Digital Implementation of Ridgelet Packets

A.G. Flesia, H. Hel-Or
A. Averbuch, E.J. Candès
R.R. Coifman, D.L. Donoho

Abstract. The Ridgelet Packets library provides a large family of orthonormal bases for functions $f(x, y)$ in $L^2(dxdy)$ which includes orthonormal ridgelets as well as bases deriving from tilings reminiscent from the theory of wavelets and the study of oscillatory Fourier integrals. An intuitively appealing feature: many of these bases have elements whose envelope is strongly aligned along specified 'ridges' while displaying oscillatory components across the main 'ridge'.

There are two approaches to constructing ridgelet packets; the most direct is a frequency-domain viewpoint. We take a recursive dyadic partition of the polar Fourier domain into a collection of rectangular tiles of various widths and lengths. Focusing attention on each tile in turn, we take a tensor basis, using windowed sinusoids in θ times windowed sinusoids in r. There is also a Radon-domain approach to constructing ridgelet packets, which involves applying the Radon isometry and then, in the Radon plane, using wavelets in θ times wavelet packets in t, with the scales of the wavelets in the two directions carefully related.

We discuss digital implementations of the two continuum approaches, yielding many new frames for representation of digital images $I(i, j)$. These rely on two tools: the pseudopolar Fast Fourier Transform, and a pseudo Radon isometry called the normalized Slant Stack; these are described in Averbuch et al. (2001). In the Fourier approach, we mimic the continuum Fourier approach by partitioning the pseudopolar Fourier domain, building an orthonormal basis in the image space subordinate to each tile of the partition. On each rectangle of the partition, we use windowed sinusoids in θ times windowed sinusoids in r. In the Radon approach, we operate on the pseudo-Radon plane, and mimic the construction of orthonormal ridgelets, but with different scaling relationships between angular wavelets and ridge wavelets. Using wavelet packets in the ridge direction would also be possible.

Because of the wide range of possible ridgelet packet frames, the question arises: what is the best frame for a given dataset? Because of the Cartesian format of our 2-D pseudopolar domain, it is possible to apply best-basis algorithms for best anisotropic cosine packets bases; this will rapidly search among all such frames for the best possible frame according to a sparsity criterion – compare N. Bennett's 1997 Yale Thesis. This automatically finds the best ridgelet packet frame for a given dataset.

Key Words and Phrases. Wavelets. Ridge function. Ridgelets. Radon Transform. Radon Isometry. Wavelet Packets, Cosine Packets. Best basis. Gabor Analysis.

Beyond Wavelets
J. Stoeckler and G. V. Welland (Eds.), pp. 1–??.

ISBN 0-12-xxxxxx-x

1. Introduction

There is considerable scientific and technological interest in representing 2-dimensional objects $f(x, y)$ – images – in terms of oscillatory $2 - d$ waveforms with anisotropic envelopes.

For example, it is widely believed in the study of computational and biological vision that early vision depends on the analysis of image data by directionally oriented waveforms [19, 35, 20, 21]. It is also believed that representations by oriented oscillatory waveforms are important for efficient representation of real image data, [2, 30, 29], and to texture simulation [22].

These scientific and engineering interests interact with the agenda of computational harmonic analysis, where an important goal in dealing with 2-D and higher-dimensional objects is to construct representations which have a wide range of orientations, aspect ratios, locations, scales, and oscillation numbers. For example, one would like the ability to decompose an object $f(x)$, $x \in R^d$, stably and compactly, into a sum of oscillatory waveforms with anisotropic envelopes:

$$\psi_{A,b,\xi}(x) = e^{i\xi'(x-b)} \exp\{-(x-b)'A(x-b)\}.$$

If such an ability were available, it would encompass in one framework most of the benefits of wavelets, Gabor transforms, ridgelets, and many other specific decompositions.

Unfortunately, no effective, stable decomposition into general collections of such atoms is available today. The present paper does make some progress in this direction in the two-dimensional case $d = 2$.

To put our approach in context, recall that an important stage in the development of computational harmonic analysis was the generalization from wavelet bases to wavelet packet and cosine packet bases [11]. Coifman and Meyer recognized that one could develop a very intuitive picture of basis construction around the notion of tilings of the time-frequency domain, and this picture would explain how to go beyond wavelets, creating a very wide variety of orthonormal bases with various interesting properties. In their approach, one had available a general family of tilings of the time-frequency plane based on recursive dyadic partitioning, and one could build a basis corresponding to each such tiling by using some simple tools deployed within an appropriately general framework. Reducing this insight to very intuitive terms, these constructions allow bases made of waveforms which, like wavelets, were localized in time and scale, but which instead of the fixed, small number of oscillations offered by wavelets, had a variable number of oscillations, controllable by the user.

In this paper, we describe a system of orthonormal bases and frames for $L^2(dxdy)$ as well as frames for two-dimensional digital data $I(i, j)$. The system we describe, *Ridgelet Packets*, has many points in common with ridgelet analysis, while allowing for a more complex and oscillatory structure. Thus, the system contains bases where the basis elements are, like ridgelets, highly orientation-selective, but which have greater degrees of oscillation either along

or across the direction of primary orientation. We also describe algorithms for two-dimensional digital data which can adaptively construct frames which are well-adapted to sparse representation of given digital data. The tools can be viewed as a contribution towards solving the problem proposed in paragraph one above: automatically decomposing objects into compact collections of anisotropic, oscillating waveforms. They bear much the same relationship to ridgelets as do wavelet packets to wavelets.

One idea driving our construction is indicated in Figure 1 below. We illustrate the tiling underlying the orthonormal ridgelet basis, and two other tilings which we have labeled 'wavelet-like' and 'FIO'. These tilings of Fourier space are all similar in that they all involve the subdivision of the radial axis dyadically. They differ in the degree of angular subdivision that they apply. For ridgelets, the frequency band $2^j < r < 2^{j+1}$ is subdivided into 2^j tiles angularly, while for the FIO tiling, the same frequency band is divided into $2^{j/2}$ tiles, and for the wavelet-like tiling, the frequency band is divided in a constant number of tiles, irrespective of the frequency range. (Inspirations for studying tilings of these forms include: for the ridgelet tiling – Candès' thesis [8, 7]; for wavelet-like tiling – work in representation of natural images, such as Watson's article on the Cortex transform [35], and Field's articles on statistics of natural images [19, 20]; for the FIO-tiling – work applying 'Second Dyadic Decomposition' in harmonic analysis going back to Fefferman's analysis of Bochner-Riesz summability; see [32, 31].)

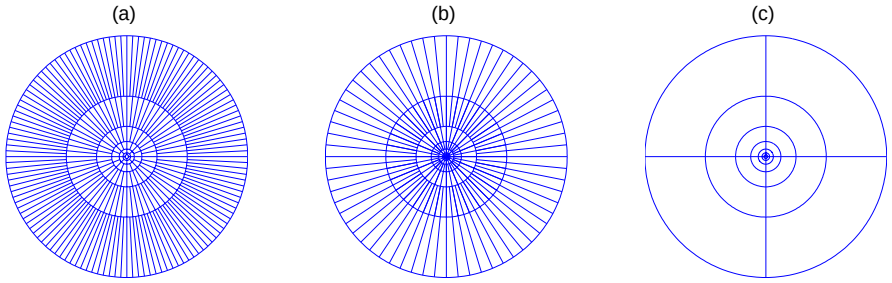

Figure 1. Three Tilings of the Frequency Domain: (a) Ridgelet tiling; (b) FIO tiling; (c) Wavelet tiling.

After tiling the frequency domain according to one of these schemes, we can create a basis where each basis element is associated with a specific frequency domain tile. If the tile is very narrow in the angular sense, then the basis element, viewed in the original spatial (x, y) domain, will exhibit a very high degree of orientational preference. If the tile is very anisotropic, the spatial envelope of the basis element, viewed in the spatial domain, will have a very

anisotropic envelope, and so a long and thin tile (long in r and thin in θ) will give rise to basis elements which are long and thin as well.

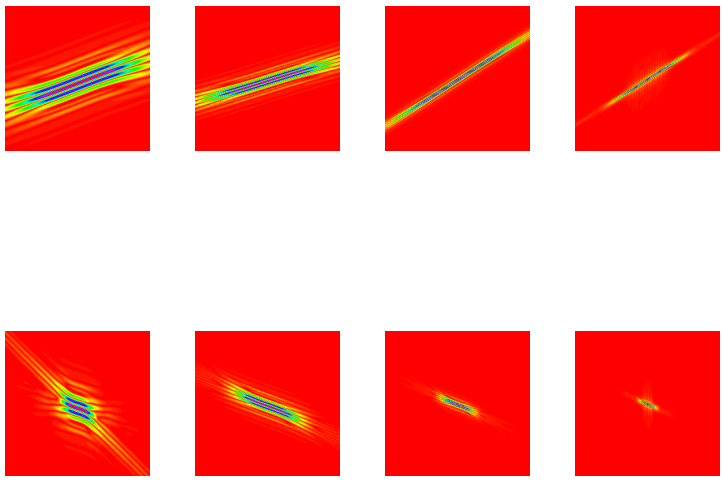

Figure 2. First Row, Left to Right: Basis Elements associated with Ridgelet tiling, from coarse to fine ridge scale, with same angular position. Second Row, Left to Right: Basis Elements associated with FIO tiling, from coarse to fine ridge scale, with same angular position. .

We have illustrated this concept in Figure 2, the first row depicts basis elements associated with the Ridgelet tiling (a) from Figure 1, that belongs to four consecutive scales in r, plotted from left to right as r increases, and same angular direction θ. The second row depicts basis elements associated with the FIO tiling (b) from Figure 1, chosen from the same scales in r as the Ridgelet basis elements. We can see that the ridgelet basis elements are generally thinner and longer than corresponding FIO-derived elements.

In discussing such tilings, we have already claimed that we can construct a basis adapted to each such tiling. To see how this can work, recall that [15] introduced orthonormal ridgelets, defined as follows. Let $(\psi_{j,k}(t) : j \in \mathbf{Z}, k \in \mathbf{Z})$ be an orthonormal basis of Meyer wavelets for $L^2(\mathbf{R})$, and let $(w^0_{i_0\ell}(\theta),\ \ell = 0, \ldots, 2^{i_0} - 1;\ w^1_{i,\ell}(\theta),\ i \geq i_0,\ \ell = 0, \ldots, 2^i - 1)$ be an orthonormal basis for $L^2[0, 2\pi)$ made of periodized Lemarié scaling functions $w^0_{i_0,\ell}$ at level i_0 and periodized Meyer wavelets $w^1_{i\ell}$ at levels $i \geq i_0$. Let $\hat{\psi}_{j,k}(\omega)$ denote the Fourier transform of $\psi_{j,k}(t)$, and define ridgelets $\rho_\lambda(x)$, $\lambda = (j, k; i, \ell, \varepsilon)$ as functions of $x \in \mathbf{R}^2$ using the frequency-domain definition

$$\hat{\rho}_\lambda(\xi) = |\xi|^{-\frac{1}{2}}(\hat{\psi}_{j,k}(|\xi|)w^\varepsilon_{i,\ell}(\theta) + \hat{\psi}_{j,k}(-|\xi|)w^\varepsilon_{i,\ell}(\theta + \pi))/2\ . \tag{1.1}$$

Here the indices run as follows: $j, k \in \mathbf{Z}$, $\ell = 0, \ldots, 2^{i-1} - 1$, $i \geq i_0$; and, if $i > i_0$, $i \geq j$. Also, if $i > i_0$ and $i > j$, then necessarily $\varepsilon = 1$. Let Λ denote the set of all such indices λ. In that article, it was shown that this collection of functions makes an orthonormal set for $L^2(\mathbf{R}^2)$.

