machines to infer rules from data.

Big Data

"Big Data" refers to the exponential growth data, with many challenging tasks of analyzing and efficiently finding the important information that is given in this complex setting.

- The roots of big data are in the data storage, database management, and data analytics for, both, commercial and non-profit applications.
- The integration of many large datasets is a primary source of big data problems present in the modern scientific and research environment, as is evident in applications ranging from 'omics' data analysis for cancer research, to studies of social networks.
- Another source of big data problems are large and heterogeneous dynamic data sets, such as those arising in the context of climate change analysis, or for the analysis of network traffic patterns.

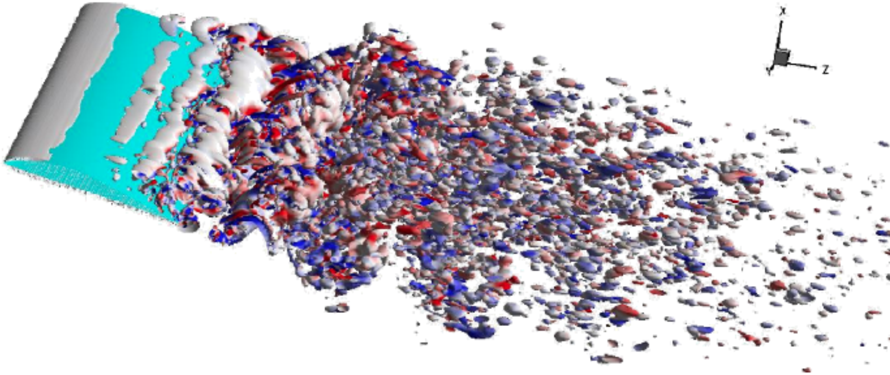
Big Data Characteristics

In view of the above, big data can be identified by the following:

- volume;
- heterogeneity;
- dynamics.

In addition to the above major characteristics, we can add: ambiguity, complexity, noise, variability, etc.

Big Data Example 1

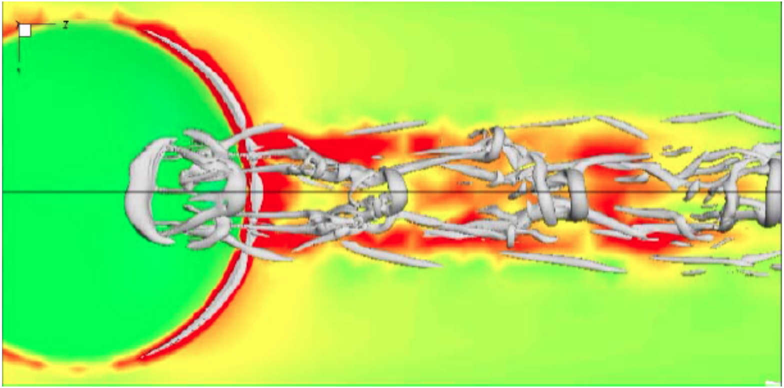


Large Eddy simulation (LES) around an Eppler foil at Re=10,000. A series of high fieldity LES of the flow around Eppler airfoils has been conducted to generate a comprehensive data base. Reynolds numbers vary from 10,000 to 120,000 and the angle of attach varies from 0 to 20 degrees.

Courtesy of Prof. Elias Balaras (GWU), via US Air Force contract FA9550-12-C-058 (2012): Learning from Massive Data Sets Generated by Physics Based Simulations

Norbert Wiener Center for Harmonic Analysis and Applications

Big Data Example 2



A simplified case of the previous LES for a 3-dimensional flow over a dimpled plate.

Courtesy of Prof. Elias Balaras (GWU), via US Air Force contract FA9550-12-C-058 (2012): Learning from Massive Data Sets Generated by Physics Based Simulations

Big Data Example Estimation

Let us provide a small numerical estimation:

- 2,000 x 1,000 x 1,000 = 2 x 10^9 grid points;
- Each grid point characterized by 3 spatial coordinates and 3 velocity components, pressure, plus possibly some other parameters;
- Flow simulation for 200 time steps;
- One way to look at it: 2 x 10^9 points in a space of dimension 1,400;
- As an example, think of computing PCA for M points in N dimensional space. The cost is $O(MN^2) + O(N^3)$;
- In our case this results in a problem with complexity on the order of 4 x 10^{15} = 4 petaFLOPs;
- Lawrence Livermore National Laboratory's IBM Sequoia reaches 16 petaFLOPS (16 x 10^{15} floating point operations per second) - it was considered to be the fastest computer in 2012, it runs 1.57 million PowerPC cores, costs approx. 250M USD.

Norbert Wiener Center for Harmonic Analysis and Applications

Another Big Data Example

- Consider the human genome. First estimates pointed at 100,000 genes. Nowadays this number has been scaled down to appprox. 45,000.
- There are many ways of representing genes. One of the more popular is by means of base pairs: approx. three billion DNA base pairs represent human genome.
- Alternatively, we could consider gene expressions (think of it as a function). There are many ways of assembling such expressions, and they are different for different individuals. Hence resulting in a much larger data set.

