Research Article

Implementational Aspects of the Contourlet Filter Bank and Application in Image Coding

Truong T. Nguyen,[1] Yilong Liu,[2] Hervé Chauris,[1] and Soontorn Oraintara[2]

[1] *Geophysical Research Center, Paris School of Mines, 35 Rue Saint-Honoré, 77305 Fontainebleau, France*
[2] *Department of Electrical Engineering, University of Texas at Arlington, 416 Yates St. Arlington, Texas 76019, USA*

Correspondence should be addressed to Truong T. Nguyen, truong.nguyen@ensmp.fr

This paper analyzed the implementational aspects of the contourlet filter bank (or the pyramidal directional filter bank (PDFB)), and considered its application in image coding. First, details of the binary tree-structured directional filter bank (DFB) are presented, including a modification to minimize the phase delay factor and necessary steps for handling rectangular images. The PDFB is viewed as an overcomplete filter bank, and the directional filters are expressed in terms of polyphase components of the pyramidal filter bank and the conventional DFB. The aliasing effect of the conventional DFB and the Laplacian pyramid to the directional filters is then considered, and the conditions for reducing this effect are presented. The new filters obtained by redesigning the PDFBs satisfying these requirements have much better frequency responses. A hybrid multiscale filter bank consisting of the PDFB at higher scales and the traditional maximally decimated wavelet filter bank at lower scales is constructed to provide a sparse image representation. A novel embedded image coding system based on the image decomposition and a morphological dilation algorithm is then presented. The coding algorithm efficiently clusters the significant coefficients using progressive morphological operations. Context models for arithmetic coding are designed to exploit the intraband dependency and the correlation existing among the neighboring directional subbands. Experimental results show that the proposed coding algorithm outperforms the current state-of-the-art wavelet-based coders, such as JPEG2000, for images with directional features.

1. INTRODUCTION

Wavelet and filter bank (FB) have been one of major research topics in signal processing for the last two decades [1]. It is shown that the continuous wavelet functions and their associated regular FBs are optimal representations of one dimensional piecewise smooth signals [2]. However, the direct extension of wavelet to two dimensions by the tensor product of one dimensional wavelet is no longer optimal for image representation due to the intrinsic geometrical structure of typical natural images [3]. In short, the separable wavelets are still optimal in representing point discontinuities in a two-dimensional signal, but not effective in capturing line discontinuities, which correspond to directional information in the image. Therefore, integrating geometric regularity in the image representation is a key challenge to improve the performances of current image coders.

Recently, Candès and Donoho constructed the curvelet transform [4], and proved that it is an essentially optimal representation of two variable functions, which are smooth except at discontinuities along $\mathbf{C}^2$ (twice differentiable) curve. The nonlinear approximation of a function f, $f_M^{(c)}$, reconstructed by M curvelet coefficients has an asymptotic decay rate of $\|f - f_M^{(c)}\|^2 \le CM^{-2}(\log_2 M)^3$. This decay rate of the approximation error is a significant theoretical improvement compared to those by wavelet and Fourier coefficients, which are $O(M^{-1})$ and $O(M^{-1/2})$, respectively [2]. Since the space of smooth functions with singularities along $\mathbf{C}^2$ curves is similar to that of natural images with regions of continuous intensity and discontinuous along smooth curves (edges), there is a strong motivation for finding a similar transform in the discrete domain [3].

In [5], Do and Vetterli proposed the pyramidal directional filter bank (PDFB) to implement the contourlet transform. The PDFB or contourlet FB is a combination of

the Laplacian pyramid [6] and the conventional directional filter bank (DFB) [7]. It unites the advantages of both systems, which are multiresolution and multidirection. The authors also show that the contourlet transform can achieve the asymptotic optimal result as the curvelet transform. Essentially, these conditions assume that the directional filters have good passband and stopband characteristics in the Fourier domain.

The contourlet transform and its associated FB has been an active subject of research in recent years [8–12]. However, there has been no comprehensive work on the detail implementation of the FB and its application in image coding.

Because of its effectiveness in representing natural image, the PDFB has been used in image coding recently [13–16]. These works, however, are based on the original implementation of the PDFB [5], which contains aliasing in the directional filters. Therefore, the achieved coding results are not comparable to state-of-the-art wavelet-based coders, such as JPEG2000.

The contribution of this work is as follows. First, we provide detail explanations of the DFB tree as it is implemented in the contourlet FB [17]. Moreover, the structure is modified to minimize the phase delay of the DFB basis. We also demonstrate how to handle size-limited images along the DFB tree and the effect of the binary tree on the frequency supports of the DFB directional filters. Our second contribution is to demonstrate several aliasing problems of the DFB tree and the PDFB. It is shown that some of these aliasing problems can be attenuated by changing the design criteria of the filters in the Laplacian pyramid of the PDFB. Finally, a novel embedded image coding system based on the image decomposition and a morphological dilation algorithm are presented. Experimental results show that the proposed coding algorithm outperforms the current state-of-the-art wavelet-based coders, such as JPEG2000, for images with directional features.

1.1. Notations and paper outline

Here, we briefly review the notations and terminology used in 2-D multirate system. For fundamental operations in multidimensional 2-D systems, we refer to [18].

(i) Uppercase and lowercase bold face letters represent 2×2 square matrices and 2×1 column vectors, respectively. For example, $h(\mathbf{n})$ is a function defined on the 2-D integer lattice $(n_1, n_2)^T$, and $\boldsymbol{\pi}$ is $(\pi, \pi)^T$. Bold letters variables are employed frequently to compress mathematical expressions. For example, $H(\boldsymbol{\omega})$ is the same as $H(\omega_1, \omega_2)$.

(ii) The superscripts T and $-T$ denote the transpose, transpose of the inverse, respectively.

(iii) $\mathcal{N}(\mathbf{M})$ is defined as the set of integer vectors of the form $\mathbf{M}\mathbf{x}$, where $\mathbf{x} \in [0, 1)^2$. $|\mathbf{M}|$ represents the determinant of the matrix $\mathbf{M}$. The number of elements in $\mathcal{N}(\mathbf{M})$ is equal to $|\mathbf{M}|$.

(iv) The Fourier transform of a 2-D filter $h(\mathbf{n})$ is defined as

$$H(\boldsymbol{\omega}) = \sum_{\mathbf{n}\in\mathbb{Z}^2} h(\mathbf{n})e^{-j\boldsymbol{\omega}^T\mathbf{n}}, \quad \text{where } \boldsymbol{\omega} = (\omega_1, \omega_2)^T. \quad (1)$$

(v) Matrix exponential notations greatly simplify 2-D multirate system expressions. The notation $\mathbf{z}^{\mathbf{M}}$ is defined as

$$\mathbf{z}^{\mathbf{M}} \triangleq \left[z_1^{m_{11}}z_2^{m_{21}}, z_1^{m_{12}}z_2^{m_{22}}\right]^T, \quad \text{where } \mathbf{M} = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}. \quad (2)$$

Therefore, $H(\boldsymbol{\omega}^{\mathbf{M}}) = H(\mathbf{M}^T\boldsymbol{\omega}) = H(m_{11}\omega_1 + m_{21}\omega_2, m_{12}\omega_1 + m_{22}\omega_2)$.

The following matrices are used to decimate subband images in the PDFB:

$$\mathbf{Q}_0 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{Q}_1 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad \mathbf{R}_0 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},$$

$$\mathbf{R}_1 = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{R}_2 = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix},$$

$$\mathbf{R}_3 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{D}_0 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix},$$

$$\mathbf{D}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \quad \mathbf{D}_2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = 2\mathbf{I}. \quad (3)$$

Paper outline

The DFB is analyzed in detail in Section 2. We first show how the DFB tree can be implemented by a binary-tree structure, as in [5, 7, 19]. The tree structure is then modified by adding delay and advanced blocks at appropriate places to make sure that the phase delays of the overall directional filters in the DFB tree are minimized. We also show how rectangular images are handled along the tree structure. In Section 3, the PDFB is viewed as an overcomplete FB, and the equivalent directional filters are expressed in terms of the polyphase components. The aliasing problems existing in the PDFB structure are analyzed in Section 3.1, where it is shown that most of the aliasing can be removed if the two lowpass filters employed in the pyramid satisfy the Nyquist criteria. Discussion and simulations in Section 3.2 demonstrate the improvement of the PDFB with new design conditions. In Section 4, we describe a novel embedded image coding scheme based on the image decomposition obtained by a combination of PDFB and wavelet filter bank. The morphological operation is employed progressively to identify clusters of significant coefficients in each bit plane. Context-based arithmetic coding is used to encode these significant coefficients. We design the context models so that the intraband and interband correlations of the overcomplete PDFB can be well exploited. The coding results and discussions are presented in Section 4.4. Section 5 concludes the paper. Parts of this work have been presented in [20, 21].

2. ITERATIVE TREE STRUCTURE FOR THE DIRECTIONAL FILTER BANK

The DFB is a 2^n-band maximally decimated, perfect reconstruction (PR) FB introduced by Bamberger and Smith in [7]. The original construction of an eight-band DFB, whose frequency partitioning when $n = 2$ (or 3) is in Figure 1, uses a two-level (or three-level) binary tree of two-channel FBs. The construction is not easy to generalize to 2^n-band DFB, $n > 3$, since two-channel FBs with different passband supports are needed at different levels of the tree. A new construction of the DFB tree is proposed in [22]; its implementation is included in the contourlet toolbox [5]. Figure 2 illustrates this new implementation of the DFB for the case n up to 3.

In order to explain the DFB tree in Figure 2, we need to define four matrices $\mathbf{P}_i$, $i = 0, 1, 2, 3$ from the previously defined matrices. The following matrix identities can be easily checked:

$$\mathbf{P}_0 = \mathbf{R}_0 \mathbf{Q}_0 = \mathbf{D}_0 \mathbf{R}_3, \qquad \mathbf{P}_1 = \mathbf{R}_1 \mathbf{Q}_1 = \mathbf{D}_0 \mathbf{R}_2, \qquad (4)$$

$$\mathbf{P}_2 = \mathbf{R}_2 \mathbf{Q}_0 = \mathbf{D}_1 \mathbf{R}_1, \qquad \mathbf{P}_3 = \mathbf{R}_3 \mathbf{Q}_1 = \mathbf{D}_1 \mathbf{R}_0, \qquad (5)$$

$$\mathbf{R}_i \mathbf{D}_1^n \mathbf{2}^n = \mathbf{D}_1^n \mathbf{R}_i^{2^n}, \quad i = 0, 1, \qquad \mathbf{R}_i \mathbf{D}_0^n = \mathbf{D}_0^n \mathbf{R}_i^{2^n}, \quad i = 2, 3, \qquad (6)$$

$$\mathbf{Q}_0 \mathbf{Q}_1 = 2\mathbf{I}, \qquad \mathbf{R}_0 \mathbf{R}_1 = \mathbf{R}_2 \mathbf{R}_3 = \mathbf{I}. \qquad (7)$$

An unimodular matrix is a matrix with determinant equal to ± 1. A resampling matrix is an unimodular matrix having integer elements. The matrices $\mathbf{R}_i$, $i = 0, \ldots, 3$, as well as $\mathbf{S}_i^n$, $n = 3$, $i = 1, \ldots, 2^n$, at the end of the DFB tree are resampling matrices. The exact values of $\mathbf{S}_i^n$ are later defined in (10) and (11). When an image is down- or upsampled by a resampling matrix, its pixels are simply reorganized (or the image is resampled). For example, the rectangular image with white line border in Figure 8(b) can be obtained by downsampling the corresponding image in Figure 8(a) by $\mathbf{R}_3$.

The DFB binary tree in Figure 2 is constructed by using fan FBs. However, these fan FBs use different decimation matrices. In the first level, the fan FB uses the decimation matrix $\mathbf{Q}_0$; the two fan FBs at the second level use the decimation matrix $\mathbf{Q}_1$. If we switch the decimation block $\mathbf{Q}_0$ at the first level and the fan FBs at the second level and apply multirate FB theory, it is straightforward to show that the overall filters have frequency supports as depicted in Figure 1(a) and the decimation ratios of all four output bands are $2\mathbf{I}$. The expressions for these four directional filters from fan filters $H_0^{(f)}(\mathbf{z})$ and $H_1^{(f)}(\mathbf{z})$ are in (12)–(15).

