# Contourlets

Minh N. Do
Martin Vetterli

**Abstract.** This chapter focuses on the development of a new "true" two-dimensional representation for images that can capture the intrinsic geometrical structure of pictorial information. Our emphasis is on the discrete framework that can lead to algorithmic implementations. We propose a double filter bank structure, named the *pyramidal directional filter bank*, by combining the Laplacian pyramid with a directional filter bank. The result is called the *contourlet transform*, which provides a flexible multiresolution, local and directional expansion for images. The contourlet transform can be designed to satisfy the anisotropy scaling relation for curves, and thus offers a fast and structured curvelet-like decomposition sampled signals. As a result, the proposed transform provides a sparse representation for two-dimensional piecewise smooth signals that resemble images. The link between the developed filter banks and the continuous-space constructions is set up precisely in a newly defined directional multiresolution analysis. Finally, we show some numerical experiments demonstrating the potential of the new transform in several image processing tasks.

## 1. Introduction and Motivation

We are interested in the construction of *efficient linear expansion* for *two-dimensional signals, which are smooth away from discontinuities across smooth curves*. Such signals resemble natural images where discontinuities are generated by *edges* – referred to the points in the image where there is a sharp contrast in the intensity, whereas edges are often gathered along smooth *contours*, which are created by typically smooth boundaries of physical objects. Efficiency of a linear expansion means that the coefficients for signals belonging to the class of interest are *sparse*, and thus it implies efficient representations for such functions using a non-linear approximation (NLA) scheme.

Over the last decade, wavelets have had a growing impact on signal processing, mainly due to their good NLA performance for piecewise smooth functions in one dimension [9, 16, 19]. Unfortunately, this is not the case in two dimensions. In essence, wavelets are good at catching point or zero-dimensional discontinuities, but as mentioned above, two-dimensional piecewise smooth functions resembling images have one-dimensional discontinuities. Intuitively, wavelets in 2-D obtained by a tensor-product of one dimensional wavelets will be good at isolating the discontinuities at edge points, but will not see the smoothness along



Beyond Wavelets
J. Stoeckler and G. V. Welland (Eds.), pp. 1–27.

ISBN 0-12-xxxxxx-x

the contours. This indicates that more powerful representations are needed in higher dimensions.

Recently, Candès and Donoho [4, 6] pioneered a new system of representation, named *curvelet*, that was shown to achieve optimal approximation behavior in a certain sense for 2-D piecewise smooth functions in $\mathbb{R}^2$ where the discontinuity curve is a $C^2$ function.[1] More specifically, an $M$-term non-linear approximation for such piecewise smooth functions using curvelets has $L^2$ square error decaying like $O(M^{-2})$, and this is the best rate that can be achieved by a large class of approximation processes [14]. An attractive property of the curvelet system is that such correct approximation behavior is simply obtained via thresholding a *fixed* transform. The key features of the curvelet elements is that they exhibit very high directionality and anisotropy.

The original construction of the curvelet transform [4] was intended for functions defined in the *continuum* space $\mathbb{R}^2$. The development of *discrete* transforms for sampled images that has all the features promised by curvelets in the continuous domain remains a challenge, especially when critical sampling is desirable. Furthermore, as the curvelet transform was defined in the frequency domain, it is not clear how curvelets are sampled in the spatial domain. In fact, in [15], one of the fundamental research challenges for curvelets was stated as: "is there a spatial domain scheme for refinement which, at each generation doubles the spatial resolution as well as the angular resolution?". This is what we will try to explore in the following.

First, we will identify the key features that make curvelets to be an efficient representation for 2-D piecewise smooth functions with smooth discontinuity curves. Based on this, we propose a filter bank structure that can deal effectively with piecewise smooth images with smooth contours. The resulting image expansion is a frame composed of contour segments, and thus is named *contourlet*. We then derive an analysis framework that connects the proposed discrete transform to the frames in the continuous-domain, which can be particularized to a curvelet-like expansion. Thus our scheme provides an effective method to implement the discrete curvelet transform. Furthermore, the resulting transform has very small redundancy, being almost critically sampled. Finally, we will show some numerical experiments demonstrating the potential of the contourlet transform in several image processing tasks.

## 2. Representing 2-D Piecewise Smooth Functions

### 2.1. Curvelet construction

In a nutshell, the curvelet transform [4] is obtained by filtering and then applying a windowed ridgelet transform [5] to each bandpass image. In $\mathbb{R}^2$, ridgelets are constant along ridge lines $x_1 \cos(\theta) + x_2 \sin(\theta) = const$ and are wavelets (with a scale $s$) along the orthogonal direction. In frequency domain, such ridgelet

[1] $C^p$ is the space of functions that are bounded and $p$-times continuously differentiable.

function is essentially localized in the corona $|\boldsymbol{\omega}| \in [2^s, 2^{s+1}]$ and around the angle $\theta$. The ridgelet transform provides a sparse representation for smooth objects with straight edges. In summary, the curvelet decomposition is composed of the following steps [4] (also see Figure 2(a)):

1. Subband decomposition of the object into a sequence of subbands.
2. Windowing each subband into blocks of appropriate size, depending on its center frequency.
3. Applying the ridgelet transform to these blocks.

The motivation behind the curvelet transform is that by smooth windowing, segments of smooth curves would look straight in sub-images, hence they can be captured efficiently by a local ridgelet transform. Subband decomposition is used to keep the number of ridgelets at multiple scales under control by the fact that ridgelets of a given scale live in a certain subband. The window's size and subband frequency are coordinated such that curvelets have support obeying the key *anisotropy scaling relation* for curves [4, 6]:

$$width \propto length^2. \tag{2.1.1}$$

### 2.2. Non-linear approximation behaviors

We next sketch illustrations on the non-linear approximation behaviors for 2-D piecewise smooth functions using different expansions. Rather than being rigorous, the following discussion aims at providing an intuition that can serve as a guideline for our construction of the pyramidal directional filter banks and contourlets latter. For a complete and rigorous discussion, we refer to [14].

Consider a simple "Horizon" model of piecewise smooth functions $f(x_1, x_2)$ defined on the unit square $[0, 1]^2$:

$$f(x_1, x_2) = 1_{\{x_2 \geq c(x_1)\}} \qquad 0 \leq x_1, x_2 \leq 1,$$

where the boundary of two pieces (or the contour) $c(x_1)$ is in $C^p$ and has finite length inside the unit square. Clearly, such a 2-D function has complexity equivalent to a 1-D function, namely its contour $c(x_1)$. The reason for studying this model is that the approximation rates for 2-D piecewise smooth functions resembling images are typically dominated by the discontinuity curves.

Let's first consider how a wavelet system performs for such function. Assume that the orthonormal wavelet transform with the separable Haar wavelet is employed. At the level $j$, wavelet basis functions have support on dyadic squares of size $2^{-j}$ (see Figure 1(a)). Let $n_j$ be the number of dyadic squares at level $j$ that intersect with the contour on the unit square. Since the contour has finite length, it follows that

$$n_j \sim O(2^j). \tag{2.2.1}$$

Thus, there are $O(2^j)$ nonzero wavelet coefficients at the scale $2^{-j}$. This is the problem of the separable wavelet transform for 2-D piecewise smooth functions. For the 1-D piecewise smooth function, the number of significant wavelet coefficients at each scale is bounded by a constant; in the 2-D case this number grows exponentially as the scale gets finer. The total number of nonzero wavelet coefficients up to the level $J$ is

$$N_J = \sum_{j=0}^{J} n_j \sim O(2^J). \tag{2.2.2}$$

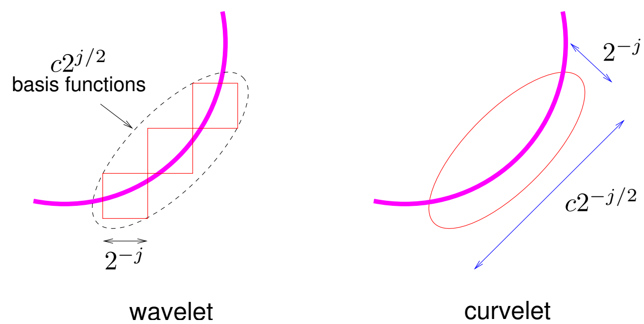



**Figure 1**. Non-linear approximation of a 2-D piecewise smooth function using wavelets and curvelets. Curvelet basis functions can be viewed as a local grouping of wavelet basis functions into linear structures so that they can capture the smooth discontinuity curve more efficiently.

