Digital Shearlet Transform: Rational Design of a Digital Parabolic Scaling Algorithm

David L. Donoho[a], Gitta Kutyniok[b], and Morteza Shahram[a]

[a]Department of Statistics, Stanford University,
Stanford, CA 94305, USA;

[b]Institute of Mathematics, University of Osnabrück,
49069 Osnabrück, Germany

ABSTRACT

Directional representation systems based on parabolic scaling such as curvelets, contourlets, or shearlets have become increasingly popular over the last years due to novel applications in, for instance, seismology or astronomy. In this paper, focussing on shearlets as one example, we present an associated digital transform which is implemented following a strict rational design in the sense that a digital theory is provided, which is an exact digitalization of the theory for continuous data. The implementation then follows this digital theory stricto sensu. Further, we introduce quantitative quality measures computed by test algorithms for general digital parabolic scaling algorithms. This testing framework aims at providing (a) a means to tune parameters of particular algorithms and (b) a standardized platform for comparing different digital parabolic scaling algorithms. We then present and discuss the values of the test measures for our algorithm and provide a recommended choice of parameters.

Keywords: Density Compensation, Digital Shearlet Theory, Fast Pseudo-Polar Transform, Monte Carlo Estimates, Parabolic Scaling, Pseudo-Polar Grid, Reproducible Research, Test Measures, Shearlet Transform, Testing Environment

1. INTRODUCTION

Algorithms based on a directional representation system such as curvelets,[2–4] contourlets,[5] or shearlets[11, 12, 14, 15] have seen an advent in recent years, since novel applications require sparse decompositions of anisotropic features. Examples are, for instance, the missing data problem in seismology[13] when missing sensors cause incomplete measurements or the geometric separation problem in astronomy when images of galaxies require separated analyses of stars, filaments, and sheets.[6, 17] Transforms based on parabolic scaling provide similar sparsity, hence are in a sense interchangeably suitable with respect to their sparsity performance.[7] However, the so far developed algorithms, which intend to implement the transforms associated with the aforementioned representation systems for continuous data, have lacked either the precise implementation of a digital version of the theory for continuous data, or they are not publicly available, and, in addition, neither of those has been reproducibly tested.

To our mind, the essence of software development – as practised in most companies in contrast to the scientific community – is a careful software design with an elaborate testing environment. In this paper we aim at providing both:

(1) A rationally designed digital parabolic scaling algorithm.

(2) An elaborate testing environment with predefined quantitative quality measures for general parabolic scaling algorithms.

Further author information: (Send correspondence to G. Kutyniok)
D.L.D.: E-mail: donoho@stanford.edu, Telephone: 1 650 723 3350
G.K.: E-mail: kutyniok@uni-osnabrueck.edu, Telephone: 49 541 969 3516
M.S.: E-mail: mshahram@stanford.edu, Telephone: 1 650 725 2236

For (1), we chose the shearlet transform as one example. Our goal will be to developed an implementation of the digital shearlet transform based on a digital shearlet theory we will introduce, which is the exact digitalization of the shearlet theory for continuous data.

For (2), our goal is three-fold. Firstly, we intend to use such quantitative quality measures to tune the parameters of our algorithm. This will allow us to speak of a 'recommended choice' of the parameters of our implementation with precise meaning of a 'good' or 'bad' choice. Secondly, other researchers might find such a testing framework useful for optimizing their digital parabolic scaling algorithms. Thirdly, the scientific community could benefit significantly from such standardized tests, since these allow quantitative comparison of different digital parabolic scaling algorithms, thereby lifting the competition of developing the 'best' code to a fair and non-objective battleground, provided that in addition the philosophy of reproducible research[10] is followed by all members of the community.

1.1. Desiderata

A fundamental task to develop a digital shearlet transform based on a rational design is the agreement on a list of desiderata such transform shall satisfy. We agreed upon the following list, which does not claim completeness.

[D1] *Algebraic Exactness.* The transform should be based on a shearlet theory for digital data, than merely being 'somewhat close' to the shearlet theory for continuous data.

[D2] *Isometry of Pseudo-Polar Transform.* We introduce oversampling and weights to obtain an isometric pseudo-polar transform, which allows us to use the adjoint as inverse transform.

[D3] *Tight Frame Property.* The shearlet coefficients computed by the transform should be based on a tight frame decomposition, which ensures an isometric relation between the input image and the sequence of coefficients as well as allows us to use the adjoint as inverse transform. This property follows by combining [D1] and [D2], and allows the comparison with other transforms in contrast to those previous two tests.

[D4] *Time-Frequency-Localization.* The spatial portrait of the analyzing elements should 'look like' shearlets in the sense that they are sufficiently smooth as well as time-localized. Recall that we have localization and smoothness in frequency domain by definition.

[D5] *True Shear Invariance.* Since the orientation-related operator of shearlets is in fact the shear operator, we expect to see a shearing of the input image mirrored in a simple shift of the transform coefficients.

[D6] *Speed.* The transform should admit an algorithm which is order $O(N^2 log(N))$ flops, where N^2 is the number of digital points of the input image.

[D7] *Geometric Exactness.* The transform should preserve geometric properties parallel to those of the continuum theory, for example, edges should be mapped to edges in the transform domain.

[D8] *Robustness.* The transform should be resilient against impacts such as (hard) thresholded and quantized coefficients.

