APPROXIMATION POWER OF DIRECTIONLETS

Vladan Velisavljević[1], *Baltasar Beferull-Lozano*[1], *Martin Vetterli*[1,2], *Pier Luigi Dragotti*[3]

[1]School of C&C Sciences, EPFL, CH-1015 Lausanne, Switzerland, [2]EECS Dep., University of California at Berkeley, CA
[3]Dep. of Electrical and Electronic Engineering, Imperial College, London, UK

ABSTRACT

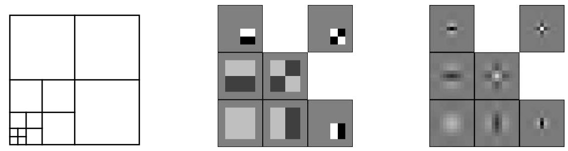
In spite of the success of the standard wavelet transform (WT) in image processing, the efficiency of its representation is limited by the spatial isotropy of its basis functions built in only horizontal and vertical directions. One-dimensional (1-D) discontinuities in images (edges and contours), which are very important elements in visual perception, intersect too many wavelet basis functions and reduce the sparsity of the representation. To capture efficiently these anisotropic geometrical structures, 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 (with the corresponding basis functions called *directionlets*) that retains the *separable* filtering and simple filter design from the standard two-dimensional (2-D) WT and imposes *directional vanishing moments* (DVM). Furthermore, we show that this novel transform has non-linear approximation efficiency competitive to the other previously proposed *over-sampled* transform constructions.

1. 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.

The 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]. The application of wavelets to image processing requires the design of 2-D wavelet filter-banks. The most common approach is to use 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.

However, the standard 2-D WT fails to provide a compact representation of 1-D discontinuities, like edges and contours. These objects are, in general, anisotropic and can have a certain dominant direction different from horizontal or vertical. Many wavelets intersect the discontinuity yielding many large magnitude coefficients. The main 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 the same number of times across the horizontal and vertical directions at each scale. As a result, the corresponding basis functions, obtained as direct products of the 1-D counterparts, are isotropic (Fig. 1).



(a) (b) (c)

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

Thus, to capture properly the geometrical coherence of contours and edges, the basis functions are required (a) to be *anisotropic* and (b) to have *multi-directional vanishing moments*. However, ensuring an efficient matching between anisotropic oriented basis functions and objects in images is a non-trivial task. This issue has already been studied and several adaptive (bandelets [2] and wedgelets [3, 4]) and non-adaptive (curvelets [5] and contourlets [6]) methods have been proposed. These methods build dictionaries of anisotropic oriented basis functions that provide a sparse representation of edges in images. Furthermore, Candès and Donoho [5] showed that the key to achieving a good non-linear approximation (NLA) behavior is the parabolic scaling relation between the length and width of anisotropic basis functions. On the other hand, the importance of M-DIR processing has been also emphasized in the cortex transform [7], the steerable pyramid [8], and the associative representation of visual information [9].

However, the implemented transforms often require *oversampling*, have *higher complexity* when compared to the standard WT, and require *non-separable* convolution and filter design. Furthermore, in some of these constructions (like curvelets) the design of the associated filters is performed in the *continuous domain* and this makes it difficult to use them directly on discrete images.

Several other approaches also analyze geometrical structures in images, like polynomial modeling with quadtree segmentation [10] and footprints and edgeprints [11]. However, all of them fail to provide a *perfect reconstruction* and *critically sampled separable* scheme while keeping a filter design completely in the *discrete domain* and with filters having DVM along *arbitrary directions*.

This work is supported by the SNF, Grant No. 200020-103729 and, in part, by EPSRC, Grant No. GR/557631/01.

