Single Image Superresolution with Directional Representations

James M. Murphy

Department of Mathematics & Information Initiative at Duke
Duke University

February 9, 2016

This is joint work with Wojciech Czaja and Daniel Weinberg

What is Superresolution?

- **Superresolution** is the process of increasing the resolution of a signal, without introducing artifacts and while preserving important signal details.
- It can be understood as a stand alone problem, or as a recovery problem.
- It can be approached both mathematically and algorithmically, with different goals.
- This talk discusses the latter approach, while taking motivation from theoretical signal processing results.

Superresolution in action

- Algorithmic superresolution is found in scientific applications and technology present in everyday life.
- Any type of image resizing feature, such as zoom, will employ a superresolution algorithm once the true resolution is exceeded.
- Bringing images to higher resolution is significant in fields where image data may contain crucial detail features, such as medical imaging and remote sensing.
- Comparing images of different resolutions, perhaps captured by different sensors, also requires superresolution.

Mathematical Formulation

- We can consider the problem of *image* superresolution as an inverse problem: we aim to recover a signal $f \in \mathbb{R}^d$ given measurements

$$y = \mathcal{L}(f) + \mathcal{N},$$

  where $\mathcal{L}$ is an operator and $\mathcal{N}$ denotes some kind of noise. For example, $\mathcal{L}$ may be a downsampling operator and $\mathcal{N}$ Gaussian noise.

- One can develop **interpolators** that recover f from y.

- These can be linear or non-linear; non-linear interpolators that incorporate directionality typically perform well. How to incorporate this directionality best is not completely understood.

- This method is *performance-driven*: few theoretical guarantees exist for superresolution cast in this light.

Mathematical Formulation

- One way to get a foothold mathematically is to restrict the class of signals under analysis.
- For example, Candès and Fernandez-Granda studied the case where signals are one-dimensional and of the form

$$f = \sum_j a_j \delta_{t_j}, \ \{t_j\}_j \subset [0, 1],$$

and the measurements are low-frequency Fourier coefficients:

$$y(k) = \sum_j a_j e^{-2\pi i k t_j}, |k| \leq K.$$

- Benedetto and Li have recently studied the generalization to measures other than sums of Diracs.

Mathematical Formulation

- In this restricted model, a condition relating the separation of the $\{t_j\}_j$ and K can be developed to prove that f can be recovered exactly via the convex optimization problem

$$f = \underset{\mathcal{F}_{2K+1} g = y}{\operatorname{argmin}} \|g\|_{TV}.$$

- This is great, but this is a *continuous* model, whereas practical image processing is done for discrete images. Moreover, the model class of sums of Dirac measures (C,F-G), or singular measures (B+L), is potentially limiting.

Engineering Sparsity

- Sparsity plays a role in algorithmic superresolution.
- At a high level, a signal is represented in a basis in which it is sparse. Intelligent interpolation can then be done based on this sparse representation.
- One approach to incorporating the power of sparsity into the superresolution problem is via *sparse mixing estimators (SME)*. This technique was developed by Mallat and Yu (2010), leading to a state-of-the-art superresolution algorithm.

Overview of SME superresolution

- In essence, the SME superresolution algorithm proceeds by decomposing a measurement y according to a (potentially) redundant frame in which it has a sparsely representation.
- For a set of angles $\{\theta\}$, directional interpolatiors U_θ are applied to blocks of frame coefficients that have a low degree of directional regularity in the direction θ.
- Sparsity is key for the application of these directional interpolators, since a signal is more likely to have low directional regularity on blocks if it is sparsely represented.

Overview of SME Superresolution

- The local directional information contained in the image is captured by studying block sparsity in an appropriate chosen basis, via directional regularity.
- This information is then exploited through directional interpolation, in order to increase the resolution of the image in a way that respects the geometry of the image.
- In order for any of this to work, a frame in which the images under consideration may be sparsely represented must be chosen. This is where harmonic analysis plays a significant role.

Harmonic Analysis for Superresolution

- Wavelets have been used for superresolution since at least the early 2000s.
- The separation of high and low frequency information that wavelets offer is crucial in efficiently recovering the high frequency information that is often lost when downsampling, or that must be generated when zooming in on an image.
- Wavelets have a variety of fast implementations, including filterbank methods, which make them a particularly useful tool for large images or batch problems.