Along the discontinuity curve $c$, it is easy to see that these nonzero wavelet coefficients decay like $O(2^{-j})$ at the $j$-th level. Next, suppose we keep only $M = N_J$ nonzero coefficients up to the level $J$ in the wavelet expansion. Then the error due to truncation of the wavelet series is

$$\begin{aligned} \|f - \hat{f}_M^{(wavelet)}\|^2 &\sim \sum_{j=J+1}^{\infty} 2^j (2^{-j})^2 \\ &\sim O(2^{-J}). \end{aligned} \tag{2.2.3}$$

Combining (2.2.2) and (2.2.3) we obtain the following non-linear approximation rate of the wavelet expansion for the "Horizon" model

$$\|f - \hat{f}_M^{(wavelet)}\|^2 \sim O(M^{-1}). \tag{2.2.4}$$

Therefore, when the discontinuity curves $c$ is sufficiently smooth, $c \in C^p$ with $p > 1$, wavelet approximation is suboptimal. It is important to note that the smoothness of the discontinuity curve is irrelevant to the performance of the wavelet approximation.

How can we improve the performance of the wavelet representation when the discontinuity curve is known to be smooth? Simply looking at the wavelet scheme in Figure 1(a) suggests that rather than treating each significant wavelet coefficient along the discontinuity curve independently, one should group the nearby coefficients since their locations are locally correlated. Recall that at the level $j$, the essential support of the wavelet basis functions has size $2^{-j}$. The curve scaling relation (2.1.1) suggests that we can group about $c2^{j/2}$ nearby wavelet basis functions into one basis function with a linear structure so that its width is proportional to its length squared (see Figure 1). This grouping operation reduces the number of significant coefficients at the level $j$ from $O(2^j)$ to $O(2^{j/2})$. Consequently, this new representation provides the same approximation error as wavelets in (2.2.3) with only $M' \sim \sum_{j=0}^{J} 2^{j/2}$ or $O(2^{J/2})$ coefficients. In other words, the $M$-term non-linear approximation using this improved wavelet representation decays like

$$||f - \hat{f}_M^{(improved-wavelet)}||^2 \sim O(M^{-2}). \tag{2.2.5}$$

Comparing with (2.2.4), we see that for $C^2$ discontinuity curves, the new representation is superior compared to wavelets and in fact achieves the optimal rate. The curvelet system achieves this optimality using a similar argument. In the original curvelet construction [4], the linear structure of the basis function comes from the ridgelet basis while the curve scaling relation is ensured by suitable combination of subband filtering and windowing.

### 2.3. A filter bank approach for sparse image expansions

The original definition of the curvelet transform as described in Section 2.1. poses several problems when one translates it into the discrete world. First, since it is a block-based transform, either the approximated images have blocking effects or one has to use overlapping windows and thus increase the redundancy. Secondly, the use of ridgelet transform, which is defined on a polar coordinate, makes the implementation of the curvelet transform for discrete images on rectangular coordinates very challenging. In [15, 25, 1], different interpolation approaches were proposed to solve the polar versus rectangular coordinate transform problem, all required overcomplete systems. Consequently, the version of the discrete curvelet transform in [25] for example has a redundancy factor equal to $16J + 1$ where $J$ is the number of multiscale levels.

Comparing the wavelet scheme with the curvelet scheme in Figure 1, we see that the improvement of curvelets can be loosely interpreted as a grouping of nearby wavelet coefficients, since their locations are locally correlated due to the smoothness of the discontinuity curve. Therefore, we can obtain a sparse image expansion by first applying a multiscale transform and then applying a local directional transform to gather the nearby basis functions at the same scale into linear structures. In essence, we first use a wavelet-like transform for *edge* detection, and then a local directional transform for *contour segment* detection.

Interestingly, this approach is similar to the popular Hough transform [17] for line detection in computer vision.

With this insight, we proposed a *double filter bank* approach for obtaining sparse expansions for typical images with smooth contours (Figure 2(b)). In our newly constructed *pyramidal directional filter bank* [11], the Laplacian pyramid [3] is first used to capture the point discontinuities, then followed by a directional filter bank [2] to link point discontinuities into linear structures. The overall result is an image expansion using elementary images like contour segments, and thus it is named the *contourlet transform*.

The contourlet transform offers a flexible multiresolution and directional decomposition for images, since it allows for a different number of directions at each scale. For the contourlet transform to satisfy the *anisotropy scaling law*, as in the curvelet transform, we simply need to impose that the number of directions is doubled at every *other* finer scale of the pyramid [11].

The contourlet transform is almost *critically sampled*, with a small redundancy factor of up to 1.33. Comparing this with a much larger redundancy ratio of the discrete implementation of the curvelet transform [25] mentioned above, the contourlet transform is much more suitable for image compression. Furthermore, the contourlet transform can be designed to be a *tight frame*, which implies robustness against the noise due to quantization or thresholding. Finally, the contourlet transform is implemented efficiently via iterated filter banks with fast algorithms. In the next section we will describe such a filter filter bank in detail.

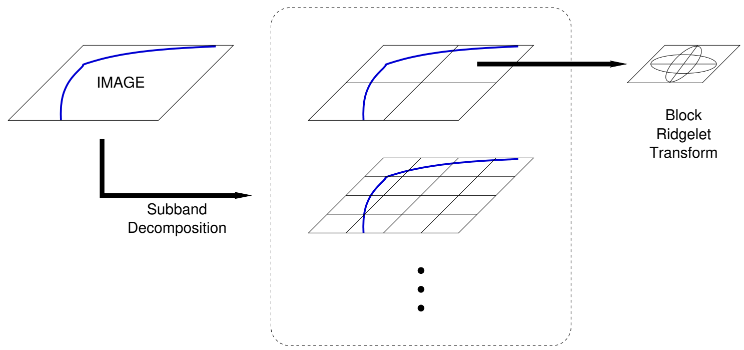
## 3. Pyramidal Directional Filter Bank

### 3.1. Multiscale decomposition

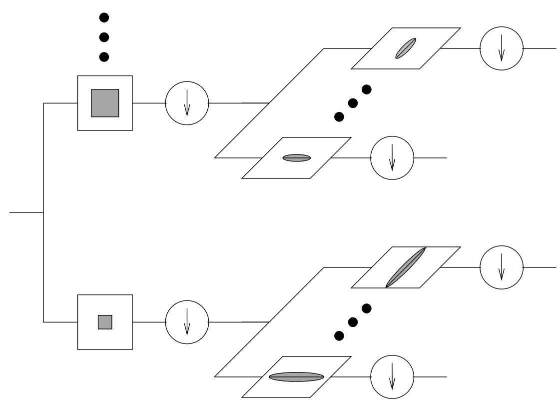
One way of achieving a multiscale decomposition is to use a Laplacian pyramid (LP) as introduced by Burt and Adelson [3]. The LP decomposition at each step generates a sampled lowpass version of the original and the difference between the original and the prediction, resulting in a bandpass image (see Figure 3(a)). The process can be iterated on the coarse version.

A drawback of the LP is the implicit oversampling. However, in contrast to the critically sampled wavelet scheme, the LP has the distinguishing feature that each pyramid level generates only one bandpass image (even for multidimensional cases) which does not have “scrambled” frequencies. This frequency scrambling happens in the wavelet filter bank when a highpass channel, after downsampling, is folded back into the low frequency band, and thus its spectrum is reflected. In the LP, this effect is avoided by downsampling the lowpass channel only.

In [13], we study the LP using the theory of frames and oversampled filter



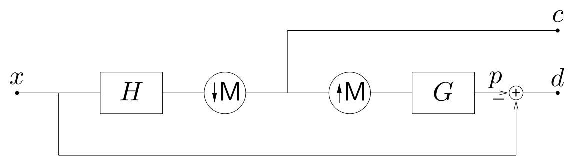

(a) Curvelet transform



(b) Contourlet transform

**Figure 2**. Two approaches for dealing with images having smooth contours. (a) *Curvelet transform:* block ridgelet transforms are applied to subband images. (b) *Contourlet transform:* image is decomposed by a double filter-bank structure, where the first one captures the edge points and the second one links these edge points into contour segments. The gray areas in the boxes represent the support sizes of the filters..

banks. We show that the LP with orthogonal filters (that is, $h[\boldsymbol{n}] = g[-\boldsymbol{n}]$ and $g[\boldsymbol{n}]$ is orthogonal to its translates with respect to the subsampling lattice) is a tight frame with frame bounds equal to 1. In this case, we suggest the use of the optimal linear reconstruction using the dual frame operator, which is symmetrical with the forward transform (see Figure 3(b)). Note that this new reconstruction is different from the usual reconstruction and is crucial for our contourlet expansion described later.