From (1.1), and the definition of 1-D Meyer wavelets in the frequency domain as having the form $e^{it_{j,k}\lambda} w(\lambda/2^j)/2^{j/2}$ of a windowed sinusoid, we can see that the construction is essentially based on selecting a tile, using a windowed sinusoid in $r = |\xi|$ and using a wavelet in θ which peaks near that tile in θ.

In the end, the construction is based on the use of a tensor basis in (r, θ) for the space of functions localized (near) the tile.

In this paper we will carefully describe a general construction that contains the idea behind the ridgelets construction as a special case. The result will be a family of bases with a variety of interesting space-frequency localization properties. For example, by modifying the ridgelet construction using wavelet packets in the ridge direction and wavelets in the angular direction, one induces on real space elements which are oscillatory ridgelets in the sense that they have angular localization features similar to the orthonormal ridgelet family – being nearly ridge functions – while being oscillatory in the ridge direction, see Figure 3. On the other hand, by using wavelets in the ridge direction together with cosine packets in the angular direction one can produce effects which are more like brush strokes – bundles of line elements with given orientation, position, and textural cross-section.

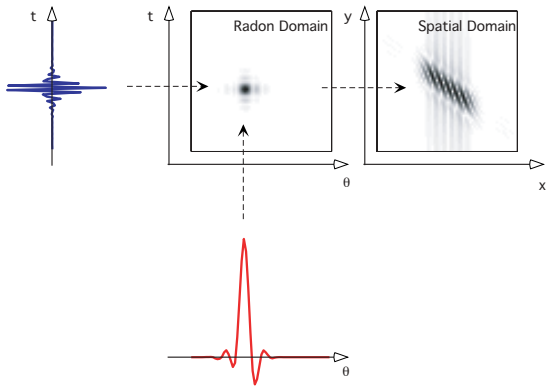

Figure 3. Construction of a Ridgelet Packet. A wavelet packet in the ridge direction and father wavelet Symmlet 8 in the angular direction are combined by tensor product, creating a waveform in the normalized Radon domain; this is backprojected into the spatial domain, creating an oriented waveform there.

We will also develop a digital realization of these ideas, based on the so-

called pseudopolar FFT. In our digital realization, we replace the polar Fourier plane (which we can view as based on concentric circles of different radii) by a pseudopolar system, based on concentric squares of different pseudo-radii. The result is that we have tilings as illustrated in Figure 4.

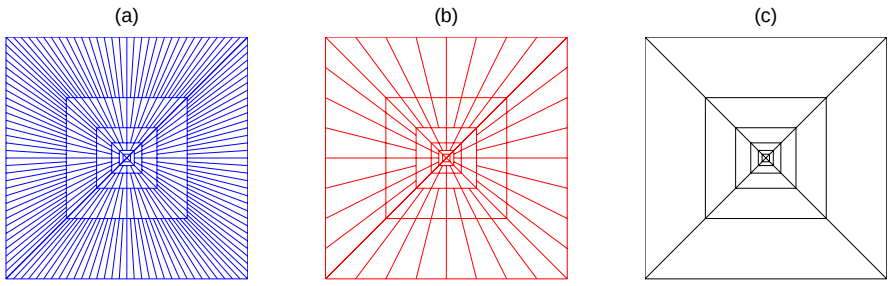

Figure 4. Three Tilings of the Digital Frequency Domain: (a) Digital Ridgelet; (b) Digital FIO; (c) Wavelet-like.

We will also describe an adaptive algorithm, based on tree pruning, that can rapidly search through a wide variety of such bases, and select one that can provide the sparsest representation of the digital data. An example of the output of this algorithm is provided in Figure 5 below.

Finally, we will describe a software package for calculating with ridgelet packets which is available as part of the Beamlab software distribution at *www.beamlab.net*.

2. Fourier Preliminaries

A function $f(x,y)$ in $L^2(dxdy)$, has a Fourier transform $\hat{f}(\xi)$ which, by Parseval/Plancherel, is isometric to it: $\|f\|_2^2 = (\frac{1}{2\pi})^2\|\hat{f}\|_2^2$ (we will typically use the term isometric in the broad sense, meaning preservation of norm up to multiplication by a fixed scalar multiple). Another isometry is given by the Cartesian-to-polar transformation $\hat{f} \mapsto \overline{F}$ defined by

$$\overline{F}(\lambda,\theta) = |\lambda|^{1/2}\hat{f}(\lambda\cdot\cos(\theta),\lambda\cdot\sin(\theta)), \tag{2.1}$$

is an isometry from $L^2(d\xi_1 d\xi_2)$ to $L^2(\mathbf{R}\times[0,2\pi))$. The composition of these two isometries, $f \mapsto \overline{F}$ maps the spatial domain into the polar Fourier domain, is again an isometry, and motivates the following

Definition 1. *By* **Polar Fourier Transform** *we mean the mapping* $\tilde{F} = \Pi(f)$ *defined by*

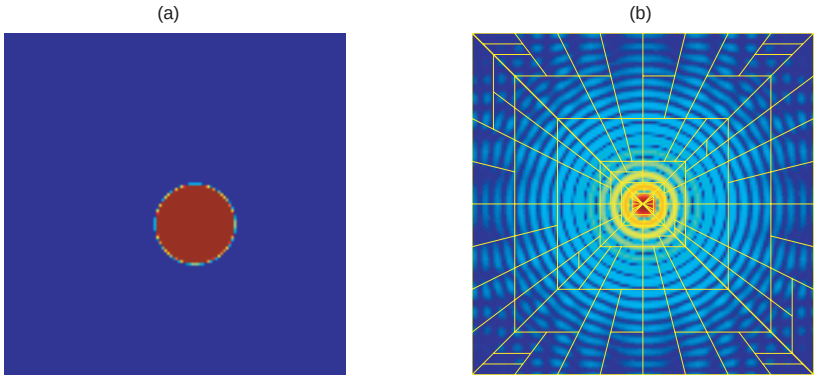

Figure 5. Disc-shaped object and Adapted Tiling. (a) Disc-shaped object. (b) Adaptive Tiling of the Fourier Plane, overlaid on the Fourier Transform of the disc. It is interesting to note that owing to the Cartesian structure of the data, the Fourier Transform of the disc does not have radial symmetry near the corners. This may explain the 'unexpected' high-frequency splits..

$$\tilde{F}(\lambda, \theta) = \hat{f}(\lambda \cdot \cos(\theta), \lambda \cdot \sin(\theta)),$$

whereas the **Polar Fourier Isometry** $\overline{F} = \overline{\Pi}(f)$ *is defined by (2.1).*

Suppose now that we take a tiling of the Fourier domain, such as one of the tilings of the introduction, and consider its equivalent in polar coordinates, as illustrated in Figure 6.

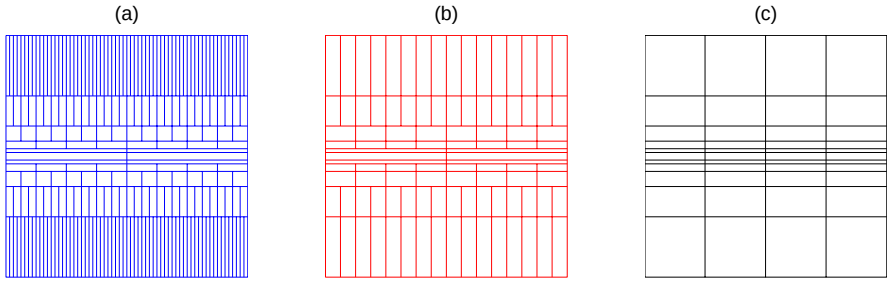

Figure 6. Tilings of Figure 1 Translated into Polar Form.

We then construct an orthonormal basis for functions on the polar strip which is the 'gluing together' of individual orthonormal bases for functions (nearly) localized to the individual tiles. Finally, we invert the Polar Fourier Isometry, producing a basis for $L^2(dx_1 dx_2)$.

It is well understood how to construct interesting bases based on tilings in dimension 1 [11, 36]. If we want to divide the frequency axis into intervals $\{I\}$ and then construct bases for the space of functions localized near those intervals, we construct smooth windows w_I localized near those intervals and obeying special properties and multiply those by sinusoids $\phi_{k,I}$ producing windowed sinusoids $w_I(t)\phi_{k,I}(t)$.

One can adapt this idea to the two-dimensional case, produce bases associated with tilings of the polar strip by a tensor product construction. We use a similar cosine packets idea in θ, partitioning the interval $[0, 2\pi)$ into a sequence of intervals $\{J\}$ and consider the tensor cosine packet as the candidate basis element

$$C_{J,\ell,I,k}(\theta, \lambda) = w_J(\theta)\phi_{\ell,J}(\theta) \cdot w_I(\lambda)\phi_{k,I}(\lambda).$$

The collection of all such elements gives (if the windows and sinusoids are set up appropriately) an orthobasis for the polar frequency domain $L^2(\mathbf{R} \times [0, 2\pi))$.

We would next attempt to use this basis to induce a frame in the spatial domain. We might be tempted to identify each $C_{J,m,I,k}$ as the image, under polar Fourier isometry, of an object $\gamma_{J,m,I,k}$ in $L^2(dx_1 dx_2)$; and then to invert the isometry, getting $\gamma_{J,m,I,k} = \overline{\Pi}^*(C_{J,m,I,k})$. The inspiration is clear. Several details are less clear:

- Can this procedure really be carried out: is such a tensor product really in the range of the polar Fourier isometry? Strictly speaking, the answer is no; some modification is required. The information in the polar Fourier strip is redundant:

  $$\overline{F}(\lambda, \theta) = \overline{F}(-\lambda, \theta + \pi);$$

  and if, in addition f is real-valued,

  $$\overline{F}(\lambda, \theta) = \overline{F}(\lambda, \theta + \pi)^*.$$

  The basis elements $C_{J,m,I,k}(\theta, \lambda)$ will have to be modified to obey these constraints. A suitable modification is:

  $$\tilde{C}_{J,m,I,k}(\theta, \lambda) = (C_{J,m,I,k}(\theta, \lambda) + C_{J,m,I,k}(\theta + \pi, -\lambda))/2$$

- Even after suitable modifications, does this lead to an interpretable basis and interpretable basis elements? Candès and Donoho, in preliminary calculations, have some theoretical insights on this matter. The digital implementations described later below provide an empirical tool for studying the resulting bases.

