

Image Texture Characterization Using the Discrete Orthonormal S-Transform

Sylvia Drabycz,^{1,4} Robert G. Stockwell,² and J. Ross Mitchell^{3,4}

We present a new efficient approach for characterizing image texture based on a recently published discrete, orthonormal space-frequency transform known as the DOST. We develop a frequency-domain implementation of the DOST in two dimensions for the case of dyadic frequency sampling. Then, we describe a rapid and efficient approach to obtain local spatial frequency information for an image and show that this information can be used to characterize the horizontal and vertical frequency patterns in synthetic images. Finally, we demonstrate that DOST components can be combined to obtain a rotationally invariant set of texture features that can accurately classify a series of texture patterns. The DOST provides the computational efficiency and multi-scale information of wavelet transforms, while providing texture features in terms of Fourier frequencies. It outperforms leading wavelet-based texture analysis methods.

KEY WORDS: 3D texture mapping, 3D wavelet transform, algorithms, biomedical image analysis, brain imaging, computer assisted detection, computer-aided diagnosis (CAD), Fourier analysis, image analysis, image processing, magnetic resonance imaging, MR imaging, pattern recognition, automated, signal processing

BACKGROUND

Texture characterization is an important problem in medical image analysis. Image texture can be defined as the spatial relationship of pixel values in an image region¹. In medical images, texture can be thought of as the local characteristic pattern of image intensity that identifies a tissue. Texture also determines local spectral or frequency content in an image; changes in local texture should cause changes in the local spatial frequency. Texture analysis is of interest in medical imaging because, as biological tissues become abnormal during a disease process, their underlying texture may also change.

Lerski et al.² and Castellano et al.³ provide good reviews of the application of texture analysis methods to medical images.

Various mathematical techniques to quantify image texture, including statistical, Fourier, and wavelet-based methods, have been applied to radiological images of numerous pathologies, such as multiple sclerosis^{4,5}, brain tumors⁶, liver diseases⁷, infarcted myocardial tissue⁸, and normal tissues of the knee⁹, and has even been used for automated detection and classification based on phase-contrast microscopy images, such as those used in cervical cancer diagnosis¹⁰.

A texture feature is a value that quantifies some characteristic of local intensity variation within an image¹. A variety of approaches exist to quantify texture. A common technique used in medical imaging is based on co-occurrence matrices¹¹. Statistical

¹From the Department of Electrical and Computer Engineering, University of Calgary, 2500 University Dr NW, Calgary, Alberta, T2N 1N4, Canada.

²From the Northwest Research Associates, Colorado Research Associates Division, 3380 Mitchell Lane, Boulder, CO, 80301, USA.

³From the Departments of Clinical Neurosciences and Radiology, University of Calgary, 2500 University Dr NW, Calgary, AB, T2N 1N4, Canada.

⁴From the Southern Alberta Cancer Research Institute, 2AA10-3280 Hospital Drive NW, Calgary, Alberta, T2N 4N1, Canada.

Correspondence to: J. Ross Mitchell, Southern Alberta Cancer Research Institute, 2AA10-3280 Hospital Drive NW, Calgary, Alberta, T2N 4N1, Canada; tel: +1-403-2106667; fax: +1-403-2108135; e-mail: rmitch@imaginginformatics.ca

Copyright © 2008 by Society for Imaging Informatics in Medicine

Online publication 2 August 2008

doi: 10.1007/s10278-008-9138-8

measures of texture are calculated based on the frequency of specific gray levels occurring between pairs of points within an image. The co-occurrence technique has been used in many studies, such as to classify benign and malignant solitary pulmonary nodules on computed tomography (CT) images¹² and to quantify pathological changes during treatment of multiple sclerosis¹³. However, the technique is limited to very small neighborhoods due to its computation complexity. Therefore, only the highest frequency textures can be analyzed. Broad, large-scale changes are difficult to detect using co-occurrence statistics. Furthermore, the resulting statistical measures are difficult to interpret and compare across patients.

A more recent method of texture analysis relies on discrete wavelet transform (DWT) (see¹⁴ for a review). Wavelets provide a multi-scale representation of an image, allowing analysis of varying degrees of detail within an image. Efficient algorithms and a solid mathematical framework make wavelets appealing for numerous applications, including texture analysis. For example, wavelet-based texture analysis has been used for automatic diagnosis and grading of breast tumor histology images¹⁵.

The S-transform¹⁶ (ST) is closely related to the continuous wavelet transform using a complex Morlet mother wavelet¹⁷ and directly measures the local spatial frequency content of each pixel in an image. The ST has been successfully used to analyze signals in numerous applications, such as seismic recordings¹⁸, ground vibrations¹⁹, hydrology²⁰, gravitational waves²¹, and power system analysis²². The one-dimensional ST has shown to be a useful tool for analysis of medical signals, such as EEG²³, functional magnetic resonance imaging²⁴, and laser Doppler flowmetry²⁵. The ST is particularly well suited to texture analysis of medical images due to its optimal space-frequency resolution and close ties to the Fourier transform (FT)—the basis of medical image reconstruction. The ST uses complex Fourier basis functions, modulated by frequency-dependent Gaussian windows. The ST preserves the phase information, uses a linear frequency scale, and can be easily inverted to recover the Fourier domain of an image.

The redundant nature of the ST algorithm has been the main obstacle in wider application of ST-based texture analysis for 2D images. Extensive processing time and a large amount of memory are required to calculate and store the texture descriptions of large

medical images. The 2D-ST of an array of size $N \times N$ has computational complexity of $O[N^4 + N^4 \log(N)]$ and storage requirements of $O[N^4]$. As a result, the ST of a typical 256×256 MR image takes approximately 1.5 h to calculate on a single computer and requires 32 GB of memory²⁶. While other researchers have developed techniques to distribute these calculations over networks of machines²⁶, the results are still difficult to manage and interpret²⁷. Therefore, previous work on 2D images has been limited to analysis of small regions of interest (ROIs) and typically collapsed to 1D spectra^{28–30}. However, small ROIs reduce the resolution of the frequency spectra, and therefore, the sensitivity to subtle texture changes. These requirements make applying the 2D-ST to clinical medical applications difficult and impractical. Clinical texture analysis requires a rapid, efficient algorithm that provides complete information about all frequency components.

Despite these limitations, the 2D-ST has shown promising results in identifying differences in texture that correlate with neurological pathology. For example, previous work has shown that ST-based texture measures can help detect sub-clinical abnormalities in normal-appearing white matter of multiple sclerosis patients³¹. The 2D-ST has also been used to identify oligodendroglial brain tumors with genetic abnormalities that make the tumors more responsive to chemotherapy treatment²⁸.

Recent work shows that a discrete, orthonormal basis can be used to accelerate calculation of the ST and eliminates the redundancy in the space-frequency domain³². The discrete, orthonormal ST (DOST) provides a spatial frequency representation similar to the DWT. However, the DOST has the additional benefits of maintaining the phase properties of the ST (and FT), retaining the ability to collapse exactly back to the Fourier domain. Furthermore, the DOST framework allows for an arbitrary partitioning of the frequency domain; this allows for a dyadic sampling similar to the discrete wavelet transform for zero redundancy, or oversampling to any extent, all the way back to the fully redundant ST.

