

The Contourlet Transform: An Efficient Directional Multiresolution Image Representation

Minh N. Do, *Member, IEEE*, and Martin Vetterli, *Fellow, IEEE*

Abstract

The limitations of commonly used separable extensions from one-dimensional transforms for images, such as the Fourier and wavelet transforms, are well known; e.g. these separable transforms offer limited directional information in representing image edges. In this paper, we pursue a “true” two-dimensional representation that can capture the intrinsic geometrical structure that is key in visual information. Unlike other transforms, such as curvelets, that were initially developed in the continuous-domain and then discretized for sampled data, our approach starts with a discrete-domain transform and then investigates its convergence to an expansion in the continuous-domain. We construct a discrete-domain multiresolution and multidirectional expansion using non-separable filter banks, in much the same way that wavelets were derived from filter banks. This construction results in a flexible multiresolution, local, and directional image expansion using contour segments, and thus it is named the *contourlet* transform. The discrete contourlet transform has a fast iterated filter bank algorithm that requires an order N operations for N -pixel images. Furthermore, we establish a precise link between the developed filter bank and the associated continuous-domain contourlet expansion in a directional multiresolution analysis framework. We prove that with parabolic scaling and sufficient directional vanishing moments, contourlets achieve the optimal approximation rate for piecewise smooth functions with discontinuities along twice continuously differentiable curves. Finally, we show some numerical experiments demonstrating the potential of contourlets in several image processing applications.

Index Terms

sparse image expansions, wavelets, contourlets, filter banks, multiresolution, multidirection, contours, geometry.

M. N. Do is with the Department of Electrical and Computer Engineering, the Coordinated Science Laboratory, and the Beckman Institute, University of Illinois, Urbana IL 61801 (email: minhdo@uiuc.edu).

M. Vetterli is with the Audiovisual Communications Laboratory, Swiss Federal Institute of Technology, Lausanne, Switzerland, and with the Department of Electrical Engineering and Computer Science, University of California, Berkeley CA 94720 (email: martin.vetterli@epfl.ch).

This work was supported in part by the National Science Foundation under Grant CCR-0237633 (CAREER).

I. INTRODUCTION

Efficient representation of visual information lies at the heart of many image processing tasks, including compression, filtering, feature extraction, and inverse problems. *Efficiency* of a representation refers to the ability to capture significant information about an object of interest using a sparse description. For image compression or content-based image retrieval, the use of efficient representation implies the compactness of the compressed file or the index entry for each image in the database. For practical applications, such an efficient representation has to be obtained by structured transforms and fast algorithms.

For *one-dimensional piecewise smooth* signals, like scan-lines of an image, *wavelets* have been established as the right tool, because they provide an optimal representation for these signals in a certain sense [1], [2]. In addition, the wavelet representation is amenable to efficient algorithms; in particular it has fast transforms and convenient tree data structures. These are the key reasons for the success of wavelets in many signal processing and communication applications; for example, the wavelet transform was adopted as the transform for the new image-compression standard, JPEG-2000 [3].

However, natural images are not simply stacks of 1-D piecewise smooth scan-lines; discontinuity points (i.e. edges) are typically positioned along smooth curves (i.e. contours) owing to smooth boundaries of physical objects. As a result of a separable extension from 1-D bases, wavelets in 2-D are good at isolating the discontinuities at *edge points*, but will not “see” the smoothness along the *contours*. In addition, separable wavelets can capture only limited *directional information* – an important and unique feature of multidimensional signals. These disappointing behaviors indicate that more powerful representations are needed in higher dimensions.

To see how one can improve the 2-D separable wavelet transform for representing images with smooth contours, consider the following scenario. Imagine that there are two painters, one with a “wavelet”-style and the other with a “X-let”-style, both wishing to paint a natural scene. Both painters apply a refinement technique to increase resolution from coarse to fine. Here, efficiency is measured by how quickly, that is with how few brush strokes, one can faithfully reproduce the scene.

Consider the situation when a smooth contour is being painted, as shown in Figure 1. Because 2-D wavelets are constructed from tensor products of 1-D wavelets, the “wavelet”-style painter is limited to using square-shaped brush strokes along the contour, using different sizes corresponding to the multiresolution property of wavelets. As the resolution becomes finer, we can clearly see the limitation of the wavelet-style painter who needs to use many fine “dots” to capture the contour.¹ The “X-let”-style painter, on the other hand,

¹Or we could consider the wavelet-style painter as a *pointillist*!

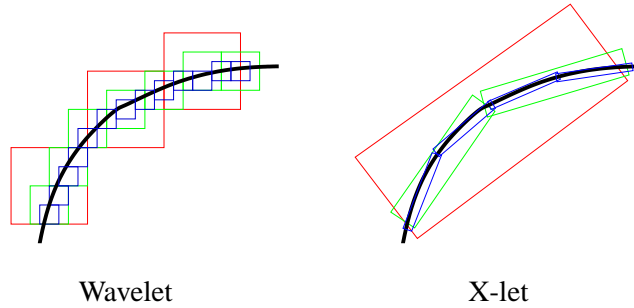


Fig. 1. Wavelet versus new scheme: illustrating the successive refinement by the two systems near a smooth contour, which is shown as thick a curve separating two smooth regions.

explores effectively the smoothness of the contour by making brush strokes with different elongated shapes and in a variety of directions following the contour. This intuition was formalized by Candès and Donoho in the *curvelet* construction [4], [5], reviewed below in Section II.

For the human visual system, it is well-known [6] that the receptive fields in the visual cortex are characterized as being *localized*, *oriented*, and *bandpass*. Recently, experiments in searching for the sparse components of natural images by Olshausen and Field [7] produced basis images that closely resemble the aforementioned characteristics of the visual cortex. This result supports the hypothesis that the human visual system has been tuned so as to capture the essential information of a natural scene using a least number of visual active cells. More importantly, this result suggests that for a computational image representation to be efficient, it should be based on a *local*, *directional*, and *multiresolution* expansion.

Inspired by the painting scenario and studies related to the human visual system and natural image statistics, we identify a “wish list” for new image representations:

- 1) **Multiresolution.** The representation should allow images to be successively approximated, starting from a coarse version and going to fine-resolution version.
- 2) **Localization.** The basis elements in the representation should be localized in both the spatial and the frequency domains.
- 3) **Critical sampling.** For some applications (e.g. compression), the representation should form a basis, or a frame with small redundancy.
- 4) **Directionality.** The representation should contain basis elements oriented at variety of directions, much more than a few directions that are offered by separable wavelets.
- 5) **Anisotropy.** To capture smooth contours in images, the representation should contain basis elements using a variety of elongated shapes with different aspect ratios.

Among these desiderata, the first three are successfully provided by separable wavelets, while the last

two, especially directionality, require new constructions.

Our aim in this paper is to construct a *discrete*-domain expansion satisfying the above desiderata and that can be applied to sampled images. After reviewing related work in Section II, we propose in Section III a multiresolution and multidirectional image expansion using non-separable filter banks. This construction results in a flexible multiresolution, local, and directional image expansion using contour segments, and thus it is named the *contourlet* transform. A question of interest is the limit behavior when such schemes are iterated over scale and/or direction, which is similar to filter banks, their iteration, and the associated wavelet construction [8], [9], [2]. This question is explored in Section IV, where we establish a precise link between the proposed filter bank and the associated continuous-domain expansion in a newly defined *directional multiresolution analysis* framework. The approximation power of the contourlet expansion is studied in Section V. We prove that with parabolic scaling and sufficient directional vanishing moments, contourlets achieve the essential optimal approximation rate, $O((\log M)^3 M^{-2})$ square error with a best M -term approximation, for 2-D piecewise smooth functions with C^2 (twice continuously differentiable) contours. Numerical experiments are presented and discussed in Section VI.

II. BACKGROUND AND RELATED WORK

Recently, Candès and Donoho [4], [5] pioneered a new expansion in the *continuous* two-dimensional space \mathbb{R}^2 using *curvelets*. This expansion achieves essentially optimal approximation behavior in a certain sense (see below) for 2-D piecewise smooth functions, which are C^2 except discontinuities along C^2 curves. The attractive property of curvelets is that such correct approximation behavior is simply obtained via thresholding (or best M -term approximation) of a *fixed* transform.

The error decay of *best M -term approximation* provides a measurement of the efficiency of an expansion. Consider a series expansion by $\{\phi_n\}_{n=1}^{\infty}$ (e.g. in a Fourier or wavelets basis) for a given function x as: $x = \sum_{n=1}^{\infty} c_n \phi_n$. Then its best M -term approximation (also commonly referred to as *nonlinear approximation* [10]) using this expansion is defined as: $\hat{x}_M = \sum_{n \in I_M} c_n \phi_n$, where I_M is the set of indexes of the M -largest $|c_n|$. It is easy to see that the quality of the approximated function \hat{x}_M relates to how *sparse* the expansion by $\{\phi_n\}_{n=1}^{\infty}$ is, or how well the expansion compacts the energy of x into a few coefficients.

For the class of 2-D piecewise smooth functions mentioned above, the best M -term approximation error $\|\hat{x}_M - x\|^2$ using curvelets has a decay rate of $O((\log M)^3 M^{-2})$ [5], while for wavelets this rate is $O(M^{-1})$ and for the Fourier basis it is $O(M^{-1/2})$ [1], [2]. Therefore, for typical images with smooth contours, we expect significant improvement of a curvelet-like method over wavelets, which is comparable to the improvement of wavelets over the Fourier basis. Perhaps equally important, the curvelet construction

demonstrates that it is possible to develop an optimal representation for images with smooth contours via a *fixed* transform. This fact has inspired our development of the contourlet transform in this paper.

The curvelet transform was developed initially in the *continuous-domain* [4] via multiscale filtering and then applying a block ridgelet transform [11] on each bandpass image. Later, the authors proposed the *second generation curvelet transform* [5] that was defined directly via frequency partitioning without using the ridgelet transform. Both curvelet constructions require a rotation operation, which makes the construction simple in the continuous-domain but causes the implementation of the curvelet transform for discrete images – sampled on a rectangular grid – very challenging. In particular, approaching critical sampling seems very difficult in such discretized constructions.

Apart from curvelets and contourlets, there have recently been several approaches in developing efficient representations of geometrical regularity. Notable examples are the bandelet representation [12], the edge-adapted multiscale transform [13], wedgelets [14], [15], edgeprints [16], and quadtree coding [17]. These approaches typically require an edge-detection step, followed by an *adaptive* representation. By contrast, curvelet and contourlet representations are *fixed* transforms. This feature allows them to be easily applied in a wide range of image processing tasks, similar to wavelets. Furthermore, we can benefit from the well-established knowledge in transform coding when applying contourlets to image compression (e.g. for bit allocation).

Several other well-known systems that provide multiscale and directional image representations include: 2-D Gabor wavelets [18], the cortex transform [19], the steerable pyramid [20], 2-D directional wavelets [21], brushlets [22], and complex wavelets [23]. The main differences between these systems and our contourlet construction is that the previous methods do not allow for a different number of directions at each scale while achieving nearly critical sampling. In addition, our construction employs iterated filter banks, which makes it computationally efficient, and there is a precise connection with continuous-domain expansions.

III. DISCRETE-DOMAIN CONSTRUCTION USING FILTER BANKS

A. Concept

Comparing the wavelet scheme with the “X-let” scheme in Figure 1, we see that the improvement of X-lets can be attributed to the grouping of nearby wavelet coefficients, since they 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 [24] for line detection in computer vision.

