Shearlets: Theory and Applications

Gitta Kutyniok[1,*], **Wang-Q Lim**[1,**], and **Gabriele Steidl**[2,***]

[1] Institut für Mathematik, Technische Universität Berlin, 10623 Berlin, Germany
[2] Fachbereich Mathematik, Technische Universität Kaiserslautern, 67653 Kaiserslautern, Germany

Received XXXX, revised XXXX, accepted XXXX
Published online XXXX

Key words Applied harmonic analysis, frames, imaging science, inverse problems, multiscale systems, shearlets, sparse approximation, wavelets
MSC (2010) 42C40, 42C15, 65J22, 94A12, 65F22, 65T60, 94A20, 68U10, 90C25, 15B52

Many important problem classes are governed by anisotropic features such as singularities concentrated on lower dimensional embedded manifolds, for instance, edges in images or shock fronts in solutions of transport dominated equations. While the ability to reliably capture and sparsely represent anisotropic structures is obviously the more important the higher the number of spatial variables is, the principal difficulties arise already in two spatial dimensions. Since it was shown that the well-known wavelets are not capable of efficiently encoding such anisotropic features, various directional representation systems were suggested during the last years. Of those, shearlets are the most widely used today due to their optimal sparse approximation properties in combination with their unified treatment of the continuum and digital realm, leading to faithful implementations. This article shall serve as an introduction to and a survey about shearlets.

1 Introduction

Recent advances in modern technology have created a new world of enormous and multidimensional data structures. In medical imaging, seismic imaging, astronomical imaging, computer vision, and video processing among others, the capabilities of modern computers and high-precision measuring devices have generated 2D, 3D, and even higher dimensional data sets of sizes that were infeasible just a few years ago. The need to efficiently handle such diverse types and huge amounts of data initiated an intense study in developing efficient multivariate encoding methodologies in the applied harmonic analysis research community.

1.1 The Applied Harmonic Analysis Approach to Data Processing

The area of applied harmonic analysis approaches problems of data processing in the following way. Given a class of data $\mathcal{C}$ in a Hilbert space $\mathcal{H}$, and a carefully constructed associated

*E-mail: kutyniok@math.tu-berlin.de, Phone: +49 30 314 25758, Fax: +49 30 314 27364
**E-mail: lim@math.tu-berlin.de, Phone: +49 30 314 25751, Fax: +49 30 314 27364
***E-mail: steidl@mathematik.uni-kl.de, Phone: +49 631 205 5320, Fax: +49 631 205 5309

collection of functions $(\psi_\lambda)_{\lambda\in\Lambda} \subset \mathcal{H}$ with Λ being a countable indexing set. On the one hand the data is *decomposed* by

$$\mathcal{C} \ni f \mapsto (\langle f, \psi_\lambda\rangle)_{\lambda\in\Lambda} \tag{1}$$

with the coefficient sequence enabling an analysis of the data f such as detection of important features, e.g., singularities. On the other hand the data is *expanded* as

$$f = \sum_{\lambda\in\Lambda} c(f)_\lambda \psi_\lambda \qquad \text{for all } f \in \mathcal{C} \tag{2}$$

with the coefficient sequence preferably being sparse in the sense of rapid decay to allow efficient encoding of the data f. It is well-known that $c(f)_\lambda = \langle f, \psi_\lambda\rangle$ for all $\lambda \in \Lambda$ in the situation of an orthonormal basis. However, often redundancy is desirable which leads to the notion of a frame.

1.2 Redundancy comes into Play

Frame theory provides a general framework for representation systems $(\psi_\lambda)_{\lambda\in\Lambda}$ to enable redundancy – in the sense of non-unique expansions – yet retaining stability. A system $(\psi_\lambda)_{\lambda\in\Lambda}$ is called a *frame* for $\mathcal{H}$, if there exist constants $0 < A \leq B < \infty$ such that

$$A\|f\|_2^2 \leq \sum_{\lambda\in\Lambda} |\langle f, \psi_\lambda\rangle|^2 \leq B\|f\|_2^2 \qquad \text{for all } f \in \mathcal{H}.$$

In case $A = B = 1$, it is referred to as a *Parsevsal frame*; and in fact, such Parseval frames are the most general systems which allow $c(f)_\lambda = \langle f, \psi_\lambda\rangle$ for all $\lambda \in \Lambda$.

Letting the *frame operator* be defined by

$$S : \mathcal{H} \to \mathcal{H}, \quad f \mapsto \sum_{\lambda\in\Lambda} \langle f, \psi_\lambda\rangle \psi_\lambda,$$

and setting $(\tilde{\psi}_\lambda)_{\lambda\in\Lambda} := (S^{-1}\psi_\lambda)_{\lambda\in\Lambda}$ to be the *canonical dual frame*, we obtain the following formulas for the reconstruction of some $f \in \mathcal{H}$ from the decomposition (1) and for a particular coefficient sequence in the expansion (2):

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

respectively. For further information on frame theory, we refer to [10].

1.3 Wavelets

One of the most well known examples of frames for $L^2(\mathbb{R}^d)$ are wavelet frames [18].

Definition 1.1 For $\psi^1, \ldots, \psi^L \in L^2(\mathbb{R}^d)$, the associated *wavelet system* is defined by

$$\{\psi^\ell_{j,m} = 2^{\frac{dj}{2}} \psi^\ell(2^j \cdot - m) : j \in \mathbb{Z}, m \in \mathbb{Z}^d, \ell = 1, \ldots, L\}.$$

In fact, finitely many generating functions $\psi^1, \ldots, \psi^L$ can be constructed so that the associated wavelet system forms an orthonormal basis (more generally, a frame) for $L^2(\mathbb{R}^d)$. In this case, the functions $\psi^1, \ldots, \psi^L$ are referred to as *wavelets*. The parameter j is typically referred to as *scale* and the parameter m as *position*.

On the more theoretical side, wavelet systems are proven to highly efficiently approximate L^2-functions which are smooth except for finitely many point singularities. On the more application oriented side, wavelets are, for instance, used in the new compression standard JPEG2000 [61].

1.4 From Wavelets to Shearlets

While wavelets can be shown to derive optimal decay rates of the error of best N-term approximation for functions governed by point singularities, it is evident that starting from dimension 2 anisotropic singularities appear such as curve-like singularities. In fact, for instance, in imaging science the discontinuity curves of an image are one of the most important features, in particular, since the human visual system is designed to react strongly to directional features. But also solutions of, for instance, transport dominated equations are governed by shock fronts. In general, one can argue that many important classes of multivariate problems are governed by anisotropic features. Moreover, in high-dimensional data most information is typically contained in lower-dimensional embedded manifolds, thereby also presenting itself as anisotropic features.

Although wavelets are highly efficient in capturing point singularities, they are isotropic basis (more generally, frame) elements due to the isotropic scaling matrix with a dyadic scaling factor 2^j. Therefore, they are only able to deliver a suboptimal approximation rate of such data. The anisotropic structures can also only be distinguished by location and orientation/direction. This indicates that our way of analyzing and representing the data should capture not only location, but also directional information – requiring anisotropically shaped elements – to overcome the drawback of wavelet systems. These two different approximation schemes with isotropic and anisotropic basis elements are illustrated in Figure 1.

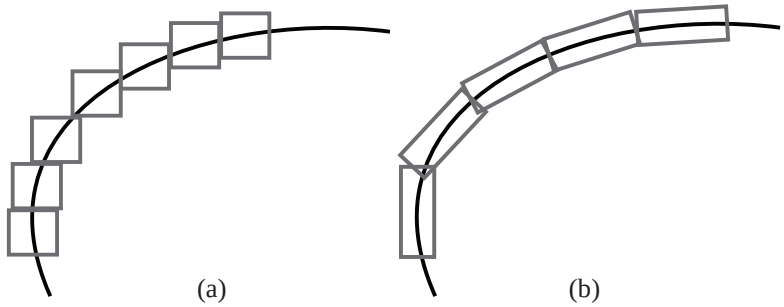


Fig. 1 (a) Approximation of a curve by isotropic basis elements. (b) Approximation of a curve by anisotropic basis elements.

Inspired by this observation, numerous approaches for efficiently representing directional features of multivariate data have been proposed in the area of applied harmonic analysis. A perfunctory list includes: *steerable pyramid* by Simoncelli et al. [60], *directional filter*

banks by Bamberger and Smith [6], *2D directional wavelets* by Antoine et al. [2], *curvelets* by Candès and Donoho [8], *contourlets* by Do and Vetterli [19], *bandelets* by LePennec and Mallat [58], and *shearlets* [35, 52]. Shearlet systems are among the most versatile and successful systems, since they, in particular, satisfy the following list of properties commonly desired for directional representation systems:

(I) Underlying group structure for availability of deep mathematical tools.
(II) Provably optimal sparse approximations of anisotropic features.
(II) Compactly supported analyzing elements for high spatial localization.
(IV) The continuum and digital realm should be treated uniformly.
(V) The associated decomposition should admit a fast implementation.

This article shall serve as an introduction to and a survey about shearlets. For additional information, we refer to the book chapter [45].

1.5 Outline

This survey paper is organized as follows. The continuous and discrete shearlet systems and their associated transforms are introduced and discussed in Sections 2 and 3, respectively. Section 4 is devoted to the optimal sparse approximation properties of shearlets for anisotropic features. A faithful digitalization is presented in Section 5. Finally, in Section 6 several imaging applications and numerical results are discussed.