We present a frequency-domain implementation of the 2D-DOST by partitioning the frequency domain in a dyadic sampling scheme. We show that the DOST can provide a pixel-by-pixel texture description of an image by creating local spectra containing the horizontal and vertical frequency

information from the FT of the image. The DOST is straightforward and fast to calculate, allowing us to analyze every pixel in a large image within seconds. Our goal in this paper is to introduce the multi-dimensional DOST, describe its application to texture classification, and characterize its performance on real and synthetic images.

THEORY

2D Discrete Orthonormal S-Transform

The process of calculating the DOST of a 1D signal in the time domain is described in³². That process involves calculating the basis functions, which are derived by taking linear combinations of the Fourier complex sinusoids in band-limited subspaces and applying appropriate phase and frequency shifts. In this paper, we describe the process of calculating the DOST of a 2D image in the frequency domain using a dyadic sampling scheme.

We begin by defining the forward 2D-FT of a discrete function $f[x,y]$, which is assumed to have a sampling interval of one in the x - and y -directions:

$$H[m, n] = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} h[x, y] e^{-2\pi i \left(\frac{mx}{M} + \frac{ny}{N} \right)} \quad (1)$$

and the inverse 2D-FT:

$$h[x, y] = \frac{1}{MN} \sum_{m=-M/2}^{M/2-1} \sum_{n=-N/2}^{N/2-1} H[m, n] e^{2\pi i \left(\frac{mx}{M} + \frac{ny}{N} \right)} \quad (2)$$

The 2D-DOST of a $N \times N$ image $h[x,y]$ is calculated by partitioning the 2D-FT of the image, $H[m,n]$, multiplying by the square root of the number of points in the partition, and performing an inverse 2D-FT,

$$S[x', y', \nu_x, \nu_y] = \frac{1}{\sqrt{2^{p_x+p_y-2}}} \sum_{m=-2^{p_x-2}}^{2^{p_x-2}-1} \sum_{n=-2^{p_y-2}}^{2^{p_y-2}-1} H[m + \nu_x, n + \nu_y] e^{2\pi i \left(\frac{mx'}{2^{p_x-1}} + \frac{ny'}{2^{p_y-1}} \right)} \quad (3)$$

where we define $\nu_x = 2^{p_x-1} + 2^{p_x-2}$ and $\nu_y = 2^{p_y-1} + 2^{p_y-2}$ as the horizontal and vertical “voice frequencies”³². One should note that the operation

to create the voice image ($S[x,y,\nu_{x0},\nu_{y0}]$) consists of an inverse fast Fourier transform (FFT) of smaller dimension (not $N \times N$). The spectrum is partitioned such that the wavenumbers (ν_{x0}, ν_{y0}) are shifted to the zero wavenumber point, and a $2^{p_x-1} \times 2^{p_y-1}$ inverse FFT is performed, resulting in a rectangular (in general) voice image of $2^{p_x-1} \times 2^{p_y-1}$ points (see Fig. 1). The total number of points in the DOST result and in the original image are the same.

The inverse transform for the 2D-DOST is similar to the 1D case in³². Since the DOST is an energy-conserving transform³², we can apply the forward 2D-FFT to each “voice” in order to reverse the spectral partitioning and reconstruct the spectrum of the image:

$$H[m, n] = \sqrt{2^{p_x+p_y-2}} \sum_{m'=-2^{p_x-2}}^{2^{p_x-2}-1} \sum_{n'=-2^{p_y-2}}^{2^{p_y-2}-1} S[m - \nu_x, n - \nu_y] e^{-2\pi i \left(\frac{m'}{2^{p_x-1}} + \frac{n'}{2^{p_y-1}} \right)} \quad (4)$$

The image can then be recovered by performing an inverse FT:

$$h[x, y] = \frac{1}{N^2} \sum_{m=-N/2}^{N/2-1} \sum_{n=-N/2}^{N/2-1} H[m, n] e^{2\pi i (mx+ny)/N} \quad (5)$$

Source code for the forward 2D-DOST is given in [Appendix](#). Note that, in the source code, the division by $\sqrt{2^{p_x+p_y-2}}$ is replaced with a multiplication by the same factor; this is to account for the fact that the inverse FFT calculation inherently divides the total number of samples in the partition ($2^{p_x-1} \times 2^{p_y-1}$), resulting in the overall scaling by $(2^{p_x+p_y-2})^{-1/2}$.

The DOST in this formulation, like the DWT, uses a dyadic sampling scheme (orders=0, 1, 2, ..., $\log_2 N - 1$; Fig. 1). However, the two transforms provide different information about the frequency content of the image. The DWT gives horizontal, vertical, and diagonal “detail” coefficients for each order, while the DOST provides information about the voice frequencies (ν_x, ν_y) that contain a bandwidth of $2^{p_x-1} \times 2^{p_y-1}$ frequencies. While the wavelet decomposes the image into “scales” of size $M \times M$, the DOST uses the minimum number of points required to describe the amplitude of a Fourier frequency component in

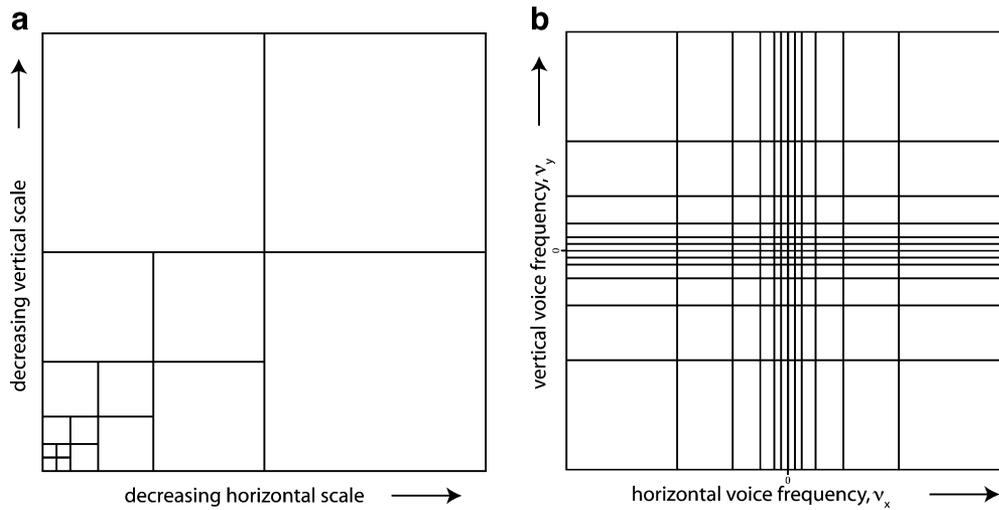


Fig. 1. Partitioning of a the DWT and b the DOST for six orders. The squares indicate the sub-images for each order. Both transforms use a dyadic sampling scheme but provide different information about the frequency content of the image. The DWT gives horizontal, vertical and diagonal “detail” coefficients for each order, while the DOST provides information about the voice frequencies (ν_x, ν_y) that contain a bandwidth of $2^{b_x-1} \times 2^{b_y-1}$ frequencies.

each of the horizontal and vertical directions. For example, a Fourier frequency component with two cycles spanning the horizontal direction and four cycles spanning the vertical direction is represented in the DOST by 4×8 points. As a result, the DOST provides more texture features for a $N \times N$ image than the DWT, while maintaining an overall size of N^2 . For example, Figure 2 shows the DOST and DWT transforms of an image containing four textures from the Brodatz texture library³³. Each texture section is size 128×128 pixels and has been normalized such that the average (DC) is zero. The DWT was calculated using the Daubechies tap-4 (db4) wavelet at five levels of decomposition. The

difference in partitioning can be seen in this example.

