

# 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 of one-dimensional transforms, such as the Fourier and wavelet transforms, in capturing the geometry of image edges are well known. In this paper, we pursue a “true” two-dimensional transform that can capture the intrinsic geometrical structure that is key in visual information. The main challenge in exploring geometry in images comes from the discrete nature of the data. Thus, unlike other approaches, such as curvelets, that first develop a transform in the continuous domain and then discretize for sampled data, our approach starts with a discrete-domain construction and then studies its convergence to an expansion in the continuous domain. Specifically, we construct a discrete-domain multiresolution and multidirection 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 via a directional multiresolution analysis framework. We show 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 representation, wavelets, contourlets, filter banks, multiresolution, multidirection, contours, geometric image processing.

## I. INTRODUCTION

Efficient representation of visual information lies at the heart of many image processing tasks, including compression, denoising, feature extraction, and inverse problems. *Efficiency* of a representation refers to the ability to capture significant information about an object of interest using a small description. For image compression or content-based image retrieval, the use of an 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 leads to 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 located along *smooth* curves (i.e. contours) owing to smooth boundaries of physical objects. Thus, natural images contain *intrinsic geometrical structures* that are key features in visual information. 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 new 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.

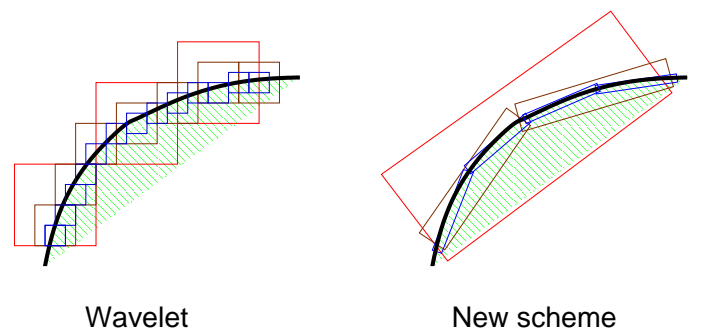


Fig. 1. Wavelet versus new scheme: illustrating the successive refinement by the two systems near a smooth contour, which is shown as a thick curve separating two smooth regions.

Consider the situation when a smooth contour is being

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

M. Vetterli is with the Audiovisual Communications Laboratory, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland, and with the Department of Electrical Engineering and Computer Science, University of California at Berkeley, Berkeley CA 94720 (email: martin.vetterli@epfl.ch).

This work was supported in part by the US National Science Foundation under Grant CCR-0237633 (CAREER) and the Swiss National Science Foundation under Grant 20-63664.00.

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 structure 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.<sup>1</sup> The new style painter, on the other hand, exploits 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*. Furthermore, experiments in searching for the sparse components of natural images produced basis images that closely resemble the aforementioned characteristics of the visual cortex [7]. 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, from coarse to fine resolutions.
- 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 a variety of directions, much more than the 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 require new constructions. Moreover, a major challenge in capturing geometry and directionality in images comes from the *discrete* nature of the data: the input is typically sampled images defined on rectangular grids. For example, directions other than horizontal and vertical look very different on a rectangular grid. Because of pixelization, the notion of smooth contours on sampled images are not obvious. For these reasons, unlike other transforms that were initially developed in the continuous domain and then discretized for sampled data, our approach

starts with a discrete-domain construction and then studies its convergence to an expansion in the continuous domain.

The outline of the rest of the paper is as follows. After reviewing related work in Section II, we propose in Section III a multiresolution and multidirection 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. It is of interest to study the limit behavior when such schemes are iterated over scale and/or direction, which has been analyzed in the connection between filter banks, their iteration, and the associated wavelet construction [8], [2]. Such a connection is studied in Section IV, where we establish a precise link between the proposed filter bank and the associated continuous-domain contourlet expansion in a newly defined *directional multiresolution analysis* framework. The approximation power of the contourlet expansion is studied in Section V. We show that with parabolic scaling and sufficient directional vanishing moments, contourlets achieve the optimal approximation rate 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

Consider a general series expansion by  $\{\phi_n\}_{n=1}^{\infty}$  (e.g. a Fourier or wavelets basis) for a given signal  $f$  as:

$$f = \sum_{n=1}^{\infty} c_n \phi_n. \quad (1)$$

The error decay of the *best  $M$ -term approximation* provides a measurement of the efficiency of an expansion. The best  $M$ -term approximation (also commonly referred to as *nonlinear approximation* [1]) using this expansion is defined as

$$\hat{f}_M = \sum_{n \in I_M} c_n \phi_n, \quad (2)$$

where  $I_M$  is the set of indexes of the  $M$ -largest  $|c_n|$ . The quality of the approximated function  $\hat{f}_M$  relates to how *sparse* the expansion by  $\{\phi_n\}_{n=1}^{\infty}$  is, or how well the expansion compacts the energy of  $f$  into a few coefficients.

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 for 2-D piecewise smooth functions that are  $C^2$  except for discontinuities along  $C^2$  curves. For this class of functions, the best  $M$ -term approximation error (in  $L_2$ -norm square)  $\|f - \hat{f}_M\|_2^2$  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 a significant improvement of a curvelet-like method over wavelets, which is comparable to the improvement of wavelets over the Fourier basis for one-dimensional piecewise smooth signals. 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.

<sup>1</sup>Or we could consider the wavelet-style painter as a *pointillist*!

The curvelet transform was developed initially in the *continuous domain* [4] via multiscale filtering and then applying a block ridgelet transform [9] 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 and correspond to a 2-D frequency partition based on the polar coordinate. This makes the curvelet construction simple in the continuous domain but causes the implementation for discrete images – sampled on a rectangular grid – to be very challenging. In particular, approaching critical sampling seems difficult in such discretized constructions.

The reason for this difficulty, we believe, is because the typical rectangular-sampling grid imposes a *prior* geometry to discrete images; e.g. strong bias toward horizontal and vertical directions. This fact motivates our development of a directional multiresolution transform like curvelets, but *directly in the discrete domain*, which results in the contourlet construction described in this paper. We would like to emphasize that although curvelet and contourlet transforms have some similar properties and goals, the latter is not a discretized version of the former. More comparisons between these two transforms are provided at the end of Section IV.

Apart from curvelets and contourlets, there have recently been several approaches in developing efficient representations of geometrical regularity. Notable examples are bandelets [10], the edge-adapted multiscale transform [11], wedgelets [12], [13], and quadtree coding [14]. These approaches typically require an edge-detection stage, 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. For example, we do not have to rely on edge detection, which is unreliable and noise sensitive. 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 [15], the cortex transform [16], the steerable pyramid [17], 2-D directional wavelets [18], brushlets [19], and complex wavelets [20]. 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 new scheme shown in Figure 1, we see that the improvement of the new scheme can be attributed to the grouping of nearby wavelet coefficients, since they are locally correlated due to the smoothness of the contours. Therefore, we can obtain a sparse expansion

for natural images by first applying a multiscale transform, followed by 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, the latter step is similar to the popular Hough transform [21] for line detection in computer vision.

With this insight, we proposed a *double filter bank* structure (see Figure 7) [22] for obtaining sparse expansions for typical images having smooth contours. In this double filter bank, the Laplacian pyramid [23] is first used to capture the point discontinuities, and then followed by a directional filter bank [24] 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*. 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 new 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 stages 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 [25], 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.

#### B. Pyramid frames

One way to obtain a multiscale decomposition is to use the Laplacian pyramid (LP) introduced by Burt and Adelson [23]. The LP decomposition at each level generates a downsampled lowpass version of the original and the difference between the original and the prediction, resulting in a bandpass image. Figure 2(a) depicts this decomposition process, where  $H$  and  $G$  are called (lowpass) analysis and synthesis filters, respectively, and  $M$  is the sampling matrix. The process can be iterated on the coarse (downsampled lowpass) signal. Note that in multidimensional filter banks, sampling is represented by sampling matrices; for example, downsampling  $x[n]$  by  $M$  yields  $x_d[n] = x[Mn]$ , where  $M$  is an integer matrix [8].

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), and this image 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 [26], we studied the LP using the theory of frames and oversampled filter banks. We showed that the LP with orthogonal filters (that is, the analysis and synthesis filters are time reversal,  $h[n] = g[-n]$ , and  $g[n]$  is orthogonal

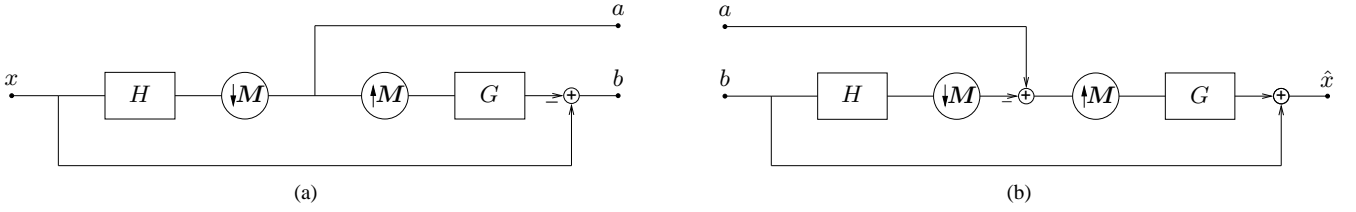


Fig. 2. Laplacian pyramid. (a) One level of decomposition. The outputs are a coarse approximation  $a[n]$  and a difference  $b[n]$  between the original signal and the prediction. (b) The new reconstruction scheme for the Laplacian pyramid [26].

to its translates with respect to the sampling lattice by  $M$ ) 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 [26] to achieve significant improvement over the usual reconstruction in the presence of noise.

### C. Iterated directional filter banks

Bamberger and Smith [24] constructed a 2-D directional filter bank (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 [24] involves modulating the input image and using quincunx filter banks with diamond-shaped filters [27]. To obtain the desired frequency partition, a complicated tree expanding rule has to be followed for finer directional subbands (e.g., see [28] for details).

In [29], 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 is a two-channel quincunx filter bank [27] with fan filters (see Figure 4) that divides a 2-D spectrum into two directions: horizontal and vertical. The second building block of the DFB is a *shearing* operator, which amounts to just reordering of image samples. Figure 5 shows an application of a shearing operator where a  $-45^\circ$  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 partition while maintaining perfect reconstruction. Thus, the key in the DFB is to use an appropriate combination of shearing operators together with two-direction partition 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 [29] (Chapter 3).