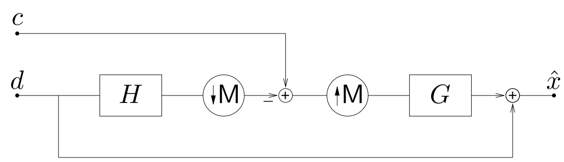




**Figure 3**. Laplacian pyramid scheme. (a) Analysis: the outputs are a coarse approximation $c$ and a difference $d$ between the original signal and the prediction. The process can be iterated by decomposing the coarse version repeatedly. (b) The proposed reconstruction scheme for the Laplacian pyramid.

### 3.2. Directional decomposition

In 1992, Bamberger and Smith [2] introduced a 2-D directional filter bank (DFB) that can be maximally decimated while achieving perfect reconstruction. The DFB is efficiently implemented via a $l$-level tree-structured decomposition that leads to $2^l$ subbands with wedge-shaped frequency partition as shown in Figure 4.

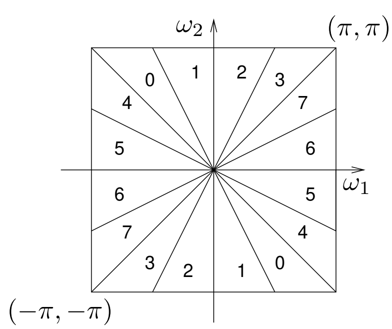



**Figure 4**. Directional filter bank frequency partitioning where $l = 3$ and there are $2^3 = 8$ real wedge-shaped frequency bands.

The original construction of the DFB in [2] involves modulating the input signal and using diamond-shaped filters. Furthermore, to obtain the desired frequency partition, an involved tree expanding rule has to be followed (see [22, 21] for details). As a result, the frequency regions for the resulting subbands do not follow a simple ordering as shown in Figure 4 based on the channel indices.

In [10, 12], we propose a new formulation for the DFB that is based only on the QFB's with fan filters. The new DFB avoids the modulation of the input image and has a simpler rule for expanding the decomposition tree. Intuitively, the wedge-shaped frequency partition of the DFB is realized by an appropriate combination of directional frequency splitting by the fan QFB's and the "rotation" operations done by resampling, which are illustrated in Figure 5 and Figure 6, respectively.

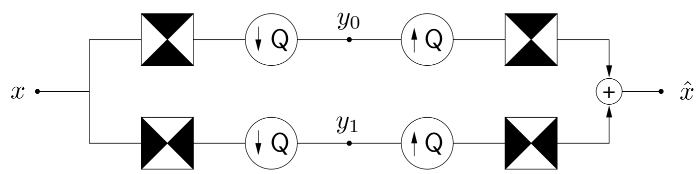



**Figure 5**. Two-dimensional spectrum splitting using the quincunx filter banks with fan filters. The black regions represent the ideal frequency supports of each filter.

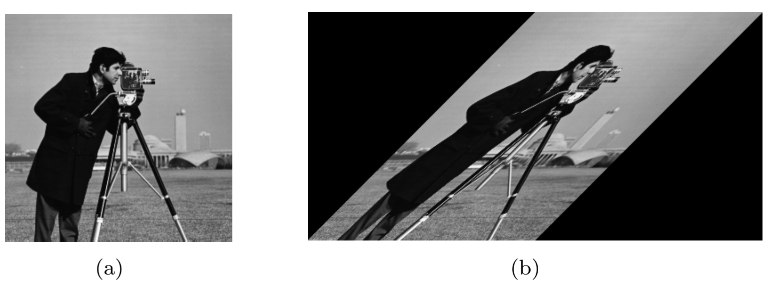



**Figure 6**. Example of a resampling operation that is used effectively as a rotation operation for the DFB decomposition. (a) The "cameraman" image. (b) The "cameraman" image after being resampled.

Using the multirate identities, we can transform a $l$-level tree-structured DFB into a parallel structure of $2^l$ channels with equivalent filters and overall sampling matrices. Denote these equivalent synthesis filters as $G_k^{(l)}$, $0 \leq k < 2^l$, which correspond to the subbands indexed as in Figure 4. The oversampling matrices have diagonal form as:

$$\mathsf{S}_k^{(l)} = \begin{cases} \operatorname{diag}(2^{l-1}, 2) & \text{for } 0 \leq k < 2^{l-1} \\ \operatorname{diag}(2, 2^{l-1}) & \text{for } 2^{l-1} \leq k < 2^l, \end{cases} \tag{3.2.1}$$

which correspond to the basically horizontal and basically vertical subbands, respectively.

With this, it is easy to see that the family

$$\left\{ g_k^{(l)}[n - \mathsf{S}_k^{(l)} m] \right\}_{0 \le k < 2^l,\ m \in \mathbb{Z}^2}, \tag{3.2.2}$$

obtained by translating the impulse responses of the synthesis filters $G_k^{(l)}$ over the sampling lattices $\mathsf{S}_k^{(l)}$, is a *basis* for discrete signals in $l^2(\mathbb{Z}^2)$. This basis exhibits both directional and localization properties. Figure 7 demonstrates this fact by showing the impulse responses of equivalent filters from an example DFB. These basis functions have linear supports in space and span all directions. Therefore (3.2.2) resembles a local Radon transform and the basis functions are referred to as *Radonlets*.

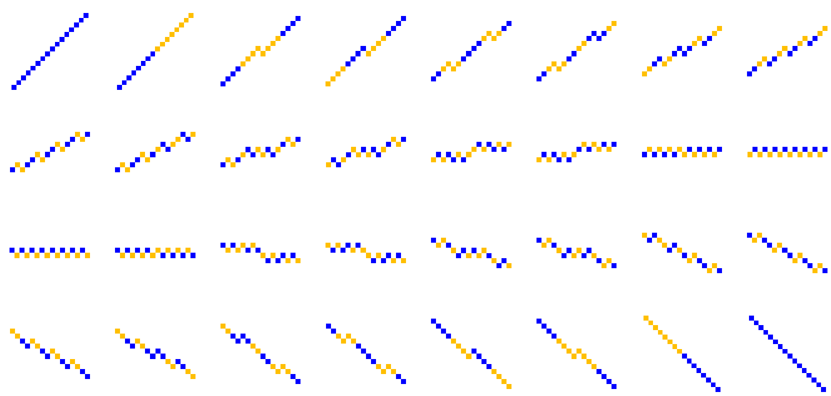


**Figure 7**. Impulse responses of 32 equivalent filters for the first half channels of a 6-levels DFB that use the Haar filters. Black and gray squares correspond to $+1$ and $-1$, respectively. Because the basis functions resemble "local lines", we call them *Radonlets*.

## 3.3. Multiscale and directional decomposition

The directional filter bank (DFB) is designed to capture the high frequency components (representing directionality) of images. Therefore, low frequency components are handled poorly by the DFB. In fact, with the frequency partition shown in Figure 4, low frequencies would "leak" into several directional subbands, hence DFB does *not* provide a sparse representation for images. To improve the situation, low frequencies should be removed before the DFB. This provides another reason to combine the DFB with a multiresolution scheme.

Therefore, the LP permits further subband decomposition to be applied on its bandpass images. Those bandpass images can be fed into a DFB so that directional information can be captured efficiently. The scheme can be iterated repeatedly on the coarse image (see Figure 8). The end result is a double iterated filter bank structure, named *pyramidal directional filter bank* (PDFB), which decomposes images into directional subbands at multiple scales. The scheme is flexible since it allows for a different number of directions at each scale.