Beginning from the third level, four types of resampled fan FBs referred to as F_0, F_1, F_2, and F_3 are used. The positions of these four resampled fan FBs at the third level of the DFB tree are illustrated in Figure 2. Each of these fan FBs has two outputs: 0 and 1. We call them resampled fan FBs because the input of F_i is resampled by the unimodular matrix $\mathbf{R}_i$ before entering the fan filters. If we switch the resampling matrices and the fan filters, the shape of the equivalent frequency response of the fan filter will be parallelograms, as demonstrated in Figure 3.

One can obtain a 2^{n+1}-band DFB from the tree structure of the 2^n-band DFB by appending fan FB's F_{2k+p} at outputs p of F_{2k} and F_{2k+1} of the nth level, where $k, p = 0, 1$. This iterative division of the frequency plane to produce finer direction can be explained as follows. Let us consider the first subband (number 0) of a four-band DFB, which is denoted as $x(\mathbf{n})$. This signal is also the input of a fan FB of type F_0 to obtain eight-band DFB (Figure 2). Since the outputs of a four-band DFB are decimated by $2\mathbf{I}$, the frequency of the subband 0 in Figure 1(a) is magnified by 2 in both ω_1 and ω_2 directions, as illustrated in Figure 4(a). The frequency regions that correspond to band 0 and 1 in an eight-band DFB (Figure 1(b)) are mapped to the two darker and the two lighter regions, respectively. On the other hand, by switching the resampling block $\mathbf{R}_0$ and the fan filters, the effective support of the fan FB type F_0 has parallelogram shapes as illustrated in Figure 3. One can see that indeed the two fan filters in fan FB type F_0 in Figure 3 separate subband 0 and subband 1 of an eight-band DFB.

Following multirate signal processing theory, when the signal $x(\mathbf{n})$ is decimated by $\mathbf{P}_0$, the frequency content of $X(\boldsymbol{\omega})$ is mapped to the support of $X(\mathbf{P}_0^{-T} \boldsymbol{\omega})$ (Figure 4(b)). By the observation of the two Figures 3(b) and 4, we can see that in order to obtain a sixteen-band DFB from eight-band DFB, we need to attach fan FB type F_0 and F_1 on the two outputs of the FB type F_0 at the third level of the DFB tree.

In order to determine the subband number at the output of the resampled fan FBs F_i for a 2^n-band DFB ($n \geq 3$), we only have to follow the path of the subband in the binary tree and collect the output number at each level of the tree. The n-digit number is the binary form of the subband number (see Figure 2 for the case of an eight-band DFB). The numbering of directional subbands in 2^n-band DFB has to start from 0, beginning from the left to right, top to bottom, as depicted in Figure 1.

After iteratively cascading fan FBs to create an n-level tree for a 2^n-band DFB, the equivalent decimation matrices for each subband will not be the same. By following the above cascading rule, the overall decimation matrix can be determined. For example, the subband i in the 2^n-band DFB whose binary representation is $i = m_1 m_2 \cdots m_n$ has the following overall decimation matrix:

$$\mathbf{P}_i^n = \begin{cases} \mathbf{Q}_0 \mathbf{Q}_1 \prod_{j=2}^{n-1} \mathbf{P}_{m_j}, & 0 \leq i < 2^{n-1}, \\ \mathbf{Q}_0 \mathbf{Q}_1 \prod_{j=2}^{n-1} \mathbf{P}_{(m_j+2)}, & 2^{n-1} \leq i < 2^n, \end{cases} \qquad (8)$$

where $\mathbf{P}_i$'s are defined in (4) and (5). Using the matrix identities from (4) to (7), we can further simplify the overall decimation matrices as

$$\mathbf{P}_i^n = \mathbf{D}_2 \prod_{j=2}^{n-1} \mathbf{D}_0 \mathbf{R}_{(3-m_j)} = \mathrm{diag}(2^{(n-1)}, 2) \prod_{j=2}^{n-1} \mathbf{R}_{(3-m_j)}^{2^{(n-j-1)}}, \qquad (9)$$

for $0 \leq i < 2^{n-1}$. The final step of the analysis FB is to resample subbands by backsampling matrices $\mathbf{S}_i^n$ to make the overall decimation matrix diagonal:

$$\mathbf{S}_i^n = \left(\prod_{j=2}^{n-1} \mathbf{R}_{(3-m_j)}^{2^{(n-j-1)}} \right)^{-1} = \prod_{j=2}^{n-1} \mathbf{R}_{(2+m_j)}^{2^{(n-j-1)}}. \qquad (10)$$
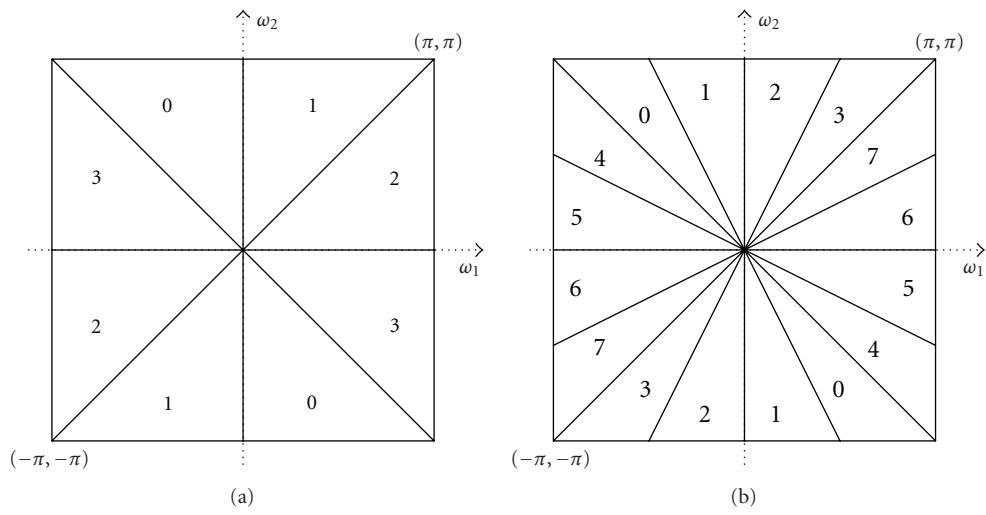
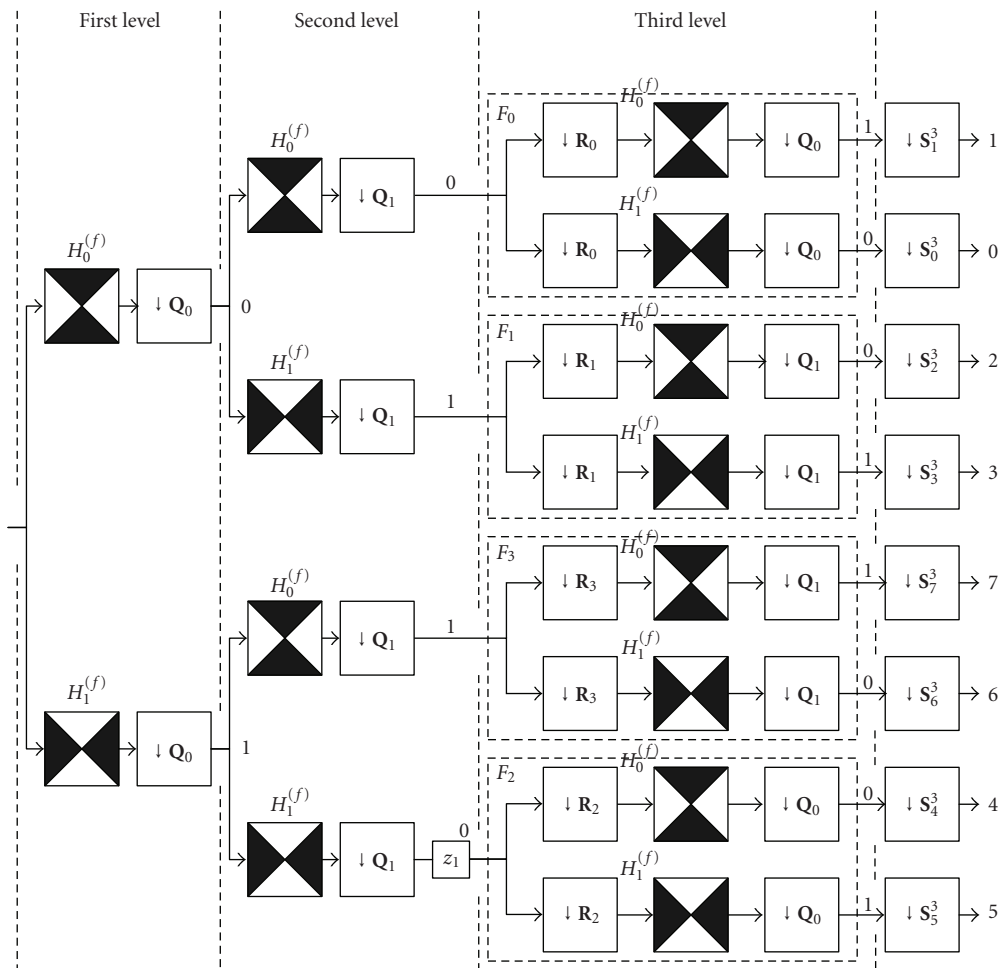FIGURE 1: Frequency divisions of the conventional DFB [7] in case of (a) four-channel DFB and (b) eight-channel DFB.

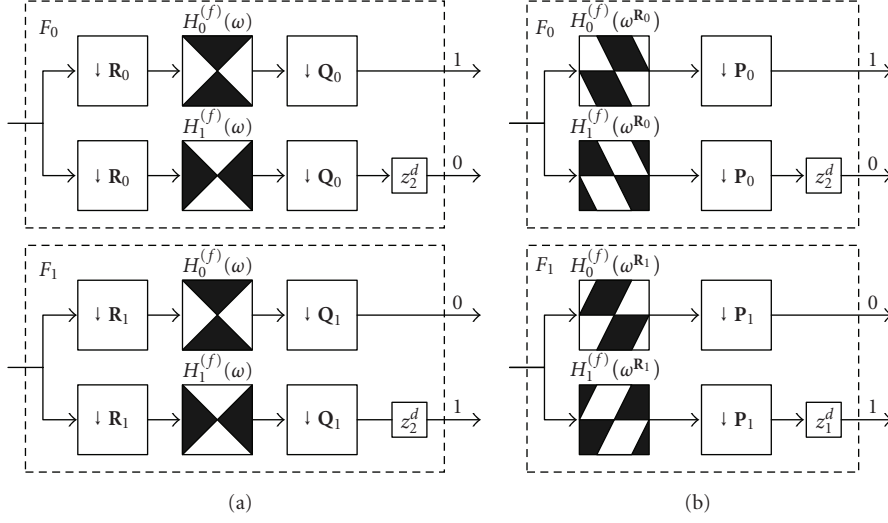FIGURE 2: The iterative tree structure for the conventional DFB [22].

FIGURE 3: Resampled fan FBs of types F_0 and F_1 used after level two of the binary DFB tree: (a) original structures and (b) equivalent structures obtained by switching the resampling blocks and the filters $H_i^{(f)}$.
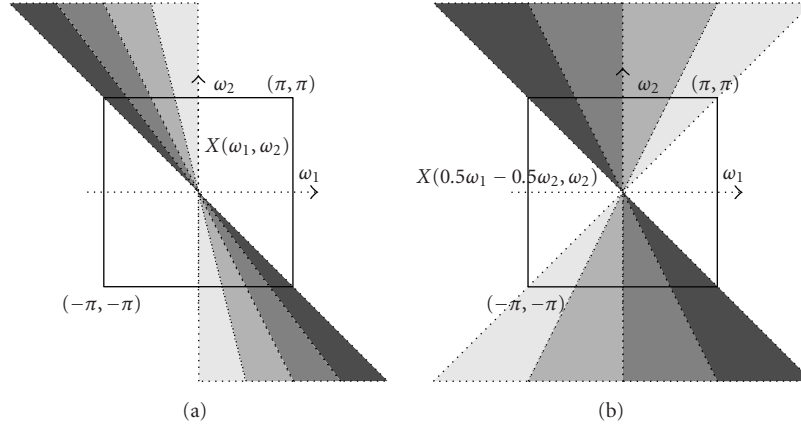

FIGURE 4: Illustration of frequency mapping when a 2-D signal $X(\boldsymbol{\omega})$ is downsampled by $\mathbf{P}_0$, (a) $X(\boldsymbol{\omega})$ is the subband 0 in a four-band DFB in Figure 1(a) decimated by $2\mathbf{I}$, (b) the support of $X(\mathbf{P}_0^{-T}\boldsymbol{\omega})$.