Using multirate identities [8], 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 (directional) synthesis filters as  $D_k^{(l)}$ ,  $0 \leq k < 2^l$ , which correspond to the subbands indexed as in Figure 3(a). The corresponding overall

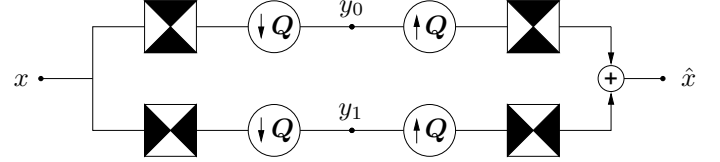


Fig. 4. Two-dimensional spectrum partition using quincunx filter banks with fan filters. The black regions represent the ideal frequency supports of each filter.  $Q$  is a quincunx sampling matrix.

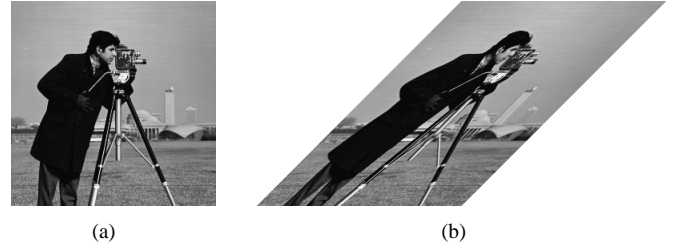


Fig. 5. Example of shearing operation that is used like a rotation operation for DFB decomposition. (a) The “cameraman” image. (b) The “cameraman” image after a shearing operation.

sampling matrices were shown [29] to have the following diagonal forms

$$S_k^{(l)} = \begin{cases} \text{diag}(2^{l-1}, 2) & \text{for } 0 \leq k < 2^{l-1}, \\ \text{diag}(2, 2^{l-1}) & \text{for } 2^{l-1} \leq k < 2^l, \end{cases} \quad (3)$$

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

$$\left\{ d_k^{(l)}[n - S_k^{(l)} m] \right\}_{0 \leq k < 2^l, m \in \mathbb{Z}^2}, \quad (4)$$

obtained by translating the impulse responses of the equivalent synthesis filters  $D_k^{(l)}$  over the sampling lattices by  $S_k^{(l)}$ , 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 (4) resembles a local Radon transform and are called *Radonlets*. Furthermore, it can be shown [29] that if the building block filter bank in Figure 4 uses orthogonal filters, then the resulting DFB is orthogonal and (4) becomes an orthogonal basis.

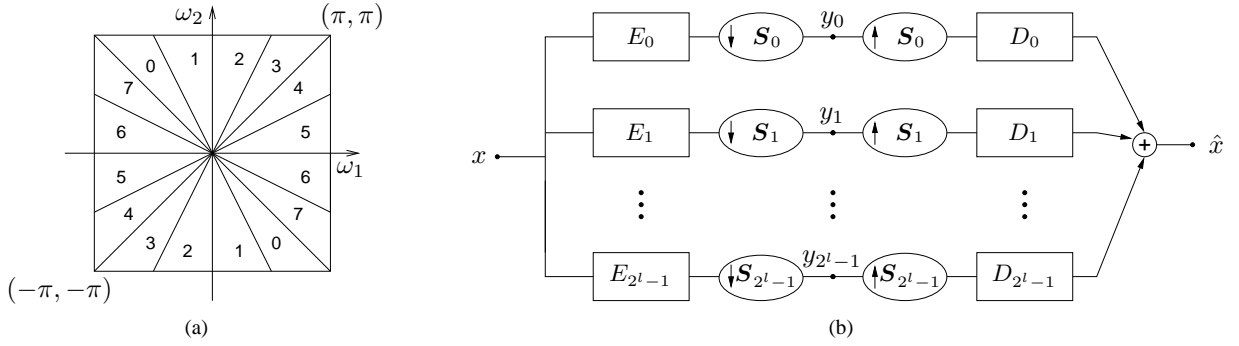


Fig. 3. Directional filter bank. (a) Frequency partitioning where  $l = 3$  and there are  $2^3 = 8$  real wedge-shaped frequency bands. Subbands 0–3 correspond to the *mostly horizontal* directions, while subbands 4–7 correspond to the *mostly vertical* directions. (b) The multichannel view of an  $l$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

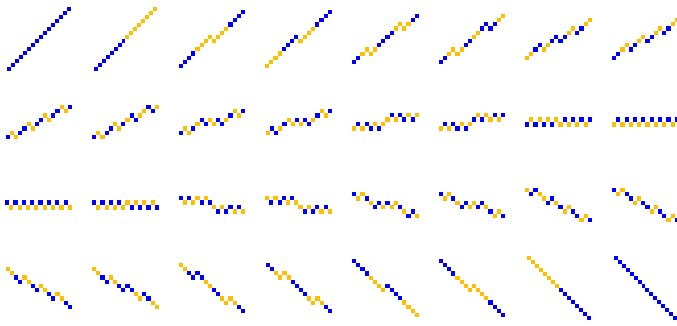


Fig. 6. Impulse responses of 32 equivalent filters for the first half channels, corresponding to the mostly horizontal directions, of a 6-levels DFB that uses the Haar filters. Black and gray squares correspond to  $+1$  and  $-1$ , respectively. Because the basis functions resemble “local lines”, we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

#### D. Multiscale and directional decomposition: the discrete contourlet transform

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.

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 *contourlet filter bank*, which decomposes images into directional subbands at multiple scales.

Specifically, let  $a_0[\mathbf{n}]$  be the input image. The output after the LP stage is  $J$  bandpass images  $b_j[\mathbf{n}]$ ,  $j = 1, 2, \dots, J$  (in the fine-to-coarse order) and a lowpass image  $a_J[\mathbf{n}]$ . That means, the  $j$ -th level of the LP decomposes the image

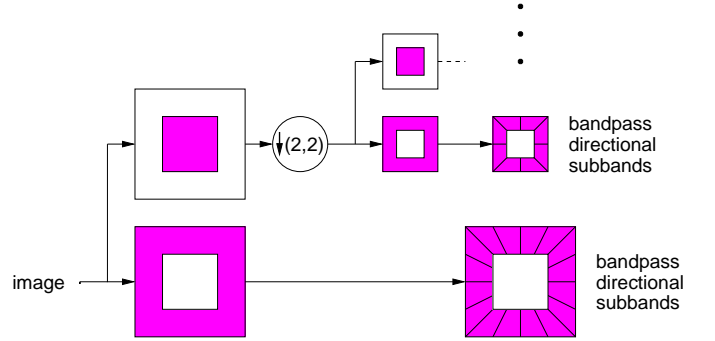


Fig. 7. The contourlet filter bank: first, a multiscale decomposition into octave bands by the Laplacian pyramid is computed, and then a directional filter bank is applied to each bandpass channel.

$a_{j-1}[\mathbf{n}]$  into a coarser image  $a_j[\mathbf{n}]$  and a detail image  $b_j[\mathbf{n}]$ . Each bandpass image  $b_j[\mathbf{n}]$  is further decomposed by an  $l_j$ -level DFB into  $2^{l_j}$  bandpass directional images  $c_{j,k}^{(l_j)}[\mathbf{n}]$ ,  $k = 0, 1, \dots, 2^{l_j} - 1$ . The main properties of the discrete contourlet transform are stated in the following theorem.

**Theorem 1:** In a contourlet filter bank, the following hold:

- 1) If both the LP and the DFB use perfect-reconstruction filters, then the discrete contourlet transform achieves perfect reconstruction, which means it provides a frame operator.
- 2) If both the LP and the DFB use orthogonal filters, then the discrete contourlet transform provides a tight frame with frame bounds equal to 1.
- 3) The discrete contourlet transform has a redundancy ratio that is less than  $4/3$ .
- 4) Suppose an  $l_j$ -level DFB is applied at the pyramidal level  $j$  of the LP, then the basis images of the discrete contourlet transform (i.e. the equivalent filters of the contourlet filter bank) have an essential support size of *width*  $\approx C2^j$  and *length*  $\approx C2^{j+l_j-2}$ .
- 5) Using FIR filters, the computational complexity of the discrete contourlet transform is  $O(N)$  for  $N$ -pixel images.

*Proof:*

- 1) This is obvious as the discrete contourlet transform is a composition of perfect-reconstruction blocks.
- 2) With orthogonal filters, the LP is a tight frame with

frame bounds equal to 1 [26], which means it preserves the  $l_2$ -norm, or  $\|a_0\|_2^2 = \sum_{j=1}^J \|b_j\|_2^2 + \|a_J\|_2^2$ . Similarly, with orthogonal filters the DFB is an orthogonal transform [29], which means  $\|b_j\|_2^2 = \sum_{k=0}^{2^{l_j}-1} \|c_{j,k}^{(l_j)}\|_2^2$ . Combining these two stages, the discrete contourlet transform satisfies the norm preserving or tight frame condition.

- 3) Since the DFB is critically sampled, the redundancy of the discrete contourlet transform 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  $C_1 2^j \times C_1 2^j$ , followed by downsampling by  $2^{j-1}$  in each dimension. For the DFB, from (3) we see that after  $l_j$  levels ( $l_j \geq 2$ ) of tree-structured decomposition, the equivalent directional filters have support of width about  $C_2 2^j$  and length about  $C_2 2^{l_j-1}$  (also see Figure 6). Combining these two stages, again using multirate identities, into equivalent contourlet filter bank channels, we see that contourlet basis images have support of width about  $C 2^j$  and length about  $C 2^{j+l_j-2}$ .
- 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 (without loss of generality we can suppose that 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.<sup>2</sup> Thus, for an  $N$ -pixel image, the complexity of the LP stage in the contourlet filter bank is:

$$\sum_{j=1}^J N \left(\frac{1}{4}\right)^{j-1} \left(\frac{L_p}{2} + 1\right) < \frac{4}{3} N \left(\frac{L_p}{2} + 1\right) \text{ (operations).} \quad (5)$$

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 stage for an  $N$ -pixel image is:

$$\sum_{j=1}^J N \left(\frac{1}{4}\right)^{j-1} L_d l_j < \frac{4}{3} N L_d \max \{l_j\} \text{ (operations).} \quad (6)$$

Combining (5) and (6) we obtain the desired result. ■

Since the multiscale and directional decomposition stages are decoupled in the discrete contourlet transform, we can have a different number of directions at different scales, thus offering a flexible multiscale and directional expansion.

