WHAT IS APPLIED HARMONIC ANALYSIS?

GITTA KUTYNIOK

Abstract. The purpose of this paper is to serve as an introduction into the new field of Applied Harmonic Analysis, which is nowadays already one of the major research area in Applied Mathematics.

1. Data, Data, Data,...

Today we are living in a data-drenched world, in which we are challenged to not only provide the methodology to process various different types of data, but – especially as mathematicians – to also analyze the accuracy of such methods and to provide a deeper understanding of the underlying structures. There is a pressing need for those tasks coming from various fields as diverse as air traffic control, digital communications, seismology, medical imaging, and cosmology. As diverse as those fields are the characteristics of the data themselves, where data are usually modeled as functions $f : X \to Y$ or just collections of points in X. Here X can, for instance, be $\mathbb{Z}^n$ or $\mathbb{R}^n$ for arbitrarily large n, a compact subset $\Omega \subset \mathbb{R}^n$, or a general Riemannian manifold, and Y can be similarly diverse concerning its mathematical structure.

Let us take a quick look at some intriguing examples of such modern data.

- *High-dimensional data* can be found in various applications where the most prominent one might be the internet. Here usually the task of search engines is to organize webpages in the widest sense. One common model for such data is to regard a collection of attributes (characterizing words, etc.) for each single webpage as one point in $\mathbb{R}^n$ usually with $n > 10.000$.
- New technology also generates new types of data not encountered before, such as *manifold-valued data*. Here we would like to mention an example coming from air plane control, where a time series of aircraft orientations (pitch, roll, yaw) can be modeled as ranging over $SO(3)$.
- Certainly, also *'common' types of data such as 2-D images appear in new fashions*, for instance, astronomical images from galaxies, medical images from MRI machines, surveillance images from military aircrafts, and so forth, each having its own specifics in, for instance, how the data are measured and what their main features are.

The task which we now face is manifold, starting from how to *measure* data in the most efficient way, especially where time is a factor such as when collecting MRI data. Since data are never 'clean', usually the next task is to *denoise* the data, which requires a suitable model for the noise. If data are missing, we face the task of *inpainting*, which certainly calls

The author would like to thank the Department of Statistics at Stanford University and the Department of Mathematics at Yale University for their hospitality and support during her visits. She acknowledges support by Deutsche Forschungsgemeinschaft (DFG) Heisenberg-Fellowship KU 1446/8-1.

for an accurate model of the class of data we are considering. Then the data need to be *analyzed* depending on the application requirements, which could involve feature detection and extraction, separation of different substructures and so on. Finally, we need to store the data which requires optimal *compression* algorithms. As mentioned before, the general task now consists in not only providing the appropriate methodologies, but even more analyzing their performance and deriving a theoretical understanding of their effectiveness.

2. Applied Harmonic Analysis enters the Stage

Let us now go back in time to the starting point of the analysis of data. Already in the late 18th century Harmonic Analysis provided us with one of the greatest achievements in the processing of signals: the Fourier Transform. This triumph was made almost complete when the Fast Fourier Transform was developed, which is still one of the most fundamental algorithms and can be found in various, even unexpected, applications such as public-key encryption. However, the Fourier Transform itself has a serious disadvantage, since it merely analyzes the global structure of a signal $f \in L^2(\mathbb{R}^n)$, which becomes apparent when observing that a local perturbation of f leads to a change of all Fourier coefficients simultaneously. However, in many signal processing tasks the location of 'peaks' of the signal needs to be detected – which might, for instance, indicate a defect in an engineering process. To phrase it more generally, it is often required to extract local information from a signal even at different resolutions, hence in a multi-scale fashion.

This deficiency led to the birth of the new field of Applied Harmonic Analysis, which is nowadays already one of the major research areas in Applied Mathematics. It exploits not only methods from Harmonic Analysis, but also borrows from areas such as Approximation Theory, Microlocal Analysis, Numerical Mathematics, and Operator Theory, to name a few.

I would now like to provide a brief overview of this exciting research area by first introducing the most fundamental concepts and then elaborating on two topics which illustrate on how those concepts are made alive.

