Directional Multiscale Statistical Modeling of Images

Duncan D.-Y. Po and Minh N. Do
Coordinated Science Lab and Beckman Institute
University of Illinois at Urbana-Champaign
Urbana IL 61801
Email: duncanpo@uiuc.edu, minhdo@uiuc.edu

ABSTRACT

The contourlet transform is a new extension to the wavelet transform in two dimensions using nonseparable and directional filter banks. The contourlet expansion is composed of basis images oriented at varying directions in multiple scales, with flexible aspect ratios. With this rich set of basis images, the contourlet transform can effectively capture the smooth contours, which are the dominant features in natural images, with only a small number of coefficients. We begin with a detail study of the statistics of the contourlet coefficients of natural images, using histogram estimates of the marginal and joint distributions, and mutual information measurements to characterize the dependencies between coefficients. The study reveals the non-Gaussian marginal statistics and strong intra-subband, cross-scale, and cross-orientation dependencies of contourlet coefficients. It is also found that conditioned on the magnitudes of their generalized neighborhood coefficients, contourlet coefficients can approximately be modeled as Gaussian variables with variances directly related to the generalized neighborhood magnitudes. Based on these statistics, we model contourlet coefficients using a hidden Markov tree (HMT) model that can capture all of their inter-scale, inter-orientation, and intra-subband dependencies. We experiment this model in the image denoising and texture retrieval applications where the results are very promising. In denoising, contourlet HMT outperforms wavelet HMT and other classical methods in terms of both peak signal-to-noise ratio (PSNR) and visual quality. In texture retrieval, it shows improvements in performance over wavelet methods for various oriented textures.

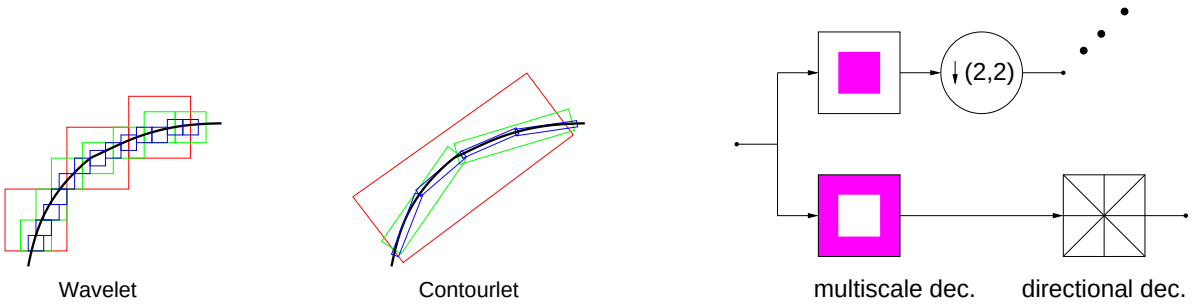
1. INTRODUCTION

In the last decade, *wavelets* have gained widespread popularity in signal processing.[1,2] Their multiresolution and time-frequency localization properties enable them to achieve sparse representations of one-dimensional signals characterized by point discontinuities. However, wavelets are defined in one-dimension. When dealing with images characterized by one-dimensional discontinuities, classical two-dimensional wavelets, which are separable products of one-dimensional wavelets, can no longer claim to be sparse.

Recently, nonseparable multiscale representations have attracted increasing attention. This type of image representation can capture the intrinsic geometrical structures such as smooth contours in natural images, which wavelets fail to capture. One such representation is the *contourlet*.[3–5] Developed by Do and Vetterli, contourlets are based on an efficient two-dimensional nonseparable filter bank that can deal effectively with images having smooth contours. Contourlets not only possess the main features of wavelets (namely, multiresolution and time-frequency localization,) but also show a high degree of directionality and anisotropy.

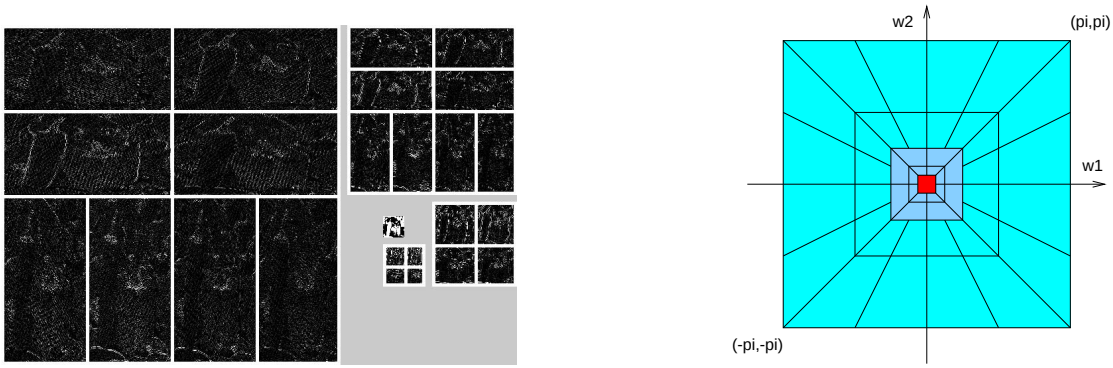
In this work, we develop and apply models for images in the contourlet domain. Similar to wavelet-based models,[6–11] contourlet-based models need to take into account the coefficients' dependencies across *scale* and *space*. However, as a "true" two-dimensional representation, contourlets allow us to also model the coefficients' dependencies across *directions*. In other words, contourlet modeling allows us to jointly model all three fundamental parameters of visual information, namely: *scale*, *space*, and *direction*.

