3D Discrete Shearlet Transform and Video Denoising

Demetrio Labate and Pooran Singh Negi
Department of Mathematics, University of Houston, Houston, TX, USA.

ABSTRACT

This paper introduces a numerical implementation of the 3D shearlet transform, a directional transform which is derived from the theory of shearlets. The shearlet approach belongs to a class of directional multiscale methods emerged during the last 10 years to overcome the limitations of traditional multiscale systems, which also include curvelets and contourlets. Unlike other methods, shearlets are derived from the theory of affine systems, which allows a very flexible mathematical structure and a natural transition from the continuous to the digital setting. Following the recent proof of the optimality of the 3D shearlet representation, in this paper we develop an algorithmic implementation of the 3D shearlet transform that follows closely the spatial-frequency pattern of the corresponding continuous transform. The performance of the algorithm is illustrated on problems of video denoising and successfully compared against other state-of-the-art multiscale techniques, including curvelets and surfacelets.

Keywords: Affine systems, curvelets, shearlets, sparsity, wavelets.

1. INTRODUCTION

The shearlet representation, originally introduced in Ref. 13, 17, has emerged in recent years as one of the most effective frameworks for the analysis and processing of multidimensional data. This approach is part of a number of multiscale methods introduced during the last 10 years with the goal to overcome the limitations of traditional wavelets by combining the standard multiscale decomposition with the ability to efficiently capture directional features. Other notable such methods include the curvelets[4] and the contourlets.[7] Both the shearlet and the curvelet representations, in particular, form a Parseval frame of $L^2(\mathbb{R}^2)$ which is (nearly) optimally sparse in the class of *cartoon-like image*, a standard model for images with edges.[4,14] Specifically, if f_M is the M term approximation obtained by selecting the M largest coefficients in the shearlet or curvelet expansion of a cartoon-like image f, then the approximation error satisfies the asymptotic estimate

$$||f - f_M^S||_2^2 \asymp M^{-2}(\log M)^3, \qquad \text{as } M \to \infty.$$

Up to the log-like factor, this is the optimal approximation rate, in the sense that no other orthonormal system or even frame can achieve a rate better than $O(M^{-2})$. By contrast, wavelet approximations can only achieve a rate $O(M^{-1})$.

Notice that, despite their similar approximations properties, shearlets and curvelets rely on a rather different mathematical structure. In fact, the shearlet approach is derived from the framework of affine systems, which provides a more flexible mathematical setting. In addition, the directionality of the shearlet systems is achieved through the use of shearing matrices rather than rotations, which are used in curvelets. This offers the advantage of preserving the discrete integer lattice and it enables a natural transition from the continuous to the discrete setting. We refer to Ref. 6, 11, 12, 20, 23 for additional observations about the numerical implementation of shearlet-based decompositions and their imaging applications.

While the theory of shearlets and curvelets has been well established in dimension 2, only very recently serious effort were made to extend this approach to higher dimensions. More precisely, the formal extension of the construction of multiscale directional systems from 2D to 3D is not a major challenge. In fact, 3D curvelets have been introduced in Ref. 3 and 3D versions of the contourlets, called *surfacelets*, in Ref. 21, with the focus in both cases being on their discrete implementations. However, the analysis of the sparsity properties of curvelets and shearlets in the 3D setting does not follow directly from the 2D argument and only very recently the sparsity

D.L. E-mail: dlabate@math.uh.edu; P.S.N. E-mail: psnegi@math.uh.edu

properties of 3D shearlets have been rigorously analyzed. Specifically, it was proved in Ref. 15 (cf. also a similar recent result announced in Ref. 19) that, for 3D functions f which are smooth away from discontinuities along C^2 surfaces, the M term approximation f_M^S obtained by selecting the N largest coefficients in the 3D Parseval frame shearlet expansion of f satisfies the asymptotic estimate

$$||f - f_M^S||_2^2 \asymp M^{-1}(\log M)^2, \qquad \text{as } M \to \infty. \tag{1}$$

Up to the logarithmic factor, this is the *optimal decay rate* for functions in this class and significantly outperforms wavelet approximations, which only yield a $M^{-1/2}$ rate.

The objective of the paper is to present a numerical implementation of the 3D Discrete Shearlet Transform which follows closely the spatial-frequency decomposition associated with the corresponding continuous transform so that it can take advantage of its sparsity properties. The performance of this numerical algorithm will be illustrated on problems of video denoising.

2. SHEARLET REPRESENTATIONS

The shearlet approach is derived from the framework of *wavelets with composite dilations*,[17,18] which are the collections of functions in $L^2(\mathbb{R}^n)$ of the form

$$\{\psi_{j,\ell,k}(x) = |\det A|^{j/2}\, \psi(B_\ell A^j x - k) : \, j \in \mathbb{Z}, \ell \in \Lambda, k \in \mathbb{Z}^n\}, \tag{2}$$

where $\psi \in L^2(\mathbb{R}^n)$, A is an expanding invertible $n \times n$ matrix, B is matrix for which $|\det B| = 1$ and Λ is a countable indexing set.

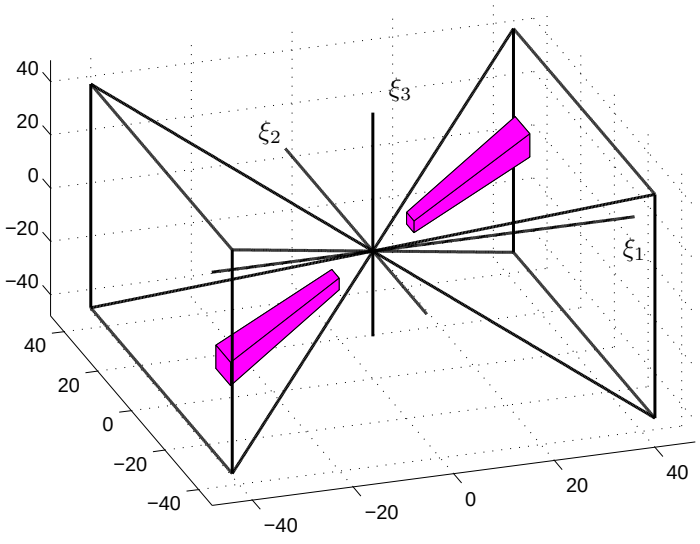


Figure 1. *Frequency support of a representative shearlet function $\psi_{j,\ell,k}$, inside the pyramidal region $\mathcal{D}_C$. The orientation of the support region is controlled by $\ell = (\ell_1, \ell_2)$; its shape is becoming more elongated as j increases ($j = 4$ in this plot).*