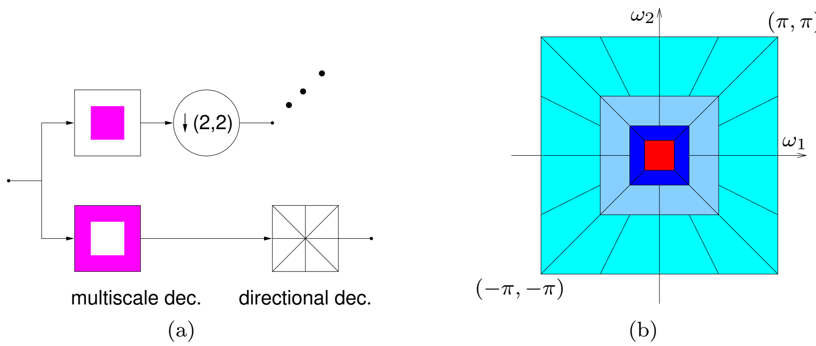



**Figure 8**. Pyramidal directional filter bank. (a) Block diagram. First, a standard multiscale decomposition into octave bands is computed, where the lowpass channel is subsampled while the highpass is not. Then, a directional decomposition with a DFB is applied to each highpass channel. (b) Resulting frequency division, where the number of directions is increased with frequency.

With perfect reconstruction LP and DFB, the PDFB is obviously perfect reconstruction, and thus it is a frame operator for 2-D signals. The PDFB has the same redundancy as the LP: up to 33% when subsampling by two in each dimension. Combining the tight frame and orthogonal conditions for the LP and DFB, respectively, it is easy to obtain the following result for the PDFB [11].

***Proposition 1.*** *The PDFB is a tight frame with frame bounds equal to* 1 *when orthogonal filters are used in both the LP and the DFB.*

Let us point out that there are other multiscale and directional decompositions such as the cortex transform [28] and the steerable pyramid [24]. Our PDFB differs from those in that it allows different number of directions at each scale while nearly achieving critical sampling. In addition, we make the link to continuous-domain construction in Section 4.

### 3.4. PDFB for curvelets

Next we will demonstrate that a PDFB where *the number of directions is doubled at every* ***other*** *finer scale in the pyramid* satisfies the key properties of curvelets discussed in Section 2.1.. That is, we apply a DFB with $\lfloor n_0 - j/2 \rfloor$ levels or $2^{\lfloor n_0 - j/2 \rfloor}$ directions to the bandpass image $b_j$ of the LP. Thus, the PDFB provides an efficient discrete implementation for the curvelet transform.

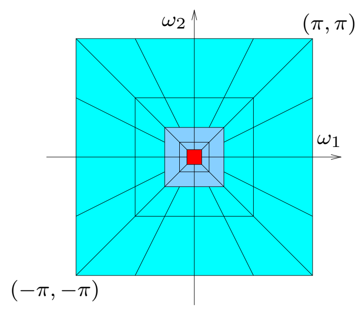



**Figure 9**. Resulting frequency division by a pyramidal directional filter bank for the curvelet transform. As the scale is refined from coarse to fine, the number of directions is doubled at every **other** octave band.

A LP, with downsampling by two in each direction, is taken at every level, providing an octave-band decomposition: the LP bandpass image $b_j$ at the level $j$ creates a subband with a corona support based on the interval $[\pi 2^{-j}, \pi 2^{-j+1}]$, for $j = 1, 2, \ldots, J$. Combining this with a directional decomposition by a DFB, we obtain the frequency tiling for curvelets as shown in Figure 9.

In terms of basis functions, a coefficient in the LP subband $b_j$ corresponds to a basis function that has local support in a square of size about $2^j$. Then, a basis function from a DFB with $\lfloor n_0 - j/2 \rfloor$ iterated levels has support in a rectangle of length about $2^{n_0 - j/2}$ and width about 1. Therefore, in the PDFB, a basis function at the pyramid level $j$ has support as:

$$width \approx 2^j \quad \text{and} \quad length \approx 2^j . 2^{n_0 - j/2} = 2^{n_0} 2^{j/2}, \tag{3.4.1}$$

which clearly satisfies the anisotropy scaling relation (2.1.1) of curvelets.

Figure 10 graphically depicts this property of a PDFB implementing a curvelet transform. As can be seen from the two pyramidal levels shown below, the support size of the LP is reduced by four times while the number of directions of the DFB is doubled. With this, the support size of the PDFB basis images are changed from one level to next in accordance with the curve scaling relation. Also note that in this representation, as the scale is getting finer, there are more directions.

## 4. Multiresolution Analysis

As for the wavelet filter bank, the iterated PDFB can be associated with a continuous-domain system, which we call *contourlet*. This connection will be

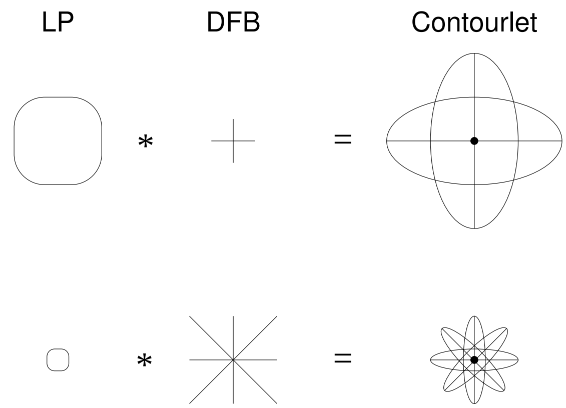



**Figure 10**. Illustration of the contourlet basis images that satisfy the curve scaling relation. From the upper line to the lower line, the scale is reduced by four while the number of directions is doubled.

made precise by studying the embedded grids of approximation as in the multiresolution analysis for wavelets [18, 20]. The new elements are multiple directions and the combination with multiscale.

### 4.1. Multiscale

Suppose that the LP in the PDFB uses orthogonal filters and downsampling by two is taken in each dimension. Under certain conditions, the lowpass filter $G$ in the LP uniquely defines an orthogonal scaling function $\phi(t) \in L^2(\mathbb{R}^2)$ via the two-scale equation [26, 19]

$$\phi(t) = 2 \sum_{n \in \mathbb{Z}^2} g[n]\phi(2t - n)$$

Denote

$$\phi_{j,n} = 2^{-j}\phi\left(\frac{t - 2^j n}{2^j}\right), \qquad j \in \mathbb{Z}, n \in \mathbb{Z}^2. \tag{4.1.1}$$

Then the family $\{\phi_{j,n}\}_{n\in\mathbb{Z}^2}$ is an orthonormal basis of $V_j$ for all $j \in \mathbb{Z}$. The sequence of nested subspaces $\{V_j\}_{j\in\mathbb{Z}}$ satisfies the following invariance properties:

$$\begin{aligned} &\text{Shift invariance:} \quad f(t) \in V_j \quad \Leftrightarrow \quad f(t - 2^j k) \in V_j, \quad \forall j \in \mathbb{Z}, k \in \mathbb{Z}^2 \\ &\text{Scale invariance:} \quad f(t) \in V_j \quad \Leftrightarrow \quad f(2^{-1}t) \in V_{j+1}, \quad \forall j \in \mathbb{Z}. \end{aligned}$$

In other words, $V_j$ is a subspace defined on a uniform grid with intervals $2^j \times 2^j$, which characterize the image approximation at the resolution $2^{-j}$. The difference image in the LP carries the details necessary to increase the resolution of an image approximation. Let $W_j$ be the orthogonal complement of $V_j$ in $V_{j-1}$ (also see Figure 11)

$$V_{j-1} = V_j \oplus W_j$$

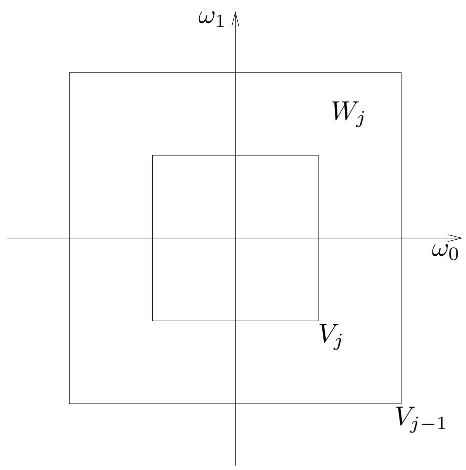



**Figure 11**. Multiscale subspaces generated by the Laplacian pyramid.

The LP can be considered as an oversampled filter bank where each polyphase component of the difference signal comes from a separate filter bank channel like the coarse signal [13]. Let $F_i(z), 0 \leq i \leq 3$ be the synthesis filters for these polyphase components. Note that $F_i(z)$ are highpass filters. As in the wavelet filter bank, we associate with each of these filters a continuous function $\psi^{(i)}(t)$ where

$$\psi^{(i)}(t) = 2 \sum_{n \in \mathbb{Z}^2} f_i[n] \phi(2t - n).$$

***Proposition 2.*** *([13]) Suppose that the LP with orthogonal filter generates an MRA. Then for a scale j,* $\{\psi_{j,n}^{(i)}\}_{0 \leq i \leq |\mathsf{M}|-1,\, n \in \mathbb{Z}^d}$ *is a tight frame of* $W_j$*. For all scales,* $\{\psi_{j,n}^{(i)}\}_{0 \leq i \leq |\mathsf{M}|-1,\, j \in \mathbb{Z},\, n \in \mathbb{Z}^2}$ *is a tight frame of* $L^2(\mathbb{R}^d)$*. In all cases, the frame bounds are equal to 1.*