2 Continuous Shearlet Systems

In the setting of continuous parameter sets, it is well-known that the short time Fourier transform and the wavelet transform are associated with square-integrable representations of the Weyl-Heisenberg group [33] and the affine group [30,31], respectively. Among all directional transforms mentioned in the introduction, the continuous shearlet transform[1] is outstanding, because it stems also from a square-integrable group representation, namely from the so-called *shearlet group*. Therefore, powerful tools of group representation theory can be exploited to study, for instance, reconstruction properties or the resolution of wavefront sets. In the following, we briefly introduce continuous shearlet systems and the associated continuous shearlet transform using this general group theoretical framework.

The main idea in the construction of (continuous) shearlet systems is to select very few generating functions to which scaling operators for different resolution levels, shearing operators as a means to change the orientation, and translation operators to select different positions are applied. The first two operators are based on the *dilation operator*, which for some $M \in \mathbb{R}^{2\times 2}$ is defined by

$$D_M : L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2), \quad (D_M f)(x) \mapsto |\det(M)|^{-1/2} f(M^{-1}x).$$

Selecting M to be a *parabolic scaling matrix* A_a, $a \in \mathbb{R}^* := \mathbb{R} \setminus \{0\}$, or a *shearing matrix* S_s, $s \in \mathbb{R}$, given by

$$A_a = \begin{pmatrix} a & 0 \\ 0 & |a|^{1/2} \end{pmatrix} \quad \text{and} \quad S_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}, \tag{3}$$

[1] 'Continuous' in the sense of continuous parameter sets.

yield the set of scaling and shearing operators, respectively. We remark that the choice of changing the orientation by shearing is superior to rotation due to the fact that in the discrete version the matrices S_k, $k \in \mathbb{Z}$, leave the digital grid $\mathbb{Z}^2$ invariant. Finally, for $t \in \mathbb{R}^2$, the *translation operator* is defined by

$$T_t : L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2), \quad (T_t f)(x) \mapsto f(x-t).$$

This leads to the definition of continuous shearlet systems.

Definition 2.1 For $\psi \in L^2(\mathbb{R}^2)$, the *continuous shearlet system* $SH(\psi)$ is defined by

$$SH(\psi) = \{\psi_{a,s,t} := T_t D_{S_s} D_{A_a} \psi = a^{-\frac{3}{4}} \psi(A_a^{-1} S_s^{-1}(\,\cdot\, - t)) : a \in \mathbb{R}^*, s \in \mathbb{R}, t \in \mathbb{R}^2\}.$$

2.1 The Shearlet Group and the Continuous Shearlet Transform

The just defined continuous shearlet system arises from the *(full) shearlet group* $\mathbb{S} := \mathbb{R}^* \times \mathbb{R} \times \mathbb{R}^2$ endowed with the group operation

$$(a, s, t) \circ (a', s', t') = (aa', s + \sqrt{|a|}s', t + S_s A_a t').$$

This is a locally compact group with left Haar measure $d_\mu(a, s, t) = da/|a|^3 ds dt$. The proofs are given in [11] and in [13] for higher dimensions. For the later, notice that the previous definitions can be naturally extended to $L^2(\mathbb{R}^d)$ by setting $A_{a,\gamma} = \mathrm{diag}(a, \mathrm{sgn}(a)|a|^\gamma I)$ with $\gamma > 0$ and appropriate generalizations of the shearing and translation operator as well as of the group operation. Instead of shearing matrices, triangular Toeplitz matrices were considered in [14]. Recently, it was shown that the shearlet group is isomorphic to the extended Heisenberg group and to a subgroup of the symplectic group [17] which leads to interesting new insights into the structure of the shearlet transform and related function spaces.

Letting the unitary representation $\pi : \mathbb{S} \to \mathcal{U}(L^2(\mathbb{R}^2))$ from $\mathbb{S}$ into the group of unitary operators on $L^2(\mathbb{R}^2)$ be defined by $\pi(a, s, t)\psi := \psi_{a,s,t}$, the continuous shearlet system $SH(\psi)$ can be written as $SH(\psi) = \{\pi(a, s, t)\psi : (a, s, t) \in \mathbb{S}\}$. This mapping is also *square-integrable*, i.e., it is irreducible and there exists a nontrivial *admissible* function $\psi \in L^2(\mathbb{R}^2)$ fulfilling, for all $f \in L^2(\mathbb{R}^2)$, the *admissibility condition*

$$\int_{\mathbb{S}} |\langle f, \pi(a, s, t)\psi\rangle|^2 \, d\mu(a, s, t) < \infty.$$

It can be proven that $\psi \in L^2(\mathbb{R}^2)$ is admissible if and only if

$$\int_{\mathbb{R}^2} \frac{|\hat{\psi}(\xi)|^2}{|\xi_1|^2} \, d\xi < \infty, \tag{4}$$

where $\hat{\psi}$ denotes the Fourier transform of ψ. For a more general approach we refer to [27].

Definition 2.2 A function $\psi \in L^2(\mathbb{R}^2)$ satisfying (4) is called a *shearlet* and the operator $\mathcal{S}_\psi : L^2(\mathbb{R}^2) \to L^2(\mathbb{S})$ given by

$$\mathcal{S}_\psi f(a, s, t) = \langle f, \pi(a, s, t)\psi\rangle$$

is referred to as *continuous shearlet transform*.

Intuitively, $\mathcal{S}_\psi$ maps f to its coefficients $\mathcal{S}_\psi f(a,s,t)$ associated with the *scale* variable $a \in \mathbb{R}^*$, the *orientation* variable $s \in \mathbb{R}$, and the *location* variable $t \in \mathbb{R}^2$. In the general group context, this transform is known as *voice transform*.

Shearlets of both band-limited and compactly supported type exist and are well-studied, see [11,15,42,50]. One prominent example of a shearlet is the *classical shearlet* $\psi \in L^2(\mathbb{R}^2)$, which is a band-limited function introduced in [52] and was the first shearlet extensively studied.

Example 2.3 A *classical shearlet* $\psi \in L^2(\mathbb{R}^2)$ is defined by

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\, \hat{\psi}_2(\tfrac{\xi_2}{\xi_1}),$$

where $\psi_1 \in L^2(\mathbb{R})$ is a wavelet, i.e., it satisfies the discrete Calderón condition given by

$$\sum_{j\in\mathbb{Z}} |\hat{\psi}_1(2^{-j}\xi)|^2 = 1 \quad \text{for a.e. } \xi \in \mathbb{R},$$

with $\hat{\psi}_1 \in C^\infty(\mathbb{R})$ and $\operatorname{supp} \hat{\psi}_1 \subseteq [-\frac{5}{4}, -\frac{1}{4}] \cup [\frac{1}{4}, \frac{5}{4}]$, and $\psi_2 \in L^2(\mathbb{R})$ is a 'bump function', namely

$$\sum_{k=-1}^{1} |\hat{\psi}_2(\xi + k)|^2 = 1 \quad \text{for a.e. } \xi \in [-1,1],$$

satisfying $\hat{\psi}_2 \in C^\infty(\mathbb{R})$ and $\operatorname{supp} \hat{\psi}_2 \subseteq [-1,1]$. Figure 2(a) illustrates the support of the associated Fourier transform $\hat{\psi}$.

Group theoretic considerations lead to the following result about invertibility of the continuous shearlet transform.

Theorem 2.4 ([11]) *Let $\psi \in L^2(\mathbb{R}^2)$ be admissible. Then $\mathcal{S}_\psi$ is an isometry.*

2.2 Shearlet Coorbit Spaces

With a square-integrable group representation at hand, there exists a very natural link to another useful concept, namely the coorbit space theory introduced by Feichtinger and Gröchenig in a series of papers [24–26, 32]. By means of coorbit space theory, it is possible to derive in a canonical way associated scales of smoothness spaces. In this framework, the smoothness of functions is measured by the decay of the associated voice transform. Moreover, by a careful discretization of the representation, it is possible to obtain (Banach) frames for these smoothness spaces.

The continuous shearlet transform does fulfill all necessary conditions for the application of coorbit space theory, which leads to novel canonical smoothness spaces, the *shearlet coorbit spaces* [12] together with their atomic decompositions and (Banach) frames for these spaces. In fact, shearlet coorbit spaces are related to Besov spaces in the sense that certain subspaces of shearlet coorbit spaces can be embedded into homogeneous Besov spaces and traces of shearlet coorbit spaces onto certain hyperplanes are contained in Besov spaces or again in shearlet coorbit spaces, see [15,16]. These subspaces resemble the concept of shearlets on the cone which is considered in the sequel.

2.3 Cone-Adapted Continuous Shearlet Systems and Resolution of the Wavefront Set

Despite the advantages of the pure group theory-based approach, one drawback is the fact that the associated continuous shearlet systems do exhibit a directional bias. For an illustration of the problem, we refer to Figure 2(b).