Although some desiderata are formulated for shearlet systems, it can easily be seen that they are relevant to any other digital parabolic scaling algorithm. Hence it is conceivable to assume that this list is a common bracket of desiderata for any digital implementation of a directional multiscale system based on parabolic scaling.

1.2. Shearlets: A Directional Multiscale Representation System based on Parabolic Scaling

The main idea for the construction of the shearlet transform for functions in $L^2(\mathbb{R}^2)$ introduced in[12] is the choice of a two-parameter dilation group, where one parameter ensures the multiscale property, whereas the second

parameter provides a means to detect directions. For each $a > 0$ and $s \in \mathbb{R}$, let D_a denote the *parabolic scaling matrix* and S_s denote the *shear matrix* of the form

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

respectively. To provide an equal treatment of the x- and y-axis, we split the frequency plane into the following four cones $\mathcal{C}_1 - \mathcal{C}_4$:

$$\mathcal{C}_\iota = \begin{cases} \{(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_1 \geq 1, |\xi_1/\xi_2| \geq 1\} & : & \iota = 1, \\ \{(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_2 \geq 1, |\xi_1/\xi_2| \leq 1\} & : & \iota = 2, \\ \{(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_1 \leq -1, |\xi_1/\xi_2| \geq 1\} & : & \iota = 3, \\ \{(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_2 \leq -1, |\xi_1/\xi_2| \leq 1\} & : & \iota = 4. \end{cases}$$

We now present the definition for a scaling of 4^j. A shearlet system can be defined similarly for a scaling of 2^j. However, in this case the odd scales have to be handled particularly carefully.

Let now $\psi_1 \in L^2(\mathbb{R})$ be a wavelet with $\hat{\psi}_1 \in C^\infty(\mathbb{R})$ and supp $\hat{\psi}_1 \subseteq [-4, -\frac{1}{4}] \cup [\frac{1}{4}, 4]$, and let $\psi_2 \in L^2(\mathbb{R})$ be a 'bump' function satisfying $\hat{\psi}_2 \in C^\infty(\mathbb{R})$ and supp $\hat{\psi}_2 \subseteq [-1, 1]$. We define $\psi \in L^2(\mathbb{R}^2)$ by

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\,\hat{\psi}_2(\tfrac{\xi_2}{\xi_1}). \tag{1}$$

For cone $\mathcal{C}_1$, at scale $j \geq 0$, orientation $k = -2^j, \ldots, 2^j$, and spatial position $m \in \mathbb{Z}^2$, the associated *shearlets* are then defined by their Fourier transforms

$$\begin{aligned} \hat{\sigma}_\eta(\xi) &= 2^{-j\frac{3}{4}} \hat{\psi}(S_k' D_{4^{-j}} \xi) \chi_{\mathcal{C}_1}(\xi) \exp\{-i(D_{4^{-j}} S_k m)' \xi\} \\ &= 2^{-j\frac{3}{4}} \hat{\psi}_1(\xi_1/4^j) \hat{\psi}_2(k + 2^j \xi_2/\xi_1) \chi_{\mathcal{C}_1}(\xi) \exp\{-i(D_{4^{-j}} S_k m)' \xi\}, \end{aligned}$$

where $\eta = (j, k, m, \iota)$ index scale, orientation, position, and cone. The shearlets for $\mathcal{C}_2 - \mathcal{C}_4$ are defined likewise by symmetry, as illustrated in Figure 1, and we denote the resulting *shearlet system* by

$$\{\sigma_\eta : \eta \in \mathbb{N}_0 \times \{-2^j, \ldots, 2^j\} \times \mathbb{Z}^2 \times \{1, \ldots, 4\}\}. \tag{2}$$

The definition shows that shearlets live on anisotropic regions of width 2^{-2j} and length 2^{-j} at various orienta-
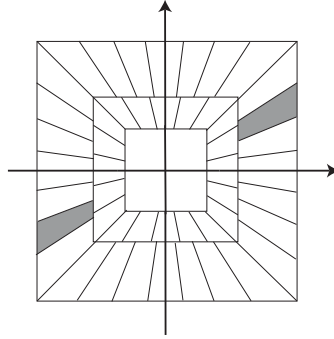
Figure 1. The tiling of the frequency domain induced by shearlets.

tions, which are parametrized by slope rather than angle as for second generation curvelets.

Setting

$$\mathcal{C} = \bigcup_{\iota=1}^{4} \mathcal{C}_\iota,$$

we have the following theorem from [12, Thm. 3] concerning the frame properties of the discrete shearlet system. Recall that $L^2(\mathcal{C}^\vee) = \{f \in L^2(\mathbb{R}^2) : \text{supp } \hat{f} \subset \mathcal{C}\}$.

THEOREM 1.1. *The system* (2) *is a Parseval frame for* $L^2(\mathcal{C}^\vee)$.

We remark that the low frequency part can be appropriately filled in to obtain a Parseval frame for $L^2(\mathbb{R}^2)$. Finally, the generating window allows in fact more freedom than (1), but in this paper we restrict ourselves to this (customary) choice.