<sup>2</sup>Here we assume all filters are implemented non-separably. For certain filters, separable filtering (maybe in polyphase domain) is possible and requires lower complexity.

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 the wavelet transform [30]. 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. In addition, from Theorem 1, part 4, we see that by altering the depth of the DFB 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.

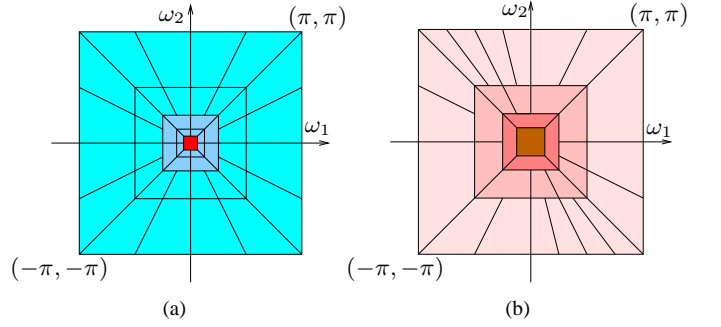


Fig. 8. Examples of possible frequency decompositions by the contourlet transform and contourlet packets.

#### IV. CONTOURLETS AND DIRECTIONAL MULTIREOLUTION ANALYSIS

As for the wavelet filter bank, the contourlet filter bank has an associated continuous-domain expansion in  $L_2(\mathbb{R}^2)$  using the *contourlet functions*. In this section, the connection between the discrete contourlet transform and the continuous-domain contourlet expansion will be made precisely via a new multiresolution analysis framework that is similar to the link between wavelets and filter banks [2]. The new elements in this framework are multidirection and its 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 contourlet filter bank 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 synthesis filter  $G$  in the iterated LP uniquely defines a unique scaling function  $\phi(\mathbf{t}) \in L_2(\mathbb{R}^2)$  that satisfies the following two-scale equation [8], [2]

$$\phi(\mathbf{t}) = 2 \sum_{\mathbf{n} \in \mathbb{Z}^2} g[\mathbf{n}] \phi(2\mathbf{t} - \mathbf{n}). \quad (7)$$



Let

$$\phi_{j,n} = 2^{-j} \phi\left(\frac{t - 2^j n}{2^j}\right), \quad j \in \mathbb{Z}, n \in \mathbb{Z}^2. \quad (8)$$

Then the family  $\{\phi_{j,n}\}_{n \in \mathbb{Z}^2}$  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  $V_{j-1}$ , or

$$V_{j-1} = V_j \oplus W_j. \quad (9)$$

In [26] we show that the LP can be considered as an oversampled filter bank where each polyphase component of the difference image  $b[n]$  in Figure 2, together with the coarse image  $a[n]$ , comes from a separate filter bank channel with the same sampling matrix  $\text{diag}(2, 2)$ . Let  $F_i(z)$ ,  $0 \leq i \leq 3$  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  $\psi^{(i)}(t)$  where

$$\psi^{(i)}(t) = 2 \sum_{n \in \mathbb{Z}^2} f_i[n] \phi(2t - n). \quad (10)$$

*Proposition 1 ([26]):* Using  $\psi^{(i)}(t)$  in (10), let

$$\psi_{j,n}^{(i)}(t) = 2^{-j} \psi^{(i)}\left(\frac{t - 2^j n}{2^j}\right), \quad j \in \mathbb{Z}, n \in \mathbb{Z}^2. \quad (11)$$

Then, for scale  $2^j$ ,  $\{\psi_{j,n}^{(i)}\}_{0 \leq i \leq 3, n \in \mathbb{Z}^2}$  is a tight frame for  $W_j$ . For all scales,  $\{\psi_{j,n}^{(i)}\}_{j \in \mathbb{Z}, 0 \leq i \leq 3, n \in \mathbb{Z}^2}$  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

$$\mu_{j,2n+k_i}(t) = \psi_{j,n}^{(i)}(t), \quad 0 \leq i \leq 3, \quad (12)$$

where  $k_i$  are the coset representatives for downsampling by 2 in each dimension

$$\mathbf{k}_0 = (0, 0)^T, \mathbf{k}_1 = (1, 0)^T, \mathbf{k}_2 = (0, 1)^T, \mathbf{k}_3 = (1, 1)^T. \quad (13)$$

With this notation, the family  $\{\mu_{j,n}\}_{n \in \mathbb{Z}^2}$  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 contourlet filter bank, the discrete basis (4) 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 DFBs in the contourlet filter bank use orthogonal filters. Although in the contourlet transform, 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:* Let

$$\rho_{j,k,n}^{(l)}(t) = \sum_{m \in \mathbb{Z}^2} d_k^{(l)}[m - S_k^{(l)} n] \phi_{j,m}(t), \quad (14)$$

for arbitrary but finite<sup>3</sup>  $l$ . Then the family  $\{\rho_{j,k,n}^{(l)}\}_{n \in \mathbb{Z}^2}$  is an orthonormal basis of a directional subspace  $V_{j,k}^{(l)}$  for each  $k = 0, \dots, 2^l - 1$ . Furthermore,

$$V_{j,k}^{(l)} = V_{j,2k}^{(l+1)} \oplus V_{j,2k+1}^{(l+1)}, \quad (15)$$

$$V_{j,k}^{(l)} \perp V_{j,k'}^{(l)} \quad \text{for } k \neq k', \quad \text{and} \quad (16)$$

$$V_j = \bigoplus_{k=0}^{2^l-1} V_{j,k}^{(l)}. \quad (17)$$

*Proof:* This result can be proved by induction on the number of decomposition levels  $l$  of the DFB, in much the same way as for the wavelet packets bases [30] (see also [2]). We only sketch the idea of the proof here. Suppose that  $\{\rho_{j,k,n}^{(l)}\}_{n \in \mathbb{Z}^2}$  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  $d_k^{(l)}$  that leads to two channels with equivalent filters  $d_{2k}^{(l+1)}$  and  $d_{2k+1}^{(l+1)}$ . This transforms the orthonormal basis  $\{\rho_{j,k,n}^{(l)}\}_{n \in \mathbb{Z}^2}$  into two orthonormal families  $\{\rho_{j,2k,n}^{(l+1)}\}_{n \in \mathbb{Z}^2}$  and  $\{\rho_{j,2k+1,n}^{(l+1)}\}_{n \in \mathbb{Z}^2}$ . Each of these families generate a subspace with finer directional resolution that satisfies the “two-direction” equation (15). With this induction, starting from an orthonormal basis  $\{\phi_{j,n}\}_{n \in \mathbb{Z}^2}$  of  $V_j$ , we obtain orthonormal bases for all directional subspaces  $V_{j,k}^{(l)}$ , hence (16) and (17). ■

### C. Multiscale and multidirection: the contourlet expansion

Applying the directional decomposition by the family (4) onto the detail subspace  $W_j$  as done by the contourlet transform, we obtain a similar result.

*Proposition 3:* Let

$$\lambda_{j,k,n}^{(l)}(t) = \sum_{m \in \mathbb{Z}^2} d_k^{(l)}[m - S_k^{(l)} n] \mu_{j,m}(t). \quad (18)$$

The family  $\{\lambda_{j,k,n}^{(l)}\}_{n \in \mathbb{Z}^2}$  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) illustrates the detail directional subspaces  $W_{j,k}^{(l)}$  in the frequency domain. The indexes  $j$ ,  $k$ , and  $n$  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  $l_j$ .

<sup>3</sup>The situation when the number of levels  $l$  of the iterated DFB goes to infinity involves a regularity study for the DFB, which will be treated elsewhere.

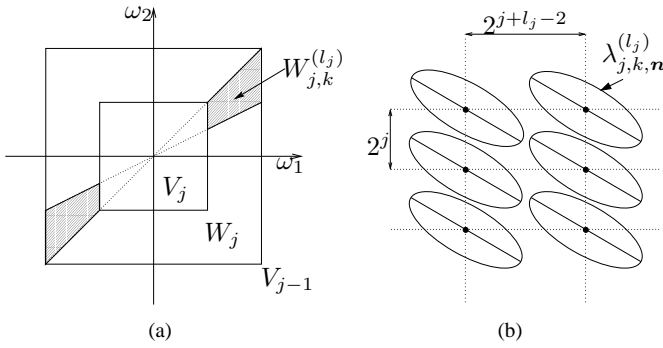


Fig. 9. Contourlet subspaces. (a) Multiscale and multidirection subspaces generated by the contourlet transform which is illustrated on a 2-D spectrum decomposition. (b) Sampling grid and approximate support of contourlet functions for a “mostly horizontal” subspace  $W_{j,k}^{(l_j)}$ . For “mostly vertical”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

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

$$\lambda_{j,k}^{(l)}(\mathbf{t}) = \sum_{\mathbf{m} \in \mathbb{Z}^2} d_k^{(l)}[\mathbf{m}] \mu_{j,\mathbf{m}}(\mathbf{t}). \quad (19)$$

Then for  $l \geq 2$ ,

$$\lambda_{j,k,\mathbf{n}}^{(l)}(\mathbf{t}) = \lambda_{j,k}^{(l)}(\mathbf{t} - 2^{j-1} \mathbf{S}_k^{(l)} \mathbf{n}). \quad (20)$$

*Proof:* The definition of  $\psi_{j,\mathbf{n}}$  in (11) implies that  $\psi_{j,\mathbf{m}+\mathbf{n}}(\mathbf{t}) = \psi_{j,\mathbf{m}}(\mathbf{t} - 2^j \mathbf{n})$ . Applying this to (12) we have  $\mu_{j,\mathbf{m}+2\mathbf{n}}(\mathbf{t}) = \mu_{j,\mathbf{m}}(\mathbf{t} - 2^{j-1} 2\mathbf{n})$ . In other words,  $\mu_{j,\mathbf{m}}$  are periodically shift-invariant with even shifts in  $\mathbf{m}$ . From (3) we see that when  $l \geq 2$ , sampling by  $\mathbf{S}_k^{(l)}$  is also even in each dimension. Thus, from (18) with a change of variable we obtain

$$\begin{aligned} \lambda_{j,k,\mathbf{n}}^{(l)}(\mathbf{t}) &= \sum_{\mathbf{m} \in \mathbb{Z}^2} d_k^{(l)}[\mathbf{m}] \mu_{j,\mathbf{m}+\mathbf{S}_k^{(l)} \mathbf{n}}(\mathbf{t}) \\ &= \sum_{\mathbf{m} \in \mathbb{Z}^2} d_k^{(l)}[\mathbf{m}] \mu_{j,\mathbf{m}}(\mathbf{t} - 2^{j-1} \mathbf{S}_k^{(l)} \mathbf{n}) \\ &= \lambda_{j,k}^{(l)}(\mathbf{t} - 2^{j-1} \mathbf{S}_k^{(l)} \mathbf{n}). \end{aligned}$$