By similar derivation, one can show that the backsampling matrices for subband i, $2^{n-1} \le i < 2^n$ are

$$\mathbf{S}_i^n = \prod_{j=2}^{n-1} \mathbf{R}_{m_j}^{2^{(n-j-1)}}. \tag{11}$$

The process in the synthesis tree is exactly the reverse of the analysis side.

2.1. Phase correction in the DFB tree

Although the construction of the DFB using iterative fan FB described above has many advantages, it also has one potential problem. Typically, the two prototype fan filters are linear phase with different phase delay. However, the resampling blocks in the tree further deviate the overall phase delay of the directional filters. In processing a size-limited image, periodic or symmetric extension is used to obtain PR without introducing additional samples. If symmetric extension is used in the DFB tree, the phase delay at each level should be less than the overall decimation ratio. For example, if the phase delay of a particular subband filter at level n is larger than the decimation ratio of that subband, then the subband image will have an artificial border inside the image. As a result, the two subband images at level $n+1$ coming from that subband will have large areas of border artifacts.

2.1.1. Phase correction in the four-band DFB tree

In order to reduce the phase difference of these linear-phase filters at each level of the tree, delay and advance blocks are inserted at appropriate branches. In order to simplify our discussion, let us assume that the first fan filter $H_0^{(f)}(\mathbf{z})$ used in the tree in Figure 2 is zero phase, and the second filter has a phase factor of $e^{-j\omega_1}$, which means that $H_1^{(f)}(\mathbf{z}) = z_1^{-1} H_{1r}^{(f)}(\mathbf{z})$, where the subscription r of any filter in this paper denotes that the transfer function is zero phase and

that the corresponding frequency response is real. Available methods for designing PR linear-phase fan FBs satisfying the condition above can be found in many previous works such as [23–25], and will not be discussed here. At the second level of the tree, the equivalent frequency responses of the four directional filters are

$$H_0(\boldsymbol{\omega}) = H_0^{(f)}(\boldsymbol{\omega})H_0^{(f)}(\boldsymbol{\omega}^{\mathbf{Q}_0}) =: H_{0r}(\boldsymbol{\omega}), \tag{12}$$

$$H_1(\boldsymbol{\omega}) = H_0^{(f)}(\boldsymbol{\omega})H_1^{(f)}(\boldsymbol{\omega}^{\mathbf{Q}_0}) =: e^{-j(\omega_1+\omega_2)}H_{1r}(\boldsymbol{\omega}), \tag{13}$$

$$H_2(\boldsymbol{\omega}) = H_1^{(f)}(\boldsymbol{\omega})H_0^{(f)}(\boldsymbol{\omega}^{\mathbf{Q}_0}) =: e^{-j\omega_1}H_{2r}(\boldsymbol{\omega}), \tag{14}$$

$$H_3(\boldsymbol{\omega}) = H_1^{(f)}(\boldsymbol{\omega})H_1^{(f)}(\boldsymbol{\omega}^{\mathbf{Q}_0}) =: e^{-j(2\omega_1+\omega_2)}H_{3r}(\boldsymbol{\omega}). \tag{15}$$

Since the equivalent decimation matrix of these filters is $\mathbf{Q}_0\mathbf{Q}_1 = 2\mathbf{I}$, the delays of the filter in n_1 and n_2 directions can be kept less than the decimation ratios in those directions. In this case, the decimation ratios in n_1 and n_2 are two. Since the delay factor in $H_3(\boldsymbol{\omega})$ is $z_1^{-2}z_2^{-1}$, an advance block z_1 is inserted at its output as in Figure 2.

2.1.2. Phase correction in the 2^n-band DFB tree

As mentioned earlier, four types of resampled FBs (F_0, F_1, F_2, and F_3) are used starting from level three of the tree (Figure 2). Using the noble identities [18], switching the resampling block and the fan filter in type F_i results in a filter whose frequency response is of parallelogram shape with corresponding decimation matrix $\mathbf{P}_i$ as defined in (4) and (5). Figures 3(a) and 3(b), respectively, show examples of the FBs before and after switching for types F_0 and F_1. It can be shown that the overall decimation matrices starting from level three to the end of the tree including the backsampling matrices $\mathbf{S}_i^n$ are $\mathbf{D}_0^{n-2}$ for subbands 0 to $2^{n-1} - 1$ and $\mathbf{D}_1^{n-2}$ for subbands 2^{n-1} to $2^n - 1$. Therefore, the task of the delay block is to keep the distance between the center of symmetry of the filter and that of the directional filter at level two less than $(2^{n-2}, 1)$ for the first half, and $(1, 2^{n-2})$ for the second half of the tree. For example, the subband 0 of a 16-band DFB can be obtained by cascading the resampled fan FB of type F_0 to output 0 of the eight-band DFB in Figure 2 (before the backsampling matrix $\mathbf{S}_0^3$), as shown in Figure 5. Hence, the equivalent filter $H_0^{2-4}(\mathbf{z})$ (the superscript $2 - 4$ represents transition from level two to level four of the tree) for output 0 can be expressed in terms of the prototype filter $H_1^{(f)}(\mathbf{z})$ as

$$H_0^{2-4}(\mathbf{z}) = H_1^{(f)}(\mathbf{z}^{\mathbf{R}_0})H_1^{(f)}(\mathbf{z}^{\mathbf{P}_0\mathbf{R}_0}). \tag{16}$$

Since the phase factor of $H_1^{(f)}(\mathbf{z})$ is $e^{-j\omega_1}$, the phase factors of $H_1^{(f)}(\mathbf{z}^{\mathbf{R}_0})$ and $H_1^{(f)}(\mathbf{z}^{\mathbf{P}_0\mathbf{R}_0})$ are $e^{-j\omega_1}$ and $e^{-j(2\omega_1+\omega_2)}$, respectively. Hence, the phase factor of $H_0^{2-4}(\mathbf{z})$ is $e^{-j(3\omega_1+\omega_2)}$. Since the equivalent decimation matrix of this path is $\mathbf{D}_0^2 = \text{diag}(4,1)$ in order to keep the center of symmetry in the range $(4 - 1)$, an advance block z_2^d with $d = 1$ is inserted before the backsampling matrix $\mathbf{S}_0^4$ as illustrated in Figure 5.

Following similar analysis, it can be shown that for an n-level tree DFB with $n \geq 4$ in order for the overall directional filters to have their centers of symmetry within the decimation ratios, a delay block z_2^d with $d = \pm 2^{(n-4)}$

must be added to the resampled fan FBs of types F_0 and F_1 at level $n \geq 4$, as depicted in Figure 3. The value of d is positive for subbands $0, 1, \ldots, 2^{n-2} - 1$ and negative for subbands $2^{n-2}, \ldots, 2^{n-1} - 1$.

For the cases of resampled fan FBs of types F_2 and F_3, we assume that the first fan filter $H_0^{(f)}(\mathbf{z})$ is zero phase and the second fan filter $H_1^{(f)}(\mathbf{z})$ has a phase factor of $e^{-j\omega_2}$. Therefore, a delay block z_1^d with $d = \pm 2^{(n-4)}$ must be added to the resampled fan FBs of types F_2 and F_3 at level $n \geq 4$. The value of d is negative for subbands $2^{n-1}, \ldots, 3 \times 2^{n-2} - 1$ and positive for subbands $3 \times 2^{n-2}, \ldots, 2^n - 1$.

In order to demonstrate the effect of phase correction, we display six basis functions of the contourlet FB without and with phase correction in Figure 6. The six basis functions are chosen to be more vertical to demonstrate horizontal shift, and created by setting a coefficient of each selected subband to one and all others to zero. The positions of these six coefficients are in the same row and equally spaced in columns. Hence, the ideal reconstruction would contain six directional impulse responses nearly equally spaced in horizontal direction (they will be exact if all filters are zero phase). The upper and lower images in Figure 6 are the reconstructions of the PDFB without and with phase correction, respectively. We can see that without phase correction, the centers of the six basis functions are shifted (from the vertical dotted lines). This is minimized when phase correction is applied. This property is very important when one would like to take advantage of the interscale and inter-band relationship between coefficients at the same location [26].

2.2. Processing size-limited images along the DFB tree

One problem related to the implementation of the DFB that so far has not received sufficient attention in the literature is how to handle border extension of size-limited images so that the transform by DFB is PR and nonexpansive. Two methods that permit such critical representation are periodic extension and symmetric extension. Generally, the boundary reflection method is preferred over periodic extension because the periodic extension method creates artificial edges at the border of the image. However, mirror-symmetry extension is not possible for the DFB. This is because mirror-symmetry extension requires that the 2-D filters to be symmetric through vertical and horizontal axes (or quadrantally 2-D symmetric [27]), which is not possible for fan filters. Therefore, in this section, we consider only periodic extension for the DFB tree.

Let us assume that the input image to the DFB tree in Figure 2 is of a rectangular size $N \times M$, where N and M are even numbers. When being convolved with $H_i^{(f)}(\mathbf{z})$, $i = 0, 1$ at the first level of DFB tree, the image is periodically extended. When the convolved images are decimated by $\mathbf{Q}_0$, the resulting images will have rectangular shapes rotated by $45°$. These rotated rectangular images are not easy to store and process at the second level of the DFB tree. Instead, after decimation of the convolved images by $\mathbf{Q}_0$, we can keep only rectangular images of size $1/2 \max(M, N) \times 1/2(M+N)$. This
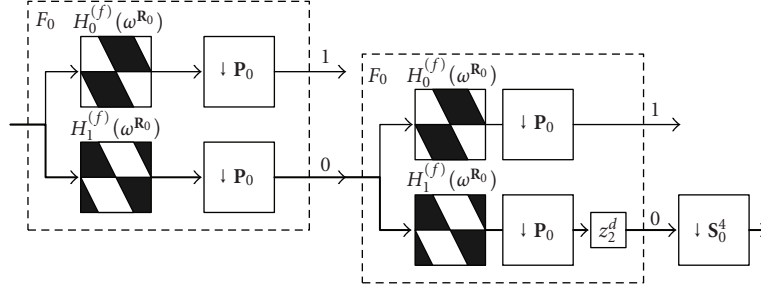
FIGURE 5: Subband 0 of the 16-band DFB. The figure shows the path (bold line) from level two to level four with corresponding transfer function $H_0^{2-4}(z)$.
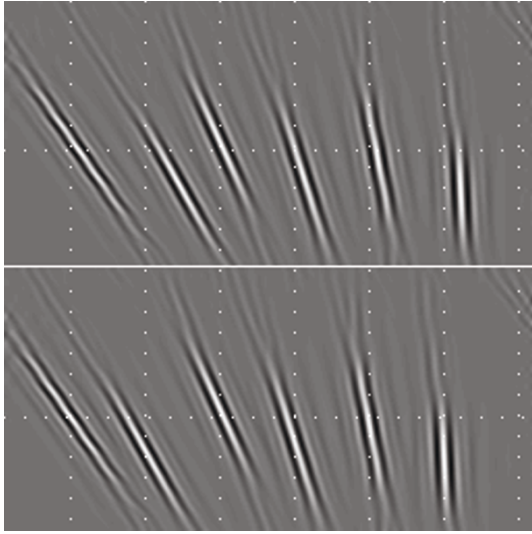

FIGURE 6: Six contourlet basis functions created by the PDFB without phase correction (upper part) and with phase correction (lower part).

operation is best explained by an example. The Boat image in the rectangular region with white line border in Figure 7(a) is of size 200×300. After downsampling by $\mathbf{Q}_0$, the rectangular image that we need to keep is of size 150×250, in the rectangular region with white line border in Figure 7(b).