Since $W_{j+1}$ is generated by four prototype functions, in general it is *not* a shift invariant subspace, unless $F_i(z)$ are shifted versions of a filter, or

$$F_i(z) = z^{-k_i} F(z) \tag{4.1.2}$$

where $k_i$ are the coset representatives of the downsampling lattice $(2, 2)$

$$k_0 = (0,0)^T,\ k_1 = (1,0)^T,\ k_2 = (0,1)^T,\ \text{and } k_3 = (1,1)^T. \tag{4.1.3}$$

Nevertheless, based on this, we can mimic $W_{j+1}$ to be a shift invariant subspace by denoting

$$\mu_{j,2n+k_i}(t) = \psi^{(i)}_{j+1,n} = \sum_{m\in\mathbb{Z}^2} f_i[m]\phi_{j,n+m}(t). \tag{4.1.4}$$

With this notation, the family $\{\mu_{j,n}\}_{n\in\mathbb{Z}^2}$ is a tight frame of $W_{j+1}$ and it resemble a uniform grid on $\mathbb{R}^2$ of intervals $2^j \times 2^j$.

## 4.2. Multiple Directions

Suppose that the DFB's in the PDFB use orthogonal filters. In the PDFB, the discrete basis (3.2.2) of the DFB can be regarded as a change of basis for the continuous subspaces from the multiscale decomposition. Although in the PDFB, the DFB is applied to the difference signal or the $W_{j+1}$ subspaces, we first show what happens when the DFB is applied to the multiresolution subspaces $V_j$.

***Proposition 3.*** *Define*

$$\theta^{(l)}_{j,k,n}(t) = \sum_{m\in\mathbb{Z}^2} g^{(l)}_k[m - \mathsf{S}^{(l)}_k n]\phi_{j,m}(t) \tag{4.2.1}$$

*The family $\{\theta^{(l)}_{j,k,n}\}_{n\in\mathbb{Z}^2}$ is an orthonormal basis of a directional subspace $V^{(l)}_{j,k}$ for each $k = 0, \ldots, 2^l - 1$. These subspaces are orthogonal with*

$$V^{(l)}_{j,k} = V^{(l+1)}_{j,2k} \oplus V^{(l+1)}_{j,2k+1}, \quad \text{and} \tag{4.2.2}$$

$$V_j = \bigoplus_{k=0}^{2^l-1} V^{(l)}_{j,k}. \tag{4.2.3}$$

**Proof:** (*Sketch*) This result is proved by induction on the number of decomposition levels $l$ of the DFB, in much the same way for the wavelet packets bases [8] (see also [19]). Assume that $\{\theta^{(l)}_{j,k,n}\}_{n\in\mathbb{Z}^2}$ is an orthonormal basis of a subspace $V^{(l)}_{j,k}$. To increase the directional resolution, an extra level of decomposition by a pair of orthogonal filters is applied to the channel represented by $g^{(l)}_k$ that leads to two channels with equivalent filters $g^{(l+1)}_{2k}$ and $g^{(l+1)}_{2k+1}$. This transforms the orthonormal basis $\{\theta^{(l)}_{j,k,n}\}_{n\in\mathbb{Z}^2}$ in two orthonormal families $\{\theta^{(l+1)}_{j,2k,n}\}_{n\in\mathbb{Z}^2}$ and $\{\theta^{(l+1)}_{j,2k+1,n}\}_{n\in\mathbb{Z}^2}$. Each of these families generates a subspace with finer directional resolution that satisfy the "two-direction" equation (4.2.3). With this, starting from the orthonormal basis $\{\phi_{j,n}\}_{n\in\mathbb{Z}^2}$ of $V_j$, all other orthonormal bases follow. $\square$

Figure 12 illustrates the "two-direction" subspace splitting by the DFB in the frequency domain. In the spatial domain, $V_{j,k}^{(l)}$ is a subspace at a scale $2^j$ and a direction $k$ among all $2^l$ directions.

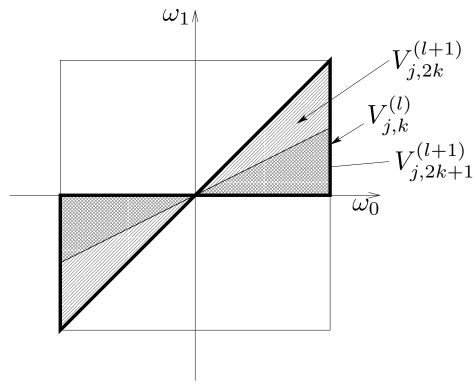



**Figure 12**. Multidirectional subspaces generated by the DFB.

Applying the directional decomposition by the family (3.2.2) onto the detail subspace $W_{j+1}$ as done by the PDFB, we obtain the similar result.

***Proposition 4.*** *Define*

$$\rho_{j,k,n}^{(l)}(t) = \sum_{m \in \mathbb{Z}^2} g_k^{(l)}[m - \mathsf{S}_k^{(l)} n] \mu_{j,m}(t) \tag{4.2.4}$$

*The family $\{\rho_{j,k,n}^{(l)}\}_{n \in \mathbb{Z}^2}$ is a tight frame of a subspace $W_{j+1,k}^{(l)}$ with frame bounds equal to 1, for each $k = 0, \dots, 2^l - 1$. These subspaces are orthogonal with*

$$W_{j+1,k}^{(l)} = W_{j,2k}^{(l+1)} \oplus W_{j,2k+1}^{(l+1)}, \quad \textit{and} \tag{4.2.5}$$

$$W_{j+1} = \bigoplus_{k=0}^{2^l-1} W_{j+1,k}^{(l)}. \tag{4.2.6}$$

**Proof:** This result is obtained by applying Proposition 2 to the subspaces in Proposition 3. □

Figure 13 shows a graphical representation of the subspaces in Proposition 4, seen in the frequency domain. The reason for $\{\rho_{j,k,n}^{(l)}\}_{n \in \mathbb{Z}^2}$ to be an overcomplete system for $W_{j+1,k}^{(l)}$ is because it uses the same sampling grid as the bigger subspace $V_{j,k}^{(l)}$.

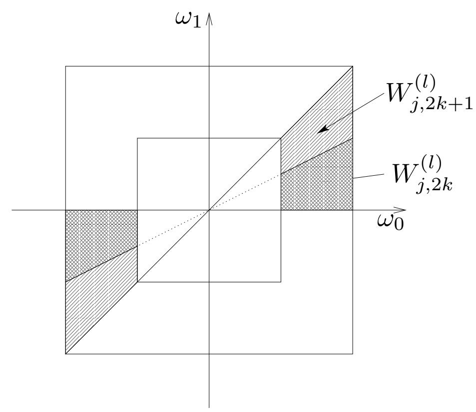



**Figure 13**. Multidirectional subspaces generated by the PDFB.

Recall that $W_{j+1}$ is *not* shift invariant but the following result establishes that its subspaces $W^{(l)}_{j+1,k}$ are since they are generated by a single prototype function.