With this insight, we proposed a *double filter bank* approach for obtaining sparse expansions for typical images having smooth contours. We called this a *pyramidal directional filter bank* (PDFB) [25], where the Laplacian pyramid [26] is first used to capture the point discontinuities, then followed by a directional filter bank [27] to link point discontinuities into linear structures. The overall result is an image expansion using basic elements like contour segments, and thus are named *contourlets*.² In particular, contourlets have elongated supports at various scales, directions, and aspect ratios. This allows contourlets to efficiently approximate a smooth contour at multiple resolutions in much the same way as the “X-let” scheme shown in Figure 1. In the frequency domain, the contourlet transform provides a multiscale and directional decomposition.

We would like to point out that the decoupling of multiscale and directional decomposition steps offers a simple and flexible transform, but at the cost of a small redundancy (up to 33%, which comes from the Laplacian pyramid). In a more recent work [28], we developed a *critically sampled* contourlet transform, which we call *CRISP-contourlets*, using a combined iterated nonseparable filter bank for both multiscale and directional decomposition. This leads to an orthonormal transform in both discrete and continuous domains.

B. Pyramid tight frames

One way of achieving a multiscale decomposition is to use the Laplacian pyramid (LP) introduced by Burt and Adelson [26]. 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 2(a)). The process is then iterated on the coarse version.

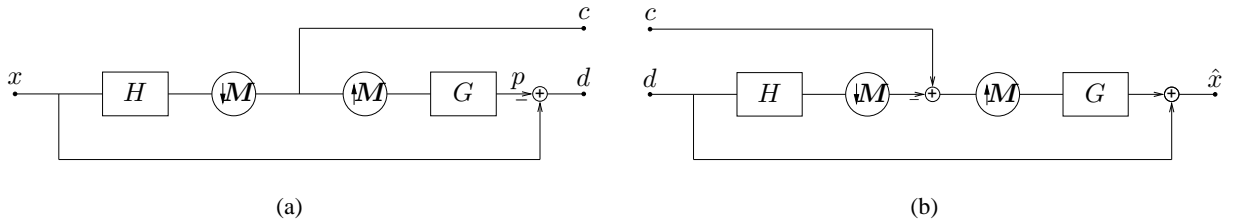


Fig. 2. (a) The Laplacian pyramid decomposition. The outputs are a coarse approximation c and a difference d between the original signal and the prediction. The process can be iterated by decomposing the coarse version repeatedly. (b) The new reconstruction scheme for the Laplacian pyramid [29]. It is the pseudo-inverse of the decomposition in (a) when the filters are orthogonal.

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

²In the sequel, we will use the terms (discrete) contourlet transform and pyramidal directional filter bank interchangeably.

(even for multidimensional cases) that 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 [29], we studied the LP using the theory of frames and oversampled filter banks. We showed that the LP with orthogonal filters (that is, $h[n] = g[-n]$ and $g[n]$ is orthogonal to its translates with respect to the subsampling lattice) provides a *tight frame* with frame bounds are equal to 1. In this case, we proposed the use of the optimal linear reconstruction using the dual frame operator (or pseudo-inverse) as shown in Figure 2(b). The new reconstruction differs from the usual method, where the signal is obtained by simply adding back the difference to the prediction from the coarse signal, and was shown [29] to achieve significant improvement over the usual reconstruction in the presence of noise.

C. Iterated directional filter banks

In 1992, Bamberger and Smith [27] constructed a 2-D directional filter banks (DFB) that can be maximally decimated while achieving perfect reconstruction. The DFB is efficiently implemented via an l -level binary tree decomposition that leads to 2^l subbands with wedge-shaped frequency partitioning as shown in Figure 3(a). The original construction of the DFB in [27] involves modulating the input image and using quincunx filter banks with diamond-shaped filters [30]. To obtain the desired frequency partition, a complicated tree expanding rule has to be followed for finer directional subbands (e.g. see [31] for details).

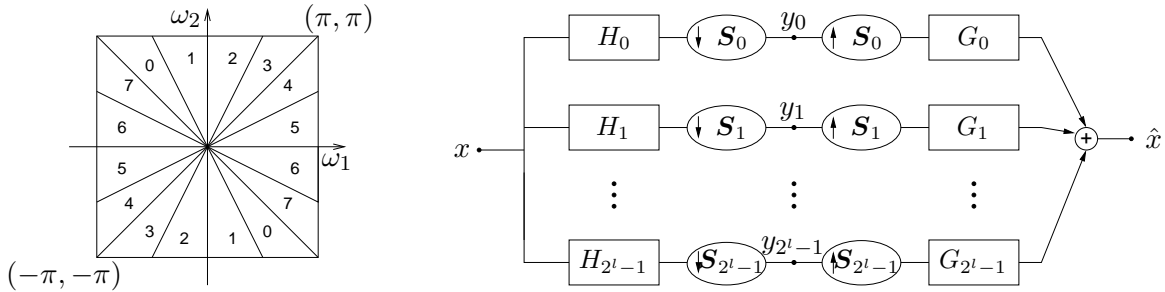


Fig. 3. (a) Directional filter bank frequency partitioning where $l = 3$ and there are $2^3 = 8$ real wedge-shaped frequency bands. Subbands 0–3 correspond to the *mostly horizontal* directions, while subbands 4–7 correspond to the *mostly vertical* directions. (b) The multichannel view of an l -level tree-structured directional filter bank.

In [32], we proposed a new construction for the DFB that avoids modulating the input image and has a simpler rule for expanding the decomposition tree. Our simplified DFB is intuitively constructed from two building blocks. The first building block are two-channel quincunx filter banks [30] with fan filters (see

Figure 4) that split a 2-D spectrum into two directions: horizontal and vertical. The second building block of the DFB are *shearing* operators, which amount to just reordering of image samples. Figure 5 shows an application of a shearing operator where a -45° direction edge becomes a vertical edge. By adding a pair of shearing operator and its inverse (“unshearing”) to before and after, respectively, a two-channel filter bank in Figure 4, we obtain a different directional frequency splitting while maintaining perfect reconstruction. Thus, the key in the DFB is to use an appropriate combination of shearing operators together with two-direction splitting of quincunx filter banks at each node in a binary tree-structured filter bank, to obtain the desired 2-D spectrum division as shown in Figure 3(a). For details, see [32] (Chapter 3).

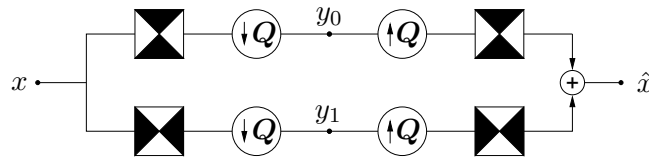


Fig. 4. Two-dimensional spectrum splitting using quincunx filter banks with fan filters. The black regions represent the ideal frequency supports of each filter. Q is a quincunx sampling matrix.

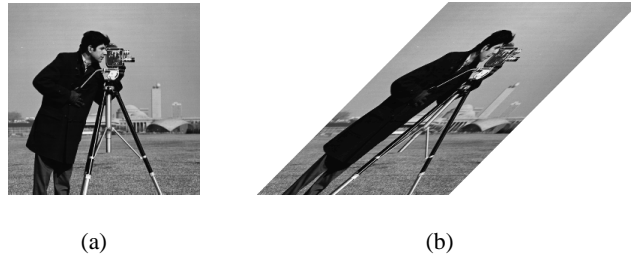


Fig. 5. Example of shearing operation that is used like rotation operation for DFB decomposition. (a) The “cameraman” image. (b) The “cameraman” image after a shearing operation.

Using multirate identities, it is instructive to view an l -level tree-structured DFB equivalently as a 2^l parallel channel filter bank with equivalent filters and overall sampling matrices as shown in Figure 3(b). Denote these equivalent synthesis filters as $G_k^{(l)}$, $0 \leq k < 2^l$, which correspond to the subbands indexed as in Figure 3(a). The corresponding overall sampling matrices were shown [32] to have the following diagonal forms

$$S_k^{(l)} = \begin{cases} \text{diag}(2^{l-1}, 2) & \text{for } 0 \leq k < 2^{l-1} \\ \text{diag}(2, 2^{l-1}) & \text{for } 2^{l-1} \leq k < 2^l, \end{cases} \quad (1)$$

which means sampling is separable. The two sets correspond to the *mostly horizontal* and *mostly vertical* set of directions, respectively.

From the equivalent parallel view of the DFB, we see that the family

$$\left\{ g_k^{(l)}[\mathbf{n} - \mathbf{S}_k^{(l)}\mathbf{m}] \right\}_{0 \leq k < 2^l, \mathbf{m} \in \mathbb{Z}^2}, \quad (2)$$

obtained by translating the impulse responses of the equivalent synthesis filters $G_k^{(l)}$ over the sampling lattices by $\mathbf{S}_k^{(l)}$, provides a *basis* for discrete signals in $l^2(\mathbb{Z}^2)$. This basis exhibits both directional and localization properties. Figure 6 demonstrates this fact by showing the impulse responses of equivalent filters from an example DFB. These basis functions have quasi-linear supports in space and span all directions. In other words, the basis (2) resembles a local Radon transform and are called *Radonlets*. Furthermore, it can be shown [32] that if the building block filter bank in Figure 4 uses orthogonal filters, then the resulting DFB is orthogonal and (2) becomes an orthogonal basis.

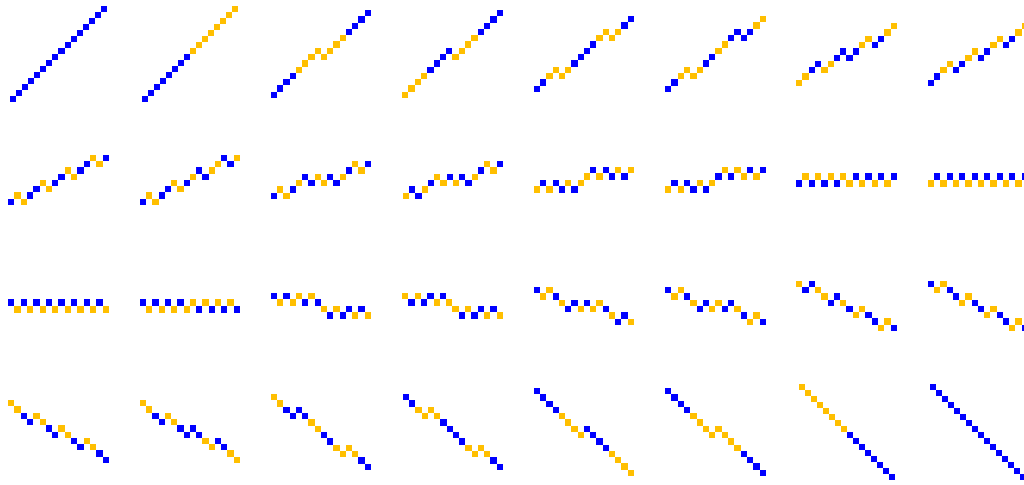


Fig. 6. Impulse responses of 32 equivalent filters for the first half channels of a 6-levels DFB that uses the Haar filters. Black and gray squares correspond to +1 and -1, respectively. Because the basis functions resemble ‘local lines’, we call them *Radonlets*.

D. Multiscale and directional decomposition: pyramidal directional filter banks

Combining the Laplacian pyramid and the directional filter bank, we are now ready to describe their combination into a double filter bank structure that was motivated in Section III-A. Since the directional filter bank (DFB) was designed to capture the high frequency (representing directionality) of the input image, the low frequency content is poorly handled. In fact, with the frequency partition shown in Figure 3(a), low frequency would “leak” into several directional subbands, hence the DFB alone does *not* provide a sparse representation for images. This fact provides another reason to combine the DFB with a multiscale decomposition, where low frequencies of the input image are removed before applying the DFB.