The rest of the paper is organized as follows. Section 2 introduces the basics of contourlets including their transform algorithm, structure, properties, and coefficient relationships. In Section 3, we study the marginal and joint statistics of contourlet coefficients of natural images via histograms and mutual information. Inspired by these results, we develop a hidden Markov tree (HMT) model for the contourlet transform in Section 4. In Section 5, we apply the contourlet HMT model in denoising and texture retrieval. Finally, a conclusion is presented in Section 6.



(a) (b) (c)

Figure 1. (a) Wavelet bases have square supports and can only capture points. (b) Contourlet bases have elongated supports and can capture line segments. Contourlets thus can effectively represent a smooth contour with fewer coefficients. (c) PDFB structure.



(a) (b)

Figure 2. (a) An example of contourlet transform of the image "Peppers." Small coefficients are colored black while large coefficients are colored white. Larger rectangles correspond to finer subbands. (b) The corresponding frequency partition.

2. BACKGROUND

2.1. Contourlets

Contourlets are a sparse efficient overcomplete expansion for two-dimensional signals that are piecewise smooth away from smooth contours.[3–5] Such signals resemble natural images of natural objects and scenes, with the discontinuities as boundaries of objects. The discontinuities, referred to as *edges*, are gathered along one-dimensional smooth contours. Two-dimensional wavelets are only good at catching zero-dimensional or point discontinuities, resulting in largely inefficient decompositions. For example, as shown in Figure 1(a), it would take many wavelet coefficients to accurately represent even one simple one-dimensional curve.

Contourlets, meanwhile, have the multiresolution and time-frequency localization properties of wavelets, but also shows a very high degree of directionality and anisotropy. Precisely, contourlet transform involves basis functions that are oriented at any power of two's number of directions with flexible aspect ratios. With such richness in the choice of basis functions, contourlets can represent any one-dimensional smooth edges with close to optimal efficiency. For instance, Figure 1(b) shows that, compared with wavelets, contourlets can represent a smooth contour with much fewer coefficients. Contourlets are implemented by the pyramidal directional filter bank (PDFB)[3–5] that is a cascade of a Laplacian pyramid[12] and a directional filter bank,[13] as shown in Figure 1(c).

Figure 2(a) shows a typical set of contourlet coefficients for the image "Peppers" divided into four scales and eight directions at the finest scale. Figure 2(b) shows the corresponding frequency division of the contourlet

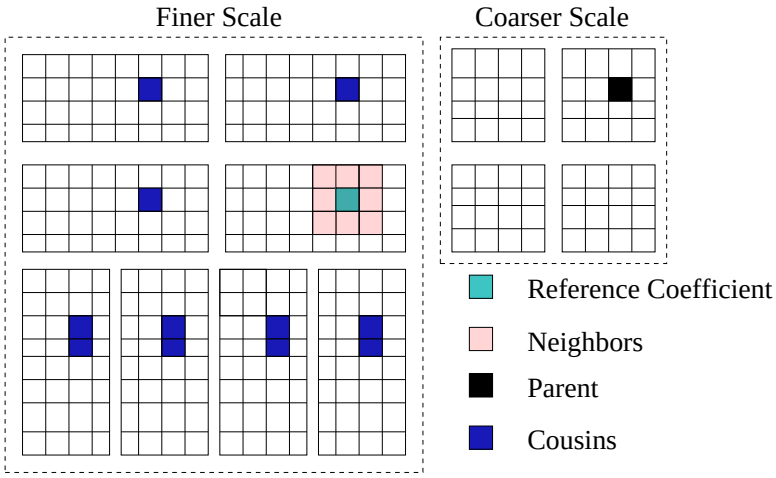


Figure 3. Contourlet coefficient relationships.

transform *.

2.2. Coefficient Relationships

We define some important contourlet coefficient relationships. Figure 3 graphically depicts some of these relationships. For each contourlet coefficient X, we define its eight adjacent coefficients in the same subband as its *neighbors* (NX). Next, the coefficient in the same spatial location in the immediately coarser scale corresponds to its *parent* (PX), and those in the same spatial location in the immediately finer scale are its child coefficients. Note that each coefficient has one parent and four children.