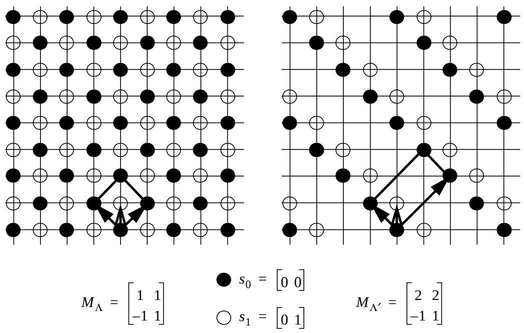


Fig. 2. The lattice Λ is determined by the generator matrix $\mathbf{M}_\Lambda = [\mathbf{d}_1, \mathbf{d}_2]^T$. 1-D filtering is applied on the pixels aligned with the vector $\mathbf{d}_1 = [1, 1]$, that is, along 45°. The pixels retained after the subsampling belong to the lattice $\Lambda' \subset \Lambda$ determined by the generator matrix $\mathbf{M}_{\Lambda'} = [2\mathbf{d}_1, \mathbf{d}_2]^T$. Filtering and subsampling are applied separately in two cosets, determined by the shift vectors $\mathbf{s}_0$ and $\mathbf{s}_1$.

Our goal is to construct an anisotropic perfect reconstruction and critically sampled basis functions with DVM, which we call *directionlets*, while retaining the simplicity of 1-D processing and filter design from the standard separable 2-D WT. Our basis construction uses the concept of integer lattices [12]. We show that our transform has good approximation properties as compared to the approximation achieved by the other overcomplete transform constructions and is superior to the performance of the standard separable 2-D WT while having the same complexity.

The outline of the paper is as follows. We give a review of integer lattices in Section 2. This concept is used in the construction of our skewed anisotropic lattice-based transforms, as presented in Section 3. In Section 4 we show that the achievable asymptotic approximation power using the skewed anisotropic transforms is $O(N^{-1.55})$. Finally, we conclude in Section 5.

2. REVIEW OF LATTICE-BASED FILTERING

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 [12]. 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

$$\mathbf{M}_\Lambda = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} = \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix}, \text{ where } a_1, a_2, b_1, b_2 \in \mathbb{Z}. \quad (1)$$

Recall that the cubic lattice $\mathbb{Z}^2$ can be partitioned into $|\det(\mathbf{M}_\Lambda)|$ cosets of the lattice Λ [12], where each coset is determined by the shift vector $\mathbf{s}_k$, for $k = 0, 1, \ldots, |\det(\mathbf{M}_\Lambda)| - 1$. Now we apply the 1-D WT (i.e. the 1-D filtering and subsampling operations) on the pixels aligned with the vector $\mathbf{d}_1$. After subsampling, the retained points belong to the sublattice Λ' of the lattice Λ ($\Lambda' \subset \Lambda$) with the corresponding generator matrix given by (see Fig. 2) [13]

$$\mathbf{M}_{\Lambda'} = \mathbf{D}_s \cdot \mathbf{M}_\Lambda = \begin{bmatrix} 2\mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix}, \mathbf{D}_s = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}.$$

Notice that both filtering and subsampling are applied in each of the cosets separately. This property allows for an efficient iteration of the transforms, as explained in [13, 14]. Furthermore, each filtering operation is purely 1-D. 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*.

3. SKEWED ANISOTROPIC WAVELET TRANSFORMS

As explained in Section 1 the standard WT produces isotropic basis functions with vanishing moments along horizontal and vertical directions, which fail to provide a sparse representation of edges and contours oriented in any other direction. However, a new modified method that we propose retains the 1-D filtering and subsampling operations and can provide both anisotropy and DVM, as we show next. We propose a construction of the anisotropic transforms with DVM along *any* two directions with rational slopes that still inherits the simplicity of processing and filter design from the standard 2-D WT. Furthermore, these anisotropic M-DIR transforms are critically sampled and lead to *perfect reconstruction*.