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

$$\left\{ \lambda_{j,k,\mathbf{n}}^{(l)}(\mathbf{t}) = \lambda_{j,k}^{(l)}(\mathbf{t} - 2^{j-1} \mathbf{S}_k^{(l)} \mathbf{n}) \right\}_{\mathbf{n} \in \mathbb{Z}^2} \quad (21)$$

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  $\{\lambda_{j,k,\mathbf{n}}^{(l)}\}_{\mathbf{n} \in \mathbb{Z}^2}$  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 (10) into (12) and then into (19), we can write  $\lambda_{j,k}^{(l)}(\mathbf{t})$  directly as a linear combination of the scaling functions

as

$$\begin{aligned} \lambda_{j,k}^{(l)}(\mathbf{t}) &= \sum_{i=0}^3 \sum_{\mathbf{n} \in \mathbb{Z}^2} d_k^{(l)}[2\mathbf{n} + \mathbf{k}_i] \left( \sum_{\mathbf{m} \in \mathbb{Z}^2} f_i[\mathbf{m}] \phi_{j-1,2\mathbf{n}+\mathbf{m}} \right) \\ &= \sum_{\mathbf{m} \in \mathbb{Z}^2} \underbrace{\left( \sum_{i=0}^3 \sum_{\mathbf{n} \in \mathbb{Z}^2} d_k^{(l)}[2\mathbf{n} + \mathbf{k}_i] f_i[\mathbf{m} - 2\mathbf{n}] \right)}_{w_k^{(l)}[\mathbf{m}]} \phi_{j-1,\mathbf{m}}(\mathbf{t}). \end{aligned} \quad (22)$$

The discrete filter  $w_k^{(l)}$  is roughly equal to the summation of convolutions between the directional filter  $d_k^{(l)}$  and bandpass filters  $f_i$ 's, and thus it is a bandpass directional filter. It can be verified that with orthogonal filters in both the LP and the DFB, for all  $l \geq 2$ ,  $j \in \mathbb{Z}$ ,  $0 \leq k < 2^l$ ,  $\mathbf{n} \in \mathbb{Z}^2$

$$\|\lambda_{j,k,\mathbf{n}}^{(l)}\|_{L_2}^2 = \|w_k^{(l)}\|_{L_2}^2 = 3/4. \quad (23)$$

Integrating the multidirection analysis over scales we obtain 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

$$\{\phi_{j_0,\mathbf{n}}(\mathbf{t}), \lambda_{j,k,\mathbf{n}}^{(l_j)}(\mathbf{t})\}_{j \leq j_0, 0 \leq k \leq 2^{l_j}-1, \mathbf{n} \in \mathbb{Z}^2} \quad (24)$$

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

$$\{\lambda_{j,k,\mathbf{n}}^{(l_j)}(\mathbf{t})\}_{j \in \mathbb{Z}, 0 \leq k \leq 2^{l_j}-1, \mathbf{n} \in \mathbb{Z}^2} \quad (25)$$

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:

$$\begin{aligned} L_2(\mathbb{R}^2) &= V_{j_0} \oplus \left( \bigoplus_{j \leq j_0} W_j \right), \quad \text{and} \\ L_2(\mathbb{R}^2) &= \bigoplus_{j \in \mathbb{Z}} W_j. \end{aligned}$$

Finally, similar to the link between wavelets and filter banks [2], the following theorem precisely connects the continuous-domain expansions by contourlet functions in (24) and (25) with the discrete contourlet transform in Section III-D.

*Theorem 3:* Suppose  $a_0[\mathbf{n}] = \langle f, \phi_{L,\mathbf{n}} \rangle$  are  $L_2(\mathbb{R}^2)$  inner products of a function  $f(\mathbf{t}) \in L_2(\mathbb{R}^2)$  with the scaling functions at a scale  $L$ . Furthermore, suppose the image  $a_0[\mathbf{n}]$  is decomposed by the discrete contourlet transform into coefficients  $\{a_J[\mathbf{n}], c_{j,k}^{(l_j)}[\mathbf{n}]\}$ ,  $j = 1, 2, \dots, J$  and  $0 \leq k < 2^{l_j} - 1$ . Then

$$a_J[\mathbf{n}] = \langle f, \phi_{L+J,\mathbf{n}} \rangle, \quad \text{and} \quad c_{j,k}^{(l_j)}[\mathbf{n}] = \langle f, \lambda_{L+j,k,\mathbf{n}}^{(l_j)} \rangle. \quad (26)$$

*Proof:* Suppose that  $a_{j-1}[\mathbf{n}] = \langle f, \phi_{L+j-1,\mathbf{n}} \rangle$  and  $a_{j-1}[\mathbf{n}]$  is decomposed by the LP into the coarser image  $a_j[\mathbf{n}]$



and the detail image  $b_j[\mathbf{n}]$ . Then using (7), (10) and (12), it is easy to verify that

$$a_j[\mathbf{n}] = \langle f, \phi_{L+j, \mathbf{n}} \rangle, \quad \text{and} \quad b_j[\mathbf{n}] = \langle f, \mu_{L+j, \mathbf{n}} \rangle.$$

Subsequently, using (19), the output of an  $l_j$ -level DFB given the input image  $b_j[\mathbf{n}]$  can be written as

$$c_{j,k}^{(l_j)}[\mathbf{n}] = \langle f, \lambda_{L+j,k,n}^{(l_j)} \rangle.$$

■

The contourlet transform has several distinguishing features that are important to emphasize.

- 1) The contourlet expansions are defined on rectangular grids, and thus offer a seamless translation (as demonstrated in Theorem 3) to the discrete world, where image pixels are sampled on a rectangular grid. To achieve this “digital-friendly” feature, the contourlet kernel functions  $\lambda_{j,k}^{(l_j)}$  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 systems in [4], [5].
- 2) As a result of being defined on rectangular grids, contourlets have 2-D frequency partition on centric squares (see Figure 8), rather than on centric circles for curvelets [4], [5] and other systems defined on polar coordinates.
- 3) Since the contourlet functions are defined via iterated filter banks like wavelets, the contourlet transform has fast filter bank algorithms and convenient tree structures.
- 4) It is easy to see that with FIR filters, the iterated contourlet filter bank leads to compactly supported contourlet frames. More precisely, the contourlet function  $\lambda_{j,k,n}^{(l_j)}$  has support of size  $width \approx C2^j$  and  $length \approx C2^{j+l_j-2}$ . In other words, at each scale and direction, the set  $\{\lambda_{j,k,n}^{(l_j)}\}_{n \in \mathbb{Z}^2}$  “tiles” the plane  $\mathbb{R}^2$  (see Figure 9(b)).
- 5) The contourlet construction provides a space-domain multiresolution scheme that offers flexible refinements for the spatial resolution and the angular resolution (see remarks at the end of Section III-D).

## V. CONTOURLET APPROXIMATION AND COMPRESSION

The proposed contourlet filter bank and its associated continuous-domain frames in 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 on contourlets to obtain a sparse 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 another paper [31].

### 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 [32].

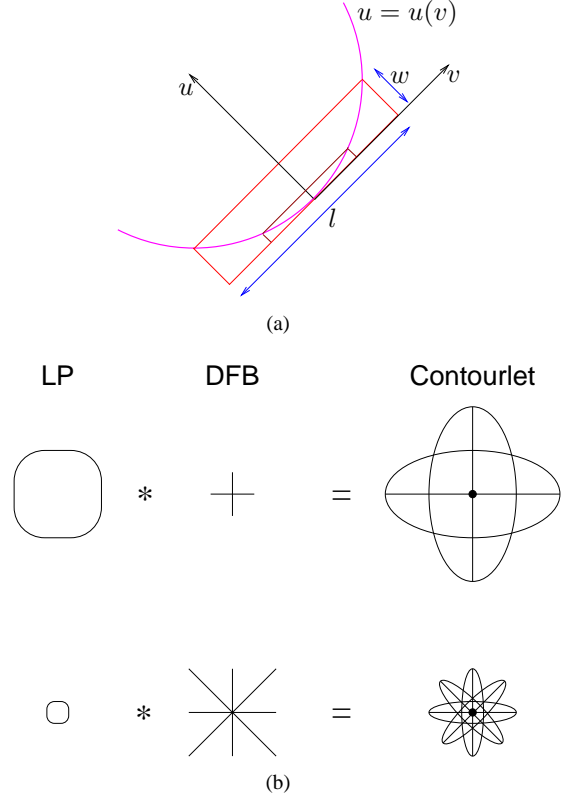


Fig. 10. Parabolic scaling relation for curves. (a) The rectangular supports of the basis functions that fit a curve exhibit the quadratic relation:  $width \propto length^2$ . (b) Illustrating the evolution of the support sizes of contourlet functions that satisfy the parabolic scaling.

The motivation behind the parabolic scaling is to efficiently approximate a smooth discontinuity curve by “laying” basis elements with elongated supports along the curve (refer to the new scheme in Figure 1). Suppose that the discontinuity curve is sufficiently smooth – specifically a  $C^2$  curve, 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(a), we can readily verify that the parametric representation of the discontinuity curve obeys

$$u(v) \approx \frac{\kappa}{2}v^2, \quad \text{when } v \approx 0, \quad (27)$$

where  $\kappa$  is the local curvature of the curve. Hence, to fit the  $C^2$  discontinuity curve at *fine scales* the width  $w$  and the length  $l$  of the basis functions have to satisfy

$$w \approx \frac{\kappa}{8}l^2. \quad (28)$$

For the contourlet frame in (24), we know that when an  $l_j$ -level DFB is applied to the pyramidal scale  $2^j$ , the contourlet functions have support size of  $width \approx C2^j$  and  $length \approx C2^{l_j+j-2}$ . Hence to make the contourlet expansion satisfy the parabolic scaling, we simply have to impose that *the number of directions is doubled at every other finer scale*. An example of such a frequency decomposition is shown in Figure 8(a).