- The above construction does not correspond to the ortho-ridgelet definition (1.1); so apparently there is more than one way to operationalize a given set of tilings.
- There is no well-defined polar Fourier transform for digital data; it is unclear whether the ideas underling this construction have serious implications for digital data.

Because of these observations, it will turn out that the polar Fourier viewpoint tells only part of the story. In broad terms, there are two viewpoints, one in the polar Frequency domain and one in the (closely related) Radon domain. It will be very instructive in the coming few sections to develop a Radon viewpoint, which provides full or partial answers to all the open issues above.

3. Radon Preliminaries

For a smooth function $f(x) = f(x_1, x_2)$ of rapid decay, let Rf denote the Radon transform of f, the integral along a line $\mathcal{L}_{(\theta,t)}$, expressed using the Dirac mass δ as

$$(Rf)(t,\theta) = \int f(x)\delta(x_1\cos\theta + x_2\sin\theta - t)\ dx\ , \tag{3.1}$$

where we permit $\theta \in [0, 2\pi)$ and $t \in \mathbf{R}$. To create a space of such objects, we let $[\ ,\ ]$ denote the pairing

$$[F,G] = \frac{1}{4\pi}\int_0^{2\pi}\int_{-\infty}^{\infty} F(t,\theta)G(t,\theta)^* dt\ d\theta\ , \tag{3.2}$$

and by $L^2(dt\ d\theta)$ norm we mean $\|F\|^2 = [F,F]$. We let R^+ denote the adjoint of the Radon transform so that for all sufficiently nice $G \in \mathcal{A}$ and all sufficiently nice $f \in L^2(dx)$,

$$[Rf, G] = \langle f, R^+G\rangle, \tag{3.3}$$

which leads to

$$(R^+G)(x) = \frac{1}{4\pi}\int_0^{2\pi} G(x_1\cos\theta + x_2\sin\theta, \theta)\ d\theta\ . \tag{3.4}$$

Define the Riesz order-1/2 fractional differentiation operator Δ^+ by

$$(\Delta^+ f)(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{it\omega}\hat{f}(\omega)|\omega|^{\frac{1}{2}}\ d\omega\ . \tag{3.5}$$

This unbounded operator is well-defined on functions which are sufficiently smooth [formally, the domain $\mathcal{D}(\Delta^+) = \{f : \int_{-\infty}^{\infty} |\hat{f}(\omega)|^2\ |\omega| d\omega\}$]. In particular, it is well-defined on every 1-D Meyer wavelet $\psi_{j,k}$, owing to $\mathrm{supp}(\hat{\psi}_{j,k}) \subset \{\omega : |\omega| \in [\frac{2}{3}\pi 2^j,\ \frac{8}{3}\pi 2^j]\}$. Moreover, on the appropriate domain, it is self-adjoint.

Definition 2. *The* **Radon Isometry** $\overline{R}$ *is the operator densely defined by the formal expression*

$$\overline{R}(f) = (\Delta^+ \otimes I)R[f].$$

The **adjoint Radon Isometry** $\overline{R}^*$ *is the operator densely defined by the formal expression*

$$\overline{R}^*(G) = R^+[(\Delta^+ \otimes I)G].$$

Now note that Rf has the antipodal symmetry

$$(Rf)(-t, \theta + \pi) = (Rf)(t, \theta) \ . \tag{3.6}$$

We adopt the convention that F (and G and variants) typically will denote a function on $\mathbf{R} \times [0, 2\pi)$ obeying the same antipodal symmetry:

$$F(-t, \theta + \pi) = F(t, \theta) \ . \tag{3.7}$$

Let $\mathcal{A}$ be the closed subspace of $L^2(dt\ d\theta)$ of functions F obeying (3.7). Let $P_{\mathcal{A}}F$ be the orthoprojector from $L^2(dt\ d\theta)$ onto $\mathcal{A}$, defined by

$$(P_{\mathcal{A}}F)(t, \theta) = (F(t, \theta) + F(-t, \theta + \pi))/2 \ . \tag{3.8}$$

4. The Ridgelet Construction, and its Properties

The Radon machinery can be used to construct orthonormal ridgelets very efficiently. Define the operator of reflection of functions of one variable $(Tf)(t) = f(-t)$ and the operator of translation by half a period by $(Sg)(\theta) = g(\theta + \pi)$. Note that the space $\mathcal{A}$ consists of objects invariant under $T \otimes S$; (3.7) can be rewritten $(T \otimes S)F = F$. In fact, $P_{\mathcal{A}} = (I + T \otimes S)/2$. Set now, for $j, k \in \mathbf{Z}$, and $i \geq i_0$, $\ell = 0, \ldots, 2^{i-1} - 1$, $\varepsilon \in \{0, 1\}$

$$W_\lambda(t, \theta) = P_{\mathcal{A}}(\psi_{j,k} \otimes w^\varepsilon_{i,\ell}) \ , \tag{4.1}$$

where $\lambda = (j, k; i, \ell, \varepsilon)$. For later reference, we spell this out:

$$W_\lambda(t, \theta) = (\psi_{j,k}(t) w^\varepsilon_{i,\ell}(\theta) + \psi_{j,k}(-t) w^\varepsilon_{i,\ell}(\theta + \pi))/2 \ . \tag{4.2}$$

It was shown in [15] that the W_λ provide an orthobasis for $\mathcal{A}$. In order to obtain orthonormality with respect to the scalar product $[,]$, a particular normalization was imposed. In that normalization, $\|\psi_{j,k}\|_{L^2(\mathbf{R})} = \sqrt{2}$, and $\|w^\varepsilon_{i,\ell}\|_{L^2[0,2\pi]} = 2\sqrt{\pi}$. In a sense, the $(W_\lambda : \lambda \in \Lambda)$ constitute a "tensor wavelet basis which has been antipodally symmetrized".

If we now let ρ_λ denote the ridgelet elements defined in the Introduction, we note that

$$\rho_\lambda = \overline{R}^*[W_\lambda], \tag{4.3}$$

this says that an orthonormal ridgelet is *isometric to a wavelet in Radon-space which has been antipodally-symmetrized.*

The ortho-ridgelet construction may be viewed as transferring a basis from Radon space to real space via an isometry $\overline{R}^*$. If we reflect on the details of the above construction, we notice that the basis we used was not completely arbitrary: it had to consist of elements both in the domain of $\overline{R}$ and the range of $\overline{R}^*$. Now the range and domain both consist of functions in $\mathcal{A}$, and the easiest elements in both range and domain to describe are functions which are bandpass in t, i.e. functions with support in the frequency domain contained in a compact set separated from the origin.

These ideas imposed the following restrictions on the construction.

- The basis on Radon space was a basis for $\mathcal{A}$ rather than $L^2(dtd\theta)$. This meant that its elements had to obey an antipodal symmetry requirement, or equivalently that an element W of the basis had to obey the invariance $P_{\mathcal{A}}W = W$.
- In order to construct such a basis, we started with an orthonormal basis for $L^2(dtd\theta)$ and operated on it by $P_{\mathcal{A}}$, creating a tight frame with antipodal symmetry. But as turned out, the tight frame was actually an orthobasis, owing to two special *closure properties* of these families we used; closure under reflection about the origin in the ridge direction:

$$\psi_{j,k}(-t) = \psi_{j,1-k}(t) \ , \tag{4.4}$$

  and closure under translation by half a cycle in the angular direction:

$$w^{\varepsilon}_{i,\ell}(\theta + \pi) = w^{\varepsilon}_{i,\ell+2^{i-1}}(\theta) \ . \tag{4.5}$$

  The closure property (4.4) would ***not*** hold for certain other prominent wavelet families, such as Daubechies' compactly supported wavelets. The significance of the closure properties was that for certain pairs $(W_\lambda, W_{\lambda'})$, $P_{\mathcal{A}}W_\lambda = P_{\mathcal{A}}W_{\lambda'}$, so that the induced frame consisted of many identical pairs. Systematically removing one element from each such pair, and rescaling the other element, we obtained an orthonormal basis.
- An 'absence of low-frequencies' restriction was imposed: the basis in the ridge-direction consisted entirely of bandpass elements, i.e. elements with frequency-domain support in an octave band disjoint from the origin.

If we weakened these conditions, the following would still be true.

- We can always start from an orthobasis for $L^2(\mathbf{R})$ and apply the projector $P_{\mathcal{A}}$, getting a tight frame for $\mathcal{A}$. We can then apply the isometry to this, getting a tight frame for real space.
- If, in addition, the original basis obeyed appropriate closure under ridge-reflection and angular translation, the tight frame in real space can be decimated by a factor of two to form an ortho-basis.

- The condition that the low-frequency terms be absent from all basis elements is simply a regularity condition on the outcome of the procedure. If certain elements in the basis have support near the origin in frequency space, then the construction can still take place; however, some of the corresponding frame elements will have poor decay.

In other words, the construction is quite general, but it might lead to a redundant set with redundancy two and it might lead to a basis where certain elements do not exhibit good spatial decay.

5. Ridgelet Packet Construction

We now propose a class of tight frames based on the remarks just given. In certain cases, these can be subsampled to form orthobases.

5.1. General Procedure

We begin with a general set of ingredients:

- An orthonormal basis $(U_\mu(t))$ for $L^2(\mathbf{R})$ for the ridge direction. If the elements are bandlimited, we call this a bandlimited basis. If the elements obey the closure condition

$$U_\mu(-t) = U_{\mu'}(t)$$

  the basis will be called a basis *closed under reflection*.

- A basis $(V_\nu(\theta))_\nu$ for $L^2[0, 2\pi)$ in the angular direction. If the elements obey the closure condition

$$V_\nu(\theta + \pi) = V_{\nu'}(\theta)$$

  the basis will be called a basis *closed under translation*.

- A collection of antipodally-symmetric functions $\mathcal{A}$ will be constructed from the two families of bases. Letting $\lambda = (\mu, \nu)$ group the indices in each of the variables,

$$W_\lambda(t, \theta) = P_{\mathcal{A}}[U_\mu \otimes V_\nu]; \tag{5.1}$$

  as the result of applying an orthonormal projector to the orthonormal basis, the W_λ make a tight frame for $\mathcal{A}$.

- A collection of functions ρ_λ will be induced by the isometry $\overline{R}^*$:

$$\rho_\lambda = \overline{R}^*(W_\lambda) \qquad \forall \lambda.$$

  As an isometry of a tight frame, the ρ_λ make a tight frame for their span. In fact their span is all of $L^2(\mathbf{R}^2)$.

We note the following.