Discrete Wavelet Decompositions

- We consider discrete wavelet frames in the continuous setting.
- Let $f \in L^2(\mathbb{R}^2)$, and let ψ be a *wavelet function*. Then f may be decomposed in the following manner, with convergence in the $L^2(\mathbb{R}^2)$ norm:

$$f = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}^2} \langle f, \psi_{m,n} \rangle \psi_{m,n},$$

where $\psi_{m,n}(x) := |\det A|^{\frac{m}{2}} \psi(A^m x - n), \ A \in GL_2(\mathbb{R})$. A typical choice for A is the dyadic isotropic matrix

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Wavelet Theory

- Wavelets decompose a signal with respect to scale and translation.
- This allows an efficient separation of high and low frequency information, which often provides a highly compressible representation of an image.
- Of course, the above formulation does not make sense for discrete, real signals. Many numerical implementations of wavelets and similar decompositions exist, including the discrete wavelet transform and steerable filterbanks. These often have different computational complexity and geometric invariance properties, but are spiritually similar.

Applications of Wavelets

- Wavelets revolutionized the fields of image compression, fusion, and registration.
- They have been implemented in a variety of ways for image superresolution.
- Wavelet methods can be implemented with fast, efficient numerical algorithms in your favorite language.

Drawbacks of Wavelets

- Wavelets are good for one dimensional jump discontinuities, but are poor in dimensions 2 or more.
- This is a major weakness, since one of the most widely-lauded applications of wavelet methods is image analysis, which is two-dimensional at its simplest.
- Basically, discontinuities destroy the sparsity of wavelets for general images; on must restrict to a class of smooth, slowly varying images to get excellent sparsity properties for a wavelet basis.

Drawbacks of Wavelets

- In higher dimensions, singularities have a *directional character*, but wavelets are fundamentally *isotropic*. This limits wavelets' effectiveness for resolving key aspects of images, such as edges. (Dahlke, De Mari, Grohs, Labate, 2015)

- What is needed are decomposition systems that are *anisotropic*, taking directionality into account.

- This weakness of wavelets partially motivated the development of *geometric multiresolution methods*, such as curvelets, ridgelets, and shearlets.

Shearlets

- Let $f \in L^2(\mathbb{R}^2)$ and ψ be a *shearlet function*. Then f may be decomposed in the following manner, with convergence in the $L^2(\mathbb{R}^2)$ norm:

$$f = \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} \sum_{m\in\mathbb{Z}^2} \langle f, \psi_{j,k,m}\rangle \psi_{j,k,m}.$$

- Here,
  - $\psi_{j,k,m}(x) := 2^{\frac{3j}{4}} \psi(S_k A_{2^j} x - m)$.
  - $A_a = \begin{pmatrix} a & 0 \\ 0 & a^{\frac{1}{2}} \end{pmatrix}$, $S_k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$.
- Note that A has been replaced with A_a, which is no longer isotropic; this will allow our new analyzing functions to be more pronounced in a particular direction.
- The new matrix S_k, a shearing matrix, lets us select the direction.
- As a becomes larger, the direction selected by S_k will be emphasized to a proportionally greater degree.

Shearlets

- Shearlets have several numerical implementations; the most popular are the fast finite shearlet transform (FFST) and Shearlab.
- These run with roughly the same computational complexity as the fast Fourier transform.
- We prefer the FFST, for its ease of use. It is less flexible than Shearlab, however.

Applications of Shearlets

- Shearlets have been applied to many of the same problems in image processing as wavelets.
- They have seen particular success in denoising, edge extraction, and image registration.
- Mathematically, their anisotropic structure has led to their use in the geometric classification of singularities.

Shearlet Optimality

- One of the theoretical benefits of shearlets is their optimality for representing a certain class of functions.

Definition

The set of *cartoon-like images* in $\mathbb{R}^2$ is

$$\mathcal{E} := \{f \mid f = f_0 + \chi_B f_1,\ f_i \in \mathcal{C}^2([0,1]^2), \|f_i\|_{\mathcal{C}^2} \le 1,\ B \subset [0,1]^2,\ \partial B \in \mathcal{C}^2([0,1])\}.$$

- The space of cartoon-like images is an attempt at a quantitative definition of signals that represent images. That is, although images are discrete, if we are to consider only continuous signals, then $\mathcal{E}$ purports to model the class of signals corresponding to images.

Shearlet Optimality

- Shearlets are known to be theoretically near-optimal for $\mathcal{E}$ over *all reasonable representation systems*.
- That is, elements of $\mathcal{E}$ may be written with near-optimally few shearlet coefficients, when compared to the number of coefficients required by other representation systems.
- From a practical standpoint, this suggest shearlets should be superior to wavelet methods in sparsity-driven problems, *when cartoon-like images are the objects of analysis*.
- This is of course a rather serious caveat, since real data do not typically fall into such tightly restricted settings.