Apart from the cross-scale and intra-subband relationships defined above, a separate category is the cross-orientation relationship. Specifically, the coefficients in the same scale and same spatial location but in different directional subbands are referred to as *cousins* (CX) of each other. This cross-orientation relationship is more important in the contourlet case than in the wavelet case as there are more directions in the contourlet representation. While wavelet coefficients are always separated into only three directions, contourlet coefficients can have an arbitrary power of two's number of directions.

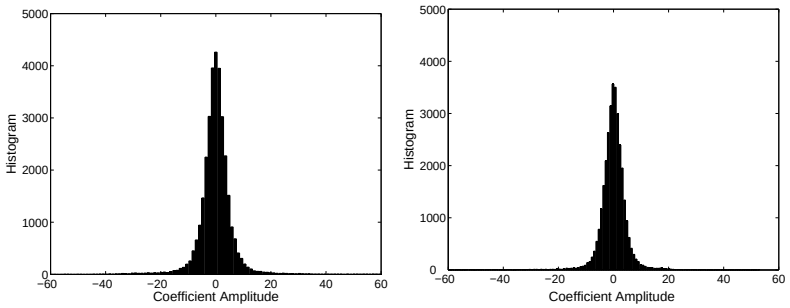
Combining the relationships across scales, space and orientations, we refer to the collective set of all parent (PX), neighbors (NX), and cousins (CX) of each coefficient X as its *generalized neighborhood* ($\widetilde{NX}$). These coefficient relationships play an important role in contourlet modeling, as will be seen in subsequent sections.

3. CONTOURLET STATISTICS

3.1. Marginal Statistics

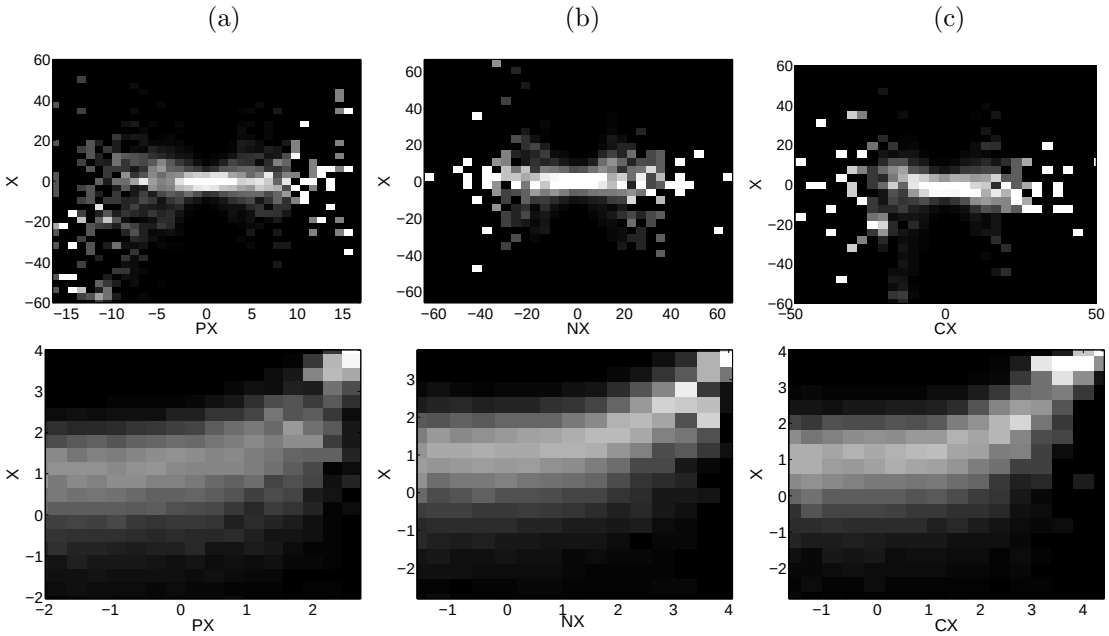
We first study the marginal statistics of the contourlet coefficients of natural images. Figure 4 plots the histograms of two finest subbands of the image "Peppers." The distributions are characterized by a very sharp peak at zero amplitude and extended tails to both sides of the peak. This implies that the contourlet transform is very sparse, as the majority of coefficients have amplitudes close to zero. The kurtoses of the two distributions are measured at 24.50 and 19.40, which are much higher than the Gaussian value of 3. It is observed that the histograms of all subbands of all natural images in our test set all yield similar distributions. Thus, the marginal distributions of natural images in contourlet domain are highly non-Gaussian.

*Recently, a modified version of the contourlet scheme that is critically sampled was developed.[14]



(a) (b)

Figure 4. Marginal statistics of two finest subbands of the image "Peppers." The kurtoses of the two distributions are measured at (a) 24.50 and (b) 19.40, showing that the coefficients are highly non-Gaussian.