In the case of shearlets, the matrices $\{B_\ell\}$ in (2) are shearing matrices and they control the orientation of the elements of the representation function system. Specifically, the 3D shearlet systems have the form:

$$\left\{\psi_{j,\ell,k}^{(d)} = |\det A^{(d)}|^{\frac{j}{2}}\, \psi^{(d)}(B_\ell^{(d)}\, (A^{(d)})^j x - k) \text{: } j \in \mathbb{Z},\, \ell = (\ell_1, \ell_2) \in \Lambda_j, k \in \mathbb{Z}^3, d = 1, 2, 3\right\},$$

where $\psi^{(d)} \in L^2(\mathbb{R}^3)$, $\Lambda_j \subset \mathbb{Z}^2$ is the set $\{(\ell_1, \ell_2) : -2^j \leq \ell_1, \ell_2 \leq 2^j\}$ and the matrices $A^{(d)}$, $B_\ell^{(d)}$ are given by

$$A^{(1)} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix},\ A^{(2)} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix},\ A^{(3)} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix},$$

$$B_\ell^{(1)} = \begin{pmatrix} 1 & \ell_1 & \ell_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ B_\ell^{(2)} = \begin{pmatrix} 1 & 0 & 0 \\ \ell_1 & 1 & \ell_2 \\ 0 & 0 & 1 \end{pmatrix},\ B_\ell^{(3)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \ell_1 & \ell_2 & 1 \end{pmatrix}.$$

The generators $\psi^{(d)}$ of the shearlet system are chosen to be the smooth band-limited function defined by

$$\hat{\psi}^{(1)}(\xi_1, \xi_2, \xi_3) = \hat{\psi}_1(\xi_1)\, \hat{\psi}_2\Big(\frac{\xi_2}{\xi_1}\Big)\, \hat{\psi}_2\Big(\frac{\xi_3}{\xi_1}\Big),\ \ \hat{\psi}^{(2)}(\xi_1, \xi_2, \xi_3) = \hat{\psi}_1(\xi_2)\, \hat{\psi}_2\Big(\frac{\xi_1}{\xi_2}\Big)\, \hat{\psi}_2\Big(\frac{\xi_3}{\xi_2}\Big),$$
$$\text{and}\quad \hat{\psi}^{(3)}(\xi_1, \xi_2, \xi_3) = \hat{\psi}_1(\xi_3)\, \hat{\psi}_2\Big(\frac{\xi_1}{\xi_3}\Big)\, \hat{\psi}_2\Big(\frac{\xi_1}{\xi_3}\Big),$$

where ψ_1 and ψ_2 satisfy the following assumptions:

(i) $\hat{\psi}_1 \in C^\infty(\widehat{\mathbb{R}})$, $\operatorname{supp} \hat{\psi}_1 \subset [-\frac{1}{2}, -\frac{1}{16}] \cup [\frac{1}{16}, \frac{1}{2}]$ and

$$\sum_{j \geq 0} |\hat{\psi}_1(2^{-2j}\omega)|^2 = 1 \quad \text{for } |\omega| \geq \frac{1}{8}; \tag{3}$$

(ii) $\hat{\psi}_2 \in C^\infty(\widehat{\mathbb{R}})$, $\operatorname{supp} \hat{\psi}_2 \subset [-1, 1]$ and

$$|\hat{\psi}_2(\omega - 1)|^2 + |\hat{\psi}_2(\omega)|^2 + |\hat{\psi}_2(\omega + 1)|^2 = 1 \quad \text{for } |\omega| \leq 1. \tag{4}$$

It was shown in Ref. 14 that there are several examples of functions satisfying these properties. Hence, the shearlet elements $\psi_{j,\ell,k}^{(d)}$ form a collection of well-localized waveforms, at various scales dependent on $j \in \mathbb{Z}$, with orientations controlled by the two–dimensional index $\ell = (\ell_1, \ell_2) \in \mathbb{Z}^2$ and spatial location $k \in \mathbb{Z}^3$. As illustrated in Fig. 1, the shearlet supports become increasingly more elongated at finer scales.

Using the assumptions above, a direct computations shows that

$$\sum_{j \geq 0} \sum_{\ell_1 = -2^j}^{2^j} \sum_{\ell_2 = -2^j}^{2^j} |\hat{\psi}^{(1)}(\xi\, A^{-j} B_{-\ell})|^2 = \sum_{j \geq 0} \sum_{\ell_1 = -2^j}^{2^j} \sum_{\ell_2 = -2^j}^{2^j} |\hat{\psi}_1(2^{-2j}\, \xi_1)|^2\, |\hat{\psi}_2(2^j\, \frac{\xi_2}{\xi_1} - \ell_1)|^2\, |\hat{\psi}_2(2^j\, \frac{\xi_3}{\xi_1} - \ell_2)|^2 = 1,$$