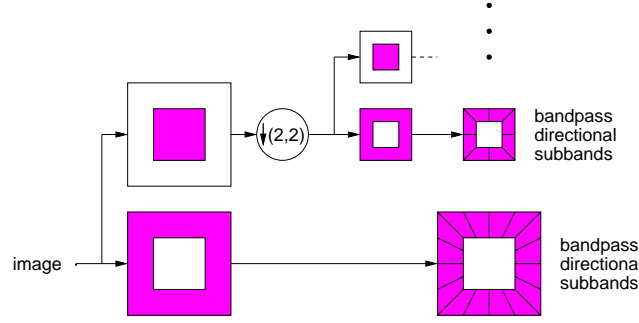


Fig. 7. Pyramidal directional filter bank that implements a discrete contourlet transform. First, a multiscale decomposition into octave bands by the Laplacian pyramid is computed, and then a directional filter bank is applied to each bandpass channel.

Figure 7 shows a multiscale and directional decomposition using a combination of a Laplacian pyramid (LP) and a directional filter bank (DFB). Bandpass images from the LP are fed into a DFB so that directional information can be captured. The scheme can be iterated on the coarse image. The combined result is a double iterated filter bank structure, named *pyramidal directional filter bank* (PDFB) [25] or *discrete contourlet transform*, which decomposes images into directional subbands at multiple scales. The main properties of the PDFB are stated in the following theorem.

Theorem 1: In a pyramidal directional filter bank, the following hold:

- 1) If both the LP and the DFB use perfect reconstruction filters, then the PDFB achieves perfect reconstruction, which means it provides a frame operator.
- 2) If both the LP and the DFB use orthogonal filters, then the PDFB provides a tight frame with frame bounds equal to 1.
- 3) The PDFB has a redundancy ratio up to $4/3$.
- 4) Suppose an l_j -level DFB is applied at the pyramidal level j of the LP ($j = 1, 2, \dots, J$, where $j = 1$ corresponds the finest scale), then the basis images of the PDFB have an essential support size of *width* $\approx 2^j$ and *length* $\approx 2^{j+l_j-2}$.
- 5) Using FIR filters, the computational complexity of the PDFB is $O(N)$ for N -pixel images.

Proof:

- 1) This is obvious as the PDFB is a composition of perfect reconstruction blocks.
- 2) Suppose an input image x to the PDFB is first decomposed into J bandpass images b_j , $j = 1, 2, \dots, J$ and a lowpass image a_J by the LP. Since with orthogonal filters, the LP is a tight frame with frame bounds equal to 1 [29], we have $\|x\|^2 = \sum_{j=1}^J \|b_j\|^2 + \|a_J\|^2$.

With orthogonal filters, the DFB is an orthogonal transform [32], which means that it transforms

each bandpass image b_j into a set of coefficients d_j with: $\|b_j\|^2 = \|d_j\|^2$. Combining two steps, the decomposition by PDFB: $x \mapsto (d_1, d_2, \dots, d_J, a_J)$ satisfies the tight frame condition: $\|x\|^2 = \sum_{j=1}^J \|d_j\|^2 + \|a_J\|^2$.

- 3) Since the DFB is critically sampled, the redundancy of the PDFB is equal to the redundancy of the LP, which is $1 + \sum_{j=1}^J (1/4)^j < 4/3$.
- 4) Using multirate identities, the LP bandpass channel corresponding to the pyramidal level j is approximately equivalent to filtering by a filter of size about 2×2^j , followed by downsampling by 2^{j-1} in each dimension. For the DFB, from (1) we see that after l_j levels ($l_j \geq 2$) of tree-structured decomposition, the equivalent directional filter would have support of width about 2 and length about 2^{l_j-1} . Combining these two steps into equivalent PDFB channels, we see that PDFB basis images would have support of width about 2^j and length about 2^{j+l_j-2} .
- 5) Let L_p and L_d be the number of taps of the pyramidal and directional filters used in the LP and the DFB, respectively (for simplicity, suppose lowpass, highpass, analysis and synthesis filters have same length). With a polyphase implementation, the LP filter bank requires $L_p/2 + 1$ operations per input sample.³ Thus, for an N -pixel image, the complexity of the LP step in the PDFB is:

$$\sum_{j=1}^J N \left(\frac{1}{4}\right)^{j-1} \left(\frac{L_p}{2} + 1\right) < \frac{4}{3} N \left(\frac{L_p}{2} + 1\right) \text{ (operations)} \quad (3)$$

For the DFB, its building block two-channel filter banks requires L_d operations per input sample. With an l -level full binary tree decomposition, the complexity of the DFB multiplies by l . This holds because the initial decomposition block in the DFB is followed by two blocks at half rate, four blocks at quarter rate and so on. Thus, the complexity of the DFB step for an N -pixel image is:

$$\sum_{j=1}^J N \left(\frac{1}{4}\right)^{j-1} L_d l_j < \frac{4}{3} N L_d \max \{l_j\} \text{ (operations)} \quad (4)$$

Combining (3) and (4) we obtain the desired result. ■

Since the multiscale and directional decomposition steps are decoupled in the PDFB or discrete contourlet transform, we can have a different number of directions at different scales, thus offering a flexible multiscale and directional expansion. Moreover, the full binary tree decomposition of the DFB in the contourlet transform can be generalized to arbitrary tree structures, similar to the *wavelet packets* generalization of

³Here we assume all filters are implemented non-separably. For certain filters, separable filtering (may be in polyphase domain) is possible and requires lower complexity.

the wavelet transform [33]. The result is a family of directional multiresolution expansions, which we call *contourlet packets*. Figure 8 shows examples of possible frequency decompositions by the contourlet transform and contourlet packets. In particular, contourlet packets allow finer angular resolution decomposition at any scale or direction, at the cost of spatial resolution.

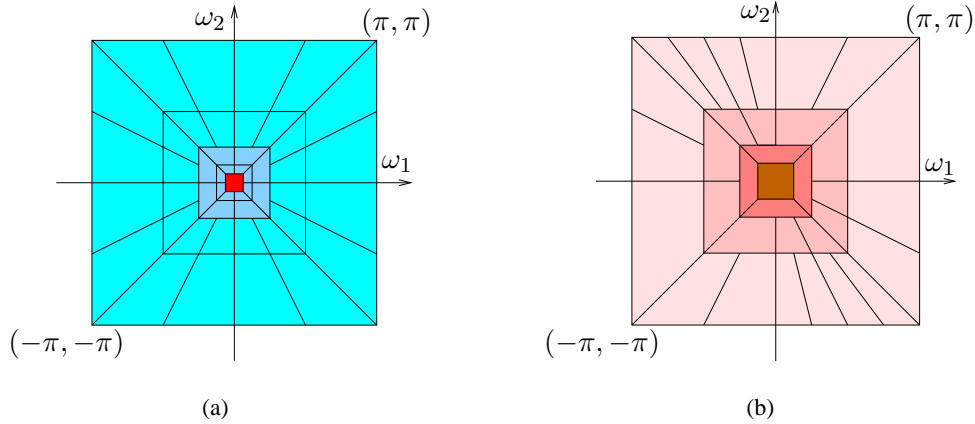


Fig. 8. Examples of possible frequency decompositions by the contourlet transform and contourlet packets.

In addition, from Theorem 1(4) we see that by altering the depth of the DFB tree decomposition tree at different scales (and even at different orientations in a contourlet packets transform), we obtain a rich set of contourlets with variety of support sizes and aspect ratios. This flexibility allows the contourlet transform and the contourlet packets to fit smooth contours of various curvatures well.

IV. CONTOURLETS AND DIRECTIONAL MULTIREOLUTION ANALYSIS

As for wavelet filter banks, the iterated PDFB has an associated continuous-domain expansion in $L_2(\mathbb{R}^2)$ using the *contourlet functions*. In this section, the connection between the discrete-domain PDFB and the continuous-domain contourlet expansion will be made precise via a new multiresolution analysis framework, similar to the link between wavelets and filter banks [34]. The new elements are multidirection and the combination with multiscale. For simplicity, we will only consider the case with orthogonal filters, which leads to tight frames. The more general case with biorthogonal filters can be treated similarly.

A. Multiscale

We start with the multiresolution analysis for the LP, which is similar to the one for wavelets. Suppose that the LP in the PDFB uses orthogonal filters and downsampling by 2 in each dimension (that means

$M = \text{diag}(2, 2)$ in Figure 2). Under certain regularity conditions, the lowpass filter G in the iterated LP uniquely defines a scaling function $\phi(\mathbf{t}) \in L_2(\mathbb{R}^2)$, which satisfies the two-scale equation [9], [2]

$$\phi(\mathbf{t}) = 2 \sum_{\mathbf{n} \in \mathbb{Z}^2} g[\mathbf{n}] \phi(2\mathbf{t} - \mathbf{n}). \quad (5)$$

Let us denote

$$\phi_{j,\mathbf{n}} = 2^{-j} \phi\left(\frac{\mathbf{t} - 2^j \mathbf{n}}{2^j}\right), \quad j \in \mathbb{Z}, \mathbf{n} \in \mathbb{Z}^2. \quad (6)$$

Then the family $\{\phi_{j,\mathbf{n}}\}_{\mathbf{n} \in \mathbb{Z}^2}$ is an orthonormal basis for an approximation subspace V_j at the scale 2^j . Furthermore $\{V_j\}_{j \in \mathbb{Z}}$ provides a sequence of multiresolution nested subspaces $\dots V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \dots$, where V_j is associated with a uniform grid of intervals $2^j \times 2^j$ that characterizes image approximation at scale 2^j . The difference images in the LP contain the details necessary to increase the resolution between two consecutive approximation subspaces. Therefore the difference images live in a subspace W_j that is the orthogonal complement of V_j in V_{j-1} .

$$V_{j-1} = V_j \oplus W_j \quad (7)$$

In [29] we show that each level in the LP can be considered as an oversampled filter bank where each polyphase component of the difference image, together with the coarse image, comes from a separate filter bank channel. Let $F_i(\mathbf{z}), 0 \leq i \leq 3$ be the synthesis filters for these polyphase components. These are highpass filters. As for wavelets, we associate with each of these filters a continuous function $\psi^{(i)}(\mathbf{t})$ where

$$\psi^{(i)}(\mathbf{t}) = 2 \sum_{\mathbf{n} \in \mathbb{Z}^2} f_i[\mathbf{n}] \phi(2\mathbf{t} - \mathbf{n}). \quad (8)$$

Proposition 1 ([29]): Using $\psi^{(i)}(\mathbf{t})$ in (8), define

$$\psi_{j,\mathbf{n}}^{(i)}(\mathbf{t}) = 2^{-j} \psi^{(i)}\left(\frac{\mathbf{t} - 2^j \mathbf{n}}{2^j}\right), \quad j \in \mathbb{Z}, \mathbf{n} \in \mathbb{Z}^2. \quad (9)$$

Then, for scale 2^j , $\{\psi_{j,\mathbf{n}}^{(i)}\}_{0 \leq i \leq 3, \mathbf{n} \in \mathbb{Z}^2}$ is a tight frame for W_j . For all scales, $\{\psi_{j,\mathbf{n}}^{(i)}\}_{j \in \mathbb{Z}, 0 \leq i \leq 3, \mathbf{n} \in \mathbb{Z}^2}$ is a tight frame for $L_2(\mathbb{R}^2)$. In both cases, the frame bounds are equal to 1.

Since W_j is generated by four kernel functions (similar to multi-wavelets), in general it is *not* a shift invariant subspace. Nevertheless, we can simulate a shift invariant subspace by denoting

$$\mu_{j-1,2\mathbf{n}+\mathbf{k}_i}(\mathbf{t}) = \psi_{j,\mathbf{n}}^{(i)}(\mathbf{t}), \quad 0 \leq i \leq 3, \quad (10)$$

where \mathbf{k}_i are the coset representatives for downsampling by 2 in each dimension

$$\mathbf{k}_0 = (0, 0)^T, \mathbf{k}_1 = (1, 0)^T, \mathbf{k}_2 = (0, 1)^T, \mathbf{k}_3 = (1, 1)^T. \quad (11)$$

With this notation, the family $\{\mu_{j-1,\mathbf{n}}\}_{\mathbf{n} \in \mathbb{Z}^2}$ associated to a uniform grid of intervals $2^{j-1} \times 2^{j-1}$ on \mathbb{R}^2 provides a tight frame for W_j .