Using integer lattices we define the new transforms, which are called *skewed anisotropic wavelet transforms* (S-AWT). Given a lattice Λ, the S-AWT consists of the 1-D transforms applied along the transform and alignment directions of the lattice Λ. We denote as S-AWT($\mathbf{M}_\Lambda$,n_1,n_2) the skewed anisotropic transform that has n_1 and n_2 transforms in one iteration steps along the transform and alignment directions, respectively. The *anisotropy ratio* $\rho = n_1/n_2$ determines the elongation of the basis functions of the S-AWT($\mathbf{M}_\Lambda$,n_1,n_2). We call the basis functions of the S-AWT *directionlets* since they are anisotropic and have a specific orientation. An example of the anisotropic construction and frequency decomposition for S-AWT($\mathbf{M}_\Lambda$,2,1) are schematically shown in Fig. 3(a) and (b). The corresponding directionlets are shown in Fig. 3(c) and (d). Notice that the S-AWT($\mathbf{M}_\Lambda$,n_1,n_2) is applied in all cosets of the lattice Λ separately. Notice also that the standard 2-D WT is equivalent to the S-AWT($\mathbf{I}$,1,1), where $\mathbf{I}$ is the identity matrix. Directionlets have DVM in *any* two directions with rational slopes.[1] The following proposition gives the number and directions of the DVM in directionlets. For reasons of lack of space, the proof is omitted here and can be found in [14].

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

The S-AWT improves the efficiency of representation of images that contains anisotropic structures in different directions as explained in the sequel.

4. NON-LINEAR APPROXIMATION

The main task of approximation is to represent a signal by a subset of transform coefficients, whereas the rest of them is set to zero. We distinguish between linear approximation and NLA. In the first, the indices of the retained coefficients are fixed, whereas in the latter, they are adapted to the signal. The mean-square error (MSE) of approximation is given by $\|\mathbf{x} - \hat{\mathbf{x}}\|^2 = \|\mathbf{y} - \hat{\mathbf{y}}\|^2$, where $\mathbf{y}$ is an orthogonal transformed version of $\mathbf{x}$. Compression using orthogonal transforms is, roughly speaking, an extension of NLA

[1] Recall that an Lth order DVM along the direction with a rational slope $r = b/a$ is equivalent to requiring the z-transform of a basis function to have a factor $(1 - z_1^{-a} z_2^{-b})^L$.

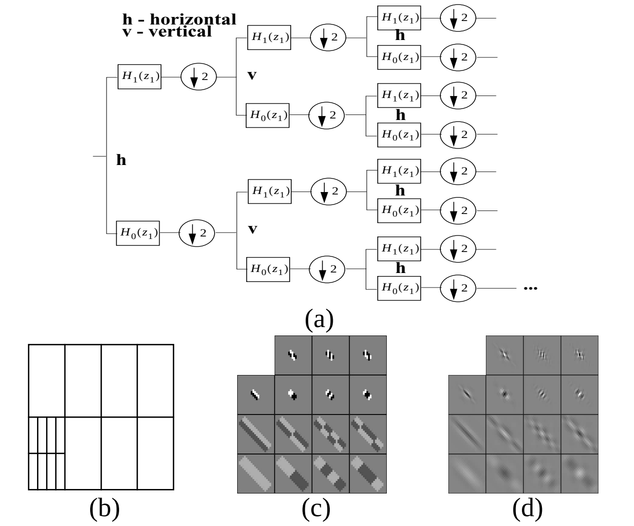


Fig. 3. The S-AWT allows for an anisotropic iteration of the filtering and subsampling operations applied along the transform and alignment directions of the lattice Λ. The transform is iterated on the low-pass, similarly as in the standard 2-D WT. (a) The filtering scheme for the S-AWT($\mathbf{M}_\Lambda$,2,1). One iteration step is shown. (b) The decomposition in frequency for two iterations. The basis functions obtained from the (c) Haar and (d) biorthogonal "9-7" 1-D scaling and wavelet functions.

that consists of (a) approximation, (b) indexing the retained coefficients, and (c) quantization of the coefficients.[2] Thus, the MSE in this case is affected by 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 very important approximation property of the transform used in NLA and compression. Mallat [15] showed that for a C^2 smooth 2-D signal $f(x_1, x_2)$ away from a C^2 discontinuity curve[3] (which we call a C^2/C^2 signal) the lower bound of the achievable MSE for any approach in construction is given by $O(N^{-2})$. Notice that the standard WT is far from optimal since its rate of decay is $O(N^{-1})$ [15]. Some other adaptive or non-adaptive methods have been shown to improve substantially the approximation power. Curvelets [5] and contourlets [6] can achieve the rate $O((\log N)^3 N^{-2})$, which is nearly optimal. Furthermore, bandelets [2] and wedgelets [3, 4] have been shown to perform indeed optimally. However, notice that none of these methods is based on critically sampled filter-banks that are very convenient for compression. Furthermore, in some cases, a complex non-separable processing is required.