(a) (b) (c)

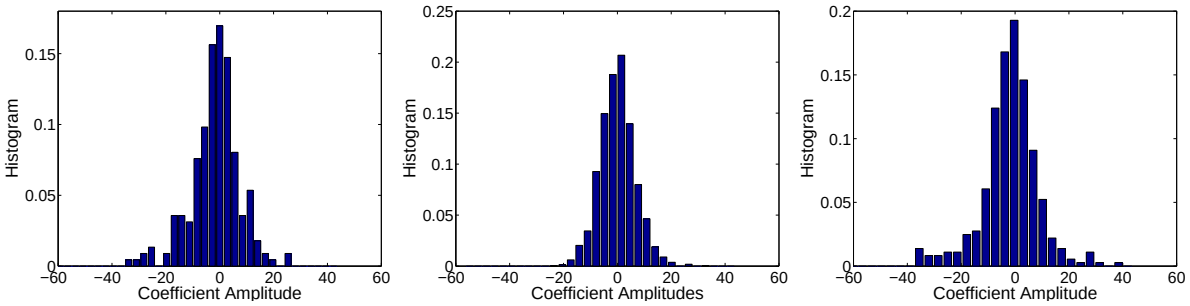
(d) (e) (f)

Figure 5. Conditional distribution of a finest subband of "Peppers," $P(X \mid .)$, conditioned on (a) parent, (b) left neighbor, (c) cousin. (d)–(f) corresponds to (a)–(c) in log-log scale.

3.2. Joint Statistics

We next study the joint statistics of contourlet images. Although the contourlet transform is a good decorrelator for images, a simple inspection of the contourlet coefficients of natural images can reveal that the coefficients are *dependent* on each other. For example, Figure 2 shows that large coefficients tend to cluster spatially around the edges of objects in the image (i.e., the peppers), and also persist to other scales and other orientations. Joint statistics are particularly important because in the wavelet case, image processing algorithms exploiting joint statistics of coefficients[7–11] show significant improvements in performance over those that exploit marginal statistics alone.[15,16] As the contourlet transform is similar to the wavelet transform, it is natural to extend this assumption to the contourlet case as well.

The top row of Figure 5 plots the conditional distributions of contourlet coefficients, $P(X \mid .)$, conditioned on their parents, left neighbors, and cousins for the image "Peppers," and the bottom row shows the same



(a) (b) (c)

Figure 6. Conditional distribution of the coefficients of a finest subband of "Peppers" on (a) their parents, (b) left neighbors, and (c) cousins. The kurtoses of the distributions are measured at 3.90, 2.90, and 2.99, respectively.

plots in log-log scale. All three conditional distributions resemble a "bow-tie" shape, while in log-log scale, they all display a horizontal distribution on the left and a unit slope distribution on the right.[6] The "bow-tie" shaped conditional distribution indicates two properties of the coefficients. First, even though the coefficients are correlated due to the overcomplete nature of the contourlet transform, they are *approximately* decorrelated with the conditional expectation $E[X \mid . \,] \approx 0$. Second, the distribution of the coefficients varies with the conditioned magnitude and therefore the coefficients are dependent on one another. From the log-log scale plots, the unit slope on the right indicates a linear relationship between the magnitudes of the coefficients and their generalized neighborhoods. Ignoring the left sides of the plots, which are caused mainly by quantization errors and other uncertainties,[9] we can conclude that contourlet coefficients of natural images are *uncorrelated* yet *linearly dependent* in magnitude on their generalized neighborhoods.

It is also of interest to investigate the vertical cross-sections of the top row of Figure 5, which correspond to the conditional distributions of contourlet coefficients, conditioned on particular magnitudes of their generalized neighborhoods. It is found that the cross-sections generally have the shape as shown in Figure 6. The kurtoses of the conditional distributions are measured at 3.90, 2.90, and 2.99, conditioned on the coefficients' parents, left neighbors, and cousins, respectively. The kurtoses are very close to the Gaussian kurtosis of 3, indicating that the conditional distributions of contourlet coefficients can be well-modeled by Gaussian distributions. Compared to the unconditional distributions in Figure 4, for which the kurtoses are highly non-Gaussian at around 20, it can be concluded that contourlet coefficients are *non-Gaussian but conditionally Gaussian*. Since in Figure 5, the vertical cross-section of each column can be approximated as a Gaussian distribution, contourlet coefficients can be accurately modeled using a mixture of Gaussian distributions of different variances.

