IEEE International Conference on Image Processing (ICIP), Thessaloniki, Greece, 2001

PYRAMIDAL DIRECTIONAL FILTER BANKS AND CURVELETS

Minh N. Do[†] *and Martin Vetterli*[†§]

[†]Audio-Visual Communications Laboratory
Swiss Federal Institute of Technology Lausanne (EPFL), 1015 Lausanne, Switzerland

[§]Department of EECS, University of California at Berkeley, Berkeley CA 94720, USA

Email: {Minh.Do, Martin.Vetterli}@epfl.ch; Web: lcavwww.epfl.ch

ABSTRACT

A flexible multiscale and directional representation for images is proposed. The scheme combines directional filter banks with the Laplacian pyramid to provides a sparse representation for two-dimensional piecewise smooth signals resembling images. The underlying expansion is a frame and can be designed to be a tight frame. Pyramidal directional filter banks provide an effective method to implement the digital curvelet transform. The regularity issue of the iterated filters in the directional filter bank is examined.

1. INTRODUCTION

Over the last decade, wavelets had a growing impact on signal processing, mainly due to their good performance for piecewise smooth functions in one dimension. Unfortunately, such is not the case in two dimensions. In essence, wavelets are good at catching zero-dimensional singularities, but two-dimensional piecewise smooth signals resembling images have one-dimensional singularities. That is, smooth regions are separated by edges, and while edges are discontinuous across, they are typically smooth curves. Intuitively, wavelets in 2-D obtained by a tensor-product of one dimensional wavelets will be good at isolating the discontinuity across the edge, but will not see the smoothness along the edge. This disappointing behavior indicates that more powerful bases are needed in higher dimensions.

In the filter bank literature, Bamberger and Smith [1] had proposed an effective filter bank for the directional decomposition of images. This directional filter bank (DFB) has the important property that it can be critically sampled while achieving perfect reconstruction. In order to obtain *sparse* image representations, where maximum information is packed into a small number of samples, we propose a combination of DFB with a multiresolution pyramid, the result is called pyramidal directional filter bank (PDFB).

Recently, Candes and Donoho [2] pioneered a new system of representations named *curvelet* that was shown to be suited for objects which are smooth away from discontinuities across smooth curves. Their initial transforms were intended for functions on the continuous space $\mathbb{R}^2$. We will demonstrate that with PDFB, one achieves a curvelet-like decomposition for discrete-time signals.

The outline of the paper is as follows. Section 2 examines the DFB, together with new results on regular and orthogonal iterated DFB. Section 3 introduces the PDFB with its frequency decomposition and frame properties. Section 4 establishes the links between PDFB and the curvelet transform. Finally, Section 5 illustrates some numerical experiments on real images.

This work was supported by a Department of Communication Systems, EPFL PhD Fellowship and the Swiss National Science Foundation grant number 21-52439.97.

2. DIRECTIONAL FILTER BANKS

The DFB realizes a division of 2-D spectrum into 2^n wedge-shaped slices as shown in Fig. 1 using an n-levels iterated tree-structured filter banks [1]. The method is to use appropriately the quincunx filter bank (QFB) [3] together with modulations and rotations. Rotations in DFB are achieved by *resampling* matrices R_i (that is, matrices with determinant equal to ± 1, so they represent a rearrangement of the input samples).

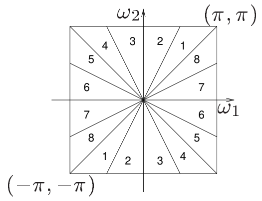


Fig. 1. Directional filter bank frequency partitioning.

More specifically, the blocks in the binary decomposition tree of the DFB are made up from two extensions of the QFB: the modulated QFB and the "skewed" QFB's (Fig. 2); the former one is used at the first two levels, while the later one is used at the remaining levels. Therefore, it can be shown that the DFB is perfect reconstruction (PR) or orthogonal *if and only if* its kernel QFB is PR or orthogonal, respectively. As a result, the design of DFB essentially amounts to the design of QFB with the desired properties.

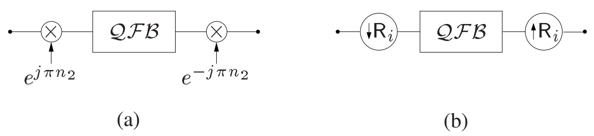


Fig. 2. Building blocks of the DFB. (a) Modulated QFB. (b) Skewed QFB's with $i = 1, 2, 3$ and 4.