Norbert Wiener Center for Harmonic Analysis and Applications

HA and Big Data

- Multiscale methods
- Compressive sensing
- Sparse representations
- Geometric and graph-based methods

Multiscale representations

- **Multiscale representation** (Multiresolution analysis (MRA), pyramid algorithms) can be described as a class of design methods in representation theory, where the input is subject to repeated transformation (filtering) in order to extract features associated with different scales.
- In image processing and computer graphics the concept of multiscale representations can be traced back to P. Burt and E. Adelson, and J. Crowley.
- In mathematics, it is associated with wavelet theory and MRA as introduced by Y. Meyer and S. Mallat

  S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation", IEEE TPAMI, 1989, Vol. 11, pp. 674–693.
- Multiscale representations found many applications to image processing and remote sensing: compression, feature detection, segmentation, classification, but also in registration and image fusion.

  G. Pajares and J. Cruz, "A wavelet-based image fusion tutorial", Pattern Recognition, 2004, Vol. 37(9), pp. 1855–1872.

Limitations of traditional wavelet representations

- Wavelets provide optimal representations for 1-dimensional signals in the sense of measuring asymptotic error with N largest coefficients in wavelet expansion, and are superior to Fourier-type representations.
- However, in dimensions higher than 1, wavelets are known to be suboptimal for representing objects with curvilinear singularities (edges), even though they outperform Fourier methods.

  D. Donoho et al., "Data compression and harmonic analysis", IEEE TIT, 1998, Vol. 44, pp. 2435–2476.
- A number of techniques have been proposed since the introduction of wavelets to address this issue, and to find better description of geometric features in images.

  L. Jacques et al., "A panorama on multiscale geometric representations, intertwining spatial, directional and frequency selectivity",

  Signal Processing, 2011, Vol. 91, pp. 2699–2730.

Norbert Wiener Center
for Harmonic Analysis and Applications

Anisotropic Harmonic Analysis

- Harmonic analysis decomposes signals into simpler elements called *analyzing functions*.
- Classical HA methods include Fourier series and aforementioned wavelets. These have proven extremely influential and quite effective for many applications.
- However, they are fundamentally isotropic, meaning they decompose signals without considering how the signal varies directionally.
- Wavelets decompose an image signal with respect to translation and scale. Since the early 2000s, there have been several attempts to incorporate directionality into the wavelet construction.

Multiscale Directional Representations

- Early attempts to make wavelets more sensitive to directionality included appropriate filter design, anisotropic scaling, steerable filters, and similar techniques.
- **Directional wavelets:** J.-P. Antoine, R. Murenzi, P. Vandergheynst, and S. Ali introduced more complicated group actions for parametrization of 2-dimesnional wavelet transforms, including rotations or similitude group. These results were later generalized to construct wavelets on sphere and other manifolds.

J.-P. Antoine, D. Rosca, P. Vandergheynst,"Wavelet transform on manifolds: old and new approaches", ACHA, 2010, Vol. 28 (2),189–202.

- Subsequently **Radon transform** has been introduced in combination with wavelet transforms to replace the angular parametrization; This results in systems such as **ridgelets** (E. Candès and D. Donoho) or **Gabor ridge functions** (L. Grafakos and C. Sansing)
- **Contourlets:** M. Do and M. Vetterli constructed a discrete-domain multiresolution and multidirection expansion using non-separable filter banks, in much the same way that wavelets were derived from filter banks.

Multiscale Directional Representations

- **Curvelets:** E. Candès and D. L. Donoho introduced the curvelets as an efficient tool to extract directional information from images. Curvelets consist of translations and rotations of a sequence of basic functions depending on a parabolic scaling parameter. The curvelet transform is first developed in the continuous domain and then discretized for sampled data.
- **Wavelets with Composite Dilations:** K. Guo, D. Labate, W.-Q. Lim, B. Manning, G. Weiss, and E. Wilson studied affine systems built by using a composition of two sets of matrices as the dilation.
- **Shearlets:** D. Labate, K. Guo, G. Kutyniok, and G. Weiss introduced a special example of the Composite Dilation Wavelets.
- **Surfacelets** (Do, Lu), **bandlets** (Le Pennec, Mallat), **brushlets** (Meyer, Coifman), **wedgelets** (Donoho), **phaselets** (Gopinath), **complex wavelets** (Daubechies), **surflets** (Baraniuk), etc, etc ...

Anisotropic Harmonic Analysis

- These constructions incorporate directionality in a variety of ways.
- To summarize, some of the major constructions include:
  - Ridgelets.

E. Candès. *Ridgelets: theory and applications*. PhD thesis. (1998).

  - Curvelets.

D. Donoho and E. Candès. *Curvelets: A surprisingly effective nonadaptive representation for objects with edges*. Curve and Surface Fitting. (1999).

  - Contourlets.

M. Do and M. Vetterli. *Contourlets*. Beyond Wavelets. (2001).

  - Shearlets.

D. Labate, W.-Q. Lim, G. Kutyniok, and G. Weiss. *Sparse multidimensional representation using shearlets*. Proc. SPIE 5914. (2005).

- Wavelets, ridgelets, curvelets, and shearlets are surprisingly related, as they all are special cases of the recently introduced $\alpha-$molecules.