More precisely, suppose that at a 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 is

$$l_j = l_{j_0} + \lfloor (j_0 - j)/2 \rfloor, \quad \text{for } j \leq j_0. \quad (29)$$

Figure 10(b) 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

For the wavelet case in 1-D, the wavelet approximation theory brought a novel 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 *highpass filter* in the wavelet filter bank must have enough “zeros at  $\omega = 0$ ”. The vanishing-moments property is the key for the sparse expansion of piecewise smooth signals by wavelets [2]. Intuitively, wavelets with vanishing moments are orthogonal to polynomial signals, and thus only a few wavelet basis functions around the discontinuities points would “feel” these discontinuities and lead to significant coefficients [33].

In the contourlet case, our target for approximation is piecewise smooth images with smooth contours. The key feature of these images is that *image edges are localized in both location and direction*. More specifically, a local region around a smooth contour can be approximated by two polynomial surfaces separated by a straight line. Thus, it is desirable that only few contourlet functions whose supports intersect with a contour *and* align with the contour local direction would “feel” this discontinuity. One 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  $\lambda(\mathbf{t})$  is said to have an  $L$ -order directional vanishing moment in the direction  $\mathbf{u} = (u_1, u_2)^T$  if all 1-D slices of that function along the direction  $\mathbf{u}$ ,  $\lambda_{\mathbf{u}, r_2}(r_1) = \lambda(r_1 \mathbf{u} + r_2 \mathbf{u}^\perp)$  where  $\mathbf{u}^\perp = (-u_2, u_1)^T$ , have  $L$  vanishing moments:

$$\int_{-\infty}^{\infty} \lambda_{\mathbf{u}, r_2}(r_1) r_1^p dr_1 = 0, \quad \text{for all } r_2 \in \mathbb{R}, p = 0, 1, \dots, L-1. \quad (30)$$

It can be shown that in the Fourier domain, the DVM condition (30) is equivalent to requiring  $\Lambda(\omega_1, \omega_2)$  and its  $L-1$  first derivatives in the direction  $\mathbf{u}$  to be zero along the line  $u_1 \omega_1 + u_2 \omega_2 = 0$  (see Figure 11). We refer to this equivalent DVM condition in the Fourier domain as having “ $L$ -order zeros along the line  $u_1 \omega_1 + u_2 \omega_2 = 0$ .”

For a contourlet function  $\lambda_{j,k}^{(l)}(\mathbf{t})$  constructed from an iterated filter bank as in (22), it has an  $L$ -order DVM along direction  $\mathbf{u}$  if the discrete-time Fourier transform  $W_k^{(l)}(e^{j\omega_1}, e^{j\omega_2})$

of the associated filter  $w_k^{(l)}[\mathbf{n}]$  also has  $L$ -order zeros along the line  $u_1 \omega_1 + u_2 \omega_2 = 0$ . This provides a condition for designing the contourlet filter bank. In an extreme case, contourlet functions with ideal frequency response (i.e. with sinc-type filters) have DVMs on every direction  $\mathbf{u}$  such that the line  $u_1 \omega_1 + u_2 \omega_2 = 0$  is outside their ideal angular frequency supports (see Figure 9(a) and Figure 11). For FIR filter  $w_k^{(l)}[\mathbf{n}]$  and when  $u_1$  and  $u_2$  are integers, the DVM condition is satisfied if the  $z$ -transform  $W_k^{(l)}(z_1, z_2)$  can be factorized as

$$W_k^{(l)}(z_1, z_2) = (1 - z_1^{u_1} z_2^{u_2})^L R(z_1, z_2). \quad (31)$$

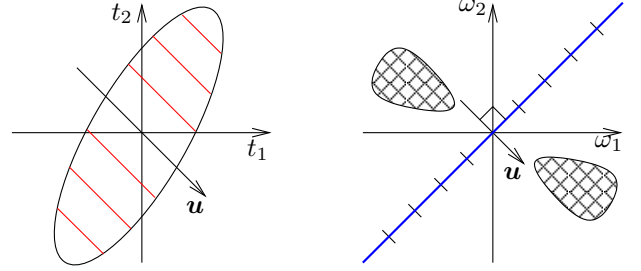


Fig. 11. Illustrating the directional vanishing moment condition in the space and Fourier domains. In the space domain (left), all the shown slices of  $\lambda(t_1, t_2)$  have vanishing moments. In the frequency domain (right),  $\Lambda(\omega_1, \omega_2)$  is “flat” along the line  $u_1 \omega_1 + u_2 \omega_2 = 0$ . The hatched regions represent the essential frequency support of  $\Lambda(\omega_1, \omega_2)$ .

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 [9], 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. Contourlet approximation

In this subsection we will show that a contourlet expansion that satisfies the parabolic scaling and has sufficient DVMs (this will be defined precisely in Lemma 1) achieves the optimal nonlinear approximation rate for 2-D piecewise  $C^2$  smooth functions with discontinuities along  $C^2$  smooth curves. As we will be interested in the asymptotic rates, not in the constants, we use the notation  $a_n \lesssim b_n$  when there exists a constant  $C$  such that  $a_n \leq C b_n$ . If  $a_n \lesssim b_n$  and  $b_n \lesssim a_n$ , then we write  $a_n \sim b_n$ .

First, notice that since the contourlet expansion is a frame, the approximation error by keeping only  $M$  coefficients as in (2) obeys

$$\|f - \hat{f}_M\|_2^2 \lesssim \sum_{n \notin I_M} |c_n|^2. \quad (32)$$

We consider compactly supported contourlets, which are obtained from iterated contourlet filter banks with FIR filters. Recall that for a contourlet frame (24) to satisfy the parabolic scaling,  $l_j$  has to follow (29). For simplicity, we set  $j_0 = 0$  and  $l_0 = 2$ .<sup>4</sup> This generates a contourlet frame that at scale  $2^j$

<sup>4</sup>Other values of  $l_0$  only changes the constant but not the exponent in the approximation rate.

( $j < 0$ ) has  $\sim 2^{-j/2}$  directions and each contourlet function  $\lambda_{j,k,n}$  has support size of *width*  $\sim 2^j$  and *length*  $\sim 2^{j/2}$ . For convenience, in this section the support size of a contourlet, *width* and *length*, is measured along the short and long dimensions of the contourlet itself (see Figure 12) instead of the horizontal/vertical spacing as shown in Figure 9(b). Based on Figure 9(b), we see that these two measures are related by a ratio between 1 and  $\sqrt{2}$ , and thus this change of measures does not affect our asymptotic analysis. Using (18) it can be verified that (also note that the support size of  $\lambda_{j,k,n}$  is  $\sim 2^j \times 2^{j/2}$ )

$$\|\lambda_{j,k,n}\|_\infty \sim 2^{-3j/4}. \quad (33)$$

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  $\mathcal{S}$ . We classify contourlets into type 1, whose support intersects with  $\mathcal{S}$ ; and type 2, whose support does not intersect with  $\mathcal{S}$ .

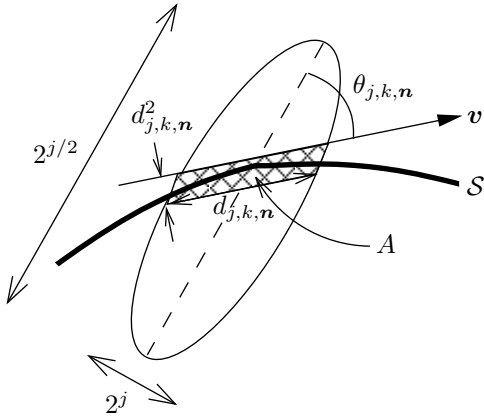


Fig. 12. Interaction between a contourlet (denoted by the ellipse) and a discontinuity curve (denoted by the thick curve).

For type 1 contourlets, ideally we would like only contourlets that “align” with the discontinuity curve  $\mathcal{S}$  to produce significant coefficients. Therefore the key issue is to characterize the decay of type 1 coefficients as contourlets “turn” away from the direction of  $\mathcal{S}$ . Let  $\theta_{j,k,n}$  be the angle between the principal direction of a type 1 contourlet  $\lambda_{j,k,n}$  and the local tangent direction  $\mathbf{v}$  of  $\mathcal{S}$  where the contourlet intersects (see Figure 12). At scale  $2^j$ , since contourlets span  $\sim 2^{-j/2}$  directions almost uniformly between 0 and  $\pi$ , within a square block of size  $2^{j/2}$  we can re-index these directions by  $\tilde{k}$ , where  $1 \leq \tilde{k} \lesssim 2^{-j/2}$ , so that  $\theta_{j,\tilde{k},n} \sim \tilde{k}2^{j/2}$ .

Let  $d_{j,\tilde{k},n}$  be the length of  $\mathcal{S}$  that intersects with the support of  $\lambda_{j,\tilde{k},n}$ . From Figure 12 we see that

$$d_{j,\tilde{k},n} \sim 2^j / \sin \theta_{j,\tilde{k},n} \sim 2^j / (\tilde{k}2^{j/2}) = 2^{j/2} \tilde{k}^{-1}. \quad (34)$$

Let  $A$  denotes the hatched region in Figure 12, which is the support region of  $\lambda_{j,\tilde{k},n}$  that contains the intersection with  $\mathcal{S}$  and is “sandwiched” between two lines parallel to  $\mathbf{v}$ . Because  $\mathcal{S}$  is  $C^2$ , using the Taylor expansion as in (27) (also see Figure 10(a)), the width of  $A$  is  $\sim d_{j,\tilde{k},n}^2$ . Thus, using (33) and (34), we can bound the integral of  $\langle f, \lambda_{j,\tilde{k},n} \rangle$  inside the

region  $A$  by

$$\begin{aligned} \left| \int_A f \lambda_{j,\tilde{k},n} \right| &\lesssim \|\lambda_{j,\tilde{k},n}\|_\infty \cdot \text{area}(A) \\ &\sim \|\lambda_{j,\tilde{k},n}\|_\infty \cdot d_{j,\tilde{k},n}^3 \\ &\sim 2^{3j/4} \tilde{k}^{-3}. \end{aligned} \quad (35)$$

The bound in (35) turns out to be critical for the desired approximation rate. Therefore we need to bound the integral of  $\langle f, \lambda_{j,\tilde{k},n} \rangle$  outside region  $A$  to be the same order. If we “remove” the region  $A$ , then we can consider the discontinuity curve as a straight line of direction  $\mathbf{v}$  that intersects the support of  $\lambda_{j,\tilde{k},n}$  with a length  $d_{j,\tilde{k},n}$ . The following lemma, which is proved in the appendix, provides a sufficient condition for the contourlet function  $\lambda_{j,\tilde{k},n}$  to obtain the desired decay rate.