In the ideal case, the Fourier transform of a DFB equivalent filter will be 1 only in its designated region. For a near horizontal subband, this is the region that is sandwiched between two lines

$\omega_1 = l\omega_2/N$ and $\omega_1 = (l+1)\omega_2/N$, where $N = 2^{n-2}$ and $l = -N, \ldots, N-1$. After some manipulations, we can compute the impulse responses of such a filter as

$$\begin{aligned} g[\boldsymbol{n}] &= \int_0^{\pi} dw_2 \int_{l\omega_2/N}^{(l+1)\omega_2/N} dw_1 \left(e^{j\boldsymbol{w}^T\boldsymbol{n}} + e^{-j\boldsymbol{w}^T\boldsymbol{n}}\right) \\ &= \frac{2\pi}{n_1}\left[\psi\left(\frac{n_1(l+1)}{N} + n_2\right) - \psi\left(\frac{n_1 l}{N} + n_2\right)\right], \end{aligned}$$

where $\psi(x) = (1 - cos(\pi x))/(\pi x)$, which behaves similarly to a 1-D sinc function. In other words, $g[\boldsymbol{n}]$ is the difference between two ridge functions with sinc-like ridge profiles, that are oriented around the line $n_2 = -n_1 l/N$ and damped by $1/n_1$.

Since all the equivalent filters in the DFB are derived from the filters in the kernel QFB, an interesting and important question is what condition should one imposes on the filters of the QFB so that iterations in DFB lead to regular 2-D signals? In general, the regularity issue for 2-D iterated filter banks is much more involved than in the 1-D case. We sketch the main results here, for a more rigorous treatments, see [4].

Intuitively, at each iterated step, the frequency responses of the equivalent filters in the DFB are thinner (by half) in one direction but the same in the other. Likewise, the overall sampling matrices are dilated (by double) in only one direction. Thus the regularity issue of the DFB iterated filters can be linked with 1-D iterated filters along the dilated direction.

Let $G_0(z_1, z_2)$ be the quincunx lowpass filter in the DFB. With a change of variable, the regularity of iterated DFB filters can be directly related with the regularity of the following filter

$$G_0^{(n)}(z_1, z_2) = \prod_{k=0}^{n-1} G_0(z_1^{2^k}, z_2). \tag{1}$$

For *a fixed value* of z_2, let us define the following 1-D filters

$$F(z_1) = G(z_1, z_2) \quad \text{and} \quad F^{(n)}(z_1) = G_0^{(n)}(z_1, z_2), \tag{2}$$

then we have

$$F^{(n)}(z_1) = \prod_{k=0}^{n-1} F(z_1^{2^k}). \tag{3}$$

By examining their impulse responses, we can show that $G_0^{(n)}(z_1, z_2)$ is smooth (for example with a certain Hölder degree) *if and only if* $F^{(n)}(z_1)$ is smooth. Equation (3) is the well-known wavelet-type iterated filters in 1-D; and thus the filter $F(z_1)$ is required to be regular of order $N > 1$ as

$$F(z_1) = (1 + z_1)^N R'(z_1). \tag{4}$$

Since (4) have to be satisfied for *all* value of z_2, the regularity condition on $G_0(z_1, z_2)$ becomes

$$G_0(z_1, z_2) = (1 + z_1)^N R(z_1, z_2). \tag{5}$$

In addition, we would like to have orthogonal DFB's, which means G_0 has to be an orthogonal quincunx filter or satisfies the following equation [5]

$$\begin{aligned} &G_0(z_1, z_2)G_0(z_1^{-1}, z_2^{-1}) \\ &\quad + G_0(-z_1, -z_2)G(-z_1^{-1}, -z_2^{-1}) = 2. \end{aligned} \tag{6}$$

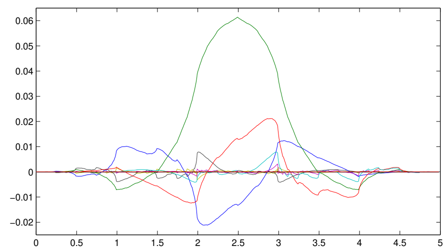


Fig. 3. Eight iteration of the smallest 2-D orthogonal quincunx filter for DFB. Each curve corresponds to a 1-D sequence of $g_0^{(8)}[n_1, n_2]$ where n_2 is fixed.