for $(\xi_1, \xi_2, \xi_3) \in \mathcal{D}_{C_1}$, where $\mathcal{D}_{C_1} = \{(\xi_1, \xi_2, \xi_3) \in \widehat{\mathbb{R}}^3 : |\xi_1| \geq \frac{1}{8}, |\frac{\xi_2}{\xi_1}| \leq 1, |\frac{\xi_3}{\xi_1}| \leq 1\}$. That is, the 3D shearlet system associated with the index $d = 1$ is a Parseval frame for the subspace of $L^2(\mathbb{R}^3)$ of functions whose frequency support is contained in pyramidal region $\mathcal{D}_{C_1}$. A similar property holds for the shearlet systems associated with $d = 2, 3$, which will generate Parseval frames for the subspaces of $L^2(\mathbb{R}^3)$ of functions whose frequency support is contained in pyramidal regions $\mathcal{D}_{C_2} = \{(\xi_1, \xi_2, \xi_3) \in \widehat{\mathbb{R}}^3 : |\xi_2| \geq \frac{1}{8}, |\frac{\xi_1}{\xi_2}| \leq 1, |\frac{\xi_3}{\xi_2}| \leq 1\}$ and $\mathcal{D}_{C_3} = \{(\xi_1, \xi_2, \xi_3) \in \widehat{\mathbb{R}}^3 : |\xi_3| \geq \frac{1}{8}, |\frac{\xi_1}{\xi_3}| \leq 1, |\frac{\xi_2}{\xi_3}| \leq 1\}$, respectively. The pyramidal regions $\mathcal{D}_{C_d}$, $d = 1, 2, 3$, in the Fourier domain, are illustrates in Fig. 2.

To obtain a Parseval frame for the whole space $L^2(\mathbb{R}^3)$, it will be sufficient to combine the shearlet system described above with a Parseval frame $\{\phi_k : k \in \mathbb{Z}^3\}$ for the subspace of $L^2(\mathbb{R}^3)$ of functions whose frequency support is contained in the low frequency region $V = \{(\xi_1, \xi_2, \xi_3) \in \widehat{\mathbb{R}}^3 : |\xi_1|, |\xi_2|, |\xi_3| \leq \frac{1}{8}\}$. Notice that, in order to avoid the overlap of the elements of the shearlet system associated with the various pyramidal regions, they are projected into the corresponding pyramidal regions. Namely, for $d = 1, 2, 3$ we define

$$\widetilde{\psi}_{j,\ell_1,\ell_2,k,d} = \begin{cases} \psi_{j,\ell_1,\ell_2,k}^{(d)} & \text{if } |\ell_1| \,\&\, |\ell_2| < 2^j \\ P_{C_d} \psi_{j,\ell_1,\ell_2,k}^{(d)} & \text{if } \ell_1 = \pm 2^j \text{ or } \ell_2 = \pm 2^j, \end{cases}$$

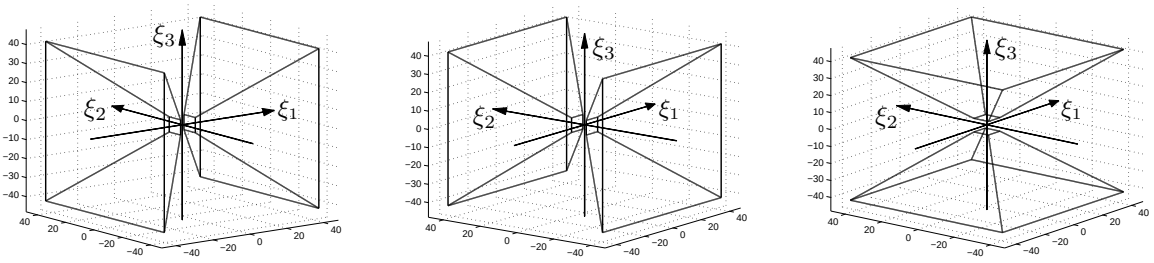


Figure 2. From left to right, the figure illustrates the pyramidal regions $\mathcal{D}_{C_1}$, $\mathcal{D}_{C_2}$ and $\mathcal{D}_{C_3}$ in the frequency space $\widehat{\mathbb{R}}^3$.

where P_{C_d} is the orthogonal projection into the subspaces of $L^2(\mathbb{R}^3)$ of functions whose frequency support is contained in pyramidal regions $\mathcal{D}_{C_d}$. This shows that the elements of the shearlet system whose frequency support overlap the boundary of the pyramidal regions get truncated. For convenience of notation, we denote the coarse scale system as $\{\widetilde{\psi}_{-1,k} = P_V\phi_k : k \in \mathbb{Z}^3\}$. Thus, we have the following result

THEOREM 2.1. *Let $\mathcal{M} = \mathcal{M}_f \cup \mathcal{M}_c$ be the indices associated with the fine-scale and the coarse scale shearlets, respectively, given by*

- $\mathcal{M}_f = \{\mu = (j, \ell_1, \ell_2, k, d) : \ j \geq 0, |\ell_1|, |\ell_2| \leq 2^j, k \in \mathbb{Z}^3, d = 1, 2, 3\}$ *(fine-scale shearlets)*
- $\mathcal{M}_c = \{\mu = (j, k) : j = -1, k \in \mathbb{Z}^3\}$ *(coarse-scale shearlets).*

The 3D system of shearlets $\{\widetilde{\psi}_\mu, \ \mu \in \mathcal{M}\}$ is a Parseval frame of $L^2(\mathbb{R}^3)$; that is, for any $f \in L^2(\mathbb{R}^3)$,

$$\sum_{\mu\in\mathcal{M}} |\langle f, \widetilde{\psi}_\mu\rangle|^2 = \|f\|^2.$$

The mapping from $f \in L^2(\mathbb{R}^3)$ into the elements $\langle f, \widetilde{\psi}_\mu\rangle$, $\mu \in \mathcal{M}$, is called the *3D shearlet transform.*

As mentioned above, by a result in Ref. 15, 16, the 3D Parseval frame of shearlets $\{\widetilde{\psi}_\mu, \ \mu \in \mathcal{M}\}$ achieves the essentially optimal approximation rate (1) for functions of 3 variables which are C^2 regular away from discontinuities along C^2 surfaces.

3. 3D DISCRETE SHEARLET TRANSFORM

We now derive a digital implementation of the 3D shearlet transform introduced above. Our approach follows essentially the same architecture as the algorithm of the 2D Discrete Shearlet Transform in Ref. 12, which consists of the cascade of a multiscale decomposition with a stage of directional filtering, based on the use of Pseudopolar FFT. The main novelty of the 3D approach is the design of a new directional filtering stage.