First, there is a simple expression for the element ρ_λ

$$\hat{\rho}_\lambda(\xi) = |\xi|^{-1/2}(\hat{U}_\mu(|\xi|) \cdot V_\nu(\theta) + \hat{U}_\mu(-|\xi|) \cdot V_\nu(\theta + \pi))/2$$

valid for $\xi = (|\xi| \cos(\theta), |\xi| \sin(\theta))$.

Second, if the elements U_μ are bandpass, with C^∞ Fourier transforms, and if the elements V_ν are C^∞ then the elements ρ_λ are likewise bandpass with smooth Fourier transforms; it follows that they are C^∞ with spatial rapid decay.

Third, the general procedure described above has been stated for tensor product bases $U_\mu \otimes V_\nu$. In general, there is no reason to restrict ourselves in this way. More generally, we may allow a semi-direct product

$$W_\lambda = P_{\mathcal{A}}[U_\mu \otimes V_{\nu|\mu}], \qquad \forall \lambda = (\mu, \nu) \tag{5.2}$$

where the basis $(V_{\nu|\mu} : \nu)$ depends on μ. The orthoridgelet basis defined in the Introduction in fact has this form, as can be seen from the constraint $i \geq j$. By and large, the freedom enabled by the rule (5.2) will only be exercised in a limited way, as exemplified by the way it is exercised in the ridgelet orthobasis; the coarsest-scale of resolution may be adjusted to the properties of the corresponding ridge element.

5.2. Bases of Ridgelet Packets

While in principle, *any* pair of bases may be used for the above construction, we are interested here in those bases deriving from applying certain principles of time-frequency localization [13, 33].

Definition 3. *We call* **ridgelet packet basis** *a basis constructed by the above procedure, where the basis U_μ is chosen from a* **wavelet packets dictionary** *and the basis V_ν (or $V_{\nu|\mu}$ if rule (5.2) is used) is chosen from a* **wavelet packets dictionary** *or a* **cosine packets dictionary**.

When the basis in the angular direction is chosen from the wavelet packets dictionary, we will sometimes speak of the *Radon-domain approach* to defining ridgelet packets, whereas when the basis in the angular direction is chosen from the cosine packets dictionary, we will speak of the *polar Fourier-domain approach*. This distinction reflects the structure of the underlying algorithms in the two situations, as we will discuss later. Admittedly, this is artificial to some extent, since the Radon and Fourier domains are related in 1-1 fashion, but we find the distinction helpful.

5.3. Radon Approach: Wavelets in both Ridge and Angular Directions

The orthonormal ridgelet basis is built using wavelets in both the ridge and angular directions. Other bases can be built within this framework, by simply

varying the base resolution level of the angular wavelets as a function of the resolution of the ridge wavelets.

In the ortho-ridgelet case, we start with ridge wavelets $\psi_{j,k}(t)$ for $j, k \in \mathbf{Z}$ and with angular wavelets $w^{\varepsilon}_{i,\ell}(\theta)$. The key decision is that we limit $i \geq j$, and we have $\varepsilon = 1$ for $i > j$, while $\varepsilon \in \{0, 1\}$ for $i = j$.

To interpret these choices, focus on the situation where $i = j$ and $\varepsilon = 0$. Hence we are looking at a tensor product based on the male-gendered wavelet at scale j, $w^0_{j,\ell}(\theta) \cdot \psi_{j,k}(t)$ Note that for $\varepsilon = 0$, $w^{\varepsilon}_{j,\ell}(\theta)$ is a "bump", integrating to $2^{j/2}$. The tensor product is thus localized near $\theta = \ell/2^j$, and has each constant-θ, varying-t profile proportional to the wavelet $\psi_{j,k}$.

This has an interpretation in terms of the tilings mentioned in the introduction. Indeed, the successive terms

$$w^0_{j,\ell}(\theta), w^1_{j,\ell}(\theta), w^1_{j+1,2\ell}(\theta), w^1_{j+1,2\ell+1}(\theta), \ldots$$

represent an orthogonal set of functions localized in the vicinity of the angular interval $[2\pi\ell/2^j, 2\pi(\ell+1)/2^j)$. Thinking now in the polar frequency domain, the orthogonal functions

$$w^0_{j,\ell} \otimes \hat{\psi}_{j,k}, w^1_{j,\ell} \otimes \hat{\psi}_{j,k}, w^1_{j+1,2\ell} \otimes \hat{\psi}_{j,k}, w^1_{j+1,2\ell+1} \otimes \hat{\psi}_{j,k}, \ldots$$

for $k \in \mathbf{Z}$, create an orthogonal set localized near the 'tile' $[2\pi\ell/2^j, 2\pi(\ell+1)/2^j) \times [2^j, 2^{j+1})$. In this way, the formula (1.1) for ridgelets implements the tiling shown in Figures 1 and 6.

This discussion suggests how we can derive a formula for basis elements which implement a quasi-FIO tiling. The idea is to use the same framework, only instead of taking the base resolution $i_0 = j + c$, we take $i_0 = j/2 + c$. The orthonormal functions

$$w^0_{i_0,\ell} \otimes \hat{\psi}_{j,k}, w^1_{i_0,\ell} \otimes \hat{\psi}_{j,k}, w^1_{i_0+1,2\ell} \otimes \hat{\psi}_{j,k}, w^1_{i_0+1,2\ell+1} \otimes \hat{\psi}_{j,k}, \ldots$$

form an orthogonal set of functions all localized near the 'tile' $[2\pi\ell/2^{i_0}, 2\pi(\ell+1)/2^{i_0}) \times [2^j, 2^{j+1})$. Hence, the angular subdivision is not nearly so fine, so that at frequency 2^j, we have tiles of height 2^j and width $2\pi \cdot 2^{-j/2}$.

Many other possibilities could be considered. Perhaps the simplest is to pick the base angular resolution fixed, independent of j: $i_0 = 3$ (say). Then the functions

$$w^0_{3,\ell} \otimes \hat{\psi}_{j,k}, w^1_{3,\ell} \otimes \hat{\psi}_{j,k}, w^1_{4,2\ell} \otimes \hat{\psi}_{j,k}, w^1_{4,2\ell+1} \otimes \hat{\psi}_{j,k}, \ldots$$

form an orthonormal set, each one localized near a tile of fixed width $2\pi/8$ and height 2^j.

All these constructions have the qualitative property that a given basis element generated from

$$w^{\varepsilon}_{i,\ell} \otimes \psi_{j,k},$$

exhibits an orientation localized to directions near $\theta_{i,\ell} = 2\pi\ell/2^i$, and a scale normal to that direction of scale 2^{-j}. To see this, note that the basis element is generated by

$$\rho_\lambda = \overline{R}^*(P_{\mathcal{A}}(w^\varepsilon_{i,\ell} \otimes \psi_{j,k}))$$

which can be written as

$$\begin{aligned}\rho_\lambda(x) &= \tfrac{1}{4\pi}\int(\psi^+_{j,k}(x_1\cos(\theta)+x_2\sin(\theta))w^\varepsilon_{i,\ell}(\theta))/2d\theta \\ &\quad +\tfrac{1}{4\pi}\int(\psi^+_{j,k}(x_1\cos(\theta+\pi)+x_2\sin(\theta+\pi))w^\varepsilon_{i,\ell}(\theta+\pi))/2d\theta,\end{aligned}$$

where $\psi^+_{j,k} = \Delta^+\psi_{j,k}$. This shows that each basis element is an angular "average" of ridge functions

$$\psi^+_{j,k}(x_1\cos(\theta)+x_2\sin(\theta))$$

over θ in a 2^{-i} vicinity of $\theta_{i,\ell}$. The only "location-like" parameter here is k, which sets the position of the underlying ridge near $x_1\cos(\theta)+x_2\sin(\theta) = t_{j,k}$, where $t_{j,k} = k/2^j$. It follows, in particular, that the system, while offering an orientation, a ridge, and a scale parameter, does not offer a traditional location parameter.

In later sections, we will give illustrations of digital frame elements inspired by these constructions.

5.4. Radon Approach: Wavelet Packets in the Ridge Direction

By *Wavelet Packets* basis in the ridge direction, we mean the use of the principle of local cosine bases of Coifman and Meyer (1989) applied in the *radial frequency variable* λ. In our opinion, the best references for understanding this construction are the article of Auscher, Weiss and Wickerhauser [1] (in English) and the monograph of Yves Meyer (in French) [26]. We note that, to avoid confusion, our notation is nonstandard, since typically the term wavelet packets refers to bases constructed by applications of special filter banks, and the specific idea we discuss now cannot be implemented through finite-length filter banks.

One chooses a partition in the 1-dimensional *frequency* variable according to the general rules of symmetric recursive dyadic partitioning. One takes the initial sequence of breakpoints $\{2^j : j \geq 0\}$ and views this as referring to the partition

$$\{(-1,0]\cup[0,1),(-2,-1]\cup[1,2),(-4,-2]\cup[2,4),\ldots\}$$

and one considers all partitions reachable from this one by repeatedly applying midpoint splits to a pair in the partition, producing a new pair. For example, we could split $[0,1)$ into $[0,1/2),[1/2,1)$ and also $(-1,0]$ into $(-1/2,0],(-1,-1/2]$ and then replace $(-1,0]\cup[0,1)$ in the initial partition by the pair $(-1/2,0]\cup[0,1/2)$ and $(-1,-1/2]\cup[1/2,1)$, producing a new partition of this sort. One

may, if one likes, impose a balance condition on partitions, allowing only partitions in which adjacent intervals differ by a factor of two in length.

Associated with any partition reachable in this way is an orthonormal basis, produced as follows. To each interval I in the partition we associate a window $w_I(\lambda)$ which is smooth and nonnegative, 1 near the center of the window, and vanishing outside a slight enlargement of the window. The squares of the windows together should form a partition of unity: $\sum_I w_I^2(\lambda) = 1$, $\forall \lambda$. Then we define a collection of trigonometric functions $\phi_{I,k}(\omega)$ associated with the window which make an orthonormal set for $L^2(I)$. If the interval does *not* abut 0, these functions are chosen from the DCT-IV system. If the window *does* abut 0, we view I and $-I$ as a single interval $\tilde{I}$ and these functions are chosen from the DST-III system. The basis is then the collection

$$\hat{U}_\mu(\lambda) = w_I(\lambda) \cdot \phi_{I,k}(\lambda), \lambda \in \mathbf{R}$$

where $\mu = (I, k)$ is an index pair unifying the indices I and k.

Some examples of this construction are quite familiar.

- *Meyer Wavelets.* If we use breakpoints $\{2^j : -\infty < j < \infty\}$, we get a partition into intervals $I_j = (-2^{j+1}, -2^j] \cup [2^j, 2^{j+1})$. The basis element U_μ with index $\mu = (I_j, k)$ is then precisely an orthonormal Meyer wavelet $\psi_{j,k}$.
- *Wilson-like Basis.* If we use breakpoints $\{1, 2, 3, \ldots\}$, we get a partition into intervals
$$I_j = (-j, -(j-1)] \cup [j-1, j), \tag{5.1}$$
and we obtain in this way elements familiar to those who understand [1] and who have studied the construction of the Wilson basis [14]. In effect, the basis elements are windowed sinusoids of frequency roughly j, exponentially localized near a position proportional to k in the time domain.