3.3. Mutual Information

We propose a quantitative study on the joint statistics of contourlet coefficients, in complement to the qualitative study previously discussed. In particular, we measure the mutual information between coefficients as a quantitative measurement of dependencies.[17,18] Mutual information, defined as[19]

$$I(X;Y) = \int\int p(x,y) \log \frac{p(x,y)}{p(x)p(y)} dxdy \tag{1}$$

between random variables X and Y, increases with increasing dependence between the two variables. Because mutual information is also easy to estimate via histograms, it makes a convenient measurement of joint statistics. Specifically, between any two contourlet coefficients X and Y, we use the estimator[20]

$$\hat{I}(X;Y) = \sum_{i,j} \frac{k_{ij}}{N} \log \frac{k_{ij}N}{k_i k_j} - \frac{(J-1)(K-1)}{2N} \tag{2}$$

where k_{ij} is the number of coefficients in a subband observed in cell (i,j), $k_i = \sum_j k_{ij}$ and $k_j = \sum_i k_{ij}$ are the marginal distribution histogram estimates, N is the total number of coefficients in the subband considered, and

Table 1. Mutual information estimate. PX refers to the parent of X; NX is a linear combination of the set of eight spatial neighbors within the same subband; CX is a linear combination of seven cousins in different directional subbands.

	Lena	Barbara	Peppers
$I(X;PX)$	0.11	0.14	0.10
$I(X;NX)$	0.23	0.58	0.17
$I(X;CX)$	0.19	0.39	0.14
$I(X;PX,NX)$	0.24	0.58	0.17
$I(X;NX,CX)$	0.26	0.59	0.20
$I(X;PX,CX)$	0.21	0.40	0.16
$I(X;PX,NX,CX)$	0.26	0.59	0.20

J and K are the number of histogram cells along the X and Y directions respectively. Note that the first term is the mutual information histogram estimate and the second term is a partial bias correction term. It can be shown that even after the bias is partially removed, the residual bias still causes the estimator to underestimate the mutual information and the estimate can only serve as a lowerbound.[20]

Table 1 shows the mutual information estimation results for three representative images "Lena," "Barbara," and "Peppers." The results are obtained using the 9-7 pyramidal filters (9-7 filters) and the McClellan transformed directional filters, proposed by Cohen and Daubechies (CD filter)[21] as the contourlet filters. Comparing the first three rows, it can be observed that the eight neighbors contain the most information about the coefficients, followed by the set of cousins, which is followed by the parent coefficient. This pattern is consistent over all images studied, and is particularly evident for highly textured images such as "Barbara." A second observation is that the joint mutual information of parent and neighbors $I(X;PX,NX)$ is approximately equal to $I(X;NX)$, while the joint mutual information of neighbors and cousins $I(X;NX,CX)$ is also only marginally higher than $I(X;NX)$ (approximately 10% to 20% increase). The gains are even smaller for images dominated by textures such as "Barbara." This indicates that the eight neighbor coefficients contain most of the available information about the coefficients, and neither parent nor cousin coefficients offer any significant additional information over the neighbors.

Mutual information estimates, however, are highly dependent on the choice of filters and the directional partition scheme used. Table 2 compares the mutual information estimates for "Lena" using different combinations of the 9-7 or Haar pyramidal filters and two different directional filters: the CD filter or the ladder filter by Phoong et al. (PKVA filter).[22] It can be observed that using the Haar pyramidal filter yields much higher mutual information than using the 9-7 filters. This suggests that the 9-7 filters are superior to the Haar filter in terms of whitening the contourlet coefficients. Similarly, comparing directional filters, the PKVA filter yields lower inter-orientation mutual information than the CD filter and therefore is more effective in whitening the coefficients. Table 2 also displays the mutual information estimates of wavelet coefficients for the same image, using different wavelet filters. Compared to contourlets, wavelet coefficients exhibit similar inter-scale and intra-subband dependencies but much lower inter-orientation dependencies. Thus for contourlets, cousin coefficients carry more significant information than in the case of wavelets.

Table 3 compares the inter-orientation and intra-subband mutual information for different directional partitioning schemes to partition the finest scale into 4, 8, and 16 orientations. Each further partition increases the inter-orientation dependency and decreases the intra-subband dependency. This observation is important in that the relative importance of inter-orientation and intra-subband dependencies can be adjusted by controlling the partitioning of each scale of the contourlet transform.

4. HIDDEN MARKOV TREE MODEL

We use the two-state hidden Markov tree (HMT) model[18, 23] to model the contourlet coefficients. The HMT models each coefficient with a mixture of two Gaussian random distributions. Each coefficient is associated

Table 2. Mutual information estimate for contourlet and wavelet representation of "Lena" using different filters.

Contourlets			
p filter; d filter	$I(X;PX)$	$I(X;NX)$	$I(X;CX)$
Haar,CD	0.18	0.33	0.32
9-7,CD	0.11	0.23	0.19
9-7,PKVA	0.11	0.24	0.15
Wavelets			
filter	$I(X;PX)$	$I(X;NX)$	$I(X;CX)$
Haar	0.20	0.27	0.14
Daubechies 4-taps	0.14	0.23	0.08
Daubechies 8-taps	0.11	0.20	0.05

Table 3. Mutual information estimate for different partitions of the finest subbands. Data are obtained using 9-7 pyramidal filters and CD directional filter on the "Lena" image.