P. Grohs, S. Keiper, G. Kutyniok, and M. Schäfer. $\alpha-$*molecules*. arXiv: 1407.4424. (2014).

Norbert Wiener Center for Harmonic Analysis and Applications

Multiscale Directional Representations

- Many of the aforementioned representations were designed specifically for dealing with images, i.e., for the case of 2-dimensional Euclidean space.
- Multiscale directional representations can also be constructed analogously for higher dimensional spaces, as well as for some manifolds.
- A different approach is needed to deal with discrete structures, such as graphs, networks, or point clouds. R. Coifman and M. Maggioni proposed to use diffusion processes on such structures to introduce the notion of scale and certain directions.

R. R. Coifman and M. Maggioni, “Diffusion wavelets,” ACHA, 2006, Vol. 21(1), pp. 53–94.

Anisotropic Harmonic Analysis

- A useful model for real images is the class of *cartoon-like images*, $\mathcal{E}^2(\mathbb{R}^2)$.
- Roughly, they are functions that are smooth away from a smooth curve of discontinuity.
- Let $f \in \mathcal{E}^2(\mathbb{R}^2)$ and let f_N be its best N-term approximation with respect to a set of analyzing functions. The optimal asymptotic decay rate of $||f - f_N||_2^2$ is $O(N^{-2})$, $N \to \infty$, achieved adaptively.
- Up to a log factor, curvelets, contourlets, and shearlets satisfy this optimal decay rate (ridgelets are only optimal for linear boundaries). Hence, these analyzing functions are *essentially optimally sparse* for cartoon-like images. Wavelets can only achieve $O(N^{-1})$. Fourier series are even worse with $O(N^{-1/2})$.
- We focus on shearlets since they have multiple, efficient numerical implementations.

Shearlets

Continuous shearlets in $\mathbb{R}^2$ depend on three parameters: the scaling parameter $a > 0$, the shear parameter $s \in \mathbb{R}$, and the translation parameter $t \in \mathbb{R}^2$, and they are defined as follows:
We define the *parabolic scaling matrices*

$$A_a = \begin{pmatrix} a & 0 \\ 0 & a^{1/2} \end{pmatrix}, \qquad a > 0$$

and the *shearing matrices*

$$S_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}, \qquad s \in \mathbb{R}.$$

Also, let D_M be the dilation operator defined by

$$D_M\psi = |\det M|^{-1/2}\psi(M^{-1}\cdot), \qquad M \in GL_2(\mathbb{R})$$

and T_t the translation operator defined by

$$T_t\psi = \psi(\cdot - t), \qquad t \in \mathbb{R}^2.$$

Norbert Wiener Center
for Harmonic Analysis and Applications

Shearlets

Definition

Let $\psi \in L^2(\mathbb{R}^2)$. The *Continuous Shearlet Transform* of $f \in L^2(\mathbb{R}^2)$ is

$$f \mapsto \mathcal{SH}_\psi f(a,s,t) = \langle f, T_t D_{A_a} D_{S_s} \psi \rangle, a > 0, s \in \mathbb{R}, t \in \mathbb{R}^2.$$

- Parabolic scaling allows for directional sensitivity.
- Shearing allows us to change this direction.
- By carefully choosing ψ and discretizing the parameter space, we can decompose $f \in L^2(\mathbb{R}^2)$ into a Parseval frame.

Norbert Wiener Center for Harmonic Analysis and Applications

- It's generally assumed that $\hat{\psi}$ splits as $\hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\hat{\psi}_2(\xi_2/\xi_1)$.
- The basic shearlet ψ is only used in a *horizontal cone*, while the reflection of ψ across the line $\xi_2 = \xi_1$ is used in a *vertical cone*. A scaling function ϕ is used for the low-pass region. This construction is known as *cone-adapted* shearlets.

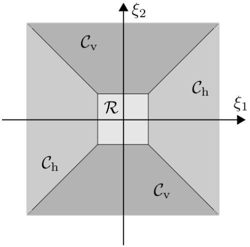


Figure : Frequency tiling for cone-adapted shearlets.

G. Kutyniok and D. Labate, eds. *Shearlets: Multiscale analysis for multivariate data.* Birkhäuser. (2012).

Shearlet Implementations

- Shearlets have several efficient numerical implementations in MATLAB that are freely available.
  - 2D Shearlet Toolbox (Easley, Labate, and Lim). [1]
  - Shearlab (Kutyniok, Shahram, Zhuang et al.). [2]
  - Fast Finite Shearlet Transform (Häuser and Steidl).[3]
- We used the last option (FFST) here, which is in many ways the most intuitive of the implementations.

---

[1] http://www.math.uh.edu/~dlabate/software.html
[2] http://www.shearlab.org/
[3] http://www.mathematik.uni-kl.de/imagepro/software/ffst/

Fast Finite Shearlet Transform (FFST)

- Consider an $M \times N$ image. Define $j_0 := \lfloor \log_2 \max\{M, N\} \rfloor$. We discretize the parameters as follows:

$$a_j := 2^{-2j} = \frac{1}{4^j}, \qquad j = 0, \ldots, j_0 - 1,$$

$$s_{j,k} := k2^{-j}, \qquad -2^j \leq k \leq 2^j,$$

$$t_m := \left(\frac{m_1}{M}, \frac{m_2}{N}\right), \quad m_1 = 0, \ldots, M-1, \; m_2 = 0, \ldots, N-1.$$

- Note that the shears vary from -1 to 1. To fill out the remaining directions, we also shear with respect to the y-axis.
- Shearlets whose supports overlap are "glued" together.
- The transform is computed through the $2D$ FFT and iFFT.