Other examples of the construction may seem more exotic:

- *Intermediate Coherence Length.* Suppose we use breakpoints
$$\{1, 2, 4, 6, 8, 12, 16, 20, 24, 28, 32, ...\},$$
where in general the $2j$-th and $2(j+1)$-th initial intervals $[2^j, 2^{j+1})$ and $[2^{j+1}, 2^{j+2})$ are recursively subdivided $j/2$ times, yielding a family of $2^j/2$ subintervals. Then we obtain a basis where the typical elements supported near high frequency ω have a frequency localized in a band of width about $\sqrt{\omega}$ and a time localization, according to the Heisenberg principle, to a correspondingly short interval of length about $1/\sqrt{\omega}$. This says that the time coherence of effects at frequency ω is not as short as in the wavelet system, where it is proportional to $1/\omega$, nor as long as in the Gabor system, where coherent effects last for about one unit of time.

- *Increasing Coherence Length.* If we use breakpoints

$$\{1, 2, 3, 4, 4\frac{1}{2}, 5, 5\frac{1}{2}, 6, 6\frac{1}{2}, 7, 7\frac{1}{2}, 8, 8\frac{1}{4}, 8\frac{1}{2}, 8\frac{3}{4}, 9, ...\},$$

  where in general the j-th initial dyadic interval $[2^j, 2^{j+1})$ is subdivided dyadically through $2j-2$ complete generations, then we obtain a basis where the typical elements supported near high-frequency ω have a frequency localized in a band of width about $1/\sqrt{\omega}$ and a time localization, according to the Heisenberg principle, to a correspondingly short interval of length about $\sqrt{\omega}$. This says that the time coherence of effects at frequency ω is not as short as in the Gabor system, where coherent effects last for about one unit of time, nor as long as in the Fourier system, where coherent effects last for infinite time.

With any of these choices, we can then subdivide the angular variable in a fashion subordinate to the ridge frequency variable, according to the same principle as in the ortho ridgelet basis. Let $V_{\nu|\mu}$ be simply the periodized Meyer wavelet as in the ortho ridgelet basis – under a low frequency constraint to be determined below – and let W_μ be a wavelet packet basis based on a different partition than the dyadic wavelet partition. Consider for example the Wilson-like basis partition (5.1) based on integer breakpoints. Choose the low-frequency constraint on $w^\varepsilon_{i,\ell}$ so that $i \geq j$ i.e. so that the angular scale is finer than the ridge frequency. It results that for $j > 0$, the ρ_λ are bandlimited and of rapid decay.

For each ρ_λ we have from (4.3) the formula

$$\rho_\lambda = \overline{R}^*[W_\lambda]$$

which gives the explicit formula

$$\begin{aligned}\rho_\lambda(x) \quad &= \tfrac{1}{4\pi}\int (U^+_\mu(x_1\cos(\theta)+x_2\sin(\theta))w^\varepsilon_{i,\ell}(\theta))/2d\theta \\ &+\tfrac{1}{4\pi}\int (U^+_\mu(x_1\cos(\theta+\pi)+x_2\sin(\theta+\pi))w^\varepsilon_{i,\ell}(\theta+\pi))/2d\theta.\end{aligned}$$

Now roughly speaking, U^+_μ, with $\mu = (j,k)$ is a sinusoid of frequency j, say localized to an interval of length ≈ 1 situated near $t \approx k$. Hence, the ridge function $U^+_\mu(x_1\cos(\theta)+x_2\sin(\theta))$ is localized near $x_1\cos(\theta)+x_2\sin(\theta) = k$. Similarly $w^\varepsilon_{i,\ell}(\theta)$ is localized near $\theta = \theta_{i,\ell} = 2\pi\ell/2^i$. It follows that the integrand is large for x in a range where $x \approx (k\cos(\theta_{i,\ell}), k\sin(\theta_{i,\ell}))$, so we may expect that for $\varepsilon = 0$ and $i = i_0(\mu)$, the function ρ_λ concentrates near $x \approx (k\cos(\theta_{i,\ell}), k\sin(\theta_{i,\ell}))$. For $\varepsilon = 1$ and $i > i_0$, one must argue by cancelation, which is more subtle.

5.5. Polar Fourier Approach: Wavelet $\otimes$ Cosine Packet

Let now (U_μ) be simply the standard Meyer wavelet basis for $\mathbf{R}$, just as in the ortho-ridgelet basis (1.1). Let $V_{\nu|\mu}$ however, be a cosine packet basis based

on a recursive dyadic partition of the angle domain. Consider for example, a partition based on dividing the angular domain into 2^j equal sectors. Use the cosine packets subordinate to this partitioning. In the polar Fourier domain, things are very simple, because Meyer wavelets are the Fourier transforms of cosine packets in the frequency domain. Hence we have cosine packets in λ times cosine packets in θ. Hence, bivariate cosine packets are being used, subordinate to a recursive dyadic partition.

For each ρ_λ we have the explicit formula

$$\begin{aligned}\rho_\lambda(x) &= \tfrac{1}{4\pi}\int(\psi^+_{j,k}(x_1\cos(\theta)+x_2\sin(\theta))V_\nu(\theta))/2d\theta \\ &+\tfrac{1}{4\pi}\int(\psi^+_{j,k}(x_1\cos(\theta+\pi)+x_2\sin(\theta+\pi))V_\nu(\theta+\pi))/2d\theta.\end{aligned}$$

Now, roughly speaking, $\psi^+_{j,k}$ is a wavelet of scale 2^{-j}, localized near $t \approx t_{j,k} = k/2^j$. Hence, the ridge function $\psi^+_{j,k}(x_1\cos(\theta)+x_2\sin(\theta))$ is localized near $x_1\cos(\theta)+x_2\sin(\theta) = t_{j,k}$. Similarly V_ν is localized to an interval $J_{m,\ell}$. It follows that the integrand is large for x in a range where $|x| \approx t_{j,k}$, and $\theta \in J_{m,l}$. We may expect that the function ρ_λ is large in the neighborhood where $x \approx \pm(|t_{j,k}|\cos(\theta_{i,\ell}), |t_{j,k}|\sin(\theta_{i,\ell}))$. Knowing the exact shape of the support requires additional insight.

Now we make the more detailed assumption that V_ν is a sinusoid in θ of frequency $2\pi\cdot k_1$ localized near the interval $J_{m,\ell}$. This allows us to study the details of ρ_λ on its support. For large $|t_{j,k}|$ and large m, the integrand is approximately of the form

$$\psi^+_{j,k}(x_1\cos(\theta_{m,l})+x_2\sin(\theta_{m,l}))w_{J_{m,l}}(\theta)\phi_{k_1}(\theta).$$

Hence it has approximately the form of a wavelet function in the ridge direction and the form of a localized sinusoid in the transverse direction.

6. Implementation on Digital Data

Ridgelet Packets bases for digital data can be constructed based on an adaptation of a circle of ideas associated to digital implementation of the Radon transform, polar Fourier transform, and ridgelet transform [3, 17].

6.1. Fast Slant Stack

Averbuch et al. (2001) [3] describe a realization of the Radon transform suited for n-by-n image data, called Fast Slant Stack, claiming that the transform is geometrically accurate and can be implemented by a fast algorithm. The geometric accuracy, for example, implies that the backprojection of a point in Radon space is a true ridge function, i.e. a true object of the form $\psi(x+sy)$, where $\psi(\cdot)$ is delta-like.

This scheme has been deployed by Donoho and Flesia [17] to produce a discrete ridgelet transform based on true ridge functions. In our work for this

paper, we have used the same scheme to provide a digital implementation of ridgelet packets.

6.2. Pseudopolar FFT

Underlying the Fast Slant Stack is a notion of digital polar transform Fourier called pseudopolar FFT in [3].

The key point is to view the digital Fourier domain not as a cartesian grid, but instead as a special pointset as shown in Figure (7). Then define the *pseudopolar Fourier transform* as the evaluation of the Fourier transform

$$\hat{I}(\xi) = \sum_{x_1,x_2=0}^{n-1} I(x_1, x_2) \exp\{-(x_1\xi_1 + x_2\xi_2)\}$$

at the $4n^2$ points of this pointset. The pointset can be viewed as a set of "concentric squares" stacked inside each other (like Chinese boxes), with equispaced points along the boundary of the box. The half-width of a side functions as a pseudo radius, and the arclength along the perimeter of the box functions as a pseudo angular variable.

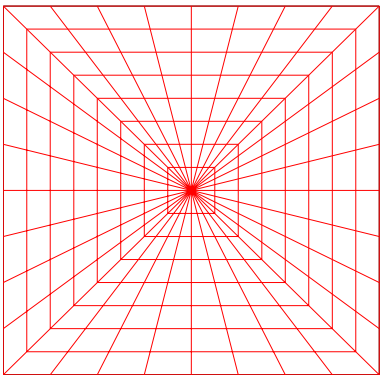

Figure 7. Pseudopolar Fourier Grid.

As shown in [3], the evaluation of the Fourier sum on this set of gridpoints can be performed in order $N \log(N)$ flops, where $N = n^2$ is the total number of pixels. The underlying ideas that allow rapid evaluation of these specific gridpoints date back to work of Pasciak [27], Edholm and Herman [18], and Lawton [24], working variously in Medical Imaging and in Synthetic Aperture Radar.

The resulting set of pseudopolar values may be viewed as a $2n$ by $2n$ array: $2n$ points on each line through the origin, and $2n$ lines through the origin,

grouped in columns as different lines through the origin, in rows as different 'radii'. We define the pseudopolar FFT $P(I)$ to be the transform from n by n arrays to $2n$ by $2n$ arrays produced in this way.

Note that the pseudopolar grid samples the region near the origin more finely than the region near the boundary. In fact the spacing between samples on line segments varies inversely with distance of the segment from the origin. Define the *normalized pseudopolar FFT* $\overline{P}(I)$ to be the result of applying a simple rescaling of entries in $P(I)$ according to the square root of the local sample spacing in the pseudopolar grid at the corresponding grid point. Since $P(I)$ is a discrete analog of $F(r,\theta) = \hat{f}(r\cos(\theta), r\sin(\theta))$, sampled at specific points in (r,θ), the definition of $\overline{P}(I)$ is very analogous to defining in the continuum case $\overline{F}(r,\theta) = r^{1/2}\hat{f}(r\cos(\theta), r\sin(\theta))$. Recall that $f \mapsto \overline{F}$ is an isometry from $L^2(dxdy)$ to $L^2(drd\theta)$; we can't get quite so much in the digital case: Instead we have

$$C_1\|I\|_2 \leq \|\overline{P}(I)\|_2 \leq C_2\|I\|_2, \tag{6.1}$$

where empirically, $C_2/C_1 < 1.1$.