	4 directions	8 directions	16 directions
$I(X;NX)$	0.26	0.23	0.20
$I(X;CX)$	0.14	0.19	0.19

with a hidden state variable, randomly distributed over its two states. Conditioned on its state, it is normally distributed with the corresponding variance of the state.

The HMT models the inter-coefficient dependencies by establishing links between the *hidden states* of parent and child coefficients in a quad-tree structure. Associated with each link is a state transition probability matrix $\mathbf{A}_{m,n}$

$$\mathbf{A}_{m,n} = \begin{pmatrix} a_{m,n}^{1,1} & a_{m,n}^{1,2} \\ a_{m,n}^{2,1} & a_{m,n}^{2,2} \end{pmatrix} \tag{3}$$

where $a_{m,n}^{k,l}$ is the probability that the child coefficient m is in state k given the parent coefficient n is in state l. The distributions of the coefficients in the root subbands are also specified in the model.

The simple quad-tree structure of HMT enables efficient training using the expectation maximization (EM) algorithm.[23] The quad-tree structure also allows for direct modeling of inter-scale dependencies, and indirect modeling of the dependencies between neighboring coefficients within the same subband, via their links to their common ancestors.

One major difference between the wavelet HMT and the contourlet HMT models is that the contourlet HMT model also accounts for the inter-orientation dependencies while the wavelet HMT does not. Figure 7 illustrates this difference. In wavelet decompositions, the four children of any parent coefficient are always in the same directional subband. As a result, each tree within the model is entirely in one direction. As the HMT model models each tree independently, dependencies across directions are not modeled. Conversely, contourlet coefficients can have their four children in two separate directional subbands. Thus, contourlet HMTs can span several adjacent directions in the finer scales and therefore, inter-orientation dependencies are modeled. In other words, the contourlet HMT model effectively captures the dependencies across all of scale, space and orientation. The software for the implemention of this model can be downloaded from http://www.ifp.uiuc.edu/˜minhdo.

5. APPLICATIONS

5.1. Denoising

We apply the contourlet HMT model in Bayesian denoising of zero-mean additive white Gaussian noise. We first obtain the contourlet HMT model of the clean image, $\boldsymbol{\theta_u}$, by subtracting the noise power from all variances in

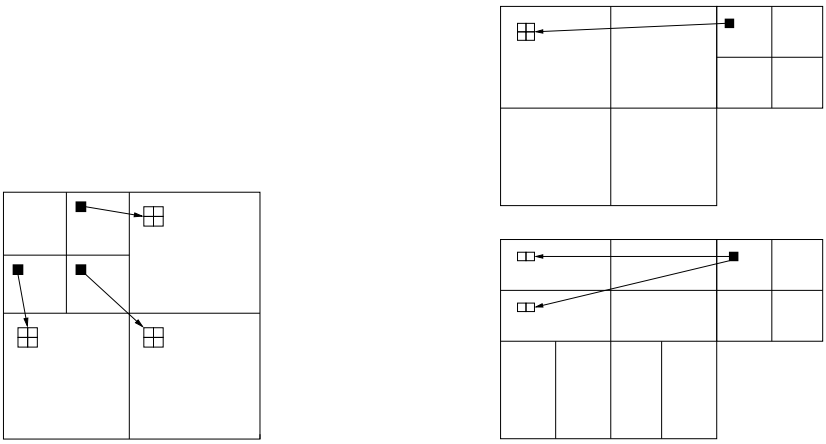


Figure 7. Parent-children relationship for (a) wavelets and (b) two possible contourlets structures. Black squares represent parent coefficients with white squares as their children. Notice for the bottom contourlet structure, a parent coefficient can have its children spread over two subbands.

Table 4. Peak signal-to-noise ratio (in decibels) of denoised images.

	noisy	wiener2 (5×5)	wavelet thresholding	wavelet HMT	contourlet HMT
Lena	14.63	24.75	24.20	**25.89**	25.73
Barbara	14.48	22.57	21.96	23.71	**23.88**
Peppers	14.74	24.60	23.82	25.28	**25.37**
Goldhill	14.52	24.32	23.51	25.18	**25.19**
Boat	14.55	23.89	23.08	**24.90**	24.83
Baboon	14.36	21.03	19.71	21.48	**21.57**
Couple	14.32	23.77	22.69	**24.19**	23.61
Zelda	14.61	25.78	26.05	27.63	**27.71**
Bridge	14.61	22.14	20.73	22.44	**22.50**
Truck	14.26	24.79	24.47	**26.49**	26.35

the noisy model, $\boldsymbol{\theta_v}$. With the model, we can perform Bayesian denoising[18, 23, 24]