***Proposition 5.*** *Let us denote*

$$\rho^{(l)}_{j,k}(t) = \sum_{m\in\mathbb{Z}^2} g^{(l)}_k[m]\mu_{j,m}(t) \tag{4.2.7}$$

*Then for* $l \geq 2$

$$\rho^{(l)}_{j,k,n}(t) = \rho^{(l)}_{j,k}(t - 2^j \mathsf{S}^{(l)}_k n) \tag{4.2.8}$$

**Proof:** By direct substitution and a change of variable. □

Consequently, the subspaces $W^{(l)}_{j+1,k}$ satisfy the following shift invariant property:

$$f(t) \in W^{(l)}_{j+1,k} \quad \Leftrightarrow \quad f(t - 2^j S^{(l)}_k n) \in W^{(l)}_{j+1,k}, \quad \forall n \in \mathbb{Z}^2. \tag{4.2.9}$$

This says that the directional multiscale subspaces $W^{(l)}_{j+1,k}$ are defined on a rectangular grid with intervals $2^{j+l-1} \times 2^{j+1}$ (or $2^{j+1} \times 2^{j+l-1}$, depending on whether it is basically horizontal or vertical). By substituting (4.1.4) into (4.2.7), we can write the prototype function $\rho^{(l)}_{j,k}(t)$ directly as a linear combination of the scaling function $\phi_{j,m}(t)$ as

$$\begin{aligned}\rho_{j,k}^{(l)}(t) &= \sum_{i=0}^{3}\sum_{n} g_k^{(l)}[2n+k_i]\left(\sum_{m\in\mathbb{Z}^2} f_i[m]\phi_{j,n+m}\right)\\ &= \sum_{m\in\mathbb{Z}^2}\underbrace{\left(\sum_{i=0}^{3}\sum_{n\in\mathbb{Z}^2} g_k^{(l)}[2n+k_i] f_i[m-n]\right)}_{c_k^{(l)}[m]}\phi_{j,m}(t). \quad (4.2.10)\end{aligned}$$

The sequence $c_k^{(l)}[m]$ resemble a summation of convolutions between $g_k^{(l)}[m]$ and $f_i[m]$, thus it is a highpass and directional filter. Equation (4.2.10) reveals the "contourlet-like" behavior of the prototype function $\rho_{j,k}^{(l)}(t)$ where it is seen as a grouping of "edge-detection" elements at a scale $j$ and along a direction $k$.

### 4.3. Multiscale and multidirection

Finally, integrating over scales we have the following result for the *contourlet frames* on the space $L^2(\mathbb{R}^2)$.

***Theorem 1.*** *For a sequence of finite positive integers $\{l_j\}_{j\le j_0}$ the family*

$$\{\phi_{j_0,n}(t),\ \rho_{j,k,n}^{(l_j)}(t)\}_{j\le j_0,\ 0\le k\le 2^{l_j}-1,\ n\in\mathbb{Z}^2} \quad (4.3.1)$$

*is a tight frame of $L^2(\mathbb{R}^2)$. For a sequence of finite positive integers $\{l_j\}_{j\in\mathbb{Z}}$, the family*

$$\{\rho_{j,k,n}^{(l_j)}(t)\}_{j\in\mathbb{Z},\ 0\le k\le 2^{l_j}-1,\ n\in\mathbb{Z}^2} \quad (4.3.2)$$

*is a directional wavelet tight frame of $L^2(\mathbb{R}^2)$. In each case, the frame bounds are equal to 1.*

**Proof:** This result is obtained by applying Proposition 4 to the following decompositions of $L^2(\mathbb{R}^2)$ into mutual orthogonal subspaces:

$$\begin{aligned}L^2(\mathbb{R}^2) &= V_{j_0}\oplus\left(\bigoplus_{j\le j_0} W_j\right), \quad \text{and}\\ L^2(\mathbb{R}^2) &= \bigoplus_{j\in\mathbb{Z}} W_j.\end{aligned}$$

□

As discussed in Section 3.4., the tight frame in (4.3.1) provides a curvelet-like expansion when the number of directions is doubled at every other finer scale. This means that if at the scale $2^{j_0}$ we start with an $l_{j_0}$-level DFB (which has $2^{l_{j_0}}$ directions) then at finer scales $2^j$, $j < j_0$, the number of decomposition levels by the DFB should be:

$$l_j = \lfloor l_{j_0} - (j - j_0)/2 \rfloor, \quad \text{for } j \le j_0. \tag{4.3.3}$$

Thus the embedded grid of approximation for the curvelet PDFB expansion at the scale $2^j$ is $2^{\lfloor n_0 + j/2 \rfloor} \times 2^j$ for basically horizontal directions and $2^j \times 2^{\lfloor n_0 + j/2 \rfloor}$ for near vertical directions, where $n_0 = l_{j_0} - j_0/2 + 2$. Figure 14 illustrates this sampling pattern at different scales and directions. The main point to note here is that in the refinement process, one spatial dimension is refined at twice the speed as the other spatial dimension.

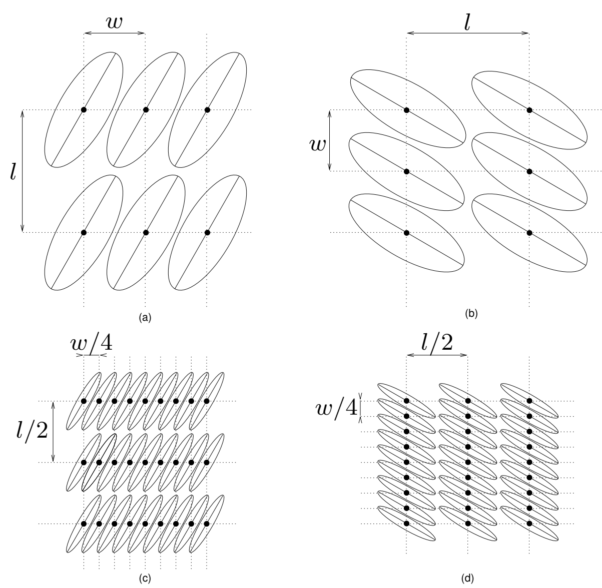



**Figure 14**. Embedded grids of approximation in spatial domain for a PDFB that implements the curvelet transform. These are four illustrative subspaces $W_{j,k}$ representing coarser vs. finer scales and basically horizontal vs. basically vertical directions. Each subspace is spanned by the shifts of a curvelet prototype function given in (4.2.7). The sampling intervals match with the supports of the prototype function, for example width $w$ and length $l$, so that the shifts would tile the $\mathbb{R}^2$ plane.

Figure 10 and Figure 14 give a complete view of the multiresolution approximation of our curvelet construction based on the PDFB. They clearly show a refinement scheme where the resolution increases in both spatial and direction domain when going from coarse to fine scale, and the basis elements exhibit the anisotropy scaling relation for curves.

## 5. Numerical Experiments

Figure 15 shows an example image that is transformed by the PDFB implementing the discrete ridgelet transform. As we can see, the coefficients in the transform domain are very sparse – significant coefficients are located around edges *and* in the right directional subbands. With non-linear approximation using the PDFB, smooth regions are represented efficiently by the small size lowpass image while smooth edges are efficiently represented by a few directional local coefficients.