Note that if we had $C_1 = C_2$ then, up to normalization, $\overline{P}$ would be an ℓ^2 isometry. In that sense, the mapping $I \mapsto \overline{P}(I)$ is *a digital analog of the polar Fourier Isometry.*

6.3. Digital Radon Domain

If we apply a 1-dimensional inverse FFT to each column of the 2-D pseudopolar FFT array, we create a new $2n$-by-$2n$ matrix. This matrix is a digital Radon transform of I; each column gives the sums of (an interpolant of) I along a family of equispaced parallel lines, where the slope of the lines in that family is indexed by the column index (which provide a pseudo-angular variable) [3]. Call the overall mapping $S(I)$ the slant stack.

If we apply instead a 1-dimensional inverse FFT to each column of the 2-D *normalized* pseudopolar FFT array, we create another new $2n$-by-$2n$ matrix. This matrix is a *preconditioned* digital Radon transform of I. Call the overall transform mapping $\overline{S}(I)$ the normalized slant stack. Because of the near-isometry property of $\overline{P}(I)$, we have

$$C_1\|I\|_2 \leq \|\overline{S}(I)\|_2 \leq C_2\|I\|_2.$$

Here C_1 and C_2 are the same as in (6.1). Again, if $C_1 = C_2$ then, up to normalization, then $\overline{S}(I)$ would be an ℓ^2 isometry. In that sense, the mapping $I \mapsto \overline{S}(I)$ is *a digital analog of the Radon Isometry.*

6.4. Strategy for Digital Implementation

We have now introduced a set of digital-data friendly tools which are analogous to the continuum-domain tools discussed in earlier sections. In a sense, we

have built up a 'dictionary' to translate between the continuum domain and the digital domain. The dictionary is summarized in the table below.

Continuum Concept	Symbol	Digital Concept	Symbol
Polar Fourier Transform	$F(\lambda, \theta)$	Pseudopolar FFT	$P(I)$
Polar Fourier Isometry	$\overline{F}(\lambda, \theta)$	Normalized Pseudopolar FFT	$\overline{P}(I)$
Radon Transform	$R(t, \theta)$	Slant Stack	$S(I)$
Radon Isometry	$\overline{R}(t, \theta)$	Normalized Slant Stack	$\overline{S}(I)$

Our strategy for digital implementation of ridgelet packets is to use this dictionary to substitute digital concepts for continuum concepts in the original definitions. Thus, if a certain orthobasis for the continuum case used wavelets 'in each direction' followed by dual Radon isometry, then we propose to work with the normalized slant stack, and use a discrete wavelet basis 'in each direction', followed by the adjoint normalized slant stack 'to return to the spatial domain'.

A remark about this strategy: before heading a great distance down this path, it is important to know that there is a 'proof of concept' which shows that at least in a special case, the strategy provides decent results. In this case, the proof of concept has been provided by the algorithm for the digital ridgelet transform in [17]. In constructing that transform, the authors have followed the strategy suggested above and carefully documented the properties of the digital domain transforms that result.

A second remark: while the continuum approach leads to the definition of various bases, in the discrete case we will only get frames, for a simple reason. The normalized pseudopolar FFT and the normalized slant stack are both transforms from $n \times n$ arrays to $2n \times 2n$ arrays. Hence the strategy must in general lead to overcomplete systems (frames) rather than orthobases. However, the frames generated in this manner can be expected to have good frame bounds, owing to the closeness of C_1 and C_2 in (6.1).

6.5. Digital Ridgelet Packets

Our implementation strategy leads to the following general schema:

Definition 4. *A* **Digital Ridgelet Packet transform** *in dimension two is a transform of n-by-n data defined as follows.*

[DRPT1] The n-by-n digital array is transformed into a digital Radon domain via the fast normalized slant stack algorithm in [3], which gives a $2n$*-by-*$2n$ *array.*

[DRPT2] These arrays are then transformed according to some specific combination of wavelet packets and cosine packets in each of the two directions (angular vs. ridge), where the combination of bases in the two directions is made according to a direct product or a semidirect product.

The **Inverse Digital Ridgelet Packet transform** *in dimension two is a transform returning from the Ridgelet Packet domain as follows.*

[IDRPT1] The Ridgelet packet coefficients are transformed back into the Radon domain by inverting the transform in step [DRPT2] above.

[IDRPT2] The Radon domain data are transformed back into the original digital spatial domain by inverting the normalized slant stack transform in step [DRPT1] above, using the algorithm in [3].

This algorithm has the following general characteristics for an image of size n by n with $N = n^2$ pixels.

- Storage Space: The algorithm requires permanent storage of order $O(N)$, and temporary storage of comparable size.
- Complexity: The forward transform algorithm requires $O(N \cdot \log N)$ flops, and the inverse transform algorithm requires $O(C(\varepsilon) \cdot N \cdot \log N)$ flops, where $C(\varepsilon)$ depends on the relative accuracy required. $C(10^{-6}) \approx 7$.

For certain purposes it may be useful to apply the adjoint of the forward digital ridgelet packet transform. This is obtained by a two-step procedure

[ADRPT1] The Ridgelet packet coefficients are transformed back into the digital Radon domain by applying the adjoint of the transform in step [DRPT2] above. If the transform is orthogonal (as would be typical) this is the same as inverting the transform.

[ADRPT2] The Radon domain data are transformed back into the original digital domain by applying the adjoint of the transform in step [DRPT1] above, using the adjoint algorithm in [3].

The adjoint uses $O(N \log(N))$ flops and $O(N)$ space.

6.6. Digital Implementation

We have developed a digital implementation of these ideas as part of a MATLAB toolbox called *BeamLab*. This toolbox contains tools for Ridgelet, beamlet, and curvelet analysis, and is available for download from

`http://www.beamlab.net`

where further information is available.

BeamLab contains scripts which can reproduce all the figures in the article you are reading. It also contains a directory named *RP_FCP* which implements ridgelet packets based on anisotropic cosine packets in the pseudopolar Fourier plane, and a directory named *RP_RWW* which implements ridgelet packets based on applying wavelets and wavelet packets in the Radon plane.

6.7. Examples of Digital Implementation

We now demonstrate the implementation indicated above by displaying some basis functions and the associated coefficient functionals.

6.8. Synthesis from Tiles

What do ridgelet packets look like? Given the emphasis of this paper on tilings, it might seem most natural to try many different tilings, and some specific basis functions associated with each. We suggest that another approach would be more directly informative: to study individual tiles.

Indeed, the system of tilings underlying our constructions has a relatively few different shapes of tiles: they are all oriented parallel to the axes in (r, θ) space, and their widths and heights vary through a dyadic set $(2^i, 2^j)$. They are located at various places in the (r, θ) plane, but most of this variation is easily visualized, as either a rescaling or a rotation.

In short, we propose that to get a good understanding of the system, we should fix a certain (r, θ) and consider objects associated to tiles at a range of different aspect ratios: $(2^i, 2^{i+h})$, for varying h, or else $(2^{j+h}, 2^j)$ for varying h. In this way we see the effect of varying the tile shape on the basis elements.

Carrying out this proposal, we begin by working in the Fourier plane with Cosine packet bases. Figure 8 gives an example of several basis functions obtained from picking a tile extending from 1/4 to 1/2 of Nyquist in the radial frequency variable, and considering various widths ranging from very narrow to very broad. In these figures, we tried to keep the base frequency as low as possible, so that we explore the shape of the envelope rather than the oscillations within the envelope.

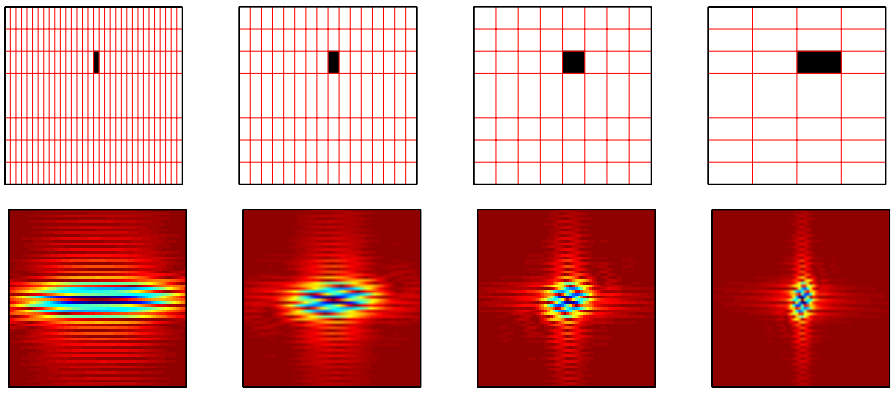

Figure 8. Basis Functions Derived from Tiles of Different Widths.

Figure 9 gives an example of several basis functions obtained by picking a tile extending from $\pi/4$ to $3\pi/8$, and considering various heights ranging from very short to very tall. In these figures, we again tried to keep the base frequency as low as possible, so that we explore the shape of the envelope rather than the oscillations within the envelope.

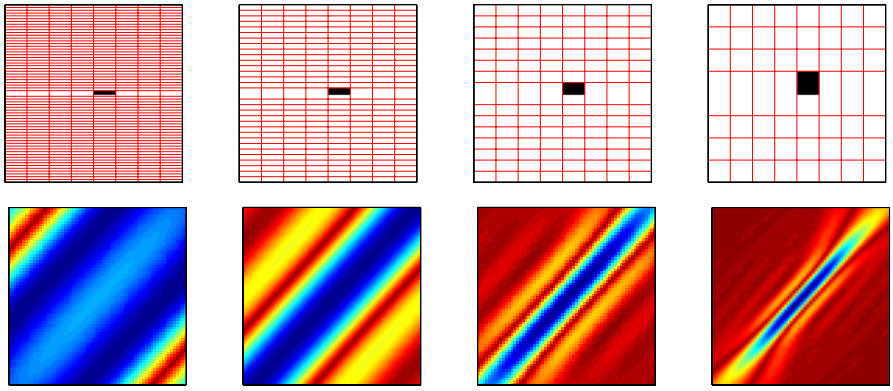

Figure 9. Basis Functions Derived from Tiles of Varying Heights.

Continuing with this proposal, we consider comparable tilings, only we implement them using the Radon plane with wavelet packet bases. Figure 10 gives an example of several basis functions obtained from picking a Radon-domain waveform made of the tensor product of a wavelet packet in the ridge direction and a wavelet in the angular direction. The figure explores the results obtained by varying the scale of the wavelet packet in the ridge direction.

6.9. Analysis

We now give an example of using the Ridgelet Packet systems to analyze images. One basic purpose of image analysis is *sparse representation:* to use as few coefficients as possible to represent the object accurately. We will take some specific images, and analyze them in several different bases, and compare the representations for sparsity.