2.1. **Decomposing the Data.** One fundamental concept in Applied Harmonic Analysis is the decomposition of data/signals using representation systems with prescribed properties for a given class of mathematical objects. Such tools offer novel ways to decompose/break down and modify/transform various such classes. Given a closed subset $\mathcal{C}$ of a Hilbert space $\mathcal{H}$, say, the basic idea consists in the construction of a representation system $\{\varphi_i : i \in I\}$ – which for now we assume constitutes an orthonormal basis – so that each signal $s \in \mathcal{C}$ admits a representation

$$s = \sum_{i \in I} \langle s, \varphi_i \rangle \, \varphi_i \tag{1}$$

with the coefficients of 'large' absolute value encoding searched-for features of the signal. This then allows us to analyze the signal s by considering the mapping to its coefficients $\{\langle s, \varphi_i \rangle : i \in I\}$, the so-called *analysis operator*.

Taking a different perspective, we might also design the representation system $\{\varphi_i : i \in I\}$ in such a way that for all elements $s \in \mathcal{C}$ only a 'few' coefficients of the expansion (1) are in fact 'large'. This leads to the notion of a *k-sparse representation*, i.e., k coefficients are non-zero. It becomes intuitively clear that such a representation should be optimal for compression, since only the (few) large coefficients need to be stored. However, it should be mentioned that k-sparse representations are an idealistic model, and we usually encounter the situation of k 'large' coefficients and many small ones, especially in infinite dimensions. Thus often nonlinear approximation rates of the partial reconstruction $\sum_{i \in I_N} \langle s, \varphi_i \rangle \varphi_i$ using the N largest coefficients $\{\langle s, \varphi_i \rangle : i \in I_N\}$ come into play.

2.2. **The Success Story of Wavelets.** One of the first highlights in Applied Harmonic Analysis was the introduction of wavelet systems in the early 1980s, which are systems of time-scale atoms capable of providing local information at different resolutions. Let us now take a look at those systems. A *wavelet* is a function $\psi \in L^2(\mathbb{R})$, which satisfies $\int_{\mathbb{R}} |\hat{\psi}(\xi)|^2/|\xi| \, d\xi = 1$, and the associated *continuous wavelet system* is given by

$$\{a^{-1/2}\psi(a^{-1}(x-t)) : a > 0, t \in \mathbb{R}\}.$$

It can clearly be seen that this collection of functions is perfectly suited to analyze a signal at each location t and at each resolution/scale a via the *Continuous Wavelet Transform*

$$\mathcal{W}_\psi f(a,t) = a^{-1/2} \int_{\mathbb{R}} f(x) \, \overline{\psi(a^{-1}(x-t))} \, dx, \quad a > 0, \, t \in \mathbb{R}.$$

The ability of the Continuous Wavelet Transform to provide local information now becomes evident by the fact that the singular support of a distribution f can be identified to be the closure of the set of points t where $|\mathcal{W}_\psi f(a,t)|$ is of 'slow decay' as $a \to 0$. Thus the Continuous Wavelet Transform resolves the singular support of f, which shows that wavelets are perfect for analyzing isotropic features.

In order to discretize the Continuous Wavelet Transform, the parameters a and t can be sampled as 2^{-j} and $2^{-j}k$, where $j, k \in \mathbb{Z}$. This leads to the well-known *discrete wavelet systems*

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

which indeed have much better nonlinear approximation rates for a smooth function $f \in L^2(\mathbb{R})$ with pointwise discontinuities than the Fourier basis. Let f_N^F be the best partial reconstruction obtained by selecting the N largest terms in the Fourier series, and likewise – assuming that $\{\psi_{j,k}\}$ forms an orthonormal basis – let f_N^W be the best partial reconstruction obtained by selecting the N largest terms from the wavelet coefficients $\{\langle f, \psi_{j,k} \rangle\}$ in absolute value. Then it can be shown that wavelets provide the optimal approximation error rate:

- Fourier approximation error: $\|f - f_N^F\|^2 \leq c \cdot N^{-1}, \quad N \to \infty.$
- Wavelet approximation error: $\|f - f_N^W\|^2 \leq c \cdot N^{-2}, \quad N \to \infty.$

The breakthrough of wavelets was then made complete by two fundamental results in the late 1980s. One milestone was the introduction of a so-called *multiresolution analysis* by Mallat [24] and Meyer [26], which basically allows a stagewise decomposition of a signal into different resolutions and their pairwise orthogonal complements, and, in particular, led to a wavelet decomposition algorithm. The second milestone was the entirely surprising construction of orthonormal bases consisting of compactly supported wavelets with minimum support size for any given number of vanishing moments by Daubechies [11], which was necessary for a fast computation of wavelet coefficients $\{\langle f, \psi_{j,k} \rangle\}$. These two results combined made wavelets a powerful tool for a local analysis of signals with an accompanying Fast

Wavelet Transform. Thus wavelets nowadays exhibit a complete package for both continuous and digital applications, and are therefore widely used. As examples we would like to mention the new compression standard JPEG-2000 – the old standard JPEG used the discrete cosine transform – and the FBI fingerprint database.

From the abundance of literature available on wavelets, we would like to refer the reader to the books by Daubechies [12] and Mallat [25] and the extensive lists of references therein.

2.3. **Gabor Systems: A Different Type of Representation System.** Let us just briefly mention one other widely used class of representation systems called *Gabor systems*, which consist of time-frequency atoms. Given a *window* $g \in L^2(\mathbb{R})$, the associated *continuous Gabor system* is defined as

$$\{g(x-t)e^{2\pi i\omega x} : t, \omega \in \mathbb{R}\},$$

and the *Continuous Gabor Transform* – for obvious reasons also called *Short-Time Fourier Transform* – of some function $f \in L^2(\mathbb{R})$ is

$$\mathcal{G}_g f(t,\omega) = \int_{\mathbb{R}} f(x)\,\overline{g(x-t)}\,e^{-2\pi i\omega x}\,dx, \quad t, \omega \in \mathbb{R}.$$

In fact, assuming that g is compactly supported, this transform computes the Fourier Transform of a local part of f, thus overcoming the problems with the aforementioned globality of the Fourier Transform in a different way than wavelets do.

As for continuous wavelet systems, we can also derive a discrete analog of continuous Gabor systems by sampling t and ω appropriately. For two given parameters $a, b > 0$, the associated *discrete Gabor system* is defined by

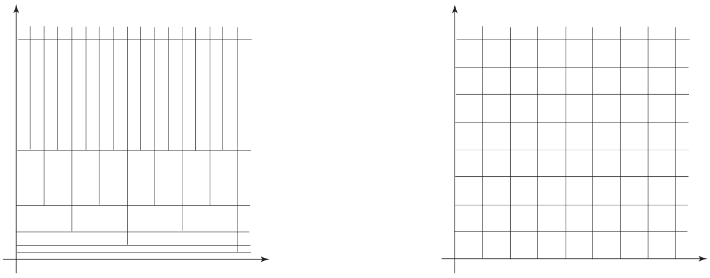
$$\{g(x-am)e^{2\pi ibnx} : m, n \in \mathbb{Z}\}. \tag{3}$$

As probably already expected, these systems are perfect, for instance, for analyzing audio signals with challenges such as extracting different instruments from a recording of a symphony.

For more information on Gabor theory we refer to the book by Gröchenig [18].

2.4. **One Secret behind Representation Systems: Tilings.** Now it is time to mention a basic concept which stands behind the design of representation systems such as wavelet and Gabor systems. Related to their abilities to provide sparse representations for special classes of signals are their *tiling* properties of the associated phase space. Figure 1 illustrates the tiling of the time-scale plane induced by wavelet systems (2) and the tiling of the time-frequency plane induced by Gabor systems (3). Certainly, from a theoretical point of view, we have lots of freedom in choosing a particular tiling following previous work in Harmonic Analysis. But motivated by particular problems, it turns out that these selected tilings are extremely useful. For a study of wavelet and Gabor systems exhibiting different tilings, we would like to refer to [20].