4.1. Anisotropic Spatial Segmentation

The approximation efficiency of directionlets is sensitive to the choice of the transform and alignment directions. However, notice that synthetic (including also C^2/C^2) and natural images have geometrical features whose orientations vary over 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 only one or a few dominant directions per segment.

[2] Some algorithms merge quantization and NLA into a single operation producing an embedded bitstream.

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

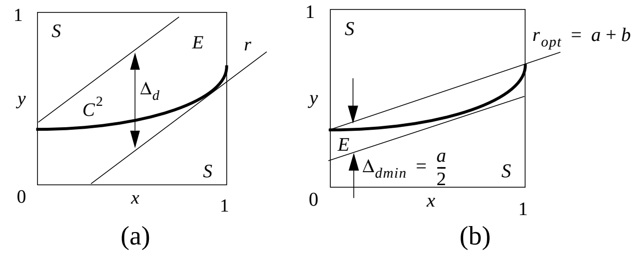


Fig. 4. The curve in the C^2/C^2 2-D function $f(x_1, x_2)$ can be locally approximated by a quadratic polynomial $y(x) = ax^2 + bx + c$. The E-type directionlets intersect the curve and have a slower decay of magnitudes across scales than the S-type directionlets, which correspond to the smooth regions. (a) The E-type directionlets lie within the strip along the slope r. (b) The width of the strip Δ_d is minimized for $r = a + b$.

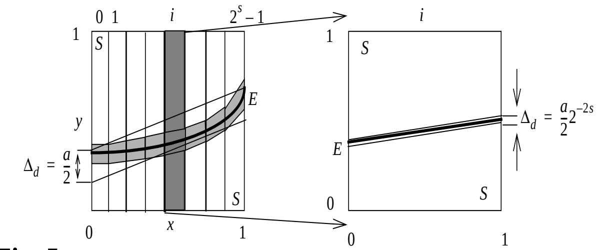


Fig. 5. Anisotropic segmentation partitions the unit square into 2^s equally wide vertical strips. The equivalent curvature is reduced in each segment by the factor 2^{2s}. Since there are 2^s segments that intersect the discontinuity, the total number of the E-type directionlets is reduced by 2^s. At the same time, the total number of the S-type coefficients is increased by the same factor.

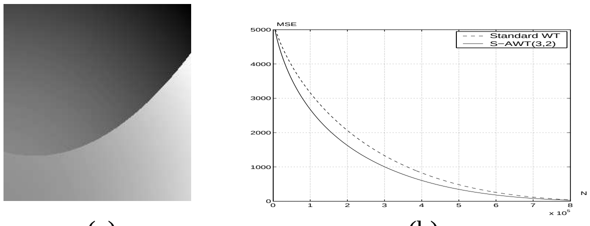
Now, recall that a C^2 curve can be locally represented by the Taylor series expansion, that is, by a quadratic polynomial $y(x) = ax^2 + bx + c$, where a and b are related to the second and first derivative of the curve (curvature and linear component), respectively. Without loss of generality, we assume that the C^2 discontinuity curve is *Horizon* [3] 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($\mathbf{M}_\Lambda$,n_1,n_2) are orthogonal and have at least two vanishing moments. Let the transform be applied along the class of straight lines defined by $\{y(x) = rx + d : d \in \mathbb{R}\}$. Here, the slope r determines the transform direction, whereas the alignment direction is vertical.

The directionlets that intersect the discontinuity curve are called *E-type*. The number of the E-type directionlets at the scale j is given by $N_e(j) = O(2^{n_2 j}\Delta_d)$. Here, Δ_d is the width of the strip along the transform direction that contains the curve (see Fig. 4). Notice that an increment in the scale index j is equivalent to an iteration step to a finer scale. It can be easily seen that the transform direction with the slope $r = a + b$ minimizes the width Δ_d (and, thereof, $N_e(j)$) on the unit square yielding $\Delta_d = a/2$.