We consider an image with oriented linear features, depicted in Figure 11

We then analyze the object in the system of anisotropic cosine packets in the frequency domain. The figure 12 below illustrates the results, by showing the pseudopolar Fourier transform with the basis tiling overlaid upon it. The various panels indicate the underlying number of coefficients exceeding a fixed threshold.

To solidify the reader's understanding, Figure 13 presents the present the

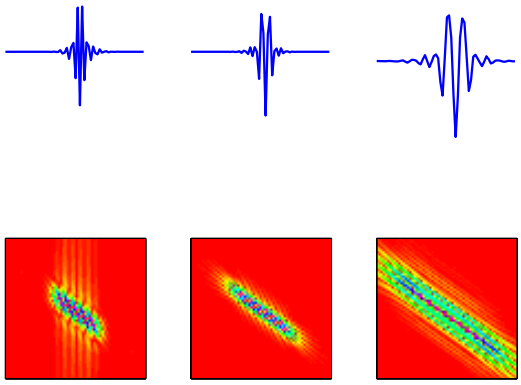

Figure 10. Basis functions derived from Radon-domain approach. Backprojections of Wavelet Packets at Different Widths in the ridge direction give different spatial waveforms.

associated graphic for the image data in original (Cartesian-product) Fourier space. The same four tilings are presented overlaid on the Fourier transform of the original image. The strong directionality of the image is evident from the concentration of energy along a line through the origin.

7. Adaptation

We have defined, in the continuum case, a rather large collection of orthonormal bases; and in the digital data, a rather large collection of frames. Which of these is 'best' for a given dataset? In this section we describing an efficient computational method for finding an optimizing frame among the ridgelet packet bases.

7.1. Background on Best Basis

This question of adaptively choosing a basis has already been thoroughly studied in the context of time-frequency decompositions. The Cosine Packets Library and the Wavelet Packets Library define rather large collections of bases for representing signals – functions of time. Coifman and Wickerhauser [10] considered a library of bases $\mathcal{L} = \{\mathcal{B}\}$ with special properties (which were exhibited by both the cosine packets and wavelet packets libraries) and showed how to rapidly compute a 'best basis', one optimizing an expression of the form

$$\max_{\mathcal{B}\in\mathcal{L}} \mathcal{E}(\mathcal{B})$$

Figure 11. Image with Oriented Linear Features.

where $\mathcal{E}(\mathcal{B})$ (the 'entropy') is a measure of the quality of the basis. The same algorithm works equally well to find the minimizing basis.

Their algorithm works under these assumptions:

- Each basis is a particular subset from a dictionary of ϕ_γ, $\gamma \in \Gamma$;
- The bases in the library have a natural tree structure, i.e. there is a certain correspondence between subtrees of a complete tree and bases in a library;
- The 'entropy' is defined by an additive measure

$$\mathcal{E}(\mathcal{B}) = \sum_{\gamma \in \mathcal{B}} e_\gamma.$$

The underlying principle that makes the Coifman-Wickerhauser algorithm work is the principle of dynamic programming, which allows to prove that bottom-up pruning of the complete tree is optimal.

The Coifman-Wickerhauser algorithm was developed originally in the context of time-frequency analysis, in connection with the Cosine Packet or Wavelet Packet bases. There the tree-ordering relationship of the bases was based on the tree-ordering property of the corresponding recursive dyadic partitions of the time and/or frequency domains, depending on whether we are considering the Cosine Packet or Wavelet Packet systems. The same general principle works also in two-dimensions, and has been used for example in constructing 2-d cosine packets bases for fingerprint image analysis [36] and in constructing 2-d brushlet bases [12]. For example, if one considers direct products of univariate 'time-frequency' bases, these correspond to direct products of pairs of univariate partitions, and to a natural tree order where one allows either vertical or

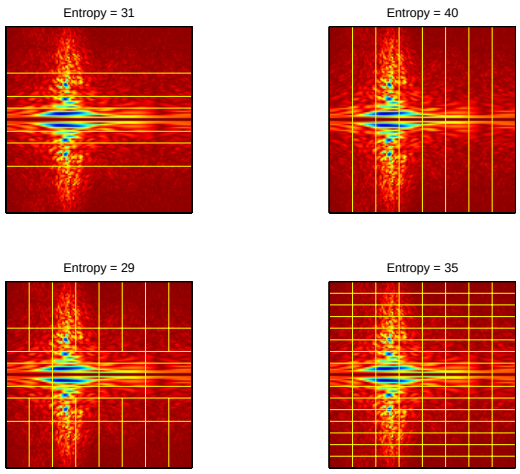

Figure 12. The pseudopolar Fourier transform overlaid with tilings defining 4 different ridgelet packet bases; the labels give the number of coefficients above threshold. Apparently, a decomposition splitting only on directions gives the sparsest representation.

horizontal splits. Other space-frequency bases can be constructed by recursive dyadic partitioning based on quadtree-splits – simultaneous splitting in both variables. Either family has the required tree property [4].

There are two general ways in which the Coifman-Wickerhauser algorithm can be applied, depending on the way in which the e_γ are specified.

First, one could be interested in finding a best-basis for an individual signal f, and then it makes sense to choose e_γ as a function of the coefficients of the signal f in basis $\mathcal{B}$. If θ_γ is the γ-th coefficient of f in basis $\mathcal{B}$, then it would be appropriate to set $e_\gamma = e(|\theta_\gamma|)$, where $e(t)$ is a concave function of t^2 – examples being $e(t) = \min(t^2, \lambda^2)$ and $e(t) = |t|^p$, $0 < p < 2$. Such are measures of the sparsity of the coefficients, and, subject to a fixed budget of coefficient energy $\sum_{\gamma\in\mathcal{B}} |\theta_\gamma|^2$ they are small when the object has sparse coefficients – a few large coefficients and many small ones. (Note that in this setting our goal would be to minimize the $\sum_{\gamma\in\mathcal{B}} e_\gamma$ rather than maximize it.)

Second, one could be interested in finding a best-basis for an ensemble of signals $\{f_m, m = 1, \ldots, M\}$. Then if θ_γ^m denotes the γ-th coefficient of f^m in basis $\mathcal{B}$, it could make sense to let $e_\gamma = Ave\{e(\theta_\gamma^m)\}$, where e is one of the functions mentioned above in the single signal case, and then to minimize $\mathcal{E}(\mathcal{B})$. Alternatively, we could take a statistical viewpoint, and look for a basis capturing the most squared variance on the diagonal; we would then let $e_\gamma = (Var_m\{\theta_\gamma^m\})^2$, and seek a basis maximizing $\sum_{\gamma\in\mathcal{B}} e_\gamma$. The basis then has an

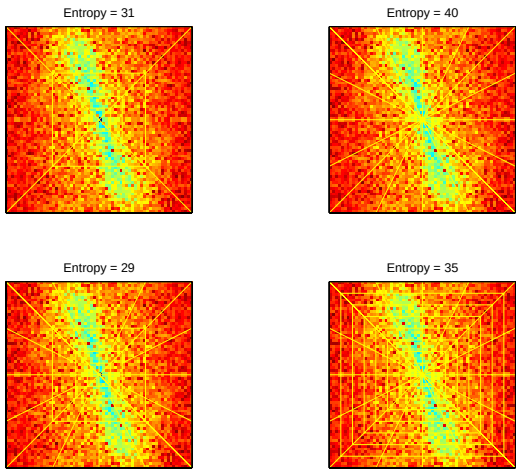

Figure 13. The two-dimensional Fourier transform overlaid with tilings defining 4 different ridgelet packet bases; the labels give the number of coefficients above threshold. Apparently, a decomposition splitting only on directions gives a larger number of coefficients exceeding threshold; while a decomposition splitting into more angular bins at higher frequencies gives fewer coefficients exceeding threshold..

interpretation as providing a best near-diagonalization within the cosine packets system [25]. In this case it also makes sense to consider maximizing the entropy derived from averaging $e(t) = t^4 - 3 \cdot t^2$ across the dataset; then $e_\gamma = Ave\{(\theta_\gamma^m)^4 - 3 \cdot (\theta_\gamma^m)^2\}$ is a measure of kurtosis, and the maximizing basis is the most kurtic basis; such bases are interesting as providing the bases which best expose the non-Gaussianity of the signal in a certain sense; compare [5, 6].

7.2. Application to Ridgelet Packets

We can adapt these existing best basis ideas to the Ridgelet Packet system in a natural way. In essence, Ridgelet Packet bases are all based on taking the data into the pseudopolar Fourier domain, viewing the pseudopolar FT as a complex valued $2n$ by $2n$ image, and then looking for the best anisotropic cosine packets basis for that image data. Bennett's thesis [4] has already explored the properties of best anisotropic cosine packets bases for real valued data; he studied the natural 2-D analog of the Coifman-Wickerhauser algorithm in which anisotropic rectangles are allowed in the tiling. The extension to complex-valued data involves no new issues. Hence, applying the Bennett algorithm to the pseudopolar FFT involves nothing substantially new.

Figure 14 below shows the results of applying best anisotropic basis algo-

rithm to the texture image of Figure 11.

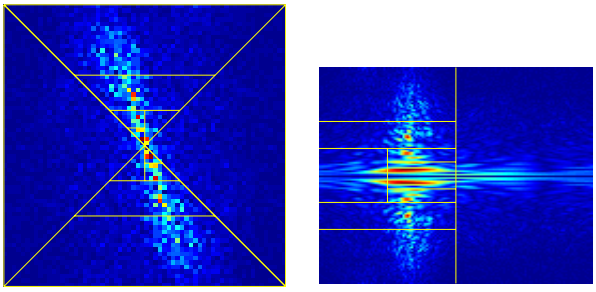

Figure 14. Left panel: Best rectangular partition overlaid on pseudopolar Fourier Plane. Right panel: Corresponding tiling of Cartesian Fourier Plane.

8. Discussion

This is obviously preliminary work; we can expect our efforts to be rapidly surpassed in various ways. We mention here some of the issues that come up in attempting to improve on the scheme developed here.

8.1. Improvements in the Digital Implementation

- *True Polar FFT.* In this article, we have based our efforts on an implementations using the pseudo Polar FFT and the Slant Stack. These are analogs, for digital data, of the continuum Fourier and Radon transforms. However, they are not precise analogs, since they do not fully represent the polar Fourier domain – using concentric squares rather than concentric circles. The discrepancy between squares and circles is particularly evident in analyzing radially symmetric objects such as a disk, which have Fourier transforms organized in a ring pattern (see Figure 4).

  In separate work, Averbuch et al. have developed a Polar FFT algorithm which starts from the pseudopolar FFT and then converts to a true polar form. While this algorithm does not offer a near isometry, the improved geometric fidelity might be important to have. A valuable next step would be to apply the Polar FFT algorithm in the present context.