**Lemma 1:** For  $p \geq 1$ , suppose the scaling function  $\phi \in C^p$  and the contourlet function  $\lambda_{j,\tilde{k}}$  has a  $(p+2)$ -order DVM along a direction  $\mathbf{u}$  that makes an angle  $\alpha$  with the discontinuity direction  $\mathbf{v}$ , where

$$\alpha \lesssim 2^{j/2} \tilde{k}^{(p-1)/(p+1)}. \quad (36)$$

Then for every  $f \in C^2$  and  $r \in \mathbb{R}$ ,

$$\left| \int_{t \in \mathbb{R}^2, t \cdot \mathbf{v} \leq r} f(t) \lambda_{j,\tilde{k},n}(t) dt \right| \lesssim \|\lambda_{j,\tilde{k},n}\|_\infty \cdot d_{j,\tilde{k},n}^3. \quad (37)$$

**Remark 1:** The condition in Lemma 1 requires contourlets to have DVMs on denser sets of directions at finer scales. In particular, for  $p = 1$ , it requires that each contourlet function  $\lambda_{j,k}$  has third-order DVMs on a set of directions with maximum angular gap  $\sim 2^{j/2}$ , which is the same order as the essential angular support of  $\lambda_{j,k}$  in the frequency domain. For larger  $p$ , this condition is relaxed at the cost of requiring higher-order DVMs. This condition is satisfied with the ideal frequency response contourlets as discussed at the end of Section V-B. At the moment, it remains as a conjecture that such a condition can be obtained at the asymptote with compactly supported contourlets. However, in practice, we only process up to a finite scale and Lemma 1 guides us to construct contourlets to have DVMs on as many directions as possible, especially around the directions of discontinuity curves. This strategy was taken in [31], which leads to good approximation performance. An alternative strategy is to design contourlets close to having the ideal frequency response, and thus make directional moments approximately vanish.

With Lemma 1, combined with (35), we have the following decay of type 1 contourlet coefficients indexed by  $j$  and  $\tilde{k}$ :

$$|\langle f, \lambda_{j,\tilde{k},n} \rangle| \lesssim 2^{3j/4} \tilde{k}^{-3}. \quad (38)$$

In addition, since the discontinuity curve  $\mathcal{S}$  has finite length, the number of type 1 coefficients with these indexes is

$$m_{j,\tilde{k}} \sim 1/d_{j,\tilde{k},n} \sim 2^{-j/2} \tilde{k}. \quad (39)$$

From (38), for a type 1 coefficient to have magnitude above a threshold  $\epsilon$ , it is necessary that  $\tilde{k} \lesssim 2^{j/4} \epsilon^{-1/3}$  and  $j \gtrsim \log \epsilon$ . Thus, the number of type 1 coefficients with magnitude above  $\epsilon$  is

$$m(\epsilon) \lesssim \sum_{j=\log \epsilon}^0 \sum_{\tilde{k}=1}^{2^{j/4} \epsilon^{-1/3}} m_{j,\tilde{k}} \sim \epsilon^{-2/3} \log(\epsilon^{-1}). \quad (40)$$

Using the last bound, we can write  $\epsilon$  as a function of the number of coefficients  $m$  as

$$\epsilon(m) \lesssim m^{-3/2}(\log m)^{3/2}.$$

Therefore the sum square error due to discarding all except  $M$  largest type 1 coefficients satisfies

$$E_1(M) \leq \sum_{m>M} |\epsilon(m)|^2 \lesssim M^{-2}(\log M)^3. \quad (41)$$

Next, consider type 2 contourlets, whose support is included in a region where  $f$  is  $C^2$ . For these contourlets, the corresponding coefficients  $\langle f, \lambda_{j,k,n} \rangle$  behave as if  $f \in C^2$ . Suppose that the scaling function  $\phi$  has accuracy of order 2, which is equivalent to requiring the filter  $G(e^{j\omega_1}, e^{j\omega_2})$  in (7) to have a second-order zero at  $(\pi, \pi)$  [34], that is for all  $p_1, p_2 \in \mathbb{Z}$ ;  $0 \leq p_1 + p_2 < 2$ ,

$$\partial_{\omega_1}^{p_1} \partial_{\omega_2}^{p_2} G(e^{j\omega_1}, e^{j\omega_2})|_{(\pi, \pi)} = 0. \quad (42)$$

Then for  $f \in C^2$ , we have

$$\|f - P_{V_j} f\|_2 \sim (2^j)^2. \quad (43)$$

Thus, the sum squared error due to discarding all except  $M \sim 2^{-2j}$  type 2 contourlet coefficients down to scale  $2^j$  satisfies

$$E_2(M) \sim (2^j)^4 \sim M^{-2}. \quad (44)$$

Combining (41) and (44), and using (32), we obtain the following result that characterizes the approximation power of contourlets.

*Theorem 4:* Suppose that a compactly supported contourlet frame (24) satisfies the parabolic scaling condition (29), the contourlet functions  $\lambda_{j,k}$  satisfy the condition in Lemma 1, and the scaling function  $\phi \in C^p$  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

$$\|f - \hat{f}_M^{(\text{contourlet})}\|_2^2 \leq C(\log M)^3 M^{-2}. \quad (45)$$

*Remark 2:* The approximation rate for contourlets in (45) is the same as the approximation rate for curvelets, which was derived in [5] and [35]. 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, which is a  $C^2$  curve, no other approximation scheme can achieve a better rate than  $M^{-2}$  [1]. In this sense, the contourlet expansion achieves the optimal approximation rate for piecewise smooth functions with  $C^2$  contours.

#### D. Contourlet compression

So far, we consider the approximation problem of contourlets by keeping the  $M$  largest coefficients. For the compression problem, we have to account for the cost to encode quantized coefficients, as well as the cost to index the retained coefficients. Fortunately, as can be seen from the last subsection, the  $M$  retained contourlet coefficients are well organized in tree structures. Specifically, from coarse to fine scales,

significant contourlet coefficients are successively localized in both location (contourlets intersect with the discontinuity curve) and direction (intersected contourlets with direction close to the local direction of the discontinuity curve). Thus, using embedded tree structures for contourlet coefficients that are similar to the embedded zero-trees for wavelets [36], we can efficiently index the retained coefficients using 1 bit per coefficient.

Suppose that contourlet coefficients are uniformly quantized with step size  $\Delta = 2^{-L}$  and coefficients with magnitude below  $\Delta$  are discarded. Instead of using fixed length coding for the quantized coefficients, a slight gain (in the log factor, but not the exponent of the rate-distortion function) can be obtained by variable length coding. In particular, we use the bit plane coding scheme [8] where coefficients with magnitude in the range  $(2^{l-1-L}, 2^{l-L}]$  are encoded with  $l$  bits. Because of (40), the number of coefficients in this range is  $\lesssim (L-l)2^{2(L-l)/3}$ . In addition, the total number of retained coefficients is  $M \lesssim L2^{2L/3}$ . Thus, the number of bits  $R$  needed for compression, which is equal to the sum of indexing and quantization costs, is bounded by

$$R \lesssim M + \sum_{l \geq 1} l(L-l)2^{2(L-l)/3} \lesssim L2^{2L/3}. \quad (46)$$

The compression distortion  $D$  is equal to the sum of the truncation distortion in (45) and the quantization distortion,

$$D \lesssim M^{-2}(\log M)^3 + M\Delta^2 \sim L2^{-4L/3}. \quad (47)$$

Combining (46) and (47) we obtain the following result for the operational rate-distortion function  $D(R)$  by contourlets.

*Corollary 1:* Under the assumption of Theorem 4, a contourlet compression system that uses embedded zero-trees and bit plane coding achieves

$$D(R) \leq C(\log R)^3 R^{-2}. \quad (48)$$

## VI. NUMERICAL EXPERIMENTS

All experiments in this section use a wavelet transform with “9-7” biorthogonal filters [37], [38] and 6 decomposition levels. For the contourlet transform, in the LP stage we also use the “9-7” filters. We choose “9-7” biorthogonal filters because they have been shown to provide the best results for images, partly because they are linear phase and are close to being orthogonal. In the DFB stage we use the “23-45” biorthogonal quincunx filters designed by Phoong et al. [39] and modulate them to obtain the biorthogonal fan filters. Apart from also 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. Figure 13 shows examples of the contourlet transform. We notice that only contourlets that match with both location and direction of image contours produce significant coefficients.

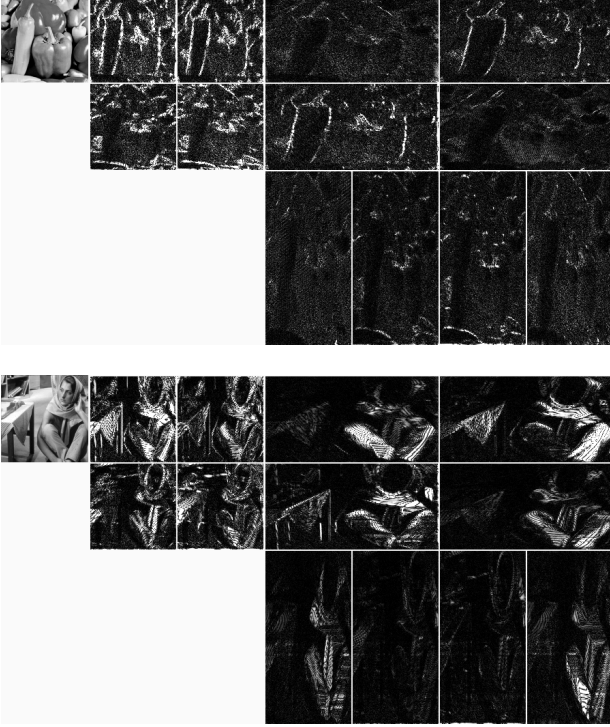


Fig. 13. Examples of the contourlet transform on the *Peppers* and *Barbara* images. For clear visualization, each image is only decomposed into two pyramidal levels, which are then decomposed into four and eight directional subbands. Small coefficients are shown in black while large coefficients are shown in white.