Fast Finite Shearlet Transform

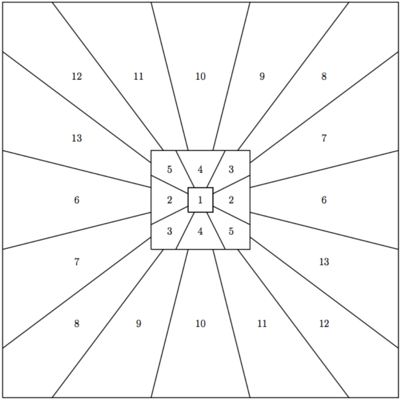
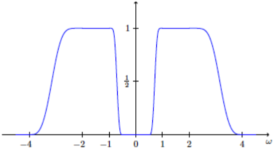


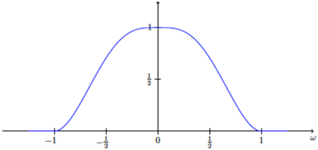
Figure : Frequency tiling for FFST.

S. Häuser and G. Steidl. *Fast finite shearlet transform: a tutorial*. arXiv:1202.1773. (2014).

Additional Picture for FFST



(a) $\hat{\psi}_1$.



(b) $\hat{\psi}_2$.

Figure : $\hat{\psi}_1$ and $\hat{\psi}_2$ for the FFST (ibid.).

Fast Finite Shearlet Transform

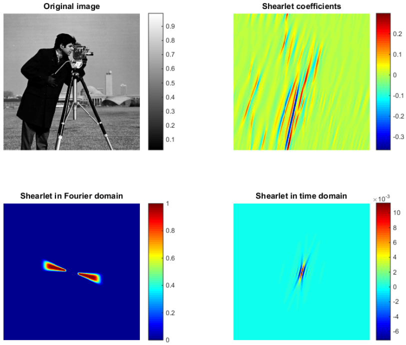


Figure : Demonstration of output from the FFST on the cameraman image. The shearlet coefficients are from scale 3 (out of 4) in the direction of slope 4.

How Well Can the FFST Resolve Directions?

We can prove that the direction of the shearlet coefficient of maximum magnitude determines the direction, at least in the ideal case.

Theorem (with D. Weinberg, 2015)

Let $f(x) = H_{y>rx}$ be a 2D Heaviside function and assume WLOG that $|r| \leq 1$. Fix a scale j and position m. Then the shearlet coefficient of the FFST $\mathcal{SH}(f)(j,k,m)$ is only nonzero for at most two consecutive values of the shearing parameter k. The value of k that maximizes $|\mathcal{SH}(f)(j,k,m)|$ satisfies

$$\left|s_{j,k} - r\right| < \frac{1}{2^j}.$$

Furthermore, for this k, $s_{j,k}$ is closest to r over all k.

D. Weinberg, Multiscale and Directional Representations of High-dimensional Information in Remotely Sensed Data, Ph.D. Thesis, University of Maryland College Park, 2015

Norbert Wiener Center for Harmonic Analysis and Applications

Sketch of Proof

- We first show by direct computation that $\int_{\mathbb{R}} \hat{\psi}(-r\omega, \omega)d\omega = \int_{\mathbb{R}} \psi(x, rx)dx$ for all $\psi \in S(\mathbb{R}^2), r \in \mathbb{R}$.
- Since $\frac{\partial}{\partial y} H_{y>rx} = \delta_{y-rx}$, $\widehat{H_{y>rx}} = \frac{1}{2\pi i\omega_2}\widehat{\delta_{y-rx}}$.
- Using the above, $\langle \widehat{\delta_{y-rx}}, \hat{\psi} \rangle = \int_{\mathbb{R}} \hat{\psi}(-r\omega, \omega)d\omega, \qquad \hat{\psi} \in C_c^\infty(\mathbb{R}^2)$.
- We compute

$$
\begin{aligned}
\mathcal{SH}(f)(j,k,m) &= \langle f, \psi_{jkm} \rangle \\
&= \langle \hat{f}, \hat{\psi}_{jkm} \rangle \\
&= \int_{\mathbb{R}^2} \frac{1}{2\pi i\omega_2} \widehat{\delta_{y-rx}}(\omega_1, \omega_2)\hat{\psi}_{jkm}(\omega_1, \omega_2)d\omega_1 d\omega_2 \\
&= \frac{1}{2\pi i} \int_{\mathbb{R}} \frac{1}{\omega_2} \hat{\psi}_{jkm}(-r\omega_2, \omega_2)d\omega_2.
\end{aligned}
$$

Norbert Wiener Center for Harmonic Analysis and Applications

Sketch of Proof

- By the way $\hat{\psi}$ decomposes,

$$\hat{\psi}_{jkm}(-r\omega_2, \omega_2) = \hat{\psi}_1(4^{-j}\omega_2)\hat{\psi}_2(-2^j r+k)\exp(-2\pi i(-r\omega_2 m_1/M+\omega_2 m_2/N$$

- Since k only appears in $\hat{\psi}_2(-2^j r + k)$, we examine that term separately.
- By assumption, $\hat{\psi}_2$ is a positive, smooth function supported on $[-1, 1]$ that is strictly increasing on $[-1, 0]$ and decreasing on $[0, 1]$.
- Hence, to obtain a nonzero shearlet coefficient, we must have $|-2^j r + k| < 1$ or $|s_{j,k} - r| < 1/2^j$.
- The shearlet slopes differ by $1/2^j$, so this can only occur at most twice.
- The coefficient is maximized when $-2^j r + k$ is closest to 0, that is, when $s_{j,k}$ is closest to r.