At the second stage of the tree, we need to pay special attention to border extension of the two outputs of the first level. We cannot use periodic extension because it will lead to disturbing border artifact along diagonal line at the center of reconstructed images. Instead, we have to extend the two rectangular subband images in such a way that the periodic extensions at the first level of the DFB tree are preserved. This extension is illustrated in Figure 7(b), where the rectangular image inside the dotted line is the extended images in the second level of the DFB tree.

The decimation ratio for the first two levels of the DFB tree is $2\mathbf{I}$, the size of subband images enters fan FB F_i in Figure 2 is of size $M/2 \times N/2$. Beginning from the third level of the tree, the decimation ratio of the fan FB F_i is $\mathbf{P}_i$, which can be interpreted as a downsampling by 2 in row (or column), followed by a resampling operation. The resampled

images need to be parallelogram periodically extended so that there are no new borders at the extension area around the rectangular subband images. Figure 8 contains the Boat image downsampling by $\mathbf{D}_0$ and downsampling by $\mathbf{P}_0$. The image in Figure 8(a) is periodically extended, but the image in Figure 8(b) is parallelogram periodically extended.

2.3. Aliasing on the binary tree of the conventional DFB

The conventional DFB realized by a binary tree of maximally decimated two-channel FBs [7] is very efficient. However, it also leads to aliasing problems to the overall directional filters, which can be divided into two types: highpass aliasing and stopband aliasing components. From Figure 2, one can see that the DFB tree needs only one prototype fan FB [19]. Following the path to obtain subband 0 of an eight-band DFB (Figure 1(b)) in the DFB tree in Figure 2 and using noble identities of multirate system theory [18], we can move all the decimation blocks to the right-hand side, after all the filtering blocks. The directional filter corresponds to subband 0 in the DFB tree is given by

$$
\begin{aligned}
H_0(\omega_1, \omega_2) &= H_0^{(f)}(\boldsymbol{\omega}) H_0^{(f)}(\boldsymbol{\omega}^{\mathbf{Q}_0}) H_1^{(f)}(\boldsymbol{\omega}^{2\mathbf{R}_0}) \\
&= H_0^{(f)}(\omega_1, \omega_2) H_0^{(f)}(\omega_1 + \omega_2, -\omega_1 + \omega_2) \quad (17) \\
&\quad \times H_1^{(f)}(2\omega_1, 2(\omega_1 + \omega_2)),
\end{aligned}
$$

where the frequency supports of $H_0^{(f)}(\omega_1, \omega_2)$, $H_0^{(f)}(\omega_1 + \omega_2, -\omega_1 + \omega_2)$, and $H_1^{(f)}(2\omega_1, 2(\omega_1 + \omega_2))$ are plotted in Figure 9(a). The dark, white, and gray areas correspond to the passband, stopband, and transition band, respectively. Assume that these filters have reasonably good frequency responses, that is, flat passband and stopband and narrow transition band. The resulting filter $H_0(\omega_1, \omega_2)$ will have its frequency support as in Figure 9(b), of which the transition bands include the areas $\omega_2 \approx \pm\pi$. By similar geometric arguments, one can show that half of the directional filters in the DFB have transition band in the high-frequency areas $\omega_2 \approx \pm\pi$; while the other half have transition band in the areas $\omega_1 \approx \pm\pi$ (see Figure 10(b)). Therefore, the impulse responses of directional filters of a DFB have a strong directional component and a high-frequency vertical (or horizontal) component. This highpass aliasing component is
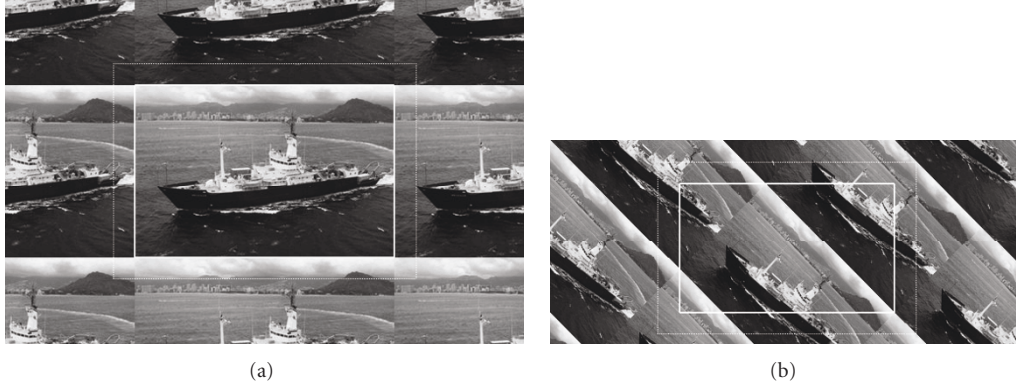
FIGURE 7: Boat image of size 200×300 (a) periodic extension of the original image and (b) after downsample by $\mathbf{Q}_0$.
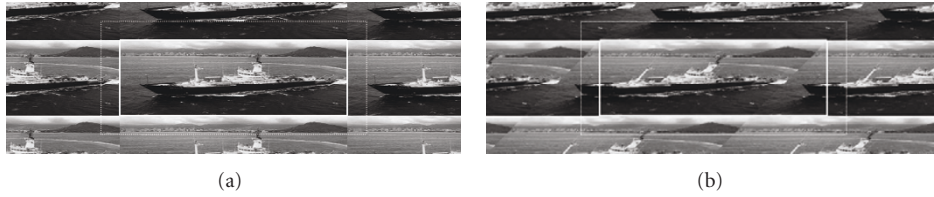

FIGURE 8: Boat image obtained from the original image of size 100×300 with periodic extension by downsampling using matrices (a) $\mathbf{D}_0$ and (b) $\mathbf{P}_0$.

not very strong, but still visible (see the leftmost image in Figure 16(b)).

This highpass aliasing problem is more pronounced in the directional filters of the second resolution of the PDFB if the lowpass filter of the Laplacian pyramid in the PDFB does not satisfy the Nyquist criterion with respect to decimation matrix $\mathbf{D}_2$, that is, if its passband and transition band are not restricted within $[-\pi/2, \pi/2]^2$. This condition will be discussed in Section 3. The equivalent directional filters of the PDFB in [5] at the second level of the pyramid are plotted in Figure 9(c), where the aliasing at high frequency is displayed very clearly.

The stopband aliasing problem of the DFB will appear when the number of directional bands increases. The frequency responses of the directional filters will have some peaks in the stopband regions. This is because when more fan filters with decimation blocks are cascaded, their transition bands will overlap. The two aliasing problems of the directional filters generated by the DFB tree are illustrated in Figure 10.

3. THE PYRAMIDAL DFB FOR THE CONTOURLET TRANSFORM

The PDFB (or contourlet FB) is created by combining the Laplacian pyramid and the DFB with 2^n orientational subbands [5]. It is shown in this section that the combination of a Laplacian pyramid and a four-band DFB is equivalent to an overcomplete five-band FB. Let us denote the two two-dimensional lowpass filters in the Laplacian pyramid in

Figure 11 as $G(\mathbf{z})$ and $F(\mathbf{z})$. The first filter can be written in type I polyphase form as follows:

$$
\begin{aligned}
G(\mathbf{z}) &= G^{(0)}(\mathbf{z}^{\mathbf{D}_2}) + z_1^{-1} G^{(1)}(\mathbf{z}^{\mathbf{D}_2}) \\
&\quad + z_2^{-1} G^{(2)}(\mathbf{z}^{\mathbf{D}_2}) + z_1^{-1} z_2^{-1} G^{(3)}(\mathbf{z}^{\mathbf{D}_2}) \\
&= \mathbf{g}^T(\mathbf{z}^{\mathbf{D}_2}) \mathbf{e}(\mathbf{z}),
\end{aligned} \tag{18}
$$

where $\mathbf{g}(\mathbf{z}) = [G^{(0)}(\mathbf{z}), G^{(1)}(\mathbf{z}), G^{(2)}(\mathbf{z}), G^{(3)}(\mathbf{z})]^T$, and $\mathbf{e}(\mathbf{z}) = [1, z_1^{-1}, z_2^{-1}, z_1^{-1} z_2^{-1}]^T$. Similarly, let $\mathbf{f}(\mathbf{z})$ be the column matrix of the (type II) polyphase components of the interpolation filter $F(\mathbf{z})$, that is,

$$
\begin{aligned}
F(\mathbf{z}) &= F^{(0)}(\mathbf{z}^{\mathbf{D}_2}) + z_1 F^{(1)}(\mathbf{z}^{\mathbf{D}_2}) \\
&\quad + z_2 F^{(2)}(\mathbf{z}^{\mathbf{D}_2}) + z_1 z_2 F^{(3)}(\mathbf{z}^{\mathbf{D}_2}) \\
&= \mathbf{f}^T(\mathbf{z}^{\mathbf{D}_2}) \mathbf{e}(\mathbf{z}^{-1}).
\end{aligned} \tag{19}
$$

The detailed output of the Laplacian pyramid is denoted as $d(\mathbf{n})$ in Figure 11(a). If one considers the four polyphase components of $d(\mathbf{n})$ as the four outputs of an FB with input $x(\mathbf{n})$ then it can be shown that the polyphase matrix of that FB is $\mathbf{I} - \mathbf{f}(\mathbf{z})\mathbf{g}^T(\mathbf{z})$.

In the PDFB depicted in Figure 11(a), a four-band DFB is applied to the detailed signal $d(\mathbf{n})$. Let $\mathbf{E}(\mathbf{z})$ be the polyphase matrix of the four directional filters of the four-band DFB, as defined in (12)–(15), that is,

$$
[H_1(\mathbf{z}), H_2(\mathbf{z}), H_3(\mathbf{z}), H_4(\mathbf{z})]^T = \mathbf{E}(\mathbf{z}^{\mathbf{D}_2}) \mathbf{e}(\mathbf{z}), \tag{20}
$$

where $\mathbf{e}(\mathbf{z}) = [1, z_1^{-1}, z_2^{-1}, z_1^{-1} z_2^{-1}]^T$. The input $x(\mathbf{n})$ goes through the Laplacian pyramid and the four-band DFB to
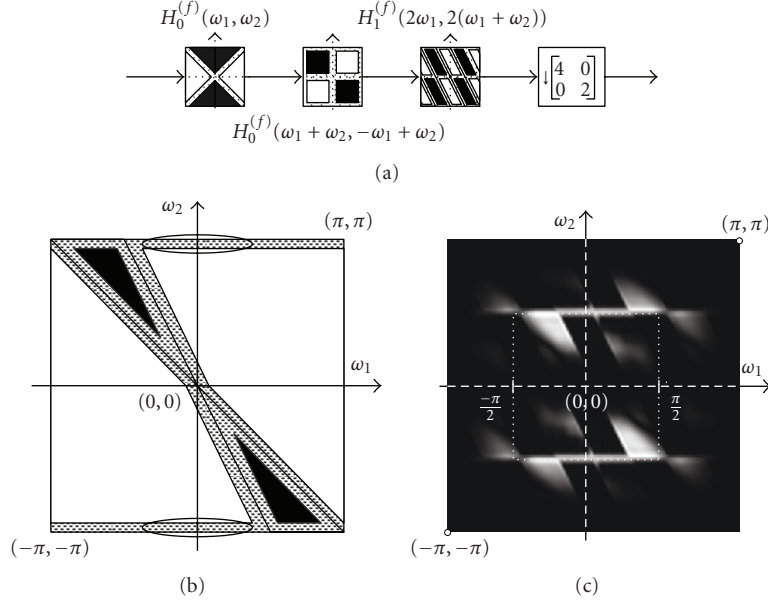
FIGURE 9: Aliasing effect on DFBs obtained by using the tree structure: (a) the equivalent structure to obtain subband 0 in eight-band DFB structure in Figure 2, (b) the equivalent directional filter $H_0(\omega_1, \omega_2)$. Black, gray and white colors denote passband, transition band, and stopband, respectively, and (c) an example of the amplitude of frequency response of a contourlet basis at the second resolution level, displayed as a gray-scale image.
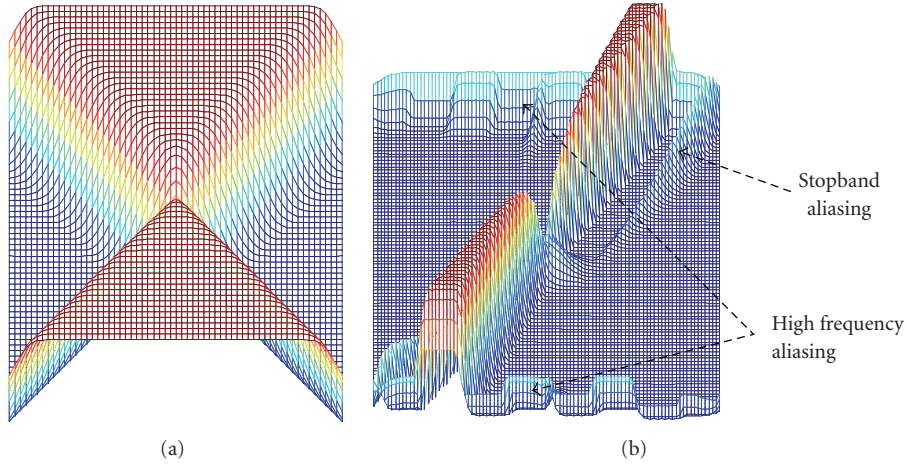

