Directionlets: Anisotropic Multi-directional Representation with Separable Filtering

Vladan Velisavljević, *Student Member, IEEE*, Baltasar Beferull-Lozano, *Member, IEEE*, Martin Vetterli, *Fellow, IEEE*, and Pier Luigi Dragotti, *Member, IEEE*

Abstract

In spite of the success of the standard wavelet transform (WT) in image processing in recent years, the efficiency of its representation is limited by the spatial isotropy of its basis functions built in the horizontal and vertical directions. One-dimensional (1-D) discontinuities in images (edges and contours) that are very important elements in visual perception, intersect too many wavelet basis functions and lead to a non-sparse representation. To capture efficiently these anisotropic geometrical structures characterized by many more than the horizontal and vertical directions, a more complex *multi-directional* (M-DIR) and *anisotropic* transform is required. We present a new lattice-based *perfect reconstruction* and *critically sampled* anisotropic M-DIR WT. The transform retains the *separable* filtering and subsampling and the simplicity of computations and filter design from the standard two-dimensional (2-D) WT, unlike in the case of some other directional transform constructions (e.g. curvelets, contourlets or edgelets). The corresponding anisotropic basis functions (*directionlets*) have directional vanishing moments (DVM) along *any* two directions with rational slopes. Furthermore, we show that this novel transform provides an efficient tool for non-linear approximation (NLA) of images, achieving the approximation power $O(N^{-1.55})$, which, while slower than the optimal rate $O(N^{-2})$, is much better than $O(N^{-1})$ achieved with wavelets, but at similar complexity.

Index Terms

Wavelets, directionlets, multiresolution, multidirection, geometry, sparse image representation, filter-banks, separable filtering, directional vanishing moments

I. Introduction

The problem of finding efficient representations of images is a fundamental problem in many image processing tasks, such as denoising, compression and feature extraction. An efficient transform-based representation requires sparsity, that is, a large amount of information has to be contained in a small portion of transform coefficients.

V. Velisavljević and B. Beferull-Lozano are with the School of Computer and Communication Sciences, EPFL, CH-1015 Lausanne, Switzerland (emails: vladan.velisavljevic@epfl.ch, baltasar.beferull@epfl.ch).

M. Vetterli is with the School of Computer and Communication Sciences, EPFL, CH-1015 Lausanne, Switzerland and with the Department of Electrical Engineering and Computer Science, University of California, Berkeley CA 94720 (email: martin.vetterli@epfl.ch).

P. L. Dragotti is with the Department of Electrical and Electronic Engineering, Imperial College, London SW7-2AZ, UK (email: p.dragotti@imperial.ac.uk).

This work is supported by the Swiss National Foundation under Grant No. 200020-103729. P. L. Dragotti is in part supported by EPSRC under Grant No. GR/557631/01.

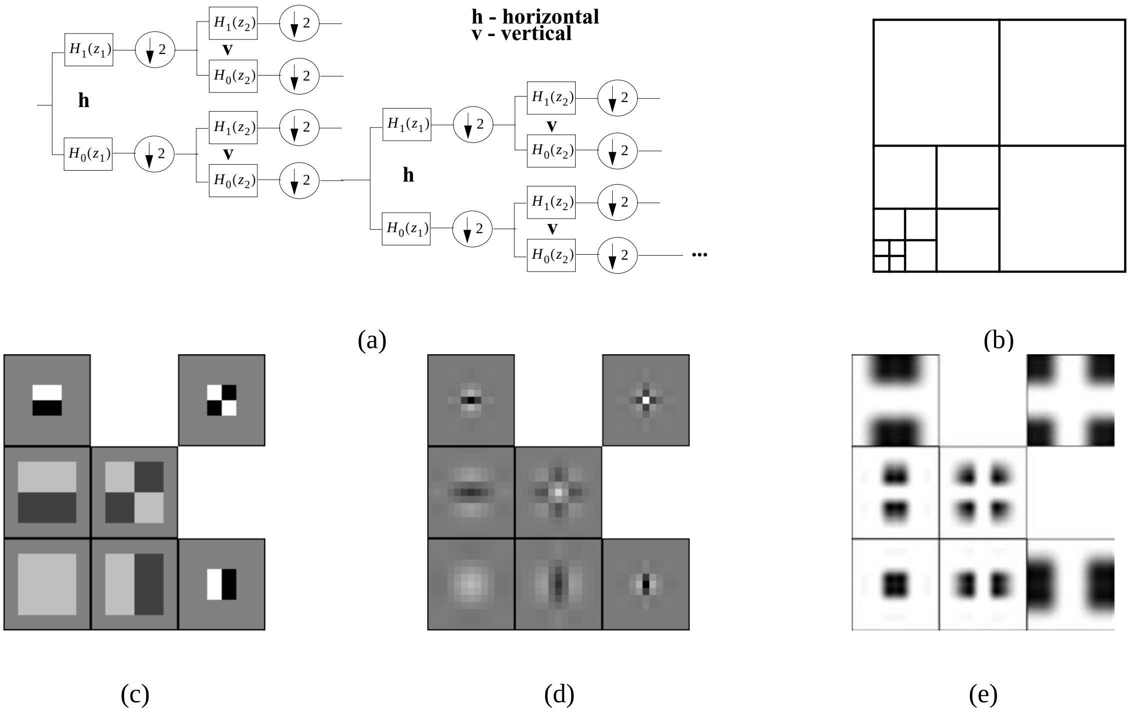
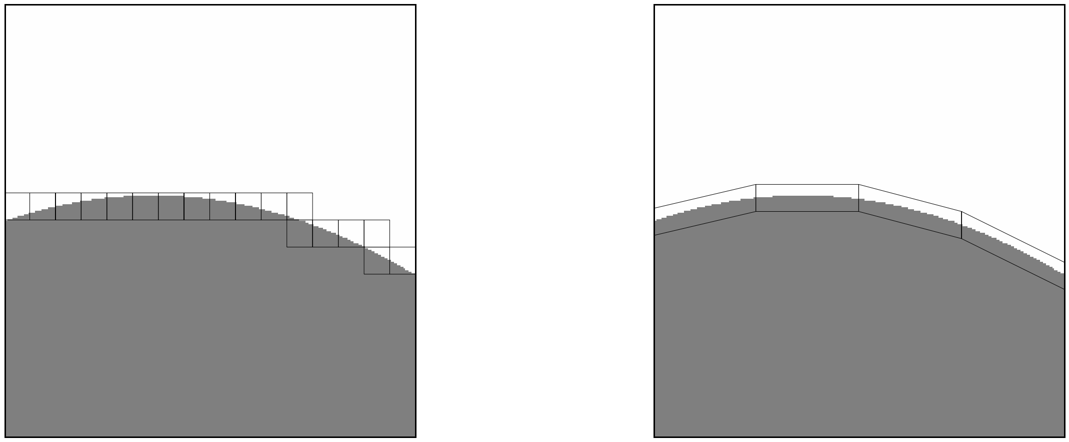


Fig. 1. The standard 2-D WT is isotropic. (a) The filtering and subsampling operations are applied equally in both directions at each scale of the transform. (b) The corresponding decomposition in frequency. The basis functions obtained in this way are isotropic at each scale as shown in (c) for Haar and in (d) for biorthogonal ”9-7” 1-D scaling and wavelet functions. (e) The corresponding Fourier transforms of the basis functions obtained from the ”9-7” 1-D filters.

The one-dimensional (1-D) WT has become very successful in the last decade because it provides a good multiresolution representation of 1-D piecewise smooth signals [1], [2]. The application of wavelets to image processing requires the design of two-dimensional (2-D) wavelet bases. The most common approach is to construct such bases using 2-D separable filter-banks, which consist of the direct product of two independent 1-D filter-banks in the horizontal and vertical directions. Filtering with high-pass (HP) filters with enough vanishing moments (or zeros at $\omega = 0$) along these two directions leads to a sparse representation of smooth signals. This method is conceptually simple and has very low complexity while all the 1-D wavelet theory carries over. These are the main reasons why it has been adopted in the image compression standard JPEG-2000 [3].

Some notable approaches use non-separable 2-D filter-banks and subsampling (e.g. quincunx) [4]–[6], but these methods are computationally complex and the design of the associated 2-D filter-banks is often challenging and involved. Also, several general multi-dimensional multi-channel filter design methods have been proposed in [7]–[10] resulting in filters with separable polyphase components. In this paper, we focus on the design and applications of 2-D separable two-channel filter-banks based on the 1-D wavelets, but allowing directionality and anisotropy.

Despite their success, the standard separable 2-D WT fails to provide a sparse representation in the presence of 1-D discontinuities, like edges or contours. These discontinuities, being highly anisotropic objects present in



(a) (b)

Fig. 2. A simple image with one discontinuity along a smooth curve is represented by the two types of basis functions: isotropic and anisotropic. The support of these basis functions is shown schematically as black rectangles. (a) Isotropic basis functions generate a large number of significant coefficients around the discontinuity. (b) Anisotropic basis functions trace the discontinuity line and produce just a few significant coefficients.

images, are characterized by a geometrical coherence that is not properly captured by the standard isotropic WT. Namely, many wavelets intersect a discontinuity and this leads to many large magnitude coefficients (Fig. 2(a)).

The reason for the inefficiency of the standard 2-D WT resides in the *spatial isotropy* of its construction, that is, filtering and subsampling operations are applied equally along both the horizontal and vertical directions at each scale (see Fig. 1(a)). As a result, the corresponding filters, obtained as direct products of 1-D filters, are isotropic at all scales (Fig. 1(c),(d)).

This motivates us to design *anisotropic basis functions* that can "match" anisotropic objects (Fig. 2(b)). However, ensuring an efficient matching between anisotropic basis functions and objects in images is a non-trivial task. Anisotropic basis functions have already been considered and exploited by adaptive (e.g. bandelets [11], [12]) or non-adaptive (edgelets and wedgelets [13]–[17], curvelets [18]–[20], contourlets [21], etc.) processing. These methods build dictionaries of anisotropic basis functions that provide a sparse representation of edges in images. Furthermore, Candès and Donoho [18] showed that the parabolic scaling relation between the length and width of basis functions is a key feature to achieve a good non-linear approximation (NLA) behavior. However, the implementation of these transforms usually requires *oversampling* having *higher complexity* when compared to the standard WT, and require *non-separable processing* (convolution) and *non-separable filter design*. Furthermore, in some of these constructions (e.g. curvelets [18]) the design of the associated filters is performed in the *continuous domain* and this makes it difficult to use them directly on discrete images and achieve perfect reconstruction.

Notice that the standard WT uses only horizontal and vertical directions and the HP filters in this transform have vanishing moments only along these directions. Since characterization of features in synthetic and natural images involves many more than these two standard directions, *multi-directionality* and *directional vanishing moments*

(DVM) play an important role in pursuing sparse representations.

Several other approaches also analyze geometrical structures in images, like polynomial modeling with quadtree segmentation [22], footprints and edgeprints [23], multiscale transform [24], etc. Apart from the goal of efficient representation exploiting geometrical coherence, multi-directional (M-DIR) processing has also been applied to image denoising and classification. Examples of such transforms are the steerable pyramids [25], the cortex transform [26], the complex wavelets [27], the directional wavelet analysis [28], the directional filter-banks [7], [8], [29], brushlets [30], and the associative representation of visual information [31]. Some other methods involve directionally adaptive processing in order to preserve edges in images [32]–[35], whereas the methods proposed in [36], [37] impose DVM in either critically sampled or oversampled filter-banks. However, all of them fail to provide a *perfect reconstruction* and *critical* and *separable sampling* while keeping filter design completely in the *discrete domain* and with filters having DVM along *arbitrary directions*.

Our goal is to construct an anisotropic perfect reconstruction and critically sampled transform with HP filters having DVM, while retaining the simplicity of 1-D processing and filter design from the standard separable 2-D WT. We propose a transform construction based on partitioning of the discrete space using integer lattices, where the 1-D filtering is performed along lines across the lattice. The corresponding basis functions are called *directionlets*. We show that our transform has good approximation properties (see also [38]) as compared to the approximation achieved by some other overcomplete transform constructions [11]–[21] and is superior to the performance of the standard separable 2-D WT having the same complexity.

The outline of the paper is as follows. We present two constructions of anisotropic transforms in Section II. In Section III, we explain the inefficiency of the M-DIR transforms built on digital lines in order to motivate the need for integer lattice-based construction. We also give a review of integer lattices and the new construction of our skewed anisotropic lattice-based transforms. In Section IV, we explore the asymptotic approximation behavior of the anisotropic M-DIR transforms. We show that the achievable approximation scaling law is $O(N^{-1.55})$, where N is the number of retained coefficients. We also present some simulation results of approximation of natural images. Finally, we conclude and give the directions of future work in Section V.

II. ANISOTROPIC 2-D WAVELET DECOMPOSITIONS

As explained in Section I, the standard WT produces isotropic basis functions, which fail to provide a sparse representation of edges and contours. However, a new modified method that we propose retains the 1-D filtering and subsampling operations and can provide anisotropy, as we show next. In the sequel of this section, we give two examples of constructions of anisotropic transforms that still inherit the simplicity of processing and filter design from the standard WT. Furthermore, these two anisotropic transforms are critically sampled and lead to perfect reconstruction.