- *Improved Coefficient Decay.* In the companion article [17], it was pointed out that the digital ridgelet transform based on ideas similar to those used in this paper does not have the same qualities of coefficient decay that were available in theory for the ortho ridgelet transform of continuous functions.

One can expect the same statement to be true for any of the frames constructed using methods of this article. Donoho and Flesia speculate that the slower decay is owing to an implicit 'image mutilation' that occurs at the heart of the Slant Stack algorithm. This interesting possibility should be pursued further; perhaps it would lead to improved sparsity of not only ridgelet but all ridgelet packet representations.

- *True Space-Frequency Atoms.* As pointed out recently by Lars Villemoes, the usual setting for application of the Coifman-Wickerhauser algorithm with Cosine produces bases where the different elements do not have uniform Hesienberg constants. Roughly speaking, in the usual application, one has to make a choice of a "window transition length", and then all windows in the whole system make the transition from 1 to 0 in that same length, no matter whether the window is short or long. The result is that the basis elements corresponding to the long intervals have poor time-frequency localization.

  To respond to this problem, Villemoes [34] has developed an algorithm that provides uniformly good time-frequency localization. If Villemoes' algorithm were used in the ridgelet packet setting, it might substantially improve the coefficient decay of the best ridgelet packet basis. This would be a valuable direction to explore.

8.2. Limitations on the Ridgelet Packet Scheme

While the Ridgelet Packet system provides an interesting collection of new bases, all these bases are missing a key ingredient: a *translation parameter*. At best, the scheme offers a ridge translation parameter, associated with displacements vertically up or down in ridge space. From wavelets and Gabor systems, we are accustomed to having a translation-like parameter.

To understand this issue better, look at a typical member of a wavelet system in Fourier space;

$$\hat{\psi}_{j,k_1,k_2}(\xi) = \exp\{i\xi'(k_1,k_2)/2^j\} \cdot \hat{\psi}(\xi/2^j)/2^{j/2}.$$

The oscillating factor, $\exp\{i\xi'(k_1,k_2)/2^j\}$ is responsible for the translation effect. By comparison, a ridgelet packet, viewed in frequency space, is made up of several terms of the form

$$\exp\{ik_1(r-a)/b\} \cdot \exp\{ik_2(\theta-c)/d\} w_r((r-a)/b) \cdot w_\theta((\theta-c)/d).$$

Now qualitatively, the product of two windows $w_r((r-a)/b) \cdot w_\theta((\theta-c)/d)$ is simply a nice smooth function. So the oscillating factor

$$\exp\{ik_1(r-a)/b\} \cdot \exp\{ik_2(\theta-c)/d\}$$

would have to be able to create a translation or at least a pseudo translation effect. However, it is evident by inspection that, for large k_1, k_2 there is no such

effect. Thus for example, under various assumptions on the behavior of a, b, c, d, the family can be shown to behave very differently than a translation family.

In response to this, it would seem worthwhile to search for a system of bases and associated algorithms with a more direct connection to translations.

An obvious way to do this would be to borrow ideas which were useful in the construction of curvelets [9]. Thus, we would separate the image into passbands based on bandpass filtering, spatially localize the bandpass images into square subimages via smooth windowing, and then apply ridgelet packet analysis to the square subimages.

Acknowledgement. This research was supported by National Science Foundation grant DMS 95–05151, DMS 00–7726 and DMS 98–72890 (KDI), by AFOSR MURI 95–P49620–96–1–0028, and by DARPA BAA-99-07. The authors would like to thank Nick Bennett and Yacov Hel-Or for helpful discussions. AGF would like to thank the Statistics Department at UC Berkeley for its hospitality.

References

[1] P. Auscher, G. Weiss and M.V. Wickerhauser. *Local sine and cosine bases of Coifman and Meyer and the construction of smooth wavelets*, Wavelets: A tutorial in Theory and Applications, Academic Press, Boston, 237-256.

[2] E.H. Adelson and E.P. Simoncelli, *Orthogonal pyramid transforms for image coding*, Pro. SPIE, **845** (October 1987), Cambridge, MA.

[3] A. Averbuch, R.R. Coifman, D.L. Donoho, M. Israeli and J. Waldén, *Fast Slant Stack: A notion of Radon Transform for Data in a Cartesian Grid which is Rapidly Computible, Algebraically Exact, Geometrically Faithful and Invertible*, to appear: SIAM J. Sci. Comp.

[4] N. Bennett, *Fast algorithm for best anisotropic Walsh bases and relatives*, Appl. Comput. Harmon. Anal., **8** (2000), no. 1, 86-103.

[5] J.B. Buckheit, *Adaptive Wavelet Methods in Signal Processing*, Ph. D. Thesis, Department of Statistics, Stanford University,1996.

[6] J.B. Buckheit and D.L. Donoho, *WaveLab and Reproducible Research*, Wavelets in Statistics, A. Antoniadis and G. Oppenheim Ed., Springer-Verlag,New York, 1995, 55-82.

[7] E. Candès, *Harmonic Analysis of Neural Networks*, Appl. Comput. Harmon. Anal., **6** (1999), no. 2, 197-218.

[8] E. Candès. *Ridgelets: Theory and Applications*, Ph.D. Thesis, Department of Statistics, Stanford University,1998.

[9] Candès, E.J. and Donoho, D.L. (2000) Curvelets: a surprisingly effective nonadaptive representation of objects with edges. in *Curve and Surface Fitting: Saint-Malo 1999* Albert Cohen, Christophe Rabut, and Larry L. Schumaker (eds.) Vanderbilt University Press, Nashville, TN. ISBN 0-8265-1357-3

[10] R.R. Coifman and M.V. Wickerhauser, *Entropy-based algorithms for best basis selection*, IEEE Transactions on Information Theory, **38** (No.2 pt.2) (March 1992), 713-18.

[11] R.R. Coifman and Y. Meyer, *Remarques sur l'Analyse de Fourier à fenêtre*, C.R. Acad Sci. Paris,**312** (1991), 259-261.

[12] R.R. Coifman and F. Meyer, *Brushlets: a tool for directional image analysis and image compression.* Applied and Computational Harmonic Analysis, **4** (No.2)(April 1997),147-87.

[13] I. Daubechies, *The wavelet transform, time-frequency localization, and signal analysis.*, IEEE Trans. IT. ,**36** (1990), 961-1005.

[14] I. Daubechies, S. Jaffard and J.L. Journé, *A simple Wilson orthonormal basis with exponential decay.* SIAM J. Math. Anal., **24** (1990), 520-527.

[15] D.L. Donoho, *Orthonormal ridgelets and linear singularities*, SIAM J. Math. Anal. **31** (2000), no. 5, 1062-1099.

[16] D.L.Donoho,*Ridge functions and orthonormal ridgelets*, J. Approx. Theory., **111** (2001), no. 2, 143-179.

[17] D.L. Donoho and A.G. Flesia, *Digital Ridgelet Transform Based on True Ridge Functions*, To appear, Beyond Wavelets, J. Schmeidler and G.V. Welland, Eds., Academic Press, 2002.

[18] P. Edholm and G.T. Herman, *Linograms in image reconstruction from projections*, IEEE Trans. Medical Imaging, MI-6 (No. 4)(1987),301-307.

[19] D.J. Field, *Relations between the statistics of natural images and the response profiles of cortical cells*, Journal of the Optical Society of America A, **4** (1987), 2379-2394.

[20] D.J. Field, *Scale-invariance and self-similar 'wavelet' transforms: an analysis of natural scenes and mammalian visual systems*, Wavelets, Fractals and Fourier Transforms: New Developments and New Applications, Oxford University Press,1993.

[21] D.J. Field, *What is the goal of sensory coding?*, Neural Computation, **6** (1994), 559-601.

[22] D. Heeger and J. Bergen, *Pyramid based texture analysis/synthesis*, Computer Graphics, (1995), 229-238, SIGGRAPH 95.

[23] P.G. Lemarié and Y. Meyer, *Ondelettes et bases Hilbertiennes*, Revista Matematica Iberoamericana, **2** (1986), 1-18.

[24] W. Lawton, *A new polar Fourier transform for computer-aided tomography and spotlight synthetic aperture radar*, IEEE Trans. Acoustics Speech Signal Process., **36** (No. 6)(1998), 931-33.

[25] S.G. Mallat, G. Papanicolaou, and Z. Zhang, *Adaptive covariance estimation of locally stationary processes.* Ann. Statist. , **26** (February 1998).

[26] Y. Meyer, *Ondelettes et algorithmes concurrents*, Hermann, Paris, 1992.

[27] J.E. Pasciak, *A note on the Fourier algorithm for image reconstruction*, Preprint AMD 896 Applied Mathematics Department, Brookhaven National Laboratory, Upton, New York 11973, 1981.

[28] K. Ramchandran and M. Vetterli, *Best wavelet packet bases in a rate-distortion sense*, IEEE Trans. Image Proc., **2** (No. 2) (1993), 160-175.

[29] E. Simoncelli and H. Farid, *Steerable wedge filters* Int'l Conf. on Computer Vision, Boston MA. June 1995.

[30] E.P Simoncelli, W.T Freeman, E.H Adelson and D.J Heeger. *Shiftable multi-scale transforms [or "What's wrong with orthonormal wavelets.* IEEE Trans. Information Theory, Special Issue on Wavelets. **38**(No. 2) (March 1992), 587-607.

[31] E.M. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton Univ. Press, 1993.

[32] C. Sogge, *Fourier Integrals in Classical Analysis*, Cambridge University Press, 1993.

[33] B. Torresani, *Time-frequency representations: wavelet packets and optimal decomposition*, Annales de l'Institut Henri Poincare (Physique Théorique), **56** (No. 2) (1992), 215-34.

[34] L.F. Villemoes, *Adapted bases of time-frequency local cosines*, Appl. Comput. Harmon. Anal. **10** (No. 2) (2001), 139-162.

[35] A.B. Watson, *The cortex transform: rapid computation of simulated neural images*, Computer Vision, Graphics, and Image Processing, **39** (No. 3) (Sept. 1987), 311-27.

[36] M.V. Wickerhauser, *Adapted Wavelet Analysis from Theory to Software*, A K Peters, Limited, 1994.

A. Averbuch
Computer Science Department
Tel Aviv University
Tel Aviv 69978, Israel
amir@math.tau.ac.il

E.J. Candès
Applied and Computational Mathematics
California Institute of Technology
1200 E. California Boulevard, MC 217-50 Pasadena, CA 91125, USA
emmanuel@acm.caltech.edu

R.R. Coifman
Department of Mathematics
Yale University
PO Box 208283 New Haven, CT 06520-8283, USA.
coifman@math.yale.edu

D.L. Donoho and A.G. Flesia
Department of Statistics
Stanford University
Sequoia Hall, 390 Serra Mall, Stanford, CA 94305-4065, USA
{donoho,flesia}@stanford.edu

H. Hel-Or
Department of Computer Science
Haifa University
Haifa 31905, Israel
hagit@cs.haifa.ac.il