B. Multidirection

In the iterated PDFB, the discrete basis (2) of the DFB can be regarded as a change of basis for the continuous-domain subspaces from the multiscale analysis in the last section. Suppose that the DFB's in the PDFB use orthogonal filters. Although in the PDFB the DFB is applied to the difference images or the detail subspaces W_j , we first show what happens when the DFB is applied to the approximation subspaces V_j .

Proposition 2: Define

$$\theta_{j,k,\mathbf{n}}^{(l)}(\mathbf{t}) = \sum_{\mathbf{m} \in \mathbb{Z}^2} g_k^{(l)}[\mathbf{m} - \mathbf{S}_k^{(l)} \mathbf{n}] \phi_{j,\mathbf{m}}(\mathbf{t}), \quad (12)$$

for arbitrary but finite⁴ l . Then the family $\{\theta_{j,k,\mathbf{n}}^{(l)}\}_{\mathbf{n} \in \mathbb{Z}^2}$ is an orthonormal basis of a directional subspace $V_{j,k}^{(l)}$ for each $k = 0, \dots, 2^l - 1$. Furthermore,

$$V_{j,k}^{(l)} \perp V_{j,k'}^{(l)} \quad \text{for } k \neq k', \quad (13)$$

$$V_{j,k}^{(l)} = V_{j,2k}^{(l+1)} \oplus V_{j,2k+1}^{(l+1)}, \quad \text{and} \quad (14)$$

$$V_j = \bigoplus_{k=0}^{2^l-1} V_{j,k}^{(l)}. \quad (15)$$

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 [33] (see also [2]). Suppose that $\{\theta_{j,k,\mathbf{n}}^{(l)}\}_{\mathbf{n} \in \mathbb{Z}^2}$ is an orthonormal basis of a subspace $V_{j,k}^{(l)}$. To increase the directional resolution, an extra level of decomposition by a pair of orthogonal filters is applied to the channel represented by $g_k^{(l)}$ that leads to two channels with equivalent filters $g_{2k}^{(l+1)}$ and $g_{2k+1}^{(l+1)}$. This transforms the orthonormal basis $\{\theta_{j,k,\mathbf{n}}^{(l)}\}_{\mathbf{n} \in \mathbb{Z}^2}$ in two orthonormal families $\{\theta_{j,2k,\mathbf{n}}^{(l+1)}\}_{\mathbf{n} \in \mathbb{Z}^2}$ and $\{\theta_{j,2k+1,\mathbf{n}}^{(l+1)}\}_{\mathbf{n} \in \mathbb{Z}^2}$. Each of these families generate a subspace with finer directional resolution that satisfies the “two-direction” equation (14). With this induction, starting from an orthonormal basis $\{\phi_{j,\mathbf{n}}\}_{\mathbf{n} \in \mathbb{Z}^2}$ of V_j , we obtain orthonormal bases for all directional subspaces $V_{j,k}^{(l)}$. ■

C. Multiscale and multidirection: the contourlet expansion

Applying the directional decomposition by the family (2) onto the detail subspace W_j as done by the PDFB, we obtain a similar result.

⁴The situation when the number of levels l of the iterated DFB goes to ∞ involves a regularity study for the DFB, which will be treated elsewhere.

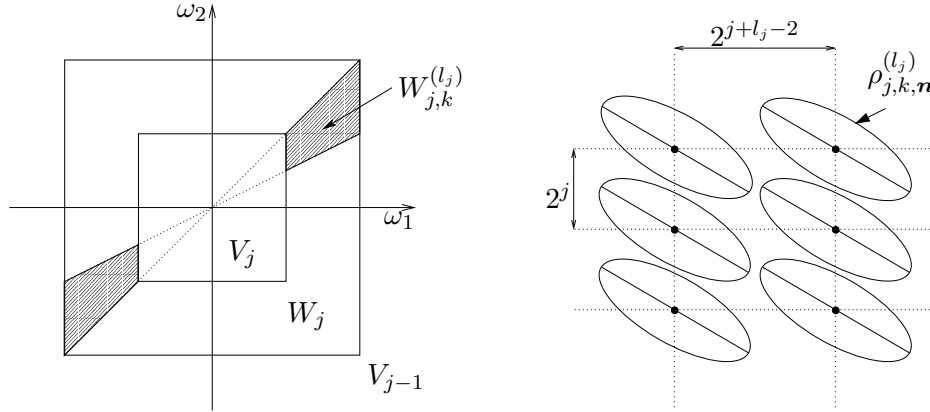


Fig. 9. (a) Multiscale and multidirectional subspaces generated by the contourlet transform which is illustrated on a 2-D spectrum decomposition. (b) Sampling grid and approximate support of contourlet functions for a ‘mostly horizontal’ subspace $W_{j,k}^{(l_j)}$. For ‘mostly vertical’ subspaces, the grid is transposed.

Proposition 3: Define

$$\rho_{j,k,n}^{(l)}(\mathbf{t}) = \sum_{\mathbf{m} \in \mathbb{Z}^2} g_k^{(l)}[\mathbf{m} - \mathbf{S}_k^{(l)} \mathbf{n}] \mu_{j-1,\mathbf{m}}(\mathbf{t}) \quad (16)$$

The family $\{\rho_{j,k,n}^{(l)}\}_{\mathbf{n} \in \mathbb{Z}^2}$ is a tight frame of a detail directional subspace $W_{j,k}^{(l)}$ with frame bounds equal to 1, for each $k = 0, \dots, 2^l - 1$. Furthermore, the subspaces $W_{j,k}^{(l)}$ are mutually orthogonal, across scales and directions.

Proof: This result is obtained by applying Proposition 1 to the subspaces in Proposition 2. ■

Figure 9(a) shows a graphical representation of the detail directional subspaces $W_{j,k}^{(l)}$ in the frequency-domain. The indexes j , k , and \mathbf{n} specify the scale, direction, and location, respectively. Note that the number of DFB decomposition levels l can be different at different scales j , and in that case will be denoted by l_j .

Recall that W_j is *not* a shift-invariant subspace. However, the following result establishes that its subspaces $W_{j,k}^{(l)}$ are, since they are generated by a single function and its translations.

Proposition 4: Let us denote

$$\rho_{j,k}^{(l)}(\mathbf{t}) = \sum_{\mathbf{m} \in \mathbb{Z}^2} g_k^{(l)}[\mathbf{m}] \mu_{j-1,\mathbf{m}}(\mathbf{t}) \quad (17)$$

Then for $l \geq 2$

$$\rho_{j,k,n}^{(l)}(\mathbf{t}) = \rho_{j,k}^{(l)}(\mathbf{t} - 2^{j-1} \mathbf{S}_k^{(l)} \mathbf{n}) \quad (18)$$

Proof: The definition of $\psi_{j,\mathbf{n}}$ in (9) implies that $\psi_{j,\mathbf{m}+\mathbf{n}}(\mathbf{t}) = \psi_{j,\mathbf{m}}(\mathbf{t} - 2^j \mathbf{n})$. Applying this to (10) we have $\mu_{j-1,\mathbf{m}+2\mathbf{n}}(\mathbf{t}) = \mu_{j-1,\mathbf{m}}(\mathbf{t} - 2^{j-1} 2\mathbf{n})$. In other words, $\mu_{j-1,\mathbf{m}}$ are periodically shift invariant with even shifts in \mathbf{m} . From (1) we see that when $l \geq 2$, sampling by $\mathbf{S}_k^{(l)}$ is also even in each dimension. Thus,

with a change of variable we obtain

$$\rho_{j,k,\mathbf{n}}^{(l)}(\mathbf{t}) = \sum_{\mathbf{m} \in \mathbb{Z}^2} g_k^{(l)}[\mathbf{m} - \mathbf{S}_k^{(l)}\mathbf{n}] \mu_{j-1,\mathbf{m}}(\mathbf{t}) = \sum_{\mathbf{m} \in \mathbb{Z}^2} g_k^{(l)}[\mathbf{m}] \mu_{j-1,\mathbf{m}}(\mathbf{t} - 2^{j-1}\mathbf{S}_k^{(l)}\mathbf{n}) = \rho_{j,k}^{(l)}(\mathbf{t} - 2^{j-1}\mathbf{S}_k^{(l)}\mathbf{n}).$$

■

Therefore the translated family of $\rho_{j,k}^{(l)}$

$$\left\{ \rho_{j,k,\mathbf{n}}^{(l)}(\mathbf{t}) = \rho_{j,k}^{(l)}(\mathbf{t} - 2^{j-1}\mathbf{S}_k^{(l)}\mathbf{n}) \right\}_{\mathbf{n} \in \mathbb{Z}^2} \quad (19)$$

is a frame of $W_{j,k}^{(l)}$. As a result, the subspace $W_{j,k}^{(l)}$ is defined on a rectangular grid with intervals $2^{j+l-2} \times 2^j$ or $2^j \times 2^{j+l-2}$, depending on whether it is mostly horizontal or vertical (see Figure 9(b)). The reason that $\{\rho_{j,k,\mathbf{n}}^{(l)}\}_{\mathbf{n} \in \mathbb{Z}^2}$ is an overcomplete frame for $W_{j,k}^{(l)}$ is because it uses the same sampling grid of the bigger subspace $V_{j-1,k}^{(l)}$.

Substituting (8) into (10) and then into (17), we can write $\rho_{j,k}^{(l)}(\mathbf{t})$ directly as a linear combination of the scaling functions as

$$\begin{aligned} \rho_{j,k}^{(l)}(\mathbf{t}) &= \sum_{i=0}^3 \sum_{\mathbf{n} \in \mathbb{Z}^2} g_k^{(l)}[2\mathbf{n} + \mathbf{k}_i] \left(\sum_{\mathbf{m} \in \mathbb{Z}^2} f_i[\mathbf{m}] \phi_{j-1,2\mathbf{n}+\mathbf{m}} \right) \\ &= \sum_{\mathbf{m} \in \mathbb{Z}^2} \underbrace{\left(\sum_{i=0}^3 \sum_{\mathbf{n} \in \mathbb{Z}^2} g_k^{(l)}[2\mathbf{n} + \mathbf{k}_i] f_i[\mathbf{m} - 2\mathbf{n}] \right)}_{c_k^{(l)}[\mathbf{m}]} \phi_{j-1,\mathbf{m}}(\mathbf{t}). \end{aligned} \quad (20)$$

The discrete signal $c_k^{(l)}$ roughly equals to the summation of convolutions between $g_k^{(l)}$ and f_i 's, and thus it resembles a highpass and directional filter. Equation (20) shows that $\rho_{j,k}^{(l)}$ can be viewed as a grouping of “edge-detection” elements $\phi_{j-1,\mathbf{m}}$ at a scale 2^{j-1} and along a direction k . This justifies the name *contourlet* for the functions $\rho_{j,k}^{(l)}$. It can be checked that

$$\|\rho_{j,k,\mathbf{n}}^{(l)}\|_2^2 = 3/4, \quad \text{for all } l \geq 2, j \in \mathbb{Z}, 0 \leq k < 2^l, \mathbf{n} \in \mathbb{Z}^2. \quad (21)$$

Finally, integrating the multidirection analysis over scales we have the following result for the *contourlet frames* of $L_2(\mathbb{R}^2)$.