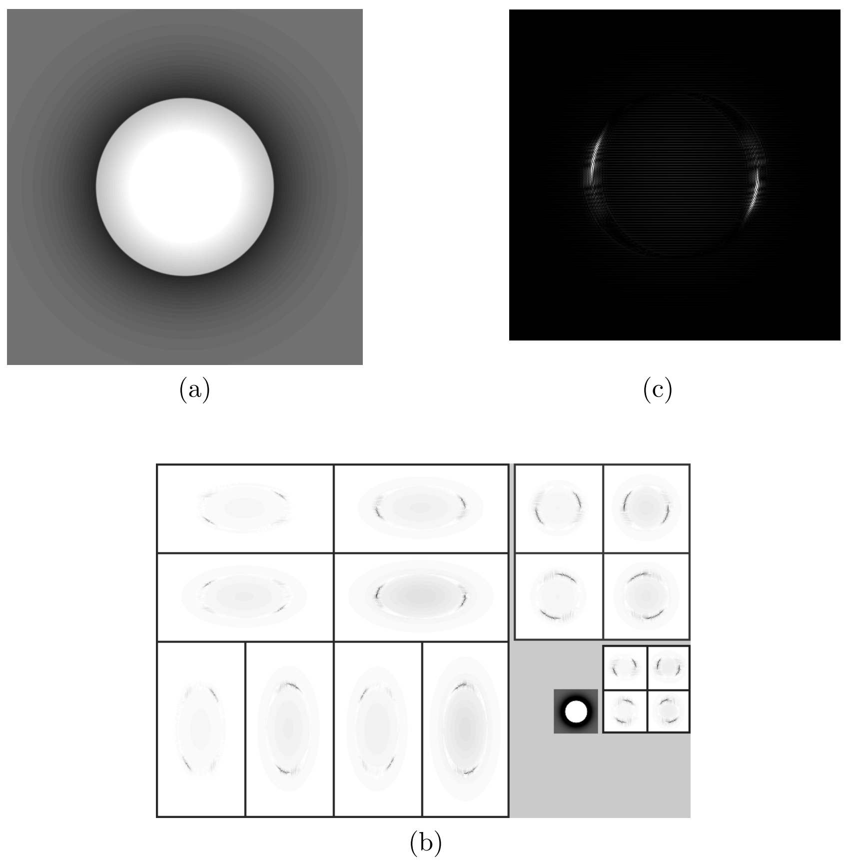



**Figure 15**. Example of PDFB. (a) Input image. (b) Magnitudes of PDFB coefficients. (c) Reconstruction from one PDFB subband. The LP uses the biorthogonal "9-7" filters, while the DFB's use the biorthogonal "23-45" quincunx filters designed by Phoong et al. [23] with support sizes equal to $23 \times 23$ and $45 \times 45$.

For comparison, Starck et al. [25] describe a different approach for the digital curvelet transform, in which they directly "discretize" the continuous definition.

Their implementation uses the discrete Radon transform on image blocks, and thus the number of represented directions, which equals the block size, is reduced by half for every finer scale. This is unlike the curvelet construction in continuous space, or our construction. Furthermore, there is a redundancy factor equals to $16J + 1$ in their implementation compared with 1.33 in ours.

We now evaluate the non-linear approximation performance of the PDFB that implements the discrete curvelet transform and compare it with the performance by the 2-D discrete wavelet transform (DWT2). In these NLA experiments, for a given value $M$, we select the $M$-most significant coefficients in each transform domain, and then compare the reconstructed images from these sets of $M$ coefficients. The wavelet transform used in the experiments is a biorthogonal transform with the "9-7" filters [7, 27] and 6 decomposition levels. The PDFB also uses the "9-7" filters in the LP decomposition. While the DFB in the PDFB uses the "23-45" biorthogonal quincunx filters designed by Phoong et al. [23]. The number of decomposition levels by the DFB at the finest pyramidal scale is 5, which leads to 32 directions.

Note that in this case, both the DWT2 and the PDFB transforms share the same multiscale detailed subspaces $W_j$ as defined in Section 4.1., which are generated by the "9-7" lowpass filters. The difference is that in the DWT2, each subspace $W_j$ is represented by a basis with three directions, whereas in the PDFB it is represented by a redundant frame with many more directions. Since the two transforms share the same detailed subspaces, it is possible to restrict the comparison in these subspaces. We expect that most of the refinement actions would happen around the image edges. Figure 16 and Figure 17 show sequences of non-linear approximated images at the finest subspace $W_j$ using the DWT2 and the PDFB, respectively. We observe that the wavelet scheme slowly refines the detailed image by isolated "dots" along the contours, while the contourlet scheme quickly refines by well-adapted "sketches". The improvement by the PDFB can be seen both in terms of visual quality and reconstruction error.

Finally, Figure 18 shows a detailed comparison of two non-linear approximated images by the DWT2 and the PDFB. We clearly see that fine contours (directional textures on cloths) are better represented by the contourlet transform compared to the wavelet transform. For more experimental results, we refer to [10].

## 6. Conclusion

In this work, we constructed a discrete transform that can offer a sparse representation for piecewise smooth images, as promised by the curvelet theory. We first identified two key features of curvelets that could lead to an improvement over the wavelet transform, namely directionality and anisotropy. From this we proposed a new filter bank structure, the pyramidal directional filter bank (PDFB), that can provide a multiscale and directional decomposition for

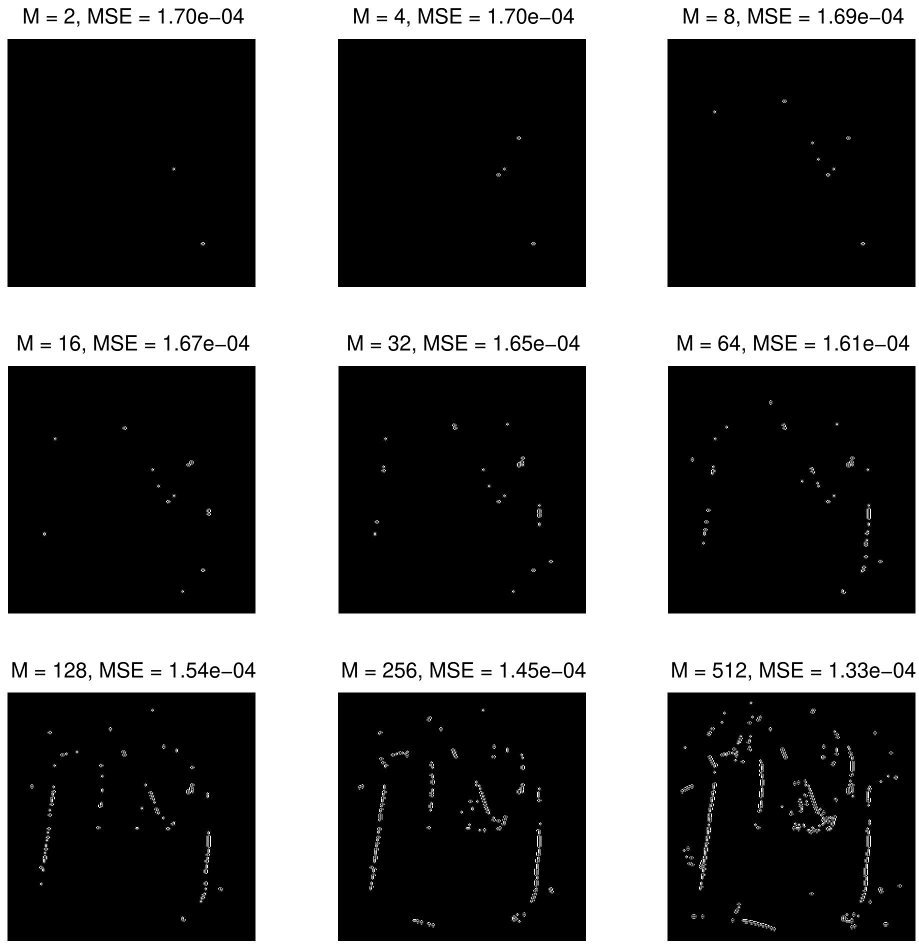



**Figure 16**. Sequence of images showing the non-linear approximation at the finest scale of the DWT2. M is the number of the most significant coefficients; MSE is the mean square error against the projection of the input image into the finest detailed subspace. The input is the "Peppers" image.

images with a small redundancy factor. The PDFB provides a frame expansion for images with frame elements like contour segments, and thus is also called the contourlet transform. The connection between the developed discrete and continuous-domain constructions was made precise via a new directional multiresolution analysis, which provides successive refinements at *both* spatial and directional resolution. The contourlet transform can be designed to satisfy the anisotropy scaling relation for curves and thus it provides a curvelet-like decomposition for images. Experiments with real images indicate the potential of