We will introduce yet another tiling – this time in the 2-dimensional situation – in Subsection 3.1 (see also Figure 2).



(a) (b)

FIGURE 1. (a): Dyadic tiling of the time-scale plane induced by wavelet systems. (b): Uniform tiling of the time-frequency plane induced by Gabor systems.

2.5. **The Role of Redundancy.** We have learned that the creation of 'nice' wavelet orthonormal bases is possible, however other representation systems such as Gabor systems encounter problems when we restrict ourselves to orthonormal bases. In fact, the Balian-Low-Theorem states that Gabor systems can *never* form an orthonormal basis if the window g has good time-frequency-localization. This is certainly a drawback, since it is exactly this property which we are interested in. Also recalling our discussion in Subsection 2.1, the design of a sparse representation system $\{\varphi_i : i \in I\}$ might not always be achievable with orthonormal bases, which calls for allowing non-unique expansions; roughly speaking to increase the possibility to have at least one sparse expansion among the set of all expansions. However, we would still like to have a stable expansion in the sense that the norm of a signal $s \in \mathcal{H}$ shall be equivalent to the ℓ_2 norm of its coefficients $\{\langle s, \varphi_i \rangle : i \in I\}$.

The origins of the following definition can actually be traced back to a 1952-publication by Duffin and Schaeffer [17] on non-uniform sampling of band-limited functions. A system $\{\varphi_i : i \in I\}$ in $\mathcal{H}$ is called a *frame*, if there exist constants $0 < A \leq B < \infty$ such that

$$A\|s\|^2 \leq \sum_{i \in I} |\langle s, \varphi_i \rangle|^2 \leq B\|s\|^2 \qquad \text{for all } s \in \mathcal{H}.$$

It is called a *tight frame* if A and B can be chosen to be equal, and a *Parseval frame*, if $A = B = 1$ is feasible. Similarly to orthonormal bases, Parseval frames also satisfy the decomposition formula (1), while providing the possibility of redundancy. A simple example of a Parseval frame are three vectors of the same length in $\mathbb{R}^2$ forming a Mercedes-Benz star, which even led to its name 'Mercedes-Benz frame'.

The power of frame theory has recently again become apparent when this theory was used to attack the 1959 Kadison-Singer Conjecture from Operator Theory and led to various new approaches [6]. A generalization of the concept of frames called *fusion frames* [4] (see also [5]) should also be mentioned, which can be viewed as frame-like collections of low-dimensional subspaces, for instance, suitable for modeling distributed processing applications requiring redundancy.

For further information about frame theory we would like to bring the book by Christensen [9] to the reader's attention.

2.6. **Taming 'Wilder' Redundancy by ℓ_1 Minimization.** We have now made ourselves familiar with wavelet and Gabor systems as representation systems, which can be designed to form an orthonormal basis or a frame. We already discussed that wavelet systems provide sparse representations for isotropic structures like peaks whereas Gabor systems are better adapted to harmonic signals. However, naturally occurring signals are usually a composition of different substructures, hence one specialized representation system alone might not lead to sufficiently sparse representations. This calls for the consideration of *dictionaries*, which are combinations of different representation systems, e.g., of wavelet and Gabor systems. However, we face the problem that such a dictionary sometimes does not constitute a frame.

Now assume $s \in \mathcal{H}$ is a signal from which we know that it has a sparse expansion in the system $\Phi = \{\varphi_i : i \in I\}$ with coefficients c^0. Notice that the associated linear system of equations $s = \Phi c$ is highly underdetermined due to the fact that Φ is an overcomplete system and also lacks any frame property. But let's assume we know that c^0 is unique in the sense that

$$s = \Phi c^0 \quad \text{with } \|c^0\|_0 < \|c\|_0 \text{ for all } c \text{ such that } s = \Phi c,$$

where $\|\cdot\|_0$ denotes the ℓ_0 'norm', i.e., the number of non-zero entries. Obviously, we can then compute c^0 by solving the optimization problem