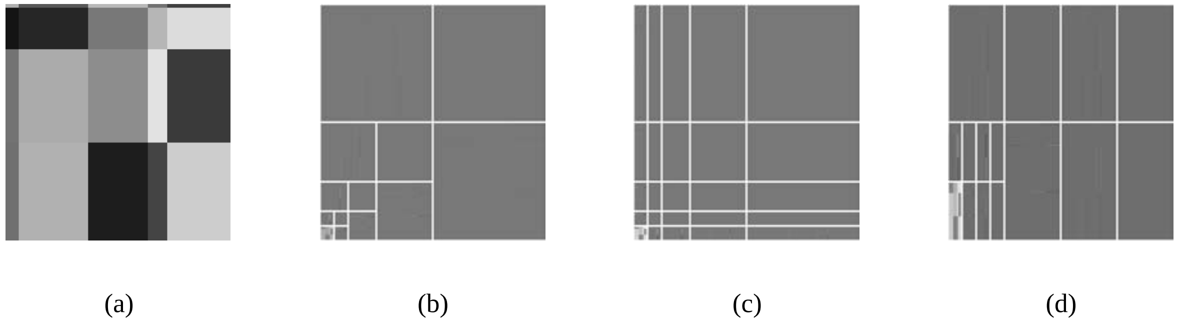


Fig. 3. (a) An image from the class $\texttt{Mondrian}(k_1,k_2)$. This class is inspired by the painting style established by Piet Mondrian (1872 - 1944). The image is transformed by the three transforms: (b) standard WT, (c) FSWT, (d) AWT(2,1) with 1-D wavelet filters having enough vanishing moments.

A. Fully Separable Decomposition

Define a simple class of piecewise polynomial images, denoted as $\texttt{Mondrian}(k_1,k_2)$ and inspired by the *geometrical period* of Piet Mondrian[1] [39].

Definition 1: The class $\texttt{Mondrian}(k_1,k_2)$ contains $M \times M$ piecewise polynomial images with k_1 horizontal and k_2 vertical discontinuities.

An example of the image from the class $\texttt{Mondrian}(k_1,k_2)$ is shown in Fig. 3(a). This class is not efficiently represented by the standard WT. The discontinuities lead to too many nonzero coefficients, as shown in the lemma below and in Fig. 3(b).

Lemma 1: Given an $M \times M$ pixel image from the class $\texttt{Mondrian}(k_1,k_2)$, the number of nonzero transform coefficients in band-pass subbands produced by the standard WT with the 1-D wavelets having enough vanishing moments[2] is given by

$$N = O\left((k_1 + k_2)M\right). \tag{1}$$

Proof: The three band-pass subbands at the jth ($1 \leq j \leq \log_2 M$) level of the standard WT contain $O(k_1 M/2^j + k_2)$, $O(k_1 + k_2 M/2^j)$, and $O(k_1 + k_2)$ nonzero coefficients. The total number of nonzero coefficients across scales is given by

$$\begin{aligned} N &= \sum_{j=1}^{\log_2 M} \left(O\left(k_1 \frac{M}{2^j} + k_2\right) + O\left(k_1 + k_2 \frac{M}{2^j}\right) + O\left(k_1 + k_2\right) \right) \\ &= O\left(2(k_1 + k_2)\log_2 M\right) + O\left((k_1 + k_2)(M - 1)\right) = O\left((k_1 + k_2)M\right). \end{aligned}$$

$\blacksquare$

To improve compactness of the representation of the class $\texttt{Mondrian}(k_1,k_2)$, we define the *fully separable WT* (FSWT). In this transform a full 1-D WT is applied in the horizontal direction (each row of image) and then, on

[1] The Dutch painter established *neoplasticism* and *De Stijl* in Europe in the beginning of the 20^{th} century. The image shown in Fig. 3(a) resembles to the paintings from his *geometrical period* (1930)

[2] A polynomial of the nth order is annihilated by a wavelet that has at least $n + 1$ vanishing moments.

each output a full 1-D WT is applied in the vertical direction (each column). The decomposition scheme is shown in Fig. 4(a). Notice that such a decomposition has already been proposed in [40] and also in [41], [42], where it is referred to as tensor wavelet basis.

The FSWT provides anisotropic basis functions (Fig. 4(c)) that are better adapted to the anisotropic objects such as the discontinuities in the class `Mondrian`(k_1,k_2). Representation efficiency is strongly improved, as can be seen in Fig. 3(c) from the resulting sparsity and it is given in Lemma 2.

Lemma 2: Given an $M \times M$ pixel image from the class `Mondrian`(k_1,k_2), the number of nonzero transform coefficients in band-pass subbands produced by the FSWT with the 1-D wavelets having enough vanishing moments is given by

$$O\left((k_1+k_2)\left(\log_2 M\right)^2\right). \tag{2}$$

Proof: Each band-pass subband is indexed by (j_1, j_2), where j_1 determines the number of the horizontal transforms, whereas j_2 enumerates the vertical transforms. The indexes are in the range $1 \leq j_1, j_2 \leq \log_2 M$.

The subband (j_1, j_2) contains $O\left(k_1+k_2\right)$ nonzero transform coefficients, therefore, the total number of nonzero coefficients is given by

$$N = \sum_{j_1=1}^{\log_2 M} \sum_{j_2=1}^{\log_2 M} O\left(k_1+k_2\right) = O\left((k_1+k_2)\left(\log_2 M\right)^2\right).$$

■

The performance of the FSWT on the class `Mondrian`(k_1,k_2), given by (2), is substantially better than the result of the standard WT, given by (1), namely, there is an exponential improvement in terms of M. The improvement is a consequence of anisotropy of the basis functions that is matched to the anisotropy of the class. However, the FSWT performs well only when it is applied on Mondrian-like images, while natural images contain features that are not well represented by straight (horizontal and vertical) lines.

Notice that if a transformed image contains a curve (or any discontinuity that is not a straight line), then the FSWT fails, as the number of nonzero coefficients grows exponentially across scales. Intuitively, the failure happens because the FSWT enforces a higher anisotropy (or elongation of the basis functions) than the one that is required in order to provide a compact representation of objects in natural images. To overcome this problem, we introduce a novel anisotropic transform, which performs better on a larger class of images.

B. Anisotropic Wavelet Decomposition

In the *anisotropic WT* (AWT) the number of transforms applied along the horizontal and vertical directions is unequal, that is, there are n_1 horizontal and n_2 vertical transforms at a scale, where n_1 is not necessarily equal to n_2. Then, the iteration is continued in the low-pass (LP), like in the standard WT. We denote such an anisotropic transform as AWT(n_1,n_2). The *anisotropy ratio* $\rho = n_1/n_2$ determines elongation of the basis functions of the AWT(n_1,n_2). An example of the construction and basis functions is shown in Fig. 5, where the AWT(2,1) is used.

Notice that both the standard WT and the FSWT can be expressed in terms of the AWT. The standard WT is simply given by AWT(1,1). However, the representation of the FSWT is more complex and is given as a concatenation

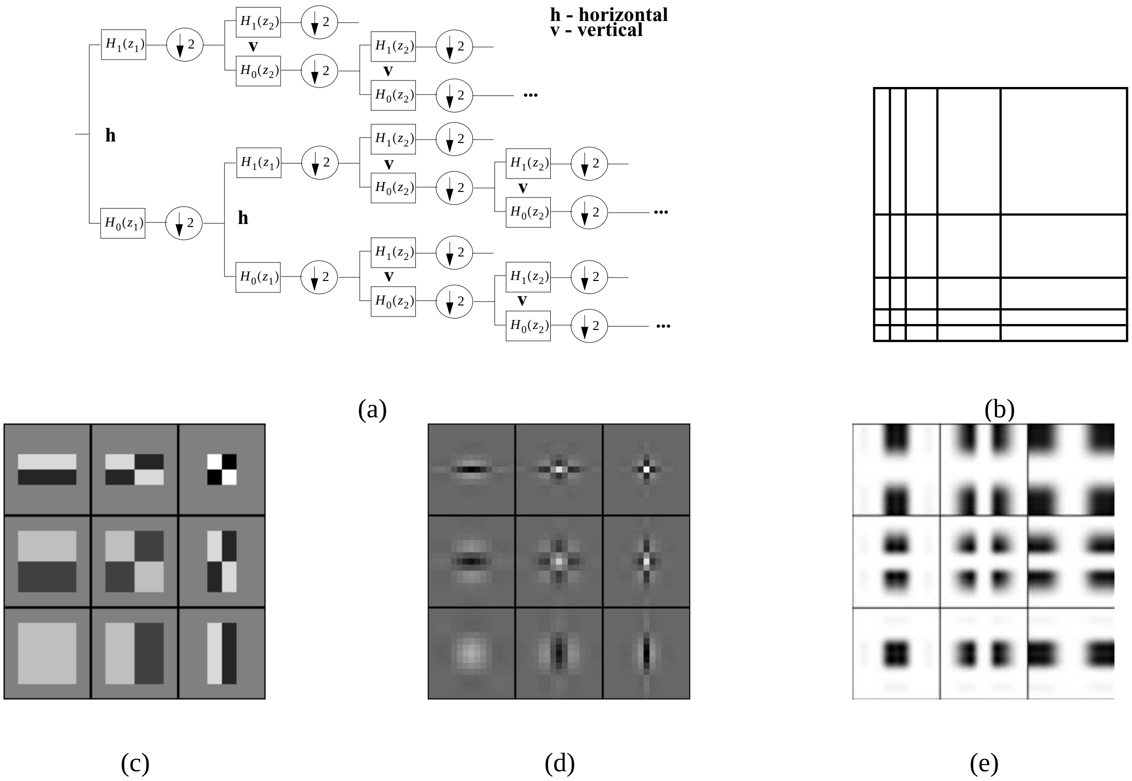


Fig. 4. The FSWT is anisotropic, as the number of 1-D transforms is not equal in the two directions. (a) An example of the transform scheme. Only 2 steps in each direction are shown. (b) The decomposition in frequency that corresponds to the construction in (a) with 4 steps in each direction. The anisotropic basis functions obtained from the (c) Haar and (d) biorthogonal "9-7" 1-D scaling and wavelet functions. (e) The corresponding Fourier transform of the basis functions obtained from the "9-7" 1-D filters.

of two AWTs. The first transform is AWT(n_{1max},0) that produces $n_{1max}+1$ subbands and it is followed by the AWT(0,n_{2max}) applied on each subband. The arguments n_{1max} and n_{2max} determine the maximal number of transforms in the two directions and depend on the size of the image.

Even though the AWT is not the most appropriate representation for the particular case of Mondrian-like images, it improves approximation of more general classes of images, as shown in Section IV. Fig. 3(d) shows the result of the AWT(2,1) of an image from the class `Mondrian`(k_1,k_2). The order of the number of nonzero coefficients is given by the following lemma.

Lemma 3: Given an $M \times M$ pixel image from the class `Mondrian`(k_1,k_2), the number of nonzero transform coefficients in band-pass subbands produced by the AWT(n_1,n_2) with 1-D wavelets having enough vanishing moments is given by

$$O\left(\left(ak_1+\frac{1}{a}k_2\right)M\right), \text{ where } a=\frac{2^{n_2}-1}{2^{n_1}-1}. \tag{3}$$

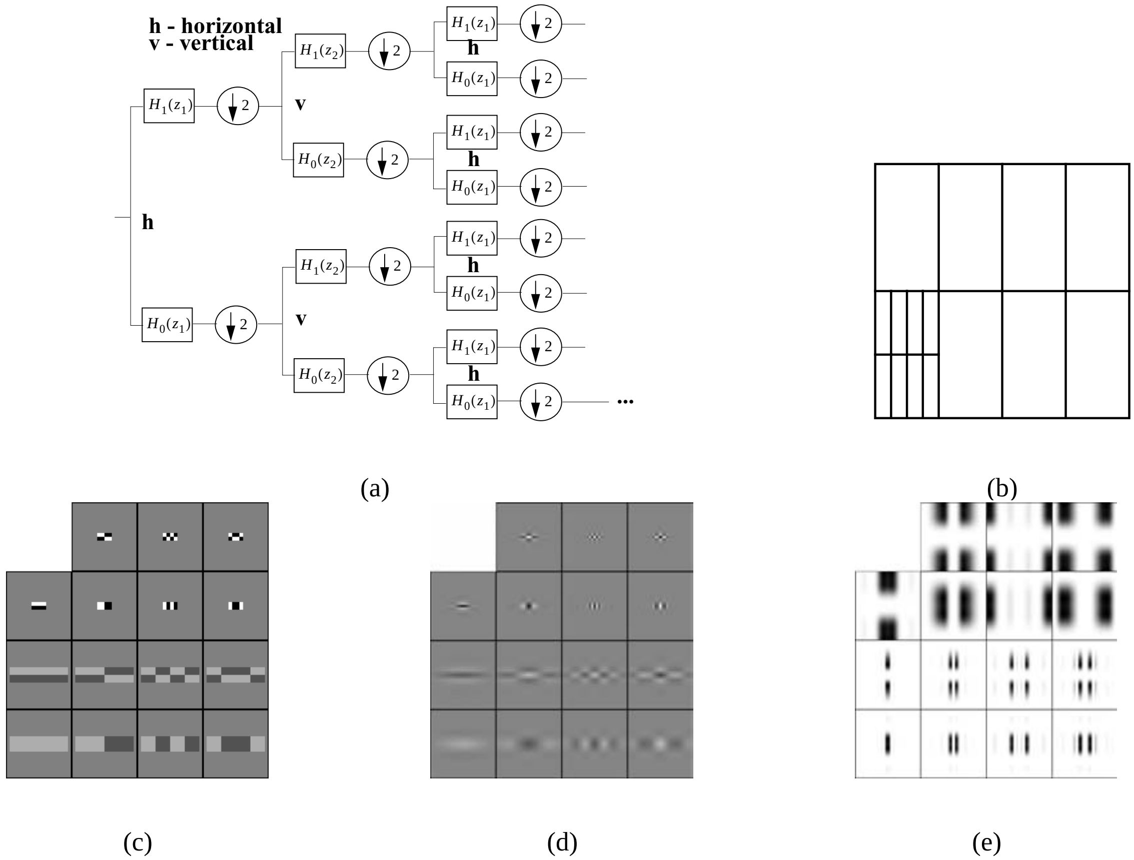


Fig. 5. The AWT allows for anisotropic iteration of the filtering and subsampling applied on the LP, similarly as in the standard WT. Although this transform does not improve approximation of the class `Mondrian`(k_1,k_2), it provides an efficient approximation tool for more general classes of images (Section IV). (a) The filtering scheme for the AWT(2,1), where one step of iteration is shown. (b) The decomposition in frequency. The basis functions obtained from the (c) Haar and (d) biorthogonal ”9-7” 1-D scaling and wavelet functions. (e) The corresponding Fourier transform of the basis functions obtained from the ”9-7” 1-D filters.