Theorem 2: For a sequence of finite positive integers $\{l_j\}_{j \leq j_0}$ the family

$$\{\phi_{j_0,\mathbf{n}}(\mathbf{t}), \rho_{j,k,\mathbf{n}}^{(l_j)}(\mathbf{t})\}_{j \leq j_0, 0 \leq k \leq 2^{l_j}-1, \mathbf{n} \in \mathbb{Z}^2} \quad (22)$$

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

$$\{\rho_{j,k,\mathbf{n}}^{(l_j)}(\mathbf{t})\}_{j \in \mathbb{Z}, 0 \leq k \leq 2^{l_j}-1, \mathbf{n} \in \mathbb{Z}^2} \quad (23)$$

is a tight frame of $L_2(\mathbb{R}^2)$. In both cases, the frame bounds are equal to 1.

Proof: This result is obtained by applying Proposition 3 to the following decompositions of $L_2(\mathbb{R}^2)$ into mutual orthogonal subspaces:

$$L_2(\mathbb{R}^2) = V_{j_0} \oplus \left(\bigoplus_{j \leq j_0} W_j \right), \quad \text{and}$$

$$L_2(\mathbb{R}^2) = \bigoplus_{j \in \mathbb{Z}} W_j.$$

■

The contourlet frames in (22)-(23) have several distinguishing features that are important to emphasize.

- 1) The contourlet expansions are defined on rectangular grids, and thus offer a seamless translation to the discrete world, where image pixels are stored on a rectangular grid. To achieve this “digital-friendly” feature, the contourlet kernel functions $\rho_{j,k}^{(l_j)}$ have to be different for different directions k and cannot be obtained by simply rotating a single function. This is a key difference between the contourlet and the curvelet [4], [5] systems. The contourlet kernel functions $\rho_{j,k}^{(l_j)}$ at different directions are defined iteratively via the DFB as described in the proof of Proposition 2.
- 2) Since the contourlet functions are defined via the iterated PDFB like wavelets, the contourlet expansions have fast filter bank algorithms and convenient tree structures.
- 3) It is easy to verify that with FIR filters, the iterated PDFB leads to compactly supported contourlet frames. More precisely, the contourlet function $\rho_{j,k,n}^{(l_j)}$ has support of size *width* $\approx C2^j$ and *length* $\approx C2^{j+l_j-2}$. Thus at each scale and direction, the set $\left\{ \rho_{j,k,n}^{(l_j)} \right\}_{n \in \mathbb{Z}^2}$ tiles the plane \mathbb{R}^2 .
- 4) The contourlet construction provides a space-domain multiresolution scheme that offers flexible refinements for the spatial resolution and the angular resolution.

V. CONTOURLET APPROXIMATION AND COMPRESSION

The proposed contourlet filter bank and its associated continuous-domain frames in the previous sections provide a framework for constructing general directional multiresolution image representations. Since our goal is to develop efficient or sparse expansions for images having smooth contours, the next important issues are: (1) what conditions should we impose in order to obtain a sparse contourlet expansion for that class of images; and (2) how can we design filter banks that can lead to contourlet expansions satisfying those conditions. We consider the first issue in this paper; the second one is addressed in a forthcoming paper.

A. Parabolic scaling

In the curvelet construction, Candès and Donoho [4] point out that a key to achieving the correct nonlinear approximation behavior by curvelets is to select support sizes obeying the *parabolic scaling* relation for curves: $width \propto length^2$. The same scaling relation has been used in the study of Fourier integral operators and wave equations, for example see [35].

The motivation behind the parabolic scaling is to efficiently approximate a smooth discontinuity curve by “laying” basis elements with linear or rectangular supports along the curve (refer to “X-let” scheme in Figure 1). Suppose that the discontinuity curve is sufficiently smooth, namely it is C^2 , then locally, by the Taylor series expansion, it can be approximated by a parabolic curve. More precisely, with the local coordinate setup as in Figure 10, it is easy to check that the parametric representation of the discontinuity curve obeys

$$u(v) \approx \frac{\kappa}{2}v^2, \quad \text{when } v \approx 0, \quad (24)$$

where κ is the local curvature of the curve. Hence, in order to match the discontinuity curve, at *fine scale* the width w and the length l of the basis functions have to satisfy (refer to Figure 10)

$$w \approx \frac{\kappa}{8}l^2. \quad (25)$$

For the contourlet frame in (22), we know that when an l_j -level DFB is applied to the pyramidal scale 2^j , the contourlet functions have support size of $width \sim 2^j$ and $length \sim 2^{l_j+j-2}$. Hence to make the contourlet expansion satisfy the parabolic scaling, we simply have to enforce that *the number of directions is doubled at every other finer scale in the pyramid*. An example of such a frequency decomposition is shown in Figure 8(a). More precisely, suppose that at the scale 2^{j_0} we start with an l_{j_0} -level DFB, then at subsequently finer scales 2^j ($j < j_0$), the number of DFB decomposition levels

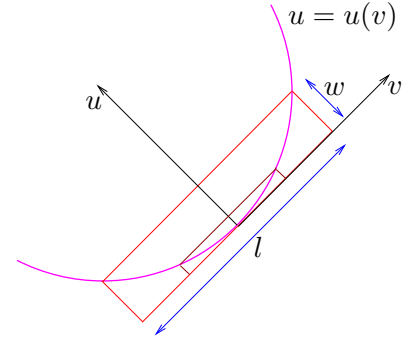


Fig. 10. Illustrating the parabolic scaling relation for curves. The rectangular supports of the basis functions that fit a curve exhibit the quadric relation: $width \propto length^2$.

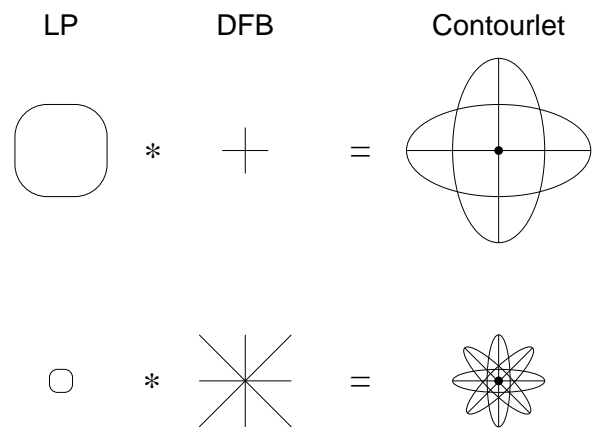


Fig. 11. Illustrating the evolution of the support size of a contourlet frame that satisfies the parabolic scaling.

has to be

$$l_j = l_{j_0} - \lfloor (j - j_0)/2 \rfloor, \quad \text{for } j \leq j_0. \quad (26)$$

Figure 11 graphically depicts a contourlet frame satisfying the parabolic scaling. As can be seen in the two pyramidal levels shown, as the support size of the basis element of the LP is reduced by four in each dimension, the number of directions of the DFB is doubled. Combining these two stages, the support sizes of the contourlet functions evolve in accordance to the desired parabolic scaling.

B. Directional vanishing moment

As an analogy, for the wavelet case in 1-D, wavelet approximation theory brought a new condition into filter bank design, which earlier only focused on designing filters with good frequency selection properties. This new condition requires wavelet functions to have a sufficient number of *vanishing moments*, or equivalently, the lowpass filter must have enough “zeros at $\omega = \pi$ ”. The vanishing-moments property is the key for the sparse expansion of piecewise smooth signals by wavelets [8], [2]. Intuitively, wavelets with vanishing moments are orthogonal with polynomial signals, and thus only a few wavelet basis functions around discontinuity points would “feel” these singularities and lead to significant coefficients [36].

In the contourlet case, our target for approximation is 2-D piecewise smooth functions with discontinuities along smooth curves. For this type of functions, *singularities are localized in both space and direction*. More specifically, a local region around a smooth contour has a singularity that can be approximated by two polynomial surfaces separated by a straight line. Thus, it is desirable that only the contourlet functions that are located *and* oriented around the discontinuity curve would “feel” this singularity. One simple way to achieve this desideratum is to require all 1-D slices in a certain direction of contourlet functions to have vanishing moments. We refer this requirement as the *directional vanishing moment* (DVM) condition.

Definition 1 (Directional vanishing moment): A 2-D function $\rho(t_1, t_2)$ is said to have an L -order directional vanishing moment along a direction $\mathbf{u} = (u_1, u_2)^T$ if all 1-D slices of that function along the direction \mathbf{u} (suppose that $u_1 \neq 0$; if $u_1 = 0$ then $u_2 \neq 0$ and we can swap the two dimensions): $\rho_{\mathbf{u},d}(t) = \rho(t, tu_2/u_1 - d)$, have L vanishing moments:

$$\int_{-\infty}^{\infty} \rho_{\mathbf{u},d}(t) t^n dt = 0, \quad \text{for all } d \in \mathbb{R}, 0 \leq n < L. \quad (27)$$

In the frequency domain, the DVM condition (27) is equivalent to requiring $\hat{\rho}(\omega_1, \omega_2)$ to have a factor $(u_2\omega_1 - u_1\omega_2)^L$. For contourlet functions $\rho_{j,k}^{(l)}(\mathbf{t})$ constructed from an iterated PDFB as in (20), taking Fourier transforms on both sides, we see that $\rho_{j,k}^{(l)}(\mathbf{t})$ has L DVMs along direction \mathbf{u} if the discrete-time Fourier

transform $C_k^{(l)}(e^{j\omega_1}, e^{j\omega_2})$ of the filter $c_k^{(l)}[\mathbf{n}]$ has L zeros along the line $u_2\omega_1 - u_1\omega_2 = 0$, which means

$$C_k^{(l)}(e^{j\omega_1}, e^{j\omega_2}) = \left(1 - e^{j(u_2\omega_1 - u_1\omega_2)}\right)^L R(e^{j\omega_1}, e^{j\omega_2}) \quad (28)$$

This requirement in turn leads to filter design conditions for the PDFB. When u_1 and u_2 are integers, (28) is equivalent to requiring $C_k^{(l)}(z_1, z_2)$ to have a factor $(1 - z_1^{u_2} z_2^{-u_1})^L$. For the ideal frequency case (i.e. with sinc-type filters), the contourlet functions have DVMs on all orthogonal directions outside the fan regions of their frequency supports.

We note that the DVM property also holds in other 2-D expansions. In particular, 2-D separable wavelets have directional vanishing moments in the horizontal and vertical directions, which make wavelets especially good in capturing horizontal and vertical edges. Ridgelets [11], which offer an optimal representation for 2-D functions that are smooth away from a discontinuity along a line, have directional vanishing moments in all but one direction.

C. Approximation rate for 2-D piecewise smooth functions

Now we show that a contourlet expansion that satisfies the parabolic scaling and has sufficient DVMs achieves the optimal nonlinear approximation behavior for 2-D piecewise smooth functions with discontinuities along smooth curves. Recall that for a contourlet frame (22) to satisfy the parabolic scaling, l_j has to follow (26). For simplicity,⁵ set $j_0 = 0$ and $l_0 = 2$. This leads to a contourlet frame which at scale 2^j ($j = -1, -2, \dots$) has about $2^{-j/2}$ directions and each contourlet function $\rho_{j,k,\mathbf{n}}$ has support size of order *width* $\sim 2^j$ and *length* $\sim 2^{j/2}$. Because of (21), it follows that the maximum amplitude of $\rho_{j,k,\mathbf{n}}$ is of order of $A_j \sim 2^{-3j/4}$.

Consider a function f defined on the unit square $[0, 1]^2$ that is C^2 except for discontinuities along a C^2 and finite length curve \mathcal{C} . Suppose that the contourlet kernel functions $\rho_{j,k}$ have two-order DVMs along a *dense enough* set of directions (this will be made precise in a moment). We classify contourlets $\rho_{j,k,\mathbf{n}}$ into type 1, whose support intersects with the discontinuity curve \mathcal{C} , and type 2, whose support is included in a region where f is C^2 .