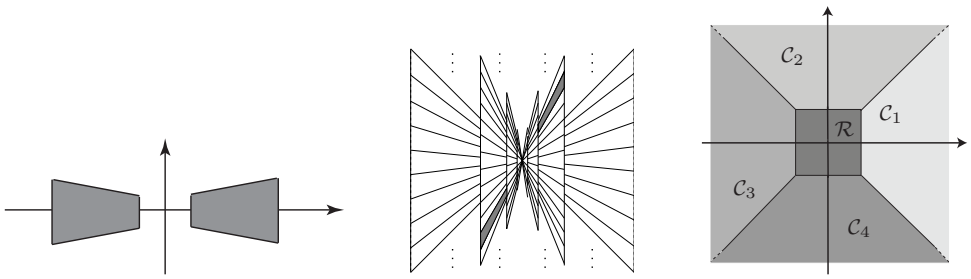


Fig. 2 (a) Support of the Fourier transform of a classical shearlet. (b) Fourier domain support of several elements of the shearlet system, for different values of a and s. (c) The cones $\mathcal{C}_1$ – $\mathcal{C}_4$ and the centered rectangle $\mathcal{R}$ in the frequency domain.

This problem can be resolved by partitioning the Fourier domain into four conic regions and considering the low frequency part separately; see Figure 2(c). This leads to the following variant of continuous shearlet systems with the function ϕ being responsible for the low frequency part as well as ψ and $\tilde{\psi}$ for the horizontal ($\mathcal{C}_1 \cup \mathcal{C}_3$) and vertical ($\mathcal{C}_2 \cup \mathcal{C}_4$) conic regions, respectively.

Definition 2.5 For $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$, the *cone-adapted continuous shearlet system* is defined by $SH(\phi, \psi, \tilde{\psi}) = \Phi(\phi) \cup \Psi(\psi) \cup \tilde{\Psi}(\tilde{\psi})$, where

$$\begin{aligned}
\Phi(\phi) &= \{\phi_t := \phi(\cdot - t) : t \in \mathbb{R}^2\},\\
\Psi(\psi) &= \{\psi_{a,s,t} = a^{-\frac{3}{4}}\psi(A_a^{-1}S_s^{-1}(\cdot - t)) : a \in (0,1],\, |s| \leq 1 + a^{1/2},\, t \in \mathbb{R}^2\},\\
\tilde{\Psi}(\tilde{\psi}) &= \{\tilde{\psi}_{a,s,t} := a^{-\frac{3}{4}}\tilde{\psi}(\tilde{A}_a^{-1}S_s^{-T}(\cdot - t)) : a \in (0,1],\, |s| \leq 1 + a^{1/2},\, t \in \mathbb{R}^2\},
\end{aligned}$$

and $\tilde{A}_a = \text{diag}(a^{1/2}, a)$.

Similar as in the situation of continuous shearlet systems, also for cone-adapted continuous shearlet systems an associated transform can be defined and isometry conditions can be determined (cf. [45]). Moreover, by considering those pairs (s,t), for which $|\langle f, \psi_{a,s,t}\rangle|$ or $|\langle f, \tilde{\psi}_{a,s,t}\rangle|$ does not decay rapidly as $a \to 0$, the wavefront set of a distribution f can be precisely determined [46].

3 Discrete Shearlet Systems

We now turn to discrete shearlet systems, which can be regarded as arising from their continuous counterparts by discretizing the associated set of parameters. Thus, in a similar way we

can derive a discrete shearlet system as well as a cone-adapted variant. Since in applications only the later one is used, we focus in the sequel on this type of shearlet systems.

3.1 2D Shearlets

Recalling the definitions of parabolic scaling and shearing from (3), a common discrete version of those operations are A_{2^j} and S_k with $j, k \in \mathbb{Z}$. Applying the discretization $(a, s, t) \mapsto (2^{-j}, -k2^{-j/2}, A_{2^j}^{-1} S_k^{-1} cm)$ (for $c \in (\mathbb{R}_+)^2$ to add flexibility), which essentially arises from the previously mentioned coorbit theory, to Definition 2.5 yields the following definition.

Definition 3.1 Let $c = (c_1, c_2) \in (\mathbb{R}_+)^2$. For $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ the *cone-adapted discrete shearlet system* $\mathcal{SH}(\phi, \psi, \tilde{\psi}; c) = \Phi(\phi; c_1) \cup \Psi(\psi; c) \cup \tilde{\Psi}(\tilde{\psi}; c)$ is defined by

$$\begin{aligned}
\Phi(\phi; c_1) &= \{\phi_m := \phi(\cdot - m) : m \in c_1 \mathbb{Z}^2\}, \\
\Psi(\psi; c) &= \{\psi_{j,k,m} := 2^{\frac{3}{4}j} \psi(S_k A_{2^j} \cdot - m) : j \geq 0, |k| \leq \lceil 2^{j/2} \rceil, m \in M_c \mathbb{Z}^2\}, \\
\tilde{\Psi}(\tilde{\psi}; c) &= \{\tilde{\psi}_{j,k,m} := 2^{\frac{3}{4}j} \tilde{\psi}(\tilde{S}_k \tilde{A}_{2^j} \cdot - m) : j \geq 0, |k| \leq \lceil 2^{j/2} \rceil, m \in \tilde{M}_c \mathbb{Z}^2\},
\end{aligned}$$

where $M_c = \text{diag}(c_1, c_2)$ and $\tilde{M}_c = \text{diag}(c_2, c_1)$.

If $\mathcal{SH}(\phi, \psi, \tilde{\psi}; c)$ is a frame for $L^2(\mathbb{R}^2)$, we refer to ϕ as a *scaling function* and ψ and $\tilde{\psi}$ as *(discrete) shearlets*. Also notice that 2D shearlets are in the spatial domain of size $2^{-j} \times 2^{-j/2}$, which shows that they will become 'needle-like' (hence more and more anisotropic) as $j \to \infty$.

3.1.1 Band-Limited Shearlets

Band-limited shearlets have the advantage of allowing the construction of Parseval frames for $L^2(\mathbb{R}^2)$. Moreover, some applications such as seismology have a natural band-limited structure which makes the associated shearlet systems preferable.

Constructions of band-limited shearlets were introduced in [35] with refinements contained in [37]. The so-called *classical shearlets*, which we already discussed in Example 2.3, are perhaps the most widely used version. In this case, typically the generator $\tilde{\psi}$ is chosen as $\tilde{\psi}(x_1, x_2) = \psi(x_2, x_1)$. The tiling of Fourier domain which they induce is illustrated in Figure 3.

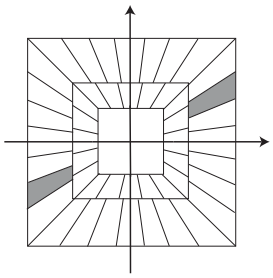


Fig. 3 Tiling of Fourier domain induced by the cone-adapted discrete shearlet system associated with classical shearlets.

For this particular choice, it was proved in [35, Thm. 3] that the associated cone-adapted discrete shearlet system $\mathcal{SH}(\phi,\psi,\tilde{\psi};(1,1))$ with a slight modification of the boundary elements and with an appropriate scaling function ϕ for the centered rectangle forms a Parseval frame for $L^2(\mathbb{R}^2)$.

3.1.2 Compactly Supported Shearlets

Despite the advantageous functional analytic properties of classical, and in general band-limited shearlets often applications seek high spatial localization. This makes compactly supported shearlets desirable. However, it is still not clear whether a cone-adapted discrete shearlet system associated with compactly supported shearlets can be introduced, which forms a Parseval frame for $L^2(\mathbb{R}^2)$. The best known construction forms a frame for $L^2(\mathbb{R}^2)$ with numerically proven ratio of the frame bounds of approximately 4, see [42].

3.1.3 Shearlet Frames

In [42], a general framework was introduced which provides sufficient conditions for a cone-adapted discrete shearlet system to form a frame with theoretical estimates for the associated frame bounds, which we will now describe.

For this, we first define

$$\Omega_0 = \{\xi \in \mathbb{R}^2 : \max\{|\xi_1|,|\xi_2|\} \leq \tfrac{1}{2}\}, \quad \Omega_1 = \{\xi \in \mathbb{R}^2 : \tfrac{1}{2} < |\xi_2| < 1, |\xi_2|/|\xi_1| < 1\},$$

and assume that

$$\operatorname*{ess\,inf}_{\xi\in\Omega_0} |\hat{\phi}(\xi)| > 0 \quad \text{and} \quad \operatorname*{ess\,inf}_{\xi\in\Omega_1} |\hat{\psi}(\xi)| > 0. \tag{5}$$

The above conditions ensure that

$$\operatorname*{ess\,inf}_{\xi\in\mathbb{R}^2} |\hat{\phi}(\xi)|^2 + \sum_{j\geq 0} \sum_{|k|\leq \lceil 2^{j/2}\rceil} \left(|\hat{\psi}(S_k^T A_{2^{-j}}\xi)|^2 + |\hat{\tilde{\psi}}(\tilde{S}_k^T \tilde{A}_{2^{-j}}\xi)|^2\right) > 0$$

where $\tilde{\psi}(x_1,x_2) = \psi(x_2,x_1)$. In fact, this inequality can be used to derive a lower frame bound for $\mathcal{SH}(\phi,\psi,\tilde{\psi},c)$. The following result shows that any functions $\phi,\psi \in L^2(\mathbb{R}^2)$ (and $\tilde{\psi}(x_1,x_2) = \psi(x_2,x_1)$), whose Fourier transform decays fast enough with sufficient vanishing moments and satisfies the lower bound conditions (5) indeed generate a shearlet frame $\mathcal{SH}(\phi,\psi,\tilde{\psi},c)$.