Norbert Wiener Center for Harmonic Analysis and Applications

Shearlets - a few results

- Shearlet multiresolution analysis theory and decomposition algorithm - G. Kutyniok and T. Sauer
- Shearlet approach to edge detection - S. Yi, D. Labate, G. Easley, and H. Krim
- Shearlet-based method to invert the Radon transform - F. Colonna, G. R. Easley, K. Guo, and D. Labate

All of the above results are in the setting of $\mathbb{R}^2$. But we are also interested in higher dimensional constructions.

Norbert Wiener Center for Harmonic Analysis and Applications

Translation-dilation-shearing group

- E. Cordero, F. DeMari, K. Nowak, and A. Tabacco introduced the following group interpretation of shearlets, by means of the **Translation-Dilation-Shearing Group**:

$$\left\{ A_{t,\ell,y} = \begin{pmatrix} t^{-1/2} S_{\ell/2} & 0 \\ t^{-1/2} B_y S_{\ell/2} & t^{1/2} (S^t_{-\ell/2}) \end{pmatrix} : t > 0, \ell \in \mathbb{R}, y \in \mathbb{R}^2 \right\},$$

where $B_y = \begin{pmatrix} 0 & y_1 \\ y_1 & y_2 \end{pmatrix}$, $y = (y_1, y_2)^t \in \mathbb{R}^2$, $S_\ell = \begin{pmatrix} 1 & \ell \\ 0 & 1 \end{pmatrix}$.

- Stems from attempts to characterize reproducing subgroups of $\mathbb{R}^{2d} \rtimes \mathrm{Sp}(2d, \mathbb{R})$.
- We aim to generalize it to $d \geq 2$.

Translation-dilation-shearing group in $\mathbb{R}^k$

- For $k \geq 1$, we define $(\mathrm{TDS})_k$ to be

$$\left\{ A_{t,\ell,y} = \begin{pmatrix} t^{-1/2} S_{\ell/2} & 0 \\ t^{-1/2} B_y S_{\ell/2} & t^{1/2} S^t_{-\ell/2}) \end{pmatrix} : t > 0, \ell \in \mathbb{R}^{k-1}, y \in \mathbb{R}^k \right\}$$

- $y = (y_1, y_2, \cdots, y_k)^t \in \mathbb{R}^k$
- $B_y = \begin{pmatrix} 0 & 0 & \dots & y_1 \\ 0 & 0 & \dots & y_2 \\ \vdots & \vdots & \ddots & \vdots \\ y_1 & y_2 & \dots & y_k \end{pmatrix}$, $S_\ell = \begin{pmatrix} 1 & 0 & \dots & 0 & \ell_1 \\ 0 & 1 & \dots & 0 & \ell_2 \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \dots & 1 & \ell_{k-1} \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}$

- The *symplectic group* $\mathrm{Sp}(d,\mathbb{R})$ is the subgroup of $2d \times 2d$ matrices $g \in \mathrm{M}(2d,\mathbb{R})$ which satisfy $g^t \mathcal{J} g = \mathcal{J}$, where $\mathcal{J} = \begin{pmatrix} 0 & I_d \\ -I_d & 0 \end{pmatrix}$.

Theorem (with E. King)

For any $k \geq 1$, $(\mathrm{TDS})_k$ is a Lie subgroup of $\mathrm{Sp}(k,\mathbb{R})$ of dimension $2k$. The left Haar measures, up to normalization, of $(\mathrm{TDS})_k$ are $d\tau = \frac{dt}{t^2} dy$ for $k = 1$ and $d\tau = \frac{dt}{t^{k+1}} dy d\ell$ for $k > 1$, where dt, dy and $d\ell$ are the Lebesgue measures over $\mathbb{R}^+$, $\mathbb{R}^k$ and $\mathbb{R}^{k-1}$, respectively.

E. J. King, Wavelet and frame theory: frame bound gaps, generalized shearlets, Grassmannian fusion frames, and p-adic wavelets, Ph.D. Thesis, University of Maryland College Park, 2009

W. Czaja, E. King, Isotropic shearlet analogs for $L^2(R^k)$ and localization operators, Numerical Functional Analysis and Optimization 33 (2012), no. 7-9, pp. 872–905

W. Czaja, E. King, Shearlet analogs of $L^2(R^d)$, Math. Nachr. 287 (2014), no. 89, pp. 903–916

Reproducing functions

- We are interested in reproducing formulas, which hold for all $f \in L^2(\mathbb{R}^d)$,

$$f = \int_H \langle f, \mu_e(h)\phi \rangle \mu_e(h)\phi dh,$$

  where H is a Lie subgroup of a particular Lie group, μ_e is a representation of that group, ϕ is a suitable window in $L^2(\mathbb{R}^d)$, and the equality is interpreted weakly.
- Define the metaplectic representation:

$$\mu(A_{t,\ell,y})f(x) = t^{k/4} e^{-i\pi \langle B_y x, x \rangle} f(t^{1/2} S_{-\ell/2} x).$$