$$\min_c \|c\|_0 \text{ subject to } s = \Phi c. \tag{4}$$

Unfortunately, this problem is NP-hard, hence computationally not feasible, and it seems hopeless to find the sparsest representation in a computationally efficient way. Then, in 1994, Chen and Donoho [7] (see also [8]) made the fundamental observation that the closest convex norm to the ℓ_0 'norm', which is the ℓ_1 norm, has the tendency to provide sparse representations for severely underdetermined systems. Hence instead of solving (4), the suggestion is to solve the ℓ_1 *minimization problem*

$$\min_c \|c\|_1 \text{ subject to } s = \Phi c, \tag{5}$$

which is now a convex optimization problem and can be solved by linear programming. An avalanche of results shows that under special conditions on the system Φ – usually phrased in terms of the *mutual coherence* $\mu = \max_{i \neq j} |\langle \varphi_i, \varphi_j \rangle| / (\|\varphi_i\|_2 \cdot \|\varphi_j\|_2)$ – and imposing constraints on the sparsity of s, the ℓ_1 minimization problem (5) computes the same solution as (4).

We can also interpret these results in a different way by assuming c^0 is a sparse vector and regarding Φc^0 as taking measurements of c^0 – typically much fewer than the dimension of c^0 – and recovering c^0 via (5). A major breakthrough was recently derived in 2006 by Donoho [13] and by Candés, Romberg, and Tao [3] who showed that using a random Gaussian matrix Φ leads to an optimally small number of necessary measurements. This approach was coined *Compressed Sensing*, and is one of the most rapidly developing topics in Applied Mathematics and Engineering.

The concept of ℓ_1 minimization can also be applied to achieve separation of signals into morphologically different subsignals, which will be extensively used in Subsection 3.2.

For further details on ℓ_1 minimization, we refer the reader to the survey paper by Bruckstein, Donoho, and Elad [1] with an extensive list of references.

3. Applied Harmonic Analysis in Action

Let us now delve into two research problems and their solutions via Applied Harmonic Analysis to illustrate on how the concepts discussed before are made alive. The first problem is focussed on the development of new efficient methodologies, whereas the second one can be regarded as the analysis of an existing methodology by combining various previously discussed concepts.

3.1. **Efficient Representations of Anisotropic Features.** Efficient and economical representation of anisotropic structures is essential in a variety of areas in Applied Mathematics. The anisotropic structure can be given either explicitly, where examples include edges or other directional objects in images, or implicitly such as, for example, in hyperbolic PDEs in the occurrence of shock fronts.

Although the Wavelet Transform outperforms the Fourier Transform for one-dimensional signals, it does not perform equally well in higher dimensions where anisotropic features such as curves begin to play a role. In fact, the Continuous Wavelet Transform is not able to identify the wavefront set of a distribution precisely, neither do discrete wavelet systems reach the optimal approximation error. Given a function f which is in $C^2(\mathbb{R}^2)$ apart from a C^2 curve, the nonlinear approximation errors satisfy:

- Fourier approximation error: $\|f - f_N^F\|^2 \leq c \cdot N^{-1/2}, \ N \to \infty.$
- Wavelet approximation error: $\|f - f_N^W\|^2 \leq c \cdot N^{-1}, \quad N \to \infty.$
- Optimal approximation error: $\|f - f_N^O\|^2 \leq c \cdot N^{-2}, \quad N \to \infty.$

The first representation system whose continuous version precisely identified wavefront sets and whose discrete version provably reached the optimal approximation error (up to a log-factor) were the *curvelets* introduced by Candès and Donoho [2] in 2004. Roughly speaking, curvelets are representation systems dependent on three parameters: scale, location, and direction. However, this system encounters the problem that the discrete version cannot be directly implemented due to the fact that the direction of objects is measured by the angle. Hence curvelets are rotation-based, and intuitively it is obvious that a rotation destroys discrete lattice structures.

Learning from the success of wavelets, we require a directional representation system which provides a complete methodology to deal with anisotropic features in the continuous, discrete, but especially also digital setting to deal with anisotropic feature as efficiently as wavelets deal with isotropic ones. Surprisingly, all these properties are indeed satisfied by the recently introduced *shearlet systems* (see [22] and references therein).