An anisotropic spatial segmentation 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 strip is 2^{-s} (see Fig. 5). The equivalent curvature in each segment is now reduced and given by $a \cdot 2^{-2s}$. The optimal transform direction is chosen for each segment independently. Notice that the total number of the E-type directionlets is reduced, that is, it is given by the sum across all the segments and it is equal to $N_e(j, s) = O(a/2 \cdot 2^{n_2 j - s})$.

The directionlets, which do not intersect the discontinuity curve are called *S-type*. The number of S-type directionlets in a segment



(a) (b)

Fig. 6. (a) An image from the class C^2/C^2 is approximated using the standard 2-D WT and the S-AWT(Λ,3,2). (b) The decay of the MSE is faster in the case of the S-AWT(Λ,3,2).

is given by $N_s(j) = 2^{(n_1+n_2)j+s} - N_e(j)$. The total number of S-type directionlets is given by the sum across all the segments, that is, $N_s(j,s) = O(2^{(n_1+n_2)j+s} - a/2 \cdot 2^{n_2 j - s})$.

4.2. Approximation Power

The S-AWT applied on a segmented image with the optimal transform direction in each segment outperforms the standard 2-D WT in both approximation and compression rate of decay of the MSE. The following theorem gives the rate of decay for C^2/C^2 images. For reasons of lack of space, we give only the key ideas of the proof (see [14] for the full proof).

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) NLA by the S-AWT using spatial segmentation and N transform coefficients achieves

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

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

(b) Compression by the S-AWT using spatial segmentation and R bits for encoding achieves

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

To approximate the function $f(x_1,x_2)$, we keep all the coefficients with the magnitudes larger or equal to 2^{-m}, where $m \geq 0$, that is, (a) the E-type directionlets at the scales $0 \leq j \leq 2m/(n_1+n_2)$ and (b) the S-type directionlets at the scales $0 \leq j \leq 2m/n_3$, where $n_3 > n_1 + n_2$ [14]. Assuming that the number of segmentation levels is given by $s = \eta m$, where $\eta \geq 0$, it can be shown [14] that optimality is achieved for the anisotropy ratio $\rho^* = n_1/n_2 = \alpha$ and the segmentation rate $\eta^* = 0$. In that case, MSE$= O(N^{-\alpha})$.

For the compression application, it can be shown that the total number of encoding bits is given by $R(m) = O(2^{\alpha m/2})$ [14]. The MSE generated by quantization is given by $O(N \cdot 2^{-2m})$. Therefore, the total MSE is equal to $O(2^{-\alpha^2 m/2})$$= O(R^{-\alpha})$ [14].

Notice that the optimal anisotropy ratio is irrational and, thus, cannot be achieved. However, the S-AWT(Λ,3,2) approximates well the optimal transform, in which case the number of segmentation levels s is defined as $s = (\log_2 N)/51$. The achievable rate of decay of the MSE is $O(N^{-1.55})$. Although this rate is slower than the ones obtained in [2]-[6], 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-based algorithms still can be applied, making it easier to have very good compression algorithms. Fig. 6 illustrates the gain obtained by NLA using the S-AWT(Λ,3,2) applied on a C^2/C^2 image when compared to the results of NLA using the standard WT.

5. CONCLUSION

We 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 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 the 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 in the case of orthogonal 1-D filter-banks since Lagrangian optimization can be implemented in a straightforward manner.

6. REFERENCES

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

[2] E. L. Pennec and S. Mallat, "Sparse geometrical image approximation with bandelets," *IEEE Trans. Image Processing*, accepted.

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

[4] 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.

[5] 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.

[6] M. N. Do and M. Vetterli, "The contourlet transform: An effi cient directional multiresolution image representation," *IEEE Trans. Image Processing*, to appear.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

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

[13] 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.

[14] V. Velisavljević, B. Beferull-Lozano, M. Vetterli, and P. L. Dragotti, "Directionlets: Anisotropic multi-directional representation with separable fi ltering," *IEEE Trans. Image Processing*, submitted for publication.

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