Pixel-Wise Local Spatial Frequency Description

Once we have the DOST description of an image, we can find the contribution of each horizontal and vertical voice to any pixel within the image. By simply choosing a set of (x, y) coordinates representing a single pixel or image region, we can determine the value of $S[x', y', \nu_x, \nu_y]$ for all (ν_x, ν_y) . Since the voice images are of varying shape and size, the value of the DOST for frequency order

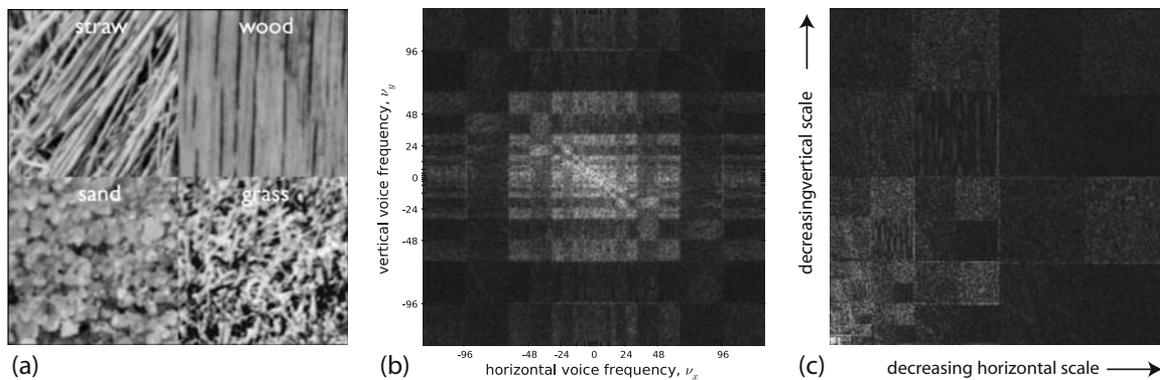


Fig. 2. a A mosaic of four texture images: straw, wood, sand and grass. b The 2D-DOST of the mosaic. c The DWT of the image using the db4 wavelet at five levels of decomposition. Note that the square root of the amplitude of both the DOST and DWT coefficients is shown for display purposes.

(p_x, p_y) at image location (x, y) can be found at $S[x/N \times 2^{p_x-1}, y/N \times 2^{p_y-1}, v_x, v_y]$. By iterating over all values of (p_x, p_y) , we can build up a local spatial frequency domain of size $2 \log_2 N \times 2 \log_2 N$ for each pixel or averaged region within an image; this domain contains the positive and negative frequency components from the DC, $(v_x, v_y) = (0, 0)$, to the Nyquist frequency, $(v_x, v_y) = (N/2, N/2)$. We refer to the local domain for a pixel (x, y) as the local spectrum, $S_{x,y}[v_x, v_y]$. Alternatively, we can refer to the components of the domain in terms of the frequency orders (p_x, p_y) , instead of the voice frequencies (v_x, v_y) , since they have a simple relationship.

Note that, since the voice images are of size $(2^{p_x-1} \times 2^{p_y-1})$, the mapping will not be unique for low-frequency components of neighboring pixels. For example, the $p_x = p_y = 2$ contribution describes the low-frequency portion of the image using 4 pixels. Therefore, all pixels in a particular quadrant of the image will contain the same low-frequency information. In the extreme case, the $p_x = p_y = 0$ order will be represented by a single point for the entire image.

The local frequency domain is analogous to a local 2D Fourier domain, often referred to as k -space. This provides a way to measure the texture characteristics by examining the contribution of each horizontal and vertical frequency bandwidth within the image. These “texture descriptors” replace the 1D texture curves of the polar ST analysis used in^{28–30}. If we use the dyadic sampling scheme presented above, the frequency scale is linear and the bandwidths do not overlap.

For example, Figure 3 shows the local domain at the center of each texture region in Figure 2a. The local domain of straw (Fig. 3a) reflects the strong diagonal stripes present in the image. The figure also shows that wood (Fig. 3b) has more high frequency components oriented horizontally than vertically, while sand (Fig. 3c) and grass (Fig. 3d) have similar frequency distributions, with grass having slightly more high frequency energy.

Primary Frequency Component

In order to more easily determine the most significant spatial frequency component at a particular point or region, we calculate the “primary” or highest amplitude frequency. This is done for a particular point by determining the value of each voice image in the DOST at that pixel location. We examine the value

of each voice image at location (x, y) and determine the component with the highest amplitude. The (v_x, v_y) pair is the “primary frequency” for the pixel.

We can extend this analysis to calculate the primary frequency at every image pixel. This results in a value of (v_x, v_y) for each value of (x, y) in the image. We can represent this as a complex image, with the real channel corresponding to the primary v_x value and the imaginary channel corresponding to the primary v_y value. The magnitude image can then be used to examine the primary radial frequency, $\nu_r = \sqrt{\nu_x^2 + \nu_y^2}$. The phase image tells us the angle of orientation of the primary frequency, $\theta_v = \arctan(v_y/v_x)$.

The DOST is normalized to preserve the length of the vector; like the FT, it satisfies a Parseval theorem. Thus, each voice is divided by the size of its partition—which attenuates the amplitude of high frequency signals when dyadic sampling is employed. To obtain a domain where the relative contribution of each partition is constant, we remove the partition factor. This provides the equivalent scaling as the original ST.

Effect of Spatial Transformations on the DOST

The DOST is rotationally variant because we calculate measures of local spatial frequency content in the horizontal and vertical directions. The result of a horizontal shift by Δx pixels and/or a vertical shift of Δy pixels results in a corresponding shift of each voice image. We can understand this effect by considering the DOST process on a shifted image. A circular shift of the image by Δx pixels in x and Δy pixels in y causes the phase of the FFT of the image to be multiplied by the complex value $e^{2\pi i(x\Delta x + y\Delta y)}$. The multiplied FFT is partitioned, and each partition is inverse Fourier-transformed. The phase ramp applied to each partition causes it to be shifted horizontally by Δx pixels and vertically by Δy pixels when inverse transformed. The multiplication of an image by a scaling factor causes the DOST to be multiplied by the same factor, as it is a linear transform.