FIGURE 10: The frequency responses of (a) a fan filter with very small transition band, and (b) the directional band of a sixteen-band DFB created from the fan filter of (a).

produce four subsampled outputs $y_i(\mathbf{n})$, $i = 1, 2, 3$, and 4. These four signals can be considered as the outputs of the analysis side of an FB with input $x(\mathbf{n})$, and it can be shown that the polyphase matrix of this overall FB is $\tilde{E}(\mathbf{z}) = \mathbf{E}(\mathbf{z})(\mathbf{I} - \mathbf{f}(\mathbf{z})\mathbf{g}^T(\mathbf{z}))$.

Therefore, the PDFB in Figure 11(a) is equivalent to the five-band analysis FB in Figure 11(b) with the same lowpass filter $G(\mathbf{z})$ and four directional filters $\tilde{H}_i(\mathbf{z})$ which are given as

$$\left[\tilde{H}_1(\mathbf{z}), \tilde{H}_2(\mathbf{z}), \tilde{H}_3(\mathbf{z}), \tilde{H}_4(\mathbf{z})\right]^T = \tilde{\boldsymbol{E}}(\mathbf{z}^{\mathbf{D}_2})\mathbf{e}(\mathbf{z}). \qquad (21)$$

3.1. Aliasing effect of the Laplacian pyramid

In order for the PDFB to achieve its potential performance, it is necessary that the equivalent directional filters have excellent frequency responses. In general, the construction of the PDFB consists of a separable Laplacian pyramid and a binary-tree conventional DFB [7]. Thus, there are two sources of aliasing that will be considered in this section: those on the DFB tree, and those caused by the pyramid structure.

The equivalent directional filters $\tilde{H}_i(\mathbf{z})$ in (21) can be expressed in terms of the lowpass filters $G(\mathbf{z})$, $F(\mathbf{z})$ and the four highpass filters $H_i(\mathbf{z})$. Let us consider a realization of
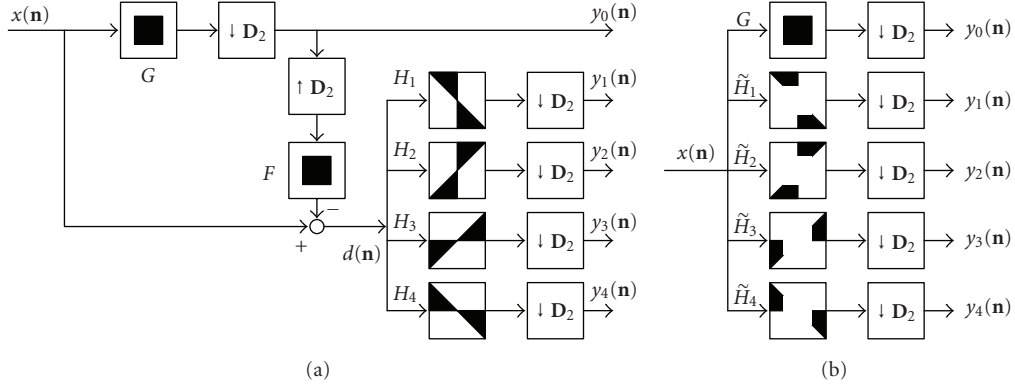
(a)

(b)

FIGURE 11: The four-band PDFB (a) the analysis side of the PDFB, (b) equivalent overcomplete FB, and (c) the synthesis FB.
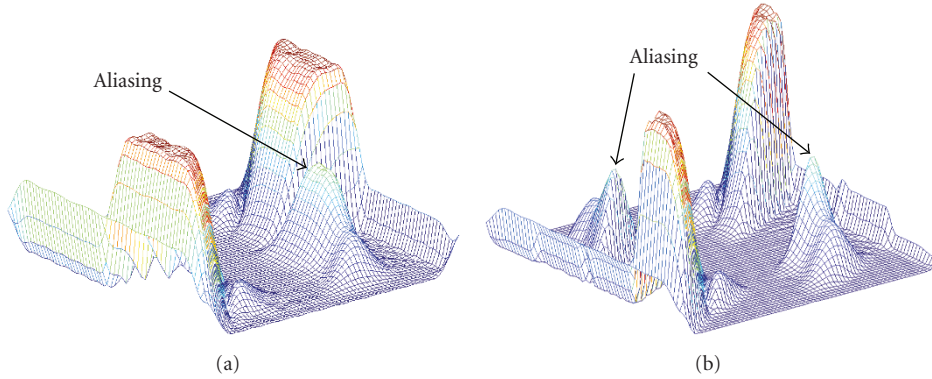

(a)

(b)

FIGURE 12: Examples of equivalent directional filters in the PDFB (a) the equivalent filter $\tilde{H}_1(\omega_1, \omega_2)$ of four-band PDFB, (b) a directional filter of eight-band PDFB, function associated with filter $\tilde{H}_1(\mathbf{z})$.

the PDFB constructed using the "9-7" biorthogonal filters as the lowpass filters $G(\mathbf{z})$ and $F(\mathbf{z})$ as in [5]. The fan FBs in the DFB tree structure are implemented using the ladder structure (with filter lengths of 21 and 41) in [23]. The first directional filter for the case of four-band DFB is plotted in Figure 12(a). It is observed that the directional filters have "bumps" in the stopband region. Similarly, an eight-band DFB can be obtained by cascading one more step of two-channel filter banks at the binary tree. Its first directional filter is presented in Figure 12(b) showing more bumps in the stopband. It will be shown later that this effect is due to aliasing resulting from decimation and interpolation of the Laplacian pyramid, and the heights of these peaks are independent from the directional filters in the DFB.

Let $FH_1(\mathbf{z}) = F(\mathbf{z})H_1(\mathbf{z})$ be written in a polyphase form as

$$FH_1(\mathbf{z}) = FH_1^{(0)}(\mathbf{z}^{\mathbf{D}_2}) + z_1^{-1}FH_1^{(1)}(\mathbf{z}^{\mathbf{D}_2}) \\ + z_2^{-1}FH_1^{(2)}(\mathbf{z}^{\mathbf{D}_2}) + z_1^{-1}z_2^{-1}FH_1^{(3)}(\mathbf{z}^{\mathbf{D}_2}), \quad (22)$$

where $FH_1^{(i)}(\mathbf{z})$ is the polyphase components of $FH_1(\mathbf{z})$. By some manipulation, it can be shown that the block diagrams in Figures 14(a) and 14(b) are equivalent, where $FH_1^{(0)}(\mathbf{z})$ is the first polyphase component of $FH_1(\mathbf{z})$. Consider the signal $y_1(\mathbf{n})$ in Figure 11(a). The corresponding block diagram

can be redrawn as in Figure 13(a), where the subsystem in the dotted rectangle is equivalent to $FH_1^{(0)}(\mathbf{z})$. Using the noble identities, the top path in Figure 13(a) can be further simplified as in Figure 13(b).

Since the PDFB is realized by FIR filters, the filters are not ideal, their frequency responses have transition bands. Figure 15(a) shows the passband support of $FH_1(\mathbf{z})$. The resulting filter $FH_1^{(0)}(\mathbf{z}^{\mathbf{D}_2})$ in Figure 13(b) is obtained by downsampling followed by upsampling the filter $FH_1(\mathbf{z})$ by $\mathbf{D}_2$. The corresponding frequency response $FH_1^{(0)}(\mathbf{D}_2^T\boldsymbol{\omega})$ whose supports are displayed in Figure 15(b) can be given by

$$FH_1^{(0)}(\mathbf{D}_2^T\boldsymbol{\omega}) = \frac{1}{|\mathbf{D}_2|}\sum_{\mathbf{k}\in\mathcal{N}(\mathbf{D}_2^T)} FH_1(\boldsymbol{\omega} - 2\pi\mathbf{D}_2^{-T}\mathbf{k}). \quad (23)$$

Therefore,

$$\tilde{H}_1(\boldsymbol{\omega}) = H_1(\boldsymbol{\omega}) - \frac{1}{|\mathbf{D}_2|}G(\boldsymbol{\omega}) \\ \times \sum_{\mathbf{k}\in\mathcal{N}(\mathbf{D}_2^T)} F(\boldsymbol{\omega} - 2\pi\mathbf{D}_2^{-T}\mathbf{k})H_1(\boldsymbol{\omega} - 2\pi\mathbf{D}_2^{-T}\mathbf{k}). \quad (24)$$

Assuming that the lowpass filter $G(\boldsymbol{\omega})$ is approximately zero in its stopband regions, two of the aliasing terms in the
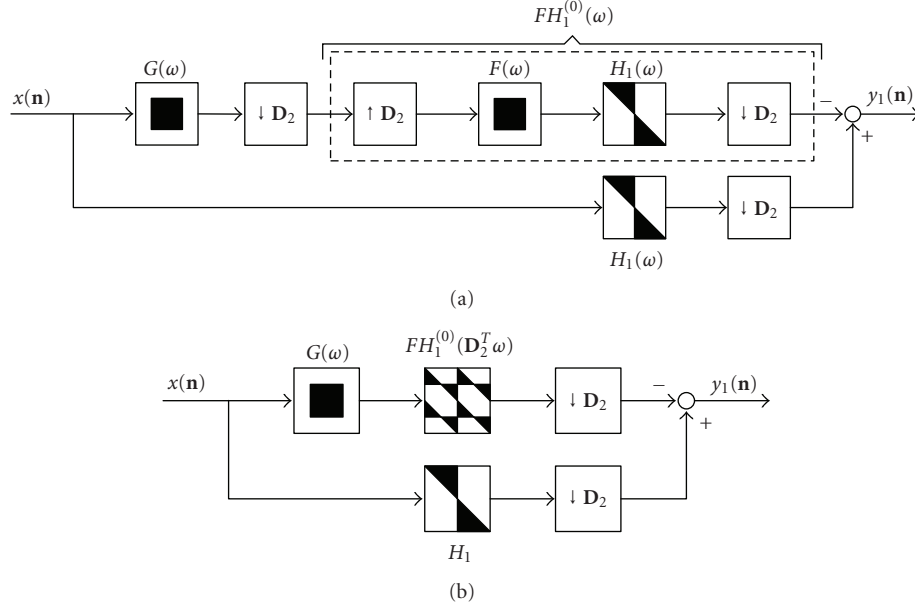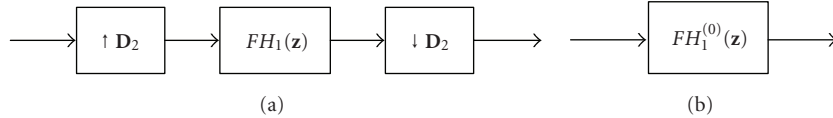
(a)

(b)

FIGURE 13: The equivalent structure to the directional filter $\widetilde{H}_1(\omega_1, \omega_2)$ in the four-band PDFB in Figure 11.