As a first step, we will express the elements of the shearlet system in a form that is more convenient to derive an algorithmic implementation of the shearlet transform. For $\xi = (\xi_1, \xi_2, \xi_3)$ in $\widehat{\mathbb{R}}^3$, $j \geq 0$, and $-2^j \leq \ell_1, \ell_2 \leq 2^j$, let the directional windowing function for the pyramidal region $\mathcal{D}_{C_1}$ be defined by

$$\widehat{W}^{(1)}_{j,\ell}(\xi) = \begin{cases} \overline{\hat\psi_2(2^j\frac{\xi_2}{\xi_1} - \ell_1)\hat\psi_2(2^j\frac{\xi_3}{\xi_1} - \ell_2)}\mathcal{X}_{\mathcal{D}_{C_1}}(\xi) + \overline{\hat\psi_2(2^j\frac{\xi_1}{\xi_2} - \ell_1)\hat\psi_2(2^j\frac{\xi_3}{\xi_2} - \ell_2)}\mathcal{X}_{\mathcal{D}_{C_2}}(\xi) & \\ \quad + \overline{\hat\psi_2(2^j\frac{\xi_1}{\xi_3} - \ell_1)\hat\psi_2(2^j\frac{\xi_2}{\xi_3} - \ell_2)}\mathcal{X}_{\mathcal{D}_{C_3}}(\xi) & \text{if } (\ell_1, \ell_2) = \pm(2^j, 2^j); \\ \overline{\hat\psi_2(2^j\frac{\xi_2}{\xi_1} - \ell_1)\hat\psi_2(2^j\frac{\xi_3}{\xi_1} - \ell_2)}\mathcal{X}_{\mathcal{D}_{C_1}}(\xi) + \overline{\hat\psi_2(2^j\frac{\xi_1}{\xi_2} - \ell_1)\hat\psi_2(2^j\frac{\xi_3}{\xi_2} - \ell_2)}\mathcal{X}_{\mathcal{D}_{C_2}}(\xi) & \text{if } \ell_1 = \pm 2^j, |\ell_2| < 2^j; \\ \overline{\hat\psi_2(2^j\frac{\xi_2}{\xi_1} - \ell_1)\hat\psi_2(2^j\frac{\xi_3}{\xi_1} - \ell_2)}\mathcal{X}_{\mathcal{D}_{C_1}}(\xi) + \overline{\hat\psi_2(2^j\frac{\xi_1}{\xi_3} - \ell_1)\hat\psi_2(2^j\frac{\xi_2}{\xi_3} - \ell_2)}\mathcal{X}_{\mathcal{D}_{C_3}}(\xi) & \text{if } \ell_2 = \pm 2^j, |\ell_1| < 2^j; \\ \overline{\hat\psi_2(2^j\frac{\xi_2}{\xi_1} - \ell_1)\hat\psi_2(2^j\frac{\xi_3}{\xi_1} - \ell_2)} & \text{otherwise.} \end{cases}$$

Similarly, we define the functions $\widehat{W}_{j,\ell}^{(2)}$ and $\widehat{W}_{j,\ell}^{(3)}$ for pyramidal region $\mathcal{D}_{C_2}$ and $\mathcal{D}_{C_3}$ by swapping $(\xi_1,\xi_2,\xi_3) \leftrightarrow (\xi_2,\xi_1,\xi_3)$ and $(\xi_1,\xi_2,\xi_3) \leftrightarrow (\xi_3,\xi_2,\xi_1)$, respectively. Notice that it follows from the properties of the shearlet construction that

$$\sum_{d=1}^{3}\sum_{j\geq 0}\sum_{\ell_1=-2^j}^{2^j}\sum_{\ell_2=-2^j}^{2^j} \widehat{W}_{j,\ell}^{(d)}(\xi) = 1, \quad \text{for } |\xi_1|,|\xi_2|,|\xi_3| \geq \tfrac{1}{8}. \tag{5}$$

Using this notation, we can write the elements of the 3D shearlet system as

$$\hat{\psi}_{j,\ell,k}^{(d)} = 2^{-2j}\, V(2^{-2j}\xi)\, \widehat{W}_{j,\ell}^{(d)}(\xi)\, e^{-2\pi i \xi A_d^{-j} B_d^{-\ell} k},$$

where $\widehat{V}(\xi) = \overline{\hat{\psi}_1(\xi_1)}\mathcal{X}_{\mathcal{D}_{C_1}}(\xi) + \overline{\hat{\psi}_1(\xi_2)}\mathcal{X}_{\mathcal{D}_{C_2}}(\xi) + \overline{\hat{\psi}_1(\xi_3)}\mathcal{X}_{\mathcal{D}_{C_3}}(\xi)$. Hence, the (fine scale) 3D shearlet transform of $f \in L(\mathbb{R}^3)$ is the mapping

$$f \to \langle f, \psi_{j,\ell,k}^{(d)} \rangle \quad = \quad \int_{\mathbb{R}^3} \hat{f}(\xi)\, \widehat{V}(2^{-2j}\xi)\, \widehat{W}_{j,\ell}^{(d)}(\xi) e^{2\pi i \xi A_d^{-j} B_d^{-\ell} k}\, d\xi, \tag{6}$$

where $j \geq 0$, $\ell = (\ell_1,\ell_2)$ with $|\ell_1|,|\ell_2| \leq 2^j$, $k \in \mathbb{Z}^3$ and $d = 1,2,3$. This expression shows that, for j,ℓ,k and d fixed, the shearlet transform of f can be computed using the following steps:

1. In the frequency domain, compute the j-th subband decomposition of f as $\hat{f}_j(\xi) = \hat{f}(\xi)\, \widehat{V}(2^{-2j}\xi)$.
2. Compute the (j,ℓ,d)-th directional subband decomposition of f as $\hat{f}_{j,\ell,d}(\xi) = \hat{f}_j(\xi)\, \widehat{W}_{j,\ell}^{(d)}(\xi)$.
3. Compute the inverse Fourier transform. This step can be represented as a convolution of the j-th subband decomposition of f and the directional filter $W_{j,\ell}^{(d)}$, that is, $\langle f, \psi_{j,\ell,k}^{(d)} \rangle = f_j * W_{j,\ell}^{(d)}(A_d^{-j} B_d^{-\ell} k)$.