In the case of a 90° rotation, the effect is a transpose or a switching of the x and y values in the image and a reversal of the y -coordinate. As a result, each voice image has the x - and y -coordinates switched and the y -coordinate reversed. Similarly, for a 180° rotation, the effect is to reverse the y -coordinates and maintain

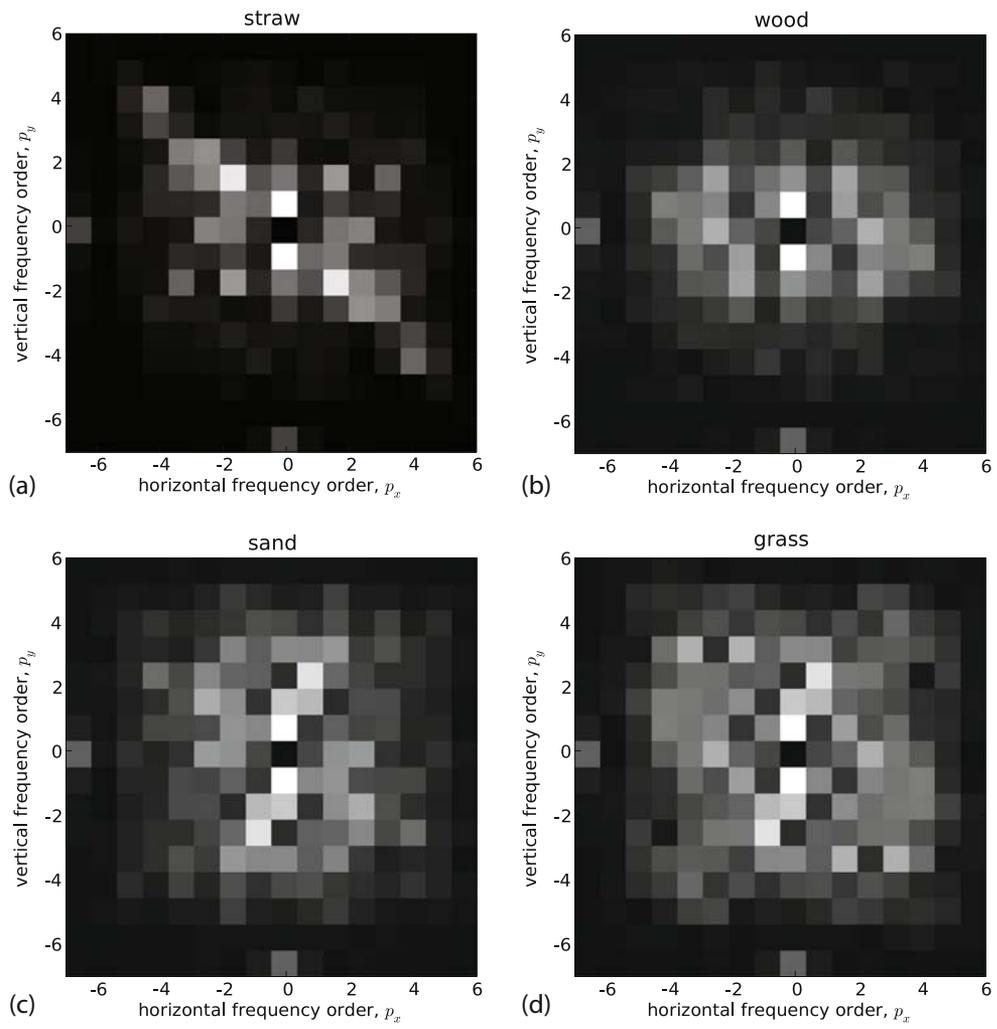


Fig. 3. The local DOST spectra of: a straw, b wood, c sand, and d grass obtained from the center pixel of each region in Figure 2a.

the x -coordinates. The result on the DOST is to reverse the y -coordinates of each voice image; the x -coordinates are not affected. In the case of rotations of $0 < \theta < 90^\circ$, the effect is more complicated. A rotation of the image implies a rotation of its FT. Therefore, when calculating the 2D-DOST of a rotated image, the partitioning grid is applied to the rotated FT. As a result, the frequency components are rotated into different partitions.

Rotationally Invariant DOST Features

In order to obtain rotationally invariant features from the DOST, we can follow a method similar to the invariant wavelet transform described in³⁴. In that approach, invariant wavelet coefficients are computed

by taking the wavelet coefficients at each level, averaged over the horizontal and vertical coefficients. Figure 4a shows the wavelet channels that are averaged together, marked with the same letter (A=level 0, B=level 1, etc.). The diagonal channels are excluded from the feature extraction, since they tend to contain the majority of the noise in the image and can degrade classification performance³⁴.

We can take a similar approach with the 2D-DOST by averaging together the magnitude of the horizontal and vertical frequency values for each order and excluding “diagonal” elements of the DOST domain where $p_x = p_y$. By combining the horizontal and vertical frequency information for the entire DOST, we can obtain an invariant domain describing the entire image. This is useful when using the

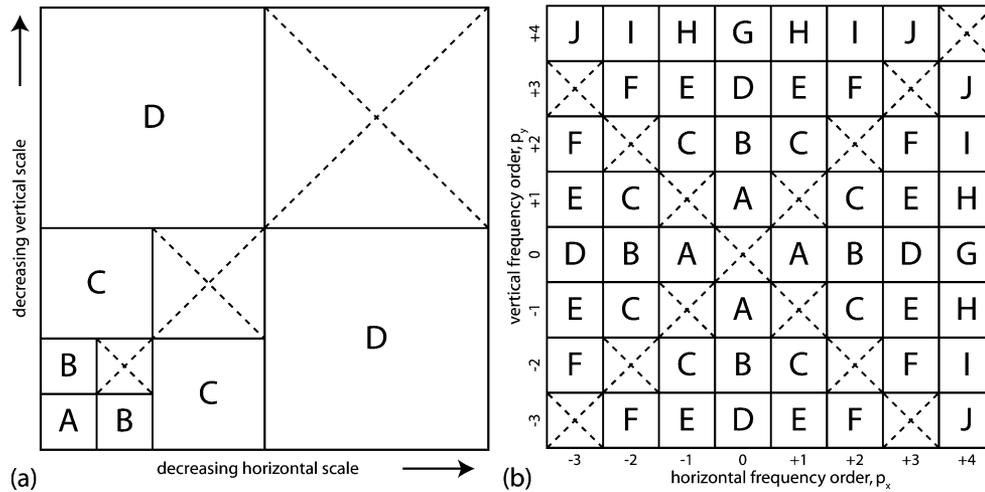


Fig. 4. a The wavelet decomposition for $N=8$ at three levels of decomposition. Invariant wavelet coefficients are calculated by averaging horizontal and vertical coefficients at each level (marked with the same letter: A level 0, B level 1, etc.). The diagonal channels are excluded from the feature extraction. b The local frequency domain generated from the 2D-DOST for $N=8$. Features marked with the same letter are averaged together to get an invariant DOST spectrum; a similar approach is taken of excluding the “diagonal” elements, where $p_x = p_y$. For a given N , the DOST approach provides more texture features (A to J = 10) than the wavelet approach (A to D = 4).