1.3. Ingredients of the Digital Shearlet Transform

The definition of the shearlet transform for continuous data (see Figure 1) indicates that the cartesian grid is not appropriate for the Fourier domain tiling. Instead the pseudo-polar grid (compare Figure 2) seems perfectly adapted to the particular tiling we aim for, and, in addition, the paper[1] provides a fast pseudo-polar transform. However, we face the problem that the pseudo-polar transform is not an isometry, which prevents us from merely employing the adjoint transform as the inverse transform. To circumvent this problem, we multiply the values on the grid points with weights which loosely speaking can be viewed as a density compensation of the sampling points. For such weights to exist, the pseudo-polar grid needs to be oversampled by a factor of 8. Once the sampling points are carefully weighted, the shearlet coefficient arrays can be computed by applying a selection of window functions on the pseudo-polar grid.

Summarizing, the digital shearlet transform of an $N \times N$ image will require the following steps:

1) Application of the pseudo-polar Fourier transform with an oversampling factor of 8 in radial direction.

2) Weighting of the function values on the pseudo-polar grid.

3) Computation of the inner products of the values on the pseudo-polar grid with a scaled and sheared generating window.

4) Application of the 2D inverse fast Fourier transform to each array.

With a careful choice of the weights and the windows, this transform is an isometry. Hence the inverse transform can be computed by merely taking the adjoint in each step.

1.4. Testing Environment

How do we tell, whether an implementation is better than another one? This is a question researcher are facing frequently. Also, how do we tune the parameters of our algorithm? A common choice is the denoising and compression of images like 'Lena', etc. However, this is far from being satisfactory for a careful researcher, since it analyses the performance of an algorithm only for a very narrowly focussed problem. What is truly seeked for are quantitative measures which allow the comparison of algorithms taking into account a wide range of performance criteria.

To tackle these urging questions in the field of digital parabolic scaling algorithms in general, we will introduce a suite of measures which quantitatively test the desiderata stated in Subsection 1.1. Each desiderata will be measured by one real value or by one curve, thus providing a standardized framework to answer the previously stated questions comprehensively.

1.5. Related Work

We also wish to give credit to several researchers who designed algorithms for transforms based on parabolic scaling, and, in particular, mention the curvelet implementation 'CurveLab',[2] the contourlet implementation,[5] as well as two shearlet implementations.[11, 16] Although performing very well in the ubiquitous denoising and compression tests, from neither of them test results beyond those were presented. Some also differ from the theory for continuous data, and others are not even publicly available.

1.6. Contribution of this Paper

Therefore, we view the contributions of this paper two-fold. Firstly, we introduce a digital shearlet transform which is *rationally designed* in the sense of it being based on a carefully defined digital shearlet theory, which is the exact digitalization of the shearlet theory for continuous data. Secondly, we provide a variety of *quantitative quality measures* for general parabolic scaling algorithms, which allow a quantitative statement about various performance aspects of such algorithms. The parameters of our digital shearlet algorithm was tuned utilizing this *testing environment*, which enabled us to provide the reader with a recommended choice of parameters within this framework.

All presented algorithms and tests as well as codes for the displayed figures and tables are provided at URL `ShearLab.org`[8] in the spirit of reproducible research.[10]

2. AN ISOMETRIC PSEUDO-POLAR TRANSFORM

2.1. General Condition on Density Compensation

Given an $N \times N$ image I, it is well known that the Fourier transform $\hat{I}$ of I evaluated on a rectangular $N \times N$ grid is an isometry:

$$\sum_{u,v=-N/2}^{N/2-1} |I(u,v)|^2 = \frac{1}{N^2} \sum_{\omega_x,\omega_y=-N/2}^{N/2-1} |\hat{I}(\omega_x,\omega_y)|^2. \tag{3}$$

We aim to obtain a similar formula for the Fourier transform of I evaluated on the pseudo-polar grid. For this, we first extend the definition of the pseudo-polar grid slightly by introducing an oversampling parameter $R > 0$ in radial direction. This new grid, which we will denote in the sequel by Ω_R, is defined by

$$\Omega_R = \Omega_R^1 \cup \Omega_R^2, \tag{4}$$

where

$$\begin{aligned}
\Omega_R^1 &= \{(-4\ell k/(RN), 2k/R) : -N/2 \le \ell \le N/2, -RN/2 \le k \le RN/2\}, \\
\Omega_R^2 &= \{(2k/R, -4\ell k/(RN)) : -N/2 \le \ell \le N/2, -RN/2 \le k \le RN/2\}.
\end{aligned}$$

This grid is illustrated in Figure 2. Notice that the 'original' pseudo-polar grid (see[1]) is a special case of this
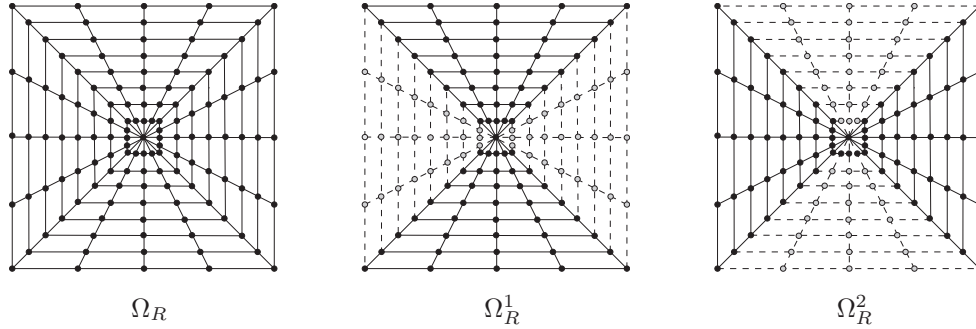
$$\Omega_R \qquad\qquad \Omega_R^1 \qquad\qquad \Omega_R^2$$

Figure 2. The pseudo-polar grid for $N = 4$ and $R = 4$.

definition for $R = 2$.