We notice that there exists an infinite number of solutions for (5)-(6) with $G_0(z_1, z_2) = G_0'(z_1)$ where $G_0'(z_1)$ is a 1-D regular order N filter (for example Daubechies filters). We refer to such solutions as "degenerated" since they are in fact 1-D filters. For general true 2-D filters, there is no simple design method like in 1-D since we lack a factorization theorem for multivariate polynomials. Using Gröbner bases technique [6], we found the smallest size non-degenerated quincunx orthogonal regular filter with $N = 2$ for DFB is

$$\frac{\sqrt{2}}{64}\begin{pmatrix} \sqrt{15} & 0 & -2\sqrt{15} & 0 & \sqrt{15} & 0 \\ -3 & 5 & 30 & 30 & 5 & -3 \\ 0 & -\sqrt{15} & 0 & 2\sqrt{15} & 0 & -\sqrt{15} \end{pmatrix}$$

It is also the unique solution (up to time reversal) with this size. The iterated filters can be checked to converge to continuous 2-D functions, which is illustrated in Fig 3.

3. PYRAMIDAL DIRECTIONAL FILTER BANKS

By its nature, the DFB is designed to capture the high frequency components (representing directionality) of images. Therefore, low frequency components are handed poorly by the DFB. In fact, with the frequency partition shown in Fig. 1, low frequencies would "leaks" into several directional subbands, hence DFB does not provide a sparse representation. To improve the situation, low frequencies should be subtracted previous to the DFB. This prompts us to combine DFB with a multiresolution scheme.

One way of achieving multiscale decomposition is to use the Laplacian pyramid (LP) introduced by Burt and Adelson [7]. The LP decomposition at each step generates a sampled lowpass version of the original and the difference between the original and the prediction, results in a bandpass image. A drawback of the LP is the implicit oversampling. However, in contrast with the critically sampled wavelet scheme, the LP has the distinguishing feature that each pyramid level generates only one bandpass image (even for multidimensional cases) which does not have "scrambled" frequencies. This frequency scrambling happens in the wavelet filter bank when a highpass channel, after downsampling, is folded back into the low frequency, and thus its spectrum is reflected. In the LP, this effect is avoided by only downsampling the lowpass channel.

Therefore the LP permits further subband decomposition to be applied on its bandpass images. Those bandpass images can be fed into a DFB so that directional information can be well captured. Fig. 4 depicts this pyramidal directional filter bank (PDFB). The scheme can be iterated repeatedly on the coarse image. The end result is a decomposition into directional subbands at multiple scales.

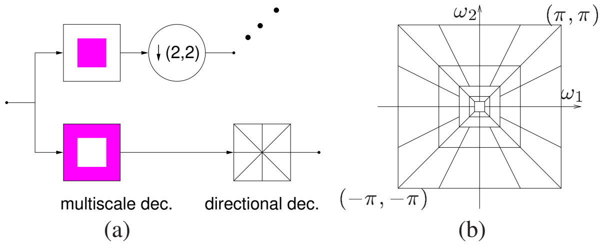


Fig. 4. Pyramidal directional filter banks. (a) Block diagram. (b) Frequency decomposition. Notice that as the scale is refined from coarse to fine, the number of directions is doubled at every other dyadic subband.

With perfect reconstructing LP and DFB, the PDFB is obviously PR, hence it is a frame operator for 2-D signals. The PDFB has the same redundancy as the LP: up to 33% when subsampling by two in each dimension. If the filters in the LP are orthogonal partners (that is, $h[\boldsymbol{n}] = g[-\boldsymbol{n}]$ and $g[\boldsymbol{n}]$ is orthogonal to its translates with respect to the subsampling lattice), then LP is shown to be a tight frame with energy conservation [8]. In that case, assuming an image x is decomposed into J bandpass images $b_j,\ j = 1, 2, \ldots, J$ and a lowpass image a_J,[1] then we have

$$\|x\|^2 = \sum_{j=1}^{J} \|b_j\|^2 + \|a_J\|^2 . \tag{7}$$

Now, in addition, we assume that the DFB is also orthogonal; it decomposes each bandpass image b_j into directional coefficients d_j with: $\|b_j\|^2 = \|d_j\|^2$. Then the decomposition by PDFB: $x \mapsto (d_1, d_2, \ldots, d_J, a_J)$ also has the energy conservation property:

$$\|x\|^2 = \sum_{j=1}^{J} \|d_j\|^2 + \|a_J\|^2 . \tag{8}$$

In other words, the PDFB is a tight frame when orthogonal filters are used in LP and DFB. Tight frame is an important property in some applications, for example, the error introduced in the transform domain is the same as the error in the reconstructed signal.