DOST to examine a small ROI. However, for medical imaging applications, we would like to retain the spatial information and the frequency information of the DOST. In this case, we can combine the diagonal elements of the local DOST spectrum, described in “Pixel-Wise Local Spatial Frequency Description.”

Figure 4b shows an example of how the values are calculated from the local frequency domain for $N=8$. The features marked with the same letter are averaged together to get an invariant DOST spectrum of texture features. For a given N , the DOST approach provides more texture features (A to H=10) than the wavelet approach (A to D=4).

METHODS

We conduct a series of experiments to determine the response of the 2D-DOST to a wide range of frequencies, texture patterns, and noise. These experiments attempt to provide a robust characterization of 2D-DOST based measures of texture and determine the limits on our ability to measure subtle texture changes in images. We also evaluate the performance of the rotationally invariant features in classifying a wide range of known textures and compare the results to those obtained using the invariant wavelet transform³⁴.

All DOST computations were carried out in Python 2.5 (<http://www.python.org>) using the Numerical Python package (<http://www.numpy.scipy.org>) and the Python Imaging Library (<http://www.pythonware.com/products/pil>). Wavelet coefficients were calculated in Matlab (The MathWorks, Inc., Natick, MA, USA) using `wavedec2.m` and `detcoef2.m` with the periodic extension mode. Rotated texture images were obtained from³³. Analysis was performed on a 2.4-GHz Intel Core 2 Duo with 2 GB of RAM on Mac OS X 10.5. Source code for the algorithm is provided in Appendix.

Transform Characterization

In order to determine the frequency response of the 2D-DOST, we calculated the point spread function. This was done by creating a “delta” image—a 256×256 image where the pixel at location (128,128) had a value of one, and all other pixels had a value of zero. We took the 2D-DOST of the delta function and then computed the local spectrum. We also measured the average time to calculate the DOST and determined the amount of memory required to store the structure, for randomly generated $N \times N$ images, where N varied from 4 to 1,024. We compared the computation times and storage requirements to the ST. These calculations were based on the calculation of real images, where only

half of the space-frequency domain needs to be calculated and stored due to Hermitian symmetry.

We tested the accuracy of the DOST inversion by taking the DOST of each of the nine Brodatz texture images shown in Figure 5, then inverting the 2D DOST domain according to Eqs. 4 and 5. We measured the L_1 norm between each inverted image, s_1 , and the original, s_0 :

$$L_1 = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N |s_1(i,j) - s_0(i,j)| \quad (6)$$

Response to Noise

To examine the changes in local DOST frequency domains with varying levels of noise, we evaluated a series of synthetic images with known frequency content and added noise. By varying both of these factors, we were able to determine the smallest change in spatial frequency, and contrast, that can be reliably detected. We expect these two factors to be related; that is, we expect that high spatial frequencies will be more reliably detected in images with high contrast-to-noise ratios. The relationship between contrast and maximum detectable frequency is akin to the modulation-transfer-function used to characterize imaging systems and should provide a robust characterization of the 2D-DOST.

We generated a series of images of size 256×256 with a single frequency component. We varied the frequency of a sinusoid of wavenumber $k=1, 10, 20, \dots, 120$, oriented either horizontally or diagonally, or both (equal components horizontally and vertically), with a minimum value of -1.0 and a maximum of $+1.0$. We also added various levels of Gaussian noise with standard deviation $\sigma=0.1, 0.2, \dots, 1.0$. This gave us a total of 260 test images (13 frequencies \times 2 orientations \times 10 noise levels).

To determine the effect of the varying frequency and noise, we measured the signal-to-noise ratio (SNR) of the local DOST spectrum, $S_{x,y}(p_x, p_y)$, for the central 5×5 pixels in the image. The SNR was defined in units of decibels (dB) as the amplitude of the frequency component being analyzed divided by the root mean square amplitude of the rest of the domain values.

$$\text{SNR}(\text{dB}) = 20 \log_{10} \left(\frac{A_{\text{signal}}}{A_{\text{noise}}} \right) \quad (7)$$

where A_{signal} is the amplitude of the voice frequency contained in the noise-free image in the local spectrum at the central pixel,

$$A_{\text{signal}} = S_{128,128}[p_x, p_y] \quad (8)$$

and A_{noise} is the RMS amplitude of the remaining frequency components, not present in the noise-

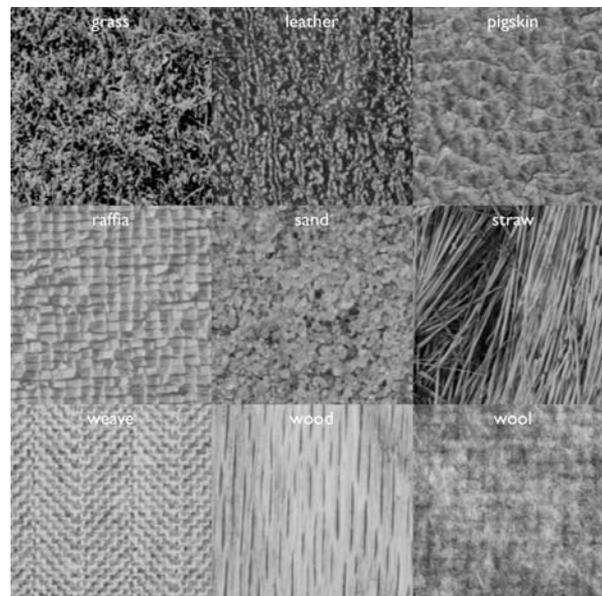


Fig. 5. The nine Brodatz texture images used for texture analysis experiments.

free image, but introduced with the addition of Gaussian noise,

$$A_{\text{noise}} = \sum_{\substack{i=0 \\ i \neq p_x}}^{\log_2 N} \sum_{\substack{j=0 \\ j \neq p_y}}^{\log_2 N} \sqrt{\frac{S_{128,128}[i,j]}{(\log_2 N)^2 - 2}} \quad (9)$$

We analyzed the accuracy of the primary frequency estimation by determining what fraction of image pixels were correctly classified in the single-frequency images. For this experiment, we used a 10×10 pixel region in the center of the image or 100 test pixels per image (for a total of 26,000 pixels). For each pixel within the central 10×10 region, we recorded the primary frequency as found by the 2D-DOST and compared it to the true frequency of the image. The number of misclassifications were recorded.

Classification Experiment

To test the ability of the rotation-invariant DOST to classify textures, we performed an experiment on nine images from the Brodatz texture library shown in Figure 5: grass, straw, herringbone weave, woolen cloth, pressed calf leather, beach sand, wood grain, raffia, and pigskin, obtained from³³. We tested the ability of the invariant DOST spectra to classify the images when: (1) trained and tested on non-rotated images, (2) trained on rotated and tested on non-rotated images, and (3) trained and tested on a variety of angles. We extracted multiple sub-images from each 512×512 texture image; each sub-image was of size 64×64 pixels.

We then calculated the invariant DOST spectrum for each sub-image.