Proof: The number of nonzero coefficients produced at the jth level of the AWT(n_1,n_2) is given by

$$\begin{aligned} n(j) =& O(k_1\,(2^{n_2}-1)\,\frac{M}{2^{n_1 j}} + k_1\,(2^{n_1}-1)\,2^{n_2} \\ &+ k_2\,(2^{n_1}-1)\,\frac{M}{2^{n_2 j}} + k_2\,(2^{n_2}-1)\,2^{n_1}). \end{aligned}$$

The total number of nonzero coefficients across scales is, therefore,

$$N = \sum_{j=1}^{\frac{\log_2 M}{\max(n_1,n_2)}} n(j) = O\left(\left(ak_1 + \frac{1}{a}k_2\right)M\right).$$

■

Notice that the result in Lemma 3 is a generalization of the result in Lemma 1. Table I summarizes the orders of numbers of nonzero coefficients in band-pass subbands produced by the three transforms applied on the class `Mondrian`(k_1,k_2).

TABLE I

ORDERS OF APPROXIMATION BY THE STANDARD WT, FSWT AND AWT APPLIED ON THE CLASS MONDRIAN(k_1,k_2).

Standard WT	**FSWT**	**AWT**
$(k_1+k_2)M$	$(k_1+k_2)(\log_2 M)^2$	$(k_1 a + k_2/a)M$

The transforms explained in this section are applied in the horizontal and vertical directions only. More general transforms can be obtained by imposing vanishing moments along different directions. These transforms provide an efficient representation of more general classes of images, involving more than only the two standard directions, as shown in the next section.

III. LATTICE-BASED SKEWED WAVELET TRANSFORMS

Several transform constructions that lead to anisotropic basis functions have been presented in Section II. However, all the constructions, including the standard WT, use only horizontal and vertical directions. Notice also that the HP filters in these transforms have vanishing moments only along these two directions. Here, we present the novel lattice-based transform, which exploits multi-directionality and retains the simplicity of computations and filter design from the standard WT.

In the continuation, we explain the problem of approximation of directions in the discrete space $\mathbb{Z}^2$ and we introduce the concept of directional interaction. Then, we propose a new lattice-based method that allows for a generalization of the transform constructions from Section II to include separable (1-D) filtering and subsampling across multiple directions, not only horizontal and vertical. We also give the polyphase analysis of the lattice-based transforms.

A. Discretization of Directions

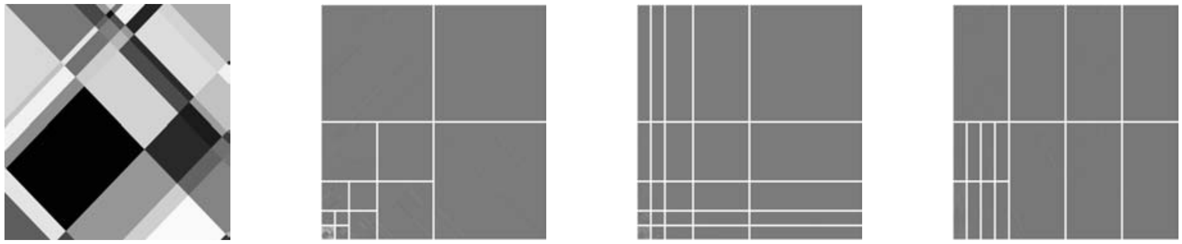
To apply a discrete transform in the discrete space $\mathbb{Z}^2$ in a certain direction, we need to define the pixels that approximate the chosen direction. This problem has been considered in computer graphics in the 1960's [43] as well as in [44], [45].

Recall that the set of points $(x, y) \in \mathbb{R}^2$ represents a continuous line with the slope r and intercept b if the following equality is satisfied:

$$y = rx + b. \tag{4}$$

The discrete approximation of (4) is called *digital line* $L(r, n)$. To preserve critical sampling in the transform, given a slope r, every pixel belongs to one and only one digital line $L(r, n)$. In that case, we say that, given a slope r, the set of digital lines $\{L(r, n) : n \in \mathbb{Z}\}$, *partitions the discrete space* $\mathbb{Z}^2$.

The definitions of digital lines proposed in [43]–[45] are similar and here we give the definition that is a variation of the one given in [43]. We show also below that such digital lines partition the discrete space $\mathbb{Z}^2$.



(a) (b) (c) (d)

Fig. 6. (a) An example of an image from the class S-Mondrian($\mathbf{M}(r_1, r_2)$,k_1,k_2), for $\mathbf{M} = [\mathbf{v}_1, \mathbf{v}_2]^T$, where $\mathbf{v}_1 = [1, 1]$ and $\mathbf{v}_2 = [-1, 1]$. The image is transformed using (b) S-WT, (c) S-FSWT, and (d) S-AWT($\mathbf{M}_\Lambda$,2,1) (directionlets), where all the transforms are built on the lattice Λ determined by the generator matrix $\mathbf{M}_\Lambda = \mathbf{M}(r_1, r_2)$.

Definition 2: Given a rational slope r, the digital line $L(r, n)$, where $n \in \mathbb{Z}$, is defined as the set of pixels (i, j) such that

$$j = \lceil ri \rceil + n, \ \forall i \in \mathbb{Z}, \text{ for } |r| \leq 1, \text{ or}$$
$$i = \lceil j/r \rceil + n, \ \forall j \in \mathbb{Z}, \text{ for } |r| > 1. \tag{5}$$

Lemma 4: Given a rational slope r, the set of digital lines $\{L(r, n) : n \in \mathbb{Z}\}$ partitions the discrete space $\mathbb{Z}^2$.

Proof: We give the proof only for the case $|r| \leq 1$. Similar arguments can be used for the other cases. For each pixel $(i, j) \in \mathbb{Z}^2$, we can find the intercept $n = j - \lceil ri \rceil$ such that the pixel belongs to the digital line $L(r, n)$. Furthermore, from (5) it follows that this intercept is unique. Therefore, the digital lines $L(r, n)$, $\forall n \in \mathbb{Z}$, partitions the discrete space $\mathbb{Z}^2$. ■

The concept of digital lines is useful for overcomplete M-DIR representation. However, in the sequel, we show why digital lines do not provide an efficient framework when transforms are applied in different directions and critical sampling is enforced.

B. Directional Interaction

To explain the problem of *directional interaction*, let us first generalize the class Mondrian allowing for more directions. The class S-Mondrian consists of the skewed Mondrian-like images along two directions with the rational slopes $r_1 = b_1/a_1$ and $r_2 = b_2/a_2$, where a_1, a_2, b_1, and b_2 are integers. To simplify notation, the two slopes are jointly denoted by the matrix

$$\mathbf{M}(r_1, r_2) = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}.$$

Definition 3: The class S-Mondrian($\mathbf{M}(r_1, r_2)$,k_1,k_2) contains $M \times M$ piecewise polynomial images with k_1 and k_2 discontinuities along the digital lines $L(r_1, n)$ and $L(r_2, n)$, respectively, where $n \in \mathbb{Z}$, $r_1 = b_1/a_1$, $r_2 = b_2/a_2$, and $a_1, a_2, b_1, b_2 \in \mathbb{Z}$.

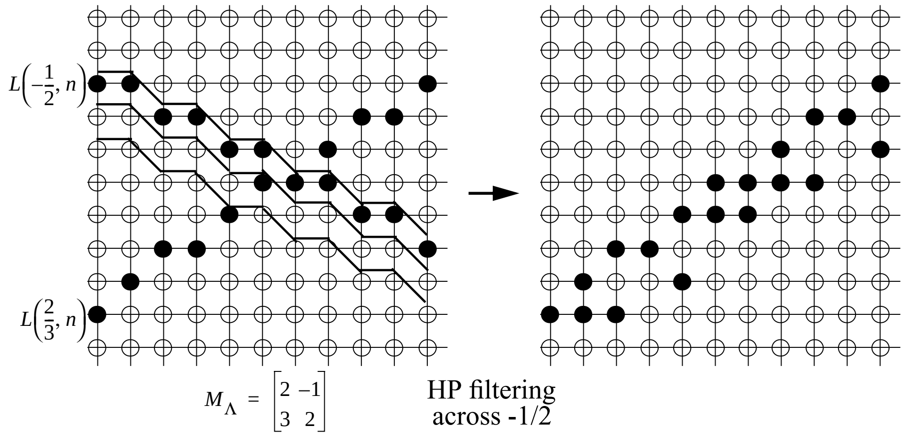


Fig. 7. A 1-D WT is applied on an image from the class S-Mondrian($M(-1/2, 2/3)$,1,1) along the digital lines $L(-1/2, n)$. The HP filtering annihilates the digital line with the slope $-1/2$. However, the nonzero coefficients produced by the other line with the slope $2/3$ are not aligned in the digital lines $L(2/3, n)$. This is called *directional interaction*. Although the transform along digital lines is efficient when applied in oversampled schemes, it fails to provide a systematic subsampling method when critical sampling is enforced.

Notice that the class Mondrian(k_1,k_2) is a special case of the larger class S-Mondrian($\mathbf{M}(r_1, r_2)$,k_1,k_2) when $\mathbf{M}(r_1, r_2) = \mathbf{I}_2$. An example of an image from the class S-Mondrian($\mathbf{M}(r_1, r_2)$,k_1,k_2) is shown in Fig. 6(a). Notice also that only the lines with rational slopes are used in the class S-Mondrian. However, in spite of this constraint, a wealth of directions is still available, as we will explain in Section III-C.

To provide a sparse representation of the class S-Mondrian($\mathbf{M}(r_1, r_2)$,k_1,k_2) and following the ideas from Section II, we apply a 1-D WT along the digital lines $L(r_1, n)$, for $n \in \mathbb{Z}$. The transform produces two types of nonzero coefficients, that is, the coefficients corresponding to the discontinuities with the slopes r_1 and r_2.

Since the HP filter has vanishing moments along digital lines with the slope r_1, the coefficients along this direction are annihilated in the HP subband, while the coefficients along the second direction with the slope r_2 are retained in both subbands. However, after subsampling, unlike in the case of the standard directions, the coefficients along the second direction are not aligned, that is, they cannot be clustered in the digital lines with the slope r_2. Therefore, the following 1-D WT applied along the digital lines with the slope r_2 does not annihilate the coefficients along the second direction and, hence, it yields a non-sparse representation. We call this phenomenon *directional interaction*. The proof is trivial and is omitted here. An example is shown in Fig. 7.

Notice also that the concept of digital lines does not provide a systematic rule for subsampling in the case of iteration of the filtering and subsampling along the directions with the slopes r_1 and r_2 when critical sampling is enforced. To overcome the directional interaction and to propose an organized iterated subsampling method we use the concept of integer lattices.

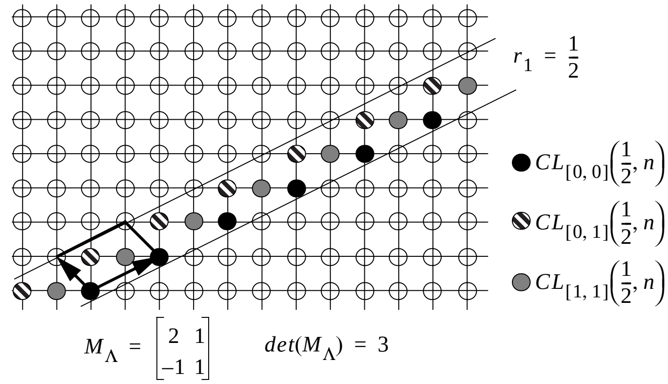


Fig. 8. The intersections between the 3 cosets of the lattice Λ given by the generator matrix $\mathbf{M}_\Lambda$ and the digital lines $L(r_1 = 1/2, n)$, where $n \in \mathbb{Z}$, are the co-lines $CL_{[0,0]}(1/2, n)$, $CL_{[0,1]}(1/2, n)$, and $CL_{[1,1]}(1/2, n)$.

C. Lattice-based Filtering and Subsampling

Instead of applying a transform along digital lines, we propose a novel method that is based on integer lattices [46]. We also prove that the lattice-based transforms can avoid directional interaction and are capable of providing the same order of approximation for the class `S-Mondrian` as the FSWT achieves for the class `Mondrian`.

A full-rank integer lattice Λ consists of the points obtained as linear combinations of two linearly independent vectors, where both the components of the vectors and the coefficients are integers. Any integer lattice Λ is a sublattice of the cubic integer lattice $\mathbb{Z}^2$, that is, $\Lambda \subset \mathbb{Z}^2$. The lattice Λ can be represented by a non-unique generator matrix

$$\mathbf{M}_\Lambda = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} = \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix}, \text{ where } a_1, a_2, b_1, b_2 \in \mathbb{Z}. \tag{6}$$

Recall that the cubic lattice $\mathbb{Z}^2$ can be partitioned into $|\det(\mathbf{M}_\Lambda)|$ cosets of the lattice Λ [46], where each coset is determined by the shift vector $\mathbf{s}_k$, for $k = 0, 1, \ldots, |\det(\mathbf{M}_\Lambda)| - 1$. Therefore, the lattice Λ with the corresponding generator matrix $\mathbf{M}_\Lambda$ given by (6), partitions each digital line $L(r_1 = b_1/a_1, n)$ into *co-lines*. Notice that a co-line is simply the intersection between a coset and a digital line. Similarly, the digital line $L(r_2 = b_2/a_2, n)$ is also partitioned into the corresponding co-lines (Fig. 8).