#### A. Wavelets vs. Contourlets

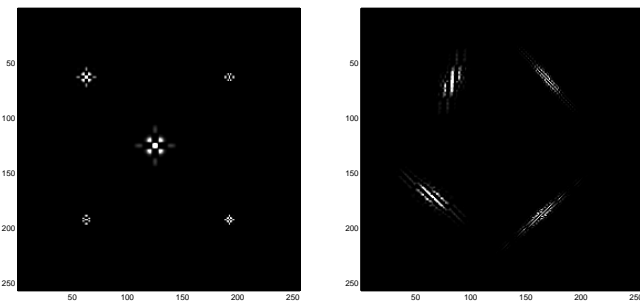


Fig. 14. Comparing a few actual 2-D wavelets (5 on the left) and contourlets (4 on the right).

To highlight the difference between the wavelet and contourlet transform, Figure 14 shows a few wavelet and contourlet basis images. We see that contourlets offer a much richer set of directions and shapes, and thus they are more effective in capturing smooth contours and geometric structures in images.

#### B. Nonlinear approximation

Next we compare the nonlinear approximation (NLA) performances of the wavelet and 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 15 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 new scheme shown in Figure 1.

Figure 16 shows a detailed comparison of two nonlinear approximated images by the wavelet and contourlet transforms using the same number of coefficients on the *Barbara* image. Contourlets are shown to be superior compared to wavelets in capturing fine contours (e.g., directional textures on cloths). In addition, there is a significant gain of 1.46 dB in peak signal-to-noise ratio (PSNR) for 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. As an example, for image denoising, random noise will generate significant wavelet coefficients just like true edges, but is less likely to generate significant contourlet coefficients. Consequently, a simple thresholding scheme [40] applied on the contourlet transform is more effective in removing the noise than it is for the wavelet transform.

Figure 17 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 smooth contours, both visually as well as in PSNR. A more sophisticated denoising scheme that takes into account the dependencies across scales, directions and locations in the contourlet domain using statistical modeling of contourlet coefficients is presented in [41] and shows further improvements.

### VII. CONCLUSION

In this work, we constructed a discrete transform that provides 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 observation, we developed a new filter bank structure, the contourlet filter bank, that can provide a flexible multiscale and directional decomposition for images. The developed discrete filter bank has a precise connection with the associated

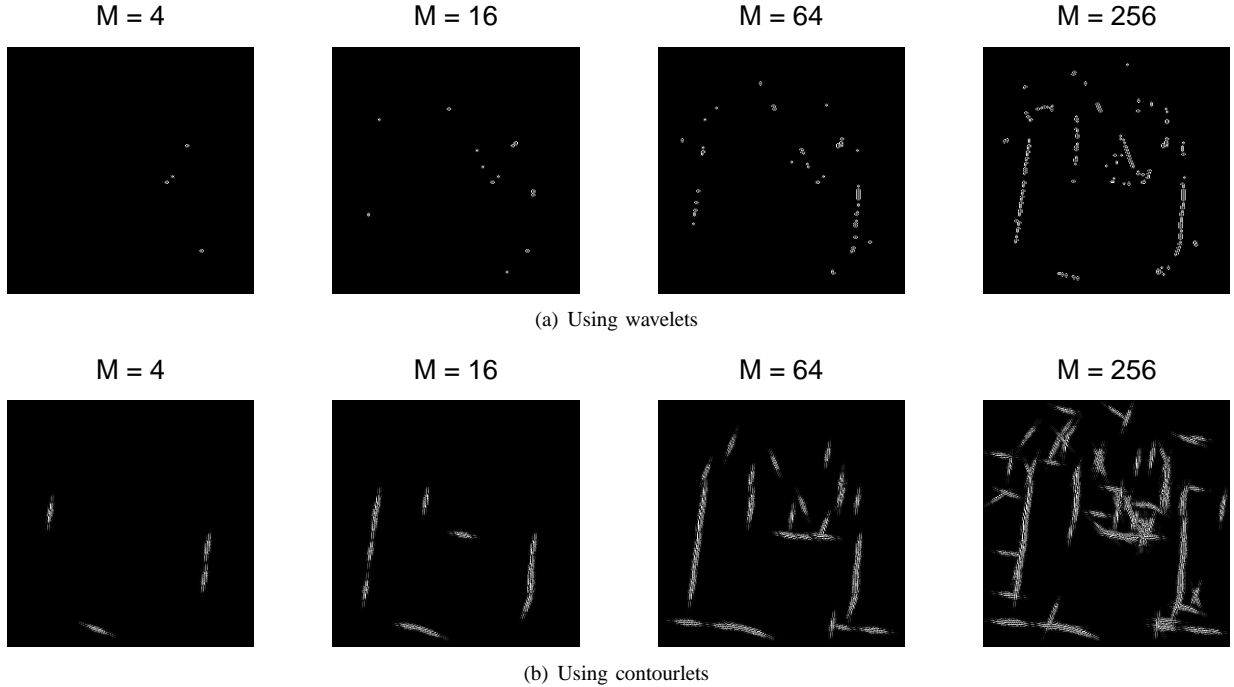


Fig. 15. Sequence of images showing the nonlinear approximations of the *Peppers* image using  $M$  most significant coefficients at the finest detailed subspace  $W_j$ , which is shared by both the wavelet and contourlet transforms.



Fig. 16. Nonlinear approximations (NLA) by the wavelet and contourlet transforms. In each case, the original image *Barbara* of size  $512 \times 512$  is reconstructed from the 4096-most significant coefficients. Only part of images are shown for detail comparison.

continuous-domain contourlet expansion. This connection is defined via a 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 approximation rate for piecewise  $C^2$  smooth images with  $C^2$  smooth contours. Experiments with real images indicate the potential of contourlets in several image processing applications. A Matlab contourlet toolbox is freely available for download from the Matlab Central ([www.mathworks.com/matlabcentral](http://www.mathworks.com/matlabcentral)).

**Acknowledgments.** The authors would like to thank Emmanuel Candès, Laurent Demanet, and Hart Smith for the fruitful discussions.

## APPENDIX

### Proof of Lemma 1:

For convenience, we make a change to a new coordinate  $(x, y)$  as shown in Figure 18, where  $\lambda_{j,\tilde{k}}$  has vanishing moments along the  $x$  direction. We drop the location index  $n$  as it is irrelevant here. Notice that since  $\alpha \lesssim 2^{j/2} \tilde{k}^{(p-1)/(p+1)} \leq 2^{j/2} \tilde{k}$ , the length of the intersection of the  $x$ -axis with the support of  $\lambda_{j,\tilde{k}}$  is  $\sim d_{j,\tilde{k}}$ . Also, for the same order, we can parameterize the discontinuity line as  $y = \alpha x$ .

The support of  $\lambda_{j,\tilde{k}}$  below the line  $y = \alpha x$  can be divided into two regions  $B$  and  $C$  as shown in Figure 18. In region  $B$ , for a fix  $y$ , locally around the support of  $\lambda_{j,\tilde{k}}$  we can use the Taylor expansion of  $f$  along the  $x$ -direction as (we write  $\partial_x^n f$  for the  $n$ -order partial derivative of  $f$  with respect to  $x$ )  $f(x, y) = f(x_0, y) + \partial_x f(x_0, y)(x - x_0) + \partial_x^2 f(\xi, y)(x - x_0)^2/2$ . Substitute this into  $\int_{-\infty}^{\infty} f(x, y) \lambda_{j,\tilde{k}}(x, y) dx$ , then because of

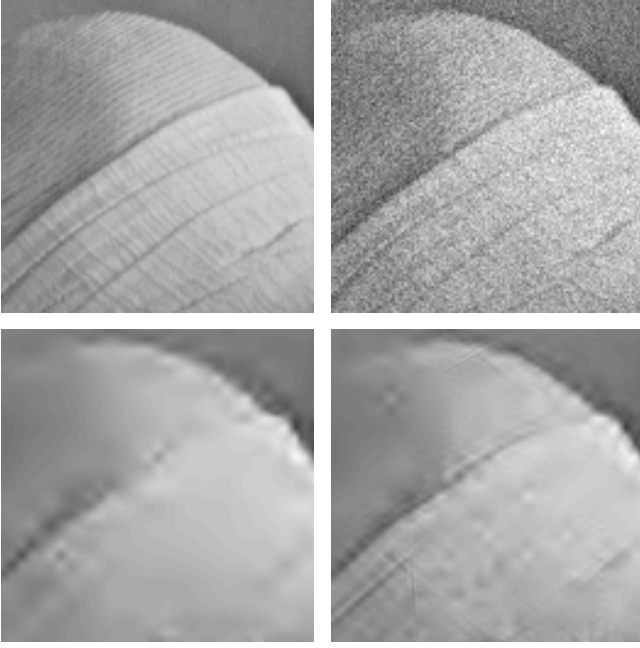


Fig. 17. Denoising experiments. From left to right, top to bottom are: original image, noisy image (PSNR = 24.42 dB), denoising using wavelets (PSNR = 29.41 dB), and denoising using contourlets (PSNR = 30.47 dB).

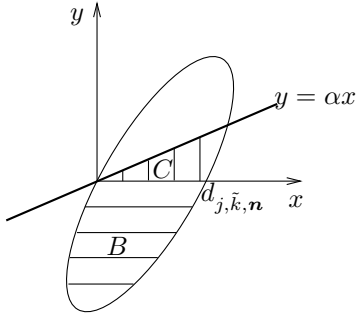


Fig. 18. Interaction between of a contourlet with a line discontinuity.

the first two vanishing moments of  $\lambda_{j,\tilde{k}}$  along the  $x$  direction, the first two terms are canceled. With only the third term left, because of the support size of  $\lambda_{j,\tilde{k}}$ , we get

$$\left| \int_B f \lambda_{j,\tilde{k}} \right| \lesssim \|\lambda_{j,\tilde{k}}\|_\infty \cdot \int_0^{2^{j/2}} \int_0^{d_{j,\tilde{k}}} x^2 dx dy \lesssim \|\lambda_{j,\tilde{k}}\|_\infty \cdot d_{j,\tilde{k}}^3. \quad (49)$$

In the region  $C$ , expand  $f$  using the Taylor expansion as  $f(x, y) = a + bx + cy + R_2$ , where  $R_2$  contains second order terms  $x^2$ ,  $xy$ , and  $y^2$ . Then similar to (49), because of the support size of  $C$  we have  $|\int_C R_2 \lambda_{j,\tilde{k}}| \lesssim \|\lambda_{j,\tilde{k}}\|_\infty \cdot d_{j,\tilde{k}}^3$ . Thus, it remains to check  $\int_C (a + bx + cy) \lambda_{j,\tilde{k}}$ .