Let us now turn to the discussion of shearlet systems to highlight the main ideas and unearth various cross-references. For each $a > 0$ and $s \in \mathbb{R}$, let A_a denote the *parabolic scaling matrix* and let S_s denote the *shear matrix* of the form

$$A_a = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix} \qquad \text{and} \qquad S_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix},$$

respectively. A *shearlet* is a function $\psi \in L^2(\mathbb{R}^2)$ satisfying $\int_{\mathbb{R}^2} |\hat{\psi}(\xi_1, \xi_2)|^2 / \xi_1^2 \, d\xi = 1$, and the associated *continuous shearlet system* generated by a shearlet ψ is then defined by

$$\{a^{-\frac{3}{4}} \psi(A_a^{-1} S_s^{-1}(x - t)) : a > 0, \ s \in \mathbb{R}, \ t \in \mathbb{R}^2\}.$$

One of the beauties of these systems lies in the fact that the parameter space $\mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^2$ can be equipped with a group multiplication [10], thus allowing us to employ representation theory of locally compact groups. The associated *Continuous Shearlet Transform* of some $f \in L^2(\mathbb{R}^2)$ is then given by

$$\mathcal{SH}_\psi f(a,s,t) = a^{-\frac{3}{4}} \int_{\mathbb{R}^2} f(x)\, \overline{\psi(A_a^{-1}S_s^{-1}(x-t))}\, dx, \quad a > 0,\ s \in \mathbb{R},\ t \in \mathbb{R}^2.$$

Let us now briefly elaborate on the different parameters involved in this transform. Letting the *scaling parameter* a converge to 0 produces needlelike functions due to the properties of parabolic scaling, which by the way was the first time employed in the 1970s by Fefferman. This fact provides the means to approximate anisotropic features with high accuracy. The *shear parameter* s allows to detect different directions now by slope rather than by angle, which is one of the main ideas to overcome the deficiency of curvelets. Finally, the *location parameter* t ensures position sensitivity. It was proven in [21] that the wavefront set of a distribution f can be identified to be the closure of the set of points (t,s) where $|\mathcal{SH}_\psi f(a,s,t)|$ is of 'slow decay' as $a \to 0$. Thus the Continuous Shearlet Transform indeed precisely resolves the wavefront set of f.

In order to approach practical applicability of shearlet systems, the parameters need to be sampled appropriately. Avoiding technical details we just mention that this can be done so that we obtain a tight frame for $L^2(\mathbb{R}^2)$ (see [19]), which tiles the frequency plane as displayed in Figure 2. The reader might want to compare this with the tilings provided by wavelet and Gabor systems in Figure 1, but should keep in mind that here we study a 2-dimensional situation. Discrete shearlet systems also have the desired property to provide

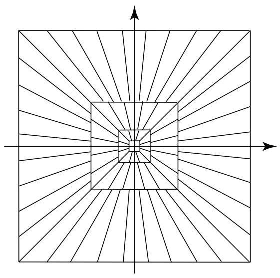


FIGURE 2. Tiling of the frequency domain induced by discrete shearlet systems.

optimally sparse representations for functions which are in $C^2(\mathbb{R}^2)$ apart from a C^2 curve, since it was shown in [15] that shearlets and curvelets have the same sparsity properties.

Now we turn our attention to an implementation of discrete shearlet systems which in contrast to curvelets does not cause many problems mainly due to the advantageous behavior of shear matrices. A software-package called *ShearLab* is currently being developed in [16]. A different approach to derive an implementation in spatial domain – similar to the Fast Wavelet Transform – was recently successfully undertaken in [23] by constructing a *shearlet multiresolution analysis* using specially designed subdivision schemes.

We conclude by observing that the interplay between 'old methods' from Classical Analysis and (Abstract) Harmonic Analysis and 'new methods' introduced by Applied Harmonic

Analysis leads to a highly effective decomposition methodology for 2-dimensional signals containing anisotropic features, which in fact is easily generalizable to higher dimensions.