Norbert Wiener Center
for Harmonic Analysis and Applications

Calderón admissibility condition

Theorem (Calderón formula, with E. King)

The following equality holds

$$\|f\|^2_{L^2(\mathbb{R}^k)} = \int_{(TDS)_k} |\langle f, \mu(A_{t,\ell,y})\phi\rangle|^2 \frac{dt}{t^{k+1}} dy d\ell$$

for all $f \in L^2(\mathbb{R}^k)$, if and only if

$$2^{-k} = \int_{\mathbb{R}^k_+} |\phi(y)|^2 \frac{dy}{y_k^{2k}} = \int_{\mathbb{R}^k_+} |\phi(-y)|^2 \frac{dy}{y_k^{2k}} \qquad (1)$$

$$0 = \int_{\mathbb{R}^k_+} \overline{\phi}(y)\phi(-y) \frac{dy}{y_k^{2k}}. \qquad (2)$$

- The case $k = 1$ was proven by DeMari and Nowak.

Norbert Wiener Center for Harmonic Analysis and Applications

Building reproducing functions

Theorem (with E. King)

Let $f : \mathbb{R} \to \mathbb{R}$ *be supported in some interval* $[0, b]$, $b > 0$ *and satisfy* $\int f^2(x)dx = \frac{1}{4}$. *For* $a > 0$, *define*

$$\phi(x) = x\left(f(x-a) - f(-x+a+2b) + f(x+a+b) + f(-x-a-b)\right).$$

Then ϕ *is a reproducing function for* $(\mathrm{TDS})_1$.

- $f = \mathbb{1}_{[0,\frac{1}{4}]}$
- $f = \frac{1}{\sqrt{\pi}} \cos \cdot \mathbb{1}_{[0,\frac{\pi}{2}]}$
- $f \in C_c^\infty(\mathbb{R})$ has support in $[0, b]$ and is scaled so that $\int f^2 = \frac{1}{4}$, then the resulting ϕ will also lie in C_c^∞

Norbert Wiener Center for Harmonic Analysis and Applications

Uniqueness of generalization

- Dahlke, Steidl, and Teschke introduced another k-dimensional shearlet transform.
- This construction does not yield a reproducing subgroup of $Sp(k, \mathbb{R})$.
- The only pseudo-TDS collection which is a reproducing subgroup of $\mathrm{Sp}(k, \mathbb{R})$ and has a representation onto the operators $T_y D_{t^{-1}(S_\ell^t)}$ is $(\mathrm{TDS})_k$.

Shearlets are not perfect

- Shearlet algorithms result in frames with high redundancy.
- There does not exist any compactly supported MRA shearlet with a desirable level of regularity (Houska, 2009), e.g., Hölder continuous in e_2 with exponent $\beta > 0.5$.
- Most common implementations of shearlets involve separable generating functions..

Composite Dilation Wavelets

- We return to the origin of wavelets: translations and dilations.
- Shearlets descend from the idea of composite wavelets as introduced by Guido Weiss.
- For any $c \in GL_n(\mathbb{R})$ we define the dilation operator $D_c : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ as $D_c f(x) = |\det c|^{-1/2} f(c^{-1}x)$.
- Let $A, B \subset GL_n(\mathbb{R})$ and $\mathcal{L} \subset \mathbb{R}^n$ be a full rank lattice, then the set $\{\psi^1, \ldots \psi^L\} \subset L^2(\mathbb{R}^n)$ forms a *Composite Dilation Wavelet* (CDW) system if the collection

$$\mathcal{A}_{AB\mathcal{L}}(\psi) = \{D_a D_b T_k \psi^i : a \in A, b \in B, k \in \mathcal{L}, 1 \leq i \leq L\}$$

forms a normalized tight frame.

Norbert Wiener Center for Harmonic Analysis and Applications

- Under the assumption that B is a finite group with $B(\mathcal{L}) \subset \mathcal{L}$ and $|\det b| = 1$, for $b \in B$, Manning (2012) proposed to group the dilations B and the lattice $\mathcal{L}$ together into a single group of shifts $\Gamma = B\mathcal{L}$.

B. Maning, Composite Multiresolution Analysis Wavelets, Ph.D. Thesis, Washington University in St. Louis, 2012

- For any $\gamma \in \Gamma$, we define the shift operator $L_\gamma : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ as $L_\gamma f(x) = f(\gamma^{-1}(x))$.
- Given $a \in GL_n(\mathbb{R})$, the set $\{\psi^1, \ldots, \psi^L\}$ forms a CDW if the system

$$\mathcal{A}_{a\Gamma}(\psi) = \{D_{a^j} L_\gamma \psi^i : j \in \mathbb{Z}, \gamma \in \Gamma, 1 \leq i \leq L\}$$

forms a tight frame.
- CDW are usual wavelets except the commutative group of translates $\mathbb{Z}^n$ is replaced with the non-commutative group of shifts Γ.