The key idea is as follows. For a type 1 contourlet $\rho_{j,k,\mathbf{n}}$, we can localize the support area that would be affected by the discontinuity curve \mathcal{C} to be “sandwiched” between two lines parallel to the tangent of the curve (see Figure 12(a)). Denote $d_{j,k,\mathbf{n}}$ to be the length of \mathcal{C} that intersects with the support of $\rho_{j,k,\mathbf{n}}$. Because the curve is twice differentiable, using Taylor expansion as in (24) (also see Figure 10), the width of the affected area is of the order $d_{j,k,\mathbf{n}}^2$. Outside that area, because f is C^2 and $\rho_{j,k,\mathbf{n}}$ has two vanishing

⁵Other values of l_0 only changes the constant but not the exponent in the approximation rate.

moments on the tangent direction, it can be shown that the line integrals along the direction of the product $f \cdot \rho_{j,k,n}$ are bounded by $A_j K d_{j,k,n}^3$. Hence, the corresponding type 1 contourlet coefficient behaves like

$$|\langle f, \rho_{j,k,n} \rangle| \sim A_j d_{j,k,n}^3 + A_j d_{j,k,n}^3 2^{j/2} \sim 2^{-3j/4} d_{j,k,n}^3. \quad (29)$$

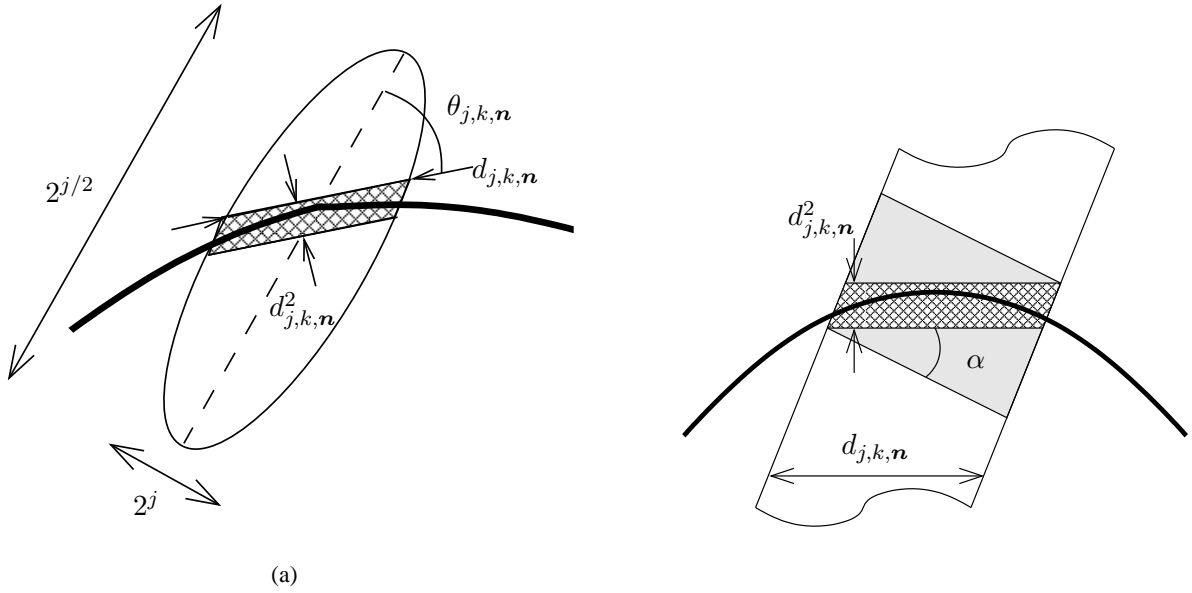


Fig. 12. (a) A contourlet function that intersects with the discontinuity curve (denoted by the thick curve). (b) Illustrating the tolerable gap in angle with directional vanishing moments. The hatched region is the ideal affected area, while the gray region is the actual affected area corresponding to DVMs on a direction that is α away from the tangent of the discontinuity curve.

Suppose that instead of having DVMs on the tangent direction of the discontinuity curve, we only find DVMs on a direction that is α degree away as shown in Figure 12(b). A simple geometrical argument shows that instead of $d_{j,k,n}^3$, the affected area would be upper bounded by $d_{j,k,n}^3 + d_{j,k,n}^2 \tan \alpha$. Thus, if $\tan \alpha \leq A_1 d$, then the decay in (29) will be unchanged. To achieve this and since $d_{j,k,n} \leq A_2 2^{j/2}$, the maximum gap α_j between directions that $\rho_{j,k,n}$ has DVMs must be of the order

$$\alpha_j \leq A_2 2^{j/2}. \quad (30)$$

We now estimate the intersection length $d_{j,k,n}$. Let $\theta_{j,k,n}$ denote the angle between the contourlet function $\rho_{j,k,n}$ and the local tangent direction of the discontinuity curve \mathcal{C} . From Figure 12(a), we see that $d_{j,k,n} \sim \max \{2^j / \sin \theta_{j,k,n}, 2^{j/2}\}$. Since there are about $2^{-j/2}$ directions, locally as the intersected contourlets slowly turn away from the discontinuity direction, the angle θ increases almost uniformly from 0 to $\pi/2$ with a step size about $2^{j/2}$. More precisely, within a square region of length $2^{j/2}$ we can reorder the direction

index to \tilde{k} , $\tilde{k} = 1, \dots, 2^{-j/2}$ so that $\theta_{j,\tilde{k},\mathbf{n}} \approx \tilde{k}2^{j/2}$. With this index, the intersection lengths behave like

$$d_{j,\tilde{k},\mathbf{n}} \sim 2^j / (\tilde{k}2^{j/2}) = 2^{j/2} \tilde{k}^{-1}.$$

Substituting this into (29), we have the following decay for type 1 contourlet coefficients at scale 2^j and \tilde{k} -th direction away from the discontinuity curve

$$|\langle f, \rho_{j,\tilde{k},\mathbf{n}} \rangle| \sim 2^{3j/4} \tilde{k}^{-3}. \quad (31)$$

In addition, since the discontinuity curve \mathcal{C} has finite length, there are $O(2^{-j/2} \tilde{k})$ corresponding contourlet coefficients.

We now turn attention to type 2 contourlets, whose support is included in a region where f is C^2 . For those contourlets, the corresponding coefficients $\langle f, \rho_{j,k,\mathbf{n}} \rangle$ behave as if f is C^2 . Because of the tight frame property in Proposition 3, we have

$$\sum_{j=-\infty}^i \sum_k \sum_{\mathbf{n}} |\langle f, \rho_{j,k,\mathbf{n}} \rangle|^2 = \|f - P_{V_i} f\|_2^2 \quad (32)$$

where $P_{V_i} f$ is the projection of f onto the approximation subspace V_i at scale 2^i . Therefore, contourlets are as effective as wavelets in approximating smooth functions.

Suppose that the scaling function ϕ has *accuracy* of order 2, which is equivalent to requiring the refinement filter $G(e^{j\omega_1}, e^{j\omega_2})$ in (5) to have a second-order zero at (π, π) [37], that is

$$\left. \frac{\partial^{k_1+k_2} G(e^{j\omega_1}, e^{j\omega_2})}{\partial^{k_1} \omega_1 \partial^{k_2} \omega_2} \right|_{(\pi,\pi)} = 0, \quad \text{for all } k_1, k_2 \in \mathbb{Z}; 0 \leq k_1 + k_2 < 2. \quad (33)$$

Then for a smooth function \tilde{f} in C^2 , we have

$$\|\tilde{f} - P_{V_i} \tilde{f}\|_2 \sim (2^i)^2. \quad (34)$$

Thus, combined with (32), the sum square error of type 2 contourlet coefficients $\langle f, \rho_{j,k,\mathbf{n}} \rangle$ due to truncation up to scale 2^i is $O(2^{4i})$. In addition, because there are $O(2^{-2j})$ type 2 contourlet coefficients at scale 2^j , it follows that type 2 contourlet coefficients at scale 2^j decay like $O(2^{3j})$.

For a best- M term approximation, based on the above rates of decay of contourlet coefficients, by setting the threshold $T = 2^{-L}$ ($L \gg 0$), we would keep all type 1 coefficients with $\tilde{k} \leq 2^{j/4+L/3}$ and all type 2 coefficients to scale $2^{-L/3}$. Moreover, since $\tilde{k} \leq 2^{-j/2}$, we have the following scheme for picking contourlet coefficients above threshold $T = 2^{-L}$.

- For $0 \geq j \geq -L/3$: keep all coefficients.
- For $-L/3 > j \geq -4L/9$: keep all type 1 (intersected) coefficients.
- For $-4L/9 > j \geq -4L/3$: keep all type 1 (intersected) coefficients with $\tilde{k} \leq 2^{j/4+L/3}$.

Then the total number of retained coefficients is:

$$M \sim 2^{2L/3} + \sum_{j=-4L/9}^{-L/3} \sum_{\tilde{k}=1}^{2^{-j/2}} 2^{-j/2} \tilde{k} + \sum_{j=-4L/3}^{-4L/9} \sum_{\tilde{k}=1}^{2^{j/4+L/3}} 2^{-j/2} \tilde{k} \sim L 2^{2L/3}. \quad (35)$$

The distortion due to truncation is

$$\begin{aligned} \|f - \hat{f}_M^{\text{contourlet}}\|^2 &\sim 2^{-4L/3} + \sum_{j=-4L/3}^{-4L/9} \sum_{\tilde{k}=2^{j/4+L/3}}^{2^{-j/2}} 2^{-j/2} \tilde{k} (2^{3j/4} \tilde{k}^{-3})^2 + \sum_{j=-\infty}^{-4L/3} \sum_{\tilde{k}=1}^{2^{-j/2}} 2^{-j/2} \tilde{k} (2^{3j/4} \tilde{k}^{-3})^2 \\ &\sim 2^{-4L/3} + \sum_{j=-4L/3}^{-4L/9} 2^j \sum_{\tilde{k}=2^{j/4+L/3}}^{2^{-j/2}} \tilde{k}^{-5} + \sum_{j=-\infty}^{-4L/3} 2^j \sum_{\tilde{k}=1}^{2^{-j/2}} \tilde{k}^{-5}. \end{aligned} \quad (36)$$

To simplify the last expression, we use the following approximation

$$\sum_{\tilde{k}=a}^b \tilde{k}^{-5} \sim \int_a^b x^{-5} dx = (a^{-4} - b^{-4})/4.$$

Substitute this back into (36) we obtain

$$\|f - \hat{f}_M^{\text{contourlet}}\|^2 \sim L 2^{-4L/3}. \quad (37)$$

Combining (35) with (37) we have

$$\|f - \hat{f}_M^{\text{contourlet}}\|^2 \sim L^3 M^{-2} \sim (\log M)^3 M^{-2}.$$

The following theorem summarizes this nonlinear approximation property of contourlets.

Theorem 3: Suppose that the contourlet frame (22) satisfies the parabolic scaling condition (26), each contourlet function $\rho_{j,k,\mathbf{n}}$ has two-order directional vanishing moments on a set of directions with maximum angular gap of $A 2^{j/2}$, and the scaling function ϕ has accuracy of order 2. Then for a function f that is C^2 away from a C^2 discontinuity curve, the M -term approximation by this contourlet frame achieves

$$\|f - \hat{f}_M^{\text{contourlet}}\|^2 \leq C(\log M)^3 M^{-2}. \quad (38)$$

Interestingly, this is exactly the same approximation behavior derived for curvelets [5], where curvelets are compactly supported in frequency and the approximation analysis was carried out in the Fourier domain. For comparison, the approximation error of the same function f by wavelets decays like M^{-1} (see for example [2]). Because the “complexity” of f is at least equal to the “complexity” of the discontinuity curve \mathcal{C} , which is a C^2 curve, no other approximation scheme can achieve a better rate than M^{-2} [10]. In this sense, the contourlet expansion achieves the optimal approximation rate for piecewise smooth functions with C^2 contours.