Hence, the shearlet transform of f can be described as a cascade of subband decomposition and directional filtering, as illustrated in Fig. 3.

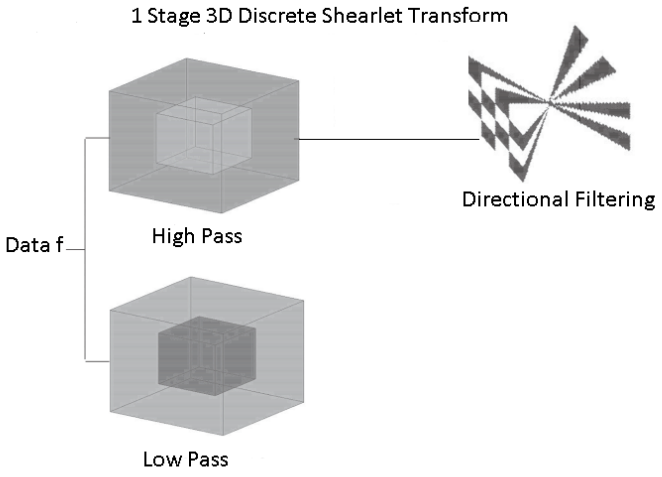


Figure 3. *The figure shows one stage of the 3D Discrete Shearlet Transform algorithm. Input data are decomposed into a low pass and high pass components. Next, the high pass component is decomposed into several directional subbands, whose number increases quadratically at finer resolution levels.*

3.1 3D DST Algorithm

Our new numerical algorithm for computing the digital values of the 3D shearlet transform, which is called 3D DST algorithm, follows closely the 3 steps indicated above.

Before describing the numerical algorithm, let us recall that a digital 3D function f is an element of $\ell^2(\mathbb{Z}_N^3)$, where $N \in \mathbb{N}$, that is, it consists of a finite array of values $\{f[n_1, n_2, n_3] : n_1, n_2, n_2 = 0, 2, \ldots, N-1\}$. Here and in the following, we adopt the convention that a bracket $[\cdot, \cdot, \cdot]$ denotes an array of indices whereas the standard parenthesis $(\cdot, \cdot, \cdot)$ denotes a function evaluation. Given a $3D$ digital function $f \in \ell^2(\mathbb{Z}_N^3)$, its Discrete Fourier Transform is given by:

$$\hat{f}[k_1, k_2, k_3] = \frac{1}{N^{3/2}} \sum_{n_1, n_2, n_3 = 0}^{N-1} f[n_1, n_2, n_3]\, e^{-2\pi i (\frac{n_1}{N} k_1 + \frac{n_2}{N} k_2 + \frac{n_3}{N} k_3)}, \qquad -\frac{N}{2} \le k_1, k_2, k_3 < \frac{N}{2}.$$

We shall interpret the numbers $\hat{f}[k_1, k_2, k_3]$ as samples $\hat{f}[k_1, k_2, k_3] = \hat{f}(k_1, k_2, k_3)$ from the trigonometric polynomial

$$\hat{f}(\xi_1, \xi_2, \xi_3) = \frac{1}{N^{3/2}} \sum_{n_1, n_2, n_3 = 0}^{N-1} f[n_1, n_2, n_3]\, e^{(-2\pi i (\frac{n_1}{N} \xi_1 + \frac{n_2}{N} \xi_2 + \frac{n_3}{N} \xi_3))}.$$

We can now proceed with the description of the implementation of the 3D DST algorithm.

First, to calculate $\hat{f}_j(\xi)$ in the digital domain, we perform the computation in the DFT domain as the product of the DFT of f and the DFT of the filters v_j corresponding to $\widehat{V}(2^{-2j}\cdot)$. This step can be implemented using the Laplacian pyramid algorithm,[2] which results in the decomposition the input signal $f \in \ell^2(\mathbb{Z}_N^3)$ into a low-pass and high-pass components, as shown in Fig. 3.

Next, to compute the directional components $\hat{f}_{j,\ell,d}$ of $\hat{f}$, the j-th subband component of f is resampled onto a psedospherical grid and a two-dimensional band-pass filter is applied. The *pseudospherical grid* is the 3D extension of the 2D pseudopolar grid and is parametrized by planes going through the origin and their slopes. That is, the pseudo-spherical coordinates $(u, v, w) \in \mathbb{R}^3$ are given by

$$(u, v, w) = \begin{cases} (\xi_1, \frac{\xi_2}{\xi_1}, \frac{\xi_3}{\xi_1}) & \text{if } (\xi_1, \xi_2, \xi_3) \in \mathcal{D}_{C_1}, \\ (\xi_2, \frac{\xi_2}{\xi_1}, \frac{\xi_3}{\xi_1}) & \text{if } (\xi_1, \xi_2, \xi_3) \in \mathcal{D}_{C_2}, \\ (\xi_3, \frac{\xi_1}{\xi_3}, \frac{\xi_2}{\xi_3}) & \text{if } (\xi_1, \xi_2, \xi_3) \in \mathcal{D}_{C_3}. \end{cases}$$