Theorem 3.2 ([42]) *Let $\phi,\psi \in L^2(\mathbb{R}^2)$ be functions such that*

$$\hat{\phi}(\xi_1,\xi_2) \leq C_1 \cdot \min\{1,|\xi_1|^{-\gamma}\} \cdot \min\{1,|\xi_2|^{-\gamma}\} \quad \textit{and}$$

$$|\hat{\psi}(\xi_1,\xi_2)| \leq C_2 \cdot \min\{1,|\xi_1|^{\alpha}\} \cdot \min\{1,|\xi_1|^{-\gamma}\} \cdot \min\{1,|\xi_2|^{-\gamma}\}, \tag{6}$$

for some positive constants $C_1, C_2 < \infty$ and $\alpha > \gamma > 3$. Define $\tilde{\psi}(x_1,x_2) = \psi(x_2,x_1)$ and assume that ϕ,ψ satisfy (5)*. Then, there exists some positive constant c^* such that $\mathcal{SH}(\phi,\psi,\tilde{\psi},c)$ forms a frame for $L^2(\mathbb{R}^2)$ for any $c=(c_1,c_2)$ with $\max\{c_1,c_2\} \leq c^*$.*

Obviously, band-limited shearlets satisfy condition (6). More interestingly, also a large class of compactly supported functions satisfies this condition, and we refer to the constructions in [42].

3.2 3D Shearlets

Continuous shearlet systems in higher dimensions than dimension 2 were first considered in [14] in the pure group theory-based approach. In the following, we now focus as before on a 'cone-adapted variant' in the 3D situation. These so-called pyramid-adapted shearlet systems again treat directions in a uniform way similar as their cone-adapted 2D counterparts. The continuous (band-limited) version of such was first introduced in [36], and it was shown that the location and the local orientation of the boundary set of certain three-dimensional solid regions can be precisely identified by the associated transform.

We now turn to the discrete setting and the definition of pyramid-adapted discrete shearlet systems. These systems will be generated by scaling according to the *paraboloidal scaling matrices* A_{2^j}, $\tilde{A}_{2^j}$ or $\breve{A}_{2^j}$, $j \in \mathbb{Z}$, and directionality will be encoded by the *shear matrices* S_k, $\tilde{S}_k$, or $\breve{S}_k$, $k = (k_1, k_2) \in \mathbb{Z}^2$, defined by

$$A_{2^j} = \text{diag}(2^j, 2^{j/2}, 2^{j/2}),\ \tilde{A}_{2^j} = \text{diag}(2^{j/2}, 2^j, 2^{j/2}),\ \breve{A}_{2^j} = \text{diag}(2^{j/2}, 2^{j/2}, 2^j)$$

and

$$S_k = \begin{pmatrix} 1 & k_1 & k_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ \tilde{S}_k = \begin{pmatrix} 1 & 0 & 0 \\ k_1 & 1 & k_2 \\ 0 & 0 & 1 \end{pmatrix},\ \breve{S}_k = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k_1 & k_2 & 1 \end{pmatrix},$$

respectively. The translation lattices will be defined through the following matrices: $M_c = \text{diag}(c_1, c_2, c_2)$, $\tilde{M}_c = \text{diag}(c_2, c_1, c_2)$, and $\breve{M}_c = \text{diag}(c_2, c_2, c_1)$, where $c_1, c_2 > 0$. Similar as for cone-adapted shearlet systems, we now partition Fourier domain into a rectangular region and six pyramids as shown in Figure 4.

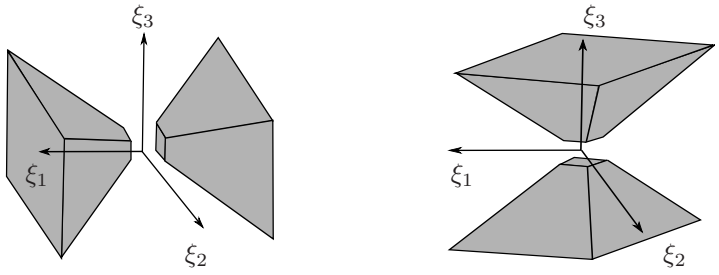


Fig. 4 The partition of Fourier domain by four of the six pyramids.

These considerations are now made precise in the following definition, where each part of the system is responsible for covering one set of pyramids.

Definition 3.3 For $c = (c_1, c_2) \in (\mathbb{R}_+)^2$, the *pyramid-adapted discrete shearlet system* $\mathcal{SH}(\phi, \psi, \tilde{\psi}, \breve{\psi}; c)$ generated by $\phi, \psi, \tilde{\psi}, \breve{\psi} \in L^2(\mathbb{R}^3)$ is defined by

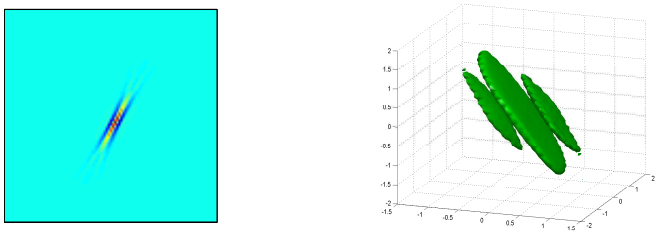
$$\mathcal{SH}(\phi, \psi, \tilde{\psi}, \breve{\psi}; c) = \Phi(\phi; c_1) \cup \Psi(\psi; c) \cup \tilde{\Psi}(\tilde{\psi}; c) \cup \breve{\Psi}(\breve{\psi}; c),$$

where

$$\begin{aligned}
\Phi(\phi; c_1) &= \{\phi_m := \phi(\cdot - m) : m \in c_1\mathbb{Z}^3\},\\
\Psi(\psi; c) &= \{\psi_{j,k,m} := 2^j\psi(S_k A_{2^j} \cdot - m) : j \geq 0, \|k\|_\infty \leq \lceil 2^{j/2}\rceil, m \in M_c\mathbb{Z}^3\},\\
\tilde{\Psi}(\tilde{\psi}; c) &= \{\tilde{\psi}_{j,k,m} := 2^j\tilde{\psi}(\tilde{S}_k \tilde{A}_{2^j} \cdot - m) : j \geq 0, \|k\|_\infty \leq \lceil 2^{j/2}\rceil, m \in \tilde{M}_c\mathbb{Z}^3\},\\
\breve{\Psi}(\breve{\psi}; c) &= \{\breve{\psi}_{j,k,m} := 2^j\breve{\psi}(\breve{S}_k \breve{A}_{2^j} \cdot - m) : j \geq 0, \|k\|_\infty \leq \lceil 2^{j/2}\rceil, m \in \breve{M}_c\mathbb{Z}^3\}.
\end{aligned}$$

The construction of pyramid-adapted shearlet systems $\mathcal{SH}(\phi, \psi, \tilde{\psi}, \breve{\psi}; c)$ runs along the lines of the construction of the 2D cone-adapted shearlet systems described in the previous subsection. Similar as in the 2D case, also compactly supported shearlet frames with controlled frame bounds can be constructed [50]. We also wish to mention that this construction can be easily generalized to even higher dimensions.

Let us finally discuss whether this is the most natural extension to 3D. For this, notice that 3D shearlets are in the spatial domain of size $2^{-j} \times 2^{-j/2} \times 2^{-j/2}$, which shows that they will become 'plate-like' as $j \to \infty$. Figure 5 illustrates some examples of 2D and 3D shearlets in the spatial domain. One could also exploit the scaling matrix $A_{2^j} = \mathrm{diag}\,(2^j, 2^j, 2^{j/2})$ with similar changes for $\tilde{A}_{2^j}$ and $\breve{A}_{2^j}$. This would lead to 'needle-like' shearlet elements instead of the 'plate-like' elements introduced in Definition 3.3. In particular, those shearlets seem well adapted for 1D singularities, while 'plate-like' shearlets seem better suited for 2D singularities. More generally, it is possible to even consider non-paraboloidal scaling matrices of the form $A_j = \mathrm{diag}\,(2^j, 2^{\alpha j}, 2^{\beta j})$ for $0 < \alpha, \beta \leq 1$. One drawback of allowing such general scaling matrices is the lack of fast algorithms for non-dyadic multiscale systems, and in case of the 'needle-like' shearlet construction the lack of frame properties. On the other hand, the parameters α and β allow us to precisely shape the shearlet elements, ranging from very 'plate-like' to very 'needle-like', according to the application at hand, i.e., choosing the shape of the shearlet which is the best *'fit'* for the geometric characteristics of the considered data. We refer to [28, 50] for more details on these general settings.



(a) (b)

Fig. 5 Examples of (a) 2D shearlets and (b) 3D shearlets in the spatial domain.

4 Quest for Optimal Sparse Approximations

As mentioned in the introduction, shearlets were introduced with the goal to provide optimally sparse approximations of anisotropic features. We now first make the notion of 'anisotropic

features' mathematically precise by introducing a mathematical model for such. Let us start with the most basic definition of this class which was also historically first, stated in [20]. We allow ourselves to phrase this definition in the d-dimensional situation for $d = 2, 3$, hence to combine the 2D with the 3D setting from [50].