Other sparsity models

- Shearlets are well-adapted for the cartoon-like regime.
- The wave atoms of Demanet and Ying are well-adapted for textures, as modelled by images with a high degree of oscillation.
- Recently, Donoho and Kutyniok have studied *joint frames of wavelets and shearlets*, building on Donoho's earlier work on *morphological component analysis*. This method aspires to capture textures with wavelets and edges with shearlets.
- This more complicated regime is interesting, but not analyzed in the present work.

Algorithm Overview

1. Decompose image I to be superresolved into a frame.
2. Perform directional interpolation on blocks of frame coefficients, depending on the directional regularity of the block. This increases the resolution of the image.
3. Apply inverse transform to modified frame coefficients, to acquire a superresolved image I.

Overview of SME

- The key insight in Mallat and Yu's basic method is how the directional interpolation is performed.
- It is performed on *blocks of frame coefficients*.
- For an angle θ and frame coefficients c, a measure of directional regularity in the direction θ is computed on a block of coefficients B:

$$R_B c = c|_B - \bar{c}|_B.$$

- Here, $\bar{c}(k, j) =$ average of the k^{th} frame coefficients in B located on the line passing through j, at angle θ, where k runs through all frame coefficients.

SME Overview

- Whether to interpolate the coefficients along the block B in the direction θ depends on whether $\|R_B c\|_2$ is small or large.
- $\|R_B c\|_2$ is the energy of the variation in coefficients, with respect to their average in the direction θ.
- If $\|R_B c\|_2$ is small, there is little difference between the average frame coefficients in the directional θ and each individual frame coefficient, so there is a strong degree of anisotropy in the direction θ.

SME Overview

- Key parameters to be set are what collection of angle $\{\theta\}$ to consider for directional interpolation, and what blocks of frame coefficients to use.
- We consider 20 equally space angles, i.e. $\{\theta\} = \{\frac{k\pi}{10}\}_{k=1}^{20}$. More angles could be included, especially for larger images, to cover significant directional phenomena that are missed by the current set-up.
- We use the block structure proposed by Mallat and Yu: 28 anisotropic rectangles of various lengths and area between 12 and 18.

Block examples

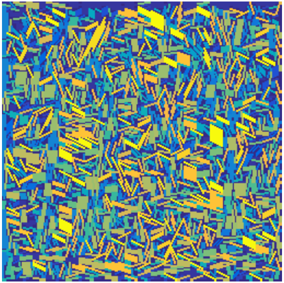

Figure: An image to be superresolved (left), covered with oriented blocks (right), as determined by directional regularity.

Role of Shearlets

- The original scheme proposed a wavelet frame be used to decompose the image I.
- We propose using a shearlet frame, as implemented by the fast finite shearlet transform (FFST).
- We hypothesized the improved anisotropy of the shearlet frame would be well-suited to correctly determining directional regularity.

Quantitative Evaluation

- In order to do a quantitative analysis of our method, we perform the following procedure on a test image I:
  1. Downsample I by a factor of 2 to acquire I_d.
  2. Superresolve I_d by a factor of 2 with a particular superresolution algorithm to acquire $\tilde{I}$.
  3. Compute the PSNR between I and $\tilde{I}$.
- In our case, $PSNR(I, \tilde{I}) = -\log_{10}(\|I - \tilde{I})\|_2)$, so a high PSNR means a low ℓ^2 error.
- In this sense, we consider superresolution as a recovery problem. Otherwise, it is difficult to quantify performance.

Shearlet SME Algorithm Overview

1. Decompose an image I (already downsampled) into a shearlet frame.
2. Compute block decomposition, based on directional regularity.
3. Directionally interpolate in blocks.
4. Apply inverse frame operator to recover a superresolved image $\tilde{I}$.

Previous Shearlet Methods

- We had previously considered a method in which directional interpolation was performed based not on block directional regularity, but simply based on which direction had maximum gradient or maximal shearlet coefficient.
- This information was then applied by blurring a naively upsampled image in the preferred direction.
- Some naive upsampling methods included isotropic interpolation, such as nearest neighbor, bilinear, or bicubic interpolation.
- This method, with naive bicubic upsampling, is included in our experimental analysis.

Synthetic Numerical Experiments

- We first consider images of simple geometric objects that fall clearly in the cartoon-like regime.
- While not a realistic test of the usefulness of a superresolution algorithm, this should give us some sense of whether the fundamental idea, of using shearlets for their optimality for cartoon-like images, is valid.
- We consider three simple synthetic images: an oriented half-plane, a circle, and a parabola.