Using this change of variables, we have that

$$\hat{f}_{j,\ell,d}(\xi) = \hat{g}_j(u, v, w)\, \widehat{W}^{(d)}(2^j v - \ell_1)\, \widehat{W}^{(d)}(2^j w - \ell_2), \tag{7}$$

where $\hat{g}_j(u, v, w)$ is the function $\hat{f}_j(\xi)$, after the change of variables, and $\widehat{W}^{(d)} = \widehat{W}_{0,0}^{(d)}$. This expression shows that the different directional subbands are obtained by simply translating the window function $\widehat{W}$ in the pseudo-spherical domain. Thus, the discrete samples $g_j[n_1, n_2, n_3] = g_j(n_1, n_2, n_3)$ are the values of the DFT of $f_j[n_1, n_2, n_3]$ on the pseudo-spherical grid. This can be computed by direct reassignment or adapting the Pseudo-polar DFT algorithm.[1]

Let $\{w_{j,\ell_1,\ell_2}^{(d)}[n_2, n_3] : n_2, n_3 \in \mathbb{Z}\}$ be the sequence whose DFT gives the discrete samples of the window function $\widehat{W}^{(d)}(2^j v - \ell_1)\, \widehat{W}^{(d)}(2^j w - \ell_2)$; that is, $\hat{w}_{j,\ell_1,\ell_2}^{(d)}[k_2, k_3] = \widehat{W}^{(d)}(2^j k_2 - \ell_1)\, \widehat{W}^{(d)}(2^j k_3 - \ell_2)$. Then, for fixed $k_1 \in \mathbb{Z}$, we have

$$\mathcal{F}_2 \left(\mathcal{F}_2^{-1} (\hat{g}_j) * w_{j,\ell_1,\ell_2}^{(d)} \right) [k_1, k_2, k_3] = \hat{g}_j[k_1, k_2, k_3]\, \hat{w}_{j,\ell_1,\ell_2}^{(d)}[k_2, k_3], \tag{8}$$

where $\mathcal{F}_2$ is two dimensional DFT, defined as

$$\mathcal{F}_2(f)[k_2, k_3] = \frac{1}{N} \sum_{n_2, n_3 = 0}^{N-1} f[n_2, n_3]\, e^{(-2\pi i (\frac{n_2}{N} k_2 + \frac{n_3}{N} k_3))}, \qquad -\frac{N}{2} \le k_2, k_3 < \frac{N}{2}.$$

Thus, equation (8) gives the algorithmic procedure for computing the discrete samples of the right hand side of (7). That is, the 3D shearlet coefficients (6) can be calculated from equation (8) by computing the inverse psedo-spherical DFT of by directly re-assembling the Cartesian sampled values and applying the inverse 3-dimensional DFT.

An alternative approach consists in mapping the filters back in Cartesian domain and then perform a convolution with band-passed data. Specifically if ϕ_P is the mapping from Cartesian domain into the pseudo-spherical domain then the 3D shearlet coefficients in the Fourier domain can be expressed as

$$\phi_P^{-1}\left(\hat{g}_j[k_1,k_2,k_3]\,\hat{w}^{(d)}_{j,\ell_1,\ell_2}[k_2,k_3]\right).$$

Following the approach in Ref. 12, this can be expressed as

$$\phi_P^{-1}\left(\hat{g}_j[k_1,k_2,k_3]\right)\,\phi_P^{-1}\left(\hat{\delta}_P[k_1,k_2,k_3]\hat{w}^{(d)}_{j,\ell_1,\ell_2}[k_2,k_3]\right),$$

where $\hat{\delta}_P$ is the DFT of the (discrete) delta distribution in the pseudo-spherical grid. Thus the 3D shearlet coefficients in the Fourier domain can be expressed as

$$\hat{f}_j[k_1,k_2,k_3]\,\hat{h}^{(d)}_{j,\ell_1,\ell_2}[k_1,k_2,k_3],$$

where $\hat{h}^{(d)}_{j,\ell_1,\ell_2} = \phi_P^{-1}\left(\hat{\delta}_P[k_1,k_2,k_3]\hat{w}^{(d)}_{j,\ell_1,\ell_2}[k_2,k_3]\right)$. Notice that the new filters $h^{(d)}_{j,\ell_1,\ell_2}$ are not obtained by a simple change of variables, but by applying a resampling which converts the pseudo-spherical grid to a Cartesian grid. Although these filters are not compactly supported, they can be implemented with a matrix representation that is smaller than the size of the data f, hence allowing to implement the computation of the 3D DST using a convolution in space domain.

The 3D DST algorithm runs in $O(N^3\log(N))$ operations.

3.2 Implementation issues.

There are many possible choices for the directional filters $W^{(d)}_{j,\ell}$ which satisfy the required condition (5). For this paper, we designed Meyer-like filters by adapting the approach used in Ref. 12 for the 2D setting. As mentioned above, by taking the inverse DFT, it is possible to implement these filters using matrix representations of size L^3 with $L \ll N$, where N^3 is the data size. In the the numerical experiment considered below, we have chosen $L = 24$, which appears to be a good compromise between localization and computation times. Finally, for the number of directional bands, our algorithm allows us to choose different number of directional bands in each pyramidal region.

4. NUMERICAL EXPERIMENTS

The ability of the 3D shearlet transform to deal with geometric information efficiently and its sparsity properties have the potential to produce significant improvement in many 3D data processing applications. As an example of these applications, we have developed an algorithm for video denoising.

In our model, we only considered the situation of zero-mean additive white Gaussian noise, which offers a good model for many practical situations. Hence, we will suppose that, for a given video f, we observe

$$y = f + n,$$

where n is Gaussian white noise with zero mean and standard deviation σ.

It is well known that the ability to sparsely represent data is very useful in decorrelating the signal from the noise. This notion has been precisely formalized in the classical *wavelet shrinkage* approach by Donoho and Johnstone,[8–10] which has lead to many successful denoising algorithms. In the following, we apply a simple video denoising strategy based on hard thresholding. Although this is a rather crude form of thresholding and more sophisticated methods are available, still hard thresholding is a good indication of the potential of a transform in denoising. Hence, we attempt to recover the video f from the observed data y as follows.

1. We compute the 3D shearlet decomposition of y as $y = \sum_\mu \langle y, \widetilde{\psi}_\mu \rangle \, \widetilde{\psi}_\mu$.