Choosing the weights carefully, a 'Plancherel theorem' similar to (3) can be proved (see[9]) for the pseudo-polar grid $\Omega_R = \Omega_R^1 \cup \Omega_R^2$ defined in (4).

THEOREM 2.1. *Let N be even, and let $w : \Omega_R \to \mathbb{R}^+$ be a weight function. Then*

$$\sum_{u,v=-N/2}^{N/2-1} |I(u,v)|^2 = \sum_{(\omega_x,\omega_y)\in\Omega_R} w(\omega_x,\omega_y) \cdot |\hat{I}(\omega_x,\omega_y)|^2$$

holds if and only if, for all $-N/2 \leq u, v \leq N/2 - 1$, the weights w satisfy

$$
\begin{aligned}
\delta(u,v) \quad = \quad & (2N+1) \cdot w(0,0) \\
& +4 \cdot \sum_{\ell=0,N/2} \sum_{k=1}^{RN/2} w(2k/R, -4\ell k/(RN)) \cdot \cos(2ku/(R(2N+1))) \cdot \cos(4\ell kv/(RN(2N+1))) \\
& +8 \cdot \sum_{\ell=1}^{N/2-1} \sum_{k=1}^{RN/2} w(2k/R, -4\ell k/(RN)) \cdot \cos(2ku/(R(2N+1))) \cdot \cos(4\ell kv/(RN(2N+1)))
\end{aligned}
$$

and, for all $(\omega_x, \omega_y) \in \Omega_R$, the weights w satisfy the symmetry conditions

$$
w(\omega_x, \omega_y) = w(\omega_y, \omega_x), \; w(\omega_x, \omega_y) = w(-\omega_y, \omega_x), \; w(\omega_x, \omega_y) = w(-\omega_x, \omega_y) \text{ and } w(\omega_x, \omega_y) = w(\omega_x, -\omega_y).
$$

Moreover, in general, R needs to be at least 8 for such weights to exist.

In fact, the stated conditions the weights have to satisfy lead to a linear system of equations with $RN^2/8 + 1/16$ variables and N^2 equations, wherefore, in general, we need the oversampling factor R to be at least 8 to enforce solvability.

2.2. Recommended Choice of Weights

To avoid high complexity, we expand the weights in terms of an (non-complete) system, compute the coefficients, and hardwire them in the algorithm. For the present algorithm, we chose 7 functions $w_1, \ldots, w_7$ on the pseudo-polar grid such that $\sum_{j=1}^{7} w_j(\omega_x, \omega_y) \neq 0$ for each $(\omega_x, \omega_y) \in \Omega_R$. The center

$$
w_1 = 1_{(0,0)} \quad \text{and} \quad w_2 = 1_{\{(\omega_x, \omega_y) : |k| = 1\}},
$$

the boundary

$$
w_3 = 1_{\{(\omega_x, \omega_y) : |k| = NR/2 \text{ and } \omega_x = \omega_y\}} \quad \text{and} \quad w_4 = 1_{\{(\omega_x, \omega_y) : |k| = NR/2 \text{ and } \omega_x \neq \omega_y\}},
$$

and the seam lines

$$
w_5(\omega_x, \omega_y) = |k| \cdot 1_{\{(\omega_x, \omega_y) : 1 < |k| < NR/2 \text{ and } \omega_x = \omega_y\}} \quad \text{and} \quad w_6 = 1_{\{(\omega_x, \omega_y) : |k| = NR/2 - 3 \text{ and } \omega_x = \omega_y\}}
$$

have to be treated with much care compared to the 'interior'

$$
w_7(\omega_x, \omega_y) = |k| \cdot 1_{\{(\omega_x, \omega_y) : 1 < |k| < NR/2 \text{ and } \omega_x \neq \omega_y\}}.
$$

Notice that here we use (ω_x, ω_y) and (k, ℓ) interchangeably.

The weighting generated by our recommended choice of coefficients for $w_1, \ldots, w_7$ is displayed in Figure 3. The different regions, in particular, the seam lines can clearly be seen.

Certainly, the isometry suffers from this restricted choice. But tuning the algorithm, we found that with respect to the trade-off between the closeness to isometry and complexity of the algorithm, this was the best choice.

3. WINDOWING THE PSEUDO-POLAR GRID

[This needs to be filled in based on the present implementation.]

4. DIGITAL SHEARLET AND ADJOINT SHEARLET TRANSFORM

[This needs to be filled in based on the present implementation.]
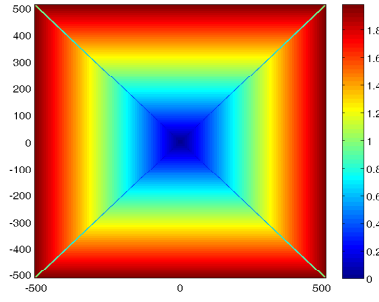
Figure 3. Recommended weighting of the pseudo-polar grid.

Figure 4. Window for [parameters].

5. QUANTITATIVE TEST MEASURES

To ensure and also prove that our implementation satisfies the previously proposed desiderata, we will define quality measures for each of those and in Section 6 provide numerical results on how accurate our implementation satisfies these. These test measures

- provided us with a framework within which we could tune our algorithm,

- are a means to present the performance of our algorithm,

- might serve as basis of comparison of parabolic scaling algorithms in general.