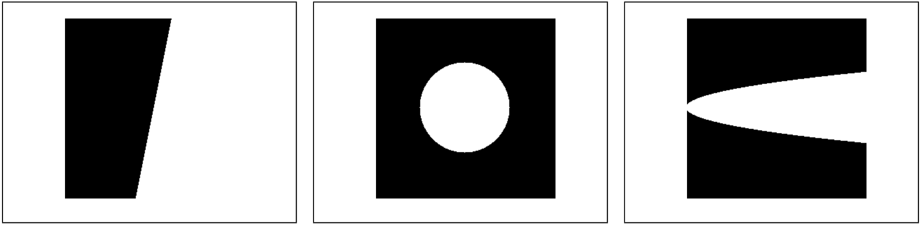

Figure: Synthetic images for superresolution algorithm evaluation, from left to right: half plane, circle, parabola. All image are of size 256×256.

Numerical Results for Synthetic Experiments

Image	PSNR linear	PSNR Shearlet Blur	PSNR SME wavelet	PSNR SME shearlet
plane	29.3764	26.9540	31.5270	**32.7285**
circle	75.9040	74.4775	77.0883	**77.4269**
parabola	26.4058	24.9647	28.2656	**28.9765**

Table: The PSNR values for synthetic experiments.

Outside the cartoon-like regime

- We now consider real test images, from remote sensing and standard image processing datasets.
- These fall outside of any simple model, and are hard to predict in terms of performance.
- In many cases, the improvement is small, even negligible, when compared to the improvements on synthetic data.

Real Data

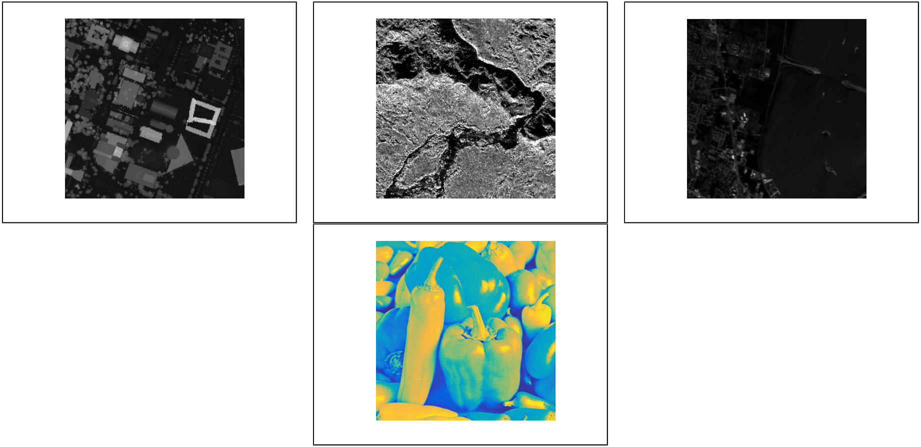

Figure: Real images for superresolution evaluation, from left to right and top to bottom: lidar, synthetic aperture radar, hyperspectral, peppers. All remotely sensed images are of size 256 $\times$ 256, while peppers is of size 512 $\times$ 512.

Real Experimental Results

Image	PSNR linear	PSNR Shearlet Blur	PSNR SME wavelet	PSNR SME shearlet
lidar	30.0101	28.6488	30.7447	**30.8486**
SAR	13.6952	13.5114	14.1380	**14.2112**
KSC	33.1207	31.5616	33.4593	**33.5295**
peppers	31.2676	27.4390	31.9322	**31.9625**

Table: The PSNR values for experiments with real images.

Conclusions

- In general, decomposing in a shearlet frame offers improved performance for the SME method, when compared to a wavelet frame.
- The change is more profound for simple images, perhaps because they fall into the cartoon-like regime, where shearlets are known to perform optimally.
- For real images, which are a mix of sharp edges, textures, and oscillating features, the results are less remarkable.

Future Work

- The work of Demanet and Ying on optimality for textures suggests the use of *wave atoms* for representing textures and oscillating features. This could be incorporated into a superresolution algorithm specifically aimed at textural images.
- The recent work of Donoho and Kutyniok on joint frames of wavelets and shearlets suggests a morphological superresolution regime might provide superior results, compared to wavelets or shearlets alone.
- Indeed, by including both types of frame elements via MCA, edges and textures can be jointly represented in an efficient way. An appropriate interpolation can then be applied separately.