Let us point out that there are other multiscale and directional decompositions such as the cortex transform [9] and the steerable pyramid [10]. Our PDFB differs with those in that it allows different number of directions at each scale while nearly achieving critical sampling. In addition, we make the link to continuous-time construction precise, both through a relation to curvelets, and by studying convergence of iterated DFB's.

4. CURVELET TRANSFORM USING PDFB

In a nutshell, the curvelet transform [2] is obtained by filtering and then applying windowed ridgelet transform on each bandpass image. In $\mathbb{R}^2$, ridgelets are constant along ridge lines $x_1 \cos(\theta) + x_2 \sin(\theta) = const$ and are wavelets (with a scale s) along the orthogonal direction. In frequency domain, such a ridgelet function is essentially localized in the corona $|\boldsymbol{\omega}| \in [2^s, 2^{s+1}]$ and around the angle θ. The ridgelet transform is shown to provide a sparse representation for smooth objects with straight edges.

The curvelet decomposition can be described in the following steps [2]:

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

The motivation for this is that by smooth windowing, segments of smooth curves with would look straight in subimages, hence they can be well captured by a local ridgelet transform. Subband decomposition is used to keep the number of ridgelets at multiple scales under control by the fact that ridgelets of a given scale live in a certain subband. The window's size and subband frequency are coordinated so that curvelets have support obeying the key anisotropy scaling relation for curves [2]:

$$width \propto length^2. \tag{9}$$

Now we will demonstrate that a PDFB where *the number of directions is double at every other finer scale in the pyramid* (that is, we apply a DFB with $(n_0 - \lfloor j/2 \rfloor)$ levels to the bandpass image b_j of the LP) satisfies those key properties of curvelets. Thus PDFB provides a filter bank approach for curvelet-like decomposition.

A LP, which downsampling by two in each direction is taken at every level, provides an octave-band decomposition: the LP bandpass image b_j at the level j serves subband $[\pi 2^{-j}, \pi 2^{-j+1}]$, $j = 1, 2, \ldots, J$. Thus, in combining with directional decomposition of DFB, we obtain the frequency tiling for ridgelets as shown in Fig. 4(b). Moreover, a coefficients from b_j generates a basis image that has local support in a square of size about 2^j; while basis images from a DFB with $(n_0 - \lfloor j/2 \rfloor)$ iterated levels has support in a rectangle of length about $2^{n_0 - j/2}$ and width 1. Together, in the PDFB, a basis image at the pyramid level j has:

$$width \approx 2^j \quad \text{and} \quad length \approx 2^j . 2^{n_0 - j/2} = 2^{n_0} 2^{j/2}, \tag{10}$$

which satisfies the anisotropy scaling relation (9) of curvelets.

Fig.5 illustrates this property by displaying some basis images from a PDFB that implements the digital curvelet transform. Those basis images have their widths contracted by half at every finer scale, while their lengths are only contracted at every other scale.

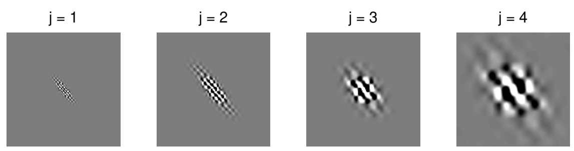
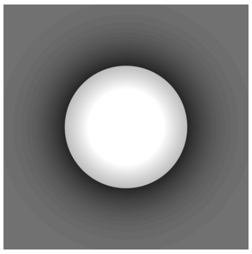


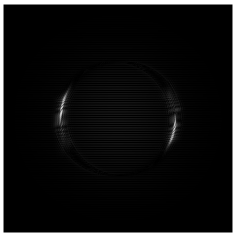
Fig. 5. Basis images from a PDFB that implements the curvelet transform.

In [11], Starck et al. describe a different approach for the digital curvelet transform, in which they directly "discretizes" the continuum definition. Their implementation uses the discrete Radon transform on image blocks, and thus the number of represented directions, which equals the block size, is reduced by half for every

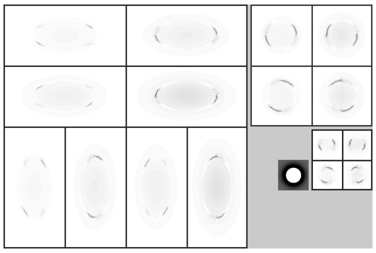
[1]The index is such the level $j = 1$ corresponds to the finest scale.