(a)

(b)

FIGURE 14: Equivalent block diagrams.

summation can be neglected, and the overall filter $\widetilde{H}_1(\omega)$ can be approximated by

$$
\widetilde{H}_1(\omega_1, \omega_2) \approx H_1(\omega_1, \omega_2)\left(1 - \frac{1}{|\mathbf{D}_2|}G(\omega_1, \omega_2)F(\omega_1, \omega_2)\right)
$$
$$
- \frac{1}{|\mathbf{D}_2|}G(\omega_1, \omega_2)F(\omega_1, \omega_2 - \pi)H_1(\omega_1, \omega_2 - \pi). \tag{25}
$$

The second term in (25) produces the peaks in stopband of $\widetilde{H}_1(\omega)$ (see Figure 12). The positions of these peaks are in the passband of the modulated directional filter $H_1(\omega_1, \omega_2 - \pi)$, so the only way to eliminate these "bumps" is to reduce the overlapping transition bands between $G(\omega_1, \omega_2)$ and $F(\omega_1, \omega_2 - \pi)$. Since both $G(\mathbf{z})$ and $F(\mathbf{z})$ are lowpass, and if the Nyquist sampling condition is satisfied, the above aliasing term will be cancelled. Therefore, the two filters in the Laplacian pyramid should satisfy the following conditions:

$$
G(\omega_1, \omega_2) \approx 0, \qquad F(\omega_1, \omega_2) \approx 0
$$
$$
\text{when } |\omega_1| > \frac{\pi}{2} \quad \text{or} \quad |\omega_2| > \frac{\pi}{2}. \tag{26}
$$

This means that the cutoff frequency of $G(\mathbf{z})$ and $F(\mathbf{z})$ must be a little less than $\pi/2$ in order to keep approximately zero response beyond $\pi/2$. For convenience, we call the conventional PDFB (with cutoff frequency at $\pi/2$) *aliasing* PDFB,

and the one satisfying the above constraints *nonaliasing* PDFB.

Recently, a new contourlet FB is presented in [28]. The new FB replaced the Laplacian pyramid by a multiresolution FB implemented in frequency domain. This new construction of contourlet FB removed the aliasing problems by requiring the lowpass filters to have frequency support strictly restricted in $[-\pi/2, \pi/2]^2$ region. In order for the new contourlet FB to be PR, its multiresolution FB has to be implemented in frequency domain, which implies periodic extension for size-limited images. This will lead to border artifact problem for image coding application.

3.2. Reduced aliasing pyramidal directional filter bank

In order to demonstrate the aliasing effect from the directional filters in the PDFB, the impulse responses of the overall directional filters at different scales are compared in Figure 16. Figure 16(b) shows the impulse responses at four scales as highlighted in Figure 16(a) for the case of aliasing PDFB, where the lowpass filters $G(\mathbf{z})$ and $F(\mathbf{z})$ are obtained by the "9-7" biorthogonal FB. Similarly, Figure 16(c) shows the impulse responses for the case of nonaliasing PDFB, where the filters $G(\mathbf{z})$ and $F(\mathbf{z})$ are designed so that the transit bands are between $\pi/4 < |\omega_i| < \pi/2$ and the frequency responses are approximately zero when $\pi/2 < |\omega_i|$. The DFB in both aliasing and nonaliasing PDFBs is identically
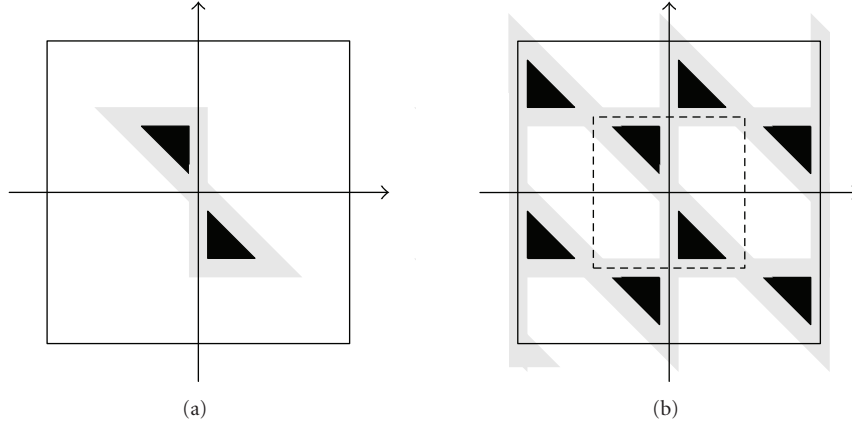
<p align="center">(a)   (b)</p>

FIGURE 15: The frequency supports of (a) $FH_1(\boldsymbol{\omega})$ and (b) $FH_1^{(0)}(\mathbf{D}_2^T\boldsymbol{\omega})$.

realized by a binary tree of two-channel fan FBs, which are implemented by a two-step ladder structure [23]. Both decompositions have 32, 16, 8, and 4 directional subbands at the first, second, third, and fourth resolutions, respectively, as displayed in Figure 16. It is evident that the filters in Figure 16(c) have much less aliasing and better directionality than those in Figure 16(b).

In practice, most of the unwanted aliasing components considered in the previous section can be reduced if the two lowpass filters in the pyramid have slightly smaller passband. Figures 17(a) and 17(b) show examples of the frequency responses of the first directional filters of the four-band and eight-band PDFBs whose $F(\mathbf{z})$ and $G(\mathbf{z})$ are designed to have a transition band $0.3\pi < |\omega_i| < 0.6\pi$. Comparing to the frequency responses presented in Figures 12(a) and 12(b), it is clear that those aliasing bumps have been significantly suppressed.

4. IMAGE CODING USING HYBRID OVERCOMPLETE PDFB AND WAVELET FB

Although the equivalent directional filters of the PDFB (or contourlet basis) are efficient in representing image contours, its performance tends to be lower than that of the traditional discrete wavelet transform (DWT) when the image sizes get smaller. In our nonlinear approximation experiment [20] with typical testing images (Lena and Barbara), the best results are achieved when the PDFB is used at higher resolution, and the DWT is used when the image sizes are less than or equal to 128×128. For the image coding application in this paper, we use a five-level image decomposition: the PDFB is used at the highest two resolutions, and the DWT is used in the next three resolutions. Figure 18(a) depicts the frequency supports of the hybrid FB. The Barbara image and its decomposition by the hybrid FB are illustrated in Figures 18(b) and 18(c), respectively.

Since significant coefficients of the PDFB are sparser than those of wavelet, utilizing the statistical properties of the transform coefficients is a crucial task in the design of high-performance PDFB-based coders. An explicit way to solve this problem is to classify the coefficients of each subband

into two subsets that separate insignificant and significant coefficients. Thus, the quantization and estimated models can be adapted to each subset independently.

A potential approach to exploit this behavior has been introduced with the morphological operation for wavelet coding [29]. Based on the observation that clusters tend to grow in both spatial and frequency domain, the previously detected significant coefficients are used as seeds for the search of new significant ones. The clustering trend of significant coefficients also exists in the PDFB bands. This suggests using a morphological dilation to identify the significant coefficients in the PDFB subbands before the coding step. Thus, different probability models can be estimated for significant and insignificant coefficients separately. On the other hand, although the overcomplete transform introduces more coefficients to be coded, strong correlations exist between neighboring directional subbands that can be adopted to improve probability modeling.

4.1. Progressive morphological dilation

Dilation is a morphological operator frequently adopted [30]. Let I be a binary-valued image where dilation will be applied. Dilation of a given set $A \subseteq I$ with set $B \subseteq I$ is defined by [31]

$$A \oplus B = \bigcup_{b \in B} A_b, \tag{27}$$

where B is a binary-valued array called a *structuring element* (SE), A_b denotes the translation of A to a point b. The dilation operation produces an enlarged set, $A \oplus B$, which can also be written as $A \cup (A \oplus B \setminus A)$, where $(A \oplus B \setminus A)$ represents the set of new points obtained by dilation. If A is the set of previously detected significant points, the points in set $(A \oplus B \setminus A)$ have a much higher probability to be recognized as significant.

Organizing and representing each subband of PDFB as irregularly shaped clusters of significant coefficients provide an efficient way for accurate probability modeling. Exploiting cross-scale dependency and correlation between neighboring
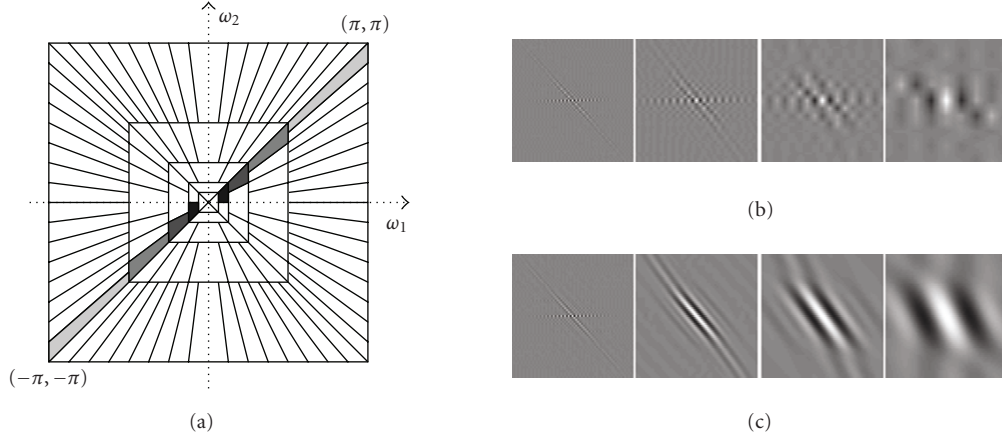
FIGURE 16: (a) The essential frequency supports of directional filters in the PDFB at different scales: level 1 with 32 bands, level 2 with 16 bands, level 3 with 8 bands, and level 4 with 4 bands. The corresponding impulse responses in case of (b) aliasing Laplacian pyramid filters and (c) nonaliasing Laplacian pyramid filters.
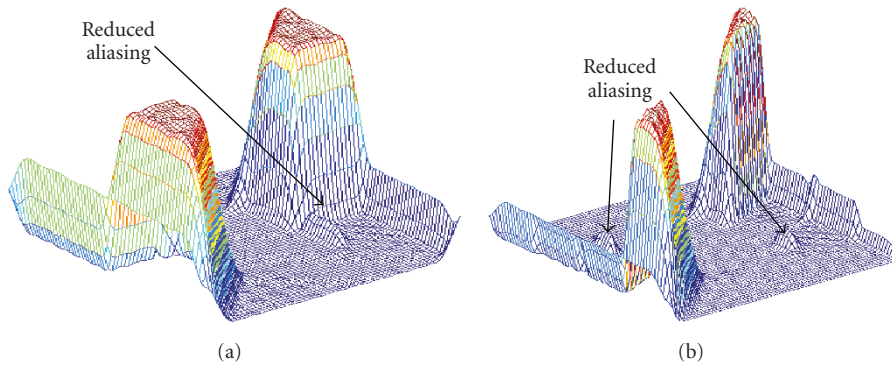

FIGURE 17: Examples of directional filters in the PDFB (a) $\tilde{H}_1(\omega_1, \omega_2)$ of a four-band PDFB and (b) a directional filter of an eight-band PDFB.

DFB bands of the overcomplete pyramid can further improve the detection accuracy of significant coefficients. Moreover, shaping the clustering boundaries with less cost should be also considered with adaptive SE's [32]. A progressive morphological dilation method is proposed to find the correlations among PDFB subbands. Six passes are performed at each bit-plane. Three different SE's, as shown in Figure 19, are adopted for variant steps of dilation.