We compared the classification results using the DOST to those obtained using the invariant wavelet transform described in³⁴. The DWT of each image was computed in Matlab with a db4 mother wavelet at five levels of decomposition using wavedec2.m. The invariant coefficients were calculated by averaging the horizontal and vertical coefficients for each level of decomposition (obtained using detcoef2.m) and computing the L1 norm, given by

$$e_n = \frac{1}{M^2} \sum_{i=1}^M \sum_{j=1}^M |w_n(i,j)| \quad (10)$$

where the w_n is the wavelet decomposition for level n of dimensions $M \times M$.

Classification was performed in Matlab using linear discriminant analysis (classify.m). We used two different types of discriminant functions: a linear classifier that fits a multivariate normal density to each group with a pooled estimate of covariance and a quadratic function that fits multivariate normal densities with covariance estimates stratified by group (texture).

For the first experiment, we extracted 64 sub-images of size 64×64 from each texture oriented at 0° . We trained the classifier on 56 sub-images of each texture (56 images \times 9 textures = 504 total) and tested the classification on eight sub-images of each texture (8 images \times 9 textures = 72 total). In the second case, we tested the ability of the DOST and wavelet methods to classify images oriented at a

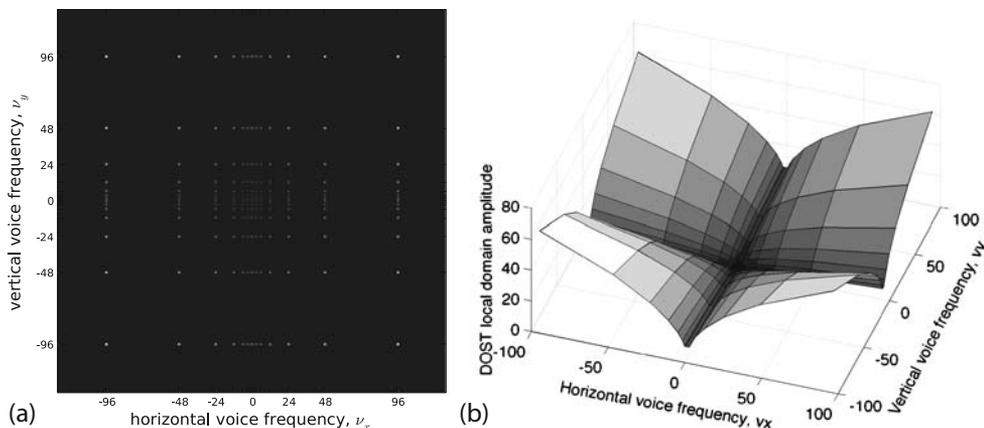


Fig. 6. a The 2D-DOST of a delta function and b its local domain at the central point of the image (128,128).

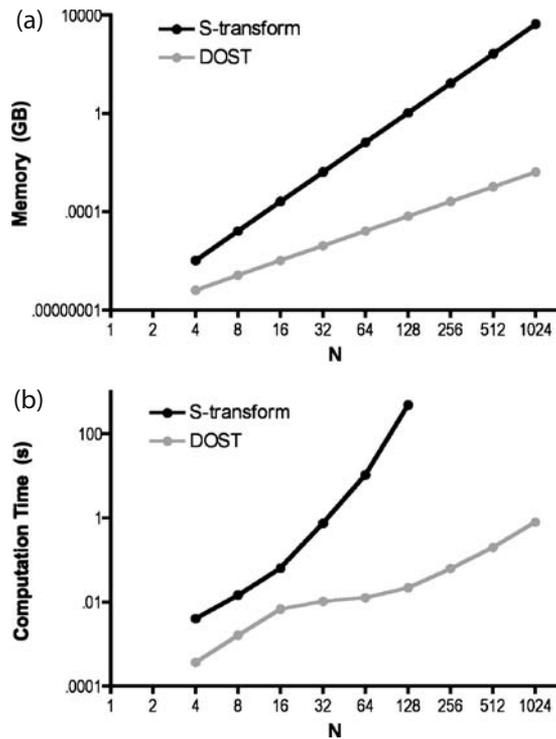


Fig. 7. a The calculated increase in memory required to store the ST and the DOST of a real $N \times N$ image. The ST requires N^4 floating point values, while the DOST requires only N^2 (double for complex images). b The measured time to compute the ST and DOST. We were unable to compute the ST of images larger than 128×128 since the memory requirements became too large.

different angle than the classifier was trained on. For this experiment, we extracted texture images oriented at angles: 0° , 30° , 60° , 90° and 120° . We extracted 25 sub-images of size 64×64 pixels from each $512 \times$

512 image at each angle (25 sub-images $\times 5$ angles $\times 9$ textures = $1,125$ images in total). We then calculated the invariant DOST spectrum for each image and the invariant wavelet spectrum. The classifier was trained on the DOST spectra from the 25 sub-images at angles 30° , 60° , 90° , and 120° (25 sub-images $\times 4$ angles $\times 9$ textures = 900 images) and tested on the 0° images (25 sub-images $\times 1$ angle $\times 9$ textures = 225 images). Finally, we trained the classifier on DOST spectra from 20 sub-images of each texture at each angle (0° , 30° , 60° , 90° , and 120° ; 20 sub-images $\times 5$ angles $\times 9$ textures = 900 images) and tested on five images of each texture at each angle (5 sub-images $\times 5$ angles $\times 9$ textures = 225 images).

RESULTS

The 2D-DOST of the delta function and the local frequency domain of the pixels at location $(128, 128)$ are shown in Figure 6. In part b of the figure, we can see the square root increase in DOST amplitude with increasing frequency. We found that the L_1 difference between the inverted and original images was close to machine epsilon (mean = 8.45×10^{-14} , median = 6.74×10^{-14}). Figure 7 illustrates the dramatic reduction in memory and computation requirements for the DOST, as compared to the ST for square images of size $N \times N$. For example, calculation of the ST of a 128×128 image required approximately 8 min and 1 GB of memory; calculation of the DOST required only 0.02 s and 65 kB of memory. We were unable to compute the

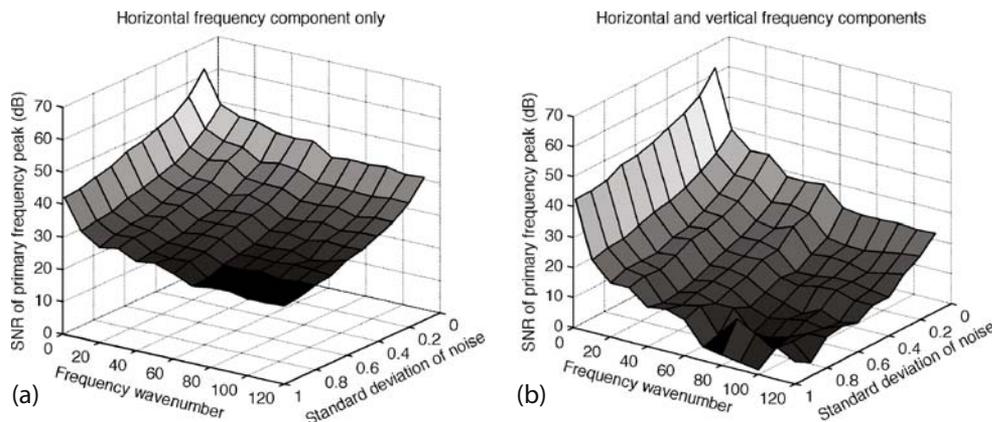


Fig. 8. a The SNR (dB) of the main frequency peak of the local DOST domain at the center of a 256×256 image containing a single horizontally oriented frequency component. b The SNR of the main peak when examining an image with a horizontally and vertically oriented single frequency component. Note that the SNR of these peaks drops more quickly with increasing noise and increasing frequency than when only one component is present.