(a)



(c)



(b)

Fig. 6. Example of PDFB. (a) Input image. (b) Magnitudes of PDFB coefficients. (c) Reconstructed image from a PDFB subband.

finer scale. This goes opposite against the curvelet construction in continuous space, and is unlike our construction. Furthermore, there is a redundancy factor equals to $16J + 1$ in their implementation compared with 1.3 in ours.

5. EXPERIMENTS AND DISCUSSION

Fig. 6 shows an example image that is transformed by the PDFB. To obtain visible directional selectivity, we use larger quincunx filters that are closer to the ideal case. As we can see, the coefficients in the transform domain are very sparse – significant coefficients are located around edges *and* in the right directional subbands. With non-linear approximation using the PDFB, smooth regions are represented efficiently by the small size lowpass image while smooth edges are efficiently represented by a few directional local coefficients.

The non-linear approximation power of PDFB is tested in denoising experiments and compared with the wavelet transform. In both cases, simple hard thresholding is performed in the transform domain. Fig. 7 displays a detail examination, where PDFB is shown to be more effective in recovering edges, as well as in signal-to-noise ratio (SNR).

For compression applications, there is the cost associated with describing the retained coefficients and their locations that must be taken into account. The success of wavelets in current image compression algorithms partly relies on the existence of embedded trees that effectively point to (in)significant coefficients in the transform domain. In PDFB, from coarse to fine scale, resolution increases in both position and direction. Thus, embedded trees similar to the zero-tree wavelet can be used to successively locate the position *and* direction of image edges. Furthermore, as smooth edges have their tangent varies slowly, significant coefficients in PDFB are localized in both position and direction.

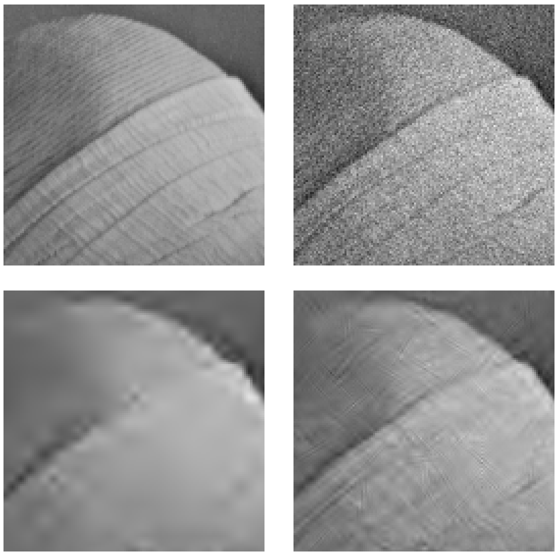


Fig. 7. Denoising experiment: original image (top left), noisy image (top right, SNR = 9.55 dB), denoising using 2-D wavelets (bottom left, SNR = 13.82 dB), and denoising using PDFB (bottom right, SNR = 15.42 dB).

Acknowledgments: The authors would like to thank Professor Ivan Selesnick for his help with the software for Gröbner bases.

6. REFERENCES

[1] 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.

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

[3] M. Vetterli, "Multidimensional subband coding: Some theory and algorithms," *Signal Proc.*, vol. 6, no. 2, pp. 97–112, February 1984.

[4] M. N. Do and M. Vetterli, "Directional local bases from multidimensional filter banks," *IEEE Trans. Image Proc.*, 2001, submitted.

[5] J. Kovačević and M. Vetterli, "Nonseparable multidimensional perfect reconstruction filter banks and wavelet bases for $\mathcal{R}^n$," *IEEE Trans. Inform. Th.*, vol. 38, no. 2, pp. 533–555, March 1992.

[6] J. Lebrun and I. Selesnick, "Gröbner bases and wavelet design," *Journal of Symbolic Computation*, 2001, to appear.

[7] 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.

[8] M. N. Do and M. Vetterli, "Frame reconstruction of the Laplacian pyramid," in *IEEE Trans. Acoust., Speech, and Signal Proc.*, Salt Lake City, May 2001.

[9] 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.

[10] 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.

[11] J. L. Starck, E. J. Candès, and D.L. Donoho, "The curvelet transform for image denoising," *IEEE Trans. Image Proc.*, 2000, submitted.