We denote as $CL_{\mathbf{s}_k}(r_1, n)$ the co-line obtained as the intersection between the kth coset of the lattice Λ and the digital line $L(r_1 = b_1/a_1, n)$. Notice that the co-line $CL_{\mathbf{s}_k}(r_1, n)$ consists of the pixels $\{c_1\mathbf{d}_1 + c_2\mathbf{d}_2 + \mathbf{s}_k : \forall c_1 \in \mathbb{Z},\ \text{fixed } c_2 \in \mathbb{Z}\}$, where $n = \lceil c_2(b_2 - r_1 a_2) + s_{k,2} - r_1 s_{k,1} \rceil$ and $\mathbf{s}_k = [s_{k,1}, s_{k,2}]$.

Now we apply the 1-D WT (including the 1-D both filtering and subsampling operations) along the co-lines $\{CL_{\mathbf{s}_k}(r_1, n) : n \in \mathbb{Z},\ k = 0, 1, \ldots, |\det(\mathbf{M}_\Lambda)| - 1\}$ (see also [47]). Notice that both filtering and subsampling are applied in each of the cosets separately. Furthermore, each filtering operation is purely 1-D. After subsampling, the retained points belong to the sublattice Λ' of the lattice Λ ($\Lambda' \subset \Lambda$) with the corresponding generator matrix

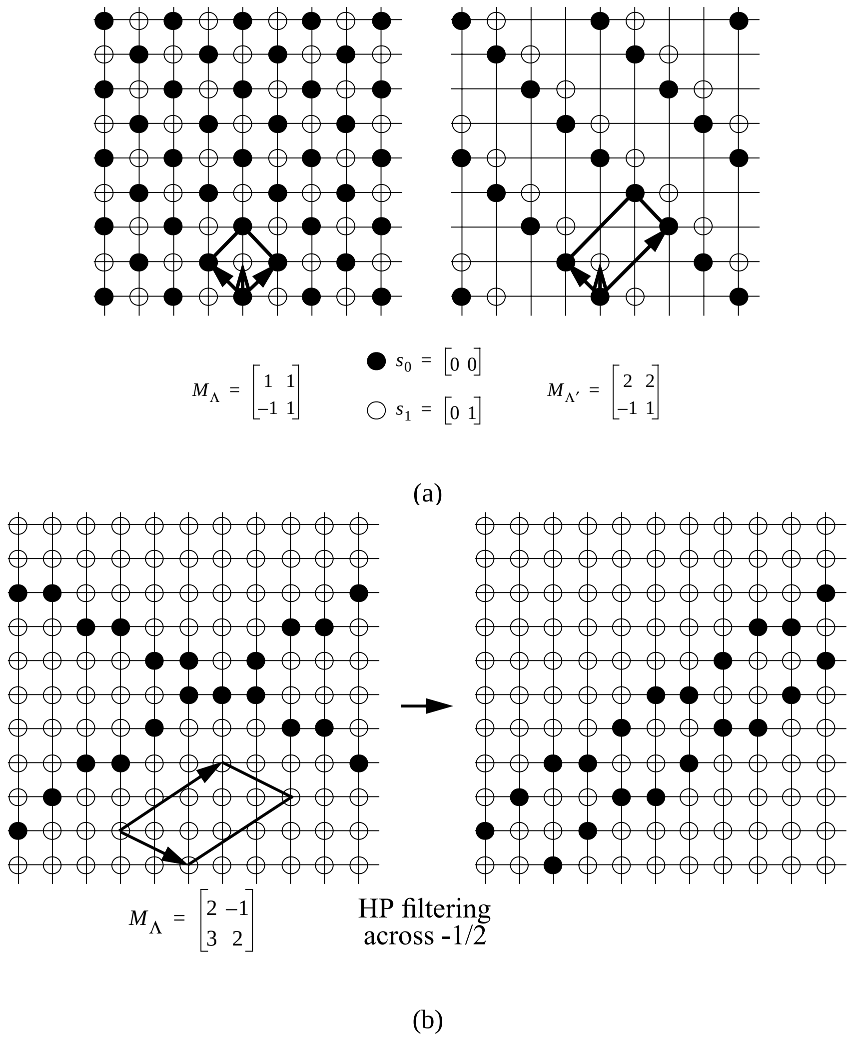


Fig. 9. (a) The lattice Λ is determined by the generator matrix $\mathbf{M}_\Lambda$. 1-D Filtering is applied along the co-lines $\{CL_{\mathbf{s}_k}(r_1, n) : n \in \mathbb{Z},\ k = 0, 1, \ldots, |\det(\mathbf{M}_\Lambda)| - 1\}$, where the slope r_1 corresponds to the vector $[1, 1]$, that is, along 45°. The pixels retained after the subsampling belong to the lattice $\Lambda' \subset \Lambda$ determined by the generator matrix $\mathbf{M}_{\Lambda'}$. Notice that filtering and subsampling are applied separately in two cosets, determined by the shift vectors $\mathbf{s}_0$ and $\mathbf{s}_1$. (b) The nonzero pixels obtained after one step of the lattice-based filtering operation applied on the same example as in Fig. 7 are clustered in the digital lines with the slope $2/3$.

given by (see Fig. 9(a))

$$\mathbf{M}_{\Lambda'} = \mathbf{D}_s \cdot \mathbf{M}_\Lambda = \begin{bmatrix} 2\mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix}.$$

Here, $\mathbf{D}_s$ is the horizontal subsampling operator, that is,

$$\mathbf{D}_s = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}.$$

We call the direction along the first vector $\mathbf{d}_1$ (with the slope $r_1 = b_1/a_1$), the *transform direction*. Similarly, the direction along the second vector $\mathbf{d}_2$ we call the *alignment direction*.

Therefore, since the filtering and subsampling are applied in each coset separately, the pixels retained after the subsampling are clustered in co-lines along the alignment direction. This property is crucial to avoid directional interaction (see Fig. 9(b)).

Lemma 5: Given a 1-D WT applied along the set of co-lines $\{CL_{\mathbf{s}_k}(r_1,n) : n \in \mathbb{Z},\ k = 0,1,\ldots,|\det(\mathbf{M}_\Lambda)|-1\}$ on an image from the class S-Mondrian($\mathbf{M}(r_1,r_2)$,k_1,k_2), the transform coefficients in band-pass subbands that correspond to the discontinuities with the slope r_2 are aligned, that is, they can be clustered in the co-lines $CL_{\mathbf{s}_k}(r_2,n)$, $n \in \mathbb{Z}$.

Proof: Recall that the co-line $CL_{\mathbf{s}_k}(r_1,n)$ consists of the pixels $\{(i,j) : i = c_1 a_1 + c_2 a_2 + s_{k,1}, j = c_1 b_1 + c_2 b_2 + s_{k,2}, \forall c_1 \in \mathbb{Z},$ fixed $c_2 \in \mathbb{Z}\}$. After the subsampling, the retained pixels belong to the lattice Λ' and, thus, the corresponding co-lines consist of the pixels (i,j) such that $i = c_1 \cdot 2a_1 + c_2 a_2 + s_{k,1}$ and $j = c_1 \cdot 2b_1 + c_2 b_2 + s_{k,2}$ for each $c_1 \in \mathbb{Z}$ and a fixed $c_2 \in \mathbb{Z}$.

Notice that the co-lines $CL_{\mathbf{s}_k}(r_2,n)$ with the other slope r_2 that correspond to the lattice Λ' consist of the same pixels. Therefore all the retained pixels are aligned in the direction with the slope r_2. ■

Combining lattices with the different constructions given in Section II, we build *skewed wavelet transforms*.

D. Skewed Wavelet Transforms

The transforms defined in Section II (the standard WT, FSWT, and AWT) are inefficient when applied on the class S-Mondrian($\mathbf{M}(r_1,r_2)$,k_1,k_2), unless $\mathbf{M}(r_1,r_2)$ is the identity matrix. Since the directions of the transforms and discontinuities in images are not matched, the transforms fail to provide a compact representation. The following lemma gives the orders of approximation that can be achieved by the three transforms with the standard directions.

Lemma 6: Given an $M \times M$ pixel image from the class S-Mondrian($\mathbf{M}(r_1,r_2)$,k_1,k_2), where $\mathbf{M}(r_1,r_2)$ *is not the identity matrix*, the standard WT, FSWT, and AWT with 1-D wavelets having enough vanishing moments provide $O((k_1+k_2)M)$ nonzero transform coefficients in band-pass subbands.

Proof: The subbands produced by the FSWT are indexed by (j_1,j_2), where $1 \le j_1, j_2 \le \log_2 M$. Each subband contains $O(k_1 M/2^{j_1} + k_2 M/2^{j_2})$ nonzero coefficients. The total number is given by

$$N = \sum_{j_1=1}^{\log_2 M} \sum_{j_2=1}^{\log_2 M} O\left(k_1 \frac{M}{2^{j_1}} + k_2 \frac{M}{2^{j_2}}\right) = O\left((k_1+k_2)M\right).$$

Notice that the standard WT, as a special case of the AWT, has the same behavior. Thus, we give the proof only for the AWT. The AWT(n_1,n_2) produces $2^{n_1+n_2}-1$ band-pass and HP subbands at each scale j. Each of these subbands contain $n(j) = O((2^{n_1+n_2}-1)M(2^{-n_1 j} + 2^{-n_2 j}))$ nonzero coefficients. Therefore, the total number of nonzero coefficients is given by

$$\sum_{j=1}^{\frac{\log_2 M}{\max(n_1,n_2)}} n(j) = O\left((k_1+k_2)M\right).$$

■

Using integer lattices, we define the three new transforms, which are *skewed versions* of the standard WT, FSWT, and AWT. Given a lattice Λ, the skewed transforms are applied along co-lines in the transform and alignment

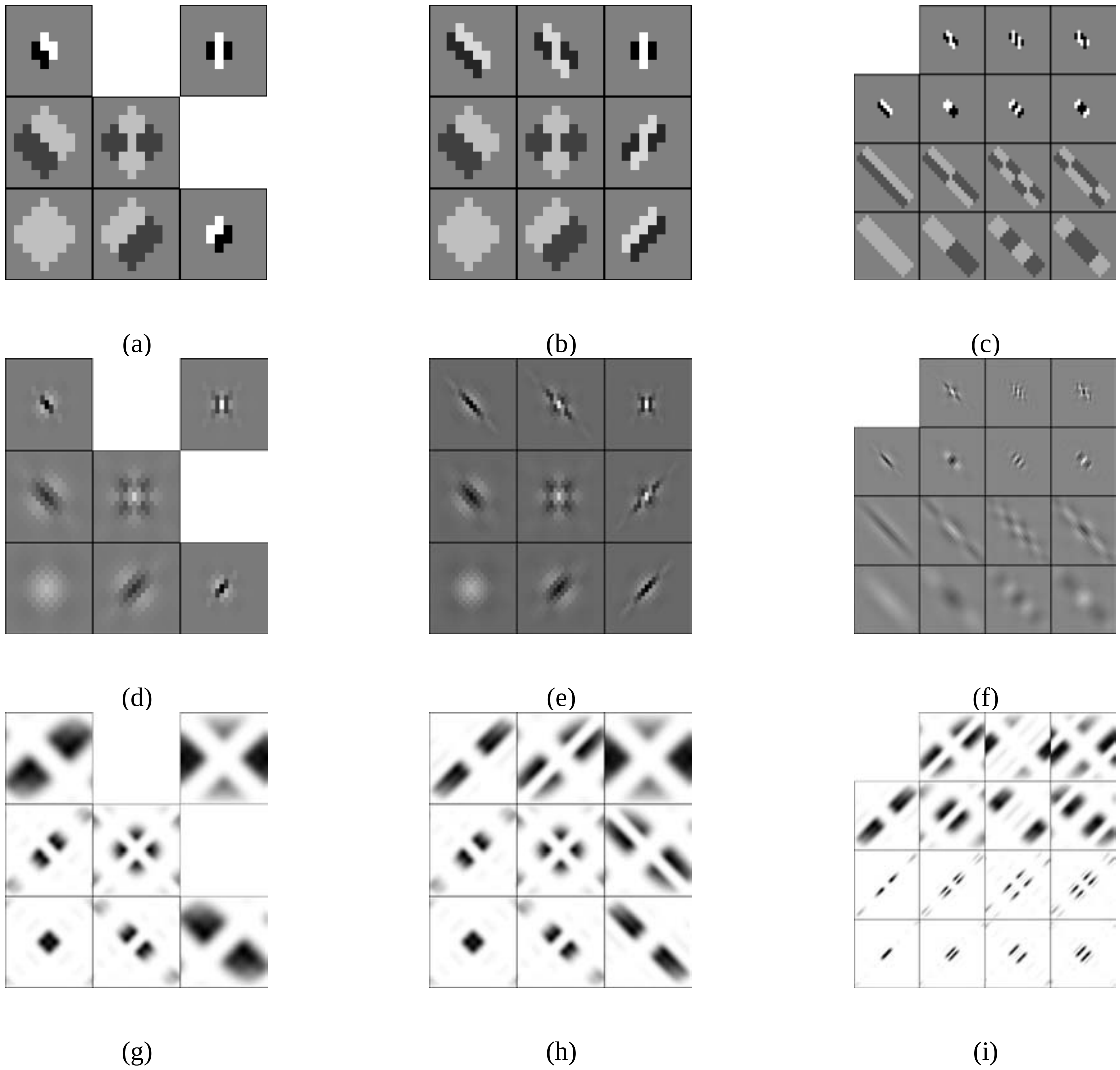


Fig. 10. The basis functions obtained by the skewed transforms using the Haar 1-D scaling and wavelet functions: (a) S-WT, (b) S-FSWT, (c) S-AWT($\mathbf{M}_\Lambda$,2,1) (directionlets). The same, but with the biorthogonal ”9-7” 1-D scaling and wavelet functions: (d) S-WT, (e) S-FSWT, (f) S-AWT($\mathbf{M}_\Lambda$,2,1) (directionlets). In all cases $\mathbf{M}_\Lambda = [\mathbf{d}_1, \mathbf{d}_2]^T$, where $\mathbf{d}_1 = [1,1]$, and $\mathbf{d}_2 = [-1,1]$. The DVMs are imposed along the vectors $\mathbf{d}_1$ and $\mathbf{d}_2$, that is, along $45°$ and $-45°$. The corresponding Fourier transforms: (g) S-WT, (h) S-FSWT, (i) S-AWT($\mathbf{M}_\Lambda$,2,1) (directionlets).

directions of the lattice Λ, retaining the same frequency decompositions as the corresponding transforms along the standard directions explained in Section II. Thus, following the notation introduced in Section II-B, we denote as S-AWT($\mathbf{M}_\Lambda$,n_1,n_2) the skewed anisotropic transform built on the lattice Λ that has n_1 and n_2 transforms in one iteration step along the transform and alignment directions, respectively. We call the basis functions of the S-AWT *directionlets* since they are anisotropic and have a specific direction. Similarly, we denote the skewed standard WT as S-WT and the skewed FSWT as S-FSWT. The corresponding basis functions are shown in Fig. 10 for the directions along the vectors $\mathbf{d}_1 = [1,1]$ and $\mathbf{d}_2 = [-1,1]$. Notice that the skewed transforms are applied in all cosets of the lattice Λ separately.