In the following, P shall denote the operator defined by the code which computes the pseudo-polar tranform, w shall denote the weighting applied to the values on the pseudo-polar grid, W shall be the windowing with additional FFT, and S shall denote the operator defined by the code which computes the shearlet transform as well as its adjoint will be denotde by $S^\star$.

[D1] *Algebraic Exactness.*
Generate a sequence of 5 random images $I_1, \ldots, I_5$ on a pseudo-polar grid for $N = 256$ and $R = 8$ with standard normally distributed entries. Our quality measure will then be the Monte Carlo estimate for the operator norm $\|W^\star W - Id\|_{op}$ given by

$$M_{[D1]} = \max_{i=1,\ldots,5} \frac{\|W^\star W I_i - I_i\|_2}{\|I_i\|_2}.$$

[D2] *Isometry of Pseudo-Polar Transform.*

- Closeness to tightness. Generate a sequence of 5 random images $I_1, \ldots, I_5$ of size 256×256 with standard normally distributed entries. Our quality measure will then be the Monte Carlo estimate for the operator norm $\|P^* w P - Id\|_{op}$ given by

$$M_{[D2],1} = \max_{i=1,\ldots,5} \frac{\|P^* w P I_i - I_i\|_2}{\|I_i\|_2}.$$

- Quality of preconditioning. Our quality measure will be the spread of the eigenvalues of the Gram operator $P^\star w P$ given by

$$M_{[D2],2} = \frac{\lambda_{\max}(P^\star w P)}{\lambda_{\min}(P^\star w P)}.$$

[D3] *Tight Frame Property.*

Generate a sequence of 5 random images $I_1, \ldots, I_5$ of size 256×256 with standard normally distributed entries. Our quality measure will then be the Monte Carlo estimate for the operator norm $\|S^*S - Id\|_{op}$ given by

$$M_{[D3]} = \max_{i=1,\ldots,5} \frac{\|S^*SI_i - I_i\|_2}{\|I_i\|_2}.$$

[D4] *Time-Frequency-Localization.*

Let I be a shearlet in a 256×256 image centered at the origin $(129, 129)$ with slope 0 of scale 5. Our quality measure will be four-fold:

- Decay in Spatial Domain. We compute the decay rates $d_1, \ldots, d_{256}$ along lines parallel to the y-axis starting from the line $[129, :]$ and the decay rates $d_{257}, \ldots, d_{512}$ with x and y interchanged. By decay rate, for instance, for the line $[129 : 256, 1]$, we first compute the smallest monotone majorant $M(x, 1)$, $x = 129, \ldots, 256$ – note that we could also choose an average amplitude here or a different 'envelope' – for the curve $|I(x, 1)|$, $x = 129, \ldots, 256$. Then the decay rate is defined to be the average slope of the line, which is a least square fit to the curve $\log(M(x, 1))$, $x = 129, \ldots, 256$. Based on these decay rates, we choose our measure to be

$$M_{[D4],1} = \max_{i=1,\ldots,512} d_i.$$

- Decay in Frequency Domain. Here we intend to check whether the Fourier transform of I is compactly supported and also the decay. For this, let $\hat{I}$ be the 2D-FFT of I and compute the decay rates d_i, $i = 1, \ldots, 512$ as before. Then we define the measure by

$$M_{[D4],2} = \max_{i=1,\ldots,512} d_i.$$

- Smoothness in Spatial Domain. We will measure smoothness by Hölder regularity and minimize $|I(u) - I(v)|/|u - v|^{\tilde{\alpha}}$ over all $u \neq v$ with respect to $\tilde{\alpha}$. The value of α, for which the minimum is reached, will define our measure

$$M_{[D4],3} = \alpha.$$

- Smoothness in Frequency Domain. We compute the smoothness now for $\hat{I}$, the 2D-FFT of I to obtain the new α, which will be our measure

$$M_{[D4],4} = \alpha.$$

[D5] *True Shear Invariance.*

Let I be an 256×256 image with a cross of the form $[129, :] \cup [:, 129]$. For each scale j we choose the set $S_j = \{s : 2^{-j}s \in \mathbb{Z} \text{ and } \hat{\psi}(S_{2^{-j}s}^T A_4^{-j} \cdot) \text{ does not touch the seam line}\}$, and generate a sequence of images $I_{s,j} = I(S_s \cdot)$, $s \in S_j$. Our quality measure will then be the curve

$$j \mapsto \max_{s \in S_j} \frac{\|SI_{s,j} - C_{j,s}SI\|_2}{\|I\|_2},$$

where $C_{j,s}$ shifts the shear parameter of the coefficients by $2^j s$.