Definition 4.1 For fixed $\nu > 0$ and $d = 2, 3$, the *class* $\mathcal{E}_\nu^2(\mathbb{R}^d)$ *of cartoon-like functions* is the set of functions $f : \mathbb{R}^d \to \mathbb{C}$ of the form

$$f = f_0 + f_1 \chi_B,$$

where $B \subset [0,1]^d$ and $f_i \in C^2(\mathbb{R}^d)$ with supp $f_i \subset [0,1]^d$ and $\|f_i\|_{C^2} \leq 1$ for each $i = 0, 1$. For dimension $d = 2$, we assume that ∂B is a closed C^2-curve with curvature bounded by ν, and, for $d = 3$, the discontinuity ∂B shall be a closed C^2-surface with principal curvatures bounded by ν.

Donoho then proved the following benchmark result in [20] for $d = 2$; its extension to $d = 3$ was derived in [50]. We wish to mention that in fact the initial statements are more general than the one for frames we present here.

Theorem 4.2 ([20, 50]) *Let $d = 2, 3$, and let $(\psi_\lambda)_{\lambda \in \Lambda}$ be a frame for $L^2(\mathbb{R}^d)$. Then the optimal asymptotic approximation error of $f \in \mathcal{E}^2(\mathbb{R}^d)$ is given by*

$$\|f - f_N\|_2^2 \leq C \cdot N^{-\frac{2}{d-1}} \text{ as } N \to \infty, \text{ where } f_N = \sum_{\lambda \in \Lambda_N} c_\lambda \psi_\lambda$$

is the (non-linear) best N-term approximation and $C > 0$.

Thus a system satisfying this decay conditions can be justifiably called a system which provides *optimally sparse approximations* of cartoon-like functions. The following result shows that shearlets indeed satisfy this benchmark result up to a (negligible) $\log$-factor. We wish to mention that this result is in fact part of a much more general framework, the framework of *parabolic molecules* [34], which provides approximation results for very general systems based on parabolic scaling.

Theorem 4.3 ([47]) *Let $c > 0$, and let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ be compactly supported. Suppose that, in addition, for all $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$, the shearlet ψ satisfies*

(i) $|\hat{\psi}(\xi)| \leq C_1 \cdot \min\{1, |\xi_1|^\alpha\} \cdot \min\{1, |\xi_1|^{-\gamma}\} \cdot \min\{1, |\xi_2|^{-\gamma}\}$ *and*

(ii) $\left|\frac{\partial}{\partial \xi_2} \hat{\psi}(\xi)\right| \leq |h(\xi_1)| \cdot \left(1 + \frac{|\xi_2|}{|\xi_1|}\right)^{-\gamma}$,

where $\alpha > 5$, $\gamma \geq 4$, $h \in L^1(\mathbb{R})$, and C_1 is a constant, and suppose that the shearlet $\tilde{\psi}$ satisfies (i) and (ii) with the roles of ξ_1 and ξ_2 reversed. Further, suppose that $\mathcal{SH}(\phi, \psi, \tilde{\psi}; c)$ forms a frame for $L^2(\mathbb{R}^2)$. Then, for any $\nu > 0$, the shearlet frame $\mathcal{SH}(\phi, \psi, \tilde{\psi}; c)$ provides (almost) optimally sparse approximations of functions $f \in \mathcal{E}_\nu^2(\mathbb{R}^2)$ in the sense that there exists some $C > 0$ such that

$$\|f - f_N\|_2^2 \leq C \cdot N^{-2} \cdot (\log N)^3 \text{ as } N \to \infty,$$

where f_N is the nonlinear N-term approximation obtained by choosing the N largest shearlet coefficients of f.

Similar results were shown for $\mathcal{E}_\nu^2(\mathbb{R}^3)$ using pyramid-adapted discrete shearlets, see [50].

5 Digital Shearlet Transforms

The first implementation for computing shearlet coefficients has been developed in [22]. This numerical algorithm implements the discrete shearlet transform associated with band-limited shearlets. For this, digital shearlet filter coefficients are obtained by approximating the inverse Fourier transform of band-limited shearlets discretized on the 2D pseudo-polar grid where the shear operator S_k is given as a translation.

We now discuss implementation strategies for computing shearlet coefficients associated with the cone-adapted discrete shearlet system based on compactly supported shearlets, as introduced in Section 3. One main focus will be on deriving a digitization which is faithful to the continuum setting. We refer to [55] and [56] for more details on this approach. For this, we will only consider shearlets $\psi_{j,k,m}$ for the horizontal cone, i.e., belonging to $\Psi(\psi, c)$. Notice that the same procedure can be applied to compute the shearlet coefficients for the vertical cone, i.e., those belonging to $\tilde{\Psi}(\tilde{\psi}, c)$, except for switching the order of variables.

We first define a separable shearlet generator $\psi \in L^2(\mathbb{R}^2)$ and an associated scaling function $\phi \in L^2(\mathbb{R}^2)$ as follows. Let $\phi_1 \in L^2(\mathbb{R})$ be a compactly supported 1D scaling function satisfying

$$\phi_1(x_1) = \sum_{n_1 \in \mathbb{Z}} h(n_1)\sqrt{2}\phi_1(2x_1 - n_1)$$

for some appropriately chosen filter h. An associated compactly supported 1D wavelet $\psi_1 \in L^2(\mathbb{R})$ can then be defined by

$$\psi_1(x_1) = \sum_{n_1 \in \mathbb{Z}} g(n_1)\sqrt{2}\phi_1(2x_1 - n_1),$$

where again g is an appropriately chosen filter. Then, the separable shearlet generator ψ is defined to be

$$\psi(x_1, x_2) = \psi_1(x_1)\phi_1(x_2).$$

The filters h and g are required to be chosen so that ψ satisfies a certain decay condition (cf. [42]) to guarantee a stable reconstruction from the shearlet coefficients. With those filters h and g, we define the scaling and wavelet filters h_j and g_j by the Fourier coefficients of the trigonometric polynomials H_j and G_j, $j > 0$, defined by

$$H_j(\xi_1) = \prod_{k=0}^{j-1} \sum_{n_1 \in \mathbb{Z}} h(n_1) e^{-2\pi i 2^k n_1 \xi_1}$$

and

$$G_j(\xi_1) = H_{j-1}(\xi_1)\Big(\sum_{n_1 \in \mathbb{Z}} g(n_1) e^{-2\pi i 2^{j-1} n_1 \xi_1}\Big),$$

respectively.

For the signal $f \in L^2(\mathbb{R}^2)$ to be analyzed, we now assume that, for $J > 0$ fixed, f is of the form

$$f(x) = \sum_{n \in \mathbb{Z}^2} f_J(n) 2^J \phi(2^J x_1 - n_1, 2^J x_2 - n_2), \tag{7}$$

where $\phi(x_1, x_2) = \phi_1(x_1)\phi_1(x_2)$. Note that the scaling coefficients $(f_J(n))_{n\in\mathbb{Z}^2}$ in (7) can be viewed as sample values of f – in fact $f_J(n) \approx f(2^{-J}n)$ with appropriately chosen ϕ. Now aiming towards a faithful digitization of the shearlet coefficients $\langle f, \psi_{j,k,m}\rangle$ for $j = 0, \ldots, J-1$, we first observe that

$$\langle f, \psi_{j,k,m}\rangle = \langle f(S_{k\cdot 2^{-j/2}}(\cdot)), \psi_{j,0,m}(\cdot)\rangle. \tag{8}$$

WLOG we will from now on assume that $j/2$ is integer; otherwise either $\lceil j/2\rceil$ or $\lfloor j/2\rfloor$ would need to be taken. Our observation (8) shows us in fact precisely how to digitize the shearlet coefficients $\langle f, \psi_{j,k,m}\rangle$, namely by applying the discrete separable wavelet transform associated with the anisotropic sampling matrix A_{2^j} to the sheared version of the data $f(S_{k\cdot 2^{-j/2}}(\cdot))$. This however requires – compare the assumed form of f given in (7) – that $f(S_{k\cdot 2^{-j/2}}(\cdot))$ is contained in the scaling space

$$V_J = \{2^J\phi(2^J\cdot - n_1, 2^J\cdot - n_2) : (n_1, n_2) \in \mathbb{Z}^2\}.$$

It is easy to see that, for instance, if the shear parameter $k\cdot 2^{-j/2}$ is non-integer, this is unfortunately not the case. The true reason for this failure is that the shear matrix $S_{k\cdot 2^{-j/2}}$ does *not* preserve the regular grid $2^{-J}\mathbb{Z}^2$ in V_J, i.e.,

$$S_{k\cdot 2^{-j/2}}(\mathbb{Z}^2) \neq \mathbb{Z}^2.$$

In order to resolve this issue, we consider the new scaling space $V^k_{J+j/2,J}$ defined by

$$V^k_{J+j/2,J} = \{2^{J+4/j}\phi_k(2^{J+j/2}\cdot - n_1, 2^J\cdot - n_2) : (n_1, n_2) \in \mathbb{Z}^2\}$$

where $\phi_k(\cdot) = \phi(S_k\cdot)$. We remark that the scaling space $V^k_{J+j/2,J}$ is obtained by refining the regular grid $2^{-J}\mathbb{Z}^2$ along the x_1-axis by a factor of $2^{j/2}$. With this modification, the new grid $2^{-J-j/2}\mathbb{Z}\times 2^{-J}\mathbb{Z}$ is now invariant under the shear operator $S_{k\cdot 2^{-j/2}}$. In fact, we have

$$2^{-J-j/2}\mathbb{Z}\times 2^{-J}\mathbb{Z} = S_{2^{-j/2}k}(2^{-J-j/2}\mathbb{Z}\times 2^{-J}\mathbb{Z}).$$

This allows us to rewrite $f(S_{k\cdot 2^{-j/2}}(\cdot))$ in (8) in the following way.