Table 1. Misclassifications (/100) of the Primary Frequency Component (Oriented Diagonally)

Frequency (diagonal)	Standard deviation of Gaussian (σ)											total
	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	
1	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0
40	0	0	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0	0	0
60	0	0	0	0	0	0	0	0	0	0	0	0
70	0	0	0	0	0	0	36	32	43	35	62	208
80	0	0	0	0	4	0	0	6	70	55	61	196
90	0	0	0	0	0	0	0	14	34	31	64	143
100	0	0	0	0	0	0	0	18	42	91	70	221
110	0	0	0	0	0	0	8	75	33	87	87	290
120	0	0	0	0	0	0	8	41	13	13	24	99
Total	0	0	0	0	4	0	52	186	235	312	386	1,157

ST of images sized 256×256 or larger since more than 17 GB of memory was required.

We found that the SNR of the main DOST peak was inversely proportional to frequency and noise, as shown in Figure 8. The SNR of the main peak dropped more quickly with increasing noise and increasing frequency when there are horizontal and vertical frequency components than when only one component is present. The misclassification rate for all horizontally oriented frequencies was zero. The diagonally oriented frequencies had a zero misclassification rate for all cases of $\sigma < 0.4$ and for all frequencies less than 70, as shown in Table 1.

The classification accuracy using the DOST is higher than that of the invariant wavelet for each case studied (Table 2). We obtained the most accurate results with the linear classifier using the DOST when training and testing on a single angle (91.7% DOST vs. 83.3% wavelet). The best results using the quadratic classifier were obtained when training on the rotated images and testing on the zero-degree images (94.7% DOST vs. 84.4% wavelet). The classification accuracy decreased slightly for the DOST when training and testing on all angles but remained higher than the wavelet case for both linear and quadratic classifiers. We observed poorer

performance from the invariant wavelet transform than previously noted³⁴. The discrepancy may be related to the difference in sub-image size (64×64 pixel as opposed to 16×16), the specific texture images used or the ratio of training to testing images used in the classification.

DISCUSSION AND CONCLUSIONS

We have presented a frequency-domain implementation of the 2D-DOST by partitioning the frequency domain in a dyadic sampling scheme. We showed that the DOST can provide a pixel-by-pixel texture description of an image by creating local spectra, containing the horizontal and vertical frequency information from the FT of the image. The DOST is straightforward and fast to calculate, allowing us to analyze every pixel in a large image under a second on standard computers. We confirmed that the DOST can be very accurately inverted to recover the original image. The DOST is robust in the presence of low to moderate noise levels and single frequency components can be accurately identified from the local spectra. Rotationally invariant texture features can be extracted

Table 2. Classification Accuracy for Invariant Wavelet (db4) and Invariant DOST on 9 Brodatz Texture Images

Method	Number of features	Classifier	Train test on 0° (%)	Train on rotated, test on 0° (%)	Train and test on all angles (%)
Invariant DOST	21	Linear	91.7	91.1	86.2
		Quadratic	91.7	94.7	88.0
Invariant wavelet	6	Linear	83.3	84.9	83.1
		Quadratic	77.8	84.4	86.7

by combining horizontal and vertical frequency information from either the entire 2D-DOST or the local DOST spectrum corresponding to a single pixel. The invariant 2D-DOST provides higher classification accuracy on texture images than the comparable invariant wavelet approach. The texture information produced by the DOST may help discriminate between normal and abnormal tissues in medical images.

One limitation of the DOST in our work is that ringing artifacts are caused due to the use of rectangular windows in the analysis. A future step will be to employ apodizing windows and oversampling algorithms, as mentioned in³², which would resolve these issues. Furthermore, our current

implementation of the DOST, given in [Appendix](#), is only suitable for images of size $2^n \times 2^n$. As outlined in³², dyadic sampling can be applied to any size data. Arbitrary sampling schemes can also be devised, for images of any size. Future work will include: extension of our code to apply dyadic sampling to arbitrary sized images, the evaluation of new sampling schemes, application of the DOST to clinical data, and comparison to previous continuous ST-based texture analysis.

APPENDIX

The following Python source code was used to calculate the 2D-DOST of a real $N \times N$ image *im*:

```

from numpy import *
from numpy.fft import *

def shift(a, n):
    if len(shape(a)) == 1:
        b = concatenate((a[n:], a[0:n]))
    elif len(shape(a)) == 2:
        c = concatenate((a[n[0]:, :], a[0:n[0], :]))
        b = concatenate((c[:, n[1]:], c[:, 0:n[1]]), 1)
    return b

N = len(im)
IM = fft2(im)
S = zeros((N,N), complex)
S[0,0] = IM[0,0]
S[N/2,0] = IM[N/2,0]
S[0,N/2] = IM[0,N/2]
S[N/2,1] = IM[N/2,1]
S[1,N/2] = IM[1,N/2]
S[N/2,N/2] = IM[N/2,N/2]
n = int(log(N)/log(2))

for py in range(1,n):
    ny=2**(py-1)
    S[0,ny:ny*2]=ifft(shift(IM[0,ny:ny*2], -ny/2))*sqrt(ny)
    S[ny:ny*2,0]=ifft(shift(IM[ny:ny*2,0], -ny/2))*sqrt(ny)
    S[N-ny*2+1:N-ny+1,0]=ifft(shift(IM[N-ny*2+1:N-ny+1,0], -ny/2))[::-1]*sqrt(ny)
    S[N/2,ny:ny*2]=ifft(shift(IM[N/2,ny:ny*2], -ny/2))*sqrt(ny)
    S[ny:ny*2,N/2]=ifft(shift(IM[ny:ny*2,N/2], -ny/2))*sqrt(ny)
    S[N-ny*2+1:N-ny+1,N/2]=ifft(shift(IM[N-ny*2+1:N-ny+1,N/2], -ny/2))[::-1]*sqrt(ny)
    for px in range(1,n):
        nx=2**(px-1)
        S[ny:ny*2, nx:nx*2]=ifftn(shift(IM[ny:ny*2, nx:nx*2], (-ny/2, -nx/2)))*sqrt(ny*nx)
        S[N-ny*2+1:N-ny+1, nx:nx*2]=ifftn(shift(IM[N-ny*2+1:N-ny+1, nx:nx*2], (-ny/2, -nx/2))[::-1])*sqrt(ny*nx)

S[1:N/2,N/2+1:] = conjugate(S[N/2+1:,1:N/2][::-1,::-1])
S[N/2+1:,N/2+1:] = conjugate(S[1:N/2,1:N/2][::-1,::-1])
S[0,N/2+1:] = conjugate(S[0,1:N/2][::-1])
S[N/2,N/2+1:] = conjugate(S[N/2,1:N/2][::-1])