[D6] *Speed.*

Generate a sequence of 5 random images I_i, $i = 5, \ldots, 9$ of size $2^i \times 2^i$ with standard normally distributed entries. Let s_i be the speed of the Shearlet Transform S applied to I_i. Our hypothesis is that the speed behaves like $s_i = c \cdot (2^{2i})^d$; 2^{2i} being the size of the input. Let now $\tilde{d}_a$ be the average slope of the line, which is a least square fit to the curve $i \mapsto \log(s_i)$. Let also f_i be the 2D-FFT applied to I_i, $i = 5, \ldots, 9$. Our quality measure will measures for complexity, the constant, and comparison with 2D-FFT defined by

$$M_{[D6],1} = \frac{\tilde{d}_a}{2 \log 2}, \quad M_{[D6],2} = \frac{1}{5} \sum_{i=5}^{9} \frac{s_i}{(2^{2i})^{M_{[D6],1}}} \quad \text{and} \quad M_{[D6],3} = \frac{1}{5} \sum_{i=5}^{9} \frac{s_i}{f_i}.$$

[D7] *Geometric Exactness.*

Let $I_1, \ldots, I_8$ be 256×256 images of a heaviside function for the edge – by interpolation – through the origin $(129, 129)$ and of slope $[-1, -0.5, 0, 0.5, 1]$ and the transpose of the middle three, and let $c_{i,j}$ be the associated shearlet coefficients for image I_i and scale j. Our quality measure will two-fold:

- Decay of significant coefficients. Consider the curve $j \mapsto \frac{1}{8} \sum_{i=1}^{8} \max |c_{i,j}|$, where the coefficients $c_{i,j}$ are associated with shearlets aligned with the line. Let d be the average slope of the line, which is a least square fit to log of this curve, and define

$$M_{[D7],1} = d.$$

- Decay of insignificant coefficients. Consider the curve $j \mapsto \frac{1}{8} \sum_{i=1}^{8} \max |c_{i,j}|$, where the coefficients $c_{i,j}$ are associated with shearlets except the ones aligned with the line. Let d be the average slope of the line, which is a least square fit to log of this curve, and define

$$M_{[D7],2} = d.$$

[D8] *Robustness.*

- Thresholding. Let I be the regular sampling of a Gaussian function with mean 0 and variance 256 on $\{-128, 127\}^2$ generating an 256×256-image. Our quality measure for $k = 1, 2$ will be the curve

$$p_k \mapsto \frac{\|S^\star \operatorname{thres}_{p_k} SI - I\|_2}{\|I\|_2},$$

where

- $\operatorname{thres}_{p_1}$ discards $100 \cdot (1 - 2^{-p_1})$ percent of the coefficients $(p_1 = [2 : 2 : 20])$,
- $\operatorname{thres}_{p_2}$ sets all those coefficients to zero with absolute values below the threshold $m/2^{p_2}$ with m being the maximal absolute value of all coefficients. $(p_2 = [0.5 : 0.5 : 5])$
- Quantization. Let I be the regular sampling of a Gaussian function with mean 0 and variance 256 on $\{-128, 127\}^2$ generating an 256×256-image. Our quality measure will be the curve

$$[5 : -0.5 : 0.5] \ni q \mapsto \frac{\|S^\star \operatorname{quant}_q SI - I\|_2}{\|I\|_2},$$

where $\operatorname{quant}_q(c) = \operatorname{round}(c/(m/2^q)) \cdot (m/2^q)$ and m being the maximal absolute value of all coefficients.

6. TEST RESULTS

6.1. Results for Tests [D1] − [D3]

Table I presents the values for the algebraic exactness $M_{[D1]}$, the isometry of the pseudo-polar transform $M_{[D2],1}$ and $M_{[D2],1}$, as well as for the tightness of the shearlet transform $M_{[D3]}$, which our recommended version achieves.

TABLE I
RESULTS FOR [D1] − [D3]

$M_{[D1]}$	$M_{[D2],1}$	$M_{[D2],1}$	$M_{[D3]}$
0.0023147827	0.016212074	1.6584499	0.017172738

The tightness of the transform is with an error of about 0.017 very accurately achieved. The first three values indicate which part of the algorithm (weighting or windowing) contribute most to this error, hence where an attempt to further improve this error shall be made. We clearly see that the weighting produces an error of about 0.016 with quotient of the eigenvalues of the Gram matrix being about 1.658, compared to the error arising from the windowing which is with 0.002 already in a different league. Thus progress on the choice of weights will have a significant impact on the tightness error. The reader should however keep in mind that there is a trade-off between the precision of the weights and the complexity of the algorithm (compare Subsection 2.2).

For illustration purposes we add a figure (see Figure 5) showing the error matrix for the three tests.
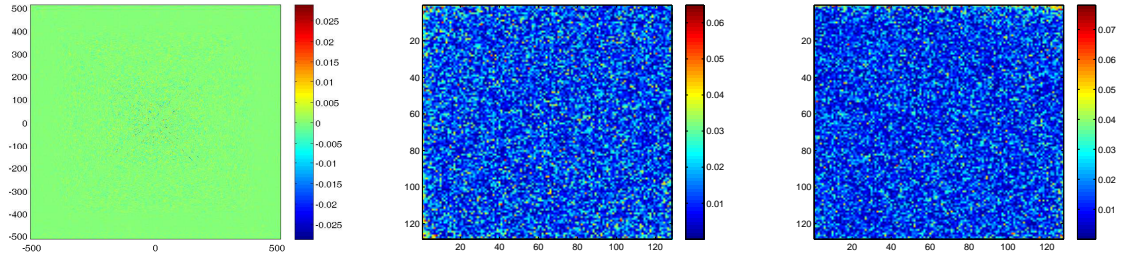
Figure 5. Error matrix for tests [D1], [D2], and [D3] with $N = 128$, $R = 8$, and $\beta = 2$.

6.2. Results for Test [D4]

Table II presents the values for the time-frequency-localization, in particular, $M_{[D4],1}$ for the decay in spatial domain, $M_{[D4],2}$ for the decay in frequency domain, $M_{[D4],3}$ for the smoothness in spatial domain, and $M_{[D4],4}$ for the smoothness in spatial domain.