2. We set to zero the coefficients $c_\mu(y) = \langle y, \widetilde{\psi}_\mu \rangle$ such that $|c_\mu(y)| < T_\mu$, where T_μ depends on the noise level.

3. We obtain an approximation $\tilde{f}$ of f as

$$\tilde{f} = \sum_\mu c^*_\mu(y) \, \widetilde{\psi}_\mu, \quad \text{where } c^*_\mu(y) = c_\mu(y) \text{ if } |c_\mu(y)| \geq T_\mu; c^*_\mu(y) = 0 \text{ otherwise.}$$

For the choice of the threshold parameter, we adopt the classical BayesShrink method,[5] consisting in choosing

$$T_{j,\ell} = \frac{\sigma^2}{\sigma_{j,\ell}},$$

where $\sigma_{j,\ell}$ is the standard deviation of the shearlet coefficients in the (j,ℓ)-th subband. For the 3D discrete shearlet decomposition, in all our tests we have applied a 3-level decomposition according to the algorithm described above. For the number of directional bands, we have chosen $n = 16, 16, 64$ (from the coarsest to the fines level) in each of the pyramidal region. Even though this does not exactly respect the rule canonical choice ($n = 4, 16, 64$) prescribed by the continuous model, we found that increasing the number of directional subbands at the coarser level produces some improvement in the denoising performance. Notice that, in our numerical implementation, downsampling occurs only at the bandpass level, and there is no anisotropic downsampling. Thus, the numerical implementation of the 3D DST which we found most effective in the denoising algorithm is highly redundant. Specifically, for data set of size N^3, a 3-level 3D DST decomposition produces $3 * \left(64 * N^3 + 16 * (\frac{2}{3}N)^3 + 16 * (\frac{2}{6}N)^3\right) + (\frac{2}{6}N)^3 \approx 208 * N^3$ coefficients.

The 3D shearlet-based thresholding algorithm was tested on 3 video sequences, called *mobile*, *coastguard* and *tempete*, for various values of the standard deviation of the noise. All these video sequences, which have been resized to $192 \times 192 \times 192$, can be downloaded from the website `http://www.cipr.rpi.edu`.

For a baseline comparison, we tested the performance of the shearlet-based denoising algorithm (denoted by 3DSHEAR) against the following state-of-the-art algorithms: 3D Curvelets (denoted by 3DCURV, cf. Ref. 3), Undecimated Discrete Wavelet Transform (denoted by UDWT, based on *symlet* of length 16), Dual Tree Wavelet Transform (denoted by DTWT, cf. Ref. 22) and Surfacelets (denoted by SURF, cf. Ref. 21). We also compared against the 2D discrete shearlet transform (denoted by 2DSHEAR), which was applied frame by frame, in order to illustrate the benefit of using a 3D transform, rather than a 2D transform acting on each frame.

As a performance measure, we used the standard *peak signal-to-noise ratio* (PSNR), measure in decibel(dB), which is defined by

$$\text{PSNR} = 20 \log_{10} \frac{255N}{\|f - \tilde{f}\|_F},$$

where $\|\cdot\|_F$ is the Frobenius norm and f is an array of size $N \times N \times N$.

The performance of the shearlet-based denoising algorithm 3DSHEAR relative to the other algorithms is shown Table I, which illustrates the denoising performance, measured in PSNR, for various video sequences and noise levels. The noise standard deviation is taken at values 30, 40 and 50, corresponding to 18.62 dB, 16.12 dB and 14.18 dB, respectively, in PSNR. Only the PSNR values are shown in the table. Notice that performance values for the algorithms 3DCURV, UDWT and DTWT are taken from Ref. 21.

The data in Table I (with the numbers in bold indicating the best performance) show that the 3D Discrete Shearlet Denoising Algorithm 3DSHEAR is highly competitive against both traditional and other state-of-the-art video denoising algorithm. In particular, 3DSHEAR consistently outperforms the curvelet-based routine 3DCURV, the wavelet-based routines UDWT and DTWT and the 2D shearlet-based algorithm. 3DSHEAR also outperforms or is essentially equivalent to the surfacelets-based denoising algorithm in all cases we tested, except for one case, namely the mobile video sequence when the noise level has PSNR = 14.18 dB (corresponding to noise with standard deviation equal to 50).

Table I: Video denoising performance using different video sequences.

PSNR (dB)	Mobile			Coastguard			Tempete		
Noise	18.62	16.12	14.18	18.62	16.12	14.18	18.62	16.12	14.18
3DCURV	23.54	23.19	22.86	25.05	24.64	24.29			
UDWT	24.02	22.99	22.23	25.95	24.95	24.2			
DTWT	24.56	23.43	22.58	26.06	25.01	24.22			
SURF	28.39	**27.18**	**26.27**	26.82	25.87	**25.15**	24.2	23.26	22.61
3DSHEAR	**28.68**	27.15	25.97	**27.36**	**26.10**	25.12	**25.24**	**23.97**	**22.81**
2DSHEAR	25.97	24.40	23.20	25.20	23.82	22.74	22.89	21.63	20.75
DWT	24.93	23.94	23.03	24.34	23.44	22.57	22.09	21.5	20.92

In Fig. 4 and 5, we compare the performance of the various video denoising routines on a typical frame extracted from the denoised video sequence. Although more subjective in nature, the figures shows that the visual quality of the shearlet-denoised frame is also superior.

The different algorithms we have compared are based on different design and computational strategies, which play an essential in their differences in performance. One of the relevant factors is the redundancy of the transform, which accounts for the worse performance of 3DSHEAR in terms of running times. This is illustrated in the Table II, which compares the running times for different 3D transforms, applied to a data set of size 193^3; all routines were run using the same system which is based on an Intel CPU 2.93GHz.

Table II: Comparison of running times for different 3D transforms.

Algorithm	Running time (data size: 192^3)
SURF	34 sec
3DSHEAR	263 sec
2DSHEAR	154 sec
3D DWT	7.5 sec

ACKNOWLEDGMENTS

The authors acknowledge support from NSF grants DMS 1005799 and DMS 1008900.