(1) *Significance detection* (SD) pass: the intraband dependency is exploited in this pass. The square SE_1 in Figure 19(a) is adopted to detect new significant coefficients based on significant neighbors in the previous bit-planes.

(2) *Cross-band prediction* (CBP) pass: based on the fact that the interband correlation exists between the successive decomposition levels, a significant cluster in a children subband can be predicted by those in the parent subband. The diamond SE_2 in Figure 19(b) is employed to dilate around the associated children coefficients corresponding to each significant coefficient in the parent subband.

(3) *Neighboring correlation prediction* (NCP) pass: this pass is designed to capture the redundancy among the overcomplete directional subbands. If the two spatial filters have small angle difference in their principal directions, their corresponding decimated subbands exhibit significant dependency between those coefficients at the same positions relative to their upper-left corners. Hence, the distribution of significant coefficients of each directional subband is highly correlated to their neighboring subbands (cousin subbands). For the current subband, those coefficients associated with the identified significant coefficients in the cousin subbands are evaluated using SE_1 to find new significant coefficients.

(4) *Boundary shaping* (BS) pass: typically, on the boundaries of a large cluster, there are a few scattered significant coefficients located in small isolated clusters. It is difficult to forecast the dimension of these isolated clusters. Hence, adaptive dilation is expected to search around a cluster. The smaller rood SE_3 in Figure 19(c) is adopted for dilation on the previously formed cluster boundaries which are identified by the
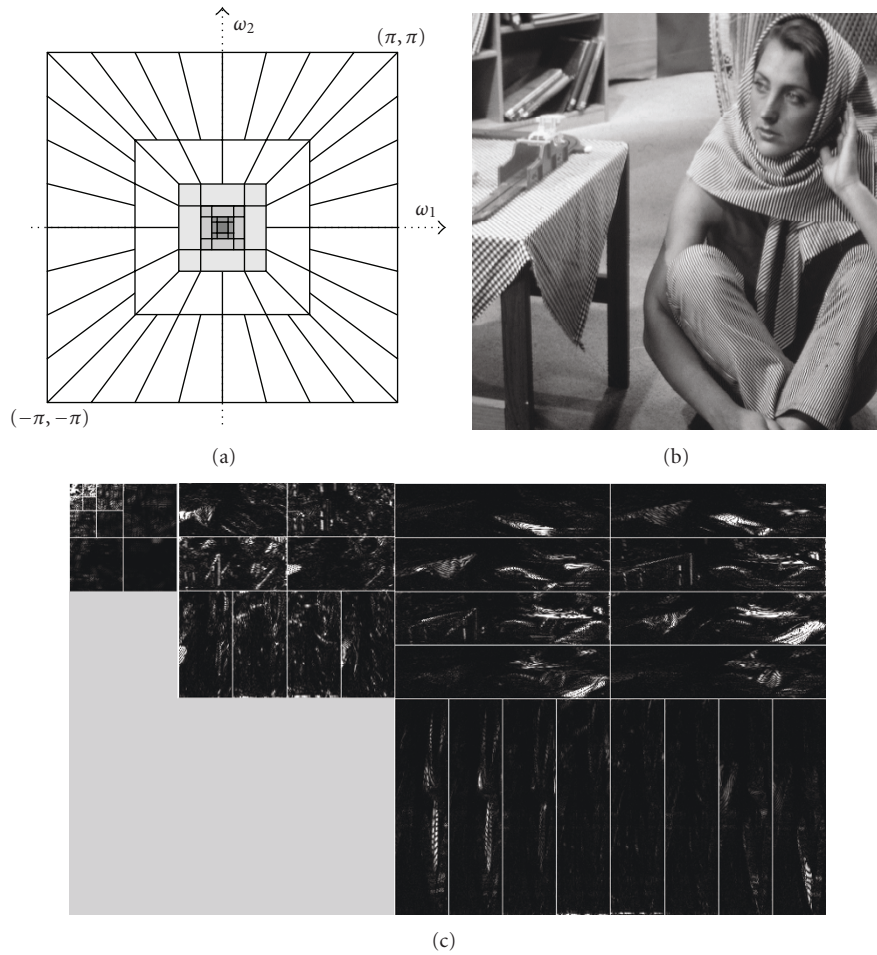
(a)

(b)

(c)

FIGURE 18: The PDFB employed in image coding. (a) The partitioning of the frequency plane by a hybrid PDFB and wavelet decomposition, (b) the Barbara testing image, and (c) the image representation by the five-level multiscale FB having the frequency supports in (a).
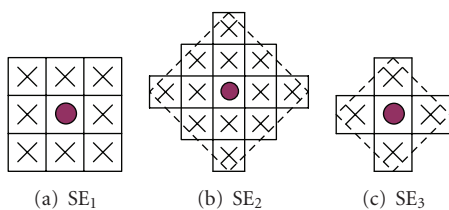

(a) SE_1 (b) SE_2 (c) SE_3

FIGURE 19: Structuring elements used in the progressive morphological dilation.

FIGURE 20: Context modeling template.

insignificant coefficients. The boundary extension is adaptively controlled based on the occurrence of new significant positions detected. It stops when the recursive dilation results in no more new significant coefficients.

(5) *Sparse significant coefficients detection* (SSD) pass: although most of the significant coefficients have been recognized by the previous four passes, there are still few sparse significant coefficients remained undetected. Those coefficients that have not been

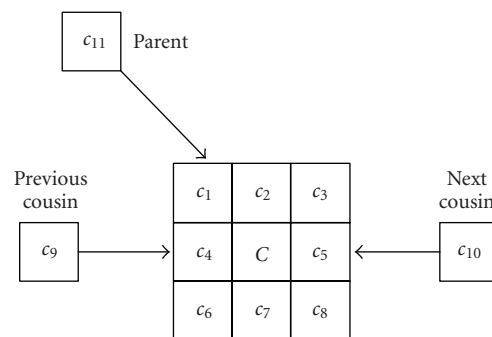processed are scanned and coded in a raster order. The dilation using SE_1 is implemented if a new significant coefficient is found.

(6) *Magnitude refinement* (MR) pass: refine those significant coefficients that have been recognized in previous bit-planes.

FIGURE 21: Reconstructed Barbara images at 0.15 bpp using (a) JPEG2000, PSNR = 25.93 dB and (b) PDFB, PSNR = 26.74 dB. (c) and (d) are zoom-in images of (a) and (b), respectively.

4.2. Context modeling

According to the conditional probability theory, context modeling can exploit the intersymbol redundancy by switching between different probability models [33]. In embedded coding systems, the coding is conducted on a series of significant maps that correspond to a set of decreasing thresholds. After each pass of coding, all coefficients are quantized to specific values and can be used as context information. Thus, a significant context template, as described in Figure 20, is defined to exploit the coefficients that have been coded by the previous passes or bit-planes.

In Figure 20, C is the current symbol to be coded. $c_1, \ldots, c_{11}$ are reconstructed values up to the current coding pass of the corresponding coefficients. We define a series of reconstruction matrices $\mathbf{Y}_s^n$, in which each element is the reconstructed value of the corresponding coefficient, n indicates the order of coding passes, s denotes the subband index. Assume C is the (i, j)th element of the binary significance map in pass n of subband s. Thus, the intraband correlation is derived from the 8 neighboring coefficients,

$$c_m = y_s^{n'}(i + a, j + b), \quad a, b = -1, 0, \text{ or } 1, \quad m = 1, \ldots, 8,$$

$$(28)$$

where $n' = n - 1$ or n since the morphological dilation does not guarantee a raster scan order for the coefficients that have

been coded. The coefficients of parent and cousin subbands are also used to exploit the interband dependency

$$c_9 = y_{s-1}^n(N_c(i), N_c(j)),$$
$$c_{10} = y_{s+1}^{n-1}(N_c(i), N_c(j)), \tag{29}$$
$$c_{11} = y_{P(s)}^n(P_c(i), P_c(j)),$$

where $P(\cdot)$ specifies the parent of band s, $N_c(\cdot)$ and $P_c(\cdot)$ are functions for the cousin and parent subbands, respectively, that specify the coordinates corresponding to the positions in the current subband. Considering the weak correlation among the cousin subbands for wavelet, c_9 and c_{10} are not used for coding wavelet coefficients in the simulation. Thus, the NCP pass is executed only for those directional subbands.

4.3. Algorithm summary

In this section, we summarize the coding algorithm. Six passes are designed in this approach. The produced progressive bitstream can be truncated at any pass. Let $w_{ij}(k)$ be the coefficient with the coordinate (i, j) relative to the upper-left corner of subband k. $LSC(k)$ and $LIC(k)$ represent the lists of significant and insignificant coefficients, respectively. $D(\text{SE}_m)$, $m = 1, 2, 3$, denotes the operation of morphological dilation with the structuring element specified by SE_m. $V[w_{ij}(k)|\text{SE}_m] = w_{ij}(k) \oplus \text{SE}_m \setminus w_{ij}(k)$,

defines the vicinity of $w_{ij}(k)$ generated by dilation with SE_m. Let $P(k)$ and $N(k)$ be the parent and cousin bands of subband k, respectively. $C[w_{ij}(k)]$ and $N[w_{ij}(k)]$ denote the corresponding coefficients in $P(k)$ and $N(k)$ associated with $w_{ij}(k)$.

The PDFB-based coding algorithm is summarized as follows.

(1) (*Initialization*): decompose the image with PDFB. Find the maximum number of bit-planes M and set $n \leftarrow M$.

(2) (*SD pass*): if $n = M$, go to step (6).

Apply $D(SE_1)$ to each entry of $LSC(k)$. Encode all $w \in V[w_{ij}(k)|SE_1]$ that have not been scanned with arithmetic coding, if w is insignificant, attach it to the end of $LIC(k)$; else encode the sign of w, and add it to the end of $LSC(k)$.

(3) (*CBP pass*): if $k = 0$, that is, the LL band, go to step (5).

For the significant coefficients of $P(k)$, apply $D(SE_2)$ to each entry of $C[w_{ij}(P(k))]$. Similar to *SD pass*, encode each newly scanned coefficient and add these coefficients to the end of either $LSC(k)$ or $LIC(k)$ depending on whether they are significant or not.

(4) (*NCP pass*): if subband k is a wavelet band, go to step (5).

For the significant coefficients of $N(k)$, apply $D(SE_1)$ to each entry of $N[w_{ij}(N(k))]$. Encode each scanned coefficient. Add the coefficient to the end of $LSC(k)$ if it is found to be a new significant coefficient; otherwise append it to $LIC(k)$.

(5) (*BS pass*): apply $D(SE_3)$ to each entry of $LIC(k)$. If a new significant coefficient is found, the recursive dilation is implemented around it until no more new significant coefficient is detected. Encode the scanned coefficients and update $LSC(k)$ and $LIC(k)$ correspondingly.

(6) (*SSD pass*): the remaining coefficients are scanned and coded in raster order. However, if one significant coefficient is found, apply $D(SE_3)$ at once, then continue to scan the following coefficients that have not been coded. Update $LSC(k)$ correspondingly.

(7) (*MR pass*): if $n = M$, go to step (8).

Encode the nth bit of those significant coefficients recognized by previous bit-planes.

(8) (*New bit-plane*): empty all LIC's, set $n \leftarrow n - 1$, and go back to step (2).

4.4. Image coding results

For the sake of comparison, the proposed PDFB-based algorithm is compared with three state-of-the-art wavelet-based coding schemes: SPIHT [34] with arithmetic coding, JPEG2000 [35], and the embedded coder of morphological representation of wavelet data (MRWD) [29]. Two 512×512

TABLE 1: Performance comparison (PSNR [dB]) for Barbara.

Rate (bpp)	SPIHT	MRWD	JPEG2000	PDFB
0.10	24.24	24.15	24.66	25.24
0.15	25.63	25.32	25.93	26.74
0.20	26.63	26.86	27.31	27.84
0.25	27.56	27.51	28.36	28.94
0.30	28.54	28.16	29.24	29.86
0.40	30.09	30.18	30.83	30.97
0.50	31.38	31.31	32.26	32.42

TABLE 2: Performance comparison (PSNR [dB]) for Lena.