Lemma 5.1 ([55]) *Let $\uparrow 2^{j/2}$ and $*_1$ denote the 1D upsampling operator by a factor of $2^{j/2}$ and the 1D convolution operator along the x_1-axis, respectively. Then, we obtain*

$$f(S_{2^{-j/2}k}(x)) = \sum_{n\in\mathbb{Z}^2} \tilde{f}_J(S_k n)2^{J+j/4}\phi_k(2^{J+j/2}x_1 - n_1, 2^J x_2 - n_2),$$

where $\tilde{f}_J(n) = ((f_J)_{\uparrow 2^{j/2}} *_1 h_{j/2})(n)$.

The second term to be digitized in (8) is the shearlet $\psi_{j,k,m}$ itself with $k = 0$, for which the following result can be employed.

Lemma 5.2 ([55]) *Retaining the notations and definitions from this section, we obtain*

$$\psi_{j,0,m}(x) = \sum_{d\in\mathbb{Z}^2} g_{J-j}(d_1 - 2^{J-j}m_1)h_{J-j/2}(d_2 - 2^{J-j/2}m_2)2^J\phi(2^J x - d).$$

As already indicated before, we will make use of the discrete separable wavelet transform associated with an anisotropic scaling matrix, which, for $j_1, j_2 > 0$ as well as $c \in \ell(\mathbb{Z}^2)$, we define by

$$W_{j_1,j_2}(c)(n_1,n_2) = \sum_{m\in\mathbb{Z}^2} g_{j_1}(m_1 - 2^{j_1}n_1)h_{j_2}(m_2 - 2^{j_2}n_2)c(m_1,m_2), \quad (n_1,n_2) \in \mathbb{Z}^2.$$

Finally, Lemmata 5.1 and 5.2 yield the following digitilized form of the shearlet coefficients $\langle f, \psi_{j,k,m}\rangle$.

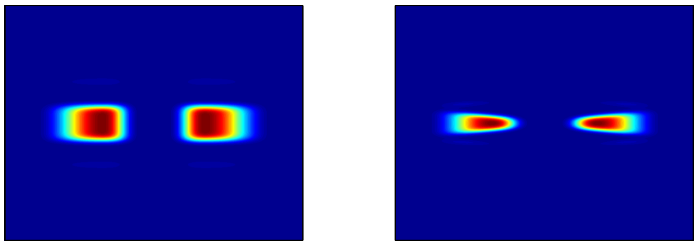
Theorem 5.3 ([55]) *Retaining the notations and definitions from this section, and letting* $\downarrow 2^{j/2}$ *be 1D downsampling by a factor of* $2^{j/2}$ *along the horizontal axis, we obtain*

$$\langle f, \psi_{j,k,m}\rangle = W_{J-j,J-j/2}\Big(\Big((\tilde{f}_J(S_k\cdot) * \Phi_k) *_1 \overline{h}_{j/2}\Big)_{\downarrow 2^{j/2}}\Big)(m),$$

where $\Phi_k(n) = \langle \phi_k(\cdot), \phi(\cdot - n)\rangle$ *for* $n \in \mathbb{Z}^2$, *and* $\overline{h}_{j/2}(n_1) = h_{j/2}(-n_1)$.

Theorem 5.3 shows that for any $f \in L^2(\mathbb{R}^2)$ of the form (7), the shearlet coefficients $\langle f, \psi_{j,k,m}\rangle$ can be faithfully computed in the digital domain, similar to wavelet theory. For more details as well as a digitalization of the inverse shearlet transform, we refer to [51,55,56].

In fact, this framework can be extended to allow a nonseparable shearlet ψ, which significantly improves the stability and directional selectivity of the shearlet transform [56]. Figure 6 shows the refined essential support of a nonseparable shearlet in Fourier domain as compared to a separable shearlet. The Matlab software package called ShearLab[2] provides code for the shearlet transforms associated with those separable and nonseparable shearlets.



(a) (b)

Fig. 6 (a) Separable shearlet. (b) Nonseparable shearlet.

Finally, the 3D shearlet transform can also be implemented faithfully by following the ideas outlined in this section. For details we refer to [51, 56] and for an implementation to ShearLab.

6 Applications

Recently, sparse and redundant representation modeling has received much attention due to its efficiency in imaging sciences, see, e.g., [1, 23]. The facts that shearlets provide (almost)

[2] `http://www.shearlab.org`

optimally sparse approximations of anisotropic features and that most imaging data and, in particular, natural images are governed by such features – the neurons in the human visual cortex are most sensitive to such edge-like structures – now allow the utilization of shearlets in this general methodological approach. In the sequel, we will discuss some exemplary applications. For other applications, we refer the interested reader to the book chapter [38].

For shearlet-based algorithms currently three toolboxes are available, which also contain code for some of the presented applications: (i) Local Shearlet Toolbox[3] was the first shearlet transform implementation for band-limited shearlets, for details see [22]. (ii) ShearLab[4] consists of three different implementations: The band-limited shearlet transform implemented on both the pseudo-polar grid and the Cartesian grid, as well as a compactly supported shearlet transform, for details see [51, 55]. (iii) Fast Finite Shearlet Transform (FFST)[5] provides fully finite and translation invariant shearlet implementation with band-limited shearlets, for details see [41].

6.1 Image Separation

The task of separating images $f \in L^2(\mathbb{R}^2)$ into morphologically distinct components has recently attracted a lot of attention, see, e.g., [3,4,62]. One particular instance of this problem is the separation of curve-like and point-like objects, which arises in several applications. In neurobiological imaging, it would, for instance, be desirable to separate spines (point-like objects) from dendrites (curve-like objects) in order to analyze them independently aiming to detect characteristics of Alzheimer disease. In astronomical imaging, astronomers would often like to separate stars from filaments for further analysis. Other applications appear in material sciences.

The mathematically precise problem can be formulated as follows: Given a point-like object $u_1^0 \in L^2(\mathbb{R}^2)$ and a curve-like object $u_2^0 \in L^2(\mathbb{R}^2)$, recover u_1^0 and u_2^0 from knowledge of $f = u_1^0 + u_2^0$. This is obviously an ill-posed inverse problem.

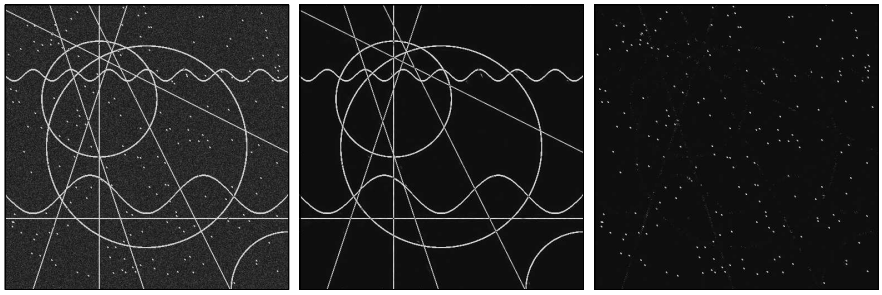
The key idea now is to approach this problem by compressed sensing methodologies, see, e.g., [44]. This mainly consists in utilizing two sparsifying systems each for one of the components followed by solving an ℓ_1 minimization problem, since the ℓ_1 norm promotes sparsity. In our situation, we choose wavelets and shearlets for the point-like and curve-like objects, respectively. The separation is then performed by solving the following convex optimization problem:

$$\arg\min_{u_1,u_2} \|(\langle u_1, \phi_\lambda\rangle)_{\lambda\in\Lambda}\|_1 + \|(\langle u_2, \psi_\gamma\rangle)_{\gamma\in\Gamma}\|_1 \quad \text{subject to} \quad \|u_1+u_2-f\|_2 \leq \sigma, \tag{9}$$

where $(\phi_\lambda)_{\lambda\in\Lambda}$ and $(\psi_\gamma)_{\gamma\in\Gamma}$ are wavelets and shearlets, respectively, and σ is the noise level. Figure 7 shows the numerical results of the separation scheme (9) with wavelets and shearlets for an artificial image (cf. [48]). Note that curve-like objects are well separated by shearlets even for regions where a curve has high curvature. A theoretical analysis of this separation scheme can be found in [21].