TABLE II
RESULTS FOR [D4]

$M_{[D4],1}$	$M_{[D4],2}$	$M_{[D4],3}$	$M_{[D4],4}$

[Complete results (m-file still needs to be written) and comment on them.]

The advantageous decay and smoothness properties are in addition visualized in Figure 6.

Figure 6. Example of a shearlet in spatial domain and its Fourier transform.

[Include figure (m-file still needs to be written).]

6.3. Results for Test [D5]

The curve which serve as our quality measures for this test is displayed in Figure [to be included].
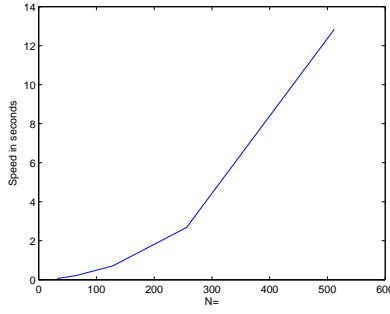
Figure 7. Result for [D5].

[Complete results (m-file still needs to be written) and comment on them.]

6.4. Results for Test [D6]

The values of the measures for the speed performance, in particular, the complexity $M_{[D6],1}$, the constant $M_{[D6],2}$, and the comparison with 2D-FFT $M_{[D6],3}$, are presented in Table III.

TABLE III
RESULTS FOR [D6]

$M_{[D6],1}$	$M_{[D6],2}$	$M_{[D6],3}$
0.42944697	1.0258062	255.77436

Figure 8. Speed of the forward Shearlet Transform for $R = 8$ and $\beta = 2$.

The precise speed curve from which these values were extracted can be found in Figure 8.

Recalling the definition of the test measures, we conclude that the complexity of our algorithm approximately equals N, where $N \times N$ is the size of the input image. Moreover, it is about 250 times slower than the 2D-FFT.

6.5. Results for Test [D7]

The values of the measures for geometric exactness, in particular, the decay of the significant coefficients $M_{[D7],1}$ and the decay of the insignificant coefficients $M_{[D7],2}$ are presented in Table IV.

TABLE IV
RESULTS FOR [D7]

$M_{[D7],1}$	$M_{[D7],2}$

[Complete results (m-file still needs to be written) and comment on them.]

For illustration purposes we add a figure showing the curve of the decay of the significant and insignificant coefficients.

Figure 9. Decay of significant and insignificant coefficients for $N = 128$, $R = 8$, and $\beta = 2$.
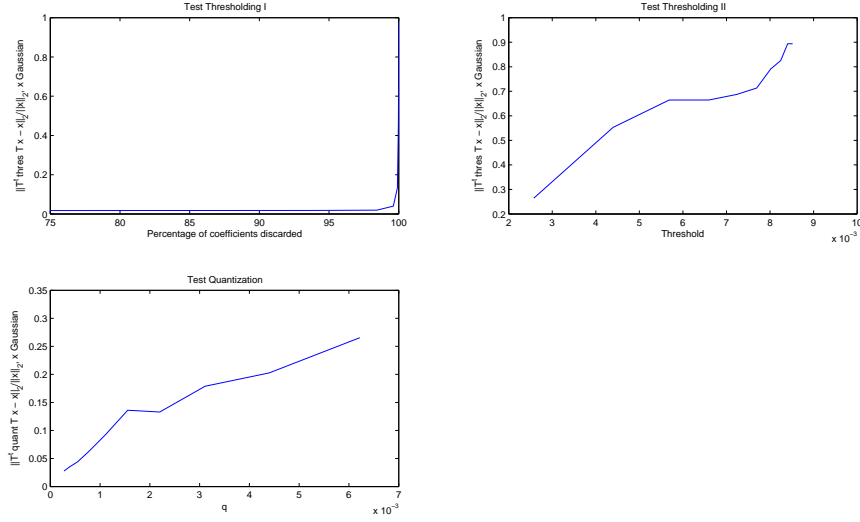
[Include figures (m-files still needs to be written).]

6.6. Results for Test [D8]

The curves which serve as our quality measures for this test are displayed in Figure 10.

[Comment on those; maybe split figures into 3 single ones and put the figures in a line.]

7. DISCUSSION

Our work makes several contributions.

Figure 10. Results for test [D8] on resilience against thresholding and quantization for $N = 128$, $R = 8$, and $\beta = 2$.

7.1. Rational Design of Digital Parabolic Scaling Algorithm

Among various existing parabolic scaling systems, we have chosen the shearlet system as an example, and implemented an associated digital algorithmic realization following a rational design in the following sense.

- A digital theory is provided, which is a exact digitalization of the theory for continuous data.

- The implementation follows this digital theory stricto sensu.

Only an implementation following such a rational design can truly reflect an elaborate theory for continuous data.

We hope that researchers adapt such a rational design strategy in the future, since otherwise each theory for continuous data – neglected by a code which does not arise from a carefully designed associated digital theory – seems in vain for applications.

Certainly, a similar strategy can be driven for other mathematically defined parabolic scaling systems such as, for instance, curvelets or contourlets.

7.2. Standardized Framework for Evaluation and Comparison

Our extensive testing environment might be useful to active researchers in the field of directional multi-scale transforms based on parabolic scaling.