$$E[\vec{u}_{i,j,k} \mid \vec{v}_{i,j,k}, \boldsymbol{\theta_u}] = \sum_m p(S_{i,j,k} = m \mid \vec{v}_{i,j,k}, \boldsymbol{\theta_u}) \times (\sigma^2_{i,j,m} I) \times (\sigma^2_{i,j,m} I + \Sigma_n)^{-1} \times \vec{v}_{i,j,k} \tag{4}$$

where $\vec{u}_{i,j,k}$ and $\vec{v}_{i,j,k}$ are respectively the set of clean and noisy coefficients that fall within an arbitrarily-sized local neighborhood window at the k^{th} coefficient in the j^{th} subband in the i^{th} scale, $p(S_{i,j,k} = m \mid \vec{v}_{i,j,k}, \boldsymbol{\theta_u})$ is the state likelihood, $\sigma^2_{i,j,m}$ is the model variance, and Σ_n is the noise covariance matrix. I denotes the identity matrix. Note that a local window of coefficients, rather than a single coefficient, is used in the equation because the overcomplete property of contourlets causes the white noise to become correlated in the contourlet domain.[24]

We applied the denoising algorithm described above to noisy test images and the results are shown in Table 4. Also in the table are the denoising results of the same noisy images using wavelet HMT and two other classical denoising methods in Matlab's "wiener2" and wavelet thresholding (with threshold = $3\sigma_n$). In terms of peak signal-to-noise ratio (PSNR), contourlet HMT gives the highest PSNR of the four methods compared. It produces better PSNR than wavelet HMT for most of the images, while easily outperforming "wiener2" and wavelet thresholding for all images. Most importantly, contourlet HMT produces superior denoised results in

Figure 8. Denoising results of "Zelda"image: (a) "Zelda" image, (b) noisy image ($\sigma_n = 51$, PSNR = 14.607), (c) wiener2 (PSNR = 25.780), (d) wavelet thresholding (PSNR = 26.046), (e) wavelet HMT (PSNR = 27.631), and (f) contourlet HMT (PSNR = 27.707).

terms of visual quality. Figure 8 shows the denoising results of the "Zelda" image. It is clear that contourlet HMT removes most of the noise and produces the smoothest image.

5.2. Texture Retrieval

The contourlet HMT model is then applied to content-based texture retrieval. We first constructed a texture image database using 64 textures from the Brodatz database.[25] Each of these textures was partitioned into 16 sub-images of size 128×128, and only 8 of the 16 were retained. Thus the database contained 512 texture images in total with 8 images from each of the 64 texture classes. For all texture images in the database, we applied the contourlet HMT model and extracted the model parameters as features. We then used each image in the database as the query image and measured, via the Monte-Carlo method, the Kullback-Liebler distances between the query image and each database image.[18] For each query, the 15 database images that gave the smallest KLD were retrieved.

We compare our results with wavelet HMT texture retrieval[5] and the resulting average retrieval rates are shown in Table 5. Note that contourlet HMT gives an average retrieval rate approximately 2.5% higher than that of wavelet HMT. Comparing the retrieval rates for individual textures, the top part of Figure 9 shows the texture images that are better retrieved by wavelets than by contourlets by at least 5%, and the bottom part shows those better retrieved by contourlets by at least 5%. From the figure, it can be seen that the textures better retrieved by wavelets are all characterized by dominant vertical or horizontal directional components. In contrast, the textures better retrieved by contourlets generally demonstrate more diverse directional components (such as circular or irregular shapes). This shows the superior quality of contourlets in capturing directional

Table 5. Average retrieval rates. All figures are percentages.

wavelet HMT	contourlet HMT
90.87	**93.29**

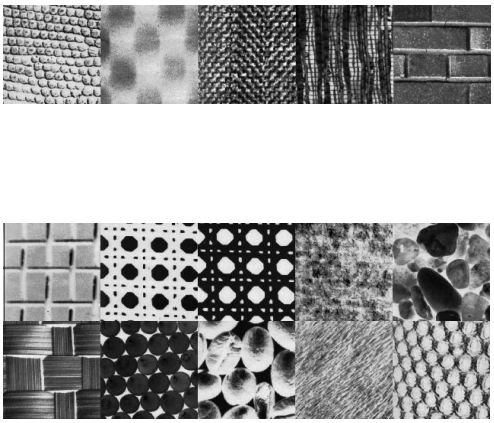


Figure 9. Texture retrieval results. The top row shows the textures that are better retrieved by wavelets than by contourlets by at least 5%. The bottom two rows show the textures that are better retrieved by contourlets than by wavelets by at least 5%.

information. For most textures, contourlet HMT texture retrieval gives satisfactory texture retrieval performance with retrieval rates typically above 80%.

6. CONCLUSION

We have studied the properties of the contourlet coefficients of natural images. It is found that similar to wavelets, contourlet coefficients are highly non-Gaussian but conditionally Gaussian conditioned on their generalized neighborhoods. Properties of persistence and clustering are apparent as the coefficients are highly dependent on their parents and neighbors, as well as cousins in different orientation subbands. Such dependencies can be quantitatively verified using mutual information, which shows contourlet coefficients exhibits highest level of dependence on their neighbors, followed by cousins, which are followed by parents.