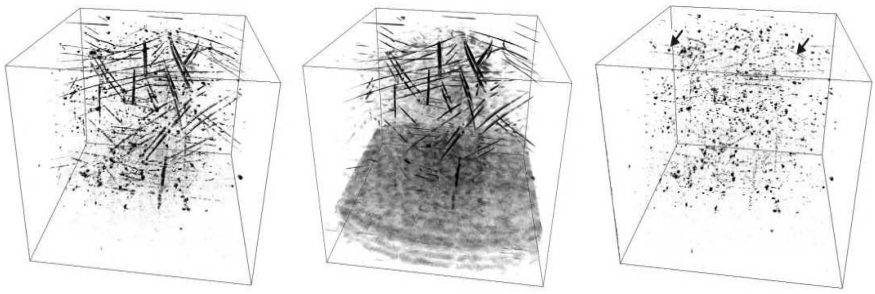
[3] http://www.math.uh.edu/~dlabate/software.html
[4] http://www.shearlab.org
[5] http://www.mathematik.uni-kl.de/imagepro/software/

In [39] the separation approach was applied in material science, more precisely to polymer composites filled with glass fibers that are used in many industrial branches. Their characteristics strongly depend on the orientation of the fibers (lines), hence the orientation distribution of the fibers is of great importance. On the other hand also the points—they represent defects in the material—have a strong influence on the performance of the material. For a further analysis, the both components have to be analyzed separately. The three-dimensional images are obtained via full-field optical coherence microscopy. Figure 8 shows the decomposition results. For technical details, we refer to [39].



(a) (b) (c)

Fig. 7 Image separation: (a) Noisy input image (curves+points). (b) Separated image (curves). (c) Separated image (points).

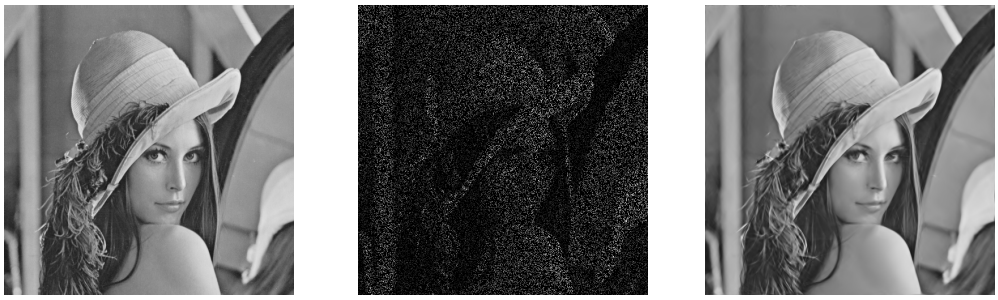


(a) (b) (c)

Fig. 8 Image separation of an FF-OCT image of polymer composites filled with glass fibers: (a) Original image (b) Lines/Fibers (c) Points/Defects. (cf. [39])

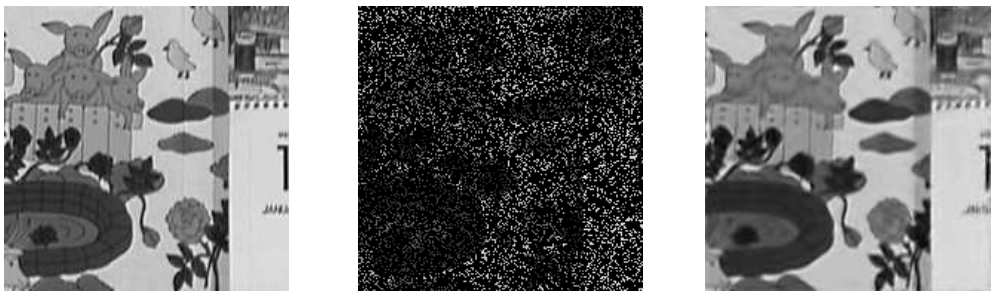
6.2 Image/Video Inpainting

Inpainting is the process of reconstructing lost or deteriorated parts of images and videos. The term ‘inpainting’ first appeared in [7], but earlier work on disocclusions was done, e.g., in [9, 57]. In this respect also interpolation, approximation, and extrapolation problems may be viewed as inpainting problems. Inpainting is a very active field of research which was tackled by various approaches. For a nice overview, we refer to [29].



(a) (b) (c)

Fig. 9 Image inpainting: (a) Original image (512×512). (b) Corrupted image with 80% missing pixels. (c) Inpainted image.



(a) (b) (c)

Fig. 10 Video inpainting: (a) Original frame of a $(192 \times 192 \times 192)$ video sequence. (b) Corrupted frame with 80% missing pixels. (c) Inpainted frame.

The mathematically precise problem can be formulated as follows: Given an image domain $\Omega_0 = \{1, \ldots, m\} \times \{1, \ldots, n\}$, the inpainting region $\Omega \subset \Omega_0$ is the subset where the pixel values $f(i,j)$, $(i,j) \in \Omega$ are unknown. The (noiseless) inpainting problem consists of finding a function u on Ω_0 from data f given on $\bar{\Omega} = \Omega_0 \backslash \Omega$ such that u is a suitable extension of f to Ω_0. With $(\psi_\gamma)_{\gamma \in \Gamma}$ a shearlet system, the shearlet based model reads as follows:

$$\arg\min_u \|(\langle u, \psi_\lambda \rangle)_{\lambda \in \Lambda}\|_1 \quad \text{subject to} \quad u|_{\bar{\Omega}} = f|_{\bar{\Omega}}.$$

Again, the backbone of this approach comes from the theory of compressed sensing.

In [28, 43], a shearlet based inpainting scheme has been theoretically analyzed. Figures 9 and 10 show inpainting results for an image and, using 3D shearlets, a video. For more extensive test results as well as comparisons with other state-of-the-art methods we refer to [51].

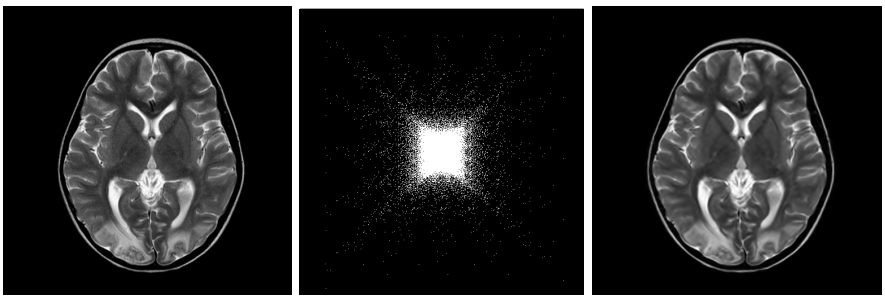
6.3 Sparse Sampling of Fourier Data

One major problem of imaging procedures in clinical research and health care such as Magnetic Resonance Imaging (MRI) or X-ray Computed Tomography (X-ray CT) is the very slow data acquisition. In those applications, the data acquisition process can also be modeled by taking Fourier measurements. One main challenge is therefore to significantly reduce the number of Fourier samples without degrading the quality of image reconstruction by developing novel sampling and reconstruction schemes. For this, we consider

$$\arg\min_u \|(\langle u, \psi_\lambda \rangle)_{\lambda \in \Lambda}\|_1 \quad \text{subject to} \quad \|P_\Omega F u - f\|_2 \leq \sigma$$

where $(\psi_\gamma)_{\gamma \in \Lambda}$ is a shearlet system, Ω is a set of sampling points in Fourier domain, P_Ω is the corresponding sampling operator, F is the Fourier operator, and σ is the noise level.

One major task is to design a sampling set Ω, which allows a successful image reconstruction using only a few number of Fourier samples. Figure 11 shows the reconstructed image by a shearlet-based sparse sampling scheme using only 5% Fourier coefficients of the original image. A detailed description and a theoretical analysis of this shearlet-based sparse sampling scheme using methods from compressed sensing can be found in [49].



(a) (b) (c)

Fig. 11 Fourier Sampling: (a) Original image (512×512). (b) Fourier mask for the subsampling operator P_Ω containing 5% sampling points in Fourier domain. (c) Reconstructed image.

6.4 Image Segmentation

Segmentation is a fundamental task in image processing which plays a role in many preprocessing steps. Recently, convex relaxation methods for image multi-labeling were addressed by several authors [5,53,54,59,63] using a total variation regularizer. For texture segmentation containing coherent curves, shearlet regularization showed superior results [40].

More precisely, to segment a given color image $f = (f(i,j))_{i,j=1}^{m,n} \in \mathbb{R}^{m,n}$ into L segments, we assign a codebook $c = (c_1, \ldots, c_L)$ to the image which contains the central colors c_k, $k = 1, \ldots, L$ of the segments (this codebook can be updated during the iterative segmentation process). For each pixel (i,j) we are searching for the probabilities $p_k(i,j)$ such that $f(i,j)$ lies in segment k. In other words, $p(i,j) = (p_1(i,j), \ldots, p_L(i,j))$ is in the probability

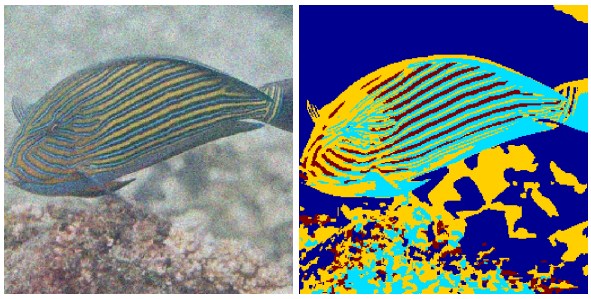
simplex

$$\triangle = \{\pi \in \mathbb{R}^L : \sum_{k=1}^{L} \pi_k = 1,\ \pi_1, \dots, \pi_L \geq 0\}.$$

Then $f(i,j)$ is assigned to those segment with the highest probability. Let $s_k := (s_k(i,j))_{i,j=1}^{m,n}$ $:= (\|f(i,j) - c_k\|_2^2)_{i,j=1}^{m,n}$. We then find p_k, $k = 1, \dots, L$ as the solution of

$$\arg\min_{p} \sum_{k=1}^{L} (\langle s_k, p_k \rangle + \alpha \|(\langle p_k, \psi_\lambda \rangle)_{\lambda \in \Lambda}\|_1) \quad \text{subject to} \quad p(i,j) \in \triangle \quad \forall (i,j)$$

with a suitable regularization parameter $\alpha > 0$. Figure 12 shows an example of a segmented image using the described method. For details, we refer to [40].