Since  $\lambda_{j,\tilde{k}}$  is a linear combination of scaling functions  $\phi_{j-1,n}$  at scale  $2^{j-1}$  as defined in (22); i.e.

$$\lambda_{j,k}(\mathbf{t}) = \sum_{\mathbf{m} \in \mathbb{Z}^2} w_k[\mathbf{m}] 2^{-j+1} \phi(2^{-j+1} \mathbf{t} - \mathbf{m}),$$

given  $\phi \in C^p$  we also have  $\lambda_{j,\tilde{k}} \in C^p$ . Moreover, by taking the derivatives of the last equation we get

$$\|\partial_y^p \lambda_{j,\tilde{k}}\|_\infty \lesssim \|\lambda_{j,\tilde{k}}\|_\infty \cdot (2^{-j})^p. \quad (50)$$

Now expand  $\lambda_{j,\tilde{k}}$  using the Taylor expansion along the  $y$  direction as

$$\lambda_{j,\tilde{k}}(x, y) = \lambda_{j,\tilde{k}}(x, 0) + \dots + \partial_y^{p-1} \lambda_{j,\tilde{k}}(x, 0) y^{p-1} / (p-1)! + \partial_y^p \lambda_{j,\tilde{k}}(x, \xi) y^p / p!. \quad (51)$$

For  $0 \leq q \leq p-1$ , because  $\lambda_{j,\tilde{k}}$  has  $(p+2)$  vanishing moments along the  $x$  direction, we have

$$\begin{aligned} & \int_C (a + bx + cy) \partial_y^q \lambda_{j,\tilde{k}}(x, 0) y^q \\ &= \int_0^{d_{j,\tilde{k}}} \partial_y^q \lambda_{j,\tilde{k}}(x, 0) \int_0^{\alpha x} (a + bx + cy) y^q dy dx \\ &= \int_{-\infty}^{\infty} \partial_y^q \lambda_{j,\tilde{k}}(x, 0) (c_1 x^{q+1} + c_2 x^{q+2}) dx \\ &= \partial_y^q \left( \int_{-\infty}^{\infty} \lambda_{j,\tilde{k}}(x, y) (c_1 x^{q+1} + c_2 x^{q+2}) dx \right) \Big|_{y=0} \\ &= 0. \end{aligned}$$

Therefore, only the last term in (51) remains, and thus

$$\begin{aligned} & \left| \int_C (a + bx + cy) \lambda_{j,\tilde{k}} \right| \\ & \lesssim \|\partial_y^p \lambda_{j,\tilde{k}}\|_\infty \cdot \int_0^{d_{j,\tilde{k}}} \int_0^{\alpha x} (1 + x + y) y^p dy dx \\ & \sim \|\partial_y^p \lambda_{j,\tilde{k}}\|_\infty \cdot \alpha^{p+1} d_{j,\tilde{k}}^{p+2}. \end{aligned}$$

The last expression  $\lesssim \|\lambda_{j,\tilde{k}}\|_\infty \cdot d_{j,\tilde{k}}^3$  because of (50), (34), and the assumption (36). This completes the proof.

## 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 surprisingly 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 piecewise  $C^2$  singularities," *Commun. on Pure and Appl. Math.*, pp. 219–266, Feb. 2004.
- [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] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*. Prentice-Hall, 1995.
- [9] E. J. Candès and D. L. Donoho, "Ridgelets: a key to higher-dimensional intermittency?" *Phil. Trans. R. Soc. Lond. A.*, pp. 2495–2509, 1999.
- [10] E. L. Pennec and S. Mallat, "Sparse geometric image representation with bandelets," *IEEE Trans. Image Proc.*, vol. 14, pp. 423–438, Apr. 2005.
- [11] 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.
- [12] D. L. Donoho, "Wedgelets: nearly-minimax estimation of edges," *Ann. Statist.*, vol. 27, pp. 859–897, 1999.



- [13] 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.
- [14] 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.*, vol. 14, pp. 343–359, Mar. 2005.
- [15] J. Daugman, "Two-dimensional spectral analysis of cortical receptive field profile," *Vision Research*, vol. 20, pp. 847–856, 1980.
- [16] 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.
- [17] 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.
- [18] J. P. Antoine, P. Carrette, R. Murenzi, and B. Piette, "Image analysis with two-dimensional continuous wavelet transform," *Signal Processing*, vol. 31, pp. 241–272, 1993.
- [19] 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.
- [20] 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.
- [21] P. V. C. Hough, "Methods and means for recognizing complex patterns," U.S. Patent 3069654, 1962.
- [22] M. N. Do and M. Vetterli, "Pyramidal directional filter banks and curvelets," in *Proc. IEEE Int. Conf. on Image Proc.*, Thessaloniki, Greece, Oct. 2001.
- [23] 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.
- [24] R. H. Bamberger 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.
- [25] Y. Lu and M. N. Do, "CRISP-contourlet: a critically sampled directional multiresolution image representation," in *Proc. SPIE Conf. on Wavelet Applications in Signal and Image Processing*, San Diego, Aug. 2003.
- [26] M. N. Do and M. Vetterli, "Framing pyramids," *IEEE Trans. Signal Proc.*, pp. 2329–2342, Sep. 2003.
- [27] M. Vetterli, "Multidimensional subband coding: Some theory and algorithms," *Signal Proc.*, vol. 6, no. 2, pp. 97–112, February 1984.
- [28] S.-I. Park, M. J. T. Smith, and R. M. Mersereau, "Improved structures of maximally decimated directional filter banks for spatial image analysis," *IEEE Trans. Image Proc.*, vol. 13, pp. 1424–1431, Nov. 2004.
- [29] 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>.
- [30] 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.
- [31] A. L. Cunha and M. N. Do, "Bi-orthogonal filter banks with directional vanishing moments," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc.*, Philadelphia, Mar. 2005.
- [32] H. Smith, "Wave equations with low regularity coefficients," *Documenta Mathematica*, vol. Extra Volume ICM 1998, II, pp. 723–730, 1998.
- [33] M. Vetterli, "Wavelets, approximation and compression," *IEEE Signal Proc. Mag.*, pp. 59–73, Sep. 2001.
- [34] R. Jia, "Approximation properties of multivariate wavelets," *Mathematics of Computation*, vol. 67, pp. 647–665, 1998.
- [35] L. Demanet, "Second-generation curvelets," Master's thesis, Université Catholique de Louvain, Belgium, 2002.
- [36] 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.
- [37] 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.
- [38] 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.
- [39] 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.
- [40] D. Donoho and I. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, Dec. 1994.
- [41] D. D.-Y. Po and M. N. Do, "Directional multiscale modeling of images using the contourlet transform," *IEEE Trans. Image Proc.*, to appear, <http://www.ifp.uiuc.edu/~minhdo/publications>.



**Minh N. Do** was born in Thanh Hoa, Vietnam, in 1974. He received the B.Eng. degree in computer engineering from the University of Canberra, Australia, in 1997, and the Dr.Sci. degree in communication systems from the Swiss Federal Institute of Technology Lausanne (EPFL), Switzerland, in 2001.

Since 2002, he has been an Assistant Professor with the Department of Electrical and Computer Engineering and a Research Assistant Professor with the Coordinated Science Laboratory and the Beckman Institute, University of Illinois at Urbana-Champaign. His research interests include wavelets, image and multidimensional signal processing, multiscale geometric analysis, and visual information representation.

He received a Silver Medal from the 32nd International Mathematical Olympiad in 1991, a University Medal from the University of Canberra in 1997, the best doctoral thesis award from the Swiss Federal Institute of Technology Lausanne in 2001, and a CAREER award from the National Science Foundation in 2003.



**Martin Vetterli** received the Dipl. El.-Ing. degree from ETH Zürich (ETHZ), Switzerland, in 1981, the MS degree from Stanford University in 1982, and the Doctorat ès Sciences degree from EPF Lausanne (EPFL), Switzerland, in 1986.

He was a Research Assistant at Stanford and EPFL, and has worked for Siemens and AT&T Bell Laboratories. In 1986, he joined Columbia University in New York, where he was last an Associate Professor of Electrical Engineering and co-director of the Image and Advanced Television Laboratory.

In 1993, he joined the University of California at Berkeley, where he was a Professor in the Department of Electrical Engineering and Computer Sciences until 1997, and now holds an Adjunct Professor position. Since 1995, he is a Professor of Communication Systems at EPF Lausanne, Switzerland, where he chaired the Communications Systems Division (1996/97), and heads the Audiovisual Communications Laboratory. Since 2001, he directs the National Competence Center in Research on mobile information and communication systems. He is also a Vice-President (International Affairs) at EPFL. He has held visiting positions at ETHZ (1990) and Stanford (1998).

He is a fellow of the IEEE, a member of SIAM, and was the Area Editor for Speech, Image, Video, and Signal Processing of the IEEE Transactions on Communications. He is also on the editorial boards of *Annals of Telecommunications*, *Applied and Computational Harmonic Analysis* and *The Journal of Fourier Analysis and Application*.

He received the Best Paper Award of EURASIP in 1984 for his paper on multidimensional subband coding, the Research Prize of the Brown Boverly Corporation (Switzerland) in 1986 for his doctoral thesis, the IEEE Signal Processing Society's Senior Awards in 1991 and in 1996 (for papers with D. LeGall and K. Ramchandran, respectively). He won the Swiss National Latsis Prize in 1996, the SPIE Presidential award in 1999, and the IEEE Signal Processing Technical Achievement Award in 2001. He was a member of the Swiss Council on Science and Technology until Dec. 2003.

He was a plenary speaker at various conferences and is the co-author, with J. Kovacevic, of the book *Wavelets and Subband Coding* (Prentice-Hall, 1995). He has published about 85 journal papers on a variety of topics in signal/image processing and communications and holds 7 patents.

His research interests include sampling, wavelets, multirate signal processing, computational complexity, signal processing for communications, digital video processing and joint source/channel coding.