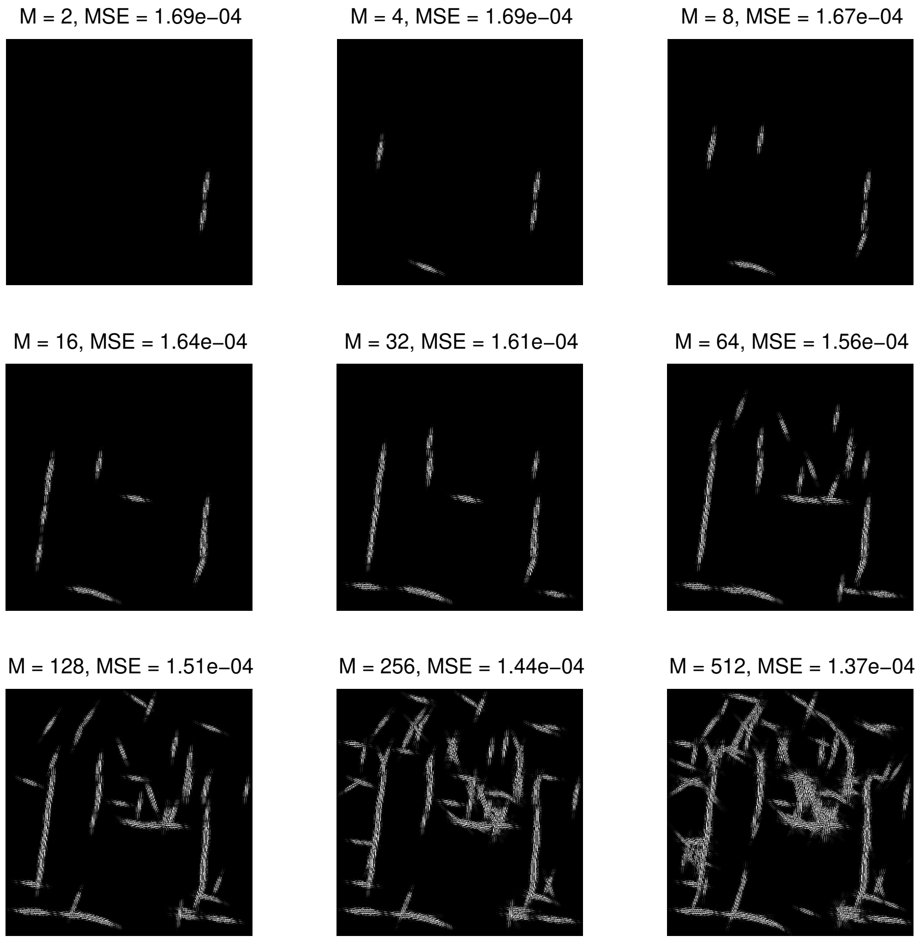



**Figure 17**. Same as in Figure 16 but with the PDFB. Note that the PDFB shares the same detailed subspace with the DWT2.

contourlets in image processing applications.

**Acknowledgement.** This work was supported in part by a Ph.D. Fellowship from the Department of Communication Systems, Swiss Federal Institute of Technology Lausanne.

## References

[1] 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 Carte-

(a) Original image

(b) DWT2: PSNR = 24.34 dB

(c) PDFB: PSNR = 25.70 dB

**Figure 18**. Detail of non-linear approximated images by the DWT2 and the PDFB. In each case, the image originally of size $512 \times 512$ is reconstructed from the 4096-most significant coefficients in the transform domain.

sian grid which is rapidly computable, algebraically exact, geometrically faithful and invertible. Technical report, 2001. Preprint, `http://www.math.tau.ac.il/~amir1/`.

[2] R. H. Bamberger and M. J. T. Smith. A filter bank for the directional decomposition of images: Theory and design. *IEEE Trans. Signal Proc.*, 40(4):882–893, April 1992.

[3] P. J. Burt and E. H. Adelson. The Laplacian pyramid as a compact image

code. *IEEE Trans. Commun.*, 31(4):532–540, April 1983.

[4] E. J. Candès and D. L. Donoho. Curvelets – a suprisingly effective non-adaptive representation for objects with edges. In A. Cohen, C. Rabut, and L. L. Schumaker, editors, *Curve and Surface Fitting*, Saint-Malo, 1999. Vanderbilt University Press.

[5] E. J. Candès and D. L. Donoho. Ridgelets: a key to higher-dimensional intermittency? *Phil. Trans. R. Soc. Lond. A.*, pages 2495–2509, 1999.

[6] E. J. Candès and D. L. Donoho. Curvelets, multiresolution representation, and scaling laws. In A. Aldroubi, A. F. Laine, and M. A. Unser, editors, *SPIE Wavelet Applications in Signal and Image Processing VIII*, volume 4119, 2000.

[7] A. Cohen, I. Daubechies, and J.-C. Feauveau. Biorthogonal bases of compactly supported wavelets. *Commun. on Pure and Appl. Math.*, 45:485–560, 1992.

[8] R. R. Coifman, Y. Meyer, and M. V. Wickerhauser. Wavelet analysis and signal processing. In M. B. Ruskai et al, editor, *Wavelets and their Applications*, pages 153–178. Jones and Barlett, Boston, 1992.

[9] R. A. DeVore, B. Jawerth, and B. J. Lucier. Image compression through wavelet transform coding. *IEEE Transactions on Information Theory, Special Issue on Wavelet Transforms and Multiresolution Signal Analysis*, 38(2):719–746, March 1992.

[10] M. N. Do. *Directional Multiresolution Image Representations*. PhD thesis, Swiss Federal Institute of Technology, Lausanne, Switzerland, December 2001.

[11] M. N. Do and M. Vetterli. Pyramidal directional filter banks and curvelets. In *Proc. IEEE Int. Conf. on Image Proc.*, Thessaloniki, Greece, Oct. 2001.

[12] M. N. Do and M. Vetterli. Contourlets. In J. Stoeckler and G. V. Welland, editors, *Beyond Wavelets*. Academic Press, New York, 2002. to appear, `http://www.ifp.uiuc.edu/~minhdo/publications`.

[13] M. N. Do and M. Vetterli. Framing pyramids. *IEEE Trans. Signal Proc.*, 2002. to appear, `http://www.ifp.uiuc.edu/~minhdo/publications`.

[14] D. L. Donoho. Sparse component analysis and optimal atomic decomposition. *Constructive Approximation*, 17:353–382, 1998.

[15] D. L. Donoho and M. R. Duncan. Digital curvelet transform: strategy, implementation and experiments. In *Proc. Aerosense 2000, Wavelet Applications VII*, volume 4056, pages 12–29. SPIE, 2000.

[16] D. L. Donoho, M. Vetterli, R. A. DeVore, and I. Daubechies. Data compression and harmonic analysis. *IEEE Trans. Inform. Th.*, 44(6):2435–2476, October 1998.

[17] P. V. C. Hough. Methods and means for recogninzing complex patterns. U.S. Patent 3069654, 1962.

[18] S. Mallat. Multiresolution approximations and wavelet orthonormal bases of $L_2(R)$. *Trans. Amer. Math. Soc.*, 315:69–87, September 1989.

[19] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 2nd edition, 1999.

[20] Y. Meyer. *Wavelets: Algorithms and Applications*. SIAM, 1992. Translated and revised by R. D. Ryan.

[21] S. Park. *New Directional Filter Banks and Their Applications in Image Processing*. PhD thesis, Georgia Institute of Technology, 1999.

[22] S. Park, M. J. T. Smith, and R. M. Mersereau. A new directional filterbank for image analysis and classification. In *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc.*, pages 1417–1420, 1999.

[23] S.-M. Phoong, C. W. Kim, P. P. Vaidyanathan, and R. Ansari. A new class of two-channel biorthogonal filter banks and wavelet bases. *IEEE Trans. Signal Proc.*, 43(3):649–665, Mar. 1995.

[24] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger. Shiftable multiscale transforms. *IEEE Transactions on Information Theory, Special Issue on Wavelet Transforms and Multiresolution Signal Analysis*, 38(2):587–607, March 1992.

[25] J. L. Starck, E. J. Candès, and D.L. Donoho. The curvelet transform for image denoising. *IEEE Trans. Image Proc.*, 11:670–684, Jun. 2002.

[26] G. Strang and T. Nguyen. *Wavelets and Filter Banks*. Wellesley Cambridge Press, Boston, 1996.

[27] M. Vetterli and C. Herley. Wavelets and filter banks: Theory and design. *IEEE Trans. Signal Proc.*, 40(9):2207–2232, September 1992.

[28] A. B. Watson. The cortex transform: Rapid computation of simulated neural images. *Computer Vision, Graphics, and Image Processing*, 39(3):311–327, 1987.


*Minh N. Do*
*Department of Electrical and Computer Engineering*

*Beckman Institute*
*University of Illinois at Urbana-Champaign*
*Urbana, IL 61801, USA*
*minhdo@uiuc.edu*

*Martin Vetterli*
*Department of Communication Systems*
*Swiss Federal Institute of Technology*
*1015 Lausanne, Switzerland*
and
*Department of Electrical Engineering and Computer Science*
*University of California at Berkeley*
*Berkeley, CA 94720, USA*
*Martin.Vetterli@epfl.ch*