D. Contourlet Compression

In the last subsection, we consider the approximation rate of contourlets by keeping the M -most significant coefficients. For the compression application, we have to account for the cost to index these M retained coefficients, as well as the cost to quantize them. Fortunately, as can be seen from the last subsection, the M retained coefficients are organized in tree-structures, where each contourlet coefficient has four children in the finer scale, at the scale location, and either in the same direction or in two finer directions (when the number of directions is doubled); see Figure 13. Indeed, in the scheme of picking M -best coefficients in the last subsection, from coarse to fine scales, we successively localize in both location (contourlets intersect with the discontinuity curve) and direction (intersected contourlets with direction close to the tangent of the discontinuity curve). These embedded tree-structures, which are similar to the embedded zero-trees for wavelets [38], allow us to efficiently index the retained coefficients using 1 bit per coefficient.

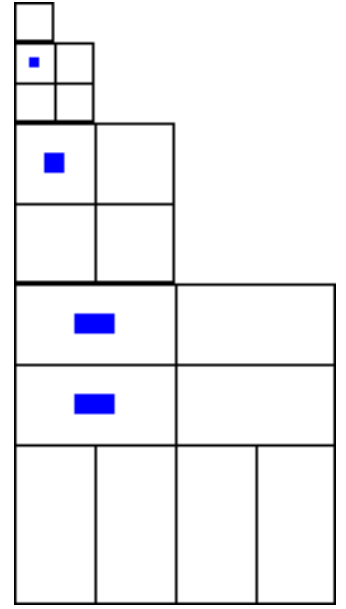


Fig. 13. A possible zero-tree of contourlet coefficients.

For quantization, instead of using fixed length coding to precision $T = 2^{-L}$, a slight gain (in the log factor, but not the exponent of the rate-distortion function) can be obtained by variable length coding. With variable length coding, we encode coefficients with magnitude in the range $(2^{l-1-L}, 2^{l-L}]$ with l bits. Such a scheme is obtained via bit plane coding [9]. In this scheme, the number of coefficients encoded with l bits is upper bounded by the number of coefficients with magnitude above the threshold 2^{l-1-L} , and according to (35) is of the order $(L - l + 1)2^{2(L-l+1)/3}$. Thus the total number of bits to encode the coefficients is

$$R \lesssim M + \sum_{l \geq 1} l(L - l + 1)2^{2(L-l+1)/3} \leq L2^{2L/3} + L2^{2L/3} \sum_{l \geq 1} l2^{-2(l-1)/3} \sim L2^{2L/3}. \quad (39)$$

The compression distortion is equal to the sum of the truncation distortion in (37) and the quantization distortion, which is $MT^2 \sim L2^{2L/3}(2^{-L})^2 = L2^{-4L/3}$. Thus the compression distortion is also of the order $L2^{-4L/3}$. Together with (39) we obtain the following estimation of the rate distortion function $D(R)$.

Corollary 1: Under the assumption of Theorem 3, a contourlet compression system that uses embedded zero-trees and bit plan coding achieves

$$D(R) \leq C'(\log R)^3 R^{-2}. \quad (40)$$

VI. NUMERICAL EXPERIMENTS

All experiments in this section use a wavelet transform with “9-7” biorthogonal filters [39], [40] and 6 decomposition levels. For the contourlet transform, in the LP stage we also use the “9-7” filters. The

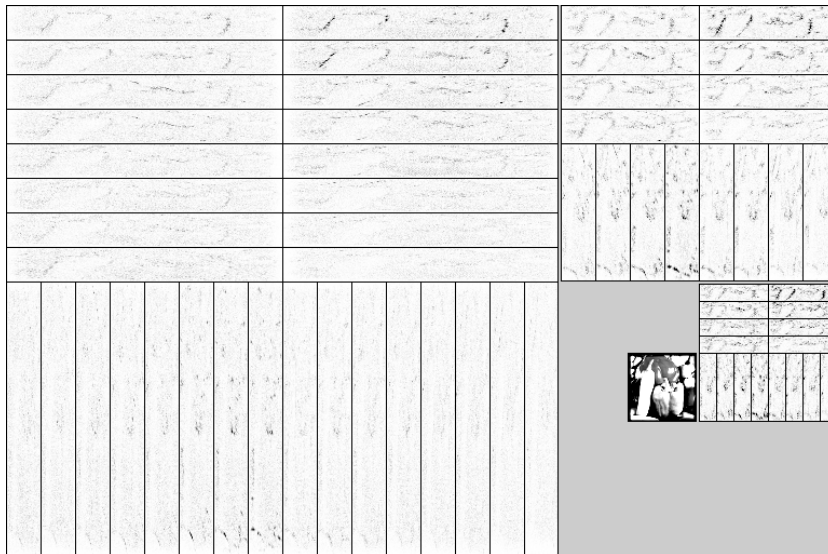


Fig. 14. Example of the discrete contourlet transform, applied to the *peppers* image.

choice of the “9-7” biorthogonal filters rather than orthogonal ones is motivated by the fact that these filters are very popular and successful in image processing practice (partly because they are linear phase) and actually they are close to being orthogonal. In the DFB stage we use the “23-45” biorthogonal quincunx filters designed by Phoong et al. [41] and modulate them to obtain the biorthogonal fan filters. Apart from being linear phase and nearly orthogonal, these fan filters are close to having the ideal frequency response and thus can approximate the directional vanishing moment condition. The drawback is that they have large support which creates a large number of significant coefficients near edges. As mentioned before, designing optimized contourlet filters is a topic to be studied further.

The number of DFB decomposition levels is doubled at every other finer scale and is equal to 5 at the finest scale. Note that in this case, both the wavelet and the contourlet transforms share the same detail subspaces W_j as defined in Section IV-A. The difference is that each detail subspace W_j in the wavelet transform is represented by a basis with three directions, whereas in the contourlet transform it is represented by a redundant frame with many more directions. An example of the contourlet transform is shown in Figure 14.

A. Wavelets vs. Contourlets

To highlight the difference between the wavelet and contourlet transform, Figure 15 shows a few wavelet and contourlet basis images. We see that contourlets offer a much richer set of directions and shapes, which help to capture geometric structures in images.

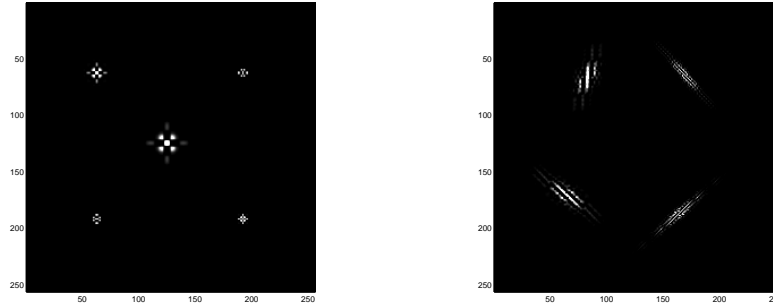


Fig. 15. Comparing a few actual 2-D wavelets (left) and contourlets (right).

B. Nonlinear approximation

Next we compare the nonlinear approximation (NLA) performances of the wavelet and the contourlet transforms. 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. Since the two transforms share the same detail subspaces, it is possible to restrict the comparison in these subspaces. We expect that most of the refinement happens around the image edges.

Figure 16 shows sequences of nonlinear approximated images at the finest detailed subspace W_j using the wavelet and the contourlet transforms, respectively, for the input *peppers* image. The wavelet scheme is seen to slowly capture contours by isolated “dots”. By contrast, the contourlet scheme quickly refines by well-adapted “sketches”, in much the same way as the “X-let” painter discussed in the Introduction.

Figure 17 shows a detailed comparison of two nonlinear approximated images by the wavelet and contourlet transforms using the same number of coefficients. Contourlets are shown to be superior compared with wavelets in capturing fine contours (directional textures on cloths). In addition, there is a significant gain of 1.46 dB in peak signal-to-noise ratio (PSNR) by contourlets.

C. Denoising

The improvement in approximation by contourlets based on keeping the most significant coefficients will directly lead to improvements in applications, including compression, denoising, and feature extraction. In particular, for image denoising, random noise will generate significant wavelet coefficients just like edges, but is less likely to generate significant contourlet coefficients. Consequently, a simple thresholding scheme [42] applied on the contourlet transform is more effective in removing the noise than it is for the wavelet transform.

Figure 18 displays a “zoom-in” comparison of denoising when applying wavelet and contourlet hard-thresholding on the *Lena* image. The contourlet transform is shown to be more effective in recovering

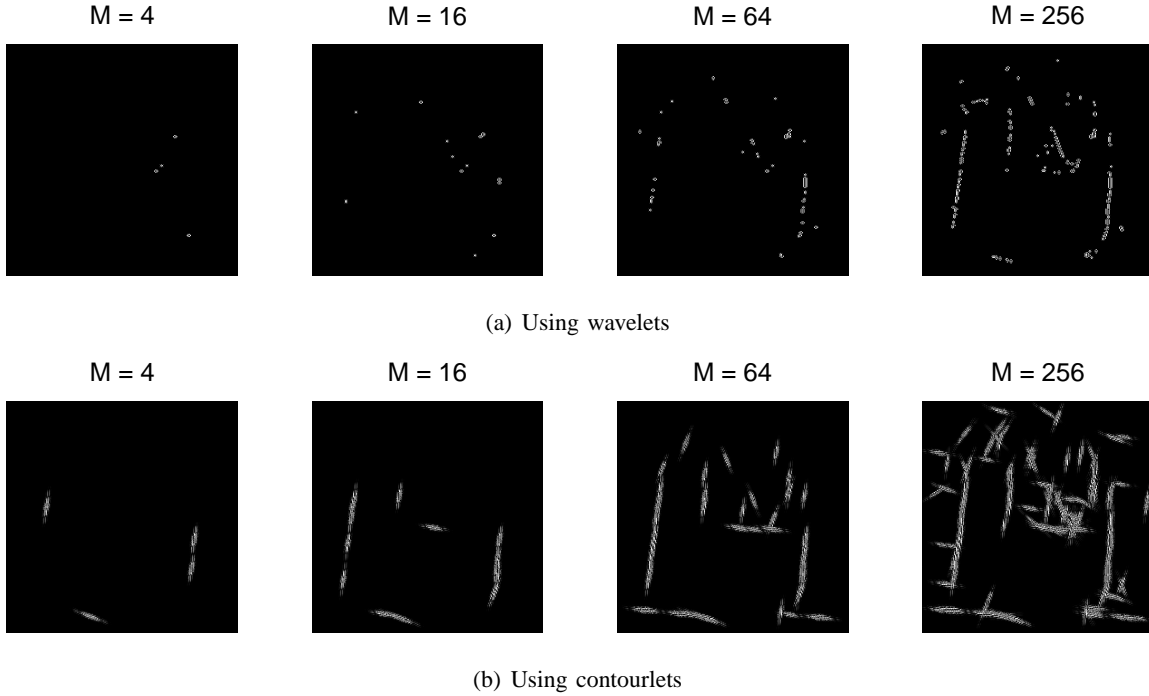


Fig. 16. Sequence of images showing the nonlinear approximations of the *peppers* image using M most significant coefficients at the finest detailed subspace W_j , which is shared by both the wavelet and contourlet transforms.

smooth contours, both visually as well as in signal-to-noise ratio (SNR). A more sophisticated denoising scheme that takes into account the dependencies across scales, directions and locations in the contourlet domain using a hidden Markov tree model is presented in [43] and shows further improvements.