(a) (b)

Fig. 12 Image segmentation: (a) Noisy image of a clown-surgeon fish, (b) Segmented image with four segments.

References

[1] M. Aharon and M. Elad, SIAM J. Imaging Sci. **1**, 228–247 (2008).
[2] J. P. Antoine, P. Carrette, R. Murenzi, and B. Piette, Signal Process. **31**, 241–272 (1993).
[3] J.-F. Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle, J. Math. Imaging Vision **22**, 71–88 (2005).
[4] J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher, J. Math. Imaging Vision **67**, 111–136 (2006).
[5] E. Bae, J. Yuan, and X.-C. Tai, Int. J. Comput. Vision **92**, 112–129 (2011).
[6] R. H. Bamberger and M. J. T. Smith, IEEE Trans. Signal Process. **40**, 882–893 (1992).
[7] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, Proceedings of SIGGRAPH 2000 (New Orleans, USA, 2000), pp. 417–424.
[8] E. J. Candès and D. L. Donoho, Comm. Pure Appl. Math. **56**, 216–266 (2004).
[9] V. Caselles, J.-M. Morel, and C. Sbert, IEEE Trans. Image Process. **7**, 376–386 (1998).
[10] O. Christensen, An Introduction to Frames and Riesz Bases (Birkhäuser, Boston, 2003).
[11] S. Dahlke, G. Kutyniok, P. Maass, C. Sagiv, H.-G. Stark, and G. Teschke, Int. J. Wavelets Multiresolut. Inf. Process. **6**, 157–181 (2008).
[12] S. Dahlke, G. Kutyniok, G. Steidl, and G. Teschke, Appl. Comput. Harmon. Anal. **27**, 195–214 (2009).

[13] S. Dahlke, G. Steidl, and G. Teschke, J. Fourier Anal. Appl. **16**, 340–354 (2010).
[14] S. Dahlke and G. Teschke, in: Group Theory: Classes, Representation and Connections, and Applications, edited by C. W. Danellis (Nova Publishers, 2010), chap. 6.
[15] S. Dahlke, G. Steidl, and G. Teschke, J. Fourier Anal. Appl. **17**, 1232–1255 (2011).
[16] S. Dahlke, S. Häuser, G. Steidl, and G. Teschke, Monatsh. Math. **169**, 15–32 (2013).
[17] S. Dahlke, F. de Mari, E. de Vito, S. Häuser, G. Steidl, and G. Teschke, ArXiv-Preprint 1404.4545 (2014).
[18] I. Daubechies, Ten Lectures on Wavelets (SIAM, Philadelphia, PA, 1992).
[19] M. N. Do and M. Vetterli, IEEE Trans. Image Process. **14**, 2091–2106 (2005).
[20] D. L. Donoho, Constr. Approx. **17**, 353–382 (2001).
[21] D. L. Donoho and G. Kutyniok, Comm. Pure Appl. Math. **66**, 1–47 (2013).
[22] G. Easley, D. Labate, and W.-Q Lim, Appl. Comput. Harmon. Anal. **25**, 25–46 (2008).
[23] M. Elad, IEEE Signal Process. Lett. **19**, 922–928 (2012).
[24] H. G. Feichtinger and K. Gröchenig, J. Funct. Anal. **86**, 307–340 (1989).
[25] H. G. Feichtinger and K. Gröchenig, Monatsh. Math. **108**, 129–148 (1989).
[26] H. G. Feichtinger and K. Gröchenig, in: Wavelets and Their Applications, edited by G. Beylkin, R. R. Coifman, and I. Daubechies (Jones and Bartlett, Boston, 1992), pp. 353–376.
[27] H. Führ and R. Raisi-Tousi, ArXiv preprint 1407.0824 (2014).
[28] M. Genzel and G. Kutyniok, SIAM J. Imaging Sci. to appear.
[29] C. Guillemot and O. Le Meur, IEEE Signal Process. Mag. **31**, 127–144 (2014).
[30] A. Grossmann, J. Morlet, and T. Paul, J. Math. Phys. **26**, 2473–2479 (1985).
[31] A. Grossmann, J. Morlet, and T. Paul, Ann. L'Institute H. Poincaré Math. Phys. **45**, 293–309 (1986).
[32] K. Gröchenig, Monatsh. Math. **112**, 1–42 (1991).
[33] K. Gröchenig, Foundations of Time–Frequency Analysis (Birkhäuser, Boston, 2001).
[34] P. Grohs and G. Kutyniok, Found. Comput. Math. **14**, 299-337 (2014).
[35] K. Guo, G. Kutyniok, and D. Labate, in: Wavelets and Splines (Athens, GA, 2005), edited by C. Guanrong and M.-J. Lai (Nashboro Press, Nashville, TN, 2006), pp. 189–201.
[36] K. Guo and D. Labate, Appl. Comput. Harmon. Anal. **30**, 231–242 (2011).
[37] K. Guo and D. Labate, Math. Model. Nat. Phenom. **8**, 82–105 (2013).
[38] G. Easley and D. Labate, in: Shearlets: Multiscale Analysis for Multivariate Data, edited by G. Kutyniok and D. Labate (Birkhäuser, Boston, 2012), pp. 283–325.
[39] B. Heise, S. E. Schausberger, S. Häuser, B. Plank, D. Salaberger, E. Leiss-Holzinger, and D. Stifter, Opt. Fiber Technol. **5**, 403–410 (2012).
[40] S. Häuser and G. Steidl, Int. J. Comput. Math. **90**, 62–81 (2013).
[41] S. Häuser, ArXiv Preprint 1202.1773 (2012).
[42] P. Kittipoom, G. Kutyniok, and W.-Q Lim, Constr. Approx. **35**, 21–72 (2012).
[43] E. J. King, G. Kutyniok, and X. Zhuang, J. Math. Imaging Vision **48**, 205–234 (2014).
[44] G. Kutyniok, GAMM Mitteilungen **36**, 79-101 (2013).
[45] G. Kutyniok and D. Labate, in: Shearlets: Multiscale Analysis for Multivariate Data, edited by G. Kutyniok and D. Labate (Birkhäuser, Boston, 2012), pp. 1–38.
[46] G. Kutyniok and D. Labate, Trans. Amer. Math. Soc. **361**, 2719–2754 (2009).
[47] G. Kutyniok and W.-Q Lim, J. Approx. Theory **163**, 1564–1589 (2011).
[48] G. Kutyniok and W.-Q Lim, in: Curves and Surfaces (Avignon, France, 2010), edited by J.-D. Boissonnat, P. Chenin, A. Cohen, C. Gout, T. Lyche, M.-L. Mazure, and L. L. Schumaker, Lecture Notes in Computer Science **6920** (Spinger, 2012), pp. 416–430.
[49] G. Kutyniok and W.-Q Lim, in preparation.
[50] G. Kutyniok, J. Lemvig, and W.-Q Lim, SIAM J. Math. Anal. **44**, 2962–3017 (2012).
[51] G. Kutyniok, W.-Q Lim, and R. Reisenhofer, ArXiv Preprint 1402.5670 (2014).
[52] D. Labate, W.-Q Lim, G. Kutyniok, and G. Weiss, Wavelets XI, SPIE Proc. **5914** (SPIE, Bellingham, WA, 2005), pp. 254–262.
[53] J. Lellmann, F. Becker, and Ch. Schnörr, Proceedings of 12th International Conference on Computer Vision (Kyoto, Japan, 2009), pp. 646–653.

[54] J. Lellmann and Ch. Schnörr, SIAM J. Imaging Sci. **4**, 1049–1096 (2011).
[55] W.-Q Lim, IEEE Trans. Image Process. **19**, 1166–1180 (2010).
[56] W.-Q Lim, IEEE Trans. Image Process. **22**, 2056–2065 (2013).
[57] S. Masnou and J.-M. Morel, Proceedings of the 1998 IEEE International Conference on Image Processing (Chicago, IL, 1998), pp. 259–263.
[58] E. L. Pennec and S. Mallat, IEEE Trans. Image Process. **14**, 423–438 (2005).
[59] T. Pock, A. Chambolle, D. Cremers, and H. Bischof, IEEE Conference on Computer Vision and Pattern Recognition (Miami, FL, 2009), pp. 810–817.
[60] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, IEEE Trans. Inform. Theory, **38**, 587–607 (1992).
[61] A. Skodras, C. Christopoulos, and T. Ebrahimi, IEEE Signal Processing Mag. **18**, 36–58 (2001).
[62] J.-L Starck, M. Elad, and D. Donoho, IEEE Trans. Image Proc. **14**, 1570–1582 (2005).
[63] C. Zach, D. Gallup, J. M. Frahm, and M. Niethammer, Proceedings of Vision, Modeling, and Visualization (Konstanz, 2008), pp. 243–252.