3.2. **The Geometric Separation Problem.** One of the most fascinating problems these days is the question how the universe will evolve in the future. Scientists attack this problem by looking at the gravitational clustering of matter today. One main task is to extract the different geometrical structures, i.e., stars, filaments, and sheets from astronomical images in order to analyze those separately and derive indicators on the accuracy of models of the universe.

Let us now consider the following situation: We are given an image containing pointlike and curvelike structures (stars and filaments), and face the task to extract two images, one containing just the pointlike phenomena and one containing just the curvelike phenomena. Although this Geometric Separation Problem seems impossible – as there are two unknowns for every datum – suggestive empirical results have been obtained by Jean-Luc Starck and collaborators (see [27] and references therein) by using Applied Harmonic Analysis techniques.

The main idea of the algorithm is to choose a deliberately overcomplete representation made of two tight frames, one is suited to pointlike structures (wavelets) and the other suited to curvelike structures (curvelets or shearlets). The decomposition principle is basically to minimize the ℓ_1 norm of the frame coefficients. This intuitively forces the pointlike objects into the wavelet components of the expansion and the curvelike objects into the curvelet or shearlet part of the expansion.

Now the main task is to analyze the class of such algorithms, study their performance, and derive a profound understanding of the underlying structures. Surprisingly, by exploiting various methods from Applied Harmonic Analysis and neighboring fields, asymptotically nearly-perfect separation of pointlike and curvelike structures can indeed be proven [14].

We now take a closer look at this result to again witness the interplay between real-world problems and the powerful methods Applied Harmonic Analysis provides us with. First, a model for the images under consideration is required, and we model the point and curve singularities by

$$\mathcal{P} = \sum_{i=1}^{P} |x - x_i|^{-1} \qquad \text{and} \qquad \mathcal{C} = \int \delta_{\tau(t)} dt, \;\; \tau \text{ a closed } C^2\text{-curve},$$

respectively. The image we observe is now given by $\mathcal{S} = \mathcal{P} + \mathcal{C}$; however, we only know $\mathcal{S}$. We next perform a subband decomposition of the image $\mathcal{S}$. More precisely, we choose a set of filters $\{F_j\}$ whose frequency supports are adapted to the frequency supports of wavelets and curvelets/shearlets at various scales j, and compute

$$\mathcal{S}_j = \mathcal{P}_j + \mathcal{C}_j, \quad \text{where } \mathcal{P}_j = \mathcal{P} \star F_j \;\text{ and }\; \mathcal{C}_j = \mathcal{C} \star F_j.$$

We then employ the following ℓ_1 decomposition technique with $\{w_\lambda\}$ being a wavelet system and $\{\gamma_\eta\}$ being a curvelet/shearlet system:

$$(W_j, C_j) = \text{ argmin } \|w\|_1 + \|c\|_1 \text{ subject to } \mathcal{S}_j = W_j + C_j, \; w_\lambda = \langle W_j, \psi_\lambda \rangle, \; c_\eta = \langle C_j, \gamma_\eta \rangle.$$

Taking now a viewpoint deriving from Microlocal Analysis, and studying the wavefront sets of $\mathcal{P}$ and $\mathcal{C}$ in phase space, a rigorous analysis of this decomposition strategy can be organized, which finally leads to the main result in [14]:

$$\frac{\|W_j - \mathcal{P}_j\|_2 + \|C_j - \mathcal{C}_j\|_2}{\|\mathcal{S}_j\|_2} \to 0, \qquad j \to \infty.$$

Thus at all sufficiently fine scales, nearly-perfect separation is achieved.

Concluding, the Geometric Separation Problem can be solved by combining 'old methods' from Microlocal Analysis and (Abstract) Harmonic Analysis such as wavefront sets, canonical transformations, oscillatory integral techniques with 'new methods' introduced by Applied Harmonic Analysis such as wavelets, curvelets, shearlets, and ℓ_1 minimization.

References

[1] A. M. Bruckstein, D. L. Donoho, and M. Elad, *From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images,* SIAM Rev., to appear.