Based on the above properties, we have developed a contourlet hidden Markov tree model (HMT) which models the contourlet coefficients using a mixture of Gaussian distributions. The model also exploits all of cross-scale, cross-orientation, and intra-subband dependencies. We applied this model on denoising and texture retrieval. The results are highly promising. In denoising, contourlet HMT produces superior peak signal-to-noise ratio (PSNR) and visual quality in the denoised images. In texture retrieval, contourlet HMT gives higher retrieval rates than wavelets for textures that show high directionality. Both results suggest that contourlets can capture directional information very well, which is a highly valuable property in image processing.

REFERENCES

1. M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*. Upper Saddle River, NJ: Prentice Hall, 1995.
2. S. Mallat, *A Wavelet Tour of Signal Processing*. San Diego, CA: Academic Press, 1999.
3. M. N. Do and M. Vetterli, "Contourlets," in *Beyond Wavelets*, J. Stoeckler and G. V. Welland, Eds. San Diego, CA: Academic Press, 2003. To Appear.
4. M. N. Do and M. Vetterli, "Contourlet: A computational framework for directional multiresolution image representation," submitted to *IEEE Transactions on Image Processing*, 2003.

5. M. N. Do, "Directional multiresolution image representations," Ph.D. dissertation, Swiss Federal Institute of Technology, Lausanne, 2001.
6. E. P. Simoncelli, "Modeling the joint statistics of images in the wavelet domain," in *Proceedings of the SPIE* 44^{th} *Annual Meeting*, 1999, pp. 188–195.
7. M. K. Mihçak, I. Kozintsev, and K. Ramchandran, "Low-complexity image denoising based on statistical modeling of wavelet coefficients," *IEEE Signal Processing Letters*, vol. 6, no. 12, pp. 300–303, December 1999.
8. J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, pp. 3445–3462, December 1993.
9. R. W. Buccigrossi and E. P. Simoncelli, "Image compression via joint statistical characterization in the wavelet domain," *IEEE Transactions on Image Processing*, vol. 8, pp. 1688–1701, December 1999.
10. S. M. Lopresto, K. Ramchandran, and M. T. Orchard, "Image coding based on mixture modeling of wavelet coefficients and a fast estimation quantization framework," in *Proceedings of IEEE Data Compression Conference*, 1997, pp. 221–230.
11. S. G. Chang, B. Yu, and M. Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 1522–1531, September 2000.
12. P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Transactions on Communications*, vol. 31, no. 4, pp. 532–540, April 1983.
13. R. H. Bamberger and M. J. T. Smith, "A filter bank for the directional decomposition of images: theory and design," *IEEE Transactions on Signal Processing*, vol. 40, pp. 882–893, April 1992.
14. Y. Lu and M. N. Do, "Crisp-contourlet: A critically sampled directional multiresolution image representation," in *SPIE Wavelet X Conference*, August 2003.
15. S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, pp. 674–693, July 1989.
16. E. P. Simoncelli and E. H. Adelson, "Noise removal via Bayesian wavelet coring," in *Proceedings of the Third International Conference on Image Processing*, 1996, pp. I. 379–382.
17. J. Liu and P. Moulin, "Information-theoretic analysis of interscale and intrascale dependencies between image wavelet coefficients," *IEEE Transactions on Image Processing*, vol. 10, no. 10, pp. 1647–1658, November 2001.
18. D. D.-Y. Po and M. N. Do, "Directional Multiscale Modeling of Images using the Contourlet Transform," submitted to *IEEE Transactions on Image Processing*, 2003.
19. T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY: John Wiley and Sons, Inc., 1991.
20. R. Moddemeijer, "On estimation of entropy and mutual information of continuous distributions," *Signal Processing*, vol. 16, no. 3, pp. 233–246, March 1989.
21. A. Cohen and I. Daubechies, "Nonseparable bidimensional wavelet bases," *Revista Matematica Iberoamericana*, vol. 9, no. 1, pp. 51–137, December 1993.
22. S.-M. Phoong, C. W. Kim, P. P. Vaidyanathan, and R. Ansari, "A new class of two-channel biorthogonal filter banks and wavelet bases," *IEEE Transactions on Signal Processing*, vol. 43, no. 3, pp. 649–665, March 1995.
23. M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models," *IEEE Transactions on Information Theory*, vol. 46, pp. 886–902, April 1998.
24. V. Melnik, I. Shmulevich, K. Egiazarian, and J. Astola, "Image denoising based on locally-adaptive filtering in the median pyramidal transform domain," in *IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing*, 2001.
25. "Brodatz database," 2002, http://www.ux.his.no/~tranden/index.html .