Rate (bpp)	SPIHT	MRWD	JPEG2000	PDFB
0.10	30.17	30.18	29.87	29.86
0.15	31.87	31.55	31.69	31.87
0.20	33.12	33.14	32.97	33.07
0.25	34.09	33.90	34.13	33.98
0.30	34.93	34.57	34.75	34.91
0.40	36.22	36.17	36.09	35.96
0.50	37.20	37.01	37.22	36.78

grayscale images, Barbara and Lena, are tested. A 5-level decomposition is used for SPIHT, MRWD, and JPEG2000 with "9/7" Daubechies wavelet filters. On the other hand, a 5-level PDFB decomposition, as described in Figure 18(a), is applied in the proposed scheme, where the four coarser levels are decomposed using "9/7" filters and the level with the finest resolution is decomposed into 16 directional subbands using PDFB.

Tables 1 and 2 show the performance of different coding systems on Barbara and Lena testing images in terms of PSNR. The performances are calculated by truncating the embedded bitstreams at different rates during decoding. In the comparison, the proposed algorithm consistently outperforms all the other three schemes. Although the over-complete transform is employed in the proposed scheme, the improvements are still remarkable for rich-edgy images. This is because the PDFB can represent the geometrical regularity of image structures with fewer coefficients and smaller magnitudes. Hence, the proposed algorithm is more efficient at very low bit-rates when a few of significant coefficients are actually used for reconstruction. For instance, the PDFB-based coder gains 0.81 dB in PSNR over JPEG2000 at 0.15 bpp for Barbara image. Figure 21 shows the reconstructed Barbara images at 0.15 bpp for JPEG2000 and the proposed coder. Moreover, with reduced aliasing artifact in the proposed PDFB structure, the textures and edges information are better preserved as shown in Figure 21. On the other hand, the PDFB-based codec exhibits competitive performance for smooth images, such as Lena. As shown in Figure 22, the performance of our codec is comparable to that of the other three coders, and the performance difference is getting smaller as the bit-rate decreases.
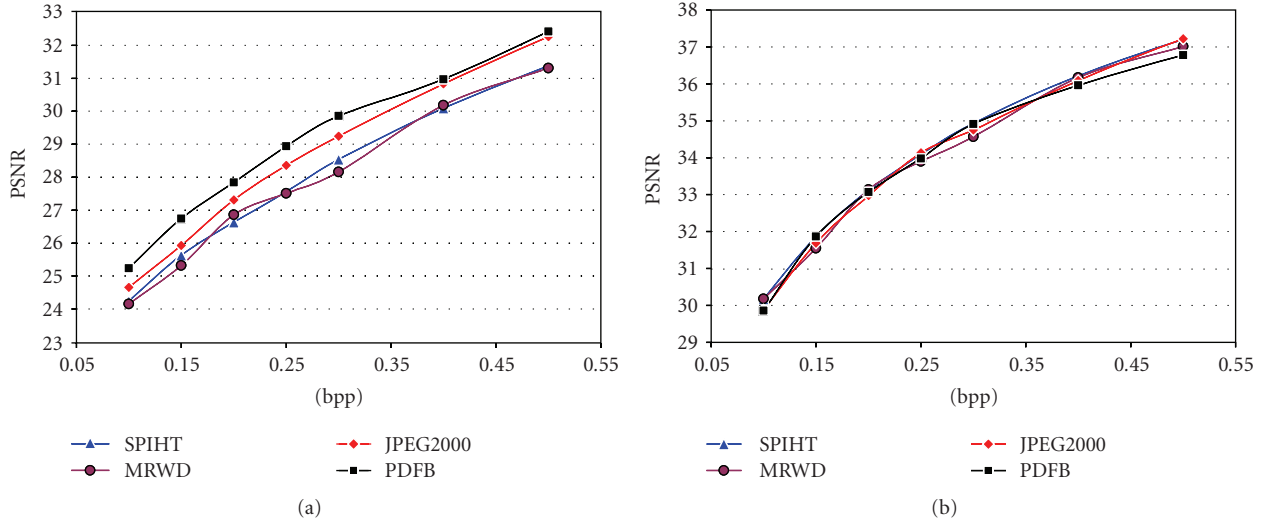
FIGURE 22: Comparison of rate-distortion curves using SPIHT, MRWD, JPEG2000, and the proposed PDFB-based codec: (a) Barbara, (b) Lena.

5. CONCLUSION

This work provides a detail examination of the contourlet FB implementation, including a modification to minimize phase delay and an explanation on image border extension along the DFB tree. The paper also discusses the PDFB for the contourlet transform as an overcomplete FB and shows that the equivalent directional filters of the PDFB suffer from the aliasing effects due to its implementation structure. By imposing the conditions that the two lowpass filters of the Laplacian pyramid should have frequency supports restricted within $[-\pi/2, \pi/2]^2$, the aliasing in the stopbands is removed. The equivalent directional filters of the PDFB have excellent passband and stopband characteristics, which is the key requirement for the contourlet to attain its asymptotic approximation power of smooth signal having discontinuities around $\mathbf{C}^2$ curve. We propose a novel embedded image coding scheme using hybrid overcomplete PDFB. The improved design of the PDFB helps to reconstruct images with directional features by fewer coefficients than other image decompositions. Based on the progressive morphological dilation of significant coefficients, the highly concentrated PDFB significant coefficients are identified by exploiting the intraband statistical dependencies. The correlation exists among the neighboring directional subbands is also adopted to reduce the high redundancy of the overcomplete transform. Experimental results justify that the proposed algorithm is superior to the state-of-the-art wavelet-based coders for rich-edgy images.

REFERENCES

[1] G. Strang and T. Q. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Wellesley, Mass, USA, 1996.

[2] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, Calif, USA, 2nd edition, 1999.

[3] M. Vetterli, "Wavelets, approximation, and compression," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 59–73, 2001.

[4] E. J. Candès and D. L. Donoho, "Curvelets: a surprisingly effective nonadaptive representation for objects with edges," in *Curve and Surface Fitting: Saint-Malo 99*, A. Cohen, C. Rabut, and L. L. Schumaker, Eds., pp. 105–120, Vanderbilt University Press, Nashville, Tenn, USA, 2000.

[5] M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2091–2106, 2005.

[6] P. J. Burt and E. H. Adelson, "The laplacian pyramid as a compact image code," *IEEE Transactions on Communications*, vol. 31, no. 4, pp. 532–540, 1983.

[7] R. H. Bamberger and M. J. T. Smith, "A filter bank for the directional decomposition of images: theory and design," *IEEE Transactions on Signal Processing*, vol. 40, no. 4, pp. 882–893, 1992.

[8] A. L. Da Cunha, J. Zhou, and M. N. Do, "The nonsubsampled contourlet transform: theory, design, and applications," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3089–3101, 2006.

[9] R. Eslami and H. Radha, "Translation-invariant contourlet transform and its application to image denoising," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3362–3374, 2006.

[10] T. T. Nguyen and S. Oraintara, "The shiftable complex directional pyramid—part 1: theoretical aspects," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4651–4660, 2008.

[11] K.-O. Cheng, N.-F. Law, and W.-C. Siu, "A novel fast and reduced redundancy structure for multiscale directional filter banks," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2058–2068, 2007.

[12] Y. M. Lu and M. N. Do, "Multidimensional directional filter banks and surfacelets," *IEEE Transactions on Image Processing*, vol. 16, no. 4, pp. 918–931, 2007.

[13] V. Chappelier, C. Guillemot, and S. Marinkovic, "Image coding with iterated contourlet and wavelet transforms," in

Proceedings of the International Conference on Image Processing (ICIP '04), vol. 5, pp. 3157–3160, Singapore, October 2004.

[14] H. Song, S. Yu, L. Song, and H. Xiong, "Contourlet image coding based on adjusted SPIHT," in *Proceedings of the 6th Pacific Rim Conference on Multimedia (PCM '05)*, Lecture Notes in Computer Science, pp. 629–640, Springer, Jeju Island, Korea, November 2005.

[15] R. Eslami and H. Radha, "On low bit-rate coding using the contourlet transform," in *Proceedings of the 37th Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1524–1528, Pacific Grove, Calif, USA, November 2003.

[16] A. Alecu, A. Munteanu, A. Piurica, J. Cornelis, and P. Schelkens, "On hybrid directional transform-based intra-band image coding," in *Advanced Concepts for Intelligent Vision Systems*, Lecture Notes in Computer Science, pp. 1049–1060, Springer, Delft, The Netherlands, 2007.

[17] M. N. Do and M. Vetterli, "Contourlets," in *Beyond Wavelets*, G. V. Welland, Ed., chapter 4, pp. 83–105, Academic Press, Amsterdam, The Netherlands, 2003.

[18] P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1993.

[19] S.-I. Park, M. J. T. Smith, and R. M. Mersereau, "Improved structures of maximally decimated directional filter banks for spatial image analysis," *IEEE Transactions on Image Processing*, vol. 13, no. 11, pp. 1424–1431, 2004.

[20] T. T. Nguyen and S. Oraintara, "On the aliasing effect of the contourlet filter banks," in *Proceedings of the 14th European Signal Processing Conference (EUSIPCO '06)*, Florence, Italy, September 2006.

[21] L. Yilong, T. T. Nguyen, and S. Oraintara, "Low bit-rate image coding based on pyramidal directional filter banks," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '06)*, vol. 2, pp. 437–440, Toulouse, France, May 2006.

[22] M. N. Do, *Directional multiresolution image representations*, Ph.D. thesis, Swiss Federal Institute of Technology, Lausanne, Switzerland, October 2001, http://www.ifp.uiuc.edu/~minhdo/publications/thesis.pdf.

[23] S.-M. Phoong, C. W. Kim, P. P. Vaidyanathan, and R. Ansari, "A new class of two-channel biorthogonal filter banks and wavelet bases," *IEEE Transactions on Signal Processing*, vol. 43, no. 3, pp. 649–665, 1995.

[24] K. S. C. Pun and T. Q. Nguyen, "A novel and efficient design of multidimensional PR two-channel filter banks with hourglass-shaped passband support," *IEEE Signal Processing Letters*, vol. 11, no. 3, pp. 345–348, 2004.

[25] J. S. Lim, *Two-Dimensional Signal and Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1990.

[26] D. D.-Y. Po and M. N. Do, "Directional multiscale modeling of images using the contourlet transform," *IEEE Transactions on Image Processing*, vol. 15, no. 6, pp. 1610–1620, 2006.

[27] H. C. Reddy, I.-H. Khoo, and P. K. Rajan, "2-D symmetry: theory and filter design applications," *IEEE Circuits and Systems Magazine*, vol. 3, no. 3, pp. 4–33, 2003.

[28] Y. Lu and M. N. Do, "A new contourlet transform with sharp frequency localization," in *Proceedings of the International Conference on Image Processing (ICIP '06)*, pp. 1629–1632, Atlanta, Ga, USA, October 2006.

[29] S. D. Servetto, K. Ramchandran, and M. T. Orchard, "Image coding based on a morphological representation of wavelet data," *IEEE Transactions on Image Processing*, vol. 8, no. 9, pp. 1161–1174, 1999.

[30] L. Vincent, "Morphological grayscale reconstruction in image analysis: applications and efficient algorithms," *IEEE Transactions on Image Processing*, vol. 2, no. 2, pp. 176–201, 1993.

[31] R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*, Addison-Wesley, Reading, Mass, USA, 1992.

[32] J. M. Zhong, C. H. Leung, and Y. Y. Tang, "Wavelet image coding based on significance extraction using morphological operations," *IEE Proceedings: Vision, Image and Signal Processing*, vol. 146, no. 4, pp. 206–210, 1999.

[33] X. Wu, "High-order context modeling and embedded conditional entropy coding of wavelet coefficients for image compression," in *Proceedings of the 31st Asilomar Conference on Signals, Systems & Computers*, vol. 2, pp. 1378–1382, Pacific Grove, Calif, USA, November 1997.

[34] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 3, pp. 243–250, 1996.

[35] T. Acharya and P. Tsai, *JPEG2000 Standard for Image Compression*, John Wiley & Sons, New York, NY, USA, 2005.