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 Journal on Scientific Computing **30**(2) (2008), 764–784.

[2] P. J. Burt, E. H. Adelson, *The Laplacian pyramid as a compact image code*, IEEE Trans. Commun. **31** (4) (1983), 532–540.

[3] E. J. Candès, L. Demanet, D. L. Donoho and L. Ying, *Fast discrete curvelet transforms*, SIAM Multiscale Model. Simul. **5** (2006), 861–899.

[4] E. J. Candès and D. L. Donoho, *New tight frames of curvelets and optimal representations of objects with C^2 singularities*, Comm. Pure Appl. Math., **57** (2004), 219–266.

[5] G. Chang, B. Yu and M. Vetterli, *Adaptive Wavelet Thresholding for Image Denoising and Compression*, IEEE Trans. Image Processing, **9** (2000), 1532–1546.

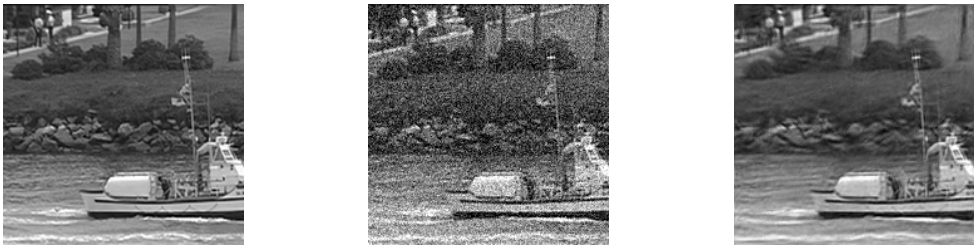
[6] F. Colonna, G. Easley, K. Guo, and D. Labate, *Radon Transform Inversion using the Shearlet Representation*, Appl. Comput. Harmon. Anal. **29**(2)2 (2010), 232–250.

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

Original frame | Noisy frame | 3DSHEAR

SURF | 2DSHEAR | DWT

Figure 4. Video Denoising of Mobile Video Sequence. The figure illustrates the denoising performance on a representative frame of the video sequence using various denoising routines. Starting from the top left: original frame, noisy frame (PSNR=18.62 dB), denoised frame using 3DSHEAR (PSNR=**28.68 dB**), SURF (PSNR=28.39 dB), 2DSHEAR (PSNR=25.97 dB) and DWT (PSNR=24.93 dB).



Original frame | Noisy frame | 3DSHEAR

SURF | 2DSHEAR | DWT

Figure 5. Video Denoising of Coast Guard Video Sequence. The figure illustrates the denoising performance on a representative frame of the video sequence using various denoising routines. Starting from the top left: original frame, noisy frame (PSNR=18.62 dB), denoised frame using 3DSHEAR (PSNR=**27.36 dB**), SURF (PSNR=26.82 dB), 2DSHEAR (PSNR=25.20 dB), DWT (PSNR=24.34 dB).

[8] D. L. Donoho, *Sparse components of images and optimal atomic decomposition*, Constr. Approx. **17**(2001), 353–382.

[9] D. L. Donoho and I. M. Johnstone, *Ideal spatial adaptation by wavelet shrinkage*, Biometrika **81**(1994), 425-455.

[10] D. L. Donoho, *Denoising via soft thresholding*, IEEE Transactions on Information Theory, **41**(1995), 613–627.

[11] G. R. Easley, D. Labate, F. Colonna, *Shearlet-Based Total Variation Diffusion for Denoising*, IEEE Trans. Image Proc. **18** (2) (2009), 260–268.

[12] G. R. Easley, D. Labate, and W. Lim, *Sparse Directional Image Representations using the Discrete Shearlet Transform*, Appl. Comput. Harmon. Anal., **25** (1) (2008), 25–46.

[13] K. Guo, G. Kutyniok, and D. Labate, *Sparse Multidimensional Representations using Anisotropic Dilation and Shear Operators,* in: Wavelets and Splines, G. Chen and M. Lai (eds.), Nashboro Press, Nashville, TN (2006), pp. 189–201.

[14] K. Guo, D. Labate, *Optimally Sparse Multidimensional Representation using Shearlets*, SIAM J. Math. Anal., **9** (2007), 298–318

[15] K. Guo, and D. Labate, *Optimally Sparse 3D Approximations using Shearlet Representations*, Electronic Research Announcements in Mathematical Sciences, **17** (2010), 126–138.

[16] K. Guo, and D. Labate, *Optimally Sparse Representations of 3D Data with C^2 Surface Singularities using Parseval Frames of Shearlets*, preprint, (2010).

[17] K. Guo, W.-Q Lim, D. Labate, G. Weiss and E. Wilson, *Wavelets with composite dilations*, Electron. Res. Announc. Amer. Math. Soc., **10** (2004), 78–87

[18] K. Guo, W-Q. Lim, D. Labate, G. Weiss and E. Wilson, *Wavelets with composite dilations and their MRA properties*, Appl. Computat. Harmon. Anal. **20** (2006), 231–249.

[19] J. Lemvig, *Optimally Sparse Approximations of Functions in $L^2(\mathbb{R}^3)$ with C^α Singularities using Shearlet*, Oberwolfach Report **44** (2010)

[20] W. Lim, The Discrete Shearlet Transform: A New Directional Transform and Compactly Supported Shearlet Frames, IEEE Trans. Image Process **19**(5) (2010), 1166–1180.

[21] Y. Lu, and M. N. Do, *Multidimensional Directional Filter Banks and Surfacelets* IEEE Transactions on Image Processing, **16**(4) (2007), 918–931.

[22] I. Selesnick, *The double-density dual-tree DWT,* IEEE Trans. Signal Proc. **52**(5) (2004), 1304-1314.

[23] S. Yi, D. Labate, G. R. Easley, and H. Krim, *A Shearlet approach to Edge Analysis and Detection*, IEEE Trans. Image Process **18**(5) (2009), 929–941.