The basis functions of the skewed transforms have DVM in *any* two directions with rational slopes. Recall that the Lth order DVM along the direction with a rational slope $r_1 = b_1/a_1$ is equivalent to requiring the z-transform of a basis function to have a factor $(1 - z_1^{-a_1} z_2^{-b_1})^L$ [21], [48]. The following lemma gives the number and directions of the DVM in directionlets.

Lemma 7: Assume that the directionlets of the S-AWT($\mathbf{M}_\Lambda$,n_1,n_2) are obtained using a 1-D wavelet with L vanishing moments. Then, at each scale of the iteration, there are:

(a) $2^{n_1} - 1$ directionlets with the Lth order DVM along the transform direction of the lattice Λ,
(b) $2^{n_2} - 1$ directionlets with the Lth order DVM along the alignment direction of the lattice Λ, and
(c) $(2^{n_1} - 1)(2^{n_2} - 1)$ directionlets with the Lth order DVM along both directions.

Proof: Recall first from [48] that 1-D filtering using the filter $H(z)$ along the transform direction of the lattice Λ is equivalent to filtering in the 2-D discrete space using $H(z_1^{a_1} z_2^{b_1})$. Similarly, filtering along the alignment direction of the lattice Λ is equivalent to filtering in the 2-D discrete space using $H(z_1^{a_2} z_2^{b_2})$. Since the 1-D HP filter has L vanishing moments, its z-transform has a factor $(1 - z^{-1})^L$. Therefore, the HP filtering along the transform and alignment directions uses the equivalent filters with the factors $(1 - z_1^{-a_1} z_2^{-b_1})^L$ and $(1 - z_1^{-a_2} z_2^{-b_2})^L$, respectively, in the z-transforms.

Filtering using the 1-D two-channel filter-bank along two directions in the construction of the S-AWT (see Fig. 5(a)) yields (a) $2^{n_1} - 1$ subbands with HP filtering along only the transform direction, (b) $2^{n_2} - 1$ subbands with HP filtering along only the alignment direction, and (c) $(2^{n_1} - 1)(2^{n_2} - 1)$ subbands with HP filtering along both directions. Thus, the statement of the lemma follows directly. ■

Efficiency of representation of the class `S-Mondrian`($\mathbf{M}(r_1, r_2)$,k_1,k_2) by the three skewed transforms depends on matching between the directions of discontinuities and the directions used in these transforms. If these directions are matched, then the orders of nonzero coefficients in band-pass subbands are equal to the orders calculated in Section II (see Table I). Otherwise, they are given by the result in Lemma 6. The following lemma formalizes this statement. The proof is omitted since it uses the same arguments as in Lemmas 1 to 3.

Lemma 8: Given an $M \times M$ pixel image from the class `S-Mondrian`($\mathbf{M}(r_1, r_2)$,k_1,k_2), the S-WT, S-FSWT and S-AWT($\mathbf{M}_\Lambda$,n_1,n_2) with 1-D wavelets having enough vanishing moments built on the lattice Λ determined by the generator matrix $\mathbf{M}_\Lambda = \mathbf{M}(r_1, r_2)$ give $O((k_1 + k_2)M)$, $O((k_1 + k_2)(\log_2 M)^2)$ and $O((k_1 a + k_2/a)M)$ nonzero coefficients in band-pass subbands, respectively. Here, $a = (2^{n_2} - 1)/(2^{n_1} - 1)$.

The transforms of the image shown in Fig. 6(a) are given in Fig. 6(b)-(d). The applied transforms are S-WT, S-FSWT, and S-AWT($\mathbf{M}_\Lambda$,2,1), where $\mathbf{M}(r_1, r_2) = \mathbf{M}_\Lambda$. Table II summarizes the orders of nonzero coefficients in band-pass subbands in the case of both matched and mismatched directions.

Notice that the lattice-based method allows for a more general construction of M-DIR transforms using more than two directions in an arbitrary order. Such M-DIR transforms and their properties are beyond the scope of this paper. More details are given in [47], [49].

TABLE II

ORDERS OF APPROXIMATION BY THE S-WT, S-FSWT AND S-AWT (DIRECTIONLETS) BUILT ON THE LATTICE Λ DETERMINED BY $\mathbf{M}_\Lambda$ APPLIED ON THE CLASS S-MONDRIAN($\mathbf{M}(r_1, r_2)$,k_1,k_2).

	$\mathbf{M}_\Lambda = \mathbf{M}(r_1, r_2)$	$\mathbf{M}_\Lambda \neq \mathbf{M}(r_1, r_2)$
S-WT	$(k_1 + k_2)M$	$(k_1 + k_2)M$
S-FSWT	$(k_1 + k_2)(\log_2 M)^2$	$(k_1 + k_2)M$
S-AWT	$(k_1 a + k_2/a)M$	$(k_1 + k_2)M$

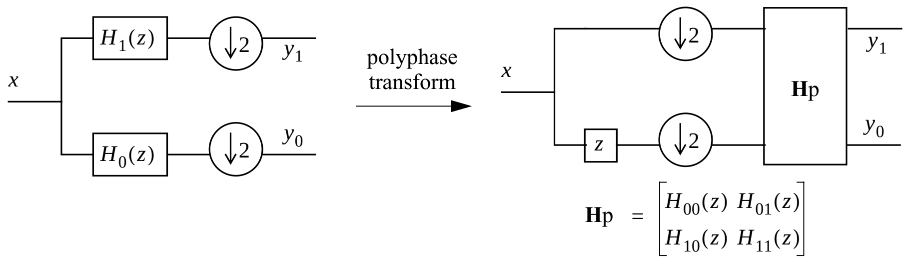


Fig. 11. A 1-D filter-bank $(H_0(z), H_1(z))$ with the subsampling factor 2 is represented in the polyphase domain with the corresponding polyphase components $H_{00}(z)$, $H_{01}(z)$, $H_{10}(z)$, and $H_{11}(z)$.

E. Polyphase Representation

Filtering and subsampling across lattices, as explained in Section III-C, can be efficiently represented in the polyphase domain. Recall first that a two-channel 1-D filter-bank $(H_0(z), H_1(z))$ followed by a subsampler by the factor 2 can be given in terms of the polyphase components as [2]

$$H_0(z) = H_{00}\left(z^2\right) + zH_{01}\left(z^2\right) \text{ and}$$
$$H_1(z) = H_{10}\left(z^2\right) + zH_{11}\left(z^2\right).$$

Here, H_{00}, H_{01}, H_{10}, and H_{11} are the polyphase components of the filters $H_0(z)$ and $H_1(z)$ that correspond to even and odd samples of the impulse response, respectively. Such a polyphase representation is shown in Fig. 11.

Similarly, we can find the equivalent polyphase components of a 2-D filter-bank $(H_0(\mathbf{z}), H_1(\mathbf{z}))$, where $\mathbf{z} = (z_1, z_2)$, applied in the lattice-based method, as explained in Section III-C. Recall that the filters $H_0(\mathbf{z})$ and $H_1(\mathbf{z})$ used in this method are purely 1-D filters, that is, $H_0(\mathbf{z}) = H_0(z_1)$ and $H_1(\mathbf{z}) = H_1(z_1)$. To illustrate this polyphase decomposition, we consider the particular example with the lattice Λ determined by the generator matrix

$$\mathbf{M}_\Lambda = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix},$$

as shown in Fig. 9(a). Recall that the lattice-based filtering and subsampling are applied in each coset of the lattice Λ separately. Thus, the equivalent scheme has two sections, which are (a) separation into two cosets and (b) 1-D filtering and subsampling in the transform direction (Fig. 12(a)). Notice that filtering in the transform direction is

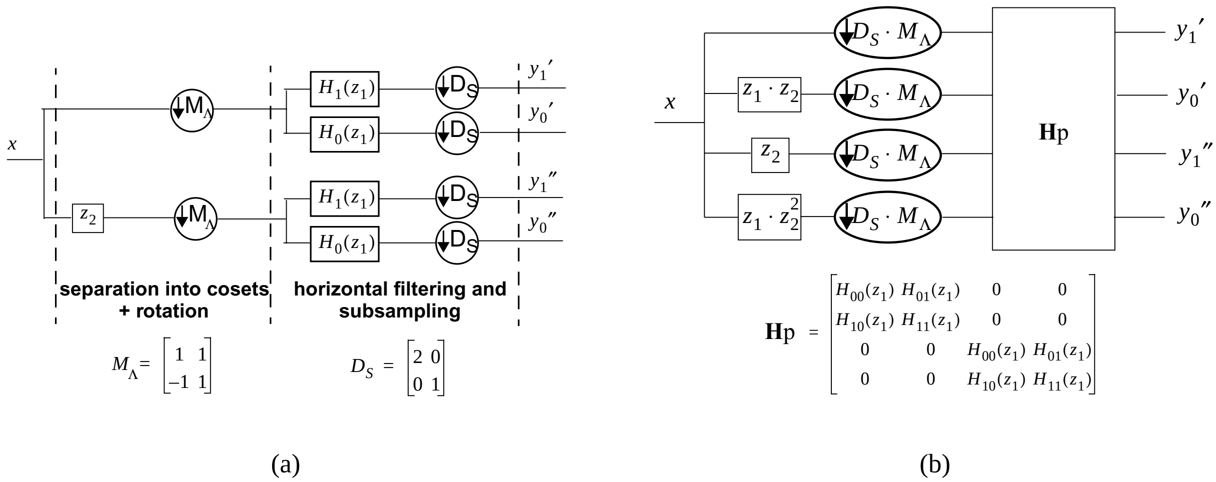


Fig. 12. (a) The 2-D two-channel filter-bank applied in the example shown in Fig. 9(a). Filtering and subsampling are applied in 2 cosets separately. (b) Equivalent polyphase representation contains 4 components. The polyphase transform $\mathbf{H}_p$ is block-diagonal.

performed as horizontal filtering preceded by rotation by the generator matrix $\mathbf{M}_{\Lambda}$.

Since the total subsampling rate is $|\det(\mathbf{D}_s \cdot \mathbf{M}_{\Lambda})| = 4$, the polyphase representation of such a filter-bank consists of 4 polyphase components. The equivalent polyphase representation is shown in Fig. 12(b), where the polyphase transform $\mathbf{H}_p$ is *block-diagonal*, that is,

$$\mathbf{H}_p = \begin{bmatrix} H_{00}(z_1) & H_{01}(z_1) & 0 & 0 \\ H_{10}(z_1) & H_{11}(z_1) & 0 & 0 \\ 0 & 0 & H_{00}(z_1) & H_{01}(z_1) \\ 0 & 0 & H_{10}(z_1) & H_{11}(z_1) \end{bmatrix}.$$

Notice that the block-diagonal polyphase transform with two identical blocks is a consequence of the separable transforms applied across cosets. This property allows for a simple filter design and computational efficiency in the polyphase domain. Such separability in the polyphase domain has also been used in other 2-D filter-bank designs [7], [8].

IV. NON-LINEAR APPROXIMATION AND COMPRESSION

The main task of approximation is to represent a signal by a portion of transform coefficients, while the rest of them is set to zero. The transform can be critically sampled (bases) or oversampled (frames). The approximation with N retained transform coefficients is also called *N-term approximation*. We distinguish between linear approximation (LA) and non-linear approximation (NLA). In the first, the indexes of the retained coefficients are fixed, whereas in the latter, they are adapted to the content of the signal.

Owing to truncation of the coefficients, the approximating signal does not match exactly the original one. The quality of the approximation is commonly measured in terms of *mean-square error* (MSE), that is, for a signal $\mathbf{x}$

and its N-term approximation $\hat{\mathbf{x}}_N$, the MSE is given by $\|\mathbf{x}-\hat{\mathbf{x}}_N\|^2$. Notice that, given a signal $\mathbf{x}$ and its transform $\mathbf{y}=\mathbf{F}\cdot\mathbf{x}$, where $\mathbf{F}$ is a tight frame or an orthogonal basis, we have the following inequality

$$\|\mathbf{x}-\hat{\mathbf{x}}_N\|^2 \leq \frac{1}{A}\|\mathbf{y}-\hat{\mathbf{y}}_N\|^2, \tag{7}$$

where $\hat{\mathbf{y}}_N$ corresponds to the truncated version of $\mathbf{y}$ with N retained coefficients, the N-term approximation $\hat{\mathbf{x}}_N$ is given by $\hat{\mathbf{x}}_N = A^{-1}\mathbf{F}^T\cdot\hat{\mathbf{y}}_N$, and A is the frame bound of $\mathbf{F}$ (for more details see Appendix I). Equality in (7) holds if the transform $\mathbf{F}$ is an orthogonal basis.