VII. CONCLUSION

In this work, we constructed a discrete transform that can offer a sparse expansion for typical images having smooth contours. Using recent results from harmonic analysis and vision, we first identified two key features of a new image representation that improves over the separable 2-D wavelet transform, namely directionality and anisotropy. Based on this, we developed a new filter bank structure, the pyramidal directional filter bank (PDFB) or the discrete contourlet transform, that can provide a flexible multiscale and directional decomposition for images. The developed discrete filter bank has an associated continuous-domain expansion similar to filter bank iterations and the associated wavelet constructions. This connection was made precise via a newly defined directional multiresolution analysis that provides successive refinements at *both* spatial and directional resolution. With parabolic scaling and sufficient directional vanishing moments, the contourlet expansion is shown to achieve the optimal nonlinear approximation behavior for piecewise smooth functions in \mathbb{R}^2 with C^2 contours. Experiments with real images indicate the potential of contourlets

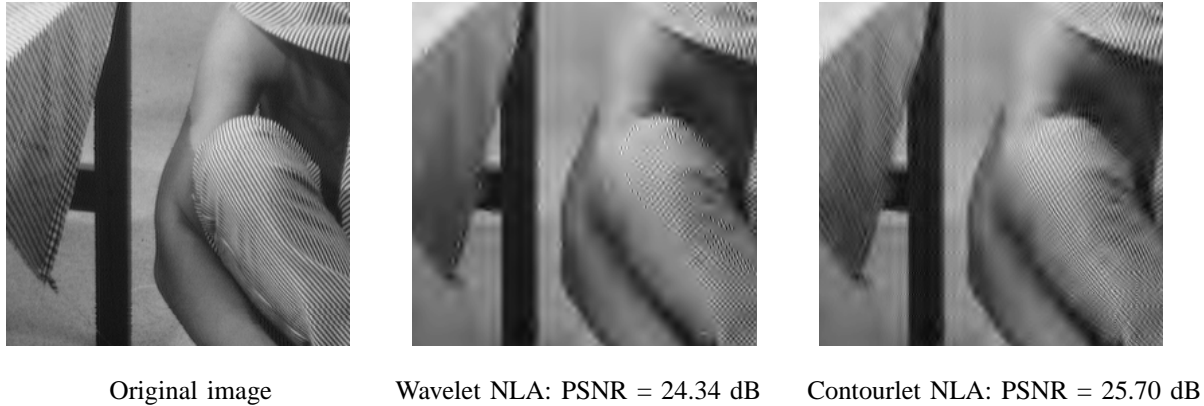


Fig. 17. Nonlinear approximations (NLA) by the wavelet and contourlet transforms. In each case, the original image *Barbara* of size 512×512 is reconstructed from the 4096-most significant coefficients. Only part of images are shown for detail comparison.

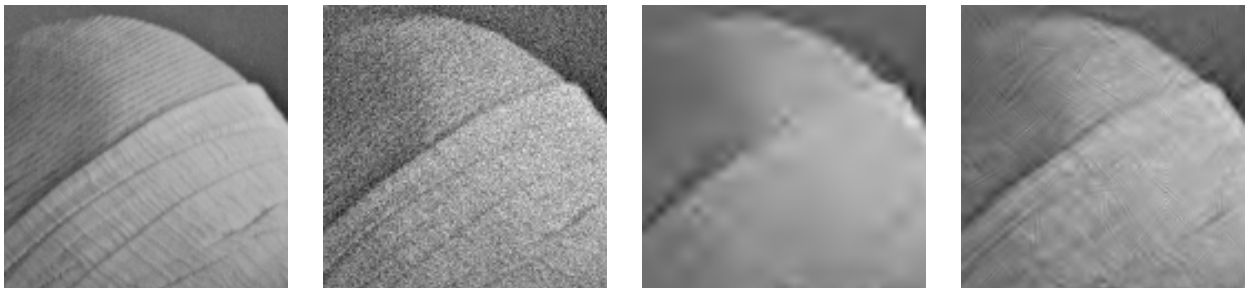


Fig. 18. Denoising experiments. From left to right are: original image, noisy image (SNR = 9.55 dB), denoising using wavelets (SNR = 13.82 dB), and denoising using contourlets (SNR = 15.42 dB).

in several image processing applications. A Matlab contourlet toolbox is available for download from www.ifp.uiuc.edu/~minhdo/software.

REFERENCES

- [1] D. L. Donoho, M. Vetterli, R. A. DeVore, and I. Daubechies, "Data compression and harmonic analysis," *IEEE Trans. Inform. Th.*, vol. 44, no. 6, pp. 2435–2476, October 1998.
- [2] S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed. Academic Press, 1999.
- [3] A. Skodras, C. Christopoulos, and T. Ebrahimi, "The JPEG 2000 still image compression standard," *IEEE Signal Processing Magazine*, vol. 18, pp. 36–58, Sep. 2001.
- [4] E. J. Candès and D. L. Donoho, "Curvelets – a suprisingly effective nonadaptive representation for objects with edges," in *Curve and Surface Fitting*, A. Cohen, C. Rabut, and L. L. Schumaker, Eds. Saint-Malo: Vanderbilt University Press, 1999.
- [5] —, "New tight frames of curvelets and optimal representations of objects with smooth singularities," Department of Statistics, Stanford University, Tech. Rep., 2002, submitted.

- [6] D. H. Hubel and T. N. Wiesel, ‘Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex,’ *Journal of Physiology*, no. 160, pp. 106–154, 1962.
- [7] B. A. Olshausen and D. J. Field, ‘Emergence of simple-cell receptive field properties by learning a sparse code for natural images,’ *Nature*, pp. 607–609, 1996.
- [8] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.
- [9] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [10] R. A. DeVore, ‘Nonlinear approximation,’ *Acta Numer.*, vol. 7, pp. 51–150, 1998.
- [11] E. J. Candès and D. L. Donoho, ‘Ridgelets: a key to higher-dimensional intermittency?’ *Phil. Trans. R. Soc. Lond. A.*, pp. 2495–2509, 1999.
- [12] E. L. Pennec and S. Mallat, ‘Image compression with geometric wavelets,’ in *Proc. IEEE Int. Conf. on Image Proc.*, Vancouver, Canada, Sep. 2000.
- [13] A. Cohen and B. Matei, ‘Compact representation of images by edge adapted multiscale transforms,’ in *Proc. IEEE Int. Conf. on Image Proc., Special Session on Image Processing and Non-Linear Approximation*, Thessaloniki, Greece, Oct. 2001, invited paper.
- [14] D. L. Donoho, ‘Wedgelets: nearly-minimax estimation of edges,’ *Ann. Statist.*, vol. 27, pp. 859–897, 1999.
- [15] M. B. Wakin, J. K. Romberg, H. Choi, and R. G. Baraniuk, ‘Rate-distortion optimized image compression using wedgelets,’ in *Proc. IEEE Int. Conf. on Image Proc.*, Rochester, New York, Oct. 2002.
- [16] P. L. Dragotti and M. Vetterli, ‘Footprints and edgeprints for image denoising and compression,’ in *Proc. IEEE Int. Conf. on Image Proc.*, Thessaloniki, Greece, Oct. 2001.
- [17] R. Shukla, P. L. Dragotti, M. N. Do, and M. Vetterli, ‘Rate-distortion optimized tree structured compression algorithms for piecewise smooth images,’ *IEEE Trans. Image Proc.*, 2002, submitted.
- [18] J. Daugman, ‘Two-dimensional spectral analysis of cortical receptive field profile,’ *Vision Research*, vol. 20, pp. 847–856, 1980.
- [19] A. B. Watson, ‘The cortex transform: Rapid computation of simulated neural images,’ *Computer Vision, Graphics, and Image Processing*, vol. 39, no. 3, pp. 311–327, 1987.
- [20] 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*, vol. 38, no. 2, pp. 587–607, March 1992.
- [21] J. P. Antoine, P. Carrette, R. Murenzi, and B. Plette, ‘Image analysis with two-dimensional continuous wavelet transform,’ *Signal Processing*, vol. 31, pp. 241–272, 1993.
- [22] F. G. Meyer and R. R. Coifman, ‘Brushlets: a tool for directional image analysis and image compression,’ *Journal of Appl. and Comput. Harmonic Analysis*, vol. 5, pp. 147–187, 1997.
- [23] N. Kingsbury, ‘Complex wavelets for shift invariant analysis and filtering of signals,’ *Journal of Appl. and Comput. Harmonic Analysis*, vol. 10, pp. 234–253, 2001.
- [24] P. V. C. Hough, ‘Methods and means for recognizing complex patterns,’ U.S. Patent 3069654, 1962.
- [25] M. N. Do and M. Vetterli, ‘Pyramidal directional filter banks and curvelets,’ in *Proc. IEEE Int. Conf. on Image Proc.*, Thessaloniki, Greece, Oct. 2001.
- [26] P. J. Burt and E. H. Adelson, ‘The Laplacian pyramid as a compact image code,’ *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 532–540, April 1983.

- [27] R. H. Bamberg and M. J. T. Smith, "A filter bank for the directional decomposition of images: Theory and design," *IEEE Trans. Signal Proc.*, vol. 40, no. 4, pp. 882–893, April 1992.
- [28] Y. Lu and M. N. Do, "CRISP-contourlet: a critically sampled directional multiresolution image representation," in *Proc. SPIE Conf. on Wavelets X*, San Diego, Aug. 2003.
- [29] M. N. Do and M. Vetterli, "Framing pyramids," *IEEE Trans. Signal Proc.*, Sep. 2003.
- [30] M. Vetterli, "Multidimensional subband coding: Some theory and algorithms," *Signal Proc.*, vol. 6, no. 2, pp. 97–112, February 1984.
- [31] 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.*, 1999, pp. 1417–1420.
- [32] M. N. Do, "Directional multiresolution image representations," Ph.D. dissertation, Swiss Federal Institute of Technology, Lausanne, Switzerland, December 2001, <http://www.ifp.uiuc.edu/~minhdo/publications>.
- [33] R. R. Coifman, Y. Meyer, and M. V. Wickerhauser, "Wavelet analysis and signal processing," in *Wavelets and their Applications*, M. B. R. et al, Ed. Boston: Jones and Barlett, 1992, pp. 153–178.
- [34] S. Mallat, "Multiresolution approximations and wavelet orthonormal bases of $L_2(R)$," *Trans. Amer. Math. Soc.*, vol. 315, pp. 69–87, September 1989.
- [35] H. Smith, "Wave equations with low regularity coefficients," *Documenta Mathematica*, vol. Extra Volume ICM 1998, II, pp. 723–730, 1998, <http://www.math.washington.edu/~hart/papers.html>.
- [36] M. Vetterli, "Wavelets, approximation and compression," *IEEE SP Mag.*, pp. 59–73, Sep. 2001.
- [37] R. Jia, "Approximation properties of multivariate wavelets," *Mathematics of Computation*, vol. 67, pp. 647–665, 1998.
- [38] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing, Special Issue on Wavelets and Signal Processing*, vol. 41, no. 12, pp. 3445–3462, December 1993.
- [39] A. Cohen, I. Daubechies, and J.-C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Commun. on Pure and Appl. Math.*, vol. 45, pp. 485–560, 1992.
- [40] M. Vetterli and C. Herley, "Wavelets and filter banks: Theory and design," *IEEE Trans. Signal Proc.*, vol. 40, no. 9, pp. 2207–2232, September 1992.
- [41] 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.*, vol. 43, no. 3, pp. 649–665, Mar. 1995.
- [42] D. Donoho and I. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, Dec. 1994.
- [43] D. D.-Y. Po and M. N. Do, "Directional multiscale modeling of images using the contourlet transform," *IEEE Trans. Image Proc.*, 2003, submitted, <http://www.ifp.uiuc.edu/~minhdo/publications>.