Firstly, it provides a means to precisely determine the performance of a given algorithm. This enables the researcher to tune the parameters of his algorithm with respect to the introduced quantitative measures. Thus vague expressions such as 'Now the algorithm seems to perform better.' will belong to the past in this field, and are replaced by standardized quantitative statements.

Secondly, researchers are now able to compare their algorithms within the presented standardized testing framework. This adds a new chapter to the philosophy of reproducible research,[10] since once implemented everybody from the community has not only access to newly designed algorithms, but also the means for a fair comparison of such.

Our general hope is to systematize the design of algorithms. Certainly, other testing measures can be imagined and added to the list. However, we believe that this present list contains and formalizes the core

problems digital parabolic scaling algorithms can encounter. Our testing environment is publicly provided at URL `ShearLab.org`.[8]

By sharing the complete code for the tests, we would like to encourage researchers, who believe that important aspects of a parabolic scaling algorithm are left out, to define additional tests and add those to our testing framework. Also we share the hope that researchers will support the idea of having a standardized elaborate test dictionary, which greatly extends the common noise suppression and compression tests using 'Lena', etc. The reader shall be reminded that our tests are specifically designed for parabolic scaling algorithms.

7.3. Results for Present Version

The aforementioned testing framework was utilized to tune our algorithm accordingly. This was proved by providing the values of the test measures our algorithm satisfies. The aspects of the algorithm which were optimized were, in particular, the choice of the weights, the tiling of the pseudo-polar grid, and choice of the window functions. The final result – the recommended version – was then presented. The code with the recommended choice of parameters is publicly available at URL `ShearLab.org`.[8]

We can report that such a framework was a significant help when tuning our algorithm. This is already common practise in software design in companies, however hardly heard of in the scientific community. This paper shall serve as a first step towards a more careful design of algorithms, since, in particular, among researchers a common 'battleground' in which exact rules are predefined should be especially welcome.

We do not claim that the test results for our algorithm presented in this paper are the best ones which can be achieved. We promote further tuning of our algorithm as well as invoking novel ideas to reduce the values of the testing measures further aiming at outperforming the present implementation.

8. CONCLUSIONS

We introduced a digital shearlet transform as one example of a parabolic scaling algorithm. The implementation was based on a carefully defined digital shearlet theory, which is the exact digitalization of the shearlet theory for continuous data. In this sense our digital shearlet transform was indeed rationally designed. The main ingredients of our algorithm are a weighted pseudo-polar transform to achieve an isometric mapping into the pseudo-polar domain, thereby using a particular oversampling strategy, as well as a careful selection of tiling windows. We also defined test measures for a general parabolic scaling algorithm, which allow a quantitative statement about various performance aspects of such algorithms. Our digital shearlet transform was tuned with respect to those measures. The values of the quality measures of our recommended version – after careful tuning – were presented.

ACKNOWLEDGEMENTS

REFERENCES

1. A. Averbuch, R. R. Coifman, D. L. Donoho, M. Israeli, and Y. Shkolnisky, *A framework for discrete integral transformations I – the pseudo-polar Fourier transform*, SIAM J. Sci. Comput. **30** (2008), 764–784.
2. E. J. Candès, L. Demanet, D. L. Donoho and L. Ying, *Fast discrete curvelet transforms,* Multiscale Model. Simul. **5** (2006), 861–899.
3. E. J. Candès and D. L. Donoho, *Continuous curvelet transform: I. Resolution of the wavefront set*, Appl. Comput. Harmon. Anal. **19** (2005), 162–197.
4. E. J. Candès and D. L. Donoho, *Continuous curvelet transform: II. Discretization of frames*, Appl. Comput. Harmon. Anal. **19** (2005), 198–222.

5. M. N. Do and M. Vetterli, *The contourlet transform: an efficient directional multiresolution image representation*, IEEE Trans. Image Process. **14** (2005), 2091–2106.

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

7. D. L. Donoho and G. Kutyniok, *Sparsity Equivalence of Anisotropic Decompositions*, preprint.

8. D. L. Donoho, G. Kutyniok, and M. Shahram, `www.ShearLab.org`.

9. D. L. Donoho, G. Kutyniok, and M. Shahram, *ShearLab: A Rational Design of a Digital Parabolic Scaling Algorithm*, preprint.

10. D. L. Donoho, A. Maleki, M. Shahram, V. Stodden, and I. Ur-Rahman, *Fifteen years of Reproducible Research in Computational Harmonic Analysis,* Computing in science and engineering **11** (2009), 8–18.

11. G. Easley, D. Labate, and W. Lim, *Sparse Directional Image Representations using the Discrete Shearlet Transform*, Appl. Comput. Harmon. Anal. **25** (2008), 25–46.

12. 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.

13. F. J. Herrmann and G. Hennenfent, *Non-parametric seismic data recovery with curvelet frames*, Geophys. J. Int. **173** (2008), 233-248.

14. G. Kutyniok and D. Labate, *Resolution of the Wavefront Set using Continuous Shearlets*, Trans. Amer. Math. Soc. **361** (2009), 2719–2754.

15. G. Kutyniok, D. Labate, and W. Lim, `www.Shearlet.org`.

16. W. Lim, *Discrete Shearlet Transform: New Multiscale Directional Image Representation* Proceedings of SAMPTA'09, Marseille 2009.

17. J.-L. Starck, E. Candès, and D. L. Donoho, *Astronomical image representation by the curvelet tansform*, Astronomy and Astrophysics **398** (2003), 785–800.