```

REFERENCES

1. Castleman KR: Digital Image Processing, Upper Saddle River, NJ, USA: Prentice Hall, 1996
2. Lerski RA, Straughan K, Schad LR, Boyce D, Bluml S, Zuna I: MR image texture analysis—an approach to tissue characterization. *Magn Reson Imaging* 11:873–887, 1993
3. Castellano G, Bonilha L, Li LM, Cendes F: Texture analysis of medical images. *Clin Radiol* 59:1061–1069, 2004
4. Yu O, Mauss Y, Zollner G, Namer IJ, Chambron J: Distinct patterns of active and non-active plaques using texture analysis on brain NMR images in multiple sclerosis patients: preliminary results. *Magn Reson Imaging* 17:1261–1267, 1999
5. Mathias JM, Tofts PS, Losseff NA: Texture analysis of spinal cord pathology in multiple sclerosis. *Magn Reson Med* 42:929–935, 1999
6. Mahmoud-Ghoneim D, Toussaint G, Constans JM, de Certaines JD: Three dimensional texture analysis in MRI: a preliminary evaluation in gliomas. *Magn Reson Imaging* 21:983–987, 2003
7. Jirak D, Dezortova M, Taimr P, Hajek M: Texture analysis of human liver. *J Magn Reson Imaging* 15:68–74, 2002
8. Kerut EK, Given M, Giles TD: Review of methods for texture analysis of myocardium from echocardiographic images: a means of tissue characterization. *Echocardiography* 20:727–736, 2003
9. Mayerhoefer ME, Breitenseher MJ, Kramer J, Aigner N, Hofmann S, Materka A: Texture analysis for tissue discrimination on T1-weighted MR images of the knee joint in a multicenter study: Transferability of texture features and comparison of feature selection methods and classifiers. *J Magn Reson Imaging* 22:674–680, 2005
10. Schilling T, Miroslaw L, Glab G, Smereka M: Towards rapid cervical cancer diagnosis: Automated detection and classification of pathologic cells in phase-contrast images. *Int J Gynecol Cancer* 17:118–126, 2007
11. Haralick RM: Statistical and structural approaches to texture. *Proc. IEEE* 67:786–808, 1979
12. McNitt-Gray MF, Wyckoff N, Sayre JW, Goldin JG, Aberle DR: The effects of co-occurrence matrix based texture parameters on the classification of solitary pulmonary nodules imaged on computed tomography. *Comput Med Imaging Graph* 23:339–348, 1999
13. Zhang Y, Zhu H, Ferrari R, Wei X, Eliasziw M, Metz LM, Mitchell JR: Texture analysis of MR images of minocycline treated MS patients. *Lect Notes Comp Sci* 2878:786–793, 2003
14. Daubechies I: Ten Lectures on Wavelets, Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 1992
15. Weyn B, van de Wouwer G, van Daele A, Scheunders P, van Dyck D, van Marck E, Jacob W: Automated breast tumor diagnosis and grading based on wavelet chromatin texture description. *Cytometry* 33:32–40, 1998
16. Gibson PC, Lamoureux MP, Margrave GF: Stockwell and wavelet transforms. *J Fourier Anal Appl* 12:713–721, 2006
17. Stockwell RG, Mansinha L, Lowe RP: Localization of the complex spectrum: the S transform. *IEEE Trans Signal Process* 44:998–1001, 1996
18. Pinnegar CR, Eaton DW: Application of the S transform to prestack noise attenuation filtering. *J Geophys Res* 108:2422, 2003
19. Aldas GU: Application of the Stockwell transform to blasting-induced ground vibration. *Int J Surf Min Reclam Environ* 19:100–107, 2005
20. Oldenborger GA, Schincariol RA, Mansinha L: Space-local spectral texture segmentation applied to characterizing the heterogeneity of hydraulic conductivity. *Water Resour Res* 38:1154, 2002
21. Clapson AC, Barsuglia M, Bizouard MA, Brisson V, Cavalier F, Davier M, Hello P, Kreckelberg S, Varvella M: A gravitational wave burst search method based on the S transform. *Classical Quant Grav* 22:S1381–S1390, 2005
22. Reddy JB, Mohanta DK, Karan BM: Power system disturbance recognition using wavelet and S-transform techniques. *International Journal of Emerging Electric Power Systems* 1:1007, 2004
23. Khosravani H, Pinnegar CR, Mitchell JR, Bardakjian BL, Federico P, Carlen PL: Increased high-frequency oscillations precede in vitro low-Mg²⁺ seizures. *Epilepsia* 46:1188–1197, 2005
24. Goodyear BG, Zhu H, Brown RA, Mitchell JR: Removal of phase artifacts from fMRI data using a Stockwell transform filter improves brain activity detection. *Magn Reson Med* 51:16–21, 2004
25. Assous S, Humeau A, Tartas M, Abraham P, L'Huillier JP: S-transform applied to laser Doppler flowmetry reactive hyperemia signals. *IEEE Trans Biomed Eng* 53:1032–1037, 2006
26. Brown RA, Zhu H, Mitchell JR: Distributed vector processing of a new local multiscale Fourier transform for medical imaging applications. *IEEE Trans Med Imaging* 24:689–691, 2005
27. Mansinha L, Stockwell RG, Lowe RP: Pattern analysis with two-dimensional spectral localisation: Applications of two-dimensional S transforms. *Physica A* 239:286–295, 1997
28. Brown R, Zlatescu M, Sijben A, Roldan A, Easaw J, Forsyth P, Pamey I, Sevcik R, Yan E, Demetrick D, Schiff D, Cairncross G, Mitchell R: The use of magnetic resonance imaging to noninvasively detect genetic signatures in oligodendroglioma. *Clin Cancer Res* 14:2357–2362, 2008
29. Zhang Y, Wells J, Buist R, Peeling J, Wee Yong V, Mitchell JR: A novel MRI texture analysis of demyelination and inflammation in relapsing-remitting experimental allergic encephalomyelitis. *Lect Notes Comp Sci* 4190:760–767, 2006
30. Zhu, H, Wei X, Zhang Y, Mayer GS, Mitchell JR: Temporal texture analysis of normal appearing white matter in multiple sclerosis. *Proc 11th Intl Soc Mag Reson Med* 277, 2003
31. Zhu H, Goodyear BG, Lauzon ML, Brown RA, Mayer GS, Law AG, Mansinha L, Mitchell JR: A new local multiscale Fourier analysis for medical imaging. *Med Phys* 30:1134–1141, 2003
32. Stockwell RG: A basis for efficient representation of the S-transform. *Digit Signal Process* 17:371–393, 2007
33. University of Southern California Signal & Imaging Processing Institute. USC-SIPI Image Database. Available at <http://sipi.usc.edu/database>. Accessed April 17, 2008.
34. Porter R, Canagarajah N: Robust rotation-invariant texture classification: wavelet, Gabor filter and GMRF based schemes. *IEE Proc Vis Image Signal Process* 144:180–188, 1997