Norbert Wiener Center for Harmonic Analysis and Applications

Composite Dilation Wavelets

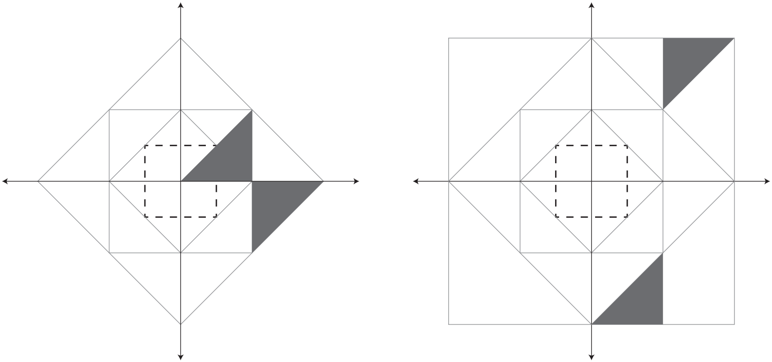


Figure : Example of a composite dilation wavelet. The picture on the left is the Fourier transform of the scaling function. The picture on the right is the Fourier transform of the wavelet.

Variational methods with directional multiscale representations

- Ginzburg-Landau energy functionals have been used by A. Bertozzi and her collaborators in a number of image processing applications that involve variational techniques.

  J. Dobrosotskaya and A. Bertozzi, "Analysis of the wavelet Ginzburg-Landau energy in image applications with edges," SIAM J. Imaging Sci., 2013, Vol. 6 (1), pp. 698–729.

- Wavelet-modified GL energies reduce blurring effects of the TV functionals, increase robustness, and sharpen edges.
- Isotropic wavelets are unable to take advantage of directional content. Anisotropy of wavelet bases and wavelet generator functions leads to biased analysis of singularities along different directions.

Shearlet Ginzburg-Landau energy

- Let
$$\psi_{i,j,t} = 2^{3i/2}\psi(S_1^j A_2^i x - t),$$
for $x \in \mathbb{R}^2$ and $\hat{\psi}(\xi) = \hat{\psi}_1(\xi_1)\hat{\psi}_2(\xi_2/\xi_1)$.
- Shearlet Besov seminorm is
$$\|u\|_S^2 = \sum_{i=0}^{\infty} 2^{2i} \sum_{j\in\mathbb{Z}} \int (|\langle u, \psi_{i,j,t}\rangle|^2 + |\langle u^*, \psi_{i,j,t}\rangle|^2)\, dt,$$
where $u^*(x,y) = u(y,x)$.
- Shearlet Ginzburg-Landau energy is
$$SGL_\alpha(u) = \frac{\alpha}{2}\|u\|_S^2 + \frac{1}{4\alpha}\int ((u^2(x) - 1)^2\, dx.$$

Norbert Wiener Center
for Harmonic Analysis and Applications

Composite Wavelet Ginzburg-Landau energy

- Let
$$\psi_{i,\gamma} = D_{a^{-i}} L_\gamma \psi,$$
for $i \in \mathbb{Z}$ and $\gamma \in \Gamma$.
- Composite Wavelet Besov seminorm is
$$\|u\|_{CW}^2 = \sum_{i=0}^{\infty} |\det a|^i \sum_{\gamma \in \Gamma} |\langle u, \psi_{i,\gamma} \rangle|^2.$$
- Composite Wavelet Ginzburg-Landau energy is
$$CWGL_\alpha(u) = \frac{\alpha}{2} \|u\|_{CW}^2 + \frac{1}{4\alpha} \int ((u^2(x) - 1)^2 \, dx.$$

Norbert Wiener Center
for Harmonic Analysis and Applications

Applications to data recovery

- In many imaging applications variational methods are based on minimizing an energy consisting of two parts: regularizing and forcing terms.
- DGL energy (D = S or CW) plays the role of a regularizer, while the forcing term is expressed as the L_2 norm between the minimizer u and the known image f on the known domain:

$$E(u) = DGL(u) + \frac{\mu}{2}\|u - f\|^2_{L^2(\Omega)}.$$

- We recover the complete image as the minimizer of the modified DGL functional.

Norbert Wiener Center
for Harmonic Analysis and Applications

Applications to data recovery

- The minimizer is a stable state solution of the respective gradient descent equation:

$$u_t = \alpha \Delta_D u - \frac{1}{\epsilon} W'(u) - \mu \chi_\Omega (u - f)$$

Here χ_Ω is the characteristic function of the known domain, $W(x) = (x^2 - 1)^2$, and

$$\Delta_{CW} u = - \sum_{i=0}^{\infty} |\det a|^i \sum_{\gamma \in \Gamma} \langle u, \psi_{i,\gamma} \rangle \psi_{i,\gamma},$$

or

$$\Delta_S(u) = - \sum_{i=0}^{\infty} 2^{2^i} u_i - \sum_{i=0}^{\infty} 2^{2^i} ((u^*)_i)^*.$$

Norbert Wiener Center for Harmonic Analysis and Applications

Shearlet inpainting

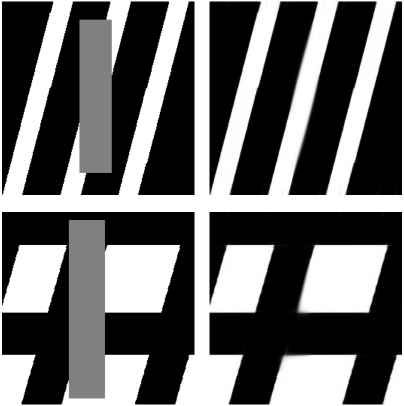


Figure : Results of anisotropic shearlet inpainting which preserves the directional distribution of the input data, as an illustration of the need to know the local and global directions in image processing.