[2] E. J. Candès and D. L. Donoho, *New tight frames of curvelets and optimal representations of objects with C^2 singularities*, Comm. Pure Appl. Math. **56** (2004), 219–266.

[3] E. J. Candès, J. K. Romberg, and T. Tao, *Stable signal recovery from incomplete and inaccurate measurements*, Comm. Pure Appl. Math. **59** (2006), 1207–1223.

[4] P. G. Casazza, G. Kutyniok, and S. Li, *Fusion Frames and Distributed Processing*, Appl. Comput. Harmon. Anal. **25** (2008), 114–132.

[5] P. G. Casazza, G. Kutyniok, S. Li, and A. Pezeshki, `www.fusionframe.org`.

[6] P. G. Casazza and J. C. Tremain, *The Kadison-Singer Problem in Mathematics and Engineering*, Proc. Nat. Acad. of Sci. **103** (2006), 2032–2039.

[7] S. S. Chen and D. L. Donoho, *Basis Pursuit,* in Conf. Rec. 28th Asilomar Conf. Signals, Syst., Comput., Pacific Grove, CA, Nov. 1994.

[8] S. S. Chen, D. L. Donoho, and M. A. Saunders, *Atomic decomposition by basis pursuit,* SIAM Rev. **43** (2001), 129–159.

[9] O. Christensen, *An Introduction to Frames and Riesz Bases*, Birkhäuser, Boston, 2003.

[10] S. Dahlke, G. Kutyniok, P. Maass, C. Sagiv, H.-G. Stark, and G. Teschke, *The uncertainty principle associated with the Continuous Shearlet Transform,* Int. J. Wavelets Multiresolut. Inf. Process. **6** (2008), 157–181.

[11] I. Daubechies, *Orthonormal bases of compactly supported wavelets*, Comm. Pure Appl. Math. **41** (1988), 909–996.

[12] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, 1992.

[13] D. L. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory **52** (2006), 1289–1306.

[14] D. L. Donoho and G. Kutyniok, *Microlocal Analysis of the Geometric Separation Problem*, preprint.

[15] D. L. Donoho and G. Kutyniok, *Sparsity Equivalence of Anisotropic Decompositions*, preprint.

[16] D. L. Donoho, G. Kutyniok, and M. Shahram, `www.ShearLab.org`.

[17] R. J. Duffin and A. C. Schaeffer, *A class of nonharmonic Fourier series*, Trans. Amer. Math. Soc. **72** (1952), 341–366.

[18] K. Gröchenig, *Foundations of Time-Frequency Analysis*, Birkhäuser, Boston, 2001.

[19] K. Guo, G. Kutyniok, and D. Labate, *Sparse Multidimensional Representations using Anisotropic Dilation and Shear Operators* Wavelets and Splines (Athens, GA, 2005), Nashboro Press, Nashville, TN (2006), 189–201.

[20] G. Kutyniok, *Affine density in wavelet analysis*, Lecture Notes in Mathematics **1914**, Springer-Verlag, Berlin, 2007.

[21] G. Kutyniok and D. Labate, *Resolution of the Wavefront Set using Continuous Shearlets*, Trans. Amer. Math. Soc., to appear.

[22] G. Kutyniok and D. Labate, `www.shearlet.org`.

[23] G. Kutyniok and T. Sauer, *Adaptive Directional Subdivision Schemes and Shearlet Multiresolution Analysis*, preprint.
[24] S. Mallat, *Multiresolution approximations and wavelet orthonormal bases of* $L^2(\mathbb{R})$, Trans. Amer. Math. Soc. **315** (1989), 69–87.
[25] S. Mallat, *A wavelet tour of signal processing*, Academic Press, Inc., San Diego, CA, 1998.
[26] Y. Meyer, *Wavelets and Operators*, Advanced mathematics, Cambridge University Press, 1992.
[27] J.-L. Starck, M. Elad, and D. L. Donoho, *Image decomposition via the combination of sparse representations and a variational approach,* IEEE Trans. Image Proc. **14** (2005), 1570–1582.

Department of Statistics, Stanford University, Stanford, CA 94305, USA
E-mail address: kutyniok@stanford.edu