In the orthogonal case, the optimal strategy to minimize the MSE is to retain the largest-magnitude transform coefficients [50]. Notice that the MSE decays as the number of retained coefficients (approximants) N grows.

Compression using orthogonal transforms is an extension of NLA that consists of (a) approximation, (b) indexing the retained coefficients, and (c) quantization of the coefficients.[3] Thus, the MSE (in this case also called *distortion*) is affected by the two factors: (a) truncation error due to NLA and (b) quantization error.

The asymptotic *rate of decay* of the MSE, as N tends to infinity, is a fundamental approximation property of the transform and this value allows us to compare approximation performance of different transforms. The higher the rate of decay, the more efficient the transform is. Similarly, the rate of decay in compression is defined as the asymptotic behavior of the distortion D, as the bitrate R tends to infinity (this is frequently called *R-D behavior*).

Mallat [50] and DeVore [53] showed that, for a 2-D piecewise C^2 smooth signal $f(x_1,x_2)$ with a 1-D C^2 smooth discontinuity curve[4] (which we call C^2/C^2 signal), the lower bound of the MSE is given by $O(N^{-2})$.

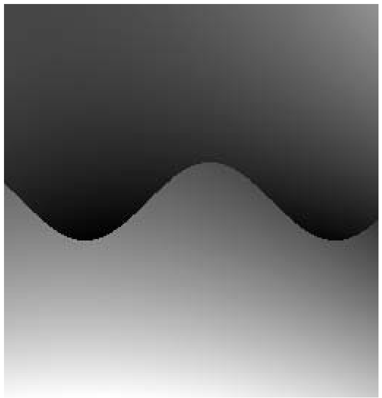
Notice that the standard WT is far from optimal since its rate of decay is $O(N^{-1})$ [1], [50]. Some other adaptive or non-adaptive methods have been shown to improve substantially the approximation power. Curvelets [18]–[20] and contourlets [21] can achieve the rate $O(N^{-2}(\log N)^3)$, which is nearly optimal. Furthermore, bandelets [11], [12] and wedgelets [13]–[17] have been shown to perform indeed optimally. However, notice that none of these methods is based on critically sampled filter-banks, which are very convenient for compression. Furthermore, a complex non-separable processing is sometimes required.

As we showed in Section II and III, anisotropy and multi-directionality improve the approximation power of the WT while keeping separability, simplicity, and critical sampling. However, the S-FSWT cannot yield a high rate of decay since it fails to provide a sparse representation of C^2/C^2 images. On the other hand, the S-AWT is capable of producing a compact representation, but it is still sensitive to the choice of the transform and alignment directions.

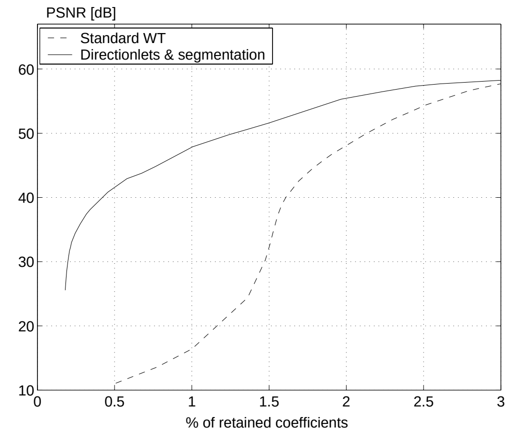
Synthetic (including also C^2/C^2) and natural images have geometrical features that vary over the space. Directionality, thus, can be considered as a local characteristic, defined in a small neighborhood. This implies the necessity for *spatial segmentation* as a way of partitioning an image into smaller segments with one or a few dominant directions per segment.

[3]Some algorithms merge quantization and NLA into a single operation producing an embedded bitstream, like zero-trees [51] or SPIHT [52].

[4]C^2 smoothness of both 1-D and 2-D functions means that the functions are twice continuously differentiable.



(a)



(b)

Fig. 13. An example of NLA of an image from the class C^2/C^2. (a) An image from the class C^2/C^2 is approximated using the standard WT and the S-AWT(Λ,2,1) with spatial segmentation. (b) The MSE expressed in terms of PSNR is significantly reduced in the case of the S-AWT(Λ,2,1).

The S-AWT is applied on a segmented image, where the transform and alignment directions are chosen independently in each segment. The transform outperforms the standard WT in both approximation and compression rate of decay of the MSE (i.e. distortion). The following theorem gives the rate of decay for C^2/C^2 images.

Theorem 1: Given a 2-D C^2/C^2 function $f(x_1, x_2)$ and $\alpha = (\sqrt{17} - 1)/2 \approx 1.562$,

(a) The N-term approximation by the S-AWT using spatial segmentation achieves

$$\text{MSE} = \|f - \hat{f}_N\|^2 = O\left(N^{-\alpha}\right).$$

In that case the optimal anisotropy ratio is $\rho^* = \alpha$.

(b) Compression by the S-AWT, using spatial segmentation and using R bits for encoding, can achieve the distortion D given by

$$D = O\left(R^{-\alpha}\right).$$

The proof of the theorem is given in Appendix II.

Notice that anisotropic segmentation is used here in the iteration, that is, an image is partitioned into vertical strips of equal widths. The number of segmentation steps depends on the anisotropy ratio, the number of approximants, the number of transform directions, and the first derivative of the C^2 curve (see the proof of Theorem 1). In particular, when the optimal anisotropy ratio $\rho^* = \alpha$ is used, the number of segmentation steps does not increase with the number of approximants. However, in reality, because of the discreteness of the transform, this anisotropy ratio cannot be exactly achieved and, in general, the number of segmentation steps has to be increased with the number of approximants. Notice that the S-AWT(Λ,3,2) approximates well the optimal transform[5] while retaining iterative

[5]There are other possible transforms with the anisotropy ratio even closer to optimal but we choose this one for the sake of simplicity.

TABLE III

DEPENDENCE OF THE APPROXIMATION RATE MSE=$O(N^{-e_1^*})$ AND THE NUMBER OF SEGMENTATION LEVELS $s = \eta_1 \log_2(N)$ ON THE GROWTH RATE OF THE NUMBER OF TRANSFORM DIRECTIONS β IN THE CASE OF THE S-AWT(Λ,3,2)

β	2	1	0.5	0.25
η_1	$1/51$	$1/26$	$2/27$	$4/29$
e_1^*	1.55	1.50	1.41	1.24

segmentation. It follows from the proof of Theorem 1 that the number of required transform directions grows with the number of segmentation steps as $O(2^{\beta s})$. Table III gives the achievable approximation and segmentation rates for the S-AWT(Λ,3,2) and different values of β.

Although the obtained approximation rate is slower than the ones obtained in [13]–[21], we want to emphasize that the S-AWT(Λ,3,2) is *critically sampled* and uses only *separable processing*. This is important for compression because, in the case of orthogonal 1-D filter-banks, the Lagrangian optimization-based algorithms still can be applied, making it easier to achieve very good compression.

In order to perform compression, the chosen transform directions in each segment have to be encoded together with the indexes and quantized values of the retained transform coefficients. The bitrate of this overhead information depends on the number of spatial segments and allowed number of transform directions per segment. Recall from Appendix II that the number of spatial segments is equal to 2^s, whereas the number of bits needed to encode the choice of directions in each segment behaves as $O(\log_2(2^{\beta s}))$=$O(\beta s)$. Thus, the number of overhead bits is given by $R_H = O(\beta s \cdot 2^s)$. However, even though this number grows exponentially with the number of segmentation steps s, the growth rate for the values of β given in Table III is smaller than the growth rate of the number of indexing and quantization bits and, thus, the dominant asymptotic behavior $D(R)$ remains the same.

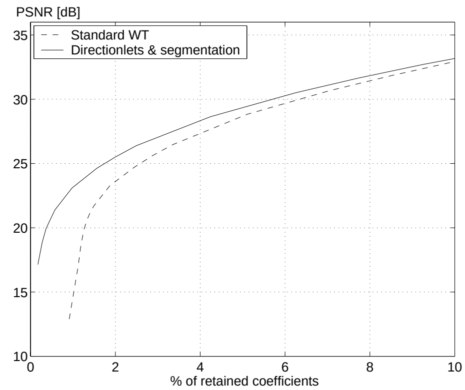
Recall also from Section III-C that the S-AWT(Λ,3,2) is applied in the $|\det(\mathbf{M}_\Lambda)|$ cosets separately. The separate filtering and subsampling in the cosets affect the order of decay of the MSE, but only up to a constant factor and, thus, the rate of decay remains the same.

Figure 13 illustrates the gain obtained by NLA using the S-AWT(Λ,2,1) with spatial segmentation applied on an image from the class C^2/C^2 when compared to the results of NLA obtained using the standard WT. Furthermore, Figure 14 shows an example of the NLA results with a natural image. The image Cameraman shown in Figure 14(a) is transformed using the standard 2-D WT without segmentation and the S-AWT(Λ,2,1) with segmentation. The MSE obtained by retaining a part of the transform coefficients is presented in Figure 14(b). The two reconstructions obtained with 0.98% of retained coefficients for the two methods are shown in Figure 14(c) and (d). Finally, the segmentation and adaptation of transform directions for the case in Figure 14(d) is illustrated in Figure 15.

V. CONCLUSION AND FUTURE WORK

We have proposed novel anisotropic transforms for images that use separable filtering in many directions, not only horizontal and vertical. The associated basis functions, called directionlets, have DVM along any two directions

(a)



(b)

(c)

(d)

Fig. 14. An example of NLA of a natural image. (a) The original image Cameraman. The image is approximated using the standard WT and the S-AWT(Λ,2,1) with spatial segmentation. For both transforms, the maximal decomposition level is 3. (b) The PSNR of the approximated image is significantly improved in the case of the anisotropic transform. (c) The reconstructed image obtained using the standard WT for 0.98% retained coefficients and quality of 13.93dB. (d) The reconstructed image obtained using directionlets with spatial segmentation for the same number of retained coefficients and quality of 23.09dB.

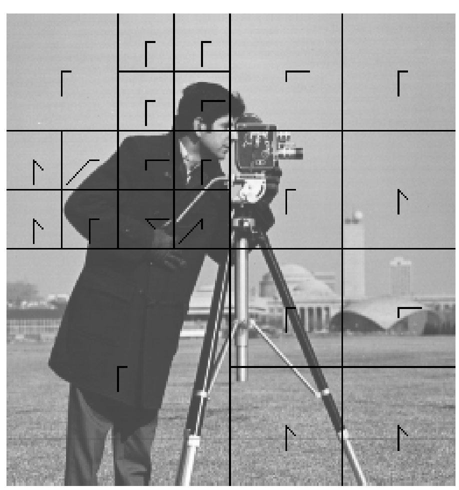


Fig. 15. The transform directions are adapted to the dominant directions in each segment of the image Cameraman shown in Figure 14(a).

with rational slopes. These transforms retain the computational efficiency and the simplicity of filter design from the standard WT. Still, multi-directionality and anisotropy overcome the weakness of the standard WT in presence of edges and contours, that is, they allow for sparser representations of these directional anisotropic features.

The NLA power of directionlets is substantially superior to that of the standard WT providing an order of decay of the MSE equal to $O(N^{-1.55})$ for the C^2/C^2 class of images. Even though this decay is slower than the one provided by some other schemes, the directionlets allow critical sampling. This is important for applications in image compression, since, in the case of orthogonal 1-D filter-banks, Lagrangian optimization can be implemented straightforwardly. For instance, the performance of the compression algorithm based on spatial-frequency quantization (SFQ) [54], [55] can be improved by replacing the standard WT with directionlets and allowing for adaptation of the transform and alignment directions and segmentation. Some details on the analysis of the applications of directionlets in image compression can be found in [49].

The directionlets built on digital lines using the 1-D oversampled transforms yield overcomplete tight frames (tightness is trivial as it follows from the tightness of the oversampled 1-D wavelet transforms). We distinguish this shift-invariant oversampling and the oversampling in directions as explained in Section IV. The redundant oversampled directionlets provide a promising framework for image denoising since they can efficiently capture geometrical structures in images [56]. An adaptive denoising algorithm that enforces coherence in images across space, scales, and directions is currently under investigation.

APPENDIX I
RELATION BETWEEN THE MSE IN THE ORIGINAL AND TRANSFORM DOMAINS

Assume that, given a frame $\mathbf{F} \in \mathbb{R}^{m\times n}$, the vector $\mathbf{y} \in \mathbb{R}^m$ is defined as $\mathbf{y} = \mathbf{F}\mathbf{x}$ for any $\mathbf{x} \in \mathbb{R}^n$. Here $m \geq n$. Recall that the inverse transform is given by $\mathbf{x} = (\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}^T \cdot \mathbf{y}$ [50]. Recall also that if the frame $\mathbf{F}$ is tight then $\|\mathbf{y}\|_2^2 = A\|\mathbf{x}\|_2^2$, where A is called the *frame bound*. Then, it also holds that

$$\mathbf{F}^T \cdot \mathbf{F} = A\mathbf{I}_n, \tag{8}$$

where $\mathbf{I}_n$ is the $n\times n$ identity matrix. In that case, the inverse transform is simplified and it is given by $\mathbf{x} = A^{-1}\mathbf{F}^T\cdot\mathbf{y}$.

Now, assume that a non-linear operator (e.g. NLA, thresholding, etc.) $T : \mathbb{R}^m \to \mathbb{R}^m$ is applied on $\mathbf{y}$ yielding $\hat{\mathbf{y}}$, that is, $\hat{\mathbf{y}} = T(\mathbf{y})$. It holds that $\hat{\mathbf{x}} = (\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}^T \cdot \hat{\mathbf{y}}$.

The MSE in the original domain is defined as $\|\mathbf{x} - \hat{\mathbf{x}}\|_2^2$ and, similarly, the MSE in the transform domain is given by $\|\mathbf{y} - \hat{\mathbf{y}}\|_2^2$. Assuming that the frame $\mathbf{F}$ is tight we can write

$$\|\mathbf{x} - \hat{\mathbf{x}}\|_2^2 = \|\frac{1}{A}\mathbf{F}^T \cdot (\mathbf{y} - \hat{\mathbf{y}})\,\|_2^2 \leq \frac{1}{A^2}\|\mathbf{F}^T\|_2^2 \cdot \|\mathbf{y} - \hat{\mathbf{y}}\|_2^2,$$

where equality holds when $\mathbf{F}$ is orthogonal.

From [57] we have that $\|\mathbf{F}^T\|_2^2 = \|\mathbf{F}\|_2^2 = A$. Hence, the MSE in the original and transform domains are related as

$$\|\mathbf{x} - \hat{\mathbf{x}}\|_2^2 \leq \frac{1}{A}\|\mathbf{y} - \hat{\mathbf{y}}\|_2^2.$$

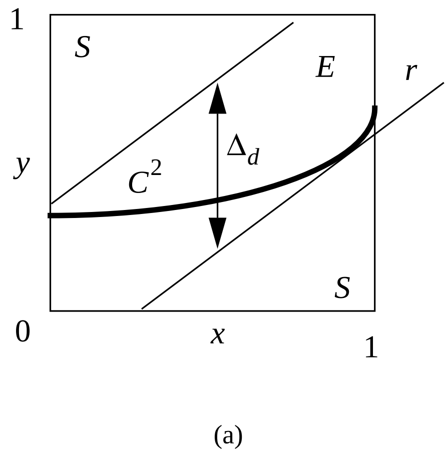


Fig. 16. The 2-D function $f(x_1, x_2)$ is C^2 smooth on the unit square away from a C^2 discontinuity curve. The curve can be locally approximated by a quadratic polynomial $y(x) = ax^2 + bx + c$. The E-type transform coefficients intersect the curve and have a slower decay of magnitudes across scales than the S-type coefficients, which correspond to the smooth regions. (a) The S-AWT produces the E-type coefficients within the strip along the slope r. (b) The width of the strip Δ_d is minimized for $r = a + b$.

APPENDIX II
PROOF OF THEOREM 1

Recall first that a C^2 curve can be locally represented by the Taylor series expansion, that is, by a quadratic polynomial

$$y(x) = ax^2 + bx + c, \tag{9}$$

where a and b are related to the second and first derivative of the curve (curvature and linear component), respectively. Without loss of generality, we assume that the C^2 discontinuity curve is *Horizon* [13] on the unit square $[0, 1]^2$.

Since the smooth regions of the function $f(x_1, x_2)$ are C^2, assume that the 1-D filters used in the S-AWT are orthogonal and have at least two vanishing moments. Let the transform be applied along the class of straight lines defined by

$$\{y(x) = rx + d : d \in \mathbb{R}\}. \tag{10}$$

Here, the slope r determines the transform direction, whereas the alignment direction is vertical. Equalizing (9) and (10) we can write

$$d(x) = ax^2 + (b - r)x + c.$$

The transform coefficients of the S-AWT that intersect the discontinuity curve are called *E-type coefficients*. The number of the E-type coefficients at the scale j is given by $N_e^{(0)}(j) = O(2^{n_2 j} \Delta_d)$. Here, n_2 is the number of transforms applied along the vertical direction, $\Delta_d = \max_{0 \le x \le 1} d(x) - \min_{0 \le x \le 1} d(x)$ is the width of the strip along the transform direction that contains the curve (see Fig. 16), and zero in the superscript of $N_e^{(0)}(j)$ denotes that no segmentation has been applied yet. The transform direction with the slope

$$r = a + b \tag{11}$$

minimizes the width Δ_d (and, thereof, $N_e^{(0)}(j)$) on the unit square. In that case the number of the E-type coefficients is given by

$$N_e^{(0)}(j) = O\left(\frac{a}{4}2^{n_2 j}\right).$$

Notice that an increment in the scale index j is equivalent to a step to a finer scale.

The transform coefficients of the S-AWT, which do not intersect the discontinuity curve are called *S-type coefficients*. The number of the S-type coefficients depends on the number of transforms n_1 and n_2 at a scale along the transform and vertical directions, respectively, as

$$N_s^{(0)}(j) = 2^{(n_1+n_2)j} - N_e^{(0)}(j) = O\left(2^{(n_1+n_2)j} - \frac{a}{4}2^{n_2 j}\right).$$

An anisotropic spatial segmentation is applied on the unit square. It partitions the unit square into vertical strips using the dyadic rule, that is, there are 2^s vertical strips at the sth level of segmentation, where the width of each is 2^{-s} (Fig. 17). The optimal transform direction, according to (11), is chosen for each segment independently. Since each segment is rescaled again to the unit square, the number of the E-type transform coefficients in a segment is reduced and is given by

$$O\left(\frac{a}{4}2^{n_2 j} \cdot 2^{-2s}\right).$$

The total number of the E-type coefficients is given by the sum across all the segments, that is,

$$N_e(j,s) = \sum_{k=0}^{2^s-1} O\left(\frac{a}{4}2^{n_2 j - 2s}\right) = O\left(\frac{a}{4}2^{n_2 j - s}\right). \tag{12}$$

Similarly, the total number of the S-type coefficients is given by

$$N_s(j,s) = \sum_{k=0}^{2^s-1} O\left(2^{(n_1+n_2)j} - \frac{a}{4}2^{n_2 j - 2s}\right) = O\left(2^{(n_1+n_2)j+s} - \frac{a}{4}2^{n_2 j - s}\right). \tag{13}$$

Notice that the exact number of the two types of coefficients given by (12) and (13) depends on the length of the 1-D filters used in the transform. However, the dependence is only up to a constant and, thus, the order of growth of these numbers across scales remains the same.

The magnitudes $|w_e(j)|$ of the E-type coefficients decay across scales as $O(2^{-(n_1+n_2)j/2})$. The S-type coefficients correspond to the smooth regions of the function $f(x_1,x_2)$ and their magnitudes $|w_s(j)|$ are upper bounded by $O(2^{-n_3 j/2})$. Notice that, since the 1-D HP filters have vanishing moments, the decay of the magnitudes of the S-type coefficients is faster than the one of the E-type coefficients, that is, $n_3 > n_1 + n_2$.

We estimate n_3 considering that the applied 1-D wavelets have at least two vanishing moments. It is shown in [50] that, the decay of the magnitudes $|w_s(j)|$ in a smooth region after two consecutive transforms with alternated transform directions is 2^{-3}. Therefore, the decay rate n_3 is given by

$$n_3 = 6 \cdot \min(n_1,n_2) + |n_2 - n_1| = \begin{cases} n_1 + 5n_2, & n_1 \geq n_2 \\ 5n_1 + n_2, & n_1 \leq n_2 \end{cases}. \tag{14}$$

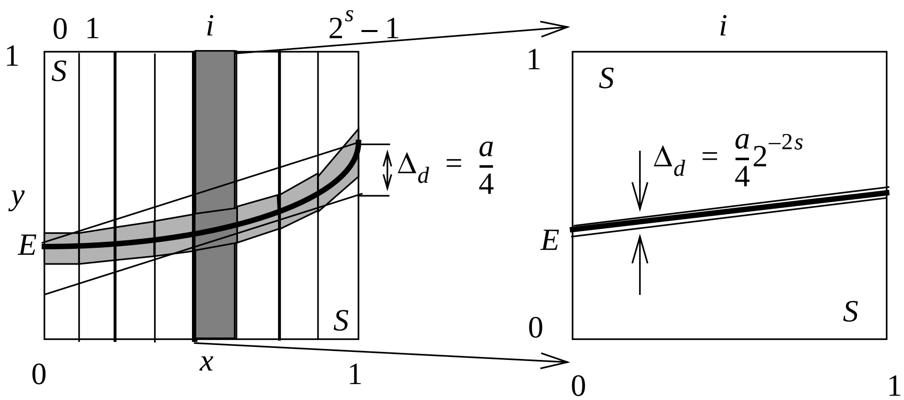


Fig. 17. Anisotropic segmentation partitions the unit square into 2^s equally wide vertical strips. After rescaling, the curvature parameter a (related to the second derivative of the C^2 curve) is reduced in each segment by the factor 2^{2s}. Since there are 2^s segments that intersect the discontinuity, the total number of the E-type transform coefficients is reduced by 2^s. At the same time, the total number of the S-type coefficients is increased by the same factor.

To approximate the function $f(x_1, x_2)$, we keep all the coefficients with the magnitudes larger than or equal to the threshold 2^{-m}, where $m \geq 0$, and discard (set to zero) the others. The retained coefficients can be divided into two groups:

(1) The E-type coefficients at the scales $0 \leq j \leq 2m/(n_1+n_2)$,

(2) The S-type coefficients at the scales $0 \leq j \leq 2m/n_3$.

From (12), (13) and decays of the magnitudes across scales, we compute the order of the total number of retained coefficients $N(m,s)$ and the corresponding MSE. The number $N(m,s)$ is the sum of the retained E and S-type coefficients:

$$\begin{aligned} N(m,s) &= \sum_{j=0}^{2m/(n_1+n_2)} N_e(j,s) + \sum_{j=0}^{2m/n_3} N_s(j,s) \\ &= O\left(2^{\frac{2n_2}{n_1+n_2}m - s}\right) + O\left(2^{\frac{2(n_1+n_2)}{n_3}m + s}\right). \end{aligned} \tag{15}$$

The MSE is given by

$$\begin{aligned} \text{MSE}(m,s) &= \sum_{j=2m/(n_1+n_2)+1}^{+\infty} N_e(j,s)|w_e(j)|^2 + \sum_{j=2m/n_3+1}^{+\infty} N_s(j,s)|w_s(j)|^2 \\ &= O\left(2^{-\frac{2n_1}{n_1+n_2}m - s}\right) + O\left(2^{-\frac{2(n_3-n_1-n_2)}{n_3}m + s}\right). \end{aligned} \tag{16}$$

Assuming that the number of segmentation levels depends on the exponent m of the threshold as $s = \eta m$, where the *segmentation rate* $\eta \geq 0$, we distinguish the two cases, as follows:

(1) The terms in (15) and (16) produced by the E-type coefficients dominate, in which case we have

$$\eta \leq \eta^* = \frac{n_2}{n_1+n_2} - \frac{n_1+n_2}{n_3} = \frac{1}{\rho+1} - \frac{\rho+1}{\rho+5},$$

where $\rho = n_1/n_2 \geq 1$. Then the MSE decays as

$$\text{MSE} = O\left(N^{-e_1}\right), \text{ where } e_1 = \frac{2n_1 + \eta(n_1+n_2)}{2n_2 - \eta(n_1+n_2)} = \frac{2\rho + \eta(\rho+1)}{2-\eta(\rho+1)}.$$

(2) The terms in (15) and (16) produced by the S-type coefficients dominate, that is, $\eta \geq \eta^*$ and

$$\text{MSE} = O\left(N^{-e_2}\right), \text{ where } e_2 = \frac{2(n_3 - n_1 - n_2) - \eta n_3}{2(n_1+n_2) + \eta n_3} = \frac{8 - \eta(\rho+5)}{2(\rho+1) + \eta(\rho+5)}.$$

Plugging (14) in the relations above and knowing that the segmentation rate η is a non-negative value, we obtain the maximal decay rate MSE= $O(N^{-\alpha})$, with $\alpha = (\sqrt{17}-1)/2 \approx 1.562$. The optimal rate is attained for the anisotropy ratio $\rho^* = n_1/n_2 = \alpha \approx 1.562$ and the segmentation rate $\eta^* = 0$.

Notice that the analysis above is based on two assumptions: (a) the optimal transform direction given by (11) is chosen and (b) the C^2 curve is globally represented by a quadratic polynomial given by (9). Here, we address these two assumptions showing that they do not constrain severely the approximation rate.

(a) Assume that the transform direction is given by the suboptimal slope $r = a + b + \epsilon$, where $|r| \leq 1$. Then it can be shown that $\Delta_d = a/4 + |\epsilon|/2 + \epsilon^2/4a = O(a)$ for $|\epsilon| \leq a$ and $\Delta_d = |\epsilon| + o(\epsilon)$ for $|\epsilon| > a$. Furthermore, assume that ϵ decays exponentially with the number of segmentation steps, that is, $\epsilon \sim 2^{-\beta s}$, where $\beta > 0$. If $\beta < 2$, then the expression of $N_e(j,s)$ given by (12) becomes $O(2^{n_2 j - (\beta-1)s})$ and the optimal segmentation rate η^* is multiplied by the factor $2/\beta$. In that case the exponent e_1 is given by

$$e_1 = \frac{2\rho + (\beta-1)\eta(\rho+1)}{2 - (\beta-1)\eta(\rho+1)},$$

whereas the exponent e_2 is unchanged. However, even though some of these parameters are changed, the optimal approximation rate remains the same, that is, MSE= $O(N^{-\alpha})$ if $\rho^* = \alpha$ and $\eta^* = 0$. On the other hand, the required number of transform directions is finite now and behaves as $1/|\epsilon| \sim 2^{\beta s}$.

(b) The analysis that leads to the approximation rate holds only for the case when the slope of the tangent direction (or, equivalently, the first derivative) of the C^2 curve is in the interval $[-1, 1]$. However, the first derivative of a general C^2 curve is not constrained on that interval and, therefore, the optimal approximation rate cannot be achieved in the same way as in the case of a quadratic polynomial. In order to be able to achieve the same rate we need to introduce an initial number of segmentation steps prior to the iteration. Recall that one step of anisotropic segmentation attenuates twice the first derivative of the C^2 curve.[6] Thus, it suffices to apply enough segmentation steps so that the maximal magnitude of the first derivative is less than or equal to 1. Then, the iterated segmentation and transform are continued on each of these initial segments and this construction results in the same optimal approximation rate. Notice that the necessity for reducing the magnitude of the first derivative below 1 is caused by the assumption that the C^2 curve is Horizon. However, if this assumption is not satisfied, then an appropriate combination of initial segmentation steps and transposition of the axes can

[6]One step of the anisotropic segmentation is equivalent to stretching the abscissa by the factor 2 and, therefore, the equivalent first derivative of the curve is also attenuated by 2.

rescale the curve so that each segment of the curve is Horizon. Therefore, the optimal approximation rate can be achieved in the case of a general C^2 curve.

For the compression application, the retained coefficients have to be indexed and quantized. For a given MSE (or *distortion*) each of these operations carries a cost in terms of the required bits.

The N retained S-AWT coefficients within a spatial segment can be organized in an embedded tree-structure, similar to the structures produced by the standard WT and exploited in the other compression algorithms (zero-trees [51], SPIHT [52], SFQ [54], [55]). The main difference between the tree-structures of the standard WT and S-AWT is in the number of descendants of each transform coefficient. While this number is fixed in the standard WT, it depends on the number of transform steps applied at each scale in the S-AWT. However, the S-AWT tree-structure allows also for indexing the retained coefficients using approximately 1 bit per transform coefficient.

A variable length coding scheme allocates l bits to encode coefficients with magnitudes in the interval $[2^{-m}2^{l-1}, 2^{-m}2^{l})$. Thus, using (15) and the optimal choice for n_1, n_2, n_3, and η, the total number of encoding bits R is given by:

$$\begin{aligned} R(m) &= N(m,0) + \sum_{l=1}^{\infty} N(m-l,0) \\ &= O\left(2^{\frac{\alpha}{2}m}\right) + \sum_{l=1}^{\infty} 2^{\frac{\alpha}{2}(m-l)} = O\left(2^{\frac{\alpha}{2}m}\right). \end{aligned} \tag{17}$$

The distortion D consists of two components: (a) the MSE resulting from the truncation of small coefficients in the approximation given by (16), and (b) distortion caused by the quantization of the retained coefficients. The second component is given by $N(m,0)\cdot 2^{-2m}$ and, thus, the total distortion is

$$D(m) = \text{MSE}(m,0) + N(m,0)\cdot 2^{-2m} = O\left(2^{-\frac{\alpha^2}{2}m}\right). \tag{18}$$

The R-D behavior follows from (17) and (18) and it is given by

$$D(R) = O\left(R^{-\alpha}\right).$$

References

[1] D. L. Donoho, M. Vetterli, R. A. DeVore, and I. Daubechies, “Data compression and harmonic analysis,” *IEEE Trans. Inform. Theory*, vol. 44, pp. 2435–2476, Oct. 1998.

[2] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*. New Jersey, NJ: Prentice Hall PTR, 1995.

[3] A. Skodras, C. Christopoulos, and T. Ebrahimi, “The JPEG 2000 still image compression standard,” *IEEE Signal Processing Mag.*, vol. 18, pp. 36–58, Sept. 2001.

[4] A. Cohen and I. Daubechies, “Non-separable bidimensional wavelet bases,” *Revista Matematica Iberoamericana*, vol. 9, no. 1, pp. 51–137, 1993.

[5] J. Kovačević, “Filter banks and wavelets: Extensions and applications,” Ph.D. dissertation, Graduate School of Arts and Sciences, Columbia Univ., New York, NY, 1991.

[6] J. Kovačević and M. Vetterli, “Nonseparable multidimensional perfect reconstruction filter banks and wavelet bases for $\mathbb{R}^n$,” *IEEE Trans. Inform. Theory*, pp. 533–555, Mar. 1992.

[7] R. H. Bamberger and M. J. T. Smith, “A filter bank for the directional decomposition of images: Theory and design,” *IEEE Trans. Signal Processing*, pp. 882–893, Apr. 1992.

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

[9] T. Chen and P. P. Vaidyanathan, “Multidimensional multirate filters and filter banks derived from one-dimensional filters,” *IEEE Trans. Signal Processing*, vol. 41, no. 5, pp. 1749–1765, May 1993.

[10] D. B. H. Tay and N. G. Kingsbury, “Flexible design of multidimensional perfect reconstruction FIR 2-band filters using transformations of variables,” *IEEE Trans. Image Processing*, vol. 2, no. 4, pp. 466–480, Oct. 1993.

[11] E. L. Pennec and S. Mallat, “Image compression with geometric wavelets,” in *Proc. IEEE International Conference on Image Processing (ICIP2000)*, Vancouver, Canada, Sept. 2000, pp. 661–664.

[12] ——, “Sparse geometric image representations with bandelets,” *IEEE Trans. Image Processing*, vol. 14, pp. 423–438, Apr. 2005.

[13] D. L. Donoho, “Wedgelets: Nearly-minimax estimation of edges,” *Ann. Statist.*, vol. 27, pp. 859–897, 1999.

[14] J. K. Romberg, M. Wakin, and R. Baraniuk, “Multiscale wedgelet image analysis: fast decompositions and modeling,” in *Proc. IEEE International Conference on Image Processing (ICIP2002)*, Rochester, NY, Sept. 2002, pp. 585–588.

[15] ——, “Approximation and compression of piecewise smooth images using a wavelet/wedgelet geometric model,” in *Proc. IEEE International Conference on Image Processing (ICIP2003)*, Barcelona, Spain, Sept. 2003, pp. 49–52.

[16] M. Wakin, J. Romberg, C. Hyeokho, and R. Baraniuk, “Rate-distortion optimized image compression using wedgelets,” in *Proc. IEEE International Conference on Image Processing (ICIP2002)*, Rochester, NY, Sept. 2002, pp. III–237 – III–240.

[17] M. Wakin, J. Romberg, H. Choi, and R. Baraniuk, “Wavelet-domain approximation and compression of piecewise smooth images,” *IEEE Trans. Image Processing*, submitted for publication.

[18] E. J. Candès and D. L. Donoho, “Curvelets - a surprisingly effective nonadaptive representation for objects with edges,” in *Curve and Surface Fitting*, A. Cohen, C. Rabut, and L. L. Schumaker, Eds. Saint-Malo: Vanderbilt University Press, 1999.

[19] E. J. Candès and D. Donoho, “Curvelets and curvilinear integrals,” Department of Statistics, Stanford University, CA, Tech. Rep., Dec. 1999.

[20] E. J. Candès and D. L. Donoho, “New tight frames of curvelets and optimal representations of objects with smooth singularities,” Department of Statistics, Stanford University, Tech. Rep., 2002.

[21] M. N. Do and M. Vetterli, “The contourlet transform: An efficient directional multiresolution image representation,” *IEEE Trans. Image Processing*, to appear.

[22] R. Shukla, P. L. Dragotti, M. N. Do, and M. Vetterli, “Rate-distortion optimized tree-structured compression algorithms for piecewise polynomial images,” *IEEE Trans. Image Processing*, vol. 14, pp. 343–359, Mar. 2005.

[23] P. L. Dragotti and M. Vetterli, “Footprints and edgeprints for image denoising and compression,” in *Proc. IEEE International Conference on Image Processing (ICIP2001)*, Thessaloniki, Greece, Oct. 2001.

[24] A. Cohen and B. Matei, “Compact representation of images by edge adapted multiscale transforms,” in *Proc. IEEE International Conference on Image Processing (ICIP2001)*, Thessaloniki, Greece, Oct. 2001, invited paper.

[25] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, “Shiftable multiscale transforms,” *IEEE Trans. Inform. Theory*, vol. 38, pp. 587–607, Mar. 1992.

[26] 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.

[27] N. Kingsbury, “Complex wavelets for shift invariant analysis and filtering of signals,” *Journal of Appl. and Comput. Harmonic Analysis*, vol. 10, pp. 234–253, 2001.

[28] R. A. Zuidwijk, “Directional and time-scale wavelet analysis,” *SIAM Journal on Mathematical Analysis*, vol. 31, no. 2, pp. 416–430, 2000.

[29] J. G. Rosiles and M. J. T. Smith, “A low complexity overcomplete directional image pyramid,” in *Proc. IEEE International Conference on Image Processing (ICIP2003)*, Barcelona, Spain, Sept. 2003.

[30] F. G. Meyer and R. R. Coifman, “Brushlets: a tool for directional image analysis and image compression,” *Journal of Appl. and Comput. Harmonic Analysis*, vol. 5, pp. 147–187, 1997.

[31] G. H. Granlund and H. Knutsson, “Compact associative representation of visual information,” in *Proc. on 10th International Conference on Pattern Recognition*, vol. ii, 16-21, June 1990, pp. 200–207.

[32] X. Li and M. T. Orchard, “New edge-directed interpolation,” *IEEE Trans. Image Processing*, vol. 10, no. 10, pp. 1521–1527, Oct. 2001.

[33] K. Hirakawa and T. W. Parks, "Adaptive homogeneity-directed demosaicing algorithm," *IEEE Trans. Image Processing*, vol. 14, no. 3, pp. 360–369, Mar. 2005.

[34] D. D. Muresan and T. W. Parks, "Prediction of image detail," in *Proc. IEEE International Conference on Image Processing (ICIP2000)*, Vancouver, BC, Canada, Sept. 2000, pp. 323–326.

[35] ——, "Adaptive optimal-recovery image interpolation," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP2001)*, Salt Lake City, UT, May 2001, pp. 1949–1952.

[36] A. L. da Cunha and M. N. Do, "Bi-orthogonal filter banks with directional vanishing moments," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP2005)*, Philadelphia, PA, Mar. 2005, pp. 553–556.

[37] A. L. Cunha, J. Zhou, and M. N. Do, "The nonsubsampled contourlet transform: theory, design, and applications," *IEEE Trans. Image Processing*, submitted for publication, 2005.

[38] V. Velisavljević, B. Beferull-Lozano, M. Vetterli, and P. L. Dragotti, "Approximation power of directionlets," in *Proc. IEEE International Conference on Image Processing (ICIP2005)*, Genova, Italy, Sept. 2005.

[39] J. Milner, *Mondrian*. London, UK: Phaidon, 2002.

[40] P. H. Westerink, "Subband coding of images," Ph.D. dissertation, Delft University of Technology, Delft, The Netherlands, 1989.

[41] R. D. Nowak and R. G. Baraniuk, "Wavelet-based transformations for nonlinear signal processing," *IEEE Trans. Signal Processing*, vol. 47, pp. 1852–1865, July 1999.

[42] C. P. Rosiene and T. Q. Nguyen, "Tensor-product wavelet vs. Mallat decomposition: A comparative analysis," in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS1999)*, Orlando, FL, June 1999, pp. 431–434.

[43] J. E. Bresenham, "Algorithm for computer control of a digital plotter," *IBM Systems Journal*, vol. 4, no. 1, pp. 25–30, 1965.

[44] J. D. Foley, A. V. Dam, S. K. Feiner, and J. F. Hughes, *Computer graphics: Principles and Practice*. Reading, MA: Adilson-Wesley Publishing Company, 1990.

[45] T. S. Chan and R. K. K. Yip, "Line detection algorithm," in *Proc. IEEE International Conference on Pattern Recognition*, vol. 2, 1996, pp. 126–130.

[46] J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices and groups*. New York, NY: Springer, 1998.

[47] V. Velisavljević, B. Beferull-Lozano, M. Vetterli, and P. L. Dragotti, "Discrete multi-directional wavelet bases," in *Proc. IEEE International Conference on Image Processing (ICIP2003)*, Barcelona, Spain, Sept. 2003, pp. 1025–1028.

[48] E. Viscito and J. P. Allebach, "The analysis and design of multidimensional FIR perfect reconstruction filter banks for arbitrary sampling lattices," *IEEE Trans. Circuits Syst.*, pp. 29–41, Jan. 1991.

[49] V. Velisavljević, "Directionlets: anisotropic multi-directional representation with separable filtering," Ph.D. dissertation, School of Computer and Communication Sciences, EPFL, Lausanne, Switzerland, 2005.

[50] S. Mallat, *A Wavelet Tour of Signal Processing*. San Diego, CA: Academic Press, 1997.

[51] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Processing*, vol. 41, pp. 3445–3463, Dec. 1993.

[52] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp. 243–250, June 1996.

[53] R. A. DeVore, "Nonlinear approximation," *Acta Numer.*, vol. 7, pp. 51–150, 1998.

[54] Z. Xiong, K. Ramchandran, and M. T. Orchard, "Space-frequency quantization for wavelet image coding," *IEEE Trans. Image Processing*, vol. 6, pp. 677–693, May 1997.

[55] ——, "Wavelet packet image coding using space-frequency quantization," *IEEE Trans. Image Processing*, vol. 7, pp. 892–898, June 1998.

[56] V. Velisavljević, P. L. Dragotti, and M. Vetterli, "Directional wavelet transforms and frames," in *Proc. IEEE International Conference on Image Processing (ICIP2002)*, Rochester, NY, Sept. 2002, pp. 589–592.

[57] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. New Jersey, NJ: Prentice Hall, 2000.