W. Czaja, J. Dobrosotskaya, B. Manning, Shearlet representations for reconstruction of missing data, Proc. SPIE Vol. 8750 (2013), no. 8750-2

Composite Dilation Wavelet Inpainting

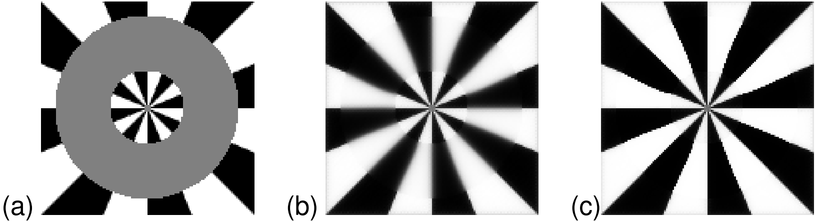


Figure : (a) Original image with gray area missing, (b) image reconstructed via minimizing CWGL after 100 iterations, $\epsilon = 1/24$, $\mu = 1600$; (c) post-processed images (b)

Composite Dilation Wavelets vs Shearlets

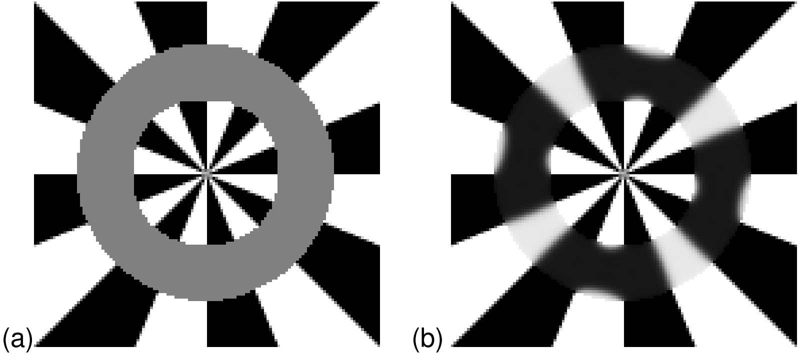


Figure : (a) Original image with gray area missing, (b) Shearlet inpainting does not show the required anisotropy, post-processing does not help

Composite Dilation Wavelets may fail too

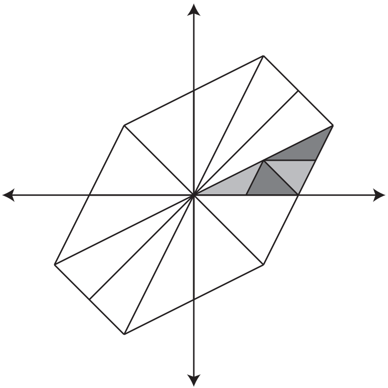


Figure : Another example of a composite dilation wavelet. The figure above describes the Haar scaling function and wavelet for the twelve element group B. The scaling function is supported and constant on both the light and dark shaded areas. There are four wavelet functions. Each wavelet function is supported on both the light and dark shaded areas and is constant on each of the smaller triangles.

Composite Dilation Wavelets may fail too

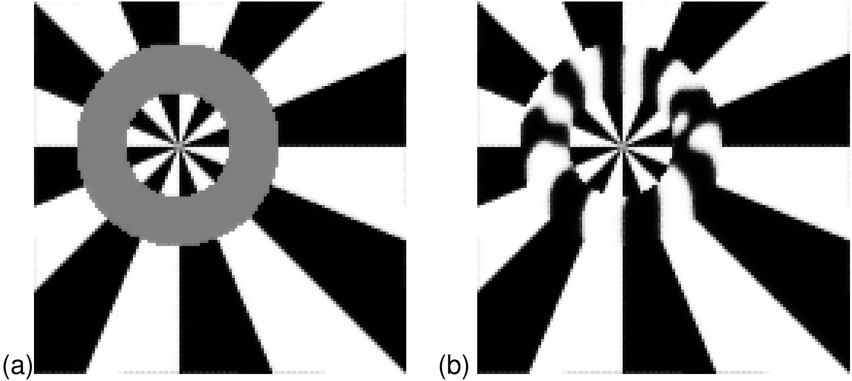


Figure : (a) Original image with a missing annulus is shown in gray, (b) the output of the inpainting simulation fails to reproduce our earlier results.

Applications to superresolution

- Jointly with researchers from NGA, we exploited applications of directional methods for image super-resolution.

E. H. Bosch,et al., “Multiscale and multidirectional tight frames for image analysis,” Proc. SPIE, Vol. 8750, 2013.

(a) Original Image

(b) Bicubic Interpolation

(c) Tight Frame Superresolution

Figure : Image taken from the MUUFL HSI-LIDAR dataset, courtesy of P. Gader (UFL) and A. Zare (UM). Spatial subset of a false-RGB combination comparing directional tight frame scheme to bicubic interpolation for doubling the resolution in each direction.

Summary and Conclusions

- We have covered some of the multiscale directional representations systems in use today.
- These systems come equipped with good approximation properties and have fast implementations.
- The difficulties arise in higher dimensions, where the computational complexity increases.
- Additionally, for more complex data structures, the a priori notion of direction may not be well defined.
- Finally, determining the set of directions without the reference to specific data is limiting. Next